An Efficient Query Mining Framework Using Spatial Hidden Markov Models for Automatic Annotation of Images

R.Ramya. M.E[1]

[1]*Student , Department of Computer Sciencea and Engineering, PPG Institute Of Technology ,Coimbatore. Tamil nadu, India.*

Abstract— **A novel method for automatic annotation of images is used with keywords from a generic vocabulary of concepts or objects combined with annotation-based retrieval of images. This can be done by using spatial hidden Markov model, in which states represent concepts. The parameters of this model are estimated from a set of manually annotated training images. An image in a large test collection is then automatically annotated with the a posteriori probability of the concepts. This annotation supports content-based search of the image-collection through keywords. The keyword relevance can be constructed using Aggregate Markov Chain (AMC). A stochastic distance between images based on their annotation and the keyword relevance are captured in the AMC is then introduced. Investigation has been made in the Geometric interpretations of the proposed distance and its relation to a clustering in the keyword space.**

Keywords— **Markovian semantic indexing, spatial hidden markov model, image annotation, query mining**

I. INTRODUCTION

Social networks are considered as online applications that Humans are likely to associate images with high level concepts, the present computer vision techniques extract from images frequently low-level features and the link between low-level features and high-level semantics of image content is vanished. Neither a single low-level feature nor a mixture of multiple low-level features has explicit semantic meaning in common. As well, the similarity measures between visual features do not essentially match human perception and therefore, retrieval results of low-level methods are usually unsatisfactory and frequently unpredictable. This is commonly called the semantic gap: the lack of chance between the information that one can extract from the visual data and the understanding that the same data have for a user in a given condition. Still, the retrieval process not succeed due to the sensory gap: the gap between the object in the world and the information in a computational description allotted to a recording of that object. Whereas the former gap brings in the subject of users' interpretations of images and how it is naturally complex to capture them in visual content, the final gap makes recognition from image content challenging due to drawbacks in recording and description capabilities. At present, only 10 percent of online image files have a annotation. Consequently, image search engines are only able to bring precision of around 42 percent and recall of approximately 12 percent, whereas 60 percent of search engine visitors use as a minimum two different search engines as they are not satisfied by the retrieved content. The most frequent complaint is that search engines do not know content semantics. In addition, about 77 percent of searchers modify keywords more than once since they cannot detect content. Annotation-Based Image Retrieval (ABIR) systems are an challenge to include more efficient semantic content into both image captions and text-based queries. The Latent Semantic Indexing (LSI)-based approaches that were originally applied with increased success in document indexing and retrieval were included into the ABIR systems to realize a more reliable concept association. Still, the level of success in these challenges is uncertain; a cause for this lies in the sparsity of the per image keyword annotation data in association to the number of keywords that are frequently assigned to documents. Markovian Semantic Indexing (MSI), a new method for automatic annotation and annotation based image retrieval has been introduced. The properties of MSI make it particularly appropriate for ABIR tasks when the per image annotation data is incomplete. The uniqueness of the method makes it also mainly applicable in the context of online image retrieval systems.

II. RELATED WORK

In [1] Z. Guo et.al presented a generative model which is called the citation-topic (CT) model for modeling linked documents that clearly judges the relations among documents. This model contains the substance of each document is a combination of two sources:

(1) The subjects of the given document and

(2) The subjects of the documents that are related to (e.g., cited by) the specified document.

This view in fact reflects the process of writing a scientific article: the authors perhaps first learn knowledge from the narrative and then merge their own creative facts with the learned knowledge to figure the content of the paper. Additionally, to capture the indirect relatives among documents, this model contains a generative process to choose related documents where the related documents are not essentially directly linked to the certain document. The problem occur in [1] is it requires high computational time.

In [2] T. Hofmann present a approach to LSA and factor analysis called Probabilistic Latent Se- mantic Analysis (PLSA) that has a solid statistical foundation, because it is based on the likelihood principle and describes a proper generative form of the data. It implies in exacting that traditional techniques from statistics can be applied for questions like model and complexity control, fitting and model combination. As well, the factor illustration obtained by PLSA allows to treat with polysemous words and to clearly differentiate between different meanings and types of word usage. However the expected performance on unseen test data is not performed. And the documents are not contained in the original training collection.

In [3] F. Monay presented and compares two easy latent space models usually used in text analysis, that is to say Probabilistic LSA (PLSA) and Latent Semantic Analysis (LSA). Annotation strategies for each form are examined. Outstandingly, 8000-image dataset found a classic LSA model described on keywords and a very essential image representation performed as well as much more complex. However Annotations in test data might not contain some correct keywords. Then for LSA and direct match, no probability is close to each ranked keyword; hence the threshold level cannot be applied in a straight way.

In [4] T. Hofmann presents confusion results for different types of linguistic data and text collections and then discusses an application in automated document indexing. The experiments indicate substantial and consistent improvements of the probabilistic method over standard Latent Semantic Analysis. However Limited computational resources are used.

In [5], R. Datta et al we surveyed almost 300 key empirical and theoretical contributions in the current decade associated to automatic image annotation and image retrieval then discuss the spawning of associated subfields in the process. It discusses important challenges concerned in the adaptation of existing image retrieval techniques to construct systems that can be helpful in the real-world. According to that, it has been achieved so far and conjectured what the future may contain for image retrieval investigation. The original depiction of an image, which is an array of pixel values matches inadequately to visual response. However this process requires high computational complexity.

III. PREVIOUS WORK

A. ANNOTATION BASED IMGE RETIEVAL

Annotation-Based Image Retrieval (ABIR) systems are an effort to include more proficient semantic content into both image captions and text-based queries. The Latent Semantic Indexing (LSI)-based approaches that were originally used with improved success in document indexing and retrieval were included into the systems of ABIR to determine a more reliable concept association.

B. MARKOV CHAIN BASED IMAGE INDEXING

Markov chain based image indexing (MSI) has been used for online image retrieval system. Considering the users' queries to construct an Aggregate Markov Chain (AMC) through which the relevance between the keywords obtained by the system. The inference engine of the Markov chain based image indexing approach lies in the clustering of the state space, because this clustering organizes the states into collections of relevance. In case of scalability, the degree of this clustering has been examined according to the size of the system. The clustering state degree in Markov Chains has been systematically studied, referred as the coupling degree. This degree quantifies the degree of state clustering in the chain and this can be assumed in terms of state connectivity..

`

IV. PROPOSED SYSTEM

The proposed framework in fig 1shows the flow of work and consist of following stages:

- ✓ Query relatedness with image
- ✓ Aggregate Markovian Chain
- ✓ Optimization step

A. QUERY RELATEDNESS WITH IMAGE

The users search for images by issuing queries, each query being an ordered set of keywords. The system responds with a list of images. The user can download or ignore the returned images and issue a new query instead. During the training phase of the system the images are considered with no annotation. As the users issue queries and pick images the system annotates the images in an automatic manner and at the same time establishes relevance relations between the keywords.

The user implicitly relates the retrieved (downloaded) images to her/his query. By assuming Spatial Hidden Markov Model (SHMM) in the order of the keywords the aim of the proposed approach is to quantify logical connections between keywords.

The Spatial Hidden Markov Model (SHMM) is a two dimensional generalization of traditional hidden Markov model (HMM). A SHMM λ is a 4-tuple $\lambda = (H, V, B, \pi)$, which specifies the number of states N, the number of observation symbols M, and the four probability measures: H(horizontal state transition matrix), V (vertical state transition matrix), B (observation symbol probability distribution) and π (initial state probability distribution). In SHMM, each image is divided into uniform blocks, from which image feature vectors are extracted. And each concept class of images is represented by a statistical model. An unknown image is classified to a specific model via estimating its feature vectors within blocks and the spatial relationship across blocks. Moreover, a concept class is associated with several keywords. The image blocks are also subsequently annotated with semantic labels.

$P(q_{x,y}|Q_{x,y-1}) = P(q_{x,y}|q_{x-1,y}, q_{x,y-1})$

Where $q_{x,y}$ denotes the state of block (x, y), and $Q_{x,y-1}$ denotes the state sequence of $q_{1,1}, q_{1,2}, \ldots, q_{x,y-1}$.

B. AGGREGATE SHMM MARKOVIAN CHAIN

Trying to compare directly the probability vectors πi and πj is a equilibrium state vector for two images, one faces the zero-frequency problem. By itself, the fact that a user puts certain keywords together in a query implicitly renders the keywords relative to each other regardless of the images that may or may not be picked by this user. We propose to use this and address the zero-frequency problem by clustering the keyword space into similar keywords. For this purpose, the Aggregate Markovian Chain AMC of all the queries asked by all users regardless of the selected images, is constructed in this step. The kernel of this process denoted by PG, is calculated in a similar to the previous step manner by the recurrent formula of (1). PG, even though a Markov kernel it will be used to cluster the keyword space rather than estimating an explicit probability distribution, hence the purpose of the AMC is to model keyword relevance.

C. OPTIMIZATION STEP

The AMC will be used to cluster the keyword space and define explicit relevance links between the keywords by means of this clustering. This clustering task is linked to the convergence characteristics of the AMC chain by evaluating the series $FG(n) = \sum_{k=0}^{n} P_G^k$ where PG is the AMC kernel. A suitable termination condition stops the series at the desired n where the slow convergence has taken over, but not before the rapid convergence has finished. The value of the determinant of $FG(n)$ is used as a termination condition since the clusters in the rows of $FG(n)$will drop its rank and the determinant will become close to zero. $FG(n)$ is the n-step expected occupancies matrix. An optimization task is related to this procedure with respect to the total variance of the columns of $FG(n)$, when projected on the direction of the eigenvectors of PG.

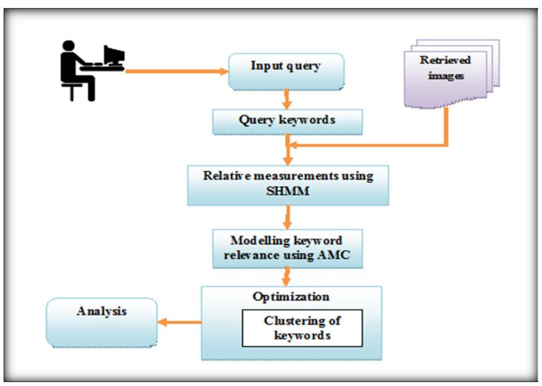

Figure 1: Proposed system achitecture

V. EXPERIMENTAL RESULTS

The proposed method is compared with the markovian semantic indexing. Comparison with MSI and SHMM in the application area of Annotation-Based Image Retrieval with Precision versus Recall diagrams on ground truth databases reveal that the proposed approach achieves better retrieval scores.

Precision

The precision rate is defined as the ratio of the number of relevant images retrieved and total number of images in the collection.

Precision=$\frac{|\{\text{relevant images}\} \cap \{\text{retrieved images}\}|}{|\{\text{retrieved images}\}|}$

Recall

Recall rate is defined as the ratio of number of relevant images retrieved and to the total number of relevant images in the collection.

Recall =$\frac{|\{\text{relevant images}\} \cap \{\text{retrieved images}\}|}{|\{\text{relevant images}\}|}$;

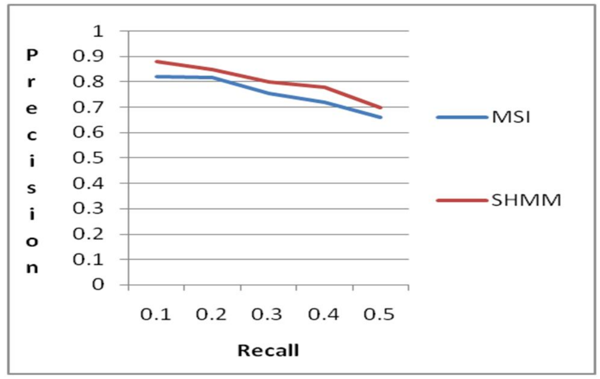

Figure 2: Precision –Recall comparision graph

Thus the above graph in figure 2 shows that proposed system of SHMM provides higher accuracy when compared with existing method of Markovian semantic Indexing (MSI).

VI. CONCLUSION

In the present work Spatial Hidden Markov Model (SHMM), a new mining process for user queries by defining connectivity measure as keyword relevance between Markovian states modeled after the user queries is proposed. This is dynamically trained by the queries of the identical users that will be provided by the system. As a result, the targeting is more accurate, when compared to other systems which use external means of no dynamic or no adaptive nature to define keyword relevance. The distance is in the form of a generalized Euclidean distance, which was constructed by means of an Aggregate Markovian Chain and established to be optimal regarding definite Markovian connectivity measures that were identified for this reason. A comparison to MSI and Proposed work has been done. Experiments have shown that proposed system achieves better retrieval results in sparsely annotated image data sets.

REFERENCES

[1] Z. Guo, S. Zhu, Y. Chi, Z. Zhang, and Y. Gong, "A Latent Topic Model for Linked Documents," Proc. 32nd Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2009.

[2] T. Hofmann, "Probabilistic Latent Semantic Indexing," Proc. 22nd Int'l Conf. Research and Development in Information Retrieval (SIGIR '99), 1999.

[3] F. Monay and D. Gatica-Perez, "On Image Auto-Annotation with Latent Space Models," Proc. ACM Int'l Conf. Multimedia (MM), 2003.4 .T. Hofmann, "Unsupervised Learning by Probabilistic Latent Semantic Analysis," Machine Learning, vol. 42, no. 1/2, pp. 177-196, 2001.

[4] T. Hofmann, "Unsupervised Learning by Probabilistic Latent Semantic Analysis," Machine Learning, vol. 42, no. 1/2, pp. 177-196, 2001.

[5] R. Datta, D. Joshi, J. Li, and J.Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age," ACM Computing Surveys, vol. 40, no. 2, pp. 1-60, 2008