Visual Search Optimization using Concept Related Re-Ranking

G. Lakshmi Narayanan
PG Student
Department of Computer Science &Engineering (PG)
National Engineering College

V. Kalaivani
Associate Professor
Department of Computer Science &Engineering (PG)
National Engineering College

ABSTRACT
Visual search re-ranking defined as re-ordering visual documents like image, videos etc. based on the initial search. Ranking the multimedia content like images, videos are a challenging research topic in the noisy visual environment. Now days, leading search engines are fully depends on the description, title, surrounding information of an image which produce irrelevant image which are not equal to visual content. In this paper, a new approach proposed to improve the visual search precision level. First, the initial ranking occurred based on the textual information like tag, description relevancy which didn't produce relevant images. Second, by using visual query examples in the search engine to filter the images based on feature. The visual equivalence between the images calculated to increase the relevance results. Mainly the Equivalence Re-ranking approach focused on the relationship between the concepts of documents considered to reorder the initial search result with higher resolution images for optimizing the list of images. And by avoiding and removing irrelevant image along with the low resolution images by re-ranking approach, will increase the performance of search engine.

Keywords
pair-wise learning, search re-ranking, visual search, example re-ranking, EP re-ranking, correlation

1. INTRODUCTION
Visual search in the noisy internet environment is a challenging research topic. Current visual search engines depend entirely on the text associated with the visual documents mostly judged by the text-based approaches, as textual information cannot able to describe the visual content. For example, when users search images with a specific image based on color or object, the images cannot be easily measured by textual description or tags. To increase the visual relevancy, re-ranking the visual documents in large visual datasets is getting more attention in recent years. It can be defined as re-ordering visual documents based on the external or secondary knowledge to improve the search precision. There are many methods used to rank the visual documents based on the external knowledge. Detecting relevant samples from the initial search results without any outward knowledge by using the concept re-ranking by person [4], [5]; Next, query examples are provided by users so that the relevant patterns can be discovered from these examples by using the secondary knowledge [16]; and another one is mining the relevant feedbacks from the gathered information available on the results. i.e., self-re-ranking, slightly relied on the outer knowledge, cannot deal with the problem "unclearness" which is derive from the text queries. Taking the query "apple" as an example, the search system cannot determine what the user is really expecting as a output or searching for, whether it is "i-pod" or "a fruit." As illustrated in Fig.1, results with different meanings but all related to "apple" can be found in the top-ranked results of keyword "apple." To address this problem, the second and the third methods leverage some auxiliary knowledge to better understand the query. Specifically, the second dimension, i.e., re-ranking based on the samples, leverages a few query examples to train the re-ranking models. However, it still cannot avoid the ambiguity problem as current visual search engines mainly support the text query. To summarize, the re-ranking from external information still cannot deal with the unclearness problem, which increase the irrelevant data at the same time leads to performance degradation. If the images are not classified correctly based on the visual features, then there is an increase in the low quality information from the dataset without any clear classification. And the above re-ranking techniques which are not dealing with the unclearness problem, because current visual search engines only support the text query along with the descriptions which are tagged with images. To address these issues, in this paper, both the query examples and textual information are combined together for re-ranking. Specifically, first feed the text query to filter the images based on the tag and description of an image.

Fig. 1 Top 20 search results for the keyword apple in the commercial search engine.

To avoid the unclearness problem, the query examples used here to filter the visual documents to get more clean results. And to re-rank the filtered results by finding the similarity between description and visual information through the concept related to the query for the purpose of ranking the document. While the visual documents re-ranked by combining both the textual concept relations like description along with the visual concept relation to increase the precision.

2. RELATED WORK

Visual search with a set of high-level concept detectors has attracted increasing attention in recent years [12], [13], [14], [15], [18]. Intuitively, if queries can be automatically mapped to related concepts, search performance will increase. For example, the “apple” concept can benefit for fruit-related queries and the “leaf” concept can also be high weighted for plant-related queries. The problem of discovering related concepts, also called “query-concept mapping,” has been focused on by many researchers recently. Kennedy et al. introduced a method to mine the top-ranked and bottom-ranked search results to discover related concepts by measuring the mutual information [14]. The basic idea is that if a concept has high mutual information with the top-ranked results and low mutual information with the bottom-ranked results, it will be considered as a related concept. Avoiding the ambiguity problem, Li et al. and Liu et al. leverage a few query examples to find related concepts; specifically, Li et al. proposed an approach to use the tf- idf like scheme [17] and Liu et al. explore the mutual information measurement [17]. Methods are motivated by the information-theoretic point of view, that is, the more query examples bear more information of a concept, the more the concept will be related to the corresponding query.

Visual search re-ranking[1], [2], [4] by producing a few query examples (e.g., images or video shots) provided by users. For instance, Yan et al. [6] proposed a method to view the query examples as pseudo positives and the bottom ranked initial results as pseudo negatives. A re-ranking model is then built based on these samples by the SVM. In a similar way, Natsev et al. [16] proposed a technique to use the same set of positive examples and randomly sample pseudo negatives a multiple number of times to build SVM models, and then fuse the ranked lists generated from these models. Liu et al. Hsu et al. proposed a new technique to formulate the re-ranking process as a random walk over a context graph, where video stories are nodes and the edges between them are weighted by multimodal similarities. The video stories with similar visual appearance and text descriptions are linked compactly, and thus, the documents ranked lower can be picked up if they are strongly linked with considerable top-ranked ones [12]. Unsupervised re-ranking [16] aims to detect repeated patterns (viewed as the relevant cues), such as video stories, image patches, and high-level concepts, in the initial search results without any outside knowledge .

Fergus et al. [8] proposed a approach to perform visual clustering on initial returned images by probabilistic latent semantic analysis (pLSA), and then learn the visual object category based on the visual clusters; finally, the images are re-ranked according to the distance to the learned categories [7].Introduced a new process to use the query examples to detect the relevant and irrelevant concepts for a given query, and then identify an optimal set of document pairs via an information theory. The final re-ranking list is directly recovered from this optimal pair set. Encouraging results have been reported in this dimension. However, the reliance on the user-provided query examples has limitations, as most users are reluctant to provide enough query examples while searching.

A new perspective in image retrieval method involving a combination of factor analysis with relevance feedback method is introduced in [7] whereas a statistical correlation model for the retrieval of relevant images is presented in [6]. In this model [6], an estimate of the correlation between two images based on the number of search sessions in which images have been marked relevant is calculated. Since in the process of relevance feedback, main emphasis is to improve the retrieval accuracy, several passes are made through the database during the retrieval process and a correlation is dynamically calculated by interacting with the user. However, establishing a correlation dynamically is not only a time consuming process but it also makes it difficult to incorporate positive and negative examples in query and/or the similarity refinement process.

To improve the initial performance [3], [5] by only mining the initial ranked list without any external knowledge. For example, Kennedy and Chang [5] proposed new method to produce the top-ranked results (positives) and bottom-ranked results (negatives) to discover the related concepts. Then the confidence scores of the related concept detectors are used as high-level features in SVM to build classifiers. Hsu et al. [4] formulate the re-ranking process as a random walk over a context graph, where video stories are nodes and the edges between them are weighted by multimodal similarities.

In this paper, first feed the keyword query to a public visual search engine and collect the visual documents along with the associated text; to avoid the ambiguity problem, the query examples used here to filter the web results and get cleaner “visual examples.” By combining the filtered visual results along with the associated text and also classifies the images on the basis of maximum correlation so that the images with more similarities and, hence, exhibiting maximum correlation with each other are grouped in the same class, are retrieved accordingly.

3. EQUIVALENCE RE-RANKING

Relation between the query image and the resultant images from the database is obtained with the help of feature equivalence. And, here the equivalence between the images has been calculated to reorder the initially ranked images. Correlation is a measure of the relation between two or more variables. Cross correlation and normalized correlation are standard statistical methods that have been successfully used in various image related applications. Given the two same size vectors or matrices, two-dimensional cross correlation between them can be calculated using the 2-D Correlation. Assuming two matrices *A* and *B* of the same dimension (*m*, *n*), the 2-D discrete normalized correlation can be computed as:

$$r = \frac{\sum_m \sum_n \left(A_{mn} - \overline{A}\right)\left(B_{mn} - \overline{B}\right)}{\sqrt{\left(\sum_m \sum_n \left(A_{mn} - \overline{A}\right)^2\right)\left(\sum_m \sum_n \left(B_{mn} - \overline{B}\right)^2\right)}} \qquad (1)$$

$\overline{A}$ and $\overline{B}$ is the mean of the values of matrix:

$$\overline{A} = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} A(i, j) \qquad (2)$$

and

$$\overline{B} = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} B(i, j) \qquad (3)$$

The calculated correlation coefficients can range from -1 to +1 such that a value of -1 represents no correlation between the matching entities whereas a value of +1 represents a perfect positive correlation or a perfect match. More than two parameters used here from the image to get the correlated value to get image between the values +1 and -1. And the decision for the retrieval process has done through the correlated value.

4. CONCEPT RELATED RE-RANKING

The proposed approach, named EQ-re-ranking (Equivalence Re-ranking), which re-rank the retrieved results from search engine through the similarity detection between the pair of images. Method is illustrated in Fig. 2. Re-ranking proceeds in four phases. First, keyword based image retrieval has done to extract the images from database based on the user specified concept and also here the description along with images also extracted. In the second phase, Query examples produced by user to the visual search engine based on the keyword concept. Simultaneously feature mean calculated based on the RGB color model and also with the use of Primitive length pixel calculation method for texture and Medium level pixel of an gray scale image.

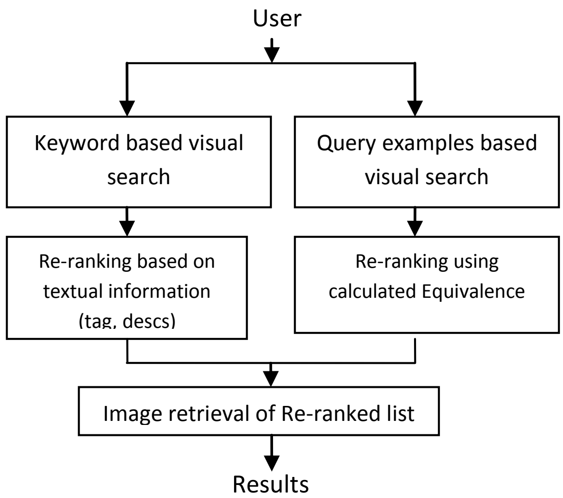

Fig. 2 Equivalence based visual re-ranking system

The third phase is to find the dependence or equivalence between the images in database and also with user given query example. Simultaneously, the description of image compared with the keyword for re-ranking based on the textual information of image. The fourth phase re-ranking is done through the correlation value along with the keyword which associated with image. And also the redundancy of information is avoided to increase performance.

4.1 Keyword Based Visual Search

Image retrieval system based on the user given keyword is the traditional approach to retrieve a relevant text document and also now a day it used to retrieve images based on the keyword. Keywords are used as descriptors to index an image however the content of an image is much richer than what any set of keywords can express [10]. Text-based image retrieval techniques employ text to describe the content of the image which often causes ambiguity and shortage in performing the image database search and query processing. This problem is due to the difficulty in specifying exact terms and phrases in describing the content of images as the content of an image is much richer than what any set of keywords can express. Since the textual annotations are based on language, variations in annotation will pose challenges to image retrieval.

Textual Equivalence

Textual equivalence between image annotations used here to find the similarity between the keyword. Three kind of text meta-data associated with each image are considered: title, tags, and description. And these are combined into a single document per image for further processing. As with the visual similarity, pair-wise text similarities between the documents (images) calculated by similarity calculation. The textual similarity between an image pair of *Ii* and *Ij* can be calculated as:

$$sim(I_i, I_j) = \frac{I_i(t).I_j(t)}{\|I_i(t)\|\|I_j(t)\|} \qquad (4)$$

where $I_i(t)$ is a textual feature vector that combines the meta-data associate with image I_i as aforementioned. $I_j(t)$ is a textual feature vector that combines the metadata associate with image I_j.

4.2 Query Example Based Visual Search

Visual search which is based on query example is the retrieval approach based on the examples, the problem of searching for digital images in large databases. "Example-based" means that the search will analyze the actual contents of an example image rather than the keywords, tags, and/or descriptions associated with the image. The term 'content' in this context might refer to RGB color model and also the relatedness between the pixels in the example and dataset and also the other information that can be derived from the image itself. Also having humans manually enter keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes the image. Query by example is a query technique that involves providing the CBIR system with an example image that it will then base its search upon. The underlying search algorithms may vary depending on the application, but result images should all share common elements with the provided example.

Options for providing example images to the system include:

[1] A pre-existing image may be supplied by the user or chosen from a random set.

[2] The user draws a rough approximation of the image they are looking for, for example with blobs of color or general shapes.

In the proposed work the pre-existing images are used here as an example to select by the user for the query example based search.

4.3 Feature Extraction

The basic idea of image retrieval by image example is to extract characteristic features from target images which are then matched with those of the query image. These features are typically derived from texture, color of a pixel along with the properties of query and target images. Feature extraction from the collection of images stored in a database and an every image have a set of pixel between x and y axis. Average of red, green, blue has been calculated for each and every pixel in an image and stored in a database. Feature extraction is one of the need process which has to be done for the process of retrieving image from the database what user needs. Extraction of features lead to evaluate the image through a comparison and find out whether is a relevant one or not. And, it is used to know the color level two images feeded by the user and also for each and every image stored in the database.

1) *Color feature:* The color spatial distributions of salient images are more concentrated than those of cluttered ones. In a image, colors of the object are less likely to be found in the background, while colors in a cluttered image are often scattered. Therefore, the feature of global color spatial distribution can be used to distinguish salient image class and gather image. Color moments [8] are a useful and convenient

feature in describing the color distribution of an image. Here calculate three moments for each of the three channels in color space and aggregate the features into one feature vector.

2) Texture feature: The primitive length pixel equivalence used here to find the texture information. And it is meaningful and enough to differentiate a image through the grey level of an images. In the implementation, the local binary pattern representation of an image is used [9] to find the equivalence.

3) Medium Level Pixel Equivalence*:* Retrieval of image corresponding to the query example consists of measuring the similarity between images represented by their grey scale features. The grey level of a image found out to classify or gather as relevant which are equivalence to query image. The value 0 to 255 used to evaluate the grey level in image. Therefore, 0 is considered as black and 255 as a white. The middle grey level of the query image is used as a filter to eliminate false examples form the multiple image datasets.

4.4 Re-ranking

Keyword based image result has been retrieved through an image retrieval process the basic re-ranking done based on the textual information which enclosed with the image. After the extraction of features from a query images lead to an extraction of relevant image from a database. To produce an optimized result the EQ re-ranking process has done based on the visual information of an image. And here re-ranking has done based on the feature (color, Medium level Pixel equivalence along with texture) equivalence of an images. Re-ordering of an initial list of images has been re-ranked based on the average of image feature dependence between the images in the database and query examples. Along with the features the description tags of images analyzed for the relevancy. Each and every image description compared to weight the image relevancy. And here re-rank based on the image description combined with the feature re-ranking to increase the relevancy using the equivalence pair wise re-ranking

The new approach equivalence re-ranking using the color, texture along with medium pixel level equivalence used to re-order the results based by joining up the visual information which calculated from pixel along with the textual information.

5. RESULTS

Visual search, where the user expected images list out based on the query concept. Here the performance of search engine relevancy and along with that the clarity of an image is needed one. Search Engine database hold high and low clarity images. And the clarity of an image depends on the level of resolution in an image. The low resolution images in database increase the irrelevant and unexpected results what user doesn't need. The below results show the performance and accuracy of the low and high clarity image database.

Dataset1: Sample set of 100 low resolution images for testing. And sample set contains buses, tiger, flowers, etc. for evaluating the relevancy and performance of search in the unclear positive examples.

Dataset2: Top 200 results of each concept like tiger, tulips, world, and apple are extracted from commercial search engine to evaluate the performance of search in the combination of high and low quality images in a dataset.

Fig 5.1 shows the comparison of different feature combination like color with PLP, MLP equivalence with PLP equivalence for the performance evaluation of search based on feature and also the level of relevancy in produced result. And here the retrieved results which are all relevant to the user query is too low. When low clarity images in dataset also get involved in the resultant list.

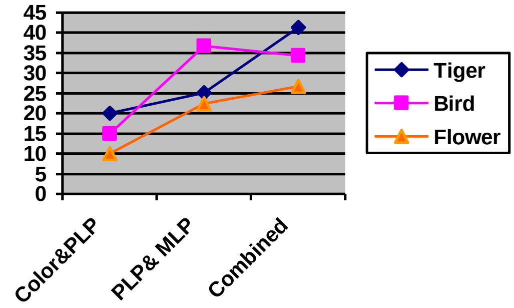

Fig 5.1 Comparison of different feature combination.

Fig 5.2 which shows the comparison between the textual and feature based retrieval where the images which holding the corresponding text get more precision in textual based ranking. Likewise, the feature based re-ranking increase the relevancy based on the query examples.

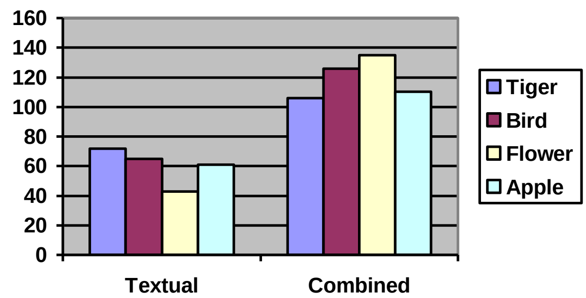

Fig 5.2 Comparison of textual and feature re-ranking

Fig 5.3 shows the Keyword Based Search where the result of initial rank produced. And the keyword flower produce relevant and irrelevant images based on query concept.

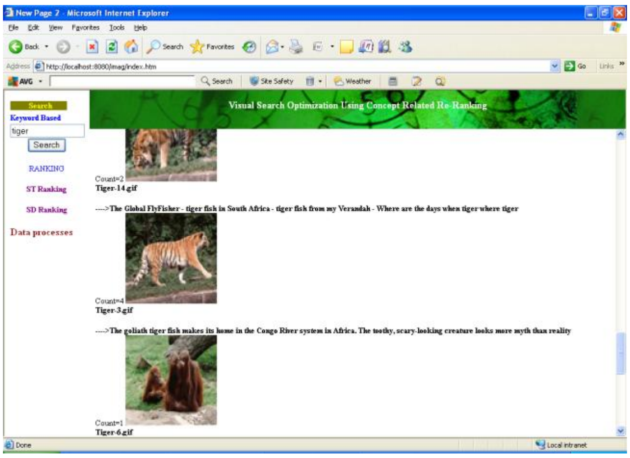

Fig. 5.3 Search by Keyword

Fig 5.4 shows the query example based search which done by the feature color to re-rank the initially searched images from a database. Which perform re-rank based on the results above color threshold value of query example.

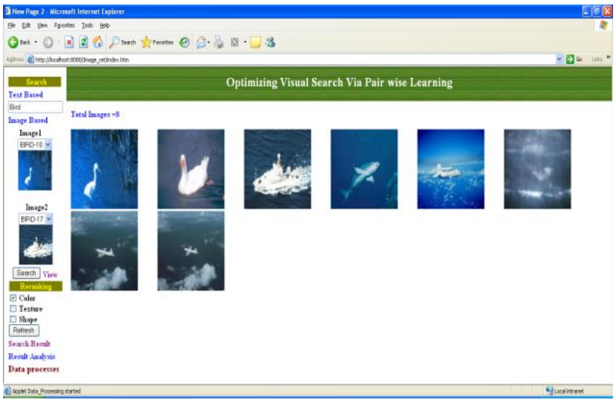

Fig. 5.4 Result based on color

Sample dataset of visual search engine holds 40 bird images and 160 other images along with low clarity images. Visual search based on color, texture and MLP extract the images which are above the given threshold value generates using the query example. So, among 8 results only 3 bird images are extracted with the other images which or not come under the bird category. And also accuracy level based on color, MLP and texture feature based retrieval decrease the irrelevancy and leads to an search result optimization by re-ranking the result by combining both textual and related concepts to reduce the irrelevancy at the time of re-ranking.

6. CONCLUSION

This paper proposed EQ re-ranking approach for increasing the accuracy of search results. Previous re-ranking approaches for visual search largely depend on the textual features (tags, description etc.) of an image. EQ re-ranking is suitable for visual search where the local features of an image and textual features also considered here to produce the re-ranked results in the comparative manner. Based on that when the re-ranking consider the local feature which represent images accurately and textual feature which increase relevant. From that re-ranked image list the accuracy and performance also increased by considering the visual relation.

7. REFERENCES

[1].Duong. B, Goertzel, J. Venuto, R. Richardson, S. Bohner, and E. Fox, “Support vector machines to weight voters in a voting system of entity extractors,” in *Proc.* IEEE Int. Joint Conf. Neural Networks, 2006.

[2].Fergus. R, L. Fei-Fei, P. Perona, and A. Zisserman, “Learning object categories from Google’s image search,” in Proc. IEEE Int. Conf. Computer Vision, 2005, pp. 1816–1823.

[3].Xu. J and W. B. Croft. Query expansion using local and global document analysis. In Proceedings of the 19th Annual International ACM SIGIR Conference, 1996.

[4].Lew. M. S, N. Sebe, C. Djeraba, and R. Jain, “Content-based multimedia information retrieval: State of the art and challenges,” ACM Transactions on Multimedia Comput., Commun., Appl., vol. 2, no. 1, pp. 1–19, Feb. 2006.

[5].Hsu. W, L.Kennedy, and S.-F. Chang, “Video search reranking through random walk over document-level context graph,” in Proc. ACM Int. Conf. Multimedia, 2007, pp. 971–980.

[6].Liu. W, and H. J. Zhang. A Statistical Correlation Analysis in Image Retrieval. Pattern Recognition, 35:2687–2693, 2002.

[7].Machado. A, C. N. J. Marinho, M. Fernando, and M. Campos. An image retrieval method based on factor analysis. In Proc. of XVI Brazilian Symp. on Computer Graphics and Image Processing (SIBGRAPI’03), pages 191–196, Oct. 2003.

[8].Stricker. M and M. Orengo, “Similarity of color images,” in Proc. SPIE Storage and Retrieval for Image and Video Databases, 1995.

[9].Ojala. T, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray scale and rotation invariant texture analysis with local binary patterns,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 971–987, Jul. 2002.

[10].Hui Hui Wang, Dzulkifli Mohamad, N.A Ismail, ”Image Retrieval: Techniques, Challenge, and Trend”, World Academy of Science, Engineering and Technology.

[11].Chuanchuan WANG†, Jun MA, “Rank Refinement for Social Images by a Random Walk Model”, Journal of Computational Information Systems, May, 2011.

[12].Neo. Y, J. Zhao, M.-Y. Kan, and T.-S. Chua, “Video retrieval using high level features: Exploiting query matching and confidence-based weighting,” in Proc. ACM Int. Conf. Image and Video Retrieval, 2006.

[13].Kennedy. L and S.-F. Chang, “A reranking approach for context-based concept fusion in video indexing and retrieval,” in Proc. ACMInt. Conf. Image and Video Retrieval, 2007.

[14].Li. X, D. Wang, J. Li, and B. Zhang, “Video search in concept subspace: a text-like paradigm,” in Proc. ACM Int. Conf. Image and Video Retrieval, 2007.

[15].Snoek. C and M. Worring, “Are concept detector lexicons effective for video search?,” in Proc. IEEE Int. Conference. Multimedia & Expo, 2007.

[16].Liu. Y, T. Mei, X. Wu, and X.-S. Hua, “Optimizing video search reranking via minimum incremental information loss,” in Proc. ACM Int. Workshop Multimedia Information Retrieval, 2008, pp. 253–259.

[17].Liu. Y, T. Mei, and X.-S. Hua, “Crowd reranking: Exploring multiple search engines for visual search reranking,” in Proc. ACM Special Interest Group on Information Retrieval, 2009, pp. 500–507.

[18].Wang. D, X. Li, J. Li, and B. Zhang, “The importance of query-concept-mapping for automatic video retrieval,” *in Proc. ACMMultimedia, 2007.*

[19].Liu. Y, T. Mei, and X.-S. Hua, “Crowdreranking: Exploring multiple search engines for visual search reranking,” in Proc. ACM Special Interest Group on Information Retrieval, 2009, pp. 500–507.

[20].Van Rijsbergen. C. J, A theoretical basis for the use of cooccurrence data in information retrieval. Journal of Documentation, 33:106–119, 197