Graph Learning System for Automatic Image Annotation

K. Aishwaryameenakshi, S. Halima Banu,
A.T.R. Krishna Priya, and S. Chitrakala

Department of Computer Science and Engineering
Easwari Engineering College
Chennai, India
{k.aishwaryameenakshi,halimabanu91,
priya.gita6,ckgops}@gmail.com

Abstract. Automating the process of annotation of images is a crucial step towards efficient and effective management of increasingly high volume of content. A graph-based approach for automatic image annotation is proposed which models both feature similarities and semantic relations in a single graph. The proposed approach models the relationship between the images and words by an undirected graph. Semantic information is extracted from paired nodes. The quality of annotation is enhanced by introducing graph link weighting techniques. The proposed method achieves fast solution by using incremental fast random walk with restart (IFRWR) algorithm, without apparently affecting the accuracy of image annotation.

Keywords: automatic image annotation, graph learning, graph link weighting, fast solution.

1 Introduction

How to index and search for the digital image collections effectively and efficiently is an increasingly urgent research issue in the multimedia community. To support this, keywords describing the images are required to retrieve and rank images. Manual annotation is a direct way to obtain these keywords. But, it is labor-intensive and error-prone. Thus automatic annotation of images has emerged as an important technique for efficient image search. An image annotation algorithm based on graph is proposed. Based on the analysis of graph structure, a fast solution algorithm is proposed.

2 Related Work

Approach using Cross-Media Relevance Models (CMRM) was introduced by Jeon et al. [1]. Feng et al. [2] modified it using multiple-Bernoulli distribution. Carneiro and Vasconcelesy presented a Supervised Multiclass Labeling (SML) [3-4]. For the past years, graph-based automatic image annotation has been proposed [5], [6]. But it considers only low-level features which are not enough to describe image semantics.

V.V. Das and J. Stephen (Eds.): CNC 2012, LNICST 108, pp. 409–412, 2012.

3 Proposed Approach

Images are segmented into different regions and features of each region are extracted. The details are appended as nodes to the graph constructed using training set. The image is annotated using graph learning method. It is proposed to extract shape context features. The entire flow described above is modeled by the figure1.

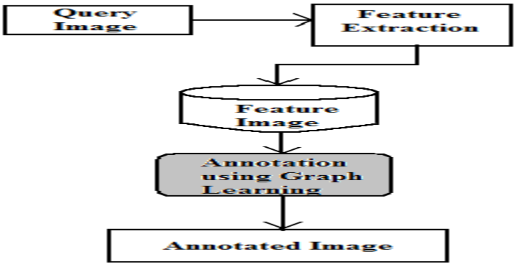

Fig. 1. System architecture

3.1 Graph Construction

Let $T = \{i_1, i_2, \ldots, i_n\}$ be training set of images, each $i \epsilon T$ is represented as visual feature f_i and $w = \{w_1, w_2, .., w_l\}$ is list of keywords. Undirected graph G=<V,E> shows relationships among images & words. Nodes are linked to its k nearest neighboring nodes based on similarity measure. $sim(f_i, f_j)$ denotes edge weight between image nodes. $sim(i, w_i)$ denotes edge weight between image and word nodes. $sim(w_i, w_j)$ denotes edge weight between word nodes. It is described by fig. 2.

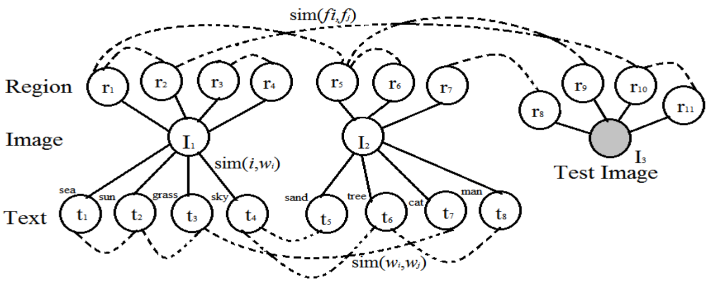

Fig. 2. Relationship among images and words

3.2 Calculating the Weights

As described in section 3.1, the weights are assigned by finding similarities using the following equations.

$$sim(f_i, f_j) = \begin{cases} \exp\left(-\frac{\|f_i - f_j\|}{\sigma^2}\right) & if\ f_i is\ KNN\ of f_j \\ 0 & otherwise \end{cases} \tag{1}$$

$$sim(w_{i,} w_j) = \frac{N(w_i, w_j)}{\min\left(N(w_i), N(w_j)\right)} \tag{2}$$

$$sim(i, w_j) = ((1-\lambda) + \lambda^* \log((1 + |w|)/df(w_j))) * \delta(i, w_j) \tag{3}$$

$N(w_i, w_j)$ is co-occurrence frequency of words w_i ,w_j , $N(w_i)$,$N(w_j)$ are occurrence frequency in training set of words w_i and w_j.|w| is annotated vocabulary size. When w_j is annotated word of image i,δ (i, w_j) is equal to 1, otherwise 0. λ is smoothing factor.

3.3 Annotating an Image Using IFRWR Algorithm

Previously systems using RWR was proposed [7] and was not scalable i.e.not suitable for large graphs. IFRWR is an iterative process as shown in the following equation.

$$R^{n+1} = (1-c)Q^{-1}Y \tag{4}$$

where Q = (*I*-*c*A). R is N dimensional vector representing transition states. N is no. of nodes. Y is N dimensional vector with all elements zero but one exception of “1” on the position which is corresponding to initial node. A is an adjacency matrix of G.

(1-c) is probability of each node back to initial node during random walking process.

Algorithm
To build the graph and annotate the image
Input : Annotated image set T={i_1,i_2,…,i_n} and the query image I_q
Output: Annotation terms of image I_q
Step 1:Let features be extracted - F={f_1,f_2,..,f_m}.Annotation word list-w={w_1,w_2,..,w_l}
Step 2:The segmented regions of I_q is denoted as $\{r_1^q, r_2^q, ..., r_N^q\}$
Step 3:Create one node for each f_i, images I_i and term w_i. Create node for image I_q
Step 4:Add edge between nodes f_i , f_j and compute sim(w_i,w_j) using the equation (2)
Step 5:Connect each image node I and its annotation word w_i , compute the edge weight using the equation (3)
Step 6:Add an edge between annotation word node w_i and its K-nearest neighbor node w_j and calculate edge weight using the equation (1)
Step 7:Initialize Y. V=0, for all its N entries except the entry of query image I_q for which it is 1
Step 8:Build adjacency matrix A and normalize columns of A
Step 9:Initialize vector R and R=Y
Step10:Use the equation (4) to find the value of R^{n+1}. Repeat this until it converges
Step11:The image I_q is annotated with words which have highest values

4 Experimental Analysis

In order to evaluate the proposed algorithm, a test is conducted on image database where 500 images are considered, in which 450 as training images and 50 as test images have been selected. It is observed that the IFRWR annotates with considerable accuracy. A comparison is made between ground truth and annotated results of proposed system as shown in the fig 3.This approach generates an average precision of 0.1501,average recall of 0.1892 and the average measure of F_{score} is calculated using the formula (7). It is estimated to be 0.1674.

$$F = 2.\frac{precision.recall}{precision+recall} \tag{7}$$

Images				
Ground Truth	Sky, Pink, Water	Lilly, Leaf	House	Puppy
Proposed Algorithm	Sky, Clouds, Pink, Water	Green leaf, Lilly, Water	House, Sky	Puppy, Grass

Fig. 3. Image annotation results

5 Conclusion

In the proposed method, initially, an undirected graph is employed to integrate correlation among low-level features, and words. Then image annotation is implemented by IFRWR which addresses scalability issue by its iterative nature. The experiments show satisfactory results of proposed algorithm. Future work includes exploring Parallel RWR and Hierarchical RWR.

References

1. Jeon, J., Lavrenko, V., Manmatha, R.: Automatic Image Annotation and Retrieval Using Cross-Media Relevance Model. In: 26th Annual International ACM SIGIR (2003)
2. Feng, S., Manmatha, R., Laverenko, V.: Multiple Bernoulli Relevance Models for Image and Video Annotation. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1002–1009 (2004)
3. Carneiro, G., Chan, A.B., Moreno, P.J., et al.: Supervised learning of semantic classes for image annotation and retrival. IEEE Transactions on Pattern Analysis and Machine Intelligence 29(3), 394–394 (2007)
4. Vasconcelos, N.: Minimum probability of error image retrieval. IEEE Transactions on Signal Processing 52(8), 2322–2336 (2004)
5. Liu, J., Li, M., Ma, W.Y., et al.: Adaptive graph model for automatic image annotation. In: Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval, pp. 61–67 (2006)
6. Pan, J.-Y., Yang, H.-J., Faloutsos, C., et al.: GCap: Graph-based Automatic Image Captioning. In: 4th International Workshop on Multimedia Data and Document Engineering (MDDE 2004), in Conjuction with CVPR 2004, pp. 146–156 (2004)
7. Guo, Y.T., Luo, B.: An Automatic Image Annotation Method Based on the Mutual K-Nearest Neighbor Graph. In: 2010 Sixth International Conference on Natural Computation, ICNC 2010 (2010)