IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 5, Oct-Nov, 2013
ISSN: 2320 - 8791
www.ijreat.org

Semi Automatic Annotation Exploitation Similarity of Pics in Personal Photo Albums

P. Subashree Kasi Thangam[1] and R. Rosy Angel[2]

[1]**Assistant Professor, Department of Computer Science Engineering College, Tuticorin.**

[2]**Assistant Professor, Department of Computer Science Engineering College, Tuticorin.**

Abstract

This paper proposed a semi-automatic tagging scheme that makes photo album tagging easy for the users. The proposed scheme involves block-based low-level feature extraction from images followed by the clustering of the feature space to form higher level, semantically meaningful patterns. The clustering of the feature space is realized by an expectation–maximization algorithm that uses an iterative approach to automatically determine the number of clusters. Then, the property of PANDA is exploited: the similarity between two clusters is estimated as a function of the similarity of both their structures and the measure components. Using the similarities of photos, exemplars are selected using AP algorithm. Then the exemplars are manually tagged. Based on the tags of exemplars, rest of the photos are automatically tagged.

1. Introduction

In online computer systems terminology, a tag is a non-hierarchical keyword or term assigned to a piece of information (such as an Internet bookmark, digital image, or computer file). This kind of metadata helps to describe an item and allows it to be found again by browsing or searching. Tags are generally chosen informally and personally by the item's creator or by its viewer, depending on the system. Tagging was popularized by websites associated with Web 2.0 and is an important feature of many Web 2.0 services. It is now also part of some desktop software.

Labeling and tagging may take the form of words, images, or other identifying marks. In 2003, the social bookmarking website Delicious provided a way for its users to add "tags" to their bookmarks (as a way to help find them later); Delicious also provided browseable aggregated views of the bookmarks of all users featuring a particular tag. Flickr allowed its users to add free-form tags to each of their pictures, constructing flexible and easy metadata that made the pictures highly searchable. The success of Flickr and the influence of Delicious popularized the concept, and other social software websites – such as YouTube, Technorati, and Last.fm – also implemented tagging. "Labels" in Gmail are similar to tags. Websites that include tags often display collections of tags as tag clouds. A user's tags are useful both to them and to the larger community of the website's users.

Automatic image annotation (also known as automatic image tagging) is the process by which a computer system automatically assigns metadata in the form of captioning or keywords to a digital image. This application of computer vision techniques is used in image retrieval systems to organize and locate images of interest from a database.

This method can be regarded as a type of multi-class image classification with a very large number of classes - as large as the vocabulary size. Typically, image analysis in the form of extracted feature vectors and the training annotation words are used by machine learning techniques to attempt to automatically apply annotations to new images. The first methods learned the correlations between image features and training annotations, then techniques were developed using machine translation to try to translate the textual vocabulary with the 'visual vocabulary', or clustered regions known as *blobs*. Work following these efforts have included classification approaches, relevance models and so on.

The advantages of automatic image annotation versus content-based image retrieval are that queries can be more naturally specified by the user. CBIR generally (at present) requires users to search by image concepts such as color and texture, or finding example queries. Certain image features in example images may override the concept that the user is really focusing on. The traditional methods of image retrieval such as those used by libraries have relied on manually annotated images, which is expensive and time-consuming, especially given the large and constantly-growing image databases in existence. Automatic tagging (or annotation) techniques rarely achieve performance that satisfies the users.

The manual image annotation is an expensive and labor intensive procedure. Although this allows semantic image retrieval manual annotations are expensive and do not always capture the content of images and videos well. Two types of manual image annotation are exhaustively annotation and batch annotation.

In this paper, a semi-automatic photo annotation scheme is proposed that is able to modulate the manual efforts and the tagging performance in a flexible way. The proposed annotation scheme works in a semi-automatic manner in which the features are extracted from the photos and the patterns are framed. And then the similarity between the photos are computed. Based on the similarities, the exemplars are selected. And the users manually tag selected exemplars only. And the rest of the photos are automatically annotated based on the tags of the exemplars.

The organization of the rest of this paper is as follows. A short review on the related work is provided in Section II and the overview of the proposed semi-automatic annotation scheme is described in Section III. Feature Extraction is described in Section IV. In Sections V and VI, the exemplar selection algorithm and the tag inference algorithm, respectively are introduced. Finally, the conclusion of the paper is given in Section VII.

2. RELATED WORK

A. Annotation of Photos

M. Ames and M. Naaman have explored the motivations for tagging changed with online photo sharing communities such as Flickr. Flickr allows annotation of photos in the form of tags, or unstructured textual labels. ZoneTag is a cameraphone application used to upload photos taken by the phone to Flickr. ZoneTag is designed to capture, annotate, store and share photos from the phone. L. Kennedy, S.-F. Chang, and I. Kozintsev have predicted the performance of search based on automatic image classifiers. D. Liu, X.-S. Hua, L. Yang, M.Wang, and H.-J. Zhang have detailed that a tag ranking is an approach, in which the tags of an image can be automatically ranked according to their relevance with the image. B. Sigurbjörnsson and R. Zwol have explained the Flickr Tag Recommendation based on Collective Knowledge.

B. Active Learning

S. Tong and E. Chang have proposed the use of a support vector machine active learning (SVM*Active*) algorithm for conducting effective relevance feedback for image retrieval. W. Jiang, S.-F. Chang, and A. Loui have explained Active Context-Based Concept Fusion With Partial User Labels to effectively exploit the contextual relations among concepts.

C. Annotation of Photo Album

B. Suh and B. Bederson have introduced two photo clustering algorithms for generating meaningful photo groups: (1) Hierarchical event clustering; and (2) Clothing based person recognition, which assumes that people who wear similar clothing and appear in photos taken in one day are very likely to be the same person. J. Cui, F.Wen, R. Xiao, Y. Tian, and X. Tang have developed interaction techniques for semi-automatic photo annotation. This approach provides the following new features: “cluster annotation” puts similar faces or photos with similar scene together to label them in one operation; “contextual re-ranking” boosts the labeling productivity by guessing the user intention; “ad hoc annotation” allows user label photos while they are browsing or searching, and improves system performance.

3. SEMI-AUTOMATIC ANNOTATION SCHEME OF PHOTO ALBUMS

The proposed annotation scheme is performed as follows. For a given photo album, low level features are extracted from all the photos. And the low level features are clustered to form patterns. Then the similarities of the patterns are computed. Using the similarities, exemplars are selected from the album. And the user manually annotate the exemplars. Based on the tags of the

exemplars, rest of the photos are automatically annotated. The architecture diagram of the proposed system is shown in below figure :

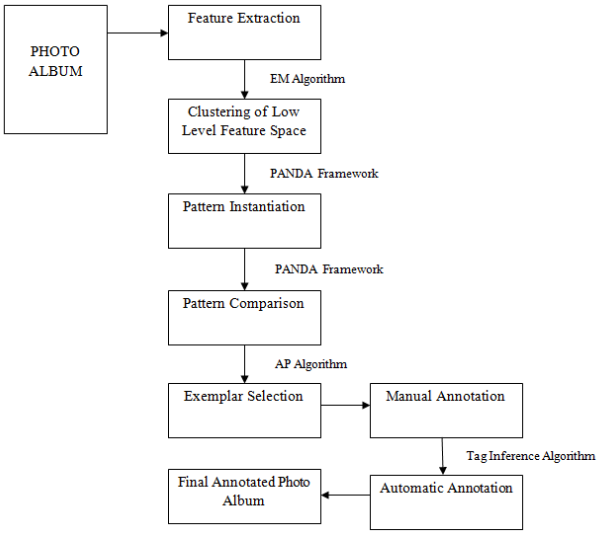

Fig. 1 Architecture diagram for semi-automatic photo annotation scheme

In the above data flow diagram (Fig. 1), for a given photo album, first low level features are extracted. Then they are clustered to form patterns and similarities are computed. Then a set of exemplars are selected and they are manually annotated. Based on these tags the rest of the photos are automatically annotated.

The Procedure of the Proposed Semi-Automatic Annotation Scheme of Photo Albums is given as follows:

Input : A given photo album.
Output : Final Annotated Photo Album.
Steps :

1) Low Level Feature Extraction.
2) Clustering of Feature Space using EM Algorithm.
3) Pattern Instantiation using PANDA Framework.
4) Computation of Pattern Similarities using PANDA Framework.
5) Selection of Exemplars using AP Algorithm.
6) Manual Annotation of Exemplars.
7) Automatic Annotation of rest of the photos using Tag Inference Algorithm.

4. FEATURE EXTRACTION

For each photos in a given photo album, low level features are extracted. The color features are extracted by extracting the first four moments of three channels of CIE Luv color space. Then the Gabor texture feature is extracted, by using six scales and six orientations of Gabor transformation to extract their means and standard deviations. And the edge direction histogram feature, to represent the shape of the photos.

The low-level feature vectors are clustered using mixture models that model the data by a number of Gaussian distributions. A cluster corresponds to a set of distributions, one for each dimension of the dataset. Each distribution is described in terms of mean and standard deviation. A probabilistic approach to assigning feature vectors to clusters is used.

For 1-D datasets, a mixture is a set of c Gaussian probability distributions, representing c clusters. The parameters of a mixture model are determined by the expectation maximization (EM) algorithm.

The EM algorithm is used to estimate the maximum likelihood L of θ given a set of features $\{x1, \ldots, xN\}$. The algorithm results in a set of distributions, a vector of pairs of means μ and standard deviations σ, each of which corresponds to a feature, and outputs the size of the cluster (the number of vectors that belong to the cluster). The vector of means μ of the distributions for every feature represents the centroid of the cluster.

The clusters resulting from the EM algorithm are considered as patterns extracted from the image database, and are represented and handled according to the PANDA formalization. A Specimen*i* is instantiated for each pattern *Pi* .

The similarity between patterns is computed by taking into account both the similarity between the patterns' structures and the similarity between the measures. The similarity between two *simple patterns* of the same type *pt* can be computed by combining, by means of an *aggregation function faggr*, the similarity between both the structure and the measure components.

The similarity between the structure of complex patterns is conceptually evaluated by using the *coupling type*, which is used to establish how component patterns can be *matched*, and the *aggregation logic*, which is used to combine the similarity scores obtained for coupled component patterns into a single overall score representing the similarity between the complex patterns.

5. EXEMPLAR SELECTION

The exemplar selection at the first round of the proposed semi-automatic tagging scheme is accomplished via a temporally consistent affinity propagation algorithm. The affinity propagation (AP) algorithm which is a similarity-based clustering algorithm that is able to group a given set of samples into several clusters as well as select an exemplar from each cluster.

Given a set of n data points $X=\{x_1,x_2,\ldots x_n\}$, the algorithm takes as input the pairwise similarity s(i,j) between any two points x_i and x_j in X. The algorithm then works by iterating the following two simple messages until convergence:

$$r(i,k) = s(i,k) - \max_{k' \neq k} [a(i,k') + s(i,k')]$$

$$a(i,k) = \begin{cases} \min [0, r(k,k) + \Sigma_{i' \in \{i,k'\}} \max[0,r(i',k)]], & i \neq k \\ \Sigma_{i' \neq k} \max [0,r\{i',k)], & i=k \end{cases}$$

The above messages have an intuitive interpretation: the "responsibility" r(i,k) sent from x_i to x_k indicates how well x_k serves as the exemplar of considering other potential exemplars for x_i, and the "availability" sent from x_k to x_i indicates how appropriate x_i chooses x_k as its exemplar considering other potential samples that may choose x_k as their exemplar. The belief that image x_i selects image x_k as its exemplar is derived as the sum of the incoming messages

$$t(i,k) = a(i,k) + s(i,k)$$

After convergence of the message updates, the exemplar of point is decided as x_k according to the criterion

$$K^* = \arg_k \max [t(i,k)]$$

And finally, the pattern similarity between the images are used to select the exemplars from the photo albums. The similarity between photos is estimated as,

$$s(i, j) = \alpha \exp (- (\|v_i - v_j\|^2) / \sigma_v^2)$$

where v_i is the low level feature vector of photo x_i, α is a weight factor between 0 and 1, and $\|.\|$ denotes l_2-norm.

The AP is selected as exemplar selection algorithm due to its advantages in the following aspects: 1) its effectiveness in clustering has been shown in many tasks; 2) it simultaneously accomplishes the clustering and the selection of exemplars. Several other methods, such as K-means and spectral clustering, only cluster samples, and the centroids of the obtained clusters may not be real samples.

6. TAG INFERENCE

Then, the rest of the photos are automatically tagged. The tags of the exemplars are taken as labeled data. Based on this, the unlabeled data are discovered i.e., the tags of the remaining photos. This method is related to a graph-based semi-supervised learning approach. This algorithm iteratively propagate the tags of each photo to others by holding the tags of exemplars.

Tag Inference Algorithm:
Input : Tags of Exemplars
Output : Tags of remaining photos.
Steps :

1) Hold the tags of exemplars.
2) Find the tags of remaining photos.

The performance of the proposed annotation scheme is evaluated by measuring precision, recall, and F1-measure of the tags obtained with the proposed annotation scheme for each photo in photo album. Then the F1-measure of all photos in an album are averaged to evaluate the tagging performance on the album. Finally, the F1-measure of all albums are averaged and it is adopted as the performance evaluation measurement in this work.

7. CONCLUSION

A semi-automatic annotation scheme for personal photo albums, is proposed which achieves a good trade-off between manual efforts and tag performance. In this scheme, features are extracted and patterns are formed. Based on the pattern similarities, exemplars are selected for manual annotation. And based on these tags, remaining photos are automatically annotated. So this scheme semi automatically annotate the personal photo albums in a conceptual manner.

REFERENCES

[1] M. Ames and M. Naaman, "Why we tag: Motivations for annotation in mobile and online media," in *Proc. ACMSIGCHI Conf. Human Factors in Computing Systems*, 2007, pp. 971–980.
[2] D. Liu, X.-S. Hua, L. Yang, M.Wang, and H.-J. Zhang, "Tag ranking," in *Proc. 18th ACM Int. Conf. World Wide Web*, 2009, pp. 351–360.
[3] D. Liu, X.-S. Hua, L. Yang, and H.-J. Zhang, "Multiple-instance active learning for image categorization," in *Proc. 15th Int. Multimedia Modeling Conf.*, 2009, pp. 239–249.

[4] W. Jiang, S.-F. Chang, and A. Loui, "Active context-based concept fusion with partial user label," in *Proc. IEEE Int. Conf. Image Processing*, 2008.
[5] B. Suh and B. Bederson, "Semi-automatic photo annotation strategies using event based clustering and clothing based person recognition," *Interact. Comput.*, vol. 19, no. 4, pp. 524–544, 2007.
[6] J. Cui, F.Wen, R. Xiao, Y. Tian, and X. Tang, "Easyalbum: An interactive photo annotation system based on face clustering and re-ranking," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, 2007, pp. 367–376.
[7] B. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 319, no. 5963, pp. 972–976, 2007.