ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization *Volume 3, Special Issue 5, July 2014*

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th -18th July -2014

Query Mining for Automatic Annotation and Annotation Based Image Retrieval Using Hidden Markov Model

Shahidha M Meeran, Bineesh V

Student, Dept of Computer Science, MEA Engineering College, Perinthalmanna, Kerala, India [1]

Assistant Professor, Dept of Computer Science, MEA Engineering College, Perinthalmanna, Kerala, India [2]

ABSTRACT: This paper introduces a method for automatic annotation of images with keywords from a generic vocabulary of concepts or objects for the purpose of annotation based retrieval of images. Automatic annotation of image can be done by using hidden Markov model, whose states represent concepts. The parameters of the model are estimated from a set of training images. Each image in a large test collection is then automatically annotated with the a posteriori probability of concepts present in it. This annotation supports annotation based search of the image-collection via keywords. The relevance of keyword can be constructed using Aggregate Markov Chain (AMC). A stochastic distance between images, based on their annotation and the keyword relevance captured in the AMC is then introduced. Geometric interpretations of the proposed distance are provided u and its relation to a clustering in the keyword space is investigated. We can use WordNet based context vectors for finding similarity between words and also find the similarity between the images. Then the images which has maximum probability to match with the query is retrieved.

KEYWORDS: Hidden Markov Model , Markovian semantic indexing, image annotation, query mining, annotation-based image retrieval .

I. INTRODUCTION

Image annotation refers to the labelling of images with a collection of predefined keywords. It is mainly used for visual information management and can be applied in different of domains such as entertainment, commerce, education, biomedicine, military, web image classification and search ,etc. In particular, image annotation can aid in image retrieval since annotated keywords greatly shrink the semantic gap between low-level features and high-level semantics. Automatic image annotation is a challenging area due to various imaging conditions, complex and hard-to-describe objects, a highly textured background, and occlusions. Image retrieval procedures can be divided into two approaches: query-by-text (QbT) and query-by-example (QbE). In QbT, queries are texts and targets are images. That is, QbT is a cross-medium retrieval. In QbE, queries are images and targets are images. Thus, QbT is a mono-medium retrieval. For practicality, images in QbT retrieval are often annotated by words. When images are sought using these annotations, such retrieval is known as annotation-based image retrieval (ABIR). In contrast, annotations in a QbE setting are not necessary, although they can be used. The retrieval is carried out according to the image contents. Such retrieval is known as content-based image retrieval (CBIR)

.

There are two reasons to support ABIR. First, CBIR has its own problems, which are probably more crucial. Second, the

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th -18th July -2014

negative effects due to the above problems in ABIR may be mitigated. It is obvious that there are many applications where the use of CBIR is advantageous. Among two major problems in ABIR, let us first consider the problem of subjectivity in annotations from the perspective of context. Actuality, subjectivity in annotations may not always cause difficulties in ABIR. Subjectivity implies that the annotations contain some contextual information derived from the annotators' view on the images. The subjectivity of annotation, which can be regarded as an obstacle to ABIR, can be useful if we can model the annotations appropriately. The word information used in IR, such as word co-occurrence frequencies, is often sparse.

In annotated images, the occurrences of words are especially limited because they must be assigned only for indexing purposes and the need for such extra effort is not appreciated. The problem of word sparseness may be mitigated by incorporating external knowledge such as thesauri that explicitly identify the relationships between words. Because of the cost, however, annotations are sometimes assigned automatically. Two types of methods are frequently used to assign textual information to images. One method is based on information extraction techniques. The other method is based on classification techniques. ABIR is powerful because it can utilize the power of natural language to represent users' search needs and the semantic contents of images. In image annotation the problem of recognizing all objects in a given image is very difficult due several invariance issues. There are two aspects of the image search and retrieval problem which make it relatively much more tractable. One is the availability of side information in the accompanying text. The other is that the search problem is far more tolerant of erroneous inferences than is usually assumed necessary for object recognition in individual images. For this first develop a generative stochastic model for images and their accompanying captions for Hidden Markov Model. Training material for such a model,is readily obtained from naturally occurring image+caption pairs. The vocabulary of the words used in the captions is used to annotate the images that may not have surrounding text. If surrounding text is available, this model is capable of using it as a prior, and computing posterior probabilities of the words in the caption given the visual evidence in the image.

Annotation-Based Image Retrieval (ABIR) systems are an attempt to incorporate more efficient semantic content into both text-based queries and image captions. Here introduce the Markovian Semantic Indexing (MSI), a method for and annotation based image retrieval. The properties of MSI make it particularly suitable for ABIR tasks when the per image annotation data is limited. In this system the users' queries are used to construct an Aggregate Markov Chain (AMC) through which the relevance between the keywords seen by the system is defined. Find different senses of user queries by using ontology database WordNet which combines gloss information of concepts with semantic relationships, and organizes concepts as high-dimensional vectors. Also find the probability of matching between annotated images and user queries, images are ranked based on markovian distance, Then the images which has maximum probability to match with the query is retrieved. The remainder of this paper is organized as follows. Section 2 describes the hidden Markov model (HMM) used for image annotation. Section 3 discuss about Annotation based image retrieval using MSI Section 4 presents some experimental results. Section 5 concluded with some discussion.

II. HIDDEN MARKOV MODELS FOR IMAGE ANNOTATION

We develop a generative stochastic model for images and their accompanying captions. Let a collection of image+caption pairs be provided and consider the problem of developing a stochastic generative model that jointly describes each pair. HMM first estimate the parameters from annotated image+caption pairs. Let I be an image, and C = {c1,c2, . . . , cN} denote the objects (concepts) present in that image, as specified by the corresponding caption C. For each image region it , t = 1, . . . , T, let xt ЄIR^d represent the color, texture, edges, shape and other visual features of the image. Let V denote the global vocabulary of the caption-words cn across the entire collection of images. We propose to model the {xt} vectors of an image I as a hidden Markov process [11], generated by an unobserved Markov chain whose states st take values in C. Specifically, each xt is generated pursant to some probability density function f(·|st) given the state st, where st is a Markov

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th -18th July -2014

chain with a known initial state s0 and transition probabilities p(st|st-1). For aligning image-regions with caption-words in an image+ caption pair, then find the state transition diagram of Markov chain of each image which generates these regions. Computing the of features present in each image by computing the joint likelihood features present in each state of an image [11]. Given an image+caption pair (I, C), the most likely alignment of image regions with caption-words is
T

$$\hat{s}_1^T = \arg\max = \Pi t=1\ f(xt|stt)\ p(st|st-1).\ _1^T \in C^T$$

Finally, the marginal likelihood of an image may be computed as
T

$f(x_1^T|s0) = f(x1\ ,V\ |s0)$ When an image with no caption is provided, this marginal likelihood enables one to use the HMM to compute the conditional probability that a particular image region it was generated by a particular concept c€V based on the total visual features $x1^T$ in the image I:

$$p(st = c|x1^T, s0) = f(x1^T, st = c|s0)/f(x1^T|s0).$$

Finally computing the probability of a caption-word being present in an image. T

$p(c€C|I, s0)\ \alpha\Sigma t=1^T\ p(st = c|x1\ , s0)$. In automatic annotation system , Once the visual features of the images are extracted and the HMM for each training image defined, estimation of the HMM parameters proceeds in a standard manner as prescribed in [12]. Once the output pdf's f(x|c) of all the states c 2 V have been estimated from the training set (annotated image+caption pairs), we proceed to decoding the HMM. This is simply a fully connected the vocabulary to HMM, one state of model corresponding to each concept in the vocabulary. Transition probabilities of the HMM are set to be uniform. Given test image, extract the visual features of test image, we perform the Balm-Welch algorithm that computes the posterior probability of (6). Once this is done for each image in the test collection, then system ready for image retrieval experiments. Note that the transition probabilities of the decoding HMM need not be uniform. . A simple variation is to compute the co-occurrence statistics of the words in the concept vocabulary, and use that frequency to adjust the transition probabilities. For instance, deer and grass to co-occur much more often than deer and water, suggesting that p(deer| grass) could be set higher than p(deer| water). In a contrastive experiment, we set

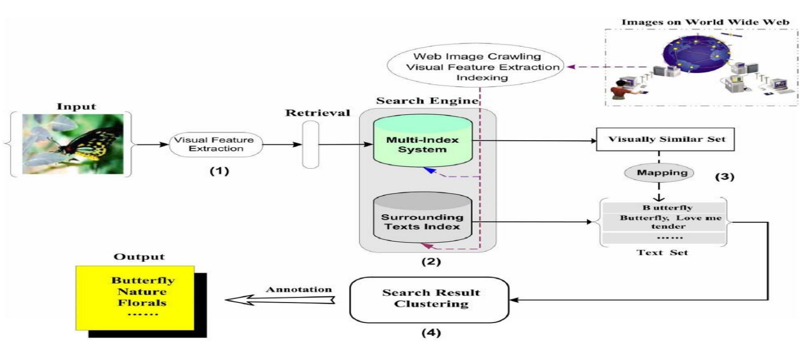


Figure 1: Automatic Image Annotation System these probabilities to be proportional to the corresponding relative frequencies, with adequate smoothing to avoid zero-probabilities.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

IJIRSET

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th -18th July -2014

III. ANNOTATION BASED IMAGE RETRIEVAL

Annotation-Based Image Retrieval systems consolidate more efficient semantic content into both text-based queries and image captions. The proposed approach will be introduced in the framework of an image retrieval system where users search for images by submitting queries that are made of keywords. Taking the user query as input to the system, find the relevance of the keywords by constructing it as AMC and measuring the semantic similarity of the keyword by using WordNet [17] database. Also discover the probability between the query and images in annotated image database. Then find the similarity between matching images by calculating the distance between them using MSI concept. Ranking the images based on this distance, and finally system responds with a list of most ranked images. The queries formed by the users of a search engine are

meaningfully refined, the keywords representing brief semantics when compared to text in documents or other vocabulary related presentations. The aim is to improve user satisfaction by returning images that have a higher probability to be accepted by the user. The expectation is that the users search for images by issuing text queries, each query being an ordered set of keywords. The system responds with a list of images.

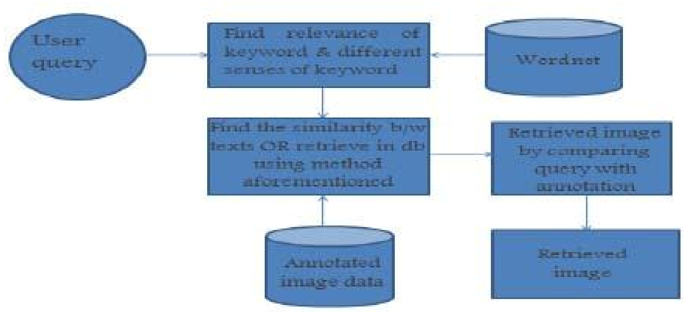


Figure 2: Image retrieval system

By assuming Markovian chain transitions in the order of the keywords the aim of the proposed approach is to quantify logical connections between keywords. If some user relates image Ii to his query qi, where keyword k2 follows keyword k1 and this occurs m times, then the one step transition probability pi(k1; k2) is being updated as follows: if pi(k1; k2) is the current probability based on M keywords then the new probability (based on M+ m keywords) is calculated by the recurrent formula pi(k1; k2)= M pi(k1; k2)+ m / M + m (1)

For finding different senses of keywords , a new measure based on the semantic ontology database WordNet is proposed which combines gloss information of concepts with semantic relationships, and organizes concepts as high-dimensional vectors. In this measure, the sense of a concept is considered highly related to the concepts in its context. Concepts frequently occurring in similar contexts are often conceptually similar. This measure use WordNet to retrieve contexts for senses because it is a general ontology with rich information about senses and their relations. That means the WordNet give synonyms of the word or phrase which we are gave. For deciding relevance of keyword, this procedure constructs a Markov chain where each keyword corresponds to a state. Each time a keyword appears in a query, its state counter value is forwarded; if another keyword coming in the same query, their interstate link counter is also forwarded. The formations of the keywords but also the sequencing of these formations are both measured in this way. The queries relating to an image are batch processed for this image, the counters are advanced, the probabilities are updated as above. Find the probability vector of each image with given query. Trying to compare directly the probability vectors of two images πi and πj

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

IJIRSET

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th -18th July -2014

respectively calculated in the previous step, one faces the zero-frequency problem. For this purpose, the Aggregate Markovian Chain (AMC) [10] of all the queries stated by all users regardless of the selected images, is constructed in this step. The kernel of this process represented by PG, is calculated in a similar to the previous step manner by the recurrent formula of (1), hence the purpose of the AMC is to model keyword relevance. The AMC will be used to group the keyword space and define explicit relevance links between the keywords by means of this grouping. The typical definition of the Markov Semantic Indexing (MSI) distance can be Let x and y two images represented by their respective steady state probability row vectors πx and πy. Let also $\Sigma(FG^T)$ be the covariance matrix of the zero-mean transpose expected fractional occupancies matrix of the Aggregate Markov Chain (AMC), calculated at the desired n. Then the MSI distance between images x and y is defined as

$$D(x, y) = (\pi x-\pi y)\,\Sigma(FG^T)\,(\pi x-\pi y)^T$$

$$=\delta^T_{xy}\,\Sigma(FG^T)\,\delta xy$$

where the dimensionality of πx and πy has been extended to that of $\Sigma(FG^T)$ by filling in with zeros the respective coordinates. Markovian Semantic Indexing give similarity between the images by finding Cross Passage Time (CPT)
[24] between the states. Geometric interpretation and optimality properties of the MSI distance described in [24]. Images are ranked based Most ranked images almost satisfy the user's expectation that are returned to the user.

IV. EXPERIMENTAL RESULTS

The COREL Image Dataset consists of 5000 images from 50 Corel Stock Photo CD's provided to us by [4]. Each CD contains 100 images with a certain theme (e.g. polar bears), of which 90 are designated to be in the training set and 10 in the test set, resulting in 4500 training images and a balanced 500-image test collection. We note that this collection and training/test division is also that used by [7] and [5]. Experimental result of automatic image annotation system is given below:

Figure 4.1: Annotated Images from HMM

Results of Annotation based image retrieval system:

5.2 Performance Matrices

The proposed method is compared with the Markovian semantic indexing. Comparison with MSI and HMM 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.

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

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

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization *Volume 3, Special Issue 5, July 2014*

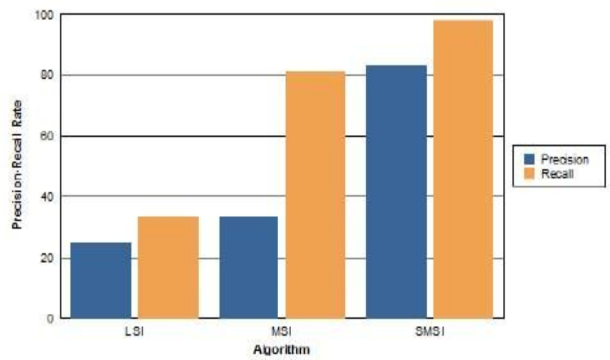
International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th -18th July -2014

|{relevant images}| :

` The experimental results for proposed system are plotted on graphs based on these formulas. The graphs shown in the figures below give better analysis perspective on the Automatic annotated online image retrieval task.



The above graph in figure 5.1 concluded that the Precision-Recall rate of SMSI is increased which will be the best one.

V. CONCLUSION

We have presented a system for annotation based image retrieval with an automatic image annotation system using HMM. The proposed the Markovian Semantic Indexing, a method for mining user queries by finding keyword relevance as a connectivity measure between Markovian states modelled after the user queries. Also find different senses of keywords by using WordNet database. The proposed system is dynamically trained by the queries of the users that will be served by the system. Accordingly, the targeting is more accurate, compared to other systems that use external means of non dynamic or non adaptive nature to define keyword relevance. A stochastic distance, was constructed by means of an Aggregate Markovian Chain and proved to be optimal with respect to certain Markovian connectivity measures that were defined for this purpose. The system was implemented with publicly available tools and tested on standard datasets.

REFERENCES

[1] A. Amir et al. IBM Research TRECVID-2003 Video Retrieval System. In Proc. TRECVID2003, November 2003.
[2] K. Barnard, P. Duygulu, N. de Freitas, D. Forsyth, D. M. Blei, and M. I. Jordan. Matching words and pictures. Journal of Machine Learning Research, 3:1107–1135, 2003.
[3] D. M. Blei and M. I. Jordan. Modeling Annotated Data. In 26th Annual International ACM SIGIR Conference, pages 127–134, 2003.
[4] P. Duygulu, K. Barnard, N. de Freitas, and D. Forsyth. Object Recognition as Machine Translation: Learning a Lexicon for a Fixed Image Vocabulary. In Seventh European Conference on Computer Vision, volume 4, pages 97–112, 2002.
[5] S. L. Feng, R. Manmatha, and V. Lavrenko. Multiple Bernoulli relevance models for image and video annotation. In Proc. IEEE Conf. On Computer Vision and Pattern Recognition, volume 2, pages II–1002–II–1009, 2004.
[6] G. Iyengar et al. Joint Visual-Test Modeling for Multimedia Retrieval. Available at: http://www.clsp.jhu.edu/ws2004/groups/ws04vstxt/, 2004.
[7] J. Jeon, V. Lavrenko, and R. Manmatha. Automatic Image Annotation and Retrieval using Cross-Media Relevance Models. In 26th Annual International ACM SIGIR COnference, pages 119–126, 2003.
[8] V. Lavrenko, S. L. Feng, and R. Manmatha. Statistical models for automatic video annotation and retrieval. In Proc. IEEE International Conf. On Acoustics, Speech and Signal Processing, volume 3, pages 17–21, May 2003.
[9]. Lorenz S, Korolkova N, Leuchs G. Continuous variable quantum key distribution using polarization encoding and post selection. *Applied Physics* 2004; **B 79**: 273--277.

[10]. Shields A, Yuan Z. Key to the quantum industry. *Physics World* 2007; **20**(3): 24--29.
[11] Arnab Ghoshal, Hidden Markov Models for Automatic Annotation and ContentBased Retrieval of Images and Video. Copyright 2005 ACM 1595930345/05/0006
[12]K. Stevenson and C. Leung, "Comparative Evaluation of Web Image Search Engines for Multimedia Applications," Proc. IEEE Int'l Conf. Multimedia and Expo, July 2005.
[13] 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.
[14]J. Li and J. Wang, "Real-Time Computerized Annotation of Pictures," Proc. ACM 14th Ann. Int'l Conf. Multimedia, 2006. [8] D. Joshi, J.Z. Wang, and J. Li, "The Story Picturing Engine – A System for Automatic Text Illustration," ACM Trans. Multimedia Computing, Comm. And Applications, vol. 2, no. 1, pp. 68-89, 2006.nd
[15]T. Hofmann, "Probabilistic Latent Semantic Indexing," Proc. 22Int'l Conf. Research and Development in Information Retrieval (SIGIR '99), 1999.
[16] *Feiyang Yu*1 *and Horace H S Ip* Automatic Semantic Annotation of Images Using Spatial Hidden Markov Model , Image Computing Group,Department of Computer Science Center for Innovative Applications of Internet and Multimedia Technologies (AIMtech Centre) City University f Hong Kong, HONG KONG, 1424403677/ 06/$20.00 ©2006 IEEE
[17] Measuring Semantic Similarity Using WordNet Based Context Vectors
[18] J. Fan and Y. Gao, and H. Luo, "Integrating Concept Ontology and Multitask Learning to Achieve More Effective Classier Training for multilevel Image Annotation," IEEE Trans. Image Processing, vol. 17, no. 3, pp. 407-426, Mar. 2008.
[19] Konstantinos A. Raftopoulos, Member, IEEE, Mining User Queries with Markov Chains:Application to Online Image Retrieval, IEEE transactions on knowledge and data engineering, vol. 25, no. 2, february 2013