Markovian Semantic Indexing:Application to Online Image Retrieval System

K. Naresh Kumar, K. Krishna Reddy

***Abstract:*-Markovsemantic indexing algorithm based on the controlled Markov chain modeling framework. Controlled Markov chain models are used to describe the temporal evolution of low-level visual descriptors extracted from the semantic indexing model. Propose a semantic indexing algorithm which uses both text and image retrieval system.The entire user Queries selected by randomly. An image retrieval system is a computer system for browsing, searching and retrieving images from a large image database. The new method, that we call Markovian Semantic Indexing (MSI), is presented in the context of an online image retrieval system.**

Index Terms:-Semantic indexing,controlled Markovchains

1. INTRODUCTION

The design of efficient indexing techniques suitable to retrieve relevant information through text based and image caption documents is necessary to enable widespread use and access to richer and novel information sources. Allowing for possible automatic procedures to semantically indexing and annotate the images. To face the problem of semantic indexing, a man uses its cognitive skills, while an automatic system can face it by adopting a two-step procedure: in the first step, some low-level indices are extracted in order to represent low-level information in a compact way; in the second step, a decision-making algorithm is used to extract a semantic index from the low-level indices.

The new method, that we call Markovian Semantic Indexing (MSI) is presented in the context of an online image retrieval system. The properties of MSI make itparticularly suitable for ABIR tasks when the per imageannotation data is limited. The characteristics of the methodmake it also particularly applicable in the context of onlineimage retrieval systems.

Annotation-Based Image Retrieval (ABIR) systems are an attempt to incorporate more efficient semantic content into both text-based queries and image captions (i.e.. Google Image Search, Yahoo! Image Search). 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.

***K. Naresh Kumar**, CSE, Anurag Engineering College, Kodad, AP, India, 9849777621.*
***K. Krishna Reddy**, CSE, Anurag Engineering College, Kodad, AP, India, 9701155482.*

2. RELATED WORK

Attempts of applying LSI/pLSI-based techniques to discover a more reliable concept association in ABIR systems have been reported in the context of online image retrieval systems. Attempts of applying LSI/pLSI-based techniques to discover a more reliable concept association in ABIR systems have been reported. The problem of automatic image annotation is closely related to that of content-based image retrieval. Since the early 1990s, numerous approaches, both from academia and the industry, have been proposed to index images using numerical features automatically-extracted from the images. Probabilistic Latent Semantic Indexing and latent semantic indexing methods to retrieve the images easily. The new method is shown to possess certain theoretical advantages and also to achieve better Precision versus Recall results when compared to Latent Semantic Indexing (LSI) and probabilistic Latent Semantic Indexing (pLSI) methods in Annotation-Based Image Retrieval (ABIR) tasks.

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.

Attempts of applying LSI/pLSI-based techniques todiscover a more reliable concept association in ABIR system the probabilistic Latent Semantic Indexing (pLSI) as an alternative to projection (LSI) or clustering methods for document retrieval. Latent Dirichlet Allocation (LDA) was proposed by Blei et al. To address the limitations of pLSI regarding generalization and over fitting while Griffiths and Steyvers incorporated a Markov chain Monte Carlo technique to LDA.

2.1 Our Contribution

The methodology proposed in this work encompasses more reliable and 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.

Furthermore, even though automatic annotation andannotation-based image retrieval systems have been presentedin the literature the proposed system is novel in theway it unifies these two tasks. Indeed both the automaticannotation and retrieval tasks are assumed in the implicituser interaction context, for dynamically mining semanticstowards qualitative probabilistic reasoning. This has implicationsin the targeting aspect since the training is performeddynamically by the same users that are actually using thesystem. The unified

Markovian setup behind the proposedsystem allows the retrieval technique to benefit from theunderlying structure of the annotation data; at the same timethe annotation data acquires concrete stochastic interpretationthrough 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.

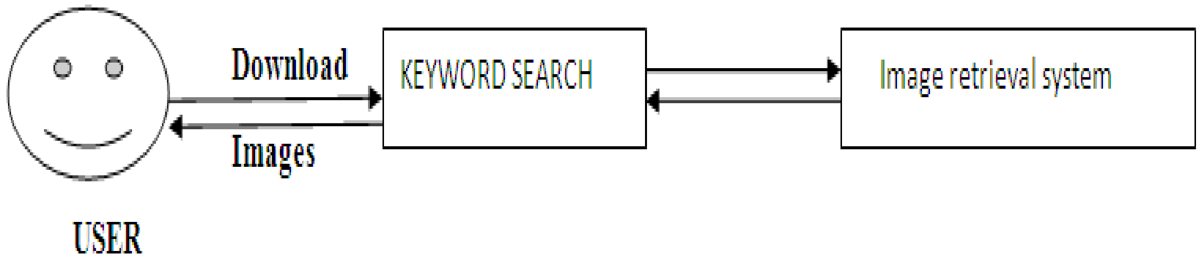


Figure 2.1.1 Diagram for content-based image retrieval system

3. THE PROPOSED APPROACH

The proposed approach will be presented in the frameworkof an online image retrieval system (similar to Googleimage search) where users search for images by submittingqueries that are made of keywords. The queries formed bythe users of a search engine are semantically refined, thekeywords representing concise semantics when comparedto text in documents or other vocabulary related presentations.The aim is to improve user satisfaction by returningimages that have a higher probability to be accepted(downloaded) by the user. The assumption is that the userssearch for images by issuing queries, each query being anordered set of keywords. The system responds with a list ofimages. The user can download or ignore the returnedimages and issue a new query instead. During the trainingphase of the system the images are considered with noannotation. As the users issue queries and pick images thesystem annotates the images in an automatic manner and atthe same time establishes relevance relations between thekeywords as will be explained later on in the manuscript.The user never annotates the images explicitly, this happensby the system transparently from the user. At the testingphase the system uses the annotations available from thetraining phase but also the keyword relevance probabilityweights also evaluated during the training phase to returnimages that better reflect the user's preferences and improveuser satisfaction. This interactive procedure has implicitconsequences that we exploit one by one in a step by stepconstruction of the proposed system.\

3.1 Random Processes

A random process is a collection of random variables indexed by some set I, taking values in some set S.

I is the index set, usually time.Ex: Z_+, R, R_+.

S is the state space, e.g.Z+, Rn , { 1, 2 , . . . , n}, { a, b, c }.

We classify random processes according to both the indexset (discrete or continuous) and the state space (finite,countable or uncountable/continuous).

3.2 Markov Processes

A random process is called a Markov Process if, conditional on the current state of the process, its future is independent of its past. More formally, X(t) is Markovian if has the followingproperty:

$P(X(t_n) = j_n \mid X(t_n-1) = j_n-1, \ldots, X(t_1) = j_1) = P(X(t_n) = j_n \mid X(t_n-1) = j_n-1)$

For all finite sequences of timest1< . . . <tn∈Iand ofstatesj1, . . . , jn∈S

3.3 Time Homogeneity

A Markov chain (X(t))is said to be time-homogeneous if P(X(s+t) =j | X(s) =i) is independent of s . When this holds, putting s = 0 gives

P (X(s + t) =j | X(s) =i) =P(X(t) =j | X(0) =i).

A Markov chain (X(t)) is said to be time-homogeneous if P (X(s + t) = j | X (s) = I) is independent of s. When this holds, putting s = 0 givesP (X (s + t) = j | X (s) = I) = P (X (t) = j | X (0) = I). Probabilities depend on elapsed time, not absolute time.

4. GEOMETRICINTERPRETATION AND OPTIMALITY PROPERTIES OF THE MSI DISTANCE

The proposed MSI distance dðx; yÞ (Definition 1) can be viewed as measuring the total variance of the rows of FGTwhen projected on the direction of the difference, ð_x _ _yÞ, of the two images. The FGT is calculated once from all thedata and thus the direction defined by the vector difference of the probability distributions of the two particular images is actually deciding their in between distance. To get further geometric and stochastic interpretations of the MSI distance one has to acquire insight into the mechanics of the Markovian convergence that produces the FGT and its relation to certain directions in the keyword space. In the remaining of this section, we will investigate the geometrical meaning of the convergence process and its relation to the 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 reader is referred to the Appendix for a brief introduction.

4.1 Discrete-time Markov chains

Consider the time-homogenous model where the transition probabilities are constant over time. The transition probability matrix P(t) contains the probabilities for the transitions. Since the probabilities for the time-homogenous model are constant, the probability matrix could simply be written as P (Figure 4-1). The rows represent the current health state and the columns represent the future state. The probabilities are described as pij where p is the probability of moving from state i to state j for any given cycle.

$$P = \left(\begin{array}{cc|cccc} & & & j & & \\ & & 0 & 1 & 2 & 3 \\ \hline & 0 & 0 & 1 & 0 & 0 \\ i & 1 & 1/3 & 0 & 2/3 & 0 \\ & 2 & 0 & 2/3 & 0 & 1/3 \\ & 3 & 0 & 0 & 1 & 0 \end{array} \right)$$

Figure 4-1 Probability transition matrix for a time homogeneous 3-state Markov model.

The sum of the row probabilities equal one since each health state is independent of the other and an animal must move to one of the three states. The diagonals represent the probability of staying in the same health state. A state is considered absorbing when the probability of leaving a state is zero. For example, being dead is an absorbing state.

4.2 Continuous Time

Continuous time is slightly more difficult to deal with as there is no real equivalent to the one-step transition matrix from which one can calculate all quantities of interest.The study of continuous-time Markov chains is based on the transition function. The transition between states is viewed as a rate for a continuous-time Markov process. The transition rate does not depend on the length of the observation interval since it is the number of transitions that occur per unit time. The transition intensity (rate) matrix Q(t) contains components qij which are transition rates from state i to j. Since the rates for a time-homogenous Markov process are constant, the rate matrix
could simply be written as Q (Figure 4-2).

$$\mathbf{Q} \overset{\text{def}}{=} \begin{bmatrix} -q_0 & q_{0,1} & \cdots & q_{0,j} & \cdots \\ q_{1,0} & -q_1 & \cdots & q_{1,j} & \cdots \\ \vdots & \vdots & \cdots & \vdots & \cdots \\ q_{i,0} & q_{i,1} & \cdots & q_{i,j} & \cdots \\ \vdots & \vdots & \cdots & \vdots & \cdots \end{bmatrix}.$$

Figure 4-2 Transition intensity (rate) matrix for a time-homogenous 3-state Markov model.

4.3 Time Non-homogeneous Markov Process

Both discrete time and continuous time Markov processes may be time non homogeneous. The relationship between the transition rates and probabilities are more complex and computing probabilities is difficult. Similar to time homogeneous Markov processes, the transition probability depends on the observation interval but the transition rate does not. Because of their complexity, time non-homogeneous Markov process will not be addressed further in this work.

5. EXPERIMENTAL RESULTS

The first experiment is a comparison to LSI, since the limited number of images used in this experiment does notpermit reliable comparison to pLSI. The full features of theproposed distance (MSI) are demonstrated in this experimentsince the generative process of the aggregate Markov chainduring the automatic annotation of images was available tous as is explained later on. Sixty four images that form twointuitive classes were used for this experiment, 32 imagesrelated to the term Greek and considered to belong to the firstclass, and 32 images related to the term Hawaiian areconsidered to belong to the second class. First, the distanceof the 64 images from the query Greek Islands is calculated andranked for both methods and the results are examined. Then,a complete distance table is built for all the in-betweendistances of these 64 images using both methods.A precisionversus recall diagram is presented for comparison.

1. Starting from the Markov kernel P of Table 1 add a small quantity to all the one-diagonal elements (elements lying on the superdiagonal) and subtract it from any random nonzero element in the same line. This way we convert the process to monodesmic (chain), according to Section 4.1 while keeping the matrix stochastic and without altering the process statistics.
2. Calculate the expected fractional occupancy matrix FG ¼ 1 ðnþ1ÞPn k¼1 Pk at the desired n, where n is estimated according to the discussion in Section 3. We present results for n ¼ 1; n ¼ 5, n ¼ 10, n ¼ 12, n ¼ 14, n ¼ 15. For calculating the powers of P there are 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.
3. Calculate the zero mean FG T by subtracting the mean row from each row of FG T and then calculate the covariance matrix of FG T and denote it _ðFG T Þ.

The Markov Kernel that Corresponds to the Chain

	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	–

5.1 Comparison to pLSI Using External Annotation

When we have access to the user queries that were used to annotate the images (if such a method has even been used), many of the advantages of the proposed method (MSI) vanish since the AMC cannot be constructed. Nevertheless, we still want to incorporate such cases in the proposed approach and compare to pLSI by modifying our method to permit AMC construction with dimensionality reduction whenever only a keywordimage matrix is available.

When we have no access to the query logs, the AMC will have to be constructed from the image annotations themselves,treating each images' set of keywords as a query related to this image. By doing this, an AMC will be constructed with dimensions equal to the total number of keywords seen by the system. For a fair comparison to LSI,we will have to apply a reduction of dimensionality to this AMC and then project the images on this reduced spacebefore we measure their distance. This can be achieved by choosing the k principal components, where k the desired dimensions after the reduction. Because AMC is a square matrix, we can proceed with the eigen decomposition and choose the k principal components that best approximate the AMC in these dimensions, ending up with an AMCk of dimension k.

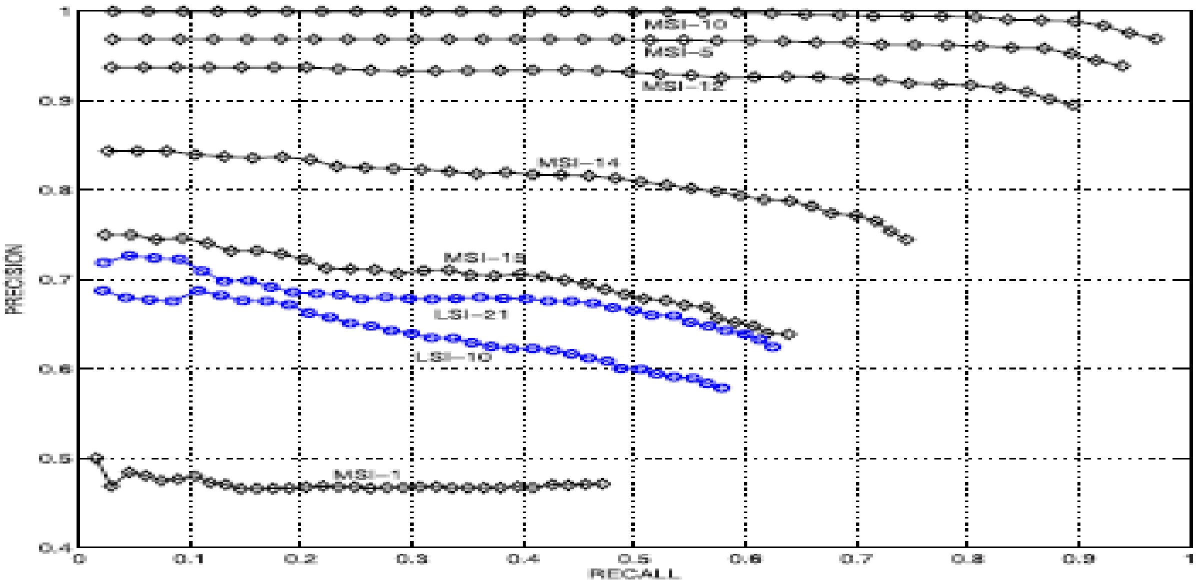
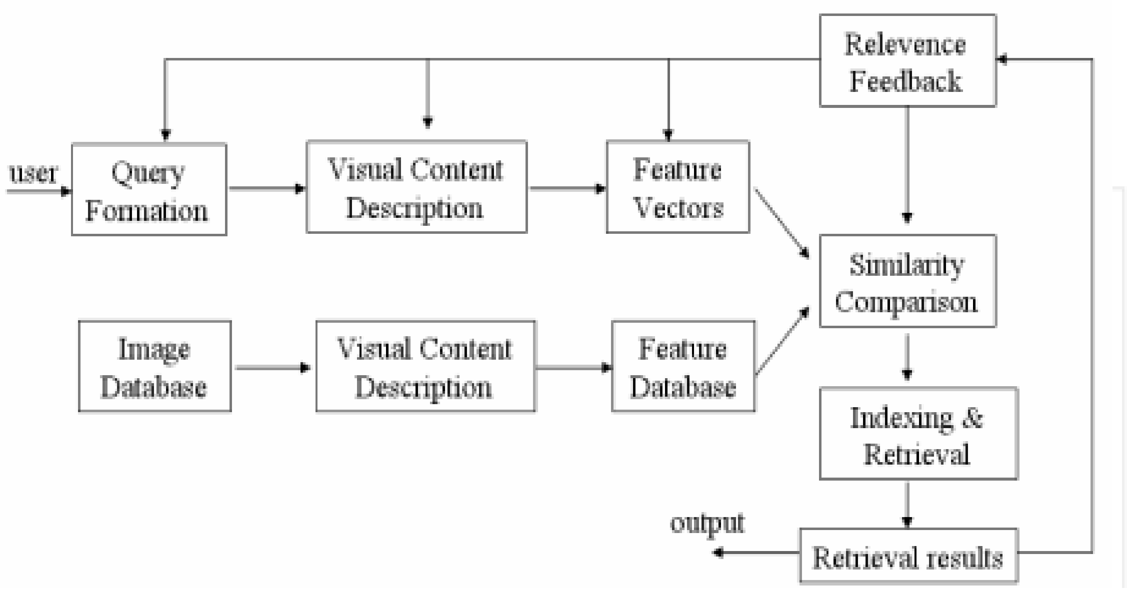


Fig 5.1: Precision versus Recall comparisons between Latent Semantic Indexing and the proposed system

5.2 Content-based image retrieval system

Content-based image retrieval uses the visual contents of an image such as color, shape, texture, and spatial layout to represent and index the image. In typical content-based image retrieval systems (Figure 4), the visual contents of the images in the database are extracted and described by multi-dimensional feature vectors. The feature vectors of the images in the database form a feature database. To retrieve images, users provide the retrieval system with example images or sketched figures. The system then changes these examples into its internal representation of feature vectors. The similarities /distances between the feature vectors of the query example or sketch and those of the images in the database are then calculated and retrieval is performed with the aid of an indexing scheme. The indexing scheme provides an efficient way to search for the image database. Recent retrieval systems have incorporated users' relevance feedback to modify the retrieval process in order to generate perceptually and semantically more meaningful retrieval results. In this chapter, weintroduce these fundamental techniques for content-based image retrieval.



5.3 Image Content Descriptors

Generally speaking, image content may include both visual and semantic content. Visual content can be very general or domain specific. General visual content include color, texture, shape, spatial relationship, etc. Domain specific visual content, like human faces, is application dependent and may involve domain knowledge. Semantic content is obtained either by textual annotation or by complex inference procedures based on visual content. This chapter concentrates on general visual contents descriptions. Later chapters discuss domain specific and semantic contents. A good visual content descriptor should be invariant to the accidental variance introduced by the imaging process (e.g., the variation of the illuminant of the scene). However, there is a tradeoff between the invariance and the discriminative power of visual features, since a very wide class of invariance loses the ability to discriminate between essential differences. Invariant description has been largely investigated in computer vision (like object recognition), but is relatively new in image retrieval.

5.4 Conceptual Comparisons and Scalability Issues

LSI achieves a reduction of dimensionality based on a L2- optimal approximation of matrices by means of Singular Value Decomposition (SVD). While SVD is well understood, its use in this context is somewhat ad hoc since there is not direct statistical interpretation in the use of the Frobenious norm for optimization. pLSI on the other hand which is based on the aspect model, defines a proper statistical model but the optimization is achieved through the Expectation Maximization algorithm which has its own problems of over fitting, sensitivity to initial conditions, and generalization on new unseen data.

Both LSI and pLSI treat the automatic indexing and thequery-based retrieval tasks by means of a standard cosinematching. While the cosine distance is widely used andgenerally accepted it has no direct interpretation in relationto the underlying models in both of these methods. Furthermore, numerical problems have to be addressed when the norm of both vectors is close to zero. The method proposed in this work (MSI/AMC) on the other hand, incorporates automatic indexing and query matching tasks within the whole framework, the distance being defined in an optimal manner, directly interpretable in relation to the clustering of the Markovian states.

The inference engine of the MSI approach lies in the clustering of the state space, since this clustering arranges the states into groups of relevance. In terms of scalability, one could examine the degree of this clustering with respect to the size of the system. The degree of state clustering in Markov Chains has been thoroughly studied, usually referred to as the coupling degree in the context of Near Completely Decomposable Markov Chains and their fasterconvergence to equilibrium. The coupling degree of the chain, quantifies the degree of state clustering one can assume in terms of state connectivity.

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.

6. CONCLUSION

The Monrovian Semantic Indexing, a new method for mining user queries by defining keywordrelevance as a connectivity measure between Monrovian states modeled after the user queries. The proposed system is 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 nondynamic or nonadaptive nature to define keyword relevance.

A Markov process is a random process in which the future is independent of the past, given the present. Thus, Markov processes are the natural stochastic analogs of the deterministic processes described by differential and difference equations. They form one of the most important classes of random processes.

APPENDIX

Constitute a Markov Chain (for example, Çinlar, 1975), which for large values of t has a stationary distribution that is precisely px. Consequently, for t large, t converges in distribution to a random vector whose distribution is px, and the elements of t converge in distribution to random variables whose distributions are the marginal distributions of px. Some regularity conditions are involved, and these are discussed by Roberts and Smith (1993) (also see Athreya, Doss and Sethuraman, 1996). Repeating the above procedure m times, making sure that the starting values, like our, are independently chosen, yields m k-tuples t1tm. These m k-tuples can be regarded as a sample of size m from the k-variate posterior distribution px. Furthermore, with t is a sample of size m from the marginal posterior distribution k. For m large this latter sample can be used to empirically construct the marginal posterior distribution ofj. Furthermore, if is any function of j, then 1jr is a consistent estimator ofE. This completes our discussion of the Gibbs sampling algorithm; some more detailscan be found in the excellent tutorial article by Casella and George (1992). However, there issome irony to the above scheme. It appears that in empirically estimating posterior distributions,and in appealing to the consistency property of estimates of functions of unknown parameters, aBayesian uses perfectly honorable frequentist procedures for easing the computational burdensthat his/her paradigm imposes. But to conclude this topic, we need discuss two issues, one amatter of concern and the other a virtue.

ACKNOWLEDGMENTS

The authors would like to thank the reviewers for their constructive comments that helped significantly in improving this paper.

REFERENCES

[1] B.J. Jansen, A. Spink, and T. Saracevic, "Real Life, Real Users, andReal Needs: A Study and Analysis of User Queries on the Web,"Information Processing and Management, vol. 36, no. 2, pp. 207-227,2000.

[2] 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.

[3]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.

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

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

[6] 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.

[7] J. Pearl, Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann, 1988.

[8] F. Monay and D. Gatica-Perez, "On Image Auto-Annotation with Latent Space Models," Proc. ACM Int'l Conf. Multimedia (MM), 2003.

Author Profile

K. Naresh Kumar received Master of Technology (Computer Science & Engineering)from Jawaharlal Nehru Technological University (JNTUH). My research interests include Information Security, Web Services, Cloud Computing, Data Mining and Mobile Computing. Presently working asAssociate Professor in CSEDepartment in Anurag Engineering College (AEC), Ananthagiri(V), Kodad(M), Nalgonda(Dt.), Andhra Pradesh, India.

K. Krishna ReddyPursuing Master of Technology (Computer Science & Engineering)from Jawaharlal Nehru Technological University (JNTUH). My research interests include Information Security, Web Services, Cloud Computing, Data Mining and Mobile Computing. Presently working as TEQIP Assistant in the department of CSEin Anurag Engineering College (AEC), Ananthagiri(V), Kodad(M), Nalgonda(Dt.), Andhra Pradesh, India.