iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 12 **Issue:** VI **Month of publication:** June 2024

**DOI:** https://doi.org/10.22214/ijraset.2024.63382

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue VI June 2024- Available at www.ijraset.com*

# Markov Chains for User Queries Mining: Image Retrieval Application using Internet

Dr. Madhav M. Bokare[1], Mr. Amol V. Suryawanshi[2]
*Department of Computer Sciencet, SRTMU Nanded*

***Abstract:*** ***We suggest a brand-new technique for automatically annotating, indexing, and retrieving photos based on annotations. We introduce the novel technique, which we term Markovian Semantic Indexing (MSI), within the framework of an image retrieval system in online mode. In the event that such a system exists, the queries entered by the users are utilised to build an Aggregate Markov Chain (AMC), which defines the significance between the terms the system sees. The photographs are automatically annotated based on the queries entered by the users. Next, depending on the annotation of each image and the keyword relevance recorded in the AMC, a stochastic distance between them is presented. The suggested distance is given geometric meanings, and its relationship to a grouping in the keyword space is examined. The optimality qualities of the suggested distance are demonstrated by the use of a novel Markovian state similarity measure, the mean first cross passage time (CPT). Images are represented as points in a vector space, and MSI is used to calculate how similar two images are. In Annotation-Based Image Retrieval (ABIR) tasks, it is demonstrated that the novel method outperforms Latent Semantic Indexing (LSI) and probabilistic Latent Semantic Indexing (pLSI) methods in terms of Precision against Recall and has some theoretical advantages.***

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

## I. INTRODUCTION

The relationship between low-level features and high-level semantics of image content is lost in current computer vision algorithms, despite the fact that humans tend to identify images with high-level concepts. In general, there is no clear semantic meaning associated with either a single low-level feature or a combination of several low-level traits. Furthermore, the similarity measures among visual features do not always correspond with human perception, which leads to usually unsatisfactory and frequently unpredictable retrieval outcomes for low-level techniques. The lack of agreement between the information that may be inferred from visual data and the meaning that the same facts have for a user in a particular circumstance is known as the "semantic gap." The sensory gap, or the difference between an object in the real world and the data in a (computational) description attached to a recording of that thing, is another reason why the retrieval process fails. The first gap deals with how consumers interpret images and how hard it is to capture them in visual material; the second gap limits the ability to record and describe images, making it difficult to recognise content from them. As of right now, barely 10% of image files available online have a formal description (annotation). Because of this, picture search engines can only provide recall and precision of about 12% and 42%, respectively, and 60% of users use two or more search engines because they are dissatisfied with the results they get. The most frequent grievance is that semantics in content are not recognised by search engines. Furthermore, 77 percent of searchers switch up their terms multiple times due to an inability to find relevant material. The goal of Annotation-Based Image Retrieval (ABIR) systems is to integrate semantic content into image captions and text-based queries (e.g., Google Image Search, Yahoo! Image Search) more effectively. In order to find a more trustworthy idea association, the Latent Semantic Indexing (LSI)-based techniques, which were first used with more success in document indexing and retrieval, were integrated into the ABIR systems. Nevertheless, the degree of success of these endeavours is debatable; one explanation for this is the sparsity of the keyword annotation data per-image relative to the total number of keywords often supplied to documents. We provide a novel approach to automatic annotation and annotation-based image retrieval: Markovian Semantic Indexing (MSI). Due to its characteristics, MSI is especially well suited for ABIR assignments where there is a dearth of annotation data for each image. The method's features also make it especially useful in the context of online picture retrieval systems. The remainder of the document is structured as follows: We first introduce related work and our contribution in the following section. In Section 3, a proximity measure (distance) and the suggested methodology (MSI) are described. Section 4 looks at the ideal characteristics and geometric interpretation of the suggested distance. In Section 5, two well-known techniques from the literature, LSI and pLSI, are experimentally compared to MSI in two different scenarios:

*a)* Of supervised MSI annotation
*b)* Of external annotations realized with unknown methods.

This section also covers scalability considerations and a conceptual comparison of the methods. Section 6 contains the Conclusion. In this work, standard letters are scalars, bold letters are vectors, and capital letters are matrices. A bold letter with a subscript indicates a column vector; the subscript indicates where the letter belongs in a matrix. All vectors are regarded as column vectors unless specifically stated otherwise, with the exception of equilibrium vectors, which are typically regarded as row vectors.

## II. RELATED WORK AND OUR CONTRIBUTION

Without a doubt, content-based retrieval has grown quickly. Over 200 content-based retrieval systems have been created recently, most of them relying on low-level features. Specifically, they fall into two primary categories:

*1)* People who carry out semantic mining by examining textual data related to images, including captions, annotations, keywords allocated to the images, alternative (alt) text on HTML sites, and surrounding text,

*2)* Systems that use low-level visual properties like colour and texture to extract them in order to perform various tasks on image data, like alignment, categorization, browsing, searching, and summarising.

The latter methods typically aren't able to capture semantics effectively, while the methods in the first category rely heavily on tedious annotation. Furthermore, some other methods accomplish content-based operations by using both low-level features—visual keywords and text annotation—but they typically require users' conscious participation in order to annotate photographs linguistically. More effective semantic material is incorporated into both text-based queries and image captions by Annotation-Based Image Retrieval systems. As a direct result, techniques originally created for document retrieval might also work well for ABIR systems. Originally, Latent Semantic Indexing was created for document retrieval.

Hofmann developed probabilistic Latent Semantic Indexing (pLSI), which is based on the Aspect Model and offers a document retrieval alternative to projection (LSI) or clustering techniques. In order to overcome the generalisation and overfitting issues with pLSI, Latent Dirichlet Allocation (LDA) was introduced by Bleiet al. Griffiths and Steyvers then combined LDA with a Markov chain Monte Carlo approach. In order to address overfitting issues, Steyvers et al. developed a novel probabilistic model for document retrieval that included Gibbs sampling to represent both authors and themes. There have been attempts to use LSI/pLSI-based methods to find a more trustworthy idea connection in ABIR systems. Joint models across text and images have already been investigated in the computer vision literature. Based on the aspect model as well, Barnard and Forsyth presented a statistical model with hybrid text/visual features. Li et al. developed a different model based on Bayesian incremental learning, and Fan et al. suggest a multilevel automatic annotation method based on both local and global visual data. Although stochastic in nature, the method shown here elevates the reasoning component of probability by clearly defining the relevance relationships between keywords. In AI systems, it is common practice to capture semantics via network representations. These are known as galleries, qualitative Markov networks, or constraint networks in the Dempster-Shafer theory; they are also known as influence diagrams, causal nets, or Bayesian networks in probability theory. It has been demonstrated that associational graphs—such as semantic networks, constraint networks, inference networks, conceptual dependencies, and conceptual structures—better capture the plausibility of human reasoning in many models of human reasoning.

### *A. Our Contribution*

This paper proposes a novel (alternative) probabilistic strategy for Annotation Based picture Retrieval that is more appropriate for sparsely annotated domains (e.g., picture databases) where the per image sparse keyword annotation is constrained, than LSI and pLSI. It tackles the zero frequency problem more naturally, which is the idea that there is usually little chance of finding similar keywords even in closely related photos because the images are not annotated with exactly the same phrases. Here, an explicit relevance link between terms with a probabilistic weight is used to solve this issue. The approach's main idea is to use a probabilistic qualitative reasoning annotation technique to convey partial views about the links between keywords, making up for the scant data. Thus, a mechanism that improves performance by mining the structure of the available data is introduced instead of adding additional data, as is the case with standard models. Moreover, the suggested system is unique in the manner it combines these two tasks, even though automatic annotation and annotation-based image retrieval systems have been published in the literature. In order to dynamically mine semantics towards qualitative probabilistic reasoning, it is envisaged that both the automatic annotation and retrieval activities will take place in the implicit user interaction context. It is demonstrated that MSI is ideal in relation to CPT. Conceptual comparisons between pLSI and LSI as well as between the two show benefits for the suggested method.

## III. THE PROPOSED APPROACH

By providing photos that are more likely to be accepted (downloaded) by the user, the goal is to increase user happiness. It is assumed that users ask questions in order to look for photos, and each question consists of an ordered list of keywords. In response, the system displays a list of pictures. The user has the option to download the returned images or to disregard them and submit a new query. The system considers the photos unannotated throughout the training phase. The system automatically annotates the photos as users select images and pose queries. It also creates relevant associations between the keywords, as will be detailed later in the article. The system annotates the photographs invisibly from the user; the user never explicitly does this. In order to return images that more accurately reflect the user's preferences and increase user satisfaction, the system utilises both the annotations that were made available during the training phase and the keyword relevance probability weights that were also assessed during that phase. We use the implicit effects of this interactive process one by one as we build the suggested system step-by-step:

Step 1: The user implicitly relates the retrieved (downloaded) images to her/his query. 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 $I_i$ to his query $q_i$ , where keyword $k_2$ follows keyword $k_1$ and this occurs 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 the current probability (before the update) based on M keywords then the new probability (based on M + m keywords) is calculated by the recurrent formula

$$pi(\mathrm{k1},\mathrm{k2}) = \frac{Mpi(\mathrm{k1},\mathrm{k2})+\mathrm{m}}{M+m} \quad (1)$$

Using this process, a Markov chain is created, with each keyword representing a state. A keyword's state counter advances each time it occurs in a query, and its interstate link counter advances if another keyword appears in the same query after it. This method is used to measure both the frequency of the keywords and their order of occurrence. Batch processing of the image-related queries results in the advancement of the counters and the updating of the probabilities as previously said. The counts are cleared before to processing the subsequent batch of requests. For every image Ii, the equilibrium state vector of the so-constructed Markov Chain will be represented by πi, which will now stand in for the image.

This modelling technique is supported by two factors: 1) the conceptual approach used in this work is qualitative; and 2) the MSI's targeting characteristics, which aim to capture aspects unique to each user, such as perception of visuals. In fact, 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, is justified by the fact that each sequence of keywords (query) originates from a specific user, filtered through her/his perception about the selected image. Our modelling approach favours this logical relationship over the conventional numerical method of estimating the distribution of images over keywords.

Step 2: One encounters the zero-frequency issue when attempting to compare the probability vectors πi and πj for the two images that were computed in the previous step directly. The very act of a user combining specific phrases in a query implicitly positions the keywords in relation to one another, independent of the specific images the user may or may not choose. By grouping related keywords together, we suggest using this to solve the zero-frequency issue. This stage creates the Aggregate Markovian Chain (AMC) of all user inquiries, regardless of the photos they have selected, for this reason. The kernel of this process denoted by PG, is calculated in a similar to the previous step manner by the recurrent formula of (1).

The aim of the AMC is to model keyword relevance since, although PG uses a Markov kernel, it will be utilised to cluster the keyword space rather than estimate an explicit probability distribution.

Step 3: Step of optimisation. The keyword space will be clustered using the AMC, and through this clustering, explicit relevance linkages between the keywords will be defined. By assessing the series FG(n)= ∑_(k=0)^n PkG, where PG is the AMC kernel, the clustering job is connected to the convergence properties of the AMC chain. Not before the rapid convergence has completed, but at the desired n where the slow convergence has taken over, is where an appropriate termination condition puts an end to the series. Since the clusters in the rows of FG(n) will lose rank and the determinant will approach zero, the value of the determinant of FG(n) is employed as a termination condition. The n-step predicted occupancies matrix is denoted by FG(n) (Appendix). This process is associated with an optimisation problem concerning the total variance of the columns of FG(n), when projected in the direction of the PG eigenvectors. We shall learn more about this idea in the following section. □(1/(n+1)) FG(n), the n-step predicted fractional occupancies matrix (Appendix), computed at the desired n, shall now be noted using only FG.

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue VI June 2024- Available at www.ijraset.com*

Step 4: At this point, the MSI distance's formal definition can be given. Definition 1. Let x and y be two pictures, denoted by the corresponding row vectors of their steady state probabilities, x and y. Let furthermore ∑(FGT) be the covariance matrix of the Aggregate Markov Chain (AMC) zero-mean transpose anticipated fractional occupancies matrix, computed at the chosen n. Next, we define the MSI distance between image x and image y as

$$d(x,y) = \frac{(\pi_{..} - \pi_{..})\Sigma(F_{\cdot}^{T})\,(\pi_x - \pi_y)^T}{\boldsymbol{\delta}_{xy}\Sigma(F_G^T)\boldsymbol{\delta}_{xy}^T,} \quad (2)$$

where the corresponding coordinates have been filled in with zeros to increase the dimensionality of πx and πy to that of ∑(FGT). Since the suggested distance is a generalised Euclidean distance function utilising a covariance matrix that is always positive definite, it is well defined.

## IV. GEOMETRIC INTERPRETATION AND OPTIMALITY PROPERTIES OF THE MSI DISTANCE

When projected on the direction of the difference, (πx - πy), between the two pictures, the proposed MSI distance d(x,y) (Definition 1) can be seen as measuring the total variance of the rows of FGT. The direction determined by the vector difference of the probability distributions of the two specific images is actually determining the distance between them because the FGT is only calculated once from all the data. Understanding the mechanics of the Markovian convergence that generates the FGT and its relationship to certain directions in the keyword space is necessary to obtain further geometric and stochastic interpretations of the MSI distance. We will explore the geometric interpretation of the convergence process and how it relates to the suggested distance in terms of state clusters and state connection metrics in the remaining portion of this section. We will make use of standard language, concepts, and notation found in the literature on stochastic processes. For a concise overview, please refer to the Appendix.

### *A. The Eigenvector Convergence Process of the Markov Kernel as a Foundation for State Partition*

This section explains how the convergence of an initial condition to its corresponding equilibrium state occurs along the eigenvectors of the Markov kernel, with each direction's relative rate of convergence determined by its corresponding eigenvalue. For a brief introduction to the notation of Ti(n), ti(n), and tg(i) and their values Tij(n), tij(n), respectively, as well as other pertinent terms utilised below, the reader is referred to the Appendix. Since the chain converges with rate 1 in this direction, as will be shown later in this section, the equilibrium state for the corresponding initial condition is the l1 normalised eigenvector, which corresponds to an eigenvalue of 1.

First, we express the transient parts of the process in terms of the projections of the initial conditions on the eigenvectors, and provide some helpful derivations. Assume that ei = Φi(0) and ej = Φj(0) represent the corresponding initial conditions for the nodes i, j, in a chain. Let us additionally consider the eigenvalues of P, denoted by λi, i=1,2,…, N, and the corresponding normalised (relative to the l1 norm) left hand eigenvectors, vi and |λ|<=1. Let be a nonincreasing ordering of the eigenvalues such that |λi|>=|λj| when i

$$\begin{aligned} \boldsymbol{e}_i = \boldsymbol{\phi}_i(0) &= v'_{1i}\boldsymbol{v}_1 + v'_{2i}\boldsymbol{v}_2 + \cdots + v'_{Ni}\boldsymbol{v}_N \\ \boldsymbol{e}_j = \boldsymbol{\phi}_j(0) &= v'_{1j}\boldsymbol{v}_1 + v'_{2j}\boldsymbol{v}_2 + \cdots + v'_{Nj}\boldsymbol{v}_N. \end{aligned} \quad (3)$$

$$\begin{aligned} \boldsymbol{\phi}_i(n) &= \boldsymbol{v}_1 + \lambda_2^n v'_{2i}\boldsymbol{v}_2 + \cdots + \lambda_N^n v'_{Ni}\boldsymbol{v}_N \\ &= \boldsymbol{v}_1 + \boldsymbol{t}_i(n) \\ \boldsymbol{\phi}_j(n) &= \boldsymbol{v}_1 + \lambda_2^n v'_{2j}\boldsymbol{v}_2 + \cdots + \lambda_N^n v'_{Nj}\boldsymbol{v}_N \\ &= \boldsymbol{v}_1 + \boldsymbol{t}_j(n). \end{aligned} \quad (4)$$

$$\begin{aligned}\bar{\boldsymbol{\eta}}_i(n) &= \sum_{k=0}^{n} \boldsymbol{\phi}_i(k) \\ &= (n+1)\boldsymbol{v}_1 + v'_{2i}\boldsymbol{v}_2 \sum_{k=0}^{n} \lambda_2^k + \cdots + v'_{Ni}\boldsymbol{v}_N \sum_{k=0}^{n} \lambda_N^k \\ &= (n+1)\boldsymbol{v}_1 + \sum_{p=2}^{N} \frac{1-\lambda_p^n}{1-\lambda_p} v'_{pi}\boldsymbol{v}_p \\ &= (n+1)\boldsymbol{v}_1 + \sum_{k=0}^{n} \boldsymbol{t}_i(k) = (n+1)\boldsymbol{v}_1 + \boldsymbol{\tau}_i(n),\end{aligned} \tag{5}$$

and the vector of the expected fractional occupancies; if started from $i$, is

$$\begin{aligned}\boldsymbol{f}_i(n) &= \frac{\bar{\boldsymbol{\eta}}_i(n)}{n+1} \\ &= \boldsymbol{v}_1 + \frac{1}{n+1} \sum_{p=2}^{N} \frac{1-\lambda_p^n}{1-\lambda_p} v'_{pi}\boldsymbol{v}_p \\ &= \boldsymbol{v}_1 + \frac{\boldsymbol{\tau}_i(n)}{n+1}.\end{aligned} \tag{6}$$

The coordinates on the eigenbase of the transient sum components are used in this formula to display their structure. When the system is started from I instead of j, the difference between the expected occupancies of each node at step n is then

$$\bar{\boldsymbol{\eta}}_i(n) - \bar{\boldsymbol{\eta}}_j(n) = \sum_{k=0}^{n} \boldsymbol{\phi}_i(k) - \sum_{k=0}^{n} \boldsymbol{\phi}_j(k) = \sum_{k=0}^{n} (\boldsymbol{\phi}_i(k) - \boldsymbol{\phi}_j(k)),$$

which, taking into account (4) and (5) becomes

$$\begin{aligned}\bar{\boldsymbol{\eta}}_i(n) - \bar{\boldsymbol{\eta}}_j(n) &= \sum_{p=2}^{N} \sum_{k=0}^{n} \lambda_p^k (v'_{pi} - v'_{pj})\boldsymbol{v}_p \\ &= \sum_{p=2}^{N} \frac{1-\lambda_p^n}{1-\lambda_p} (v'_{pi} - v'_{pj})\boldsymbol{v}_p \\ &= \boldsymbol{\tau}_i(n) - \boldsymbol{\tau}_j(n).\end{aligned} \tag{7}$$

Thus, the difference represented on the eigenbase of the transient sum components.Observe that the term (n+1)v1, in (5),vanishes in the difference of (7) since I and j are in the same chain, and as a result, the result converges. This isn't always the case, though, because if I and j are in distinct chains, their equilibrium vectors—let's say v1 and u1—are different. As a result, the corresponding term becomes (n+1)(v1-u1), going towards infinity. If the multidesmic processes are changed to monodesmic by adding tiny, nearly zero probabilities between the various chains, we can solve this issue. Since monodesmic processes will always be taken into consideration in the sequel, the term Markovian chain will be employed hereinafter.

In the event of a monodesmic process, the difference in the expected fractional occupancies always converges.

$$\begin{aligned}\boldsymbol{f}_i(n) - \boldsymbol{f}_j(n) &= \frac{\bar{\boldsymbol{\eta}}_i(n) - \bar{\boldsymbol{\eta}}_j(n)}{n+1} \\ &= \frac{1}{n+1} \sum_{p=2}^{N} \frac{1-\lambda_p^n}{1-\lambda_p} (v'_{pi} - v'_{pj})\boldsymbol{v}_p \\ &= \frac{\boldsymbol{\tau}_i(n) - \boldsymbol{\tau}_j(n)}{n+1}.\end{aligned} \tag{8}$$

As $n$ goes to infinity, (7) becomes

$$\bar{\boldsymbol{\eta}}_i(\infty) - \bar{\boldsymbol{\eta}}_j(\infty) = \sum_{p=2}^{N} \frac{v'_{pi} - v'_{pj}}{1-\lambda_p} \boldsymbol{v}_p = \boldsymbol{t}_g(i) - \boldsymbol{t}_g(j), \tag{9}$$

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue VI June 2024- Available at www.ijraset.com*

And on the eigenbase, this is how the transitory sum difference is expressed. Since this expression connects the mechanics of convergence to the real process statistics, it will be helpful in proving optimality concepts of the suggested distance in the next sections. Let j be recurrent and i and j be two states of the same chain. The anticipated number of transitions the process requires to reach state j for the first time, starting from i, is the mean first passage time I j from state i to state j.

$$\bar{\theta}_{ij} = \frac{t_g(jj) - t_g(ij)}{\pi_j}, \qquad (10)$$

where $\boldsymbol{\pi}$ = (π1,π2,π3…πN) the equilibrium state of the chain. Equations (9) and (10) show that the mean first

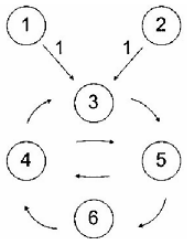



Fig.1.A Markov chain showing two nonrecurrent states with identical future.

The coefficients of the starting conditions, ei and ej, in their formulation as a linear combination of the eigenvectors of P, are connected to the elements of V' and the passing time from j to i. Specifically, we obtain from (9) and (10)

$$\pi_j \bar{\theta}_{ij} = \frac{v'_{2j} - v'_{2i}}{1 - \lambda_2} v_{j2} + \cdots + \frac{v'_{Nj} - v'_{Ni}}{1 - \lambda_N} v_{jN}, \qquad (11)$$

However, as Fig. 1 illustrates, in the case of the nonrecurrent states 1 and 2, the process never visits both states, therefore even though both states lead to identical futures, the first passing time between them cannot be determined. Regardless of how the process accesses these stages, we require a connection measure between any pair of states to handle this and other similar non-recurrent state scenarios. Rather than directly connecting the states, we propose to accomplish this via relating the states based on the statistics of the entire process after visiting/starting from these states.

*B. The Mean First Cross Passage Time*

Then, we will need to compare them based on how connected they are to other states. To that end, we are going to present a new metric for state similarity, which we are calling mean first cross passage time (CPT). The difference in these two states' passages compared to all other states will be measured by the mean initial cross passage time between states I and J. Owing to the memoryless property, the process loses its ability to respond to different initial conditions (i or j) once it reaches the same state, say k, from i and j. As a result, the mean first cross passage time between two states represents the expected amount of time needed for the process to reverse the effect of one of these two states being the initial condition rather than the other. It makes sense that the CPT would be wide between states if there were no states that both of them could connect to by brief crossings.

We start by selecting a random state, k. The process from i visited k for the first time after an average of transitions, and the process from j also visited k for the first time after an average of transitions, if the first passage crossing from i and from j occurred at k. Given that the initial crossing of passages occurred at state k, the difference between the mean first passage timings of i and j to k is the estimated time k will have to wait for the second passage after it was visited by the first passage. This is the CPT.

Put formal terms on this, if I, j, and k are arbitrary and k is recurring, and the mean first passage time between i and k and between j and k equals

$$\bar{\psi}_{ij} = \sum_{k \in N} |\bar{\theta}_{ik} - \bar{\theta}_{jk}| \pi_k, \qquad (12)$$

*which from (11) becomes:*

$$\begin{aligned} \bar{\psi}_{ij} &= \sum_{k \in N} |\pi_k \bar{\theta}_{ik} - \pi_k \bar{\theta}_{jk}| \\ &= \sum_{k \in N} |\frac{v'_{2j} - v'_{2i}}{1 - \lambda_2} v_{k2} + \cdots + \frac{v'_{Nj} - v'_{Ni}}{1 - \lambda_N} v_{kN}| \end{aligned} \qquad (13)$$

To address the issue of nonrecurrent states and eliminate the requirement that the first crossing of passages occur at state k, we can assume zero k for nonrecurrent states k and add up the first cross passage times between i and j at k, weighted by the likelihood that the process will be at state k. This yields, respectively, the mean first cross passage time between i and j given that the first crossing of passages will occur at state k. The existence of a Perron root z‚2 (modulus close to one) indicates that, for any arbitrary pair of Markovian states i and j, their expected first Cross Passage Time is indicated by. This means that, regardless of where the process starts, let's say from i, at a large time n, its state vector will be v 1 á n 2 v0 2iv 2 because, in accordance with (4), the other terms will vanish. Consequently, the chain will move towards equilibrium at the rate that 2 shrinks the coefficient v0 2i, which is slow if 2 is close to one. In the event that the process was initiated from j rather than i, its state vector at a big n would be v 1 Í n 2 v0 2jv 2, and the difference in state vectors at this point resulting from launching the process from j instead of i would be in j v0 2j v0 2ij. In terms of clustering, this shows that a clustering of the states with regard to their CPT and the ultimate convergence to equilibrium when begun from these states corresponds to a 1D clustering of the projections v0 2i of the initial conditions ei on the eigenvector v 2.

This link calculates the variations in keywords between the two photographs and restores the rationale behind the suggested method. Moreover, the clustering of the states attained by the AMC's convergence process determines the CPT between two states (keywords). Consequently, the definition of the relevance linkages evolves from the clustering of the state space. By extrapolating the foregoing, we can conclude that, in the event that more keywords are utilised to describe the photos, their MSI distance is weighing the relevance link for each keyword pair according to the difference in their likelihood of representing the two images. We will now demonstrate that this occurs in the best possible way.

### *C. Optimality of the Proposed Distance*

In addition to measuring the relevance of keywords, which is described by the MSI distance, the CPT quantifies the importance of the coordinates of v0 2 in terms of connectivity between states, providing a stochastic interpretation. The difference between any pair of coordinates of this vector is the best greatest approximation to the CPT between the respective states, as was demonstrated in (13) and described at the end of Section 4.2. As a result, the direction in which the CPT projects with the greatest variance is v0 2, as the largest best estimate of the total CPT between all pairs of states in the network is the sum of all the pairwise absolute differences in this vector's coordinates. Now, we'll demonstrate that this is also the path that maximises the suggested separation. Proposition 1. The proposed distance d(x,y)=(πx-πy) ∑(FGT) (πx-πy)T is maximised on the direction of maximal CPT. Let ∑(FGT) be the covariance matrix of the transpose zero-mean expected fractional occupancies matrix of a Markovian chain with kernel P.

Evidence. Simple examination reveals that d(x,y) is the total variance of the FGT rows projected on the direction; as a result, factor 1 n 2 1 2 decreases more slowly as n rises than all other factors 1n k 1k. Remember that the series is terminated by the desired n in the optimisation step of the FG convergence process once the fast-shrinking eigenvectors have reached convergence. Since the rows of the FGT are spread by the maximum factor of 1n2 12 in this direction, the direction of v0 2 is the one with the total highest variance at this time (the other factors having converged to 1 1k ). However, as the recommended distance is maximised in the direction of the largest projected initial cross passage time, it can be seen that this is also the direction of the overall maximum CPT from above. According to the last claim, a set of coordinates in x and y, respectively, that deviate by the same amount are penalised more when the associated states have higher CPT. When (πx - πy) is in the direction of v' 2, where the corresponding states have the maximum CPT, the maximum penalty occurs.

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue VI June 2024- Available at www.ijraset.com*

## V. EXPERIMENTAL RESULTS

In two scenarios, we evaluate the suggested method against the LSI and pLSI methodologies. Since there aren't enough images utilised in this experiment to make a reliable comparison to pLSI, the first experiment compares to LSI. Since the generative process of the aggregate Markov chain during the automatic annotation of photos was available to us, as will be discussed later, the entire features of the suggested distance (MSI) are illustrated in this experiment. For this experiment, sixty-four photographs were employed, which correspond to two intuitive classes: 32 images were associated with the term Greek and are regarded as belonging to the first class, and 32 images associated with the term Hawaiian are regarded as belonging to the second class. Initially, the 64 photos' distance from the Greek Islands inquiry is determined, ranked, and analysed using both approaches. Next, utilising both techniques, a comprehensive distance table is constructed for all the intermediate distances of these 64 photos. For comparison, a precision versus recall graphic is shown. Since the second experiment used a publicly accessible ground-truth database, over whose annotation we had no influence, it is unable to fully illustrate the capabilities of our approach. It is not possible to generate the Aggregate Markov Chain with sufficient reliability for our strategy. However, this experiment compares to pLSI in terms of its capacity to extract latent features from databases that have already been annotated, even though the annotation was done using unidentified techniques. In these situations, we suggest a tweak to the conventional MSI method that includes a clear dimensionality reduction phase (since the implicit dimensionality reduction achieved by grouping the keyword space cannot be utilised). Using the ground-truth database, this comparison to the pLSI is carried out once more, this time with precision versus recall diagrams.

### *A. Comparison to LSI When the Proposed Markovian Annotation is Available*

We requested permission to record click-through data (query q, returned ranking r, and clicks c) from a set of university students when they used Google Image Search in English. Since the suggested system needed to be tested for its capacity to adapt to the preferences of the user, the students were requested to play the roles of several target groups with particular interests (tourists, sports fans, artists, etc.) at different times. To store a log file containing the query-ID, URL, and click-log, a straightforward proxy was employed. Only the chosen photos from the images that Google Image Search returned were taken into consideration. Each picked image's unique Markov chain (based on the query terms as previously mentioned) was constructed from the recorded sessions over a period of six months. Each image was then automatically annotated using the equilibrium state vector of the corresponding Markovian chain. Similarly, a Markov chain representing the Aggregate Markov Chain was built from the total number of queries from all users, irrespective of the photos selected, acting as a dynamically adaptable taxonomy for the keywords already present in the user queries. Figure 2 illustrates a portion of this network that clustered around the terms Greek and Hawaiian. We excluded the transition probabilities that were less than 0.1. Table 2 presents a rating of 64 photos for each of the two approaches under examination based on their semantic distance from the query Greek Islands.

The initial enumeration for each image is contained in the first column. Columns 4–24 contain the equilibrium state vector of the Markov chain that represents each image. The MSI and LSI methods' rankings with respect to distance from the query Greek Islands are contained in the second and third columns of this table, respectively. The vector [Greek (0:5), Islands (0:5)] represents the query Greek Islands. Intelligent retrieval, comprising dependencies beyond simple keyword co-occurrences, is made possible by the semantic importance of the keywords inferred by the Markovian network of the aggregate Markov chain, as demonstrated by the ranking of the images retrieved by the proposed system (MSI) in Table 2. While photos (1), (2), and (6), for example, are the most likely to be related to the query since they include the same keywords, images (3), (4), and (5), while having the same exact keywords as images (1), (2), and (6), are ranked 29th, 25th, and 17th, respectively. Prior to (3), (4), and (5), we encounter numerous photos with supplementary keywords in locations ranging from 4th to 16th. This occurs as a result of the network being closer, as demonstrated by the Aggregate Markov Chain of Figure 2, which begins the process from initial conditions (or transitory states) equal to the vector representing each of these images (e.g., (15), (14), (7), (8), (11), (9), etc.

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue VI June 2024- Available at www.ijraset.com*

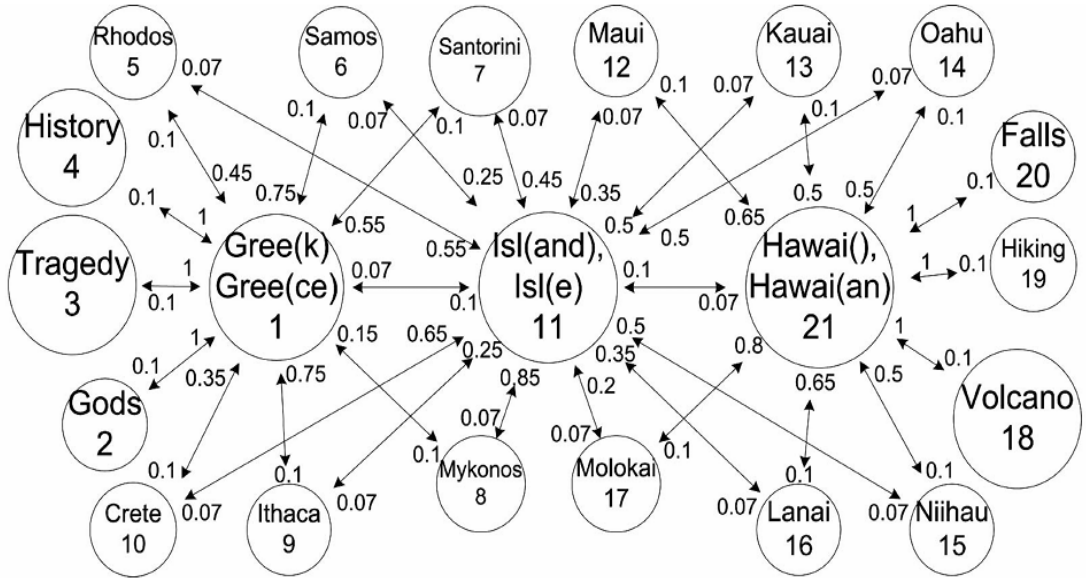



Fig.2.The Aggregate Markovian Process of the experiment in Section 5.1 is shown in this figure. Only the part of the network forming a cluster around the concepts of Greek, Hawaiian, and Islands is graphically shown. The transition probabilities leading outside the displayed portion of the network have been set to zero and their weight distributed evenly to the respective connections within the displayed portion of the network. Transition probabilities less than 0.05 have been omitted.

Predicted number of transitions to arrive at the state that the query represents, as opposed to beginning with the initial conditions that are represented by the images (3), (4), or (5). Similar circumstances can be seen further down the ranking, as image (14) for example comes above many photos that contain both query terms, even if they only contain one of the two. This is further clarified by looking at Fig. 2's aggregate process, which shows that most users are more likely to ask island or Greek query keywords after requesting information about Santorini than they are to ask any other island name. We also note that, as should be the case and as shown by the AMC Markov chain, all photos including Greek subjects appear before all images featuring Hawaiian subjects in this order. This is the case because Greek islands are more likely than Hawaiian islands to lead to both query phrases. It is also evident that the images featuring any of the three terms (God, Tragedy, and History) that are only associated with Greek but not with islands appear after those featuring Greek islands but before those featuring Hawaiian islands. This is as it should be, as even if the images featuring Hawaiian islands are only associated with one keyword, all of the keywords associated with Hawaiian islands most likely lead away from at least one of the query keywords. When it comes to the LSI rankings, we see that while at first it returns more logical results—for instance, images (1)–(6) are the closest to the query because they contain exactly the same keywords—further down the ranking, we encounter nonlogical results. For instance, image #40, a Hawaiian island, comes in at number 10, and it is closer to the query Greek Island than many other images featuring Greek islands. The LSI ranking loses the coherence we saw in the MSI, which states that every Greek island came before every Hawaiian island. In order to investigate the behaviour of the two approaches further, we created a distance table between each pair of the 64 images. Taking into account that images (1)–(32) form the class Greek and images 33–64 form the class Hawaiian, we then created a Precision versus recall diagram for each of the two approaches. The following are the procedures involved in creating the distance table for the suggested (MSI) approach:

*1)* A tiny amount is added to each one-diagonal element (elements on the super diagonal) starting with the Markov kernel P of Table 1 and subtracted from each random nonzero element in the same line. In this method, we maintain the matrix's stochastic nature and adhere to Section 4.1's conversion of the process to monodesmic (chain) without changing the process statistics.

*2)* Substitute the desired n into the predicted fractional occupancy matrix calculation FG =1/((n+1)) ∑_(k=1)^n⁡〖Pk〗 , where n is calculated based on the discussion in Section 3. Results for n=1, n=5, n=10, n=12, n=14, and n=15 are shown. In order to determine the powers of P, there is no need for matrix multiplication since, according to (6), an eigenvalue decomposition of P is enough to calculate FG at any n, only the powers of the eigenvalues need to be calculated.

Table 1
The Markov Kernel that Corresponds to the Chain of Fig. 2

| | K1 GRE | K2 GOD | K3 TRA | K4 HIS | K5 RHO | K6 SAM | K7 SAN | K8 MYK | K9 ITH | K10 CRE | K11 ISL | K12 MAU | K13 KAU | K14 OAH | K15 NIH | K16 LAN | K17 MOL | K18 VOL | K19 HIK | K20 FAL | K21 HAW |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GRE | - | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | - | - | - | - | - | - | - | - | - | - |
| GOD | 1.0 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| TRA | 1.0 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | |
| HIS | 1.0 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | |
| RHO | 0.45 | - | - | - | - | - | - | - | - | - | 0.55 | - | - | - | - | - | - | - | - | - | - |
| SAM | 0.75 | - | - | - | - | - | - | - | - | - | 0.25 | - | - | - | - | - | - | - | - | - | - |
| SAN | 0.55 | - | - | - | - | - | - | - | - | - | 0.45 | - | - | - | - | - | - | - | - | - | - |
| MYK | 0.15 | - | - | - | - | - | - | - | - | - | 0.85 | - | - | - | - | - | - | - | - | - | - |
| ITH | 0.75 | - | - | - | - | - | - | - | - | - | 0.25 | - | - | - | - | - | - | - | - | - | - |
| CRE | 0.35 | - | - | - | - | - | - | - | - | - | 0.65 | - | - | - | - | - | - | - | - | - | - |
| ISL | 0.07 | - | - | - | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | - | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | - | - | - | 0.07 |
| MAU | - | - | - | - | - | - | - | - | - | - | 0.35 | - | - | - | - | - | - | - | - | - | 0.65 |
| KAU | - | - | - | - | - | - | - | - | - | - | 0.5 | - | - | - | - | - | - | - | - | - | 0.5 |
| OAH | - | - | - | - | - | - | - | - | - | - | 0.5 | - | - | - | - | - | - | - | - | - | 0.5 |
| NIH | - | - | - | - | - | - | - | - | - | - | 0.5 | - | - | - | - | - | - | - | - | - | 0.5 |
| LAN | - | - | - | - | - | - | - | - | - | - | 0.35 | - | - | - | - | - | - | - | - | - | 0.65 |
| MOL | - | - | - | - | - | - | - | - | - | - | 0.2 | - | - | - | - | - | - | - | - | - | 0.8 |
| VOL | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 1.0 |
| HIK | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 1.0 |
| FAL | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 1.0 |
| HAW | - | - | - | - | - | - | - | - | - | - | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | - |

*3)* Subtract the mean row from each row of FGT to get the zero mean FGT. Next, compute the covariance matrix of FGT and mark it as ∑(FGT).

*4)* For each pair of rows $r_i$, $r_j$ of Table 2 calculate their distance $(r_i - r_j) \sum(F_G{}^T) (r_i - r_j)^T$

Assuming that the query q = [Greek (0.5), Islands (0.5)] is a row vector, its distance from each row (ri) in Table 2 can be computed using the formula (q – ri ) ∑(FGT ) (q – ri )T, where the dimensionalities are equalised by adding zeros to the relevant locations. The stages involved in creating the distance table using the LSI approach are as follows:

1. Starting from the keyword-image matrix A, which is the transpose of Table 2, we perform a singular value decomposition $A = USV^T$ of this matrix for the desired dimensionality k. Since the dimensions in this experiment are not many we used the full dimensionality k = 21 and just a reduction to k = 10.

2. Since the columns of $SV^T$ represent the images in their LSI representation, we follow the standard LSI approach and calculate the distance between the columns of $SV^T$ as the cosine of their angle.

In the case of the query q = [Greek(0.5), Islands(0.5)], considered as a column vector, first it is projected on the image space by $q= q^TUS^{-1}$ and then its distance to all the images is measured by the cosine of its angle with each column of $SV^T$ .

Figure 4 displays the Precision versus Recall graphs for the LSI for k = 21 (no dimensionality reduction) and k = 10, as well as the MSI for n ¼ 1; 5; 10; 12; 14; and 15. The fact that the best LSI result, obtained as anticipated when no dimensional reduction is used, is incredibly subpar in comparison to the MSI results indicates that LSI is unable to acquire the latent features that are caught by way of the AMC utilised in MSI. When it comes to MSI, we can see how important the choice of n is to the method's accuracy since, as n grows, accuracy increases quickly until a maximum is reached at n = 10, at which point the best MSI result places nearly all of the photos in the correct class. The accuracy gradually decreases for n > 10 and barely approaches the best LSI result for n > 15. The distance table between the first 15 Greek (listed as G1-G15) and the first 15 Hawaiian (listed as H33-H47) images of Table 2 is shown in Fig. 3. Here, we can see the perfect score of the MSI, for n ¼ 10, where all 30 images are matched in the correct class for the first 9 rankings, in contrast to the LSI rankings where many images are matched in the incorrect of the two classes. Our attempt to apply a probabilistic method for semantic inference in sparse environments is justified by these results. Since the annotation was done implicitly by the click-through data, the Precision versus Recall measure of the suggested model yields nearly perfect scores for some mixing values. As a result, we observe that the proposed system dynamically adapts to the particular user group that is actually utilising the system. With MSI, there is less need to turn to external taxonomy systems in order to assign relevance metrics between keywords. As a result, there is less need to assess how well those systems work with the semantics of the real users of the system.

| IMG # | RANK MSI | RANK LSI | KEYWORDS | | | |
|---|---|---|---|---|---|---|
| 1 | 1 | 1 | GRE(0.5) | - | ISL(0.5) | - |
| 2 | 2 | 2 | GRE(0.4) | - | ISL(0.6) | - |
| 3 | 29 | 6 | GRE(0.3) | - | ISL(0.7) | - |
| 4 | 25 | 3 | GRE(0.8) | - | ISL(0.2) | - |
| 5 | 17 | 4 | GRE(0.2) | - | ISL(0.8) | - |
| 6 | 3 | 5 | GRE(0.7) | - | ISL(0.3) | - |
| 7 | 6 | 29 | GRE(0.2) | - | ISL(0.3) | CRE(0.5) |
| 8 | 7 | 25 | GRE(0.1) | - | ISL(0.3) | CRE(0.6) |
| 9 | 9 | 24 | GRE(0.3) | - | ISL(0.2) | CRE(0.5) |
| 10 | 15 | 11 | GRE(0.2) | - | ISL(0.1) | CRE(0.7) |
| 11 | 8 | 31 | GRE(0.4) | - | ISL(0.2) | ITH(0.4) |
| 12 | 20 | 58 | GRE(0.2) | - | ISL(0.3) | ITH(0.5) |
| 13 | 10 | 62 | GRE(0.1) | - | ISL(0.1) | MYK(0.8) |
| 14 | 5 | 63 | - | - | ISL(0.1) | MYK(0.9) |
| 15 | 4 | 38 | GRE(0.3) | - | ISL(0.1) | MYK(0.6) |
| 16 | 26 | 61 | GRE(0.1) | - | ISL(0.2) | MYK(0.7) |
| 17 | 16 | 12 | GRE(0.4) | - | ISL(0.1) | MYK(0.5) |
| 18 | 12 | 9 | GRE(0.4) | - | ISL(0.1) | SAN(0.5) |
| 19 | 13 | 7 | GRE(0.5) | - | - | SAN(0.5) |
| 20 | 18 | 18 | - | - | ISL(0.4) | SAN(0.6) |
| 21 | 11 | 17 | GRE(0.1) | - | ISL(0.1) | SAN(0.8) |
| 22 | 24 | 36 | GRE(0.3) | - | - | SAN(0.7) |
| 23 | 27 | 32 | GRE(0.3) | - | ISL(0.1) | SAM(0.6) |
| 24 | 28 | 30 | GRE(0.4) | - | ISL(0.2) | SAM(0.4) |
| 25 | 14 | 19 | GRE(0.3) | - | ISL(0.4) | SAM(0.3) |
| 26 | 19 | 40 | GRE(0.2) | - | ISL(0.1) | RHO(0.7) |
| 27 | 22 | 59 | - | - | ISL(0.1) | RHO(0.9) |
| 28 | 21 | 15 | GRE(0.2) | - | - | RHO(0.8) |
| 29 | 23 | 23 | GRE(0.4) | - | ISL(0.3) | RHO(0.3) |
| 30 | 31 | 8 | GRE(0.5) | - | - | HIS(0.5) |
| 31 | 30 | 20 | GRE(0.7) | - | - | TRA(0.3) |
| 32 | 32 | 55 | GRE(0.5) | - | - | GOD(0.5) |
| 33 | 58 | 64 | - | HAW(0.4) | - | FAL(0.6) |
| 34 | 62 | 41 | - | HAW(0.3) | - | HIK(0.7) |
| 35 | 38 | 50 | - | HAW(0.5) | - | VOL(0.5) |
| 36 | 61 | 49 | - | HAW(0.1) | ISL(0.5) | MOL(0.4) |
| 37 | 49 | 37 | - | HAW(0.2) | ISL(0.3) | MOL(0.5) |
| 38 | 46 | 16 | - | HAW(0.1) | ISL(0.6) | LAN(0.3) |
| 39 | 63 | 26 | - | HAW(0.5) | - | LAN(0.5) |
| 40 | 40 | 10 | - | - | ISL(0.5) | LAN(0.5) |
| 41 | 59 | 56 | - | HAW(0.3) | ISL(0.3) | LAN(0.4) |
| 42 | 47 | 22 | - | HAW(0.4) | ISL(0.2) | NIH(0.4) |
| 43 | 36 | 60 | - | HAW(0.4) | ISL(0.1) | NIH(0.5) |
| 44 | 55 | 42 | - | HAW(0.3) | - | NIH(0.7) |
| 45 | 64 | 47 | - | HAW(0.6) | ISL(0.2) | OAH(0.2) |
| 46 | 48 | 54 | - | - | ISL(0.1) | OAH(0.9) |
| 47 | 50 | 45 | - | HAW(0.3) | ISL(0.2) | OAH(0.5) |
| 48 | 41 | 13 | - | HAW(0.3) | - | OAH(0.7) |
| 49 | 56 | 28 | - | HAW(0.2) | ISL(0.3) | OAH(0.5) |
| 50 | 45 | 43 | - | HAW(0.4) | ISL(0.3) | KAU(0.3) |
| 51 | 42 | 52 | - | HAW(0.7) | ISL(0.1) | KAU(0.2) |
| 52 | 53 | 57 | - | HAW(0.2) | ISL(0.1) | KAU(0.7) |
| 53 | 60 | 51 | - | - | ISL(0.1) | KAU(0.9) |
| 54 | 52 | 21 | - | HAW(0.3) | ISL(0.2) | MAU(0.5) |
| 55 | 43 | 46 | - | - | ISL(0.4) | MAU(0.6) |
| 56 | 54 | 14 | - | - | ISL(0.3) | MAU(0.7) |
| 57 | 37 | 53 | - | HAW(0.7) | ISL(0.1) | MAU(0.2) |
| 58 | 44 | 27 | - | - | ISL(0.7) | MAU(0.3) |
| 59 | 51 | 48 | - | HAW(0.5) | ISL(0.5) | - |
| 60 | 57 | 34 | - | HAW(0.7) | ISL(0.3) | - |
| 61 | 39 | 44 | - | HAW(0.4) | ISL(0.6) | - |
| 62 | 35 | 33 | - | HAW(0.2) | ISL(0.8) | - |
| 63 | 33 | 35 | - | HAW(0.2) | ISL(0.8) | - |
| 64 | 34 | 39 | - | HAW(0.6) | ISL(0.4) | - |

Each row represents an image. The first column holds the image tag number, the second and third columns hold the ranking number with respect to the MSI and LSI distances from the query "Greek Islands," respectively for this image. The rest of the entries correspond to the steady state probabilities of the respective Markov chain representing each image. Only the zero entries that help for better illustration appear as dashes, the other zero entries have been omitted. The ranking for the MSI distance is implied by the AMC of (Fig. 2). Thirty of these images are shown in the distance table of (Fig.3) for rows 1-15 and 33-47of this table.

In order to provide a more compact representation of keywords and images in a space with fewer dimensions, a singular value decomposition method is used to start from a keyword/image frequency matrix. This mechanism correlates keywords that appear at the same image and images when they are annotated with the same keyword. To infer associations between keywords exclusively, our approach requires a keyword/keyword 1 step transition probability square matrix (AMC), which we then simply employ.

Fig.3.The distance table between 30 images of two classes. ImagesG1 G15 are the first 15 images of Table2 and belong to the class Greek whereas images H33-H47 are represented by the corresponding 15 rows of the same table and belong to the class Hawai. The proposed distance (MSI) (upper number) and the latent semantic indexing (lower number) are compared. The first nine best matches are marked with the respective super scripts. MSI achieves better results since all the first nine matches are correct (the images match to images in the same class), where many of the first nine matches are not correct for the LSI. The precision versus recall diagram for this experiment for all the 64 images of Table2 is shown in Fig.4

The covariance matrix will be projected onto the difference between the keyword/image frequency matrix and the picture vectors that are represented. The way the system behaves in relation to power n reveals the mechanism underlying MSI. We observe that outcomes get better as n grows, with n=10 yielding the best Precision Recall result. This result demonstrates that by propagating relevant connections based on the available data, rather than introducing new observations to the system, we were able to increase retrieval. The suggested system leverages the convergence of the Markov chain (by increasing the AMC to the proper power) to build on the current observations by mining deeper qualitative inferences between keywords at the point when LSI or pLSI approaches would need more data to improve outcomes.

*B. Comparison to pLSI Using External Annotation*

Many of the benefits of the suggested method (MSI) disappear since the AMC cannot be formed when we have no control on the annotation technique used for the photos or access to the user queries that were used to annotate the images (if such a method has even been employed). However, we still wish to tweak our method to allow AMC creation with dimensionality reduction whenever just a keyword image matrix is provided, and include such scenarios in the proposed methodology and compare with pLSI. Our approach is based on the hypothesis that the AMC was progressively built through user inquiries. Techniques for aggregating and disaggregating the Markovian states—which arise when groups of states materialise as blocks—can be used in this case to minimise dimensionality.

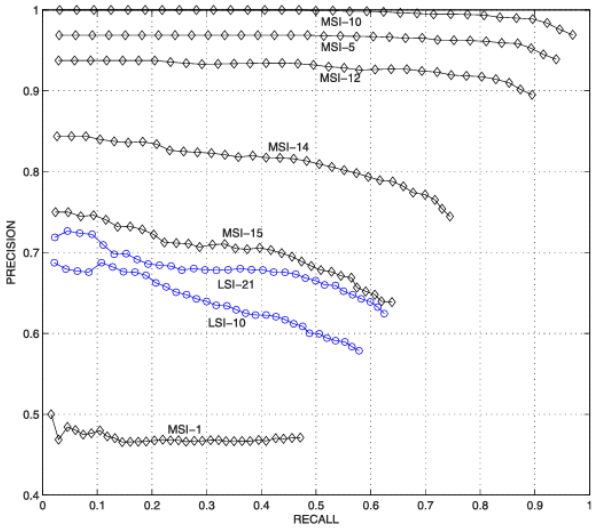


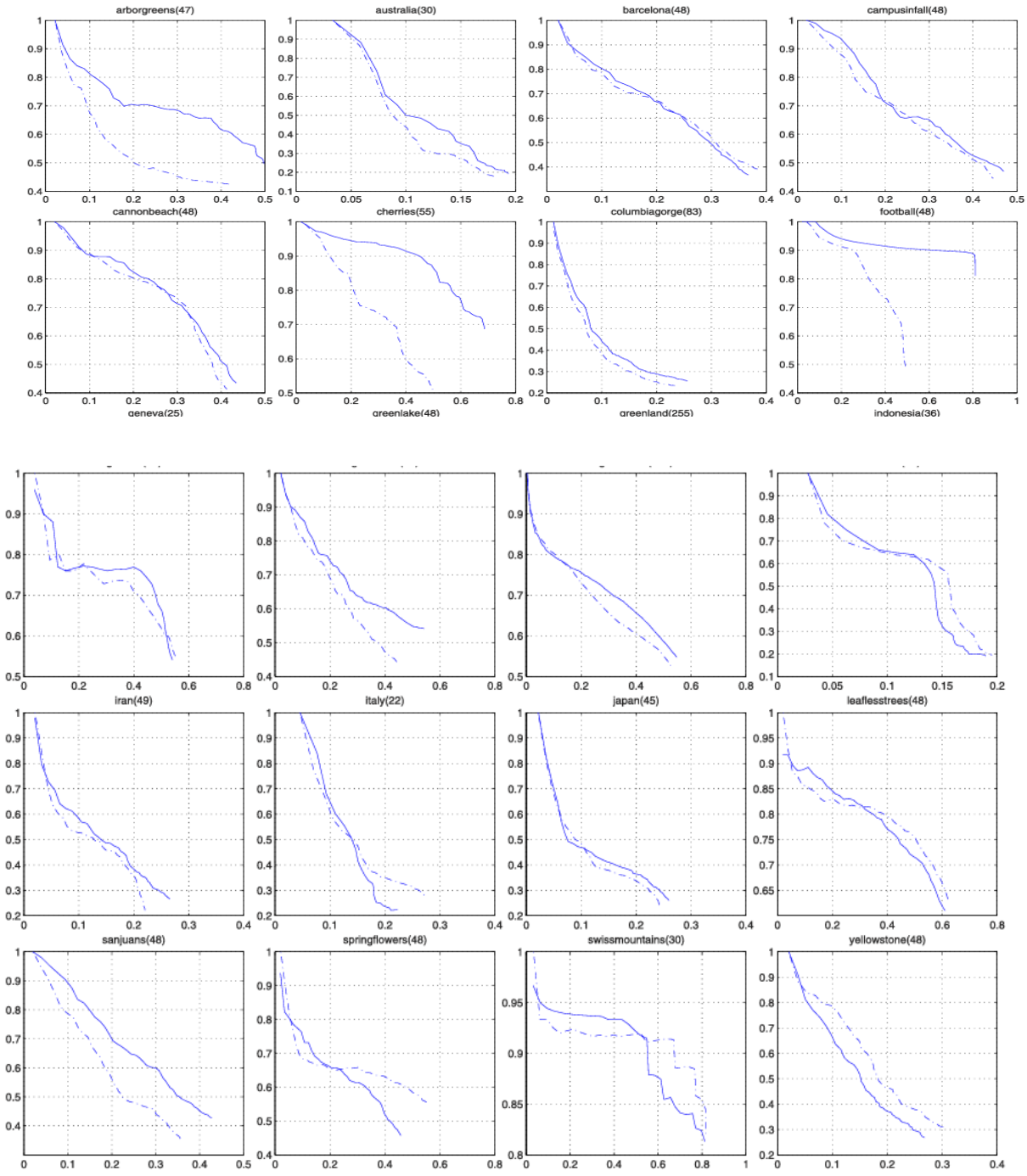

Fig. 4. Precision versus Recall comparisons between Latent Semantic Indexing and the proposed system for various parameters of the two algorithms. The experiment is performed on the 64 images of Table 2. The first 32 images belong to the class Greek and the next 32 to the class Hawaiian. The diamond graphs represent the results of the proposed distance (implied by the general process of Table 1) for the cases of n=1, n=5, n=10, n=12, n=14, n=15. The circle graphs represent the results when the LSI distance is used with the figure/ keyword frequency matrix of Table 2 for the parameters of k=10 and k=21. See text for more.

Diagonal AMC. Details on fast such methods can be found.

In the event that we do not have access to the query logs, we will need to build the AMC directly from the image annotations, considering the keyword set for each image as a query unique to that image. This will result in the construction of an AMC whose dimensions match the total number of keywords the system has seen. We must apply a reduction of dimensionality to this AMC and project the images onto this reduced space before measuring their distance in order to provide a fair comparison to LSI. Selecting the k principal components—where k is the desired dimension following reduction—will help you do this. Given that AMC is a square matrix, we can use the eigen decomposition method to determine which k primary components, or AMC k of dimension k, best approximate the AMC in these dimensions. Thus, the processes for creating the distance table for the modified MSI technique are as follows:

Step 2 above can be skipped if raising to a power is not necessary. In this scenario, MSI operates similarly to LSI as the annotation's origin is unclear and there is no Markovian connection to the keyword data. When no power is delivered, the updated suggested method and LSI differ primarily in the AMC matrix. To decrease the dimensionality of the keyword space, the MSI approach does, in fact, use a square keyword/keyword matrix rather than the non-square keyword/image matrix of LSI. In comparison to LSI, the suggested approach (MSI) has the benefit that the keyword/keyword matrix is square, allowing for the use of more optimal algorithms using eigen decomposition (as opposed to singular value decomposition, which LSI employs on non-square matrices). In the event of external annotations with an unclear provenance, we compare the modified suggested distance (MSI) to the commonly used and highly recognised probabilistic Latent Semantic Indexing approach for indexing based on latent variables. We use the ground-truth database, which consists of 1,109 images divided into 20 classes with roughly 50 images apiece. The images are annotated with a total of 437 keywords, with each annotation consisting of a text string with up to 25 keywords.

The classes Cambridge and Barcelona2 were left out because the images in those classes are only there without any annotations. We compare the results using Precision versus Recall graphs using this data set. According to Peter Gehler, our pLSI implementation was accessible online at. The best matching outcomes for every class are displayed in Fig. 5. At k = 200, both approaches produced the best matching curve, and it is here that we see that MSI (solid line) outperforms the other way for the majority of the classes. Figure 6 shows the overall Precision versus Recall for all classes combined. It is possible to confirm that, for k = 200 (the number of dimensions after reduction), both approaches produce the greatest results, but MSI performs better. Both approaches yield curves in the region below k = 300 for k > 200, and for k < 200, the rates of degradation of the findings are equal for both approaches. Even if the annotation process is not controlled in this experiment, making it impossible to fully disclose the advantages of the suggested system, we can nevertheless observe that MSI outperforms pLSI at all dimensionalities, even at 200 dimensions.



*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 12 Issue VI June 2024- Available at www.ijraset.com*

*C. Conceptual Comparisons and Scalability Issues*

By using Singular Value Decomposition (SVD) to obtain an L2-optimal matrix approximation, LSI reduces the dimensionality of the system. Even though SVD is widely known, its application in this situation is a little haphazard because the Frobenious norm's usage for optimisation lacks a clear statistical interpretation. However, the optimisation is accomplished via the Expectation Maximisation method, which has its own issues with overfitting, sensitivity to initial circumstances, and generalisation on new, unseen data. In contrast, pLSI, which is based on the aspect model, establishes a correct statistical model. Conversely, the suggested MSI/AMC method includes an optimisation phase based on the convergence properties of the Markovian chain as represented by the AMC kernel.

Since these can be found using an eigen decomposition of the AMC, the optimisation step in the suggested model can be directly interpreted in the keyword-relevance linkages in addition to being optimal and stochastic with regard to the connectivity of the Markovian states.

The automatic indexing and query-based retrieval activities are handled by both LSI and pLSI using a typical cosine matching. Despite being often used and acknowledged, the cosine distance has no clear relationship to

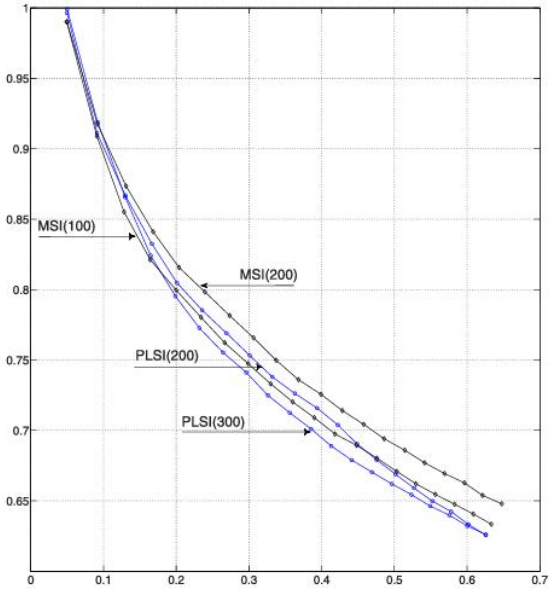



Fig. 6. Precision versus Recall comparisons between Probabilistic Latent Semantic Indexing and the proposed system (MSI) for various parameters of the two algorithms. The experiment is performed on the 1,109 images of the ground-truth database. The diamond graphs represent the results of the proposed distance for the cases of k=100 and k=200. The circle graphs represent the results when the pLSI distance is used with k=200 and k=300. In Fig. 5, the Precision versus Recall is shown for each individual class separately.

To the fundamental models in these two approaches. Furthermore, when the norm of both vectors is near zero, numerical issues need to be resolved. On the other hand, the approach suggested in this study (MSI/AMC) integrates automatic indexing and query matching duties across the entire framework, with the distance calculated in an optimal way that is easily interpreted in connection with the Markovian state clustering.

The process of thresholding all but the k greatest singular values in LSI reduces dimensionality. In the aspect model formulation of pLSI, k is represented by the bottleneck random variable. However, the value of k is selected a priori; typically, several values of k are analysed and combined. The suggested MSI model allows for a consideration of a reduction in dimensionality in the keyword clusters, which is accomplished by the convergence technique. In that scenario, the number of clusters is the analogue of k and is determined automatically from the convergence threshold.

Additionally, a comparison based on computational complexity points to benefits for MSI. Since, according to (6), only the AMC kernel's eigenvalues need to be raised to the appropriate power, the powers of the AMC kernel in the first implementation and the PCA treatment in the second implementation that are involved in the corresponding MSI optimisation steps are both computed accurately and efficiently. On the other hand, the EM algorithm is an iterative process that simply ensures the likelihood function's local maximum.

The clustering of the state space, which groups the states into relevant groups, serves as the inference engine for the MSI technique. One might investigate the degree of this clustering in relation to the system's size in order to assess scalability. In the context of Nearly Completely Decomposable Markov Chains and their rapid convergence to equilibrium, the degree of state clustering in Markov Chains has been extensively investigated and is commonly referred to as the coupling degree. The degree of state clustering that can be assumed in terms of state connectedness is quantified by the coupling degree of the chain.

Higher coupling degrees in these situations result in slower convergence, which is undesirable if one is interested in the chain's equilibrium state. However, in terms of the MSI distance, higher coupling degrees indicate better connectivity differentiation between states and, consequently, better system performance. In most real-life applications, the Markov chains we come across are sparse and they more or less possess some structure, that is, the ratio of the number of nonzero elements to the total number of elements in the underlying chain is small; moreover, the magnitude and location of these nonzero elements is not random, hence higher coupling degrees can be expected as the size of the system increases.

## VI. CONCLUSION

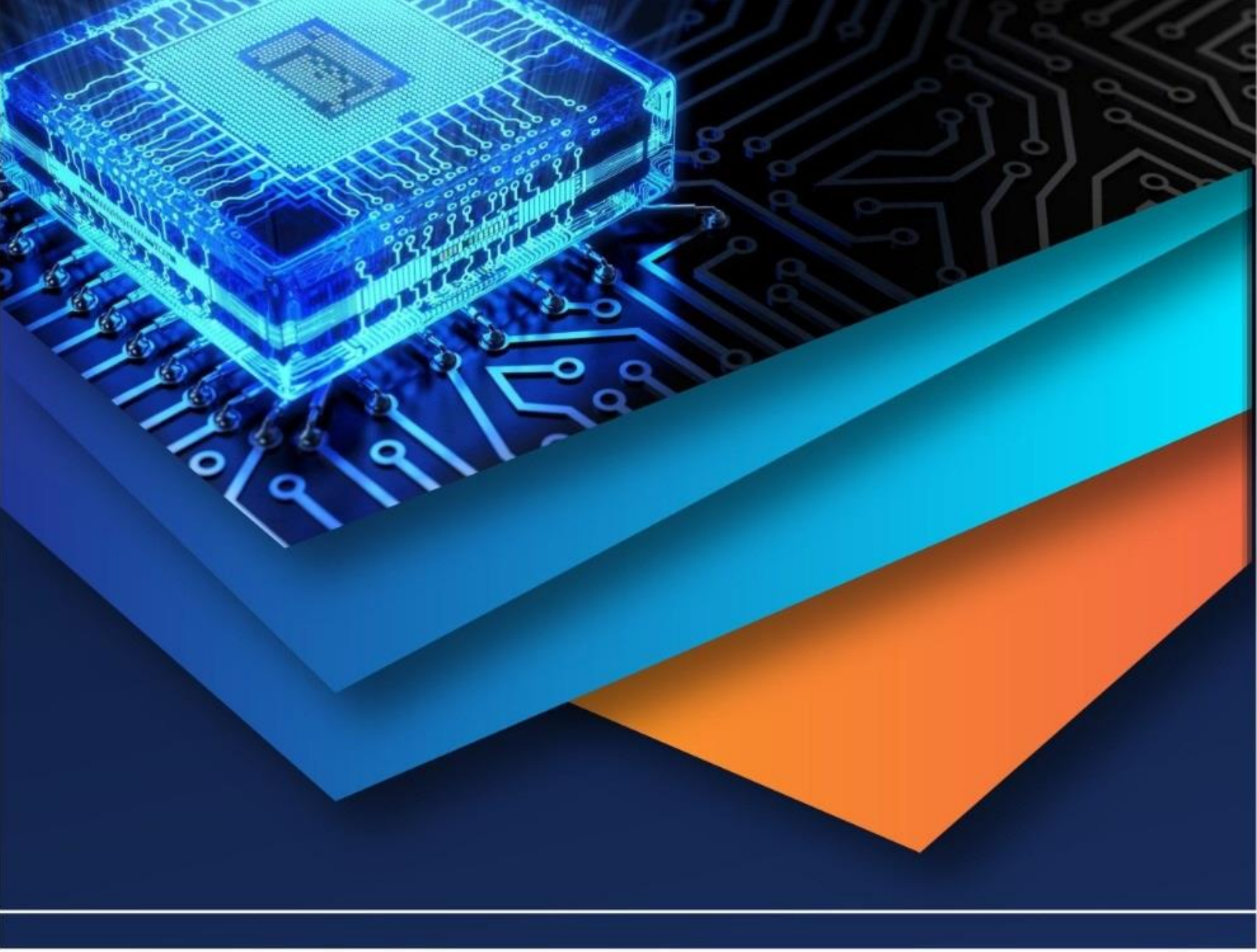
By characterising keyword importance as a connection measure between Markovian states that are modelled after user queries, we introduced the Markovian Semantic Indexing, a novel approach to mining user queries. The queries of the same consumers who would be serviced by the proposed system are used to dynamically train it. As a result, targeting is more precise than in other systems that define keyword relevance through external, non-dynamic, or non-adaptive methods. Using an Aggregate Markovian Chain, a stochastic distance was created in the form of a generalised euclidean distance, and it turned out to be optimal with regard to specific Markovian connectivity measures that were designed specifically for this purpose. A theoretical comparison between the suggested method (MSI) and Latent Semantic Indexing and Probabilistic Latent Semantic Indexing indicated several advantages.

Studies have demonstrated that MSI performs better when retrieving information from poorly annotated picture data sets. When 64 photos were collected from Google Image Search and transparently annotated using the suggested system, the results showed that the MSI method had certain advantages over LSI, primarily in terms of collecting images with more complex connections than just keyword cooccurrence. After adapting the suggested approach to include AMC creation and dimensionality reduction in external annotations, a second comparison to pLSI was carried out using the ground-truth annotated collection of 1,109 images. For this experiment, the Precision versus Recall findings showed that MSI outperforms pLSI in all dimensionalities up to 200 dimensions.

## REFERENCES

[1] S. Santini and R. Jain, “Similarity Measures,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, no. 9, pp. 871-883, Sept. 1999.
[2] 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.
[3] comScore’s Report Article, “Comscore’s Qsearch 2.0 Service,” comScore’s Report Article, www.comscore.com, 2007.
[4] B.J. Jansen, A. Spink, and T. Saracevic, “Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web,” Information Processing and Management, vol. 36, no. 2, pp. 207-227, 2000.
[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.
[6] 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.
[7] 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.
[9] M.W. Berry, S.T. Dumais, and G.W. O’Brien, “Using Linear Algebra for Intelligent Information Retrieval,” SIAM Rev., vol. 37, no. 4, pp. 573-595, 1995.
[10] T. Hofmann, “Probabilistic Latent Semantic Indexing,” Proc. 22nd Int’l Conf. Research and Development in Information Retrieval (SIGIR ’99), 1999.
[11] T. Hofmann, “Unsupervised Learning by Probabilistic Latent Semantic Analysis,” Machine Learning, vol. 42, no. 1/2, pp. 177- 196, 2001.
[12] D.M. Blei and A.Y. Ng, and M.I. Jordan, “Latent Dirichlet Allocation,” J. Machine Learning Research, vol. 3, pp. 993-1022, 2003.
[13] T.L. Griffiths and M. Steyvers, “Finding Scientific Topics,” Proc. Nat’l Academy of Sciences USA, vol. 101, no. suppl. 1, pp. 5228-5235, 2004.
[14] M. Steyvers, P. Smyth, M. Rosen-Zvi, and T. Griffiths, “Probabilistic Author-Topic Models for Information Discovery,” Proc. 10th ACM SIGKDD Conf. Knowledge Discovery and Data Mining, 2004.
[15] 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.
[16] T.-T. Pham, N.E. Maillot, J.-H. Lim, and J.-P. Chevallet, “Latent Semantic Fusion Model for Image Retrieval and Annotation,” Proc. 16th ACM Conf. Information and Knowledge Management (CIKM), 2007.
[17] 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.
[18] F. Monay and D. Gatica-Perez, “On Image Auto-Annotation with Latent Space Models,” Proc. ACM Int’l Conf. Multimedia (MM), 2003.

[19] K. Barnard and D. Forsyth, "Learning the Semantics of Words and Pictures," Proc. Int'l Conf. Computer Vision, vol. 2, pp. 408-415, 2001.
[20] L.-J. Li and G. Wang, and L. Fei-Fei, "OPTIMOL: Automatic Online Picture Collection via Incremental Model Learning," Int'l J. Computer Vision, vol. 88, no. 2, pp. 147-168, 2010.
[21] 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.
[22] G. Shafer, P.P. Shenoy, and K. Mellouli, "Propagating belief Functions in Qualitative Markov Trees," Int'l J. Approximate Reasoning 1, vol. 4, pp. 394-400, 1987.
[23] L.G. Shapiro, "GroundTruth Database," http://www.cs. washington.edu/research/imagedatabase/groundtruth/, Univ. of Washington, 2012.
[24] J. Pearl, Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann, 1988.
[25] L.D. Lowrance, T.D. Garvey, and T.M. Strat, "A Framework for Evidential Reasoning Systems," Proc. Fifth Nat'l Conf. Artificial Intelligence (AAAI '86), pp. 896-901, 1986.
[26] U. Montanari, "Networks of Constraints, Fundamental Properties and Applications to Picture Processing," Information Science, vol. 7, pp. 95-132, 1974.
[27] W. Woods, Representation and Understanding, D. Bobrow and A. Collins, eds. Academic Press, 1975.
[28] R.O. Duda, P.E. Hart, and N.J. Nilsson, "Subjective Bayesian Methods for Rule-Based Inference Systems," Proc. Nat'l Computer Conf. and Exposition (AFIPS), vol. 45, pp. 1075-1082, 1976.
[29] R. Schank, "Conceptual Dependency: A Theory of Natural Language Understanding," Cognitive Psychology, vol. 4, pp. 552- 631, 1972.
[30] J.F. Sowa, Conceptual Structures: Information Processing in Mind and Machine. Addison-Wesley, 1984.
[31] O. Tuzel, F. Porikli, and P. Meer, "Pedestrian Detection via Classification on Riemannian Manifolds," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30, no. 10, pp. 1713-1727, Oct. 2008.
[32] R. Howard, Dynamic Probabilistic Systems. John Wiley and Sons, Inc., 1971.
[33] G. Zhen, Z. Shenghuo, C. Yun, Z. Zhongfei, and G. Yihong, "A Latent Topic Model for Linked Documents," Proc. 32nd Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '09), 2009.
[34] W.J. Stewart, Numerical Solution of Markov Chains. Princeton Univ. Press, 1994.
[35] http://people.kyb.tuebingen.mpg.de/pgehler/code/index.html, 2012.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY