IJREB

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 (Issue 4) Pages (27-36)

Journal home page: www.ijreb.org

AUTOMATIC IMAGE ANNOTATION: A NEW TECHNIQUE FOR IMAGE RETRIEVAL USING MSI

Jinesh Melvin. Y.I and Kathiresan. V.

Department of Computer Science and Engineering, Narayanaguru College of Engineering, Manjalumoodu, K.K Dist, Tamil Nadu

ABSTRACT

The Latent Semantic Indexing (LSI) and Probabilistic Latent Semantic Indexing (pLSI) in the annotation based image retrieval (ABIR) task it gives better result. In the existing systems technical operations for online image retrieval are of i) automatic annotation ii) indexing iii) annotation based retrieval of images. The keyword space with the images and their annotation also the keyword relevance are captured in the AMC. The problems addressed with these are it brings large gap between the users query and the image description. It gives confused state and collaboration in selecting documents, also limitation in image description. Here to introduce the new method called Markovian chain with ranking method. A stochastic distance between images, based on their annotation and the keyword relevance captured in the AMC. By means of a new measure of Markovian state similarity, the mean first cross passage time (CPT), optimality properties of the proposed distance are proved. This method is shown to possess certain theoretical advantages. Optimal ranking is set through a counter for each customer response process.

KEYWORDS: Markovian semantic indexing, image annotation, query mining, annotation-based image retrieval

IJREB

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 (Issue 4) Pages (27-36)

Journal home page: www.ijreb.org

INTRODUCTION

Even though humans tend to associate images with high level concepts, the current computer vision techniques extract from images mostly low-level features and the link between low-level features and high-level semantics of image content is lost. Neither a single low-level feature nor a combination of multiple low-level features has explicit semantic meaning in general. In addition, the similarity measures between visual features do not necessarily match human perception [1] and, thus, retrieval results of low-level approaches are generally unsatisfactory and often unpredictable.

This is what is called the semantic gap: the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation. However, the retrieval process fails also due to the sensory gap: the gap between the object in the world and the information in a (computational) description assigned to a recording of that object. While the former gap brings in the issue of users' interpretations of images and how it is inherently difficult to capture them in visual content, the latter gap makes recognition from image content challenging due to limitations in recording and Description capabilities. Currently, only 10 percent of online image files have a professional description (annotation). Annotation-Based Image Retrieval (ABIR) systems are an attempt to incorporate more efficient semantic content in to both text-based queries and image captions (i.e. Google Image Search, Yahoo! Image Search). The Latent Semantic Indexing (LSI)-based approaches that were initially applied with increased success in document indexing and retrieval were incorporated into the ABIR systems to discover a more reliable concept association. However, the level of success in these attempts is questionable; a reason for this lies in the sparsity of the per-image keyword annotation data in comparison to the number of keywords that are usually assigned to documents. Markovian Semantic Indexing (MSI), a new method for automatic annotation and annotation based image retrieval. The properties of MSI make it particularly suitable for ABIR tasks when the per image annotation data is limited. The characteristics of the method make it also particularly applicable in the context of online image retrieval systems.

IJREB

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 (Issue 4) Pages (27-36)

Journal home page: www.ijreb.org

RELATED WORKS AND CONTRIBUTION

Growth in content-based retrieval has been unquestionably rapid. In the recent years, more than 200 content-based retrieval systems have been developed, the majority which are based on low level features. In particular, they can be classified into two main categories: 1) those that perform semantics mining based on the analysis of textual information associated to images, such as annotations, assigned keywords, captions, alternative (alt) text in html pages or surrounding text, and 2) those that are based on the extraction of low-level visual features such as color, texture in order to perform alignment, classification, browsing, searching, summarization, etc. in image collections. Methods of the first category depend on laborious annotation, while the latter methods usually cannot effectively capture semantics. Additionally, some other techniques use both low-level features in the form of visual keywords and text annotation to perform content based operations but they usually demand the explicit involvement of users for linguistic annotation of pictures. Annotation-Based Image Retrieval systems incorporate more efficient semantic content into both text-based queries and image captions. A direct consequence is that methods initially developed for document retrieval may be suitable for ABIR systems, as well. Latent Semantic Indexing was initially developed for document retrieval. Attempts of applying LSI/pLSI-based techniques discover a more reliable concept association in ABIR systems has been reported. In the computer vision literature joint models over image and text have been already explored. The approach proposed herein, while stochastic in nature, raises the reasoning aspect of probabilities, as it defines explicit relevance relationships between keywords.

Contribution

The methodology proposed in this work encompasses a novel (alternative) probabilistic approach for Annotation- Based Image Retrieval that, compared to LSI and pLSI, is better suited to sparsely annotated domains, like in image databases where, the per image sparse keyword annotation is also limited. It addresses in a more natural way the zero frequency problems, defined as the fact that the probability to find common keywords even in closely related images is typically small because the images are not annotated with exactly the same keywords. This problem is addressed here by means of an explicit relevance link between keywords that carries a probabilistic weight. We show that assigning logical connections between keywords by means of a Markovian model, permits better generalization over a sparsely annotated domain hence the proposed approach raises the reasoning aspect next to the numerical aspect of probabilities. The key idea behind the approach is to compensate for the sparse data by incorporating an annotation procedure of probabilistic qualitative reasoning that will propagate partial beliefs regarding connections between keywords. A mechanism that gains performance from mining the structure of the existing data rather that incorporating new data, as it happens with traditional models, is hence introduced.

Furthermore, even though automatic annotation and annotation-based image retrieval systems have been presented in the literature the proposed system is novel in the way it unifies these two tasks. Indeed both the automatic annotation and retrieval tasks are assumed in the implicit user interaction context, for dynamically mining semantics towards qualitative probabilistic reasoning. This has implications in the targeting aspect since the training is performed dynamically by the same users that are actually using the system. The unified Markovian setup behind the proposed system allows the retrieval technique to benefit from the underlying structure of the annotation data; at the same time the annotation data acquires concrete stochastic interpretation through the way it is treated by the retrieval process. A comparison with LSI and pLSI in the application area of ABIR with Precision versus Recall diagrams on ground truth databases reveal that the proposed approach achieves better retrieval scores. A new measure of similarity between Markovian states, the expected first Cross Passage Time (CPT) is defined and its properties are examined. MSI is shown to be optimal with respect to CPT. Conceptual comparisons with LSI but also with pLSI reveal advantages for the proposed approach.

THE PROPOSED APPROACH

A vector space is available on the basis of which will propose a distance that renders certain Markovian connectivity measures and has an optimal stochastic interpretation. The proposed approach will be presented in the framework of an online image retrieval system (similar to Googleimage search) where users search for images by submitting queries that are made of keywords. The queries formed by the users of a search engine are semantically refined, the keywords representing concise 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 (downloaded) by the user. The assumption is that the user 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 as will be explained later on in the manuscript. The user never annotates the images explicitly, this happens by the system transparently from the user. At the testing phase the system uses the annotations available from the training phase but also the keyword relevance probability weights also evaluated during the training phase to return images that better reflect the users preferences and improve user satisfaction. This interactive procedure has implicit consequences that we exploit one by one in a step by step construction of the proposed system:

Step 1: The user implicitly relates the retrieved (downloaded)images to her/his query. By assuming Markovianchain 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 I_i to hisqueryq_i, where keyword K_2follows keyword K_1 and thisoccurs m times, then the one step transition probability$p_i(K_1, K_2)$is being updated as follows: if $p_i(K_1, K_2)$is thecurrent probability (before the update) based on M keywords then the new probability (based on $M + m$ keywords)is calculated by the recurrent formula

$$p_i(K_1, K_2) = \frac{Mp_i(K_1, K_2) + m}{M + m} \qquad (1)$$

This procedure constructs a Markov chain where each keyword corresponds to a state. Each time a keyword appears in a query, its state counter is advanced; if another keyword follows in the same query, their interstate link counter is also advanced. The occurrences of the keywords but also the sequencing of these occurrences is both measured this way. The queries pertaining to an

ISSN 2321-743X
International Journal of Research in
Engineering and Bioscience
Volume 2 (Issue 4) Pages (27-36)
IJREB
Journal home page: www.ijreb.org

image are batch processed for this image, the counters are advanced, and the probabilities are updated as above. Before the next set of queries is processed, the counters are cleared. The equilibrium state vector of such constructed Markov Chain, for each image I_i, will be denoted by $\pi_i$and will representthe image from now on. This modeling approach isjustified: 1) by the qualitative character of the conceptualapproach that is adopted in this work, and 2) the targeting aspects of the MSI, whereby the objective is to capture user specific aspects (e.g., perception of images). Indeed, the fact that each sequence of keywords (query) comes from specific user, filtered through her/his perception about the selected image, justifies the assignment of a logical connection of relevance between these two keywords, in addition to the individual connections between each keyword and the selected image. It is this logical connection that is favored by our modeling approach rather than computing the distribution of images over keywords in the traditional numerical fashion.

Step 2: Trying to compare directly the probability vectorsπ_i and π_j calculated in the previous step 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. Here the 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 P_G, is calculated in a similar to the previousstep manner by the recurrent formula of (1). P_G, eventhough a Markov kernel it will be used to cluster thekeyword space rather than estimating an explicit probability distribution, hence the purpose of the AMC is to model keyword relevance.

Step 3: 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 $F_G(n) = \sum_{k=0}^{n} P_G^k$where $P_G$the AMC kernel is. A suitable terminationcondition stops the series at the desired n where the slow convergence has taken over, but not before the

IJREB

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 (Issue 4) Pages (27-36)

Journal home page: www.ijreb.org

rapidconvergence has finished. The value of the determinant of$F_G(n)$is used as a termination condition since the clusters inthe rows of $F_G(n)$will drop its rank and the determinantwill become close to zero. $F_G(n)$is the n-step expectedoccupancies matrix (Appendix). An optimization task is related to this procedure with respect to the total variance of the columns of $F_G(n)$, when projected on the direction ofthe eigenvectors ofP_G. Further insight in this concept willbe gained in the next section. From now on, we will use justF_G to note$\frac{1}{n+1}F_G(n)$, the n-step expected fractional occupanciesmatrix (Appendix), calculated at the desired n.

Step 4: The formal definition of the MSI distance can beprovided at this step.

GEOMETRIC INTERPRETATIONS AND OPTIMALITY PROPERTIES OF THE MSI DISTANCE

The proposed MSI distance d(x, y) can be viewed as measuring the total variance of the rows of $F_G^T$when projected on the direction of the difference, $(\pi_x - \pi_y)$of the two images. The F_G^T is calculated once from all thedata and thus the direction defined by the vector differenceof the probability distributions of the two particular images is actually deciding their in between distance. To get furthergeometric and stochastic interpretations of the MSI distance one has to acquire insight into the mechanics of the Markovian convergence that produces the $F_G^T$and itsrelation to certain directions in the keyword space. The geometricalmeaning of the convergence process and its relation tothe proposed distance with respect to state clusters and state connectivity measures. Standard terminology, concepts and notation from the stochastic processes literature will be used.

The Mean First Cross Passage Time

The Mean First Cross Passage Time has arranged certain concepts regarding the directions of convergence and their stochastic meaning we will attempt to connect these concepts to the mechanics of the proposed MSI distance. First, we need a connectivity measure between the Markovian states that will serve as the explicit probabilistic relevance link between the keywords. We can see the states 1 and 2 being isolated from each other but identical as far as the process is concerned after starting from these states (from fig). We cannot define any metric that involves directly these two states, like the first passage time, since the two

IJREB

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 (Issue 4) Pages (27-36)

Journal home page: www.ijreb.org

states are isolated from each other and a passage between them does not exist. We will have then to compare them through their connectivity to other states and for this purpose, we now introduce a new measure of state similarity that we call mean first cross passage time (CPT). The mean first cross passage time between states i and j will measure the difference of the passages of these two states to all the other states. Due to the memory less property, once the process reach the same state, lets say k, from i and from j then the different initial condition (i or j) has no effect for the process any more, therefore, the mean first cross passage time between two states will be the expected time it takes for the process to cancel out the fact of one of these two states being the initial condition rather than the other. Intuitively, the CPT will be large between states that there is no state to which both of them connect with short passages.

Optimality of the Proposed Distance

The CPT is not only a measure of the relevance between keywords that is modeled in the MSI distance but also gives stochastic interpretation by quantifying the importance of the coordinates of $v_2'$with respect to connectivity between states. The difference between any pair of the coordinates of this vector is the best largest approximation to the CPT between the corresponding states. Consequently, the sum of all the pair wise absolute differences of the coordinates of this vector is the largest best approximation of the total CPT between all pairs of states of the network and thus$v_2'$is the direction where the CPT projects with maximumvariance. We will now show, that is also the direction thatmaximizes the proposed distance.

PRIORITY BASED RANKING ALGORITHM

In Priority Based Ranking method, the Ranking counter is set to add every count for the select of image. After get the set of images, once user select a image the counter corresponds to that image is add by one. For every time the selection of image the counter value increases. And the higher value annotation is set with the priority list. According to the ranking value the priority list set the image annotation. Image annotates with higher priority value. Every time the selection of image the counter value increases. For every new query counter value sets, the higher value annotation is set with the priority list.

IJREB

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 (Issue 4) Pages (27-36)

Journal home page: www.ijreb.org

According to the ranking value the priority list set the image annotation.

EXPERIMENTAL RESULTS

Compared the proposed method to the LSI and pLSI approaches in two scenarios. The first experiment is a comparison to LSI, since the limited number of images used in this experiment does not permit reliable comparison to pLSI. The full features of the proposed distance (MSI) are demonstrated in this experiment since the generative process of the aggregate Markov chain during the automatic annotation of images was available

In the second experiment, the full features of our method cannot be demonstrated since the scenario involves a publicly available ground-truth database, during the annotation of which we had no control. The Aggregate Markov Chain, necessary for our method cannot be reliably constructed. Nevertheless this experiment serves as a comparison to pLSI in the ability to extract latent features in the case of already annotated databases, the annotation having been performed with unknown methods. For such cases, we propose a modification to the standard MSI approach that involves an explicit step for dimensionality reduction

1. CONCLUSIONS

Markovian Semantic Indexing, a new method for mining user queries by defining keyword relevance as a connectivity measure between Markovian states modeled after the user queries. The proposed systemic dynamically trained by the queries of the same users that will be served by the system. Consequently, 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, in the form of generalized euclidean distance, was constructed by meansof an Aggregate Markovian Chain and proved to be optimal with respect to certain Markovian connectivity measures that were defined for this purpose. A comparison to Latent Semantic Indexing and probabilistic Latent semantic Indexing revealed certain theoretical advantages of the proposed method (MSI). Experiments have shown that MSI achieves better retrieval results in sparsely annotated image data sets.

REFERENCES

1. R. Datta, D. Joshi, J. Li, and J.Z. Wang, “Image Retrieval: Ideas, Influences, and Trends of the New Age,” ACM

ISSN 2321-743X
International Journal of Research in
Engineering and Bioscience
Volume 2 (Issue 4) Pages (27-36)
IJREB
Journal home page: www.ijreb.org

Computing Surveys, vol. 40, no. 2, pp. 1-60, 2008.

2. T. Hofmann, “Probabilistic Latent Semantic Indexing,” Proc. 22nd Int’l Conf. Research and Development in Information Retrieval (SIGIR ’99), 1999.
3. T. Hofmann, “Unsupervised Learning by Probabilistic Latent Semantic Analysis,” Machine Learning, vol. 42, no. 1/2, pp. 177-196, 2001.
4. D.M. Blei and A.Y. Ng, and M.I. Jordan, “Latent Dirichlet Allocation,” J. Machine Learning Research, vol. 3, pp. 993-1022, 2003.
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, article 5, pp. 1-60, 2008.
6. 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.
7. A. Bhattacharya, V. Ljosa, J.-Y. Pan, M.R. Verardo, H. Yang, C. Faloutsos, and A.K. Singh, “Vivo Visual Vocabulary Construction for Mining Biomedical Images,” Proc. IEEE Fifth Int’l Conf. Data Mining, Nov. 2005.
8. 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.