Collective Evolutionary Indexing of Multimedia Objects

C.H.C. Leung, W.S. Chan, and J. Liu

Department of Computer Science,
Hong Kong Baptist University,
Kowloon Tong, Hong Kong
{clement,wschan,jiming}@comp.hkbu.edu.com

Abstract. The evolution of multimedia technology and the Internet boost the multimedia sharing and searching activities among social networks. The requirements of semantic multimedia retrieval goes far beyond those provided by the text-based search engines technology. Here, we present an collaborative approach that enables the semantic search of the multimedia objects by the collective discovery and meaningful indexing of their semantic concepts. Through the successive use of our model, semantic concepts can be discovered and incorporated by analyzing the users' search queries, relevance feedback and selection patterns. Eventually, through the growth and evolution of the index hierarchy, the semantic index can be dynamically constructed, validated, and naturally built-up towards the expectation of the social network.

Keywords: Collaborative Indexing, Multimedia Retrieval, Social Networks.

1 Introduction

With the rapid development of multimedia technology, the Internet, and cloud computing, numerous multimedia data objects are created in various forms and formats. However, multimedia information search is far more difficult than searching text-based documents since the content of text-based documents can be extracted automatically while the content of multimedia objects cannot be automatically determined [12,13].

Research in image retrieval has been divided into two main categories: "concept-based" image retrieval, and "content-based" image retrieval [1,3,4,5,10,15,16,21,24] The former focuses on higher-level human perception using words to retrieve images (e.g. title, keywords, captions), while the latter focuses on the visual features of the image (e.g. size, colour, texture). In an effective "concept-based" multimedia retrieval system, efficient and meaningful indexing is necessary [7,8]. Due to current technological limitations, it is impossible to extract the semantic content of multimedia data objects automatically [20,23,26]. Meanwhile, the discovery and insertion of new indexing terms are always costly and time-consuming. Therefore, novel indexing mechanisms are required to support their search and retrieval.

O. Gervasi et al. (Eds.): ICCSA 2009, Part I, LNCS 5592, pp. 937–948, 2009.

Community users are usually critical to a multimedia searching system. The relevance feedback which collected from users' interactions through a system can be used to adapt and evolve the system behaviour to fit their needs. Different communities of users may have different expected relevance to a search term, depending on time, geographic and cultural interests. The relevance feedback which provided by the user are essential. However, it is hard to convince users to rate the numerous multimedia data objects. Furthermore, the concern of user privacy issues is rising [18,25], and therefore we should explore other means for discovering multimedia data objects. The history of user behaviour and the time spent on the query sessions may reflect user feedback implicitly, while the analysis of query sessions is beyond the scope of this study. By analyzing both of the explicit and implicit relevance feedback from user, the search system can be tuned based on user preferences.

Our proposed approach can be seen as a form of machine learning, while it is dissimilar from a typical supervised learning approach, since there is no separate training phase. By the continuous interactive processes between users and the system, user knowledge and judgement of the search terms, relevance to the data objects will evolve over the time. Thus, we will also study the index convergence characteristics and behaviour.

2 Our Collaborative Indexing Approach and Hierarchical Evolution

Our approach concentrates on the indexing of semantic contents of multimedia objects and will exclude metadata from consideration, since indexing by metadata is relatively straightforward and less meaningful than semantic contents as perceived by humans [13]. This indexing approach enables user to search multimedia objects conceptually by the semantic visual features.

2.1 Index Basic Element and Hierarchial Structure

We consider a set of multimedia data objects $\{O_j\}$, such as images, video, or music, where their semantic characteristics and contents cannot be extracted automatically. Each O_j has an index set I_j, which consists of a number of elements $\{e_{j1}, e_{j2}, ..., e_{jM_j}\}$. Each basic index element e is a triple, which is composed of an index term ID T_{j-ID}, a corresponding index score S_{jk}, and an object ID O_{j-ID}. Each index set I_j indicates the relationships between index terms $\{T\}$ and the multimedia object O_j associated with the index scores S. The index scores reflect the significance of an index term to the object; the higher the index score, the more important is the index term to the object. In other words, the lower the index score means less important is the index term to the object. Fig. 1 shows the relationship of the multimedia data objects, index sets and index terms.

The index hierarchy refers to the collective index sets I_1, I_2, ..., I_m of all the objects O_1,O_2, ..., O_m in the database [13]. Fig. 2 shows that the index set $\{I\}$

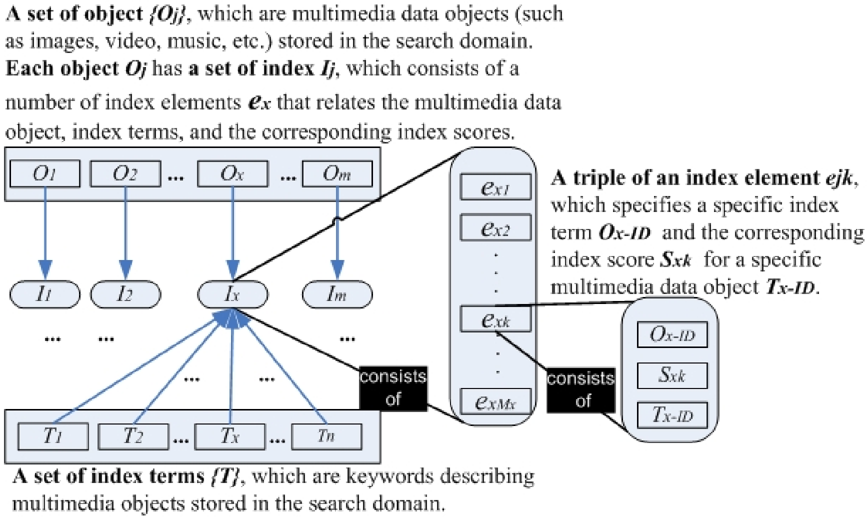


Fig. 1. Relationship Between Multimedia Objects, Index sets, and Index Terms

is partitioned into N levels L_1, L_2, ..., L_N by partitioning the score value S_{jk} with a set of parameters P_1, P_2, ..., P_N. For a given index term with score x, the index term will be placed in level L_i if $P_i \leq x < P_{i+1}$, where $i = 1, ..., N-1$. Otherwise, it would be placed in level N if $P_N \leq x$. In this index hierarchy, the higher the level, the more important it is.

2.2 Index Evolution and Index Score Update

In order to be discovered by users, each multimedia object may be minimally indexed initially. When an object O_j is minimally indexed, it means that O_j has only a single index term T where T consists of a single word only. Through successive usage of the system, the index set of an object would be grow, such that an object which is minimally indexed with term $T = T_1$ may become indexed with multiple terms $T = T_1, T_2, ..., T_n$. Consider an object O_J, which is minimally indexed with an index term T_1. Consider a user input query $Q(T_1, T_2)$, the system would return an answer vector V_{ans} which consists a set of objects that is indexed with T_1 or T_2. When the user select O_J in V_{ans}, T_2 would be added to O_J at the low level of the index hierarchy. If many queries that contain T_2 also select O_J continuously, the index score of T_2 of O_J would be increased and promoted to a high level of the index hierarchy. Consequently, T_2 of O_J would be properly indexed. Meanwhile, T_1 of O_J may drop to the lower level of the index hierarchy since it would be affected by the user relevance feedback.

The index scores are directly affected by the user relevance feedback, either positive or negative. By the continuous use of the system, our system can collect and analyze both explicit and implicit relevance feedback from users. When the system receives a positive feedback from user, the index score(s) that relate to

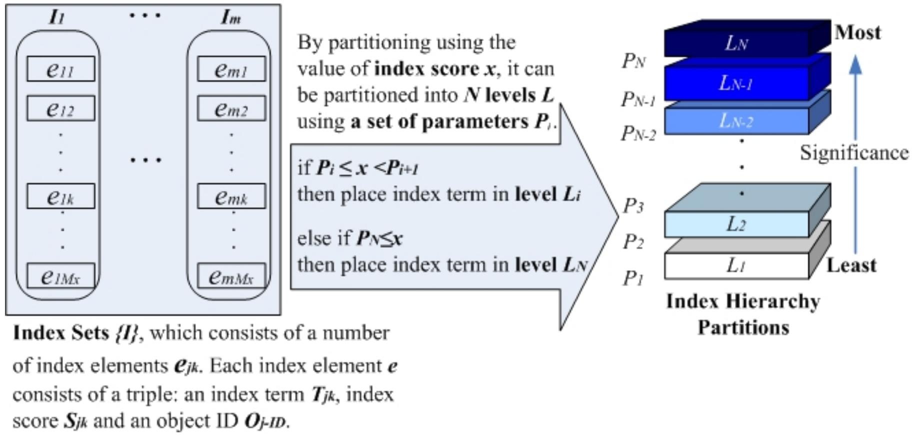


Fig. 2. Index Hierarchy

the search terms of the query would be increased. Similarly, the index score(s) would be decreased when the a negative feedback is received. Those positive and negative feedbacks can be the relevance feedback that are collected directly or indirectly from users.

Consider an example of a user input search query $Q(T_1, T_2)$ that consists of two search terms T_1 and T_2, and suppose there are k multimedia objects $O_1, O_2, ..., O_k$ returned in the answer vector V_{ans} disregarding the object rankings. When the user provides a positive feedback or selects the desired object O_x in the answer vector, the index scores of T_1 and T_2 of O_x would be increased by a predefined value Δ_+. In contrast, when the user provides a negative feedback on O_x, the index scores of T_1 and T_2 of O_x would be decreased by a predefined value Δ_-. Moreover, when a user do not select any object in the answer vector, the index scores of T_1 and T_2 for all objects in the answer vector $(O_1, O_2, ..., O_k)$ would also be decreased.

2.3 Object Ranking Strategies

In the information explosion era, search rankings has become significant [28]. Usually, the top-ranked objects should be more relevant to the search query. Every submitted query which consists of one or more search terms, is expected to return an answer vector $V_{ans} = [O_1, O_2, ...O_k]$, where k is the number of objects returned that is relevant to the search terms. Our object ranking approach relies on the index scores, since this score implies the relevance of an index term to an object. The higher the score, the more relevant is the index term to the object. There are two approaches for ranking the multimedia data objects returned in the query result lists.

Naïve Strategy. This strategy is to return the best k result objects ordered by their index scores, related to the search term(s) of the object, in descending order. By this strategy, the top ranked object O_1 in the answer vector V_{ans} is the most relevant to the search query, while O_k is the least. It is a typical strategy for building hot links, such as "top ten list of most clicked links" which can be found in many portals homepage. It implies that the probability that an object O_j being clicked would be directly proportional to the rank of object O_j. Since the "top ten" are more likely to be seen, therefore, those top ranked objects are more likely to be selected. Thus, the index scores of those objects have higher chance to be promoted, such that there would be an initial bias, in the top tens links, which can easily lead to a local maxima problem. Consequently, some significant objects would be hard to show up or ranked very low in the result lists. In the worst case, some objects may never been shown to users, such that, those objects may be nearly "hidden" or never receive positive feedback from users.

Randomized Strategy with Genetic Algorithms. This strategy provides variations in query results by using Genetic Algorithms (GA), and thus it is designed for overcoming the local maxima problem. It returns k result objects in the answer vector V_{ans} by random extractions from the index. The random extraction process involves a series of random selection among the object set O_j. It performs again and again until k distinct objects are selected to be shown. By the randomness of the GA, the best rated objects would have proportionally higher chances to appear in the answer vector; meanwhile, those "hidden" objects would have a non-zero probability of being promoted to appear in the answer vector. Therefore, those "hidden" objects would have a chance to be ranked in a higher ranking position and thus discovered eventually.

Although the problem of local maxima can be solved by our randomized strategy, it also introduces some noise which tend to lower the overall system performance. The system performance quality is degraded since the answer vector would consist of both good relevant objects and some irrelevant objects. However, those irrelevant objects are essential in the discovery of the "hidden" objects. Therefore, we adopt elitism, a technique from GA, to reduce the noise that is induced by the randomized strategy. By adjusting the elite E (number of best objects in an answer vector), where $E \in [0, k]$, it guarantees the quality of the answer returned. There are two extreme cases when considering the value of E; $E = 0$ means there is no any elitism such that all objects in the answer vector are generated by the randomized strategy. On the other way around, $E = k$ means there is no noise object in the answer vector, such that it is equivalent to the mentioned naïve strategy.

2.4 Measuring Indexing Performance

There are many different measures for evaluating the performance of information retrieval (IR) systems [17]. For a given collection of documents, every document can be classified as relevant or irrelevant in relation to a particular query. In particular, *precision* and *recall* have been used regularly to measure the performance of information retrieval and information extraction systems. Precision deals with substitution and insertion errors while recall deals with substitution and deletion errors [14].

In our proposed indexing model, we assume that each index score of an index term for a specific multimedia data object has an hidden ideal index score S_H. The higher the S_H value means the more relevant an index term to a multimedia data object. The values of hidden ideal index score S_H can take the form of various of distribution (e.g., normal distribution, uniform distribution), and is fixed at the initial stage and lies between 0 and 1. To determine the relevance of a multimedia data object to a query, we can consider whether the index score has reached the threshold of the hidden ideal index score.

3 Influence of User Relevance Feedback

Relevance feedback (RF) is a classical information retrieval (IR) technique where users relay their agreement with the system's evaluation of document relevance back to the system, which then uses this information to provide a revised list of search results [22]. It allows user to mark relevant (positive feedback) or irrelevant (negative feedback) to the object(s) of the result list by their relevance judgments. The user relevance feedback collected would be useful for refining the index scores, such that it helps tuning the index hierarchy to fit user preferences.

Our model collects both explicit and implicit relevance feedback from the user community. The explicit feedback refers to the relevance, indicating the relevance of the object retrieved for a query, and is collected directly from user judgements. Our model enables users to indicate relevance explicitly using a binary relevance system. Binary relevance feedback indicates that a multimedia data object is either relevant or irrelevant for a specific query. Once a user submits a query, our system will return a list of query results to the user. In order to maintain the spirit of Web 2.0 [2,6,27] with collaborative users involvement, our system allows users to provide their relevance feedback for the multimedia objects of the query results. Their feedback can be either positive or negative.

Although the idea of exploiting user's feedback to rate relevance seems promising, it is not easy to convince a community of users to spend their time to explicitly rate objects. Therefore, our model also collects implicit relevance feedback from them. The implicit feedback is inferred from user behaviour and their history, such as noting which object(s) that users do and do not select for viewing, and the duration of time spent in viewing an object. All such information can be collected automatically and would reflect user satisfaction and expectation of the query result. When users click on an object in the answer vector, we can infer that the selected object may be relevant to the user query. Our system will treat it as a kind of positive feedback from the user implicitly. On the contrary, when users do not select any object in the answer vector, we can infer that they may think that the objects in the answer vector are irrelevant to their input query or they are not interested in those objects. Our system will treat it as a kind of negative feedback from the user implicitly.

4 Modelling Index Convergence Behaviour

Since the measurement of the relevance of an index term to an object is based on the related index score, the index scores of the system are expected to evolve to an ideal situation. We assume that each index score for an index term of a specific object would have a hidden ideal score value S_H. When the actual index score reaches S_H, this index can be considered as having convergent. By the continuous usage of the system, the indexes would be eventually convergent.

4.1 Index Convergence Behaviour

In theory, indexes would evolve to an ideal state by the index continuous convergence processes. Consider there are J objects in the search space, each of the objects are indexed with m initial index terms, and M be the number of the maximal index terms.

Let N_t be the state of the system which signifies the number of terms remaining to be indexed. N_t is a random variable that changes over time. Let the process starts at $t = 0$. Thus initially, we have

$$N_0 = J(M - m). \tag{1}$$

As time goes on, N_t will gradually decrease. N_t will decrement by 1 whenever a potential indexable term is being indexed. We assume that the random indexing pattern for a given term follows a Poisson process[11] with indexing rate μ, where in a small time interval Δh, a potential indexable term has a probability of $\approx \mu \Delta t$ of being actually indexed. The rate is dependent on the usage and indexing frequency of objects in the collection. Thus, over time, each potential indexable term is gradually being deleted as they are become indexed.

From the property of the Poisson distribution[9,19], the probability that a potential indexable term remaining unindexed at time t is $e^{-\mu t}$. Therefore, we obtain the following binomial distribution[19] for N_t

$$Prob[N_t = k] = \binom{N_0}{k} (1 - e^{-\mu t})^{N_0 - k} e^{-\mu t k}, \tag{2}$$

which gives

$$E(N_t) = N_0 e^{-\mu t}, \tag{3}$$

$$Var(N_t) = N_0 (1 - e^{-\mu t}) e^{-\mu t}, \tag{4}$$

Adopting a time unit of days, $\frac{1}{\mu}$ can be taken as the average time elapsed to install the index term. For example, if $\mu = 0.1$, this means that the average time to install the index term is 10 days. Fig. 3 plots the number of remaining index terms over time for $N_0 = 100,000$, and $\mu = \frac{1}{30}, \frac{1}{60}$. We see that the number of indexable terms drops quickly at first, then do so slowly as time goes on. As $t \to \infty$, we see from equation (3) that the collection tends to be fully indexed with $E(N_t) \to 0$, irrespective of the initial number of potential indexable terms.

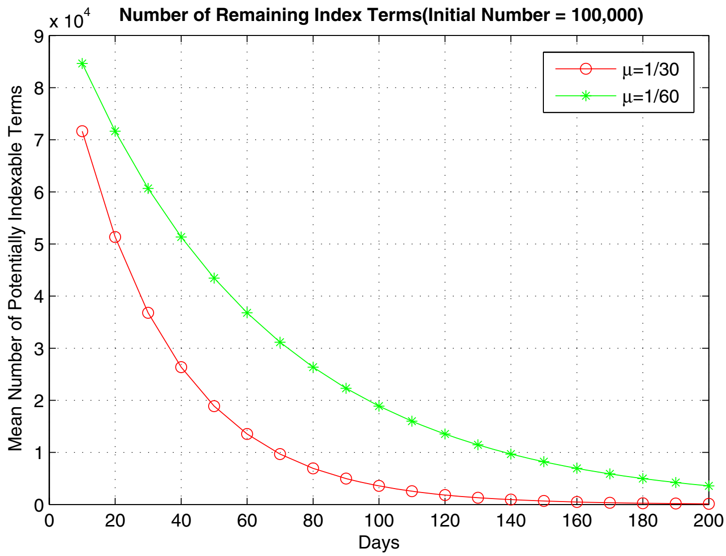


Fig. 3. Number of Remaining Index Terms

Also, from equation (4), $Var(N_t) \to 0$ as $t \to \infty$, which indicates that the effect of stochastic fluctuation would be small; this implies that, over a long period of time, the process may be viewed as a deterministic one.

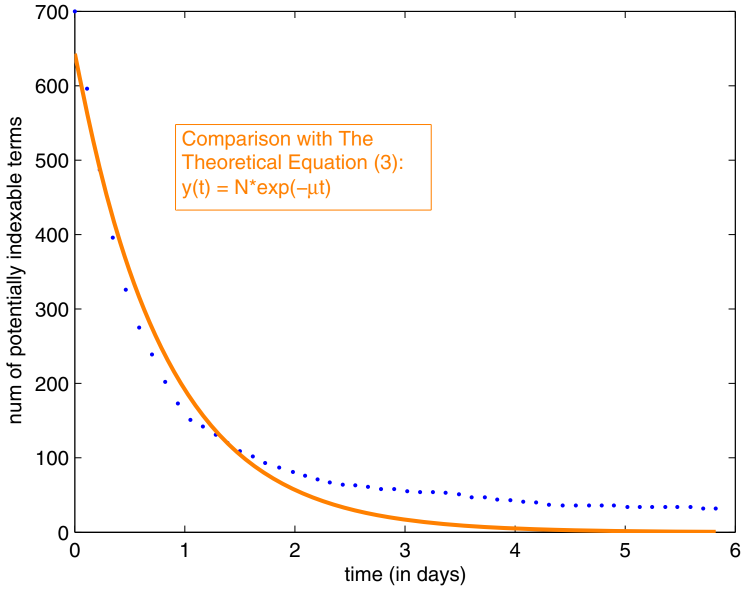
5 Experiments on Index Convergence

We performed experiments to examine the convergence behaviour of the index hierarchy. In order to evaluate the convergence behaviour of the model, we simulated the search processes including query submission from user, searching the relevant multimedia data objects, ranking the results, and collecting user relevance feedback. The goal of the series of experiments is to investigate the convergence behaviour of the index hierarchy.

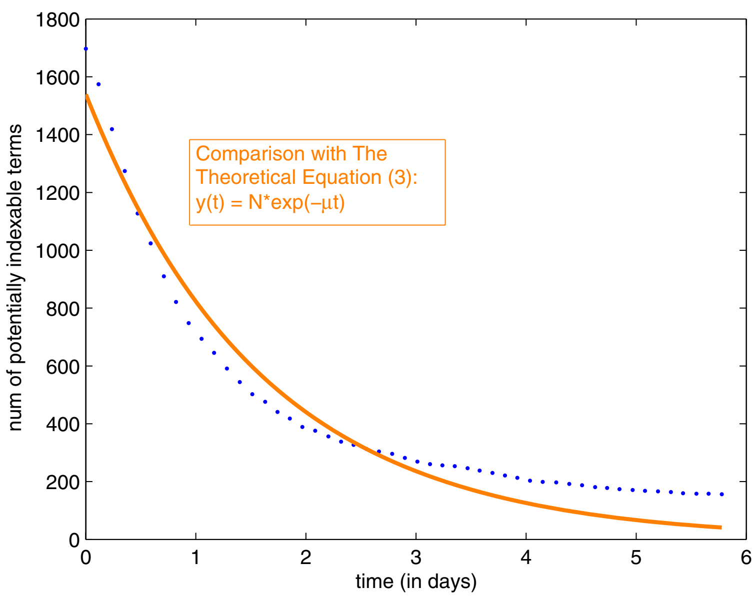
5.1 Decay Behaviour of Number of Remaining Index Terms and Index Convergence

In our simulation model, we assume the arrival of user queries follow poisson distribution. We tested the runs in the same environment with the same initial settings, such as number of queries (i.e., 50,000), number of initial index terms per object (i.e., 3), score increment / decrement after reeving user relevance feedback, indexing threshold (i.e., 50%) etc. By varying some variables, such as maximum number of index terms per object t_{max} and number of multimedia data objects O in the search domain, we test its effect on the convergence behaviour.

In the first series of runs, we performed the queries with 100 data objects and $d = 80$. We collected the number of potentially indexable terms N_t for every



(a) t_{max}=10 and N_0=700



(b) t_{max}=20 and N_0=1700

Fig. 4. Test Results

1,000 queries. Initially, each data object O is indexed with 3 index terms. In each query, the number of search terms is fixed as 2 and the number of objects returned in search results are fixed as 10. We tested it with different number of maximum number of indexable terms t_{max}. Fig. 4 (a) (i.e., $t_{max} = 10$ and $N_0 = 700$) and (b) (i.e., $t_{max} = 30$ and $N_0 = 1700$) show the points that are the number of potentially indexable terms N_t collected from the runs. Then, we fit the data points exponentially with the equation (3). Our results show clearly that the number of potentially indexable terms 'decay' exponentially through the time.

6 Conclusions and Future Works

We presented a collaborative indexing approach for enabling multimedia retrieval within a large collection of multimedia data objects. Our indexing approach helps to discover multimedia resources systematically by keeping track of the user query behaviour. By analyzing the search information, the user relevance feedback helps the index hierarchy to evolve towards to users' desired preferences. Thus, user satisfaction would be maximized. Our experimental results show that the index converges successfully after successive use. In the future, we will further focus on examining the behavioral pattern and stochastic modelling of index convergence characteristics.

References

1. Azzam, I.A., Leung, C.H.C., Horwood, J.F.: Implicit concept-based image indexing and retrieval. In: Proceedings of the 10th International Multimedia Modeling Conference (MMM 2004), Brisbane, Australia, January 5-7, 2004, p. 354. IEEE Computer Society, Los Alamitos (2004)
2. Boll, S.: Multitube - where web 2.0 and multimedia could meet. IEEE Multimedia 14(1), 9–13 (2007)
3. Finkelstein, L., Gabrilovich, E., Matias, Y., Rivlin, E., Solan, Z., Wolfman, G., Ruppin, E.: Placing search in context: the concept revisited. ACM Transactions on Information Systems 20(1), 116–131 (2002)
4. Funkhouser, T., Min, P., Kazhdan, M., Chen, J., Halderman, A., Dobkin, D., Jacobs, D.: A search engine for 3D models. ACM Transactions on Graphics 22(1), 83–105 (2003)
5. Gevers, T., Smeulders, A.: Image Search Engines: An Overview. In: Emerging Topics in Computer Vision, Prentice Hall, Englewood Cliffs (2004)
6. Giannakidou, E., Kompatsiaris, I., Vakali, A.: Semsoc: Semantic, social and content-based clustering in multimedia collaborative tagging systems. In: ICSC 2008: Proceedings of the 2008 IEEE International Conference on Semantic Computing, Washington, DC, USA, pp. 128–135. IEEE Computer Society, Los Alamitos (2008)
7. Go'mez, J., Vicedo, J.L.: Next-generation multimedia database retrieval. IEEE Multimedia 14(3), 106–107 (2007)

8. Goth, G.: Multimedia search: Ready or not? IEEE Distributed Systems Online 5(7) (2004)
9. Haight, F.A.: Handbook of the Poisson Distribution. Wiley, Chichester (1967)
10. Hawarth, R., Buxton, H.: Conceptual-description from monitoring and watching image sequences. Image and Vision Computing 18, 105–135 (2000)
11. Kingman, J.F.C.: Poisson processes. Oxford University Press, Oxford (1993)
12. Leung, C., Liu, J., Chan, W.S., Milani, A.: An architectural paradigm for collaborative semantic indexing of multimedia data objects. In: Sebillo, M., Vitiello, G., Schaefer, G. (eds.) VISUAL 2008. LNCS, vol. 5188, pp. 216–226. Springer, Heidelberg (2008)
13. Leung, C.H.C., Liu, J.: Multimedia data mining and searching through dynamic index evolution. In: VISUAL 2007: Proceedings of the 9th International Conference on Visual Information Systems, Shanghai, China, pp. 298–309 (2007)
14. Makhoul, J., Kubala, F., et al.: Performance measures for information extraction. In: Proceedings of DARPA Broadcast News Workshop, Herndon, VA (1999)
15. Mu'ller, H., Mu'ller, W., Squire, D.M., Marchand-Maillet, S., Pun, T.: Performance evaluation in content-based image retrieval: Overview and proposals. Pattern Recognition Letters 22(5) (2001)
16. Over, P., Leung, C.H.C., Ip, H.H.-S., Grubinger, M.: Multimedia retrieval benchmarks. IEEE Multimedia 11(2), 80–84 (2004)
17. van Rijsbergen, C.J.K.: Getting into information retrieval. In: Agosti, M., Crestani, F., Pasi, G. (eds.) ESSIR 2000. LNCS, vol. 1980, pp. 1–20. Springer, Heidelberg (2001)
18. Saint-Jean, F., Johnson, A., Boneh, D., Feigenbaum, J.: Private web search. In: WPES 2007: Proceedings of the 2007 ACM workshop on Privacy in electronic society, pp. 84–90. ACM, New York (2007)
19. Sleeper, A.: Six Sigma Distribution Modeling. McGraw-Hill, New York (2006)
20. Snoek, C.G.M., Worring, M., van Gemert, J.C., Geusebroek, J.-M., Smeulders, A.W.M.: The challenge problem for automated detection of 101 semantic concepts in multimedia. In: MULTIMEDIA 2006: Proceedings of the 14th annual ACM International Conference on Multimedia, pp. 421–430. ACM, New York (2006)
21. Tam, A.M., Leung, C.H.C.: Structured natural-language descriptions for semantic content retrieval of visual materials. Journal of the American Society for Information Science and Technology 52(11), 930–937 (2001)
22. Vinay, V., Wood, K., Milic-Frayling, N., Cox, I.J.: Comparing relevance feedback algorithms for web search. In: WWW 2005: Special interest tracks and posters of the 14th International Conference on World Wide Web, pp. 1052–1053. ACM, New York (2005)
23. Wei, X.Y., Ngo, C.W.: Fusing semantics, observability, reliability and diversity of concept detectors for video search. In: MM 2008: Proceedings of the 16th ACM International Conference on Multimedia, pp. 81–90. ACM, New York (2008)
24. Wong, R.C.F., Leung, C.H.C.: Automatic semantic annotation of real world web images. IEEE Transactions on Pattern Analysis and Machine Intelligence 30, 1933–1944 (2008)
25. Xu, Y., Wang, K., Zhang, B., Chen, Z.: Privacy-enhancing personalized web search. In: WWW 2007: Proceedings of the 16th International Conference on World Wide Web, pp. 591–600. ACM, New York (2007)

26. Yang, B., Hurson, A.R.: Ad hoc image retrieval using hierarchical semantic-based index. In: AINA 2005: Proceedings of the 19th International Conference on Advanced Information Networking and Applications, Washington, DC, USA, pp. 629–634. IEEE Computer Society, Los Alamitos (2005)
27. Yesilada, Y., Harper, S.: Web 2.0 and the semantic web: hindrance or opportunity?: W4a - International Cross-disciplinary Conference on Web Accessibility 2007. SIGACCESS Access. Comput. (90), 19–31 (2008)
28. Yoshida, Y., Ueda, T., Tashiro, T., Hirate, Y., Yamana, H.: What's going on in search engine rankings? In: AINAW 2008: Proceedings of the 22nd International Conference on Advanced Information Networking and Applications - Workshops (aina workshops 2008), Washington, DC, USA, pp. 1199–1204. IEEE Computer Society, Los Alamitos (2008)