Concept-Based Query Expansion and Interactive Visualization for Web Image Search

by

© *Enamul Hoque*

A thesis submitted to the
School of Graduate Studies
in partial fulfilment of the
requirements for the degree of
Master of Sciences

Department of Computer Science
Memorial University of Newfoundland

December 2011

St. John's Newfoundland

Abstract

Even though Web image search queries are often ambiguous, traditional search engines promote only the most common interpretations of the query, which makes it difficult for the searcher to find the desired images. For addressing this problem, a concept-based query expansion technique is used to generate a diverse range of images covering multiple interpretations of the query. Then, a multi-resolution extension of a Self-Organizing Map is used to group conceptually and visually similar images. The resulting interface, named concept-based image search system (CBIS), allows the searcher to interactively highlight and filter images based on the concepts and zoom into an area within the image space to show additional images that are similar. Finally, a query refinement approach is proposed, which enables enhancement of queries based on explicitly selecting concepts or extracting concepts from example images. A series of evaluations illustrate the potential benefits of the system.

Acknowledgements

First of all, I would like to thank my supervisors sincerely, Dr. Minglun Gong and Dr. Orland Hoeber, who demonstrated how research should be done, provided insightful advices, and most importantly constantly challenged me to do better.

This thesis would not have been possible without the generous resources provided by the Department of Computer Science and Memorial University. My research is financially supported by my supervisor's grant (Research Development Corporation (RDC) Industrial Research Innovation Fund and NSERC Discovery Grants) and scholarship from the University's School of Graduate Studies.

I am thankful to Grant Strong, a doctoral student within our research group, for his active support during the conduct of my research work. Thanks also go to the other graduate students in the Department of Computer Science who provided me insightful comments and suggestions.

My dear friends in St. John's always made my life fun and interesting. I am especially grateful to all of them.

Last but not the least, I owe my deepest gratitude to my loving parents, Fazlul Hoque and Nargis Khanam, who have been my heros, friends, strongest supporters, and the reason I am me.

Refereed publications from this thesis

During the course of this research, a number of peer-reviewed publications have been made. The research work on concept-based query expansion technique (Chapter 3 and Chapter 4) was first published in [34]. Later, the enhanced method along with the evaluation depicting the key benefits of the developed system (Chapter 3, Chapter 4, and Section 6.3) was presented in [32]. Also, the trade-offs between diversity and precision arises in the context of the proposed query expansion technique (Chapter 3 and Section 6.2) is studied in [31]. As an outcome of this study, an automatic method for tuning the diversification parameter is proposed based on the ambiguity level of the original query. Furthermore, the approach for organizing the image search results based on conceptual and visual similarity (Chapter 4) was published in [74]. Finally, the major interactive visualization features of the system were summarized in [33].

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Web image retrieval has traditionally followed an approach that is an extension of Web document search techniques [42]. The textual information surrounding and associated with a particular image is used as the core information that describes the image. Queries are matched to this text to produce a set of search results which are then organized in paged grids. This approach can work well when images are concisely and accurately described within Web pages, and when searchers provide good descriptions of their image needs. Unfortunately, the accuracy of the descriptions given to images on the Web cannot be guaranteed; nor can we rely upon searchers to provide clear textual descriptions of the images they seek. When these conditions are not met, the search results may include many non-relevant images.

Image search tasks can be fundamentally categorized into two types: rediscovery tasks and discovery tasks [73]. Within rediscovery tasks, the searcher knows precisely what image they are looking for. By contrast, for discovery tasks, the mental model of the image that is being sought may be incomplete or vague; searchers may find that many images match their needs to various degrees. Often, Web image search tasks fall into this second category. It has also been noted that many Web image search queries are associated with conceptual domains that include proper nouns (e.g., people's names, locations, landmarks) [2, 37]. Searches of this type will often have a less specific aim with respect to resolving the image need (i.e., many different images may be viewed as relevant).

Further complications might arise due to the fact that Web image search queries are often very short [2], and open to many different interpretations. It is possible that different searchers may enter the same query, but their intentions and needs may vary significantly from each other. For example, the query "Beetle" might be interpreted as Volkswagen Beetle Car, Beetle (insect), or Abner Jenkins (a popular fictional character in comics who was known as Beetle). Even within a more specific query, searchers might have an interest in finding images that are related to but not explicitly identified in the query. For example, if a searcher submits a query such as "Washington, D.C.", they may wish to see some representative images of different landmarks situated within the city of Washington, D.C.

In such task scenarios, traditional search engines return results that contain many irrelevant images with respect to the searcher's interpretation. This is especially true if the searcher's meaning for the query is different than the most common interpretation. As a consequence, browsing the image set becomes tedious, frustrating, and unpredictable. In order to illustrate this kind of scenario, a sample of image search results from Google Image Search is provided in Figure 1.1. In this example, after

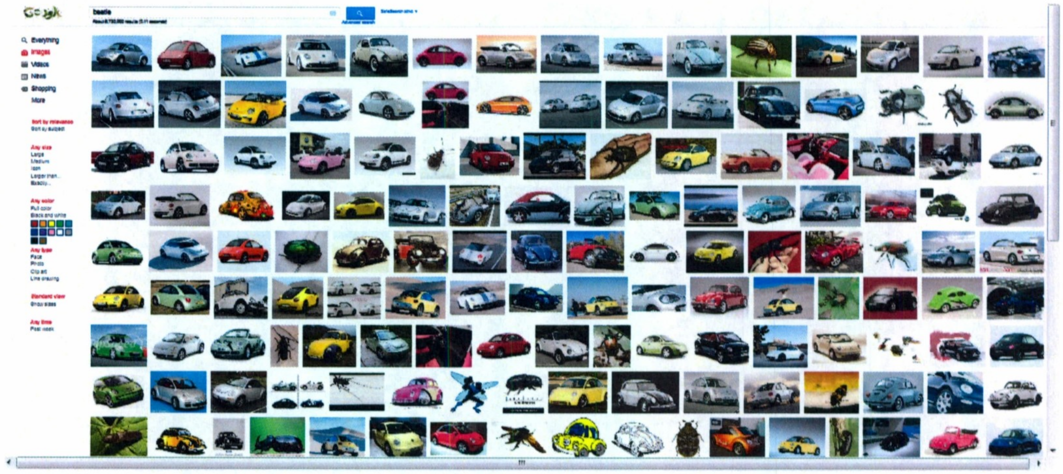


Figure 1.1: A screenshot showing the results of Google Image Search for query "Beetle" (July 20, 2011).

the searcher enters the query "Beetle", Google Image Search returns images from different interpretations including Beetle cars, Beetle insects, and the comic character. It is clear from the top search results that the interpretation of Beetle as car model dominates the image set, while less common interpretations like the comic character can rarely be found.

At this point, searchers can sequentially examine the search results to find images that are relevant to their needs. However, it is understandable from the sample results that the searchers' needs may not be fulfilled, particularly when their meaning is different than the common interpretation that the search engine promotes. Generally, the next step would be to attempt to modify the query by adding terms to make it more specific. However, Google Image Search itself does not provide any significant support for refining the query. As such, the searcher is responsible for manually modifying the query, which may be a difficult task for users who have difficulty articulating what it is they are looking for.

To alleviate this potential problem, this thesis outlines a new approach to image search that is able to extract the various conceptual aspects of a query, presents a diverse range of images covering those aspects to the searcher, and allows the searcher to actively explore within the image search results and refine the query. Such an approach moves beyond the traditional information retrieval focus of matching queries to documents, instead providing a Web Information Retrieval Support Systems (WIRSS) that aims to enhance the human element of Web search [85, 26]. The goal is to allow searchers to take an active role within the search process, which is particularly valuable for search tasks that are ambiguous or exploratory in nature [26]. By taking advantage of human decision-making processes, additional information can be provided at a high level to improve the searcher's ability to find images that are relevant.

1.2 Approach

The primary objective of this thesis is to address the fundamental issues related to the shortcomings of current Web image search methods by establishing a novel approach that introduces conceptual information into the process of retrieving and presenting the images; and the methods by which the searcher can interact with the system in exploration and refinement tasks. The overall research work can be divided into three different stages as discussed in the rest of this section.

The first stage is an automatic process, where a novel concept-based query expansion approach is used to retrieve a diversified set of search results when the query is ambiguous. Within this process, given a user supplied query, a number of concepts pertaining to the query are discovered from within Wikipedia. These concepts are ranked according to the semantic relatedness to the original query, and the top few concepts are used within the query expansion process to retrieve a range of images, providing a broad view of what is available. Here, the number of concepts to be used is dynamically determined by the degree of ambiguity of the query (based on how many different interpretations it matches within Wikipedia). As an example, if the searcher provides the query "Beetle", the system automatically discovers the related concepts with respect to the different meanings of "Beetle", ranging from different related models of Volkswagen cars such as Volkswagen New Beetle, Volkswagen Type 1, and Volkswagen Type 2, to the numerous species of beetle insects such as Rhinoceros beetle, Stag beetle, and Ground beetle. When, these concepts are used in the query expansion process, the search results become explicitly diverse.

Diversification of the image search results is highly desirable, particularly for ambiguous queries. However, showing such a broad image set using traditional grid-based representations causes the images from each of the different interpretations of the query to be scattered throughout the search results set. As such, finding images of a particular interpretation by linear evaluation becomes difficult and time-consuming. The second stage of the research addresses this limitation by organizing the images based on their visual and conceptual similarity, and designing a set of visualization and interaction techniques that support the searcher in exploring the image set.

Within the visualization approach, the search results set are organized based on the visual and conceptual similarity of images using a Self-Organizing Map (SOM). At the same time, all of the concepts used in the query expansion process are presented in a hierarchical manner. Used together, these two components provide the searcher with a visual summary of the image search results and an overview of the concepts related to the query. This organization technique allows the searcher to explore the images based on the conceptual and visual features. The exploration of images can be supported via a set of interactive features: concept-based focusing and filtering, which allows highlighting and filtering of semantically similar images; and zooming and panning, which supports exploration within the image search space.

The interactive features for exploring the image set provides an opportunity for searchers to identify specific concepts and images that are of particular interest. To extend further support, in the third stage of the research, the interactive query refinement features have been designed to create a query that is an accurate representation of the desired information needs. Unlike traditional models of image search, this interactive step supports the searcher to enhance the original query, or to deviate towards a new goal. This query refinement loop is facilitated based on the selection of concepts acquired through the query expansion process or the example images from the search results.

The resultant user interface is named as concept-based image search system (CBIS). Figure 1.2 shows a screenshot of CBIS for an example query "Beetle", demonstrating its interactive visualization features. From this figure, it can be seen that the images from the semantically related concepts are grouped together (e.g., images from Volkswagen cars are close to each other), while other images are far from each other (e.g., images of Beetle insects are separated from Volkswagen cars). The *concept hierarchy* organizes all the related concepts of the query into a tree view (left panel).

Figure 1.2: A screenshot of the CBIS interface, showing the results for the sample query "Beetle".

Searcher can interactively highlight and filter images based on the concepts in the concept hierarchy and zoom into an area within the image space to show additional images that are more conceptually and visually similar. Also, they can either choose a concept from the concept hierarchy or add a few example images into the *query panel* for further refinement (right panel). Used together, these operations support the searcher in exploring the image set and refining the query.

1.3 Research Questions

The key features in the image search approach developed in this thesis research include automatically expanding the original query using related concepts to retrieve a broad range of images, organizing the images and concepts, and supporting the searcher in actively exploring and refining the search results. Since this approach moves beyond the traditional image search paradigm, it leads to some fundamental research questions, which will be addressed in this thesis:

What is the impact of the proposed query expansion method on diversity and precision for short and ambiguous queries?

The primary aim for concept-based query expansion is to produce a diversified set of images that cover the various interpretations and aspects of the queries under the expectation that at least one of the interpretations matches the searchers' intent. Allowing searchers to then interactively explore images of a particular interpretation may allow them to perform their searches more effectively and efficiently.

However, in doing so there is a danger in broadening the query too much, resulting in a potentially significant decrease in precision. That is, the more broad and diverse the search results are, the less chance that a particular search result will be relevant to the searcher's information needs. As such, this trade-off between diversity and precision must be studied in order to understand the situations where more or less diversification is beneficial.

Does the approach for organization and exploration of image search results improve the user performance for image search tasks?

Our expectation is that for the search tasks that are rather complex or exploratory in nature, the proposed visualization and interaction techniques incorporated in CBIS may effectively support the searcher in exploring images. As such, this may lead to improving searchers' performance in terms of accuracy and time taken for completing the tasks. These claims will be validated via analysis and user evaluations.

Does the approach for organization and exploration of image search results improve the user's perceptions of usefulness, ease-of-use, and satisfaction while performing image search tasks?

Sometimes quantitative measures (such as accuracy and time taken for completing the tasks) may not be good indicators of the user experience, particularly when the search tasks are vague or ambiguous [27]. Since the approach for organization and exploration of image search results is particularly targeted at such complex tasks, alternative subjective measures such as perceived usefulness, ease-of-use, and satisfaction in completing image search tasks can be helpful in validating the potential benefits that CBIS may provide.

1.4 Primary Contributions

One of the fundamental contributions of this research is the novel technique of concept-based query expansion for Web image retrieval using Wikipedia. For short and ambiguous queries, traditional approaches generally promote the most common meaning, making it very difficult for searchers to find images that match their needs. On the contrary, the query expansion technique described in this thesis automatically diversifies the search results with the expectation that at least one interpretation of the query will match the searcher's needs. When such diverse search results are presented within the visual interface, searchers can effectively explore the search results to find images of interest. Within the query expansion process, the novelty comes from the use of Wikipedia as a knowledge base, as well as the methods by which relevant information is extracted and used. Although there are a few approaches that use

Wikipedia for query expansion to improve image retrieval accuracy in general [53]; to our knowledge, using Wikipedia for query expansion in the context of diversifying the image search results and supporting searchers in the exploration of Web image search results has not been done before.

The second major contribution is the approach of using visualization and interaction techniques for exploring within the image search results. The primary method of organization is motivated by similarity-based image organization based on a multi-resolution Self-Organizing Map (SOM) [70, 73]. Prior work in this area has focused on organizing images using visual features only; in this research the conceptual features derived from the query expansion are introduced into the image organization process. Moreover, the previous approach [73] mainly considers browsing and exploration activities within image collections, while this research includes a number of novel interactive visualization features with the goal of providing a comprehensive Web image retrieval support system that provides interactive support to the searcher for exploring within the image search space.

The third major contribution is the approach for interactive query refinement. Most of the previous query refinement methods attempt to extract the visual features from the images to find more images that are similar to the given images, which often lead to semantic gap during the retrieval process [13]. However, in this work, knowledge of the source concept for each image is known. Therefore, when a searcher selects an image and asks for "more like this", the concept is used during the query refinement process, allowing the retrieval of more semantically similar images to the given images.

An additional contribution arises from the outcome of the evaluations that are used to validate the proposed approach. A study of the trade-offs between diversity and precision for Web image search using concept-based query expansion examines the potential side-effect of precision due to diversification of the search results. As an outcome of this research, a method for tuning the degree of diversification parameter is suggested based on the degree of ambiguity of the original query. In another empirical study, two new precision metrics are presented to measure the effectiveness of the exploration features of CBIS. This study demonstrates that the precision of the image search results increases as a result of concept-based focusing and filtering, as well as visual zooming operations. Finally, a user study was conducted to in order to measure the potential benefits provided by CBIS in comparison to a traditional Web image search interface. This study shows potential improvement in searcher's performance for image search tasks, and how this results in a positive search experience.

1.5 Organization of the Thesis

The remainder of the thesis is organized as follows. In the next chapter, an overview of related work is provided. Chapter 3 outlines the techniques used for query expansion. Chapter 4 explains the approach used for organizing the search results and describe the interactive exploration features that enhance the image search experience. In Chapter 5, the interactive query refinement feature is described in detail. In Chapter 6, three different evaluations are conducted: the effect of diversity over precision is studied via a set of queries having different degrees of ambiguity; the effectiveness of the exploration features are analyzed via an empirical study; and the potential benefits of image set organization and exploration features are evaluated through a user study. The thesis concludes with a summary of the research contributions and an overview of future work in Chapter 7.

Chapter 2

Related Work

2.1 Fundamentals of Web Image Retrieval

2.1.1 Traditional Image Retrieval Model

The primary method of image retrieval used on the Web is based on keyword search [42]. Search engines have adapted their document retrieval algorithms to the metadata (keywords, tags, and/or associated descriptions) of images and present the results in a scrollable list that is ranked based on relevance to the query. In addition, keyword search relies on the assumptions that the contents of images are accurately described by the metadata, and that the searcher is able to provide a concise description of what they are seeking; these assumptions are not always valid.

In the situation where the query is ambiguous, the search engines generally promote the most popular interpretation of the query. However, if searchers have different meanings than the popular one, then they need to go through the list sequentially until they find the desired images. These may become very difficult and time consuming tasks for searchers, as they need to go through a large amount of irrelevant images with respect to their image needs. The lack of support for refining the queries may also add more difficulty to the searchers. That is, they need to modify the query manually to create accurate representations of the image needs, which is difficult for the common searcher [2].

Another approach to image retrieval that has been explored within the research literature is content-based image retrieval (CBIR), which uses the visual features of the images for searching purposes [13, 67]. According to this approach, searchers need to provide one or more images, which are used as the query. Alternatively, they can draw sketches as visual queries. As a consequence, the system returns a set of images that are visually similar to the given images/sketches. However, they often lead to a semantic gap: the gap between the way a person finds similarities between images at the conceptual level and the way the system generates similarity based on pixel statistics [13].

Despite the above mentioned limitations, popular Web image search engines have started to adopt CBIR techniques. Google Similar Image allows the searcher to provide an image as query, and returns the images that are visually similar to this given image (see Figure 2.1). But due to the aforementioned semantic gap, it often returns poor search results. As an example, we present the search results provided by Google Similar Image that are visually similar to a query image of a 'black shoe'.

Clearly, according to the results shown in Figure 2.1, the system fails to recognize the given object as a shoe, instead returning images of other objects that are similar in colour only. This illustrative example provides us the evidence that existing CBIR techniques may not be able to adequately recognize the rich semantic meaning of the

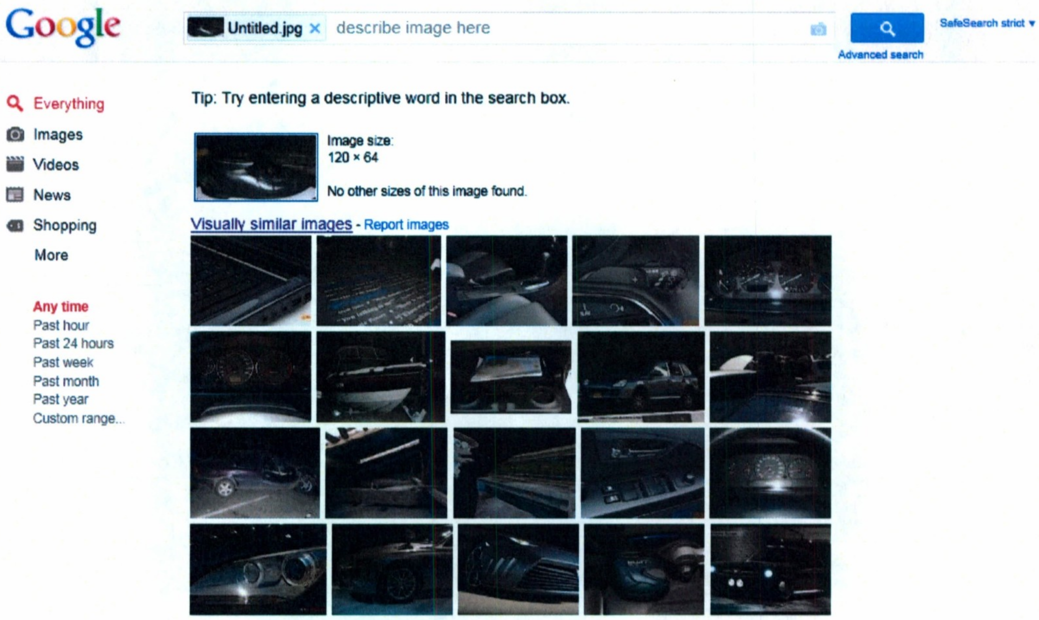


Figure 2.1: A screenshot showing the results of Google Similar Image for an image query of black 'shoe' (August 5, 2011).

queries.

2.1.2 Web Image Search Behaviour

Although current image search systems are mainly based on document search techniques, a number of studies have identified image search behaviour as being significantly different than document search behaviour [2, 23, 37]. It has been observed from both query and session-level search statistics that image searchers view more pages of search results, they spend more time looking at those pages of search results, and they click on more results than Web document searchers [2]. Overall, this indicates that image search tasks are generally more exploratory in nature than document search tasks. Possible reason for this user behaviour is that there is often no definitive answer to an image query; the sought after image could be one of many, and is based on subjective opinion. The other reason could be the visual nature of image searching that may lead searchers to visit images solely for their aesthetic value.

Web image search queries are often very short and ambiguous [2, 23, 82], a characteristics which is also common with document search [61]. Goodrum and Spink found that image queries contained on average 3.74 words (although commonly at least one of these was the image request term, e.g. 'picture' or 'image') [23]. It has also been noted that Web image search queries commonly include many proper nouns (e.g., people's names, locations, landmarks) [2, 37]. Searches of this type will often have a specific focus, but a less specific aim with respect to resolving the image needs (i.e., many different images may be viewed as relevant).

Another important characteristics obtained from session-level statistics is that the majority of query sessions for image search comprise a trail of *related* queries. This indicates that query refinement tasks are more common in image search. However, such refinement is particularly challenging in image search where a searchers' needs (a mental picture of the desired image) may be far more difficult to specify with traditional keyword queries. Furthermore, the visual nature of image search makes it very easy to become side-tracked when something else of interest catches the searcher's eye, even if the initial query is task-focused. This phenomena described as *serendipity*, or the act of unexpectedly encountering something fortunate, has long been identified as valuable. Particularly, if a searcher is browsing for entertainment, this kind of tangent is rather desired [2].

2.2 Query Expansion

2.2.1 Fundamentals

Query expansion is the process of adding a number of meaningful terms to an initial query in order to produce a better set of search results [18, 7]. This process of adding terms can either be manual, interactive, or automatic. Manual query expansion relies on a searcher's expertise and knowledge to identify additional terms to add to their query. Interactive query expansion identifies and presents to the searcher an ordered list of potential query expansion terms, allowing them to explicitly choose which to add to their query. Automatic query expansion calculates and assigns weights to a set of candidate terms; those with the highest weights are added to the initial query without any effort or intervention by the user. Since different weighting functions produce different results, retrieval performance depends on how the weights have been calculated.

Query expansion techniques may also be classified according to the type of information used in order to generate the expansion, including relevance feedback, corpus dependent knowledge models, and corpus independent knowledge models [7]. Relevance feedback is an established technique for modifying an initial query using words from user-identified relevant documents, or top-ranked documents (pseudo-relevance). Techniques based on corpus dependent knowledge models generate query expansion by matching the original query to a knowledge base that was created based on some or all of the collection being queried. The problem with traditional relevance feedback techniques and corpus dependent query expansion is that they are content driven. The corpus content is analyzed to extract candidate terms for query expansion. This can only work if there are sufficient relevant documents to work with and also that these documents contain a reasonable set of terms that represent the subject area for the query.

Corpus independent knowledge models eliminates this problem by taking advantage of an external knowledge base. Such a knowledge base can be in the form of a thesaurus or an ontology. Generic thesauri (e.g., WordNet [48]) may be too broad and shallow to provide comprehensive coverage of specific topics. However, Wikipedia has much greater domain coverage compared to WordNet and and has shown to improve retrieval performance for a variety of retrieval tasks [51]. There are also some domain specific ontologies that could be use as knowledge bases (e.g., ACM Computer Classification System for the computer science domain [1]). The major problem with domain-specific ontologies is that they are expensive to construct and may not be available in many domains [7].

2.2.2 Query Expansion for Image Retrieval

Unlike the general domain of document-centric information retrieval, query expansion has been studied in only a few instances within the context of Web image retrieval. Those that have explored such techniques have shown them to be promising [53]. Since query expansion can be an effective tool in order to promote diversity and provide assistance in query refinement, it can be beneficial for image search. This is especially true for searchers conducting discovery tasks where their information needs are not clear enough, and many different images may be viewed as relevant.

However, there are a number of challenges associated with using a query expansion process for image retrieval. The first problem is finding a suitable knowledge base that has sufficient coverage of the conceptual domains that are common in image search. Often the textual medata and the noisy textual descriptions do not provide enough information for expanding the query, which suggests that it may be better to use an external knowledge base. While, WordNet [48] has been used for this purpose in some image retrieval research [41], it does not contain information pertaining to the proper nouns that are common in image search queries. Using Wikipedia to extract the related concepts for reformulating queries has shown promise [53], since it has much better coverage of conceptual domains. Another knowledge base named Yago [76], that automatically extracts relational facts about entities from Wikipedia and unifies these with WordNet, has also been used to reformulate the queries. This approach has been evaluated using photo retrieval tasks with less popular entities, resulting in significant improvement in retrieval performance [77].

The second challenge is to design efficient and effective algorithms that can process such semi-structured knowledge to derive and rank the meaningful terms to be used in the query expansion process. A useful approach to this problem is to measure the semantic relatedness between the original query and each of the concepts derived from that query. A number of different methods have been devised to use Wikipedia for this purpose, including WikiRelate! [75], Explicit Semantic Analysis (ESA) [21], and Wikipedia Link-based Measure (WLM) [49]. Among these methods, WLM has been reported to be computationally more efficient and accurate [49].

The third issue to resolve is to decide whether to make the query expansion process interactive or automatic. While interactive query expansion has been extensive studied for the document-centric information retrieval [20, 30, 40], it has not been well studied in the domain of image retrieval. The challenge in applying interactive query expansion in image retrieval is that by allowing the searcher to explicitly choose terms from a given list will force them to continue to think about how to describe their image needs, when what they really want to do is look at the images. Alternatively, the query can be automatically expanded using the most similar concepts to the original query [53]. In doing so, the query expansion and image search result diversification processes become seamless to the searcher, allowing them to delve into the exploration of the images without delay.

The last challenge is to choose a useful method for organizing the image search results. By expanding the query, the number of images within the search results can grow very large. Further, by broadening the search results for a somewhat ambiguous query, many images that are not relevant to a particular interpretation of the query may be included in the search results. As such, the rank of the search results from a particular expanded query may not be as important as the visual features of the images [53]. A useful approach is to take advantage of the visual and semantic features of the images when organizing the search results, and then allow the searcher to browse and explore within the search results space. Examples of such approaches for organizing image search results are discussed in more detail in Section 2.3.

2.2.3 Issue of Diversification

One of the main purposes of automatic query expansion is to enhance diversity in the search results set (i.e., to ensure that different interpretations of an ambiguous query are retrieved). This problem of image search diversification has been recognized as an important topic in the research community. Most of the previous approaches to search results diversification can be categorized as either implicit or explicit [62]. Implicit search results diversification approaches can be viewed as some form of clustering by performing direct comparison between the retrieved images, under the assumption that similar documents will cover similar aspects [12, 87]. The alternative approach to diversification is to explicitly utilize different aspects or interpretations associated with a query by directly modelling these aspects [63]. Automatic query expansion can be regarded as an explicit way of diversifying search results, where each of the query aspects are represented as *sub-queries.*

While such explicit diversification is often desirable, broadening the query too much may result in a potentially significant decrease in precision. That is, the more sub-queries that are used in the query expansion, the less chance that a given sub-query will be relevant to the searcher's intents for the original query. Therefore, it is important to understand the situations where queries should be expanded more and where they should be expanded less.

In this context, some have begun to study of how to minimize the trade-off between precision and diversification within the context of image search. For example, Song et al. [68] propose a re-ranking method based on topic richness analysis to enrich topic coverage in retrieval results, while maintaining acceptable retrieval performance. A topic richness score was computed by analyzing the degree of mutual topic coverage between an image pair based on the assumption that images are annotated by several words. Van Zwol et al. [79] proposed a method for optimizing precision-diversity trade-off by estimating a query model from the distribution of tags that are associated with the images.

Maintaining a good balance between diversity and precision is challenging, since it requires an automatic modelling of the searcher's query to determine an appropriate degree of diversification to promote. In many cases, image search queries are inherently ambiguous (e.g., "Beetle"). In other cases, a query might be very specific, and the scope for broadening the query may be limited. For example, queries for a particular landmark within a specific setting like "Eiffel Tower Bastille Day" may leave little room for diversifying the search results. Therefore, it is necessary to automatically determine the level of ambiguity for the given query. To deal with this issue, Song et al. classify queries into three types: ambiguous queries, broad queries, and clear queries. They propose a machine learning approach to automatically identify ambiguous queries based on the query and its top returned documents [69].

2.3 Organizing and Visualizing Image Search Results

Among the popular modern search engines, the presentation techniques of image search results have changed very little over the last decade: a paged grid layout is used to organize images based on their rank of relevance. While such grid interfaces are easy to use, they provide limited ability to manipulate and explore the search results. Further, simply ordering by rank of relevance does not promote diversity within the search results set. Rather, similar images from the most common interpretation of the query are grouped near the top of the grid. Considering the limitations of traditional approaches, a number of organization and visualization techniques are proposed including clustering, tree-based visualization, focus+context, three dimensional visualization, and similarity-based image browsing.

2.3.1 Clustering

A number of works have considered clustering images based on visual features alone [47, 78], or the combination of tags and visual content [22, 52, 16] to visualize image search results. The premise in these approaches is that by clustering images, the diversity within the collection can be shown. Within image search, the problem has been approached from the perspective of jointly optimizing the search precision and diversity using dynamic programming [15]. Unfortunately, such approaches may not be sufficient to capture semantic diversity, which requires a deeper level of knowledge about the images. Cai et al. organized the results into different semantic clusters with the goal of facilitating browsing activities [11]. They proposed a hierarchical clustering method using visual, textual, and link analysis that is mainly targeted at clustering the search results of ambiguous queries. According to this algorithm, the image search results are clustered at two different levels. At the first level, images are clustered based on their textual and link representations. Each resultant cluster is regarded as a semantic category. Then, at the next level, low level visual features from images are used to cluster each semantic category. Several images from each category are selected as representative images according to their image ranks, which enables the user to understand the main topics of the search results quickly.

Wang et al. argued that clustering hundreds of images using high dimensional visual features might not be efficient for practical use [82]. They presented IGroup, an alternative system for semantic clustering of image search results, based on only textual analysis of the search results (see Figure 2.2). Through a user study, they reported a significant improvement in terms of efficiency, coverage, and satisfaction. The difficulty with such explicit clustering approaches is that they partition the images, requiring the searcher to explore each cluster separately.

To address the shortcomings of the previous research works, another cluster based image search interface is proposed that employs both textual and visual analysis. They cluster the key phrases of the textual descriptions into semantic clusters followed by grouping images corresponding to each key phrase into some visually coherent clusters. An user interface is designed to explore the hierarchical clustering results [16]. However, navigating through the hierarchical clustering results might be difficult as the searcher needs to click on a semantic cluster followed by clicking on one of its visual clusters to ultimately examine an individual image.

Clustering techniques have also shown some promise in promoting diversity in the search results because of their discriminative power [78]. Also, with the help of a user friendly interface, they may provide more meaningful representations of image search results [60]. Most of the clustering algorithms, however, share some common challenging issues including determining a suitable similarity measure, the efficiency of clustering algorithms, determining an appropriate number of clusters to use, and deciding how to order or organize the clusters.

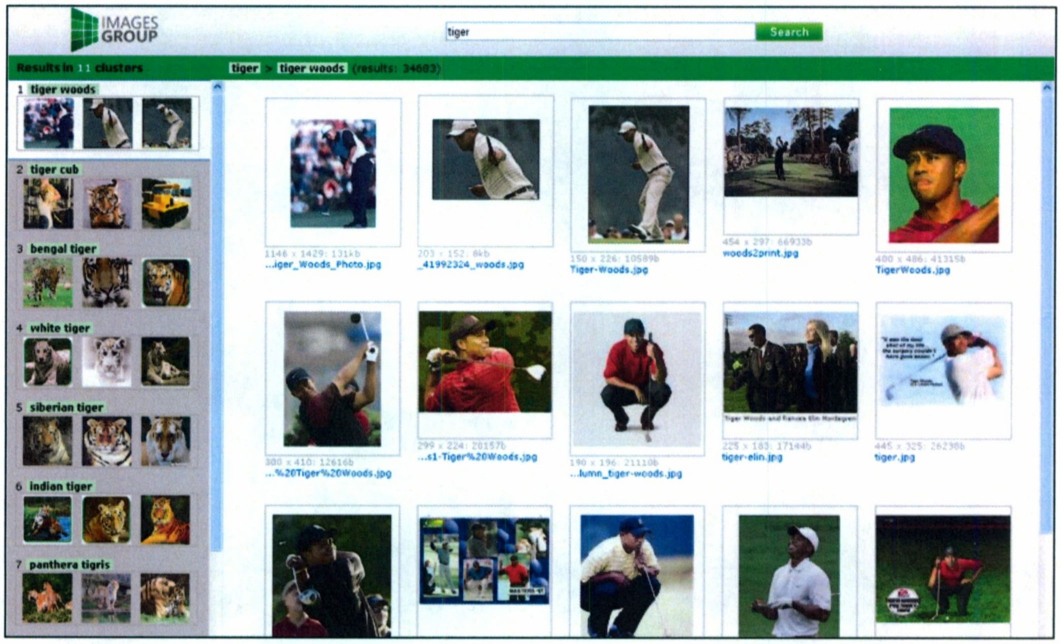


Figure 2.2: A screenshot showing the results of IGroup for the query "Tiger". Images are grouped into semantic clusters based on textual features [82].

2.3.2 Tree-Based Visualization

Organizing the images in an hierarchical manner, and presenting them using tree visualization has been applied in seminal works as well as in more recent works. In PhotoMesa [5], a tree-map algorithm was used to present large numbers of images grouped by directory, or other available metadata, to better deal with the space constraint. Hierarchical Cellular Tree (HCT) has also been proposed to enhance browsing, which enables the user to navigate through the image database [43]. Kustanowitz and Shneiderman introduce a bi-level radial quantum layout where one region is designated for primary content, which can be a single photo, text, graphic, or combination.

Adjacent to that primary region, groups of photos are placed radially in an ordered fashion, such that the relationship of the single primary region to its many secondary regions is immediately apparent [45].

Recently, Google Image Swirl presented a visualization method for hierarchically clustered images using a radial layout (see Figure 2.3), in which each layer of the tree is arranged radially around its parent. When the user selects a branch of the tree to explore, it is separated from the parent and expanded, while the parent is shrunk to make space. This allows the user to quickly navigate to the images of interest [39]. This is a promising approach, but the number of images shown at any one time is limited. Also, the image search results were pre-organized [38], which required significant amount of off-line pre-processing of the data to generate the trees.

2.3.3 Focus + Context

Focus + Context techniques provide both overview and detail in one unified interface by spatially distorting a data representation, giving more room to designated points of interest and decreasing the space given to regions away from those points [83]. For exploring large collection of images, Focus + Context techniques are regarded to be promising (e.g., hyperbolic trees [19] or fisheye views [47, 59, 88]). In JustClick, the *topic network* is automatically generated from popular image topics at Flickr and their inter-topic contextual relationships [19]. Then, the *hyperbolic visualization* method is used to enable interactive navigation and exploration of the topic network, so that users can gain insights of large-scale image collections at a glance, build up their mental query models interactively, and specify their queries (i.e., image needs) more

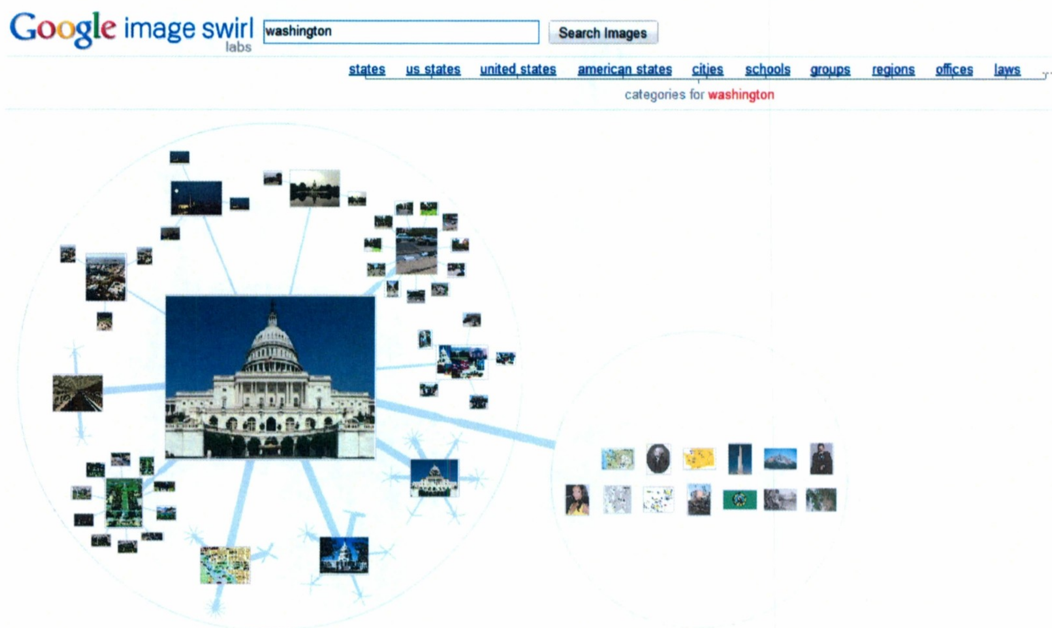


Figure 2.3: A screenshot showing the results of Google Image Swirl for the query "Washington". The images are hierarchically clustered and presented using a radial layout [39].

precisely by selecting the image topics on the topic network directly (see Figure 2.4).

A number of research works utilize fisheye views for image search results presentation [47, 59, 88] to allow users to see local details and global perspective simultaneously. Liu et al. propose a fisheye view for visualizing image search results [47]. Users can interact with the overview by dragging a slider to adjust the global overlapping ratio, and click on interesting images to generate a detail view. A *fisheye view* is then generated to provide this detailed view function. The fisheye view is also utilized in exploring an annotated image collections via multiple foci and multiple contexts [59].

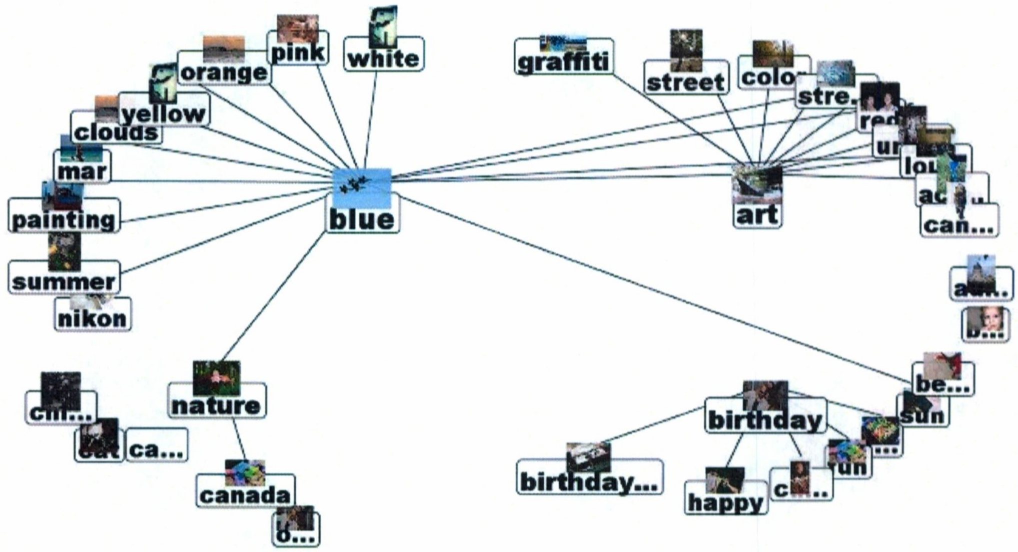


Figure 2.4: JustClick allows interactive exploration of the large-scale manually-annotated Flickr image collections at the topic level by changing focus [19].

As the user changes the focus by navigating between images, keywords, and concepts, the semantic fisheye view dynamically resizes the images in the collection to reflect the user's interest.

More recently, a combination of concept map, Venn diagram, and fisheye views has been proposed to enable exploratory querying and browsing of a semantically annotated image database [36]. It allows the user to initially define a query by selecting one or more concepts. Then, the system employs Venn diagrams to visualize the relations between the query concepts (see Figure 2.5). To examine the content of the query concepts they employ a fisheye view. Within this view, distortion of the images in the display is applied according to their distance from the focus, while easy navigation is obtained by showing the map of the selected concepts supporting the

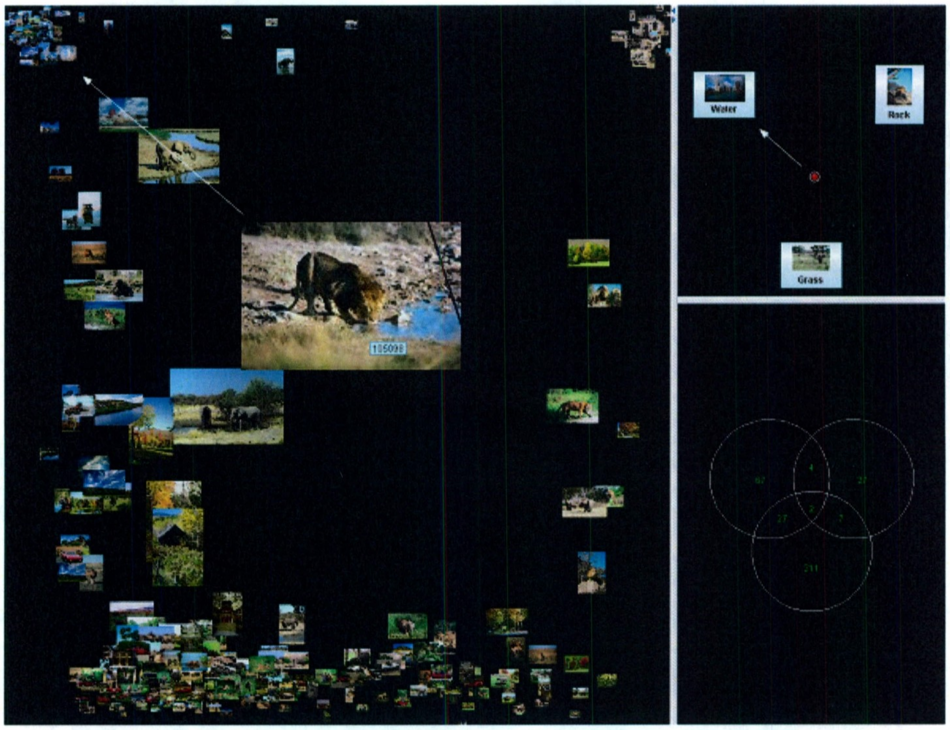


Figure 2.5: Fisheye view of the concept space while browsing the images. Here, two arrows indicate the movement of the users cursor and the corresponding movement of the red dot icon in the concept space [36].

movements through the distorted space.

2.3.4 Three Dimensional Visualization

Recently, three dimensional visualization techniques have been proposed for exploring image set [58, 64, 35]. Porta developed a number of three dimensional visualization techniques including shot display, spot display, cylinder display, rotor display, tornado display, and tornado of planes display [58]. These methods may be useful in the situations where users need to look at the entire image collection, because they do not know exactly what to search. However, these visualizations may make it difficult to see individual image in detail, due to occlusion and distorted views of the images.

Motivated by Porta's visualization techniques, shot display is used within an image search interface named ImageFlow [35]. Here, images are visualized like bullets fired by a virtual gun towards the screen. Semantic features (e.g., relevance, related queries) are mapped to the spatial dimensions (e.g., x, y, z) in a way that allows for several levels of engagement: from passively viewing a stream of images, to seamlessly navigating through the semantic space, to actively collecting images for sharing and reuse (see Figure 2.6). This is an interesting approach for visualizing large collection of images, however, decoding the semantic features that are mapped in different dimensions may require extra cognitive effort.

Three dimensional visualization techniques are generally expected to be aesthetically pleasing. They can be also useful in providing a global overview of a large collection of images by creating the extra space that the third dimension provides. However, occlusion and irregular size of the images might become an important concern for most of the three dimensional visualization techniques. There are also often interaction difficulties in three-dimensional visualizations, making the interface less easier to use.

2.3.5 Similarity-Based Image Browsing

Similarity-Based Image Browsing (SBIB) is an approach that takes advantage of the fundamental aspects of CBIR (described in Section 2.1), but eliminates the need for

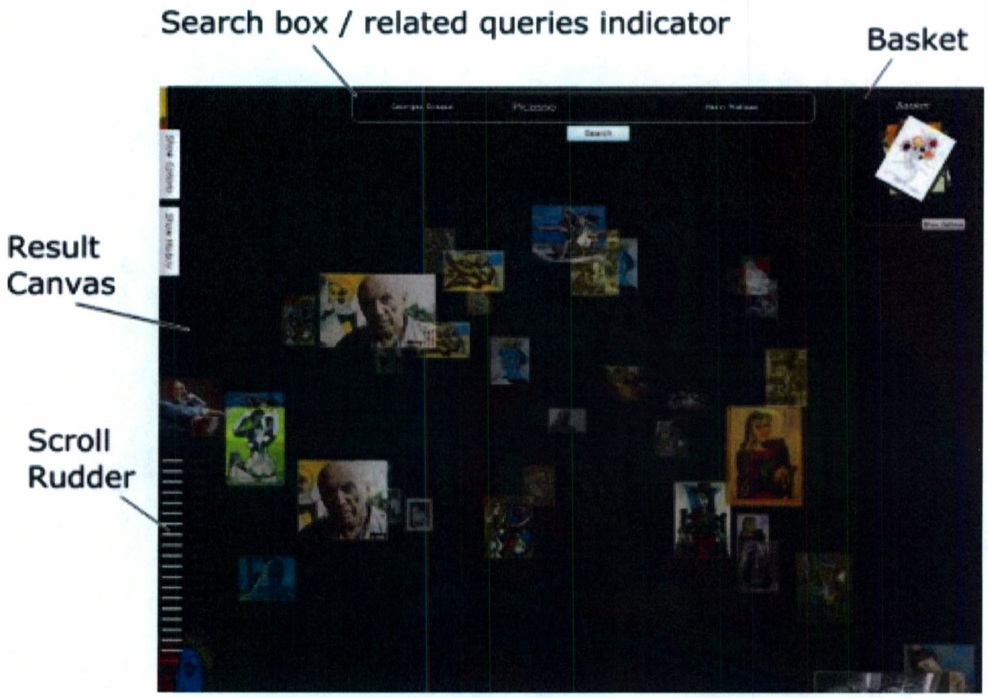


Figure 2.6: A screenshot from ImageFlow that maps semantic features of the images to the three spatial dimensions [35].

the searcher to identify a set of relevant images as *a priori*. Images are organized based solely on their features, allowing searchers to explore the collection even if they do not have a clearly defined information needs [24]. The challenge of SBIB is to arrange images based on visual similarities in such a way as to support the browsing and exploration experiences.

Several SBIB approaches have been proposed, all of which use color, structure, and texture features as the basis of similarity measures for their organizations, but differ in how those similarity measures are used to relate images [24]. Heesch et al. model relations between images using nearest neighbor networks adapted from document browsing techniques [25]. Search, in their case, is the process of following a path through the series of connections by clicking relevant images. Nguyen and Worring propose a system to meet the three requirements of overview, structure preservation (as it relates to image similarity), and visibility in [54]. They organize the images using non-linear probabilistic methods and k-means clustering to determine overview images while browsing. Pecenovic et al. use Sammon's projection to map the images onto 2D space for visualization and use a heuristic balanced k-means algorithm for determining representative centroids [56]. Strong and Gong's approach also maps images onto 2D space, but uses a multi-resolution Self-Organizing Map. This approach provides both an organizing structure for the images and a measure of importance that can be used to determine which images to display when space is limited [70, 71]. Through a user study, the interactive features within this approach have been shown to be very useful and easy to use [73].

2.4 Query Refinement

Query refinement is an essential information retrieval tool, that presents system-generated terms to assist the user in articulating their query when it is poorly defined or vague [18, 81]. A number of works in query refinement for document searching have focused on the selection of refinement terms, and how they are presented to the user. Anick and Tipirneni suggest a method for generating query refinement terms using lexical dispersion of a word - the number of different lexical compounds in which the word occurs [4]. The algorithm takes a list of documents as input, and returns the most disperse words as facets and the most frequent lexical compounds as value of the facets. Joho et al. investigate interactive query refinement using hierarchical query term presentation [40]. They show that hierarchical presentation requires significantly less effort and is superior in terms of user satisfaction. Fonseca et al. present a query refinement strategy based on mining association rules from query logs [20]. The refinement terms are clustered into concepts where similar terms appear together.

The effectiveness of the visualization support for the interactive query refinement has also been studied in a few works [28, 17]. WordBars creates a visual representation of the frequencies of the terms found in the top ranked document surrogates returned from an initial query, in the form of a histogram. Terms from the histogram can be easily added or removed from the query, generating a new set of search results. One of the key benefits of this visualization is that it allows the users to recognize terms from the list, rather than having to recall relevant query terms for a given topic [28].

Query refinement is no less necessary in image search than in other search applications. This is due to the fact that the ambiguous nature of the queries (which are more common in this domain) requires further interactions from the user to narrow down the scope within the broad ranges of search results. However, one of the fundamental challenges here is that in document centric approaches, refinement terms are most often generated from the underlying corpus [4] or from the set of retrieved documents [4, 40]; while in Web image searching, only textual information surrounding and associated with the retrieved images are available for use, which is not sufficient and reliable enough to produce accurate refinement terms.

Some other alternative methods have been studied for query refinement in the context of content-based image retrieval. One of these approaches analyzes the content of the retrieved images to provide sufficient clues for query refinement. This strategy is often based on relevance feedback techniques, which shift from single-round queries to navigational queries where a single retrieval instance consists of multiple rounds of query reformulation. This can be considered as supervised learning to adjust the subsequent retrieval process by using information gathered from user feedback. An image retrieval system named PicSOM adopted this technique using multiple parallel Self-Organizing Maps (SOMs). Here, the relevance feedback technique is based on spreading the user responses to local SOM neighborhoods by convolution with a kernel function [46].

Although relevance feedback techniques in image retrieval have been studied for over a decade, we do not see many real world implementations. This is potentially due to the feedback process that the users must go through, which can be viewed as obtrusive and inconvenient [13]. Moreover, simply providing binary feedback over some retrieved images may not be adequate to capture the underlying high level semantics that prompt users to provide the feedback.

One way to enhance the semantic information in the query refinement process is to make use of an external knowledge base, which can be processed to discover the underlying aspects of the queries. Such a knowledge base enables users to narrow down the scope of search results based on the provided relevant knowledge. In this context, limiting the scope of the search results can be done by allowing users to browse through system-generated facets. Yee et al. [86] present an interface for searching and browsing images by making use of hierarchical faceted metadata and dynamically generated query previews. They apply the system for searching an art catalogue where the faceted meta-data is partially provided by the collection itself and partially extracted using WordNet [48]. User evaluations have shown that this faceted browsing of image search results is particularly useful for exploratory search tasks.

Recently, the researchers from Yahoo! describe a system that enables faceted exploration of Web image search results [80]. The system processes semi-structured information sources to extract objects and facets. Then the facets are ranked based on a statistical analysis of image search query logs and the tagging behaviour of users annotating photos in Flickr, and finally presented to the users to perform query refinement (see Figure 2.7). Such faceted exploration may be regarded as a promising approach for interactively refining the query. However, the system allows only one level of refinement, which may not be sufficient enough for the queries when the searcher starts with queries that are too ambiguous.

2.5 Discussion

In this chapter, the traditional image retrieval models were briefly reviewed and their limitations in addressing the unique image search behaviour were identified. An overview of the seminal works as well as some of the new approaches were surveyed that aim to improve the traditional image search methods. In particular, three different aspects of Web image search were discussed that are relevant to this research: query expansion, organizing and visualizing image search results, and query refinement.

Query expansion has been found to be a promising approach in improving the

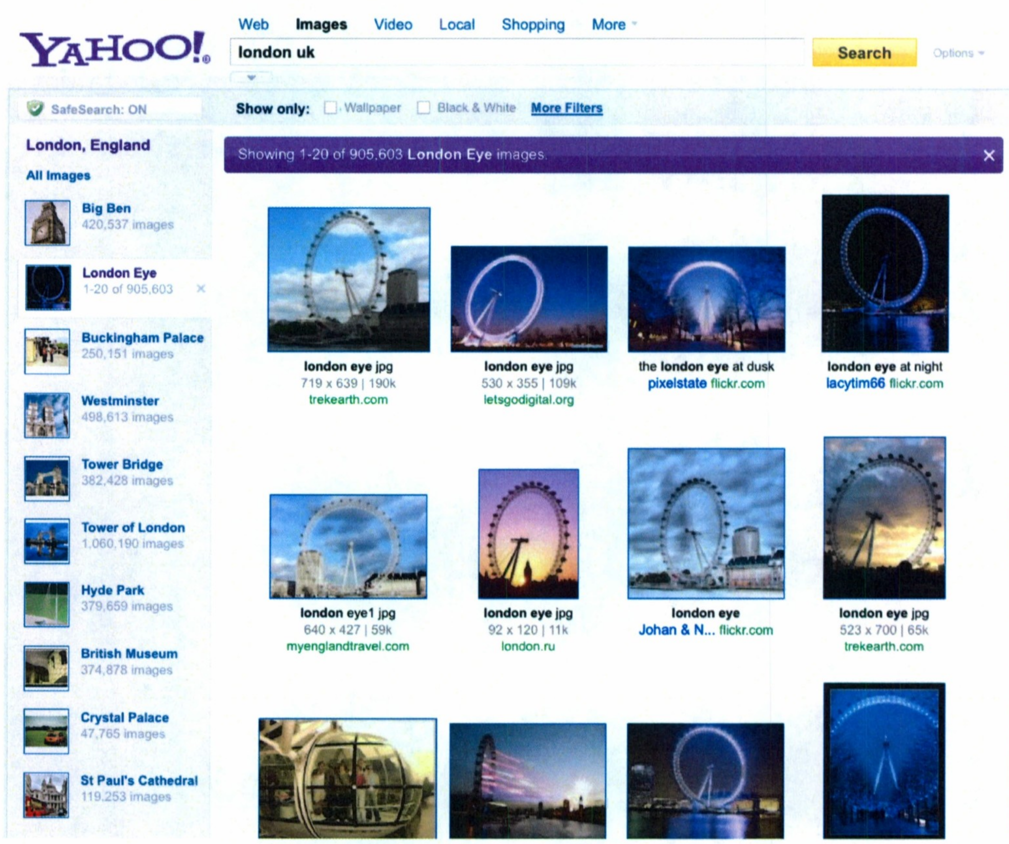


Figure 2.7: A screenshot from Yahoo image search results, when the user queries for "London UK" and has clicked on the "London Eye" facet [80].

image search results and assisting the searcher in refining the query by suggesting system-generated terms. However, from the literature review on query expansion, a number of challenges were discussed. One of the fundamental challenges was to find a suitable knowledge base for choosing the relevant terms that can be used in query expansion. In this direction, different knowledge sources for query expansion were identified as well as their potential advantages and limitations were discussed. These aspects of the research served as the guiding principles in the design of the query expansion technique to be introduced in Chapter 3.

In Section 2.3, a number of organization and visualization techniques were discussed that aim to address the limitations of the traditional list or grid-based approaches to organizing images. Each of these techniques have their own advantages and limitations. Although clustering techniques may provide an intuitive form of image search results representation, designing an efficient and effective clustering algorithm is challenging. Tree-based visualizations can provide hierarchical representations of images, but possibly at the cost of limiting the number of images to be shown at once. Focus + context techniques and three dimensional visualizations suffer from the irregular size of the images and some interaction difficulties, despite the fact that they may be useful in providing the overview of a large image collection. Similarity-based image browsing has been regard as a promising approach; however, some scopes of research exist with respect to defining the similarity between images and addressing the display space constraints. The study of all the above stated methods have provided useful insights into designing the interactive visualization features for image search to be discussed in Chapter 4.

Finally, in Section 2.4, a set of techniques for query refinement were studied with respect to the the domains of text-based and content-based image retrieval. Relevance feedback techniques are commonly found in the domain of content-based image retrieval which attempt to capture the user's precise needs through iterative feedback. In the domain of text-based image retrieval, a number of research works have been conducted with an aim to find suitable terms and create their meaningful representation for the selection of query refinement. The advantages and disadvantages of these methods were taken into consideration while designing the interactive query refinement features to be discussed in Chapter 5.

Chapter 3

Concept-Based Query Expansion for Image Search

3.1 Motivation

Even though Web image search queries are often ambiguous, traditional search engines retrieve and present results solely based on relevance ranking, where only the most common and popular interpretations of the query are considered [42]. Rather than assuming that all users are interested in the most common meaning of the query, a more sensible approach may be to produce a diversified set of images that cover the various aspects of the query, under the expectation that at least one of these interpretations will match the searchers' information needs. Motivated by this hypothesis, a novel query expansion technique has been designed that explicitly diversifies the image search results. The system automatically understands the different aspects of the given query which are defined as *concepts*. Then, instead of using only the original query, a number of sub-queries are created based on these concepts, resulting in the retrieval of a diverse set of images.

In this chapter, the proposed query expansion technique is described in detail. The steps involved in the process of deriving concepts pertaining to the query, and how to apply these in the query expansion process are explained. Where necessary, illustrative examples are provided to depict how the approach works. At the end of the chapter, the potential advantages, issues, and limitations of the work are discussed.

3.2 Knowledge Base for Query Expansion

As identified in Section 2.2, one of the fundamental problems for query expansion is how to select the most suitable or relevant terms that are to be used in the process. In the absence of accurate textual descriptions of the images, an alternate idea is to use an external knowledge base that contains sufficient information about the given query so that relevant terms for expanding the query can be obtained.

In this work, Wikipedia is used as the knowledge base for the query expansion process. With over three million articles in its English version, Wikipedia is a rich semi-structured resource. It covers a large number of articles (topics) describing people, places, landmarks, animals, and plants [53], which are very common in image search queries [2, 37]. Most of the articles also contain various representative images and associated textual captions that might represent valuable aspects of the queries. As such, Wikipedia can be used as an excellent source of information for the purposes query expansion in image search.

The structure of Wikipedia is utilized in the development of the core technique in this work. As such, before explaining the detailed methodology of query expansion, the organization of Wikipedia is briefly described in the remainder of the section.

Article Links

Wikipedia is structured as an interconnected network of articles. Each article can link to several Wikipedia articles. A contributor can insert a hyperlink between a word or phrase that occurs in the article being edited and any other Wikipedia article. If each Wikipedia article is denoted as a node, and each hyperlink between articles as an edge pointing from one node to another, then Wikipedia articles form a directed graph [66].

Category Links

In Wikipedia, each article belongs to at least one category (e.g., the article of "Volkswagen Beetle" belongs to different categories such as "Volkswagen vehicles", "Rear-engined vehicles ", "1980s automobiles", etc.). These categories can be further classified by associating them with one or more parent categories.

Redirect Links

Wikipedia guarantees that there is only one article for each concept by using *Redirect Pages* to link equivalent concepts to a preferred one. A redirect page exists for each alternative name that can be used to refer to a Wikipedia concept. For example, a Wikipedia article of "Washington, D.C." is redirected from the articles of "Washgìnton DC", "Washington D.C.", "District of Columbia", and so on.

Disambiguation Pages

In Wikipedia, disambiguation pages allow users to choose among several Wikipedia concepts for an ambiguous query. For instance, consider "Beetle" as an example, which can be denoted as an insect, a car, a comic character, and so on. In this thesis, all the concepts mentioned in disambiguation pages are referred as the different possible interpretations (or senses) of the original query.

3.3 Approach

The process of performing concept-based query expansion of image search queries follows three steps: extracting concepts from Wikipedia, ranking the extracted concepts, and generating the expanded queries. This overall process of query expansion is illustrated in Figure 3.1. In the first step, different possible interpretations are derived, and for each interpretation a number of candidate concepts representing various aspects of the query are discovered from within Wikipedia. In the following step, these concepts are ranked according to the semantic relatedness to the original query, and only a limited number of concepts are chosen. Finally, the selected concepts are used to generate expanded queries and retrieve a range of images that provides a broad view of what is available. The detailed description of the above mentioned steps are provided in the remainder of this section.

3.3.1 Extracting Concepts from Wikipedia

For this work, a dump of the Wikipedia collection [84] was obtained in June 2010, and was pre-processed to support the type of knowledge extraction required for query expansion. During this pre-processing step, Wikipedia collection was split into individual articles and stored in the concept knowledge base in order to process the information faster.

Matching a user-supplied query Q to this knowledge base is simply a matter of

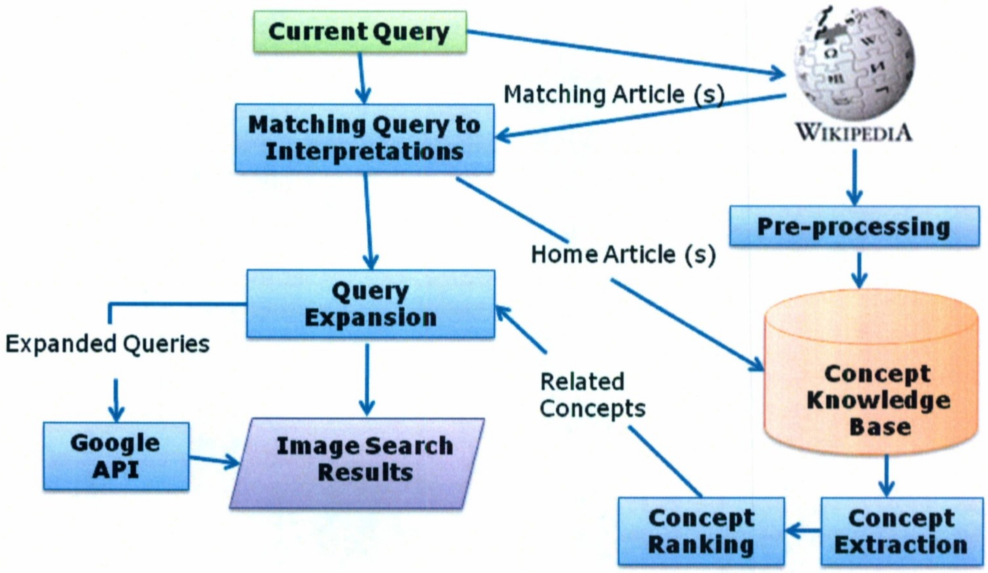


Figure 3.1: The steps in the concept-based query expansion process.

selecting all of the matching articles (referred to as the home articles) from Wikipedia using its search feature. In the case where the query is ambiguous and Wikipedia suggests multiple interpretations, a filtering step is performed. In this step, the *commonness* value of each interpretation is calculated based on the number of times it is referred by the other articles in Wikipedia [50]. At the end, the ones with higher commonness values ($commonness > 0.01$) are used as the home articles.

In analyzing Wikipedia, it has been observed that the in-going link articles (ones having links to a home article) and out-going link articles (ones to which a home article links) often provide meaningful information that is closely related to the concept of the home article, and hence the user-specified query. Therefore, for each article (concept) within the collection, these linked articles are located and their titles extracted as candidates for related concepts.

It also has been found that the captions surrounding the images present within a given article can often provide a valuable perspective on the visual features associated with the article's concept. To ensure the inclusion of all the concepts associated with the image captions, Wikifier [50] is used to augment the captions with links to relevant Wikipedia articles that may have been missed by the author of the article. These links are used to extract the concepts associated with the image captions.

The end result of this process is the selection of a set of home article(s) $\{h_s | 1 \leq s \leq q\}$ (for q interpretations of given query Q), along with a list of all the candidate articles C_{h_s} for each home article h_s based on the in-going links, out-going links, and image captions. These concepts provide the basis for the automatic query expansion process.

3.3.2 Ranking the Extracted Concepts

Due to the richness of Wikipedia, the number of concepts obtained in the process described above may become very large. While it is good that so much information is available for the query expansion process, there is a risk in expanding the query too broadly resulting in a significant negative impact on precision for a given interpretation of the query. In order to address this potential problem, the extracted concepts are ranked and only the top-N concepts that are most related to the home articles are chosen. Here, the value of N controls how many concepts are used in the query expansion process. The effect of this parameter is discussed in Section 3.5 and 6.2.

In order to promote diversification, these top-N concepts should be distributed among the candidate concepts extracted from each home article. However, the interpretations of the query (represented by each of the home articles h_s) may have different degrees of importance with respect to the original query. That is, a home article having rich amount of information with a large set of candidate articles is considered to be more important than those that do not contain many links. For example, the interpretations of the query "Beetle" are "Volkswagen Beetle", "Beetle (insect)", and "Beetle (comics)". Here, the number of candidate articles extracted for the interpretations Volkswagen Beetle and Beetle (insect) are much more than the candidate articles for Beetle (comics), indicating the difference in the importance of each of the interpretations. Considering this difference, the top-N concepts are non-uniformly distributed among the different interpretations based on the number of the candidate concepts $\{C_{h_s} | 1 \leq s \leq q\}$. As such, the number of related concepts N_{h_s} that are to be selected for a particular home article h_s is determined as follows:

$$N_{h_s} = \frac{|C_{h_s}| \times N}{\sum_{j=1}^{q} |C_{h_j}|} \tag{3.1}$$

Note that the sum of all N_{h_s} values equals N.

To select these N_{h_s} concepts for each home article h_s, it is necessary to rank the candidate concepts C_{h_s} based on their relevance to the home article h_s. Here, the approach to this problem is to measure the semantic relatedness between the home article and each of the candidate concepts. The purpose here is to ensure that the highly related concepts and the corresponding expanded queries remain focused to the interpretation that they belong to. WLM [49] is used to measure the semantic relatedness, taking advantage of the hyperlink structure of the associated articles to determine how much they share in common. For each of the candidate articles $c_i \in C_{h_s}$ extracted from the home article, WLM is applied between the home article h_s and the candidate articles. However, the concepts that have been extracted from the image captions within the home article can be more valuable comparing to the in-going and out-going links from the image retrieval perspective. As such, a re-weighting function is used to produce the relatedness score so that more importance can be provided to the concepts originating from captions:

$$r(c_i, h_s) = \min(WLM(c_i, h_s)(1 + \alpha_s), 1) \tag{3.2}$$

Since WLM provides a value in the range [0,1], the relatedness score is restricted to that range using the *min* function. The re-weighting factor α_s is provided according to the following function:

$$\alpha_s = \begin{cases} k\frac{|C_{h_s}|}{N_{h_s}} & \text{if concept } c_i \text{ originates from a caption} \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

Here, C_{h_s} and N_{h_s} are defined as above, and k is a system parameter that controls the importance of the concepts derived from the captions. In the prototype implementation $k = 0.01$. This results in a 10 - 20% increase in the score for the concepts derived from the captions, with proportionally more importance given when there are more concepts extracted from the home articles.

The outcome of this process is that the top-N concepts are selected from among the candidate articles, such that those from the image captions are given preference over those from the in-going links and out-going links of the home articles. Further, they are distributed across all of the interpretations of the query (as provided by matching the query to Wikipedia home articles). These concepts are used as the source for the query expansion.

3.3.3 Generating Expanded Queries

In order to ensure that the expanded queries remain focused on the topic of the query itself, the top-N related concepts $\{c_r | 0 \leq r \leq N\}$ are prepended with their associated home article h_r, resulting in queries of the form $< h_r, c_r >$. We define c_0 to be null and h_0 to the the original query Q, producing the original query plus N expanded queries.

Given that individual expanded queries have differing degrees of relevance to the original query, it is necessary to ensure that more images are retrieved for concepts that are most similar to the original query, even when the original query has multiple meanings. In this regard, others who have explored the use of query expansion in image retrieval have retrieved an equal number of images for each sub-query, and then ranked all the image by providing different degree of relevance to different query terms and their resultant images as a post-retrieval step [53, 77]. However, in this research rather than retrieving equal number of images for each query, different degree of relevance of the expanded queries were considered during the image retrieval step. As such, the number of images to retrieve for each expanded query is dynamically determined based on their relatedness score to the home articles according to the following formula:

$$I_r = \frac{r(c_r, h_s) \times I_t}{\sum_{k=0}^{N} r(c_k, h_s)} \tag{3.4}$$

Here, r is the same function provided in equation 3.2, and I_t is the total number of images to be retrieved by all of the queries. Since the null expanded query (c_0) is the original query, $r(c_0, h_s) = 1$ is used in the above calculation. Each query is sent to the Google AJAX Search service, and the desired number of images are retrieved. Duplicate images are detected and resolved based on the URL of the source image provided by the search engine. At the end of the process, for each of the images, a tag is assigned that represents the source concept (the concept that was used to retrieve this image). This concept information is stored along with each image, and used later for the purpose of image organization (Chapter 4).

3.4 Example

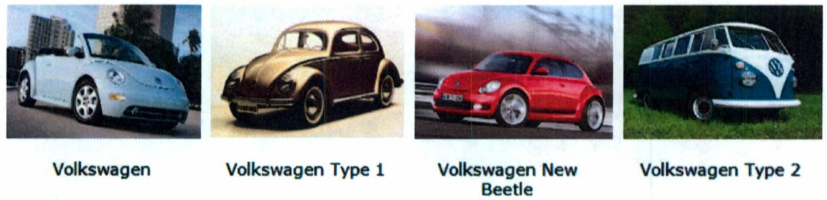
In order to illustrate the whole process of the concept-based query expansion, an example is provided here. When the searcher submits a query, the system first uses Wikipedia to find article(s) that match the given query. For example, if the user enters the query "Beetle", the system uses the query to perform a search in Wikipedia to find the multiple articles about "Beetle". Articles having commonness scores higher than the threshold value (0.01), namely "Volkswagen Beetle", "Beetle (insect)", and "Beetle (comics)" are then selected as the home articles, representing different interpretations of the query.

For each home article, all the articles that are from the in-going links, out-going links, and captions are extracted from Wikipedia. Due to the richly interconnected nature of the Wikipedia articles that are associated with the chosen home articles, a large number of concepts are found for the query "Beetle". In fact, 923 different concepts are extracted from Wikipedia, which are regarded as the candidates for relevant concepts to be used for query expansion.

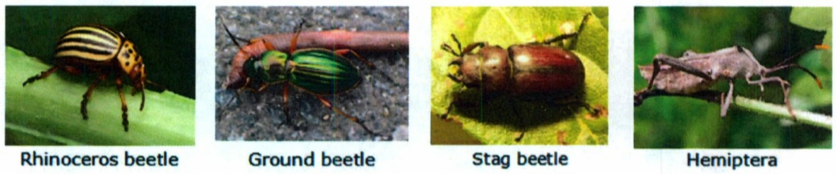
These candidate concepts extracted for the query "Beetle" are ranked and only the top-N concepts that are most related to the home articles are chosen. The top-N concepts are non-uniformly allocated to different interpretations based on the number of candidate concepts extracted for each home article. As a result, the number of concepts to be used within the query expansion process for some interpretations can be much more than the others that do not have many candidate concepts. In this case, Volkswagen Beetle and Beetle (insect) articles contain many more links than the article for Beetle (comics), leading to an uneven distribution of concepts among the interpretations.

In order to select the top-N related concepts, the concepts are ranked based on a decreasing order of semantic relatedness. For this purpose, the semantic relatedness score between each extracted concept and its home article is computed using WLM. At the end of this process, the relevant concepts that are to be used in the query expansion are selected. In this example, the value of N is set to 9, which leads to the selection of 9 different concepts for generating the sub-queries (Section 6.2 will describe how this value is chosen). Among them, four concepts are selected for Volkswagen Beetle and another four concepts are selected for Beetle (insect). Since Beetle (insect) has quite less number of links comparing to the other two interpretations, only one concept is selected for this interpretation.

In the query expansion procedure, the sub-queries are generated by combining each concept and its corresponding interpretations. The queries are sent to the Google AJAX Search service, and the result set is retrieved. In the current prototype im-



(a) Concepts for Volkswagen Beetle



(b) Concepts for Beetle (insect)

(c) Concepts for Beetle (comics)

Figure 3.2: List of the selected concepts and their representative images for the query "Beetle".

plementation, the system is configured to retrieve a total of 300 images, dynamically distributed across all of the expanded queries based on their relatedness to their home article (such that those that are more related retrieve more images).

Figure 3.2 shows the concepts that are chosen for query expansion along with representative images. As can be seen, most of the relevant concepts selected for Volkswagen Beetle are mainly related models of cars manufactured by Volkswagen. Similarly, for the meaning Beetle as insect, different types of beetle insects are selected. The overview of the search results, as provided in Figure 3.2, demonstrates the enhanced diversity in the search results by employing the query expansion.

3.5 Discussion

In this chapter, an approach for performing concept-based query expansion has been described, wherein the goal is to produce a diversified set of image search results. In particular, the approach uses knowledge contained within Wikipedia to extract concepts related to the source query, using these concepts to expand the query and retrieve a diverse range of images.

In Section 2.2, it was identified that finding a suitable knowledge base for the query expansion process is challenging, particularly in the domain of image retrieval. In this research, the system takes advantage of the rich structure of Wikipedia to extract relevant information to be used in the process. While others have used Wikipedia for query expansion in the context of general Web search [51], the approach described in this chapter is novel in that it takes advantage of specific aspects of image search by providing more weight to the concepts originated from the article's captions. It can provide sufficient conceptual information regarding the query, so that the expanded queries can be broader than the original query. The expectation is that when such a broad range of interpretations of the query are used to produce the search results, then at least one of them would most likely satisfy the user.

However, as noted in Section 2.2, there is a danger in broadening the query too much, resulting in a potentially significant decrease in precision. To address this, it is necessary to control how many concepts should be used in the query expansion process by setting the appropriate value of N. Here, the value of N is an explicit indicator of the degree of diversification, which is defined as the diversification parameter. With a smaller value of N, fewer concepts will be used, and the search results will remain more focused. If N is increased, then more concepts will be used for query expansion, and in turn the search results will be more diversified covering more aspects pertaining to the query.

The fundamental trade-off against precision is based on the fact that as N is increased, there is a higher chance that a concept will be selected for the query expansion process that is not relevant to the searcher's information needs, resulting the inclusion of images that are not relevant. Hence, to fulfil the diversification objective, the goal is to increase N without sacrificing too much on precision. How to select the suitable value for N is studied in detail in Section 6.2. As a result of this study, an automatic method for tuning the diversification parameter is suggested based on the degree of ambiguity of the original query.

Another important issue that can affect precision is the method by which the extracted concepts are ranked, since the relevance of the top-N concepts to the given query have a direct impact on precision. In Section 2.2, different methods for measuring the semantic relatedness were mentioned. Among them, WLM [49] is chosen in this work over some other methods [75, 21] for the purpose of measuring the semantic relatedness of the concepts to the query. It uses the hyperlink structure of Wikipedia to measure the semantic relatedness, which is computationally more efficient and accurate than the other two methods. However, besides sematic relatedness, some other parameters (such as popularity of the concept) can be considered for the purpose of ranking concepts. For example, van Zwol et al. ranked the facets based on a statistical analysis of image search query logs, and the tagging behaviour of users annotating photos in Flickr [80]. If such valuable resources are utilized effectively, they may provide an alternative way to select the concepts to be used in the query expansion process.

The process of retrieving concepts from Wikipedia, as described in this chapter, is robust in nature, although it may be vulnerable to potential vandalism. In some cases it may be possible to alter or vandalize one or few articles. However, it is much less likely that a large number of extracted concepts (in-going and out-going links) will be changed drastically for many articles. As such, the concepts to be chosen for expanding the query as well as the generated image search results will not be significantly affected by any potential vandalism.

The performance of the query expansion process in terms of speed may vary with respect to different queries. For a broad query, it may not possible to generate the image search results in real time according to the computational resources and Internet facilities of a personal computer. However, for different topics found in Wikipedia, the interpretations and concept information can be easily extracted and pre-processed before the query expansion is performed. This will allow the query expansion process to become more efficient. As such, when implemented in a search engine, generating expanded search results can be possible in real time.

One of the important design decisions made within the approach is choosing the query expansion process as automatic over interactive. As noted in Chapter 2, interactive query expansion may have some value regarding document search [20]. However, it forces the searcher to choose the concepts for expanding the query, which might be cumbersome for image search. Instead, automatic query expansion allows the image search results to be immediately available for the searcher, allowing them to explore the images quickly.

When coupled with a visual interface that allows searchers to explore the images from each concept, this approach can actively support to find images that they are seeking for. In the next chapter, such a user interface along with the interactive features for exploring the image search space will be introduced.

Chapter 4

Organization and Exploration of Image Search Results

4.1 Motivation

Although applying concept-based query expansion is an explicit way of diversifying search results, it introduces some new challenges. The difficulty with retrieving a broad and diverse range of images is how to present them in such a way that the searcher can focus on the specific aspects of the query in which they are interested. A naïve approach would be to use a traditional paged grid layout of the images, ordered by the rank in the search results list, and perhaps the semantic relatedness between the top-N concepts and the home articles. However, such an approach may not be effective in supporting image search tasks, as the meaning of the organization of the images may be rather obscure. Since the expanded query produces a broad range of images representing the various interpretations of the query, it is necessary to provide an interface that can allow the searcher to easily find the aspect of their query that matches their search intentions. A visual method for organizing the images retrieved is well suited to this task.

As noted in Chapter 2, there is merit to 2D similarity-based organization, since it can improve the searcher's performance compared to traditional list-based methods for organizing image search results [73]. Most of the previous works on SBIB arrange images based on visual similarity only [25, 54, 56, 70]. However, many Web image search queries are associated with conceptual domains that include nouns, people's names, and locations [2]. It is advantageous for the search results of such queries to be diverse in nature, both from different interpretations of the query as well as the visual features. In situations such as this, it would be more beneficial to consider not only visual similarities but also conceptual relatedness between images in organizing the results. The goal is to place images from related concepts in nearby locations, instead of scattered throughout the whole image space. This will allow the searcher to focus on a specific area to explore conceptually related images.

An additional issue with the SBIB approach is that when the image collection is very large, it becomes impractical to display all the images without any occlusion on the limited screen space. A promising approach to deal with this issue could be to present only the representative images initially, while keeping the other images hidden. Then, the searcher will be allowed to zoom into an area, so that the additional space created between images can be used to show the images that were previously hidden [73].

While images can be regarded as the primary data for organizing and visualizing the search results, the concepts and interpretations derived from the query expansion process might also provide valuable means in exploring the image set. Since, each image in the search results is associated with a concept that resulted in its retrieval, this information can be effectively use by organizing the concepts in a meaningful way. As found in some previous research works, presenting the list of related facets or concepts of the given query and utilizing them in supporting the image set exploration can be useful [48, 86]. By designing appropriate exploration features based on the available concepts, the searcher can be facilitated to explore the results at a more semantic level.

In this chapter, CBIS, a novel image search interface is introduced for organizing and exploring of image search results. The organization method is based on a previous SBIB approach proposed by Strong and Gong [70, 71]. However, unlike all the previous works, this thesis introduce conceptual features along with the visual features in the organization process. A new set of interactive exploration features are also developed to support the searcher in finding images.

CBIS takes the image search results along with the top-N related concepts that produce those images (through the query expansion process described in the previous chapter), and then organizes images within the image space, so that they are grouped together, based on their conceptual and visual similarities. At the same time, the top-N related concepts selected for query expansion are also presented to the searcher in a hierarchical manner based on an ontology. The resulting interface is highly interactive, allowing the searcher to dynamically highlight and filter images based on the concepts, and zoom into an area within the image space to show additional images that are more conceptually and visually similar.

4.2 Approach

The overall approach for organizing and exploring image search results is illustrated in Figure 4.1. CBIS takes a set of images I_t and the top-N related concepts as input, where each image is associated with the concept from the query expansion process that resulted in its retrieval. Then, the relations between the concepts are used to derive a conceptual feature vector for each image, which is used in conjunction with the visual feature vector extracted from the image to form a hybrid feature vector. A multi-resolution Self-Organizing Map (SOM) based approach is then used to map images within the image space, so that the ones with similar hybrid features are placed close to each other. The searcher can then visually explore the search results using the following interactive features: concept-based focusing and filtering, which allows highlighting and filtering of semantically similar images; zooming into an area of interest, to unveil more conceptually and visually similar images by moving images that are distant from the focal point out of viewport; and panning, which allows the searcher to move within the image space. Detailed descriptions of the approach are provided in the remainder of this section.

4.2.1 Organization of Image Search Results

4.2.1.1 Hybrid Feature Generation

SOM is a type of artificial neural network that is trained using unsupervised learning to produce a low-dimensional representation of the high-dimensional input space of the training samples [44]. To use it for image organization, it is necessary to represent all the images using feature vectors. Different ways for generating feature vectors from

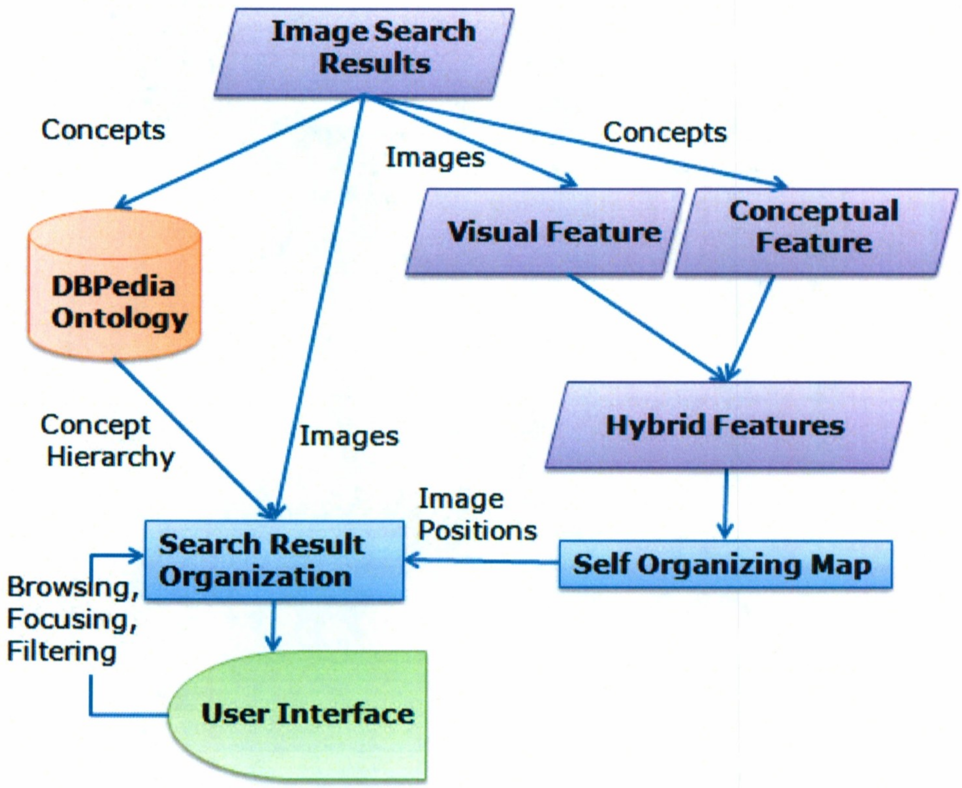


Figure 4.1: The steps in organization and exploration of image search results.

visual information and their performances have been studied previously [71]. While these types of feature vectors can be used to organize images based on color and/or shape similarities, they cannot group conceptually related, but visually different, images together. To address this problem, a hybrid feature vector is proposed.

The hybrid feature vector for an image I contains two portions: a conceptual feature vector $C(I)$, determined using the concept tag c_r carried by the image I; and a visual feature vector $V(I)$, extracted from the color features. For the visual portion $V(I)$, a 64 dimensional colour-gradient correlation vector is used since it is efficient to calculate, and offers improved organization results relative to comparable feature extraction methods, such as the colour histogram [71].

Extracting conceptual feature vectors, on the other hand, is not as straightforward. The main difficulty is that images themselves do not provide any useful conceptual information; all the information available are the associated concept tags with images. As such, instead of computing for each individual image, conceptual feature vector is computed for its associated concept. To simplify the problem, it is assumed that different images retrieved carrying the same concept tag are conceptually the same; therefore, they have the same conceptual feature vector.

It is also difficult to convert concepts into feature vectors directly. The only available information that can be extracted from the concepts is the pairwise semantic distances between all the concepts. By converting these distances into the feature vector, it will be possible to map the images from the semantically related concepts to be placed nearer to each other. As such, to encode the semantic relatedness between concepts, an $N \times N$ semantic distance matrix is computed for the top-N concepts using equation 4.1:

$$D = \begin{bmatrix} 0 & 1 - r(c_0, c_1) & ... & 1 - r(c_0, c_N) \\ 1 - r(c_1, c_0) & 0 & ... & 1 - r(c_1, c_N) \\ \cdots & \cdots & \cdots & \cdots \\ 1 - r(c_N, c_0) & 1 - r(c_N, c_1) & ... & 0 \end{bmatrix} \tag{4.1}$$

In this equation, the relatedness between any two given concepts, $r(c_i, c_j)$ is computed using WLM [49]. Here, by definition, $r(c_i, c_i) = 1$, and $r(c_i, c_j) = r(c_j, c_i)$. Hence, this matrix is symmetric.

The above matrix encodes the relatedness information among different concepts, which is then used to generate a set of vectors $C_1, ..., C_N$, having m-dimensions. The vectors are needed to model the relatedness information as closely as possible (i.e., the distance between any two vectors is approximate to, if not the same as, the semantic distance between the corresponding concepts). This is the same as minimizing the following least-squares function:

$$C_1, ..., C_N = argmin_{C_1,....,C_N} \sum_{1 \leq j,k \leq N} (\|C_j - C_k\| - D_{j,k})^2 \tag{4.2}$$

where $\|C_j - C_k\|$ is the Euclidean distance between the two vectors, and $D_{j,k}$ is the relatedness score between the two corresponding concepts as encoded in equation 4.1.

As a result, the task of finding a set of vectors $C_1, ..., C_N$ based on a given distance matrix D can be easily solved by classical multi-dimensional scaling [9]. The dimensions of vector C are chosen as $m = 4$. Given the fact that the number of concepts used in the query expansion will be rather limited (as a result of setting the value of N according to the method described in Section 6.2), it is expected that 4-dimensional feature vector will be sufficient to encode distance between any two given concepts with reasonable accuracy. For example, even for a highly ambiguous query having four different interpretations the value of N will be 12, resulting in 12 different feature vectors $C_1, ..., C_{12}$, which can be easily encoded by 4-dimensional vectors.

As noted previously, it is assumed that the images carrying the same concept tag are assigned with the same conceptual feature vector. That is, the feature vector C(I) for an individual image I carrying the concept tag c_r is assigned with the feature vector C_r. In the end, the hybrid feature vector $H(I)$ for a given image I is formed as $< C(I), V(I) >$, where c_r is the concept used to retrieve image I. Since the conceptual portion has $m = 4$ dimensions and the visual part has $n = 64$ dimensions, the total dimensions of a hybrid feature vector is 68.

4.2.1.2 Image Organization Using Self-Organizing Map

The method of image organization using SOM is designed based on a previous work by Strong and Gong [70, 71]. Within this approach, the SOM consists of a 2D lattice of interconnected nodes or units, each of which has a weight vector that is of the same dimension as those extracted from the images. Initially, the weight vectors are assigned random values.

The SOM training process requires multiple iterations. During each iteration, all the hybrid feature vectors are exposed to the SOM in a random order. When a particular feature vector is exposed to the SOM for training, it finds the Best Match Unit (BMU) that is closest to that vector in terms of Euclidean distance. The SOM then updates all of the units in the BMU's neighbourhood to be similar to the image vector. The amount that the image vector influences the units within the neighbourhood varies with the distance from the BMU. Over time, the neighbourhood shrinks to one unit, forcing the training to converge. At this point relative positions of the images in 2D can be determined by using the positions of their BMUs in the map.

The topology-favouring property of the SOM ensures that images with similar feature vectors are mapped to units within the SOM that are closer to each other, and vice versa. One thing to note is that if the SOM is too small with respect to the number of images to be organized, multiple images may be mapped to the same unit. To avoid this, it has been ensured that there are approximately three times as many SOM units as there are images. This gives enough room for the images to find unique positions within the SOM.

Due to the large difference between the size of the SOM and the number of images, using the positions from the SOM directly in the organization of the images produces a layout that follows the messy-desk metaphor. It has been observed that using this layout, searchers sometimes experienced difficulties scanning the irregularly placed images within the display. A potential solution to this difficulty is to align the images in the messy-desk arrangement into a more structured neat-desk layout [73]. This alternative layout provides less of a departure from what searchers expect in the presentation of image search results, while still maintaining a visual encoding of the degree of similarity. As such, neat-desk layout was preferred for this work. In order to align the images to a structured grid, a kd-tree [6] was generated using the positions provided by the SOM.

At this stage of the organization, each image has a position on a grid that makes sense based on its similarity to its neighbours. However, very often there will be more images retrieved than can be shown at a reasonable size within the available display space. For example, if three hundred images are retrieved, and each of them requires at least 100×100 pixels, only total of 84 images (7×12) can be shown on a screen having resolution of 1280×768 pixels (with minimal spaces between the images). To address this problem, a priority order for the images is provided. The priority order is obtained by generating multiple grids at different resolutions. This multi-resolution structure is constructed in a bottom-up approach from the grid of images generated by the SOM, whereby the dimensions of the grid are repeatedly divided in half. In each step, groups of 2×2 adjacent images become mapped to the same location in the higher-level grid. In order to assign the priority order which determines which image to show at the given grid resolution, the average feature vector of the images for each group is generated, and the image whose feature vector is closest to the average is chosen to be most representative of this space.

As a result, only images with high display priority are shown, when there is insufficient space to display all images. The amount of space available is relative to the screen resolution, as well as two interactive parameters: zoom level and image size. As the searcher zooms into a particular area of interest, the viewport becomes smaller, more space is made available available between the images, and images with a lower display priority can be shown. This zoom operation, and the associated panning, allow the searcher to identify an area in the search results that matches the query intent, and then zoom to view other similar images that are hidden below. Once the searcher has zoomed into a level that shows all of the images that are available, further zooming increases the sizes of the images. As a result, there is a smooth zoom effect that moves from creating more space to show more images, to showing more image detail once all of the images are visible.

In addition to aiding in the arrangement of images from the search results, the approach also uses the top-N related concepts from which the expanded queries were derived to support focusing and filtering operations. Each of these concepts is mapped to an ontology using DBPedia [8]. According to the mapped ontology, concepts having the same type are grouped together under that corresponding type (e.g., people, location, company, automobile, etc.). Then, the organization of concepts is displayed to the searcher in a hierarchical manner. Within the concept hierarchy, the query itself represents the root node, while each interpretation of the query are placed as the child of the root nodes. At the next level, concepts are grouped by their types under the interpretation that they belong to. Images retrieved for a particular concepts are connected as the children of that concept. All the images available in the image space are represented as leaves of the concept hierarchy. The concept hierarchy is designed to support the searcher in exploring the images as described in the next section.

4.2.2 Interactive Search Results Exploration

The query expansion process described in the previous chapter may result in the introduction of many non-relevant images within the search results, especially for ambiguous queries where there are multiple interpretations of the searcher's intent. As such, it is important to provide a mechanism by which the searcher can easily narrow down the search results to those that match their interests. This work allows the searcher to do this in two complimentary ways: they may use a concept hierarchy to focus and filter the search results, or they may perform visual filtering through pan and zoom operations.

4.2.2.1 Concept-Based Focusing and Filtering

The searcher can use the hierarchy of concepts for both focusing and filtering operations. By clicking on any of the concepts, all the images that were retrieved as a result of this concept are temporarily assigned a high display priority, bringing them to the foreground of the image organization; the remaining images are dimmed giving the focused images more visual prominence. Similarly, by clicking on an interpretation, all the images that were connected to the concepts under that interpretations are highlighted. An example of this is shown in Figure 4.2.

In addition, the searcher can use checkboxes associated with each node in the concept hierarchy to filter the search results, removing the associated images from the display. This feature allows the searcher to quickly scan the names of the concepts, removing those that are obviously not relevant to their meaning for the query.

Together, these two features can allow the searcher to quickly inspect a particular interpretation or concept. That is, clicking on it will focus the display on the images that were retrieved as a result of this interpretations or concept. The searcher can then decide whether to keep it as part of the search results set or remove it. This interactivity allows searchers to easily explore the search results, based on the Wikipedia concepts that generated the expanded queries.

4.2.2.2 Visual Focusing and Filtering

Since the images comprising the search results are organized based on similarity using a multi-resolution SOM, a logical method for exploring within this image space is through pan and zoom operations. As the searcher zooms into an area of interest

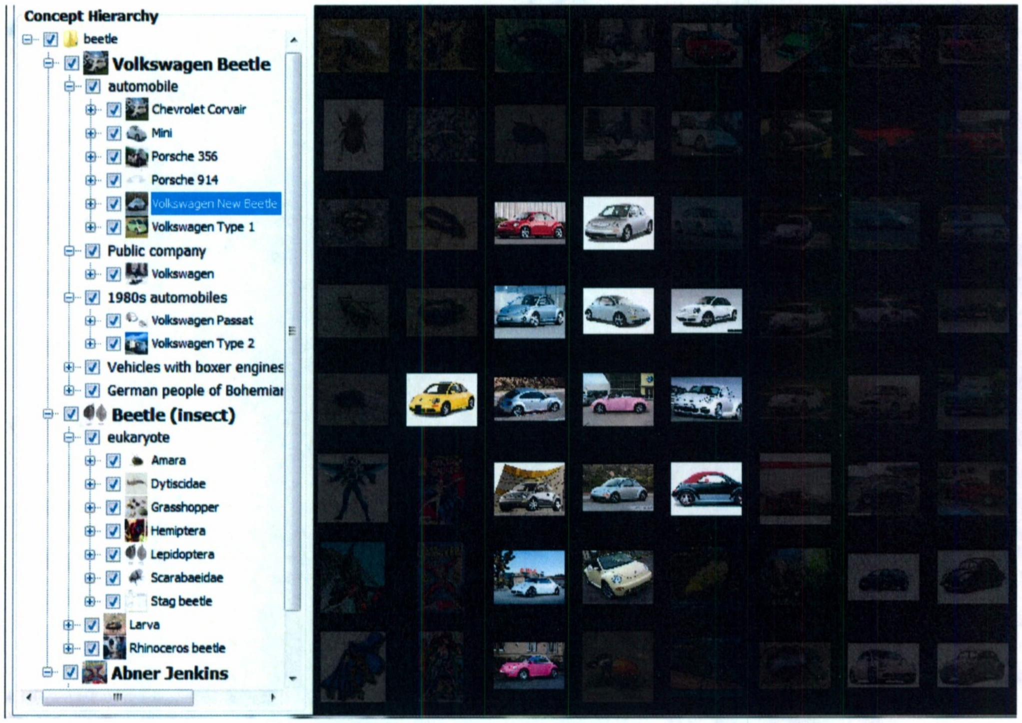


Figure 4.2: Concept-based focusing allows searchers to inspect the images that were retrieved as a result the concepts derived from Wikipedia for the query. In this particular example using the query "Beetle", once the searcher selects "Volkswagen New Beetle", images from this concept are given a high display priority and are pulled to the front; all others are dimmed.

and more space is created between the images that are visible, additional images that were previously hidden are shown. At the same time, images that are distant from the focal point of the zoom operation are pushed out of the viewport. A number of steps in the zoom operation are illustrated in Figure 4.3.

This operation can be viewed as a both a focusing operation (where hidden images that are in the vicinity of the focal point are shown) and a filtering operation (where

(a) Top-level zoom. (b) Mid-level zoom. (c) Bottom-level zoom.

Figure 4.3: For search results retrieved using the query "Beetle", as the searcher zooms into the region that contains automobiles, more images of "Volkswagen Beetle" are shown and images that are relevant to other interpretations are pushed out of the viewport. The red box in (a) and (b) shows the zoom areas in (b) and (c), respectively.

images that are not of interest to the searcher are removed). Searchers are able to take advantage of their ability to easily scan the visual features of the images to identify those that look like what they are seeking. By zooming into these regions, the searchers can then explore other images that look similar but were previously hidden due to space constraints of the display.

4.3 Example

In order to illustrate the benefits the system provides to the users as they seek to interactively explore the image search results, an example is provided here. Suppose the searcher starts with the query "Beetle", although the intention was to find images from "Volkswagen Beetle car". As a consequence, the query is expanded and a broad range of search results are retrieved.

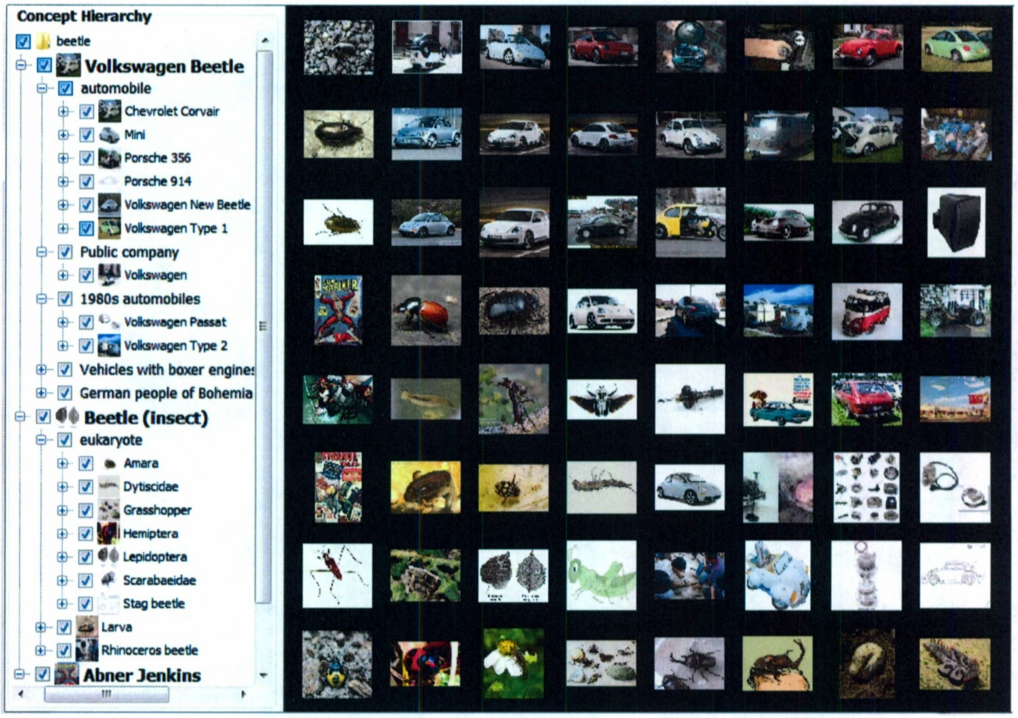


Figure 4.4: The screenshot showing the image search results of the query "Beetle". Here, the concept hierarchy organizes all the related concepts of the query into a tree view, and the image space organizes images based on their conceptual and visual similarity (in the right side).

Then, the image search results are organized and presented to the searcher (see Figure 4.4). Here, the concept hierarchy presents the overview of the top-N related concepts as a tree. The image space presents search results based on their conceptual and visual similarity using SOM. As a result, images from similar concepts are grouped together (e.g., images from "Volkswagen Beetle cars" are located in the top-right side of the image space).

The query "Beetle" is ambiguous, and so the search results contain images ranging from different topics such as insects, comic characters, and cars. The organization of the images in Figure 4.4 shows that using conceptual information in addition to visual information not only groups the images of the same concept together, but also places images from related concepts at nearby locations. Hence, once users identify the images of interest, they can easily zoom into the area to find more conceptually related images. As they zoom further into the area of interest, the visual similarities between the images at the neighbouring locations becomes apparent to the searcher.

The searcher can quickly browse images by using the concept-based focusing operation. Since the interpretations and the top-N related concepts for the query are presented as the nodes of the concept hierarchy, the searcher can select either an interpretation or a concept from the tree that invokes the focusing operation. Figure 4.5 shows the results of conceptual focusing operation as the searcher selects the interpretation "Volkswagen Beetle" from the concept hierarchy. This causes images related Beetle cars to be pulled in front of the display and other images are dimmed, so that the searcher can easily find the area, where images belonging to Beetle are located. It is clear from Figure 4.5 that all images related to "Volkswagen Beetle" are grouped together due to the organization of images based on both conceptual and visual similarity. As such, the searcher can easily identify the area of interest within the image space.

As the searcher performs the focusing operation, images from other interpretations get dimmed but still remain within the image space. The searcher can filter them from the display by unchecking the checkboxes in the concept hierarchy. Figure 4.6 shows the results of the concept-based filtering operation, as the searcher unchecked the concepts from Beetle (insect). It can be seen that as a result of the filtering

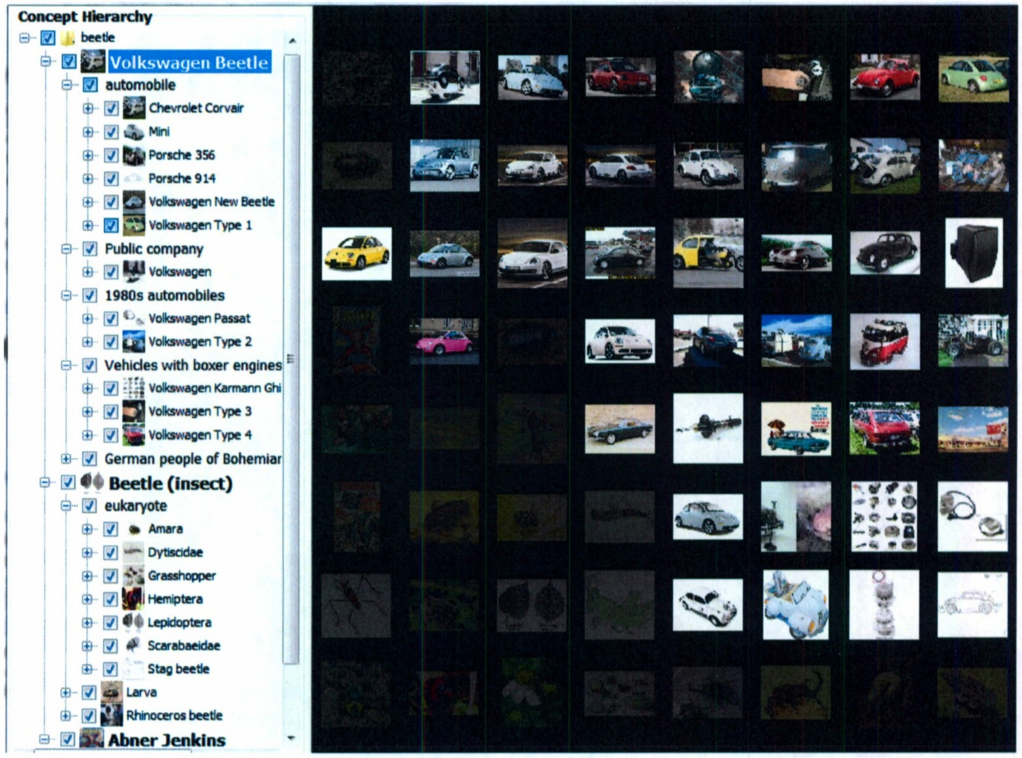


Figure 4.5: The screenshot showing the results for the query "Beetle" after performing concept-based focusing operation. Here, as the searcher selects "Volkswagen Beetle", images from this interpretation are highlighted.

operation, all the images from this interpretation are removed from the display. As a result, it can be seen that some gaps are created, where these images were located. This gaps also illustrate the ability of the organization method to group conceptually similar images.

At this point, the searcher has found the area of interest within the image space (through conceptual focusing operation) and eliminated the images that were irrelevant (through the filtering operation). Now, the searcher can zoom into the area

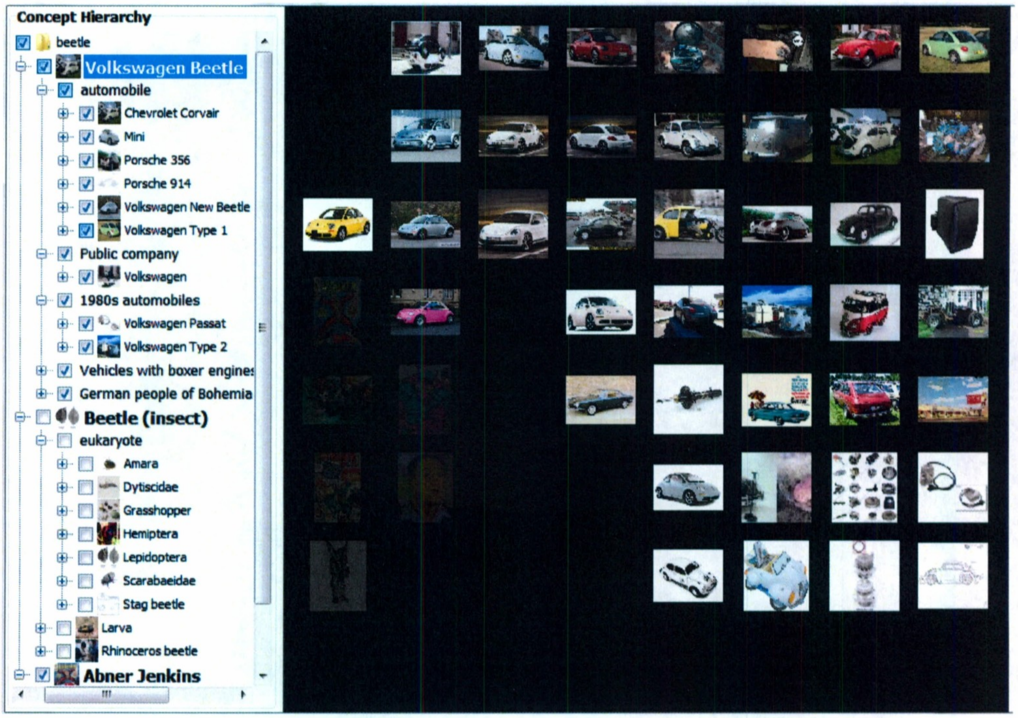


Figure 4.6: The screenshot showing the results for the query "Beetle" after performing filtering operation. Here, as the searcher unchecked "Beetle (insect)", related images are removed from the display.

of Volkswagen Beetle images, and pan around different areas of canvas to examine the images of interest at the deepest zoom level. Figure 4.7 shows such a screenshot, where the searcher has reached the bottom level zoom, and the canvas contains the images mainly from "Volkswagen Beetle". At this zoom level, the searcher can not only find all the images from similar concepts grouped together, but also see the visually similar images having closer proximity (e.g., blue cars are mainly placed on the middle portion of the screen). Finally, searchers can examine each individual image that seems relevant to their image needs. Note that it is possible to zoom out to the

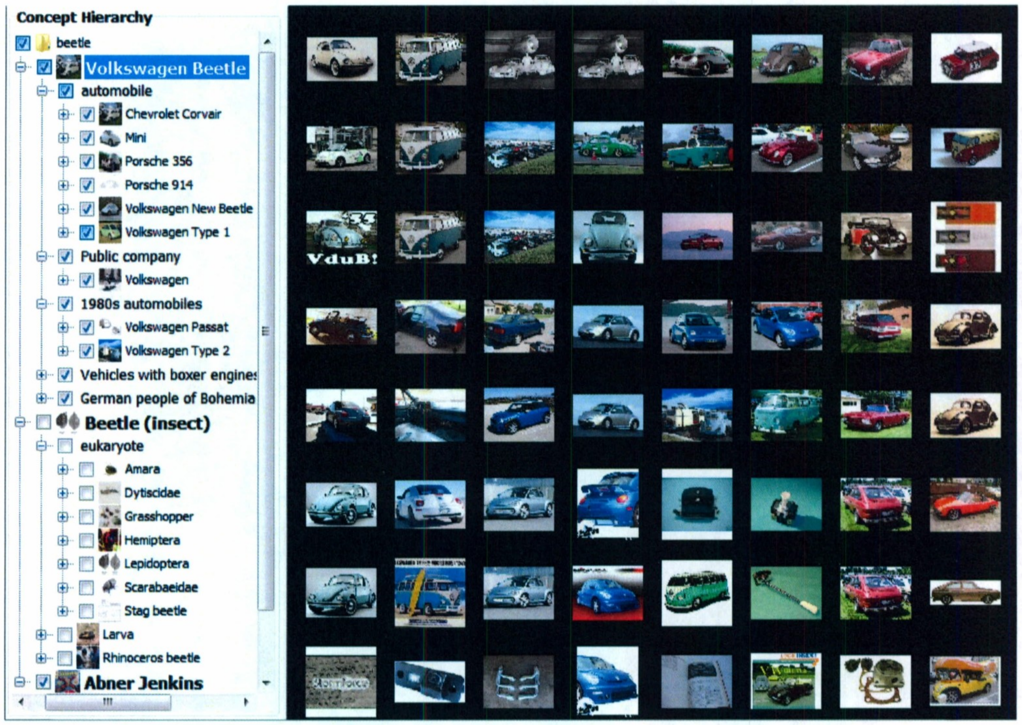


Figure 4.7: The screenshot showing the image search results for the query "Beetle", as the searcher found an area of interest and zoomed in to that region.

initial display at any point to once again see the overview of the image space.

4.4 Discussion

As noted previously, query expansion has the side effect of introducing potentially irrelevant images within the search results, especially when a query is ambiguous and can be interpreted in multiple different ways. To address this problem, CBIS allows searchers to perform conceptual filtering and focusing operations using a hierarchical representation of the concepts. In addition, the images themselves are organized using a multi-resolution grid layout derived from a SOM, which not only groups visually and conceptually similar images, but also provides a solution to the problem when there are more images to show than display space allows. Zooming into the image space results in a visual focusing and filtering operation, displaying more images as spaces between the images are created, and moving images that are distant from the focal point out of the viewport.

Although the illustrative examples presented in this chapter demonstrate the potential benefits of the approach, evaluations should be conducted to validate the interactive features. In this regard, a detailed evaluation is conducted to measure the changes in precision of the image search results as a result of concept-based focusing and filtering, as well as zooming operations (Section 6.3). In addition, the potential benefits provided by the interactive features for exploring images is evaluated through a user study (Section 6.4).

The approach for organizing Web image search results, as described in this chapter, is motivated by a similarity-based image organization method proposed by Strong and Gong [70, 71]. However, a key difference between the approach described in this thesis and all existing SBIB techniques is that here images are organized not only by visual information, but through the extraction and utilization of concept information obtained during the query expansion process. The benefit of such organization is that it groups images from the same or related concepts together, while simultaneously grouping visually similar objects, allowing users to explore particular areas of interest based on both conceptual and visual features of the images. This visualization technique can be particularly helpful for dealing with ambiguous queries, by separating images of different concepts from each other into their own areas. The benefits of incorporating conceptual information in addition to visual information will be studied through the experiments described in Section 6.4.

The approach used here places equal weight on the concepts and the visual features during the training process of SOM. This resulted in reasonably good image organization results. However, it might be interesting to experiment with different weighting schemes. In some cases, a searcher may be more interested in visual similarity first; in others conceptual similarity first. Dynamically adjusting the weights on the vectors may be a useful feature for the user to control (or for the the system to automatically set based on features of the query or the search results set).

Previously, two alternative canvas layouts were proposed for organizing the images based solely on their similarity (following a messy-desk metaphor), and in a more structured layout (following a neat-desk metaphor) [73]. This thesis uses a grid-based layout that follows the neat-desk metaphor, since it provides a familiar interface while visually encoding the degree of similarity between images and avoiding the overlapping of images found in the messy-desk metaphor. However, messy-desk layouts can reveal the grouping or clustering of images more clearly since they do not shift the original SOM positions of images. As such, further investigation to choose between the messy or neat desk metaphors might be worthwhile.

The approach of organization described here moves beyond the traditional linear browsing of ordered images into a nonlinear browsing paradigm. In order to deal with insufficient space, it introduces display priority to determine the images to be shown at the current stage of SOM. Currently, such display priority is computed based on the feature vector of the images. It might be interesting to involve the rank produced in the results by Google, so that images having a popular rank would get a higher display priority.

As noted in Chapter 2, a fundamental issue with organizing images in general is that organizing of hundreds of images using high dimensional feature vectors is computationally intensive [82]. This remains the case for organizing images using a SOM. As the size of the feature vectors and the number of images grows, the performance of SOM degrades due to the dramatic increase in computational cost. In this context, some researchers computed the organization results for some popular queries prior for experiment purpose. For example, in work that was a pre-cursor to Google Image Swirl, the image search results were pre-organized (for 2000 different popular queries), for the purpose of user evaluation [38]. Similarly, in this work CBIS organizes the image search results using SOM, and cached such organization information (SOM positions) prior to their use by the searcher, to ensure the browsing activity is possible in real time. However, the study of other alternative organization methods to SOM, and enhancing performance further using Graphics Hardware (GPU) [70, 72] should be investigated so that time need for training SOM can be reduced to near-real time.

Chapter 5

Concept-Based Interactive Query Refinement

5.1 Motivation

Traditional Web image retrieval mainly consists of two steps: query formulation, followed by linear evaluation of the search results. Within this process, the search engine relies upon the assumption that the users would be able to formulate the query accurately according to their information needs. When this assumption is not valid, the search results may not satisfy the users. In that case, searchers will be responsible for manually refining the query. As noted in Chapter 2, the majority of query sessions for image search comprise a trail of *related* queries, suggesting the fact that query refinement tasks are more frequent in image search than document search [2].

Searchers commonly have difficulties in refining their queries manually. They are required to describe something that is abstract (an idea of an image) with keywords, which is quite challenging. Furthermore, they can get sidetracked from the original information need when they find some images of interests but not the ones they intended to look for initially. This phenomena described as *serendipity*, or the act of unexpectedly encountering something fortunate, has long been identified as valuable [3]. Serendipitous discoveries might be useful from different perspectives of information seeking: reinforcing an existing problem or solution or taking it in a new direction, identifying information relevant to a latent goal, or just finding information of interest.

According to the approaches described in the previous two chapters, searchers can start with a rather broad query, and then through the interactive exploration features, identify specific concepts and images that are of particular interest. In some cases, this might lead users to become interested in some related topics which they were unaware of. For example, searchers may start with an intention to find Beetle cars. However, as they are provided with some other car models in the concept hierarchy, such as *Porsche*, they might suddenly become interested in Porsche cars.

In this scenario, it might be useful to interactively support one or more query refinement loops, with the goal of creating a query that is an accurate representation of the desired information needs. Unlike traditional models of image search, these interactive steps might help the searcher to enhance their original query, or become interested about a new topic. Ultimately, this can provide a mechanism by which the searchers can more accurately describe to the search engine what it is they are looking for. Furthermore, it can also be used to allow the focus of the search task to be easily changed based on serendipitous discoveries.

In this chapter, two different methods for interactive query refinement are proposed. Used together with the exploration features of CBIS (as discussed in Chapter 4), these operations empower the searcher to take an active role in the image retrieval process, supporting their ability to refine their image needs as they explore the image search space.

5.2 Approach

The overall approach developed for interactive query refinement is illustrated in Figure 5.1. As the searcher initializes a query, the system performs concept-based query expansion to retrieve images, followed by organizing them in the methods previously explained. Searchers can then explore the search results within the image space, which will allow them to find images and concepts that are of interest. At this point, searchers can interactively refine their image needs in two different ways: either by choosing a concept, or selecting example images. Selecting a concept for query refinement will cause the system to run it as a new refined query, so that this concept might be expanded to retrieve results as broad as the original query. On the other hand, query refinement can also be performed via example images by taking advantage of the searcher's ability to find topics of interest from the images. When the searcher finds images from the search results that look interesting, they can be added to the query panel. Then, by using the linked concepts with these images as the sub-queries (but not expanding them further), a new refined set of search results can be produced that are semantically relevant to the chosen images. The detailed description of these two interaction techniques are provided in the remainder of this section.

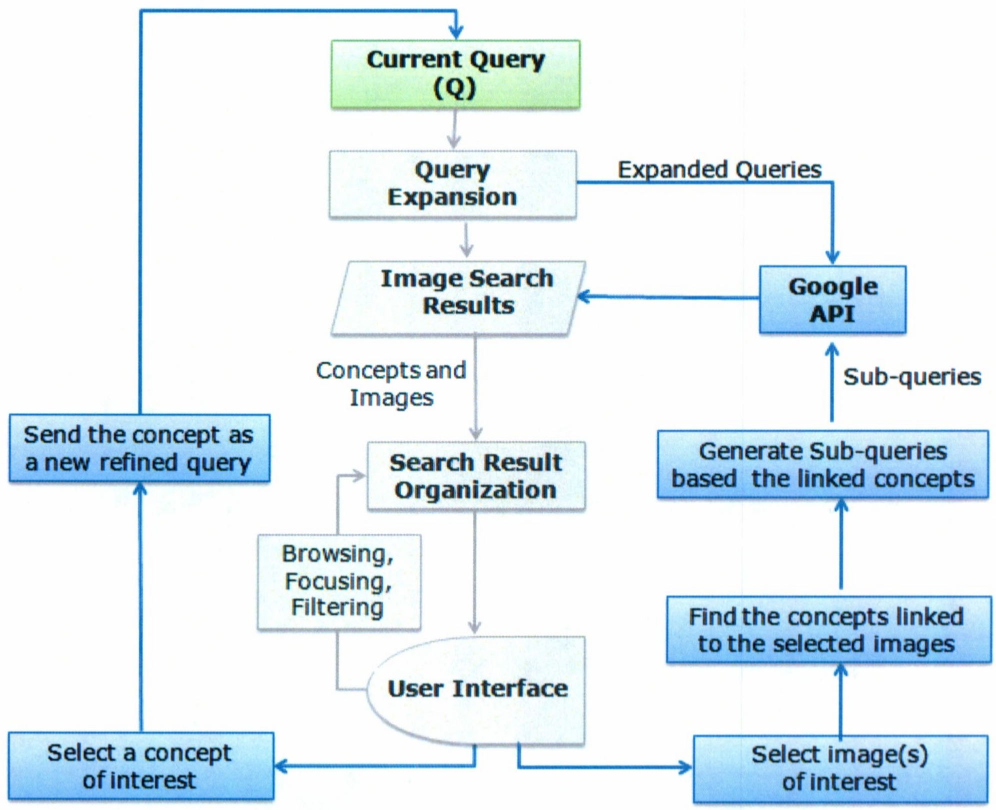


Figure 5.1: The steps in concept-based interactive query refinement. Note that the steps belonging to the approaches described in the previous two chapters are shown in the grey boxes.

5.2.1 Refinement by Expanded Concepts

The fundamental design decision behind this query refinement method is based on the fact that as searchers start browsing, they may want to accurately describe their specific image needs, but may not be able to recall the appropriate query terms. This is particularly evident for the approach of exploring image search results described in Chapter 4. That is, as searchers browse the concept hierarchy and the images associated with the concepts, a specific concept may be identified as a more accurate description of the information needs. In this situation, the representation of the concepts in the concept hierarchy can be an effective way to interactively support the searcher in the query refinement process. This will allow the users to begin with a vague initial query, and select the appropriate terms to refine the query. The resulting query refinement feature supports a fundamental usability principle given by Nielsen: *recognition rather than recall*, implying that the searchers should be able to see the information they need, rather than having to remember it [55].

The searcher can initiate this query refinement step by choosing a concept from the concept hierarchy. By right-clicking on this concept C_i, it is placed in the query panel. Consequently, a new search is initiated based on the query as $Q = C_i$. This query is automatically expanded, and a broad range of images are retrieved. During this process, the top-N related concepts to this new query Q are used in the query expansion.

As the images are retrieved, they are organized according to the conceptual and visual similarity, and then displayed within the image space. In addition, a new concept hierarchy is formed based on the concepts used in the query expansion process for this refined query. This allows the searcher to start with a vague query, and then refine it into something more specific based on the concepts extracted from Wikipedia. From here, the searcher can continue the search activities using the provided exploration and refinement features.

5.2.2 Refinement by Query Images

As searchers explore the search results from different concepts, they may discover a set of images that are relevant to their information needs. In fact, the visual nature of the images can easily catch their eyes which helps them to find images that are very similar to what it is they are seeking. As such, once searchers find some interesting or relevant images, they may wish to retrieve more images that are like these. To take advantage of this behaviour, query refinement is facilitated based on the selection of a set of images.

The searcher can initiate this query refinement loop by selecting a set of images. Once the searcher becomes interested in any particular image, it can be added to the query panel by right-clicking on it (Figure 5.2). Upon clicking the search button, the system automatically finds the associated concepts for the images within the query panel, and uses these concepts as the basis for refining the queries.

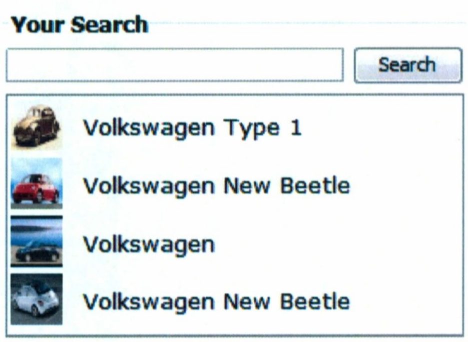


Figure 5.2: The query panel showing the selected images along with the associated concepts to be used for refining the query.

Formally stated, let us consider that the searcher starts with the query Q, which is expanded based on top-N related concepts. The returned image set is then organized and presented to the searcher. As the searcher explores the search results, they may choose any number of images and add them to the query panel. Once the refined search is initiated, the system finds the concepts that are linked with the images. If there are M different concepts $\{c_r|0 \leq r \leq M\}$ associated with the selected images, then the system will generate M different queries based on the associated concepts as $< h_r, c_r >$. Here h_r is the home article to which a particular concept c_r belongs to.

Each of these M different queries is sent to the Google AJAX Search service, and new search results are produced. During this process, the number of images to be retrieved I_t, remains the same as for the previous search results (300). However, instead of running all the $(N+1)$ sub-queries, this time only M related sub-queries are used (a subset of queries from the previous query expansion). As such, for each of the sub-queries, a larger set of images are retrieved than the set of images retrieved for the same sub-query in the previous search results. The number of images returned for each sub-queries is determined based on the semantic relatedness to its home article. This calculation is based on Equation 3.4, but has been modified by replacing the value of N with M:

$$I_r = \frac{r(c_r, h_s) \times I_t}{\sum_{k=1}^{M} r(c_k, h_s)} \tag{5.1}$$

Unlike the previous approach to query refinement where query expansion is performed for a particular selected concept, this query refinement process narrows the scope of the query to just the images (and associated concepts) chosen by the searcher.

As such, the query expansion process is not performed; instead more images are retrieved only based on the specific concepts associated with the query images. The retrieved images are organized and displayed within the image space, and only the concepts that produced these images are shown in the concept hierarchy. From here, the search results can be explored, and further query refinement activities can be performed.

5.3 Example

In order to illustrate the benefits of interactive query refinement, two examples are provided based on the two approaches discussed previously: refinement by expanded concepts and refinement by query images.

5.3.1 Refinement by Expanded Concepts

Suppose the searcher starts with the query "Beetle" and explores the images from different concepts of Volkswagen Beetle cars. During this exploration process, any particular concept from Beetle car may seem to be highly interesting to the user. For example, the searcher may have an interest in not just Beetles, but other vehicles from VolksWagen that were built in the same era. As such, by seeing the Volkswagen Type 2 in the concept hierarchy, they may identify this as being relevant to their information needs. The searcher can then select this particular concept of interest from the concept hierarchy to the query panel, so that this concept can be expanded. The system then performs query expansion on Volkswagen Type 2, returning a set of images from the closely related concepts.

The overview of the search results for the refined query "Volkswagen Type 2" are shown in Figure 5.3. Note that for this new query, the top-N related concepts are presented in the concept hierarchy, most are some variants or highly related models of Volkswagen Type 2, such as Volkswagen Type 2 (T2), Volkswagen Type 2 (T3), Volkswagen Transporter (T4). It could be possible that the searchers were interested about the different models of Volkswagen Type 2 before, but they did not know the words to use to describe what they were looking for. As such, when these concepts were presented in the concept hierarchy, it was easier for them to find their images of interest. From here, the searcher can again explore the new search results, such as images from different types of Volkswagen Type 2 models. If needed, further query refinement can also be performed.

In this scenario, the searcher was able to start with an ambiguous query, and then refine the query by choosing a specific concept of interest. As the new expanded search results are provided, the searcher was able to learn even more about the specific concept by exploring its different sub-topics (top-N concepts presented in the concept hierarchy). From this example, it is easy to see the value of being able to discover more knowledge about a specific concept of interest supported by the interactive query refinement.

5.3.2 Refinement by Query Images

Let us again consider the scenario where the searcher begins with the vague query "Beetle". The search results returned from this query contains images from insects, cars and comic characters. However, the searcher wants to explore the images of

Figure 5.3: New search results based on the refinement by expanding the concept "Volkswagen Type 2".

different Beetle car models in great detail. As users browse the images of Beetle cars, some images may appear as interesting or look like something that they are seeking. The searcher can immediately add each image by right-clicking on the image and adding it to the query panel, as shown in Figure 5.4. From this figure, it can be seen that the searcher chose some images from Volkswagen Beetle models such as Volkswagen New Beetle and Volkswagen Type 1 for further query refinement.

Once the search button is clicked, concepts associated with the selected images are used to retrieve a new set of images. As such, the new image search results contain images only from the concepts associated with the chosen images. Images from other concepts related to the previous query are no longer available, which allows the system to retrieve more images from the associated concepts (since the total number of images to be retrieve remains same as the previous query). The

Figure 5.4: A screenshot of the image search interface, showing the search results for the query "Beetle". As the searcher selects some images for refining the query, they are added to the query panel (the right side).

new search results are shown in Figure 5.5, where the concept hierarchy contains only the concepts associated with the chosen images, and the search results contain a new larger set of images for each concept than the previous query. As images from other interpretations (such as insects and comic characters) are eliminated from the search results, the searcher can easily focus on the associated concepts presented in the concept hierarchy and corresponding images in the image space.

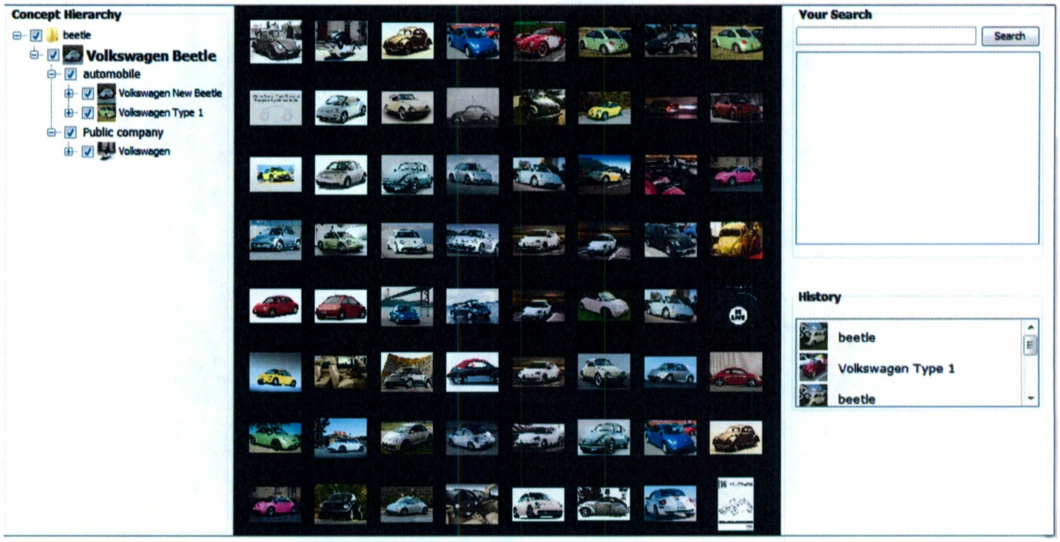


Figure 5.5: A screenshot of the image search interface, which shows the new search results based on the refinement by query images.

5.4 Discussion

For ambiguous queries, query expansion followed by search results organization provides an appealing way to support the searcher in discovering the concepts and images that they are looking for. In addition, it is also beneficial to allow them to further refine the query, and find more images according to their image needs. To facilitate this feature, two alternative approaches for query refinement are proposed. Fundamentally, both of these approaches narrows the scope of the original query.

As noted in Section 2.4, one of the main challenges for all interactive query refinement approaches is providing the searcher with sufficient information upon which to refine the query. In this work, the concept information derived in the query expansion process was used as the primary source for designing query refinement features. When such semantically enriched information is used in the process of refining the query, then the returned results provide more conceptually related perspectives, which are often expected by the searchers.

Query refinement by expanding a concept allows searchers to enhance the query by exclusively focusing on only one concept, treating it as a new query. This new query is expanded, which may allow the searcher to easily deviate from the initial goal of the query. In other words, this feature may support searchers to find something serendipitously, which are of interest to the searcher but are not particularly relevant to the current search activity. As discussed previously, enhancing this phenomena in the query refinement was one of the primary motivation for designing this approach.

The method of refining the query by images takes advantage of the visual processing capabilities of human beings, allowing the searchers to select some images instantly as they catch their eyes. Unlike query refinement by expanding a concept, this method keeps the focus of the results on the original query, as the initial query is prepended to each of the concepts found from the selected images. As such, there is less chance to shift from the primary goal.

Query refinement by images might seem to be similar to the relevance feedback techniques used in the context of CBIR [46, 13]. As mentioned in Section 2.4, relevance feedback techniques allow the searcher to provide some images as positive examples and then analyze the content of the these images, so that more search results that are visually similar to the given examples could be retrieved. Although this approach might be helpful in retrieving visually similar images, it lacks semantic information in the retrieval process [13]. Query refinement by images (as explained in this chapter) can be viewed as a relevance feedback technique. However, the main advantage of this approach over traditional relevance feedback techniques is that it uses conceptual information associated with the images of interest, allowing the retrieval of more semantically related images, which is more desirable in the context of Web image search.

One of the important design aspects of this interactive query refinement method is its iterative nature. In Section 2.4, a related method for interactive query refinement for image search has been mentioned, which promotes faceted exploration of the image search results [80]. Using their system, searchers were able to refine the initial query by exploring different related facets (concepts). However, the system does not provide further query refinement support, which would require the system to run the facet as a new query and perform query expansion on it. On the contrary, the methods described in this chapter allows a searcher to start with a rather vague query and then perform a number of query refinement loops as required, so that the refined query may become more and more specific.

Although the examples presented in this chapter reveal the potential benefits of interactive query refinement, one may expect further evaluations on this approach. In this regard, a possible way could be to conduct user studies. However, it is very difficult to measure the benefits of interactive query refinement features in a controlled laboratory setting. The main difficulty is that the ability to effectively refine a query is largely dependent on the individual's prior knowledge on the query topic and their ability to learn about the topic during the search task. Also, it is very difficult to decide when the tasks have been successfully completed [27]. As such, no user study was conducted to validate this feature; instead illustrative examples were provided in the previous section to demonstrate its potential benefits.

Chapter 6

Evaluation

6.1 Introduction

In the previous three chapters the primary approaches of this thesis were discussed. Although the examples and analysis provided within these chapters showed some promise regarding the benefits that the system provides, further comprehensive evaluations are required to validate the value of the proposed approaches. The main goal of these evaluations is to address the fundamental research questions that were asked regarding the proposed approaches (as mentioned in Chapter 1).

In this chapter, three different evaluations are described that address various aspects of the proposed approaches. In the first evaluation, the effect of query expansion on the retrieval efficiency of image search results is examined. Query expansion is an effective way to promote diversity within the image search results; however, broadening the query aggressively may lead to significant drop in precision. In this context, a set of experiments varying the degree of diversification were preformed to illustrate the trade-off between diversity and precision.

In the second evaluation, the value of the different interactive exploration features of the developed image search interface (CBIS) is examined through an empirical study. This study shows how the precision of the search results changes as a result of interactive exploration activities (as explained in Chapter 4).

Finally, a user study was conducted in order to measure the potential benefits provided by CBIS compared to a traditional Web image search interface (Google Image Search). The primary objective of this study was to analyze whether real users can benefit from the proposed system, and to measure their subjective reactions and impressions in using the image search interface.

6.2 Evaluating the Precision-Diversity Trade-offs for Query Expansion

The goal of the query expansion process is to automatically diversify the images retrieved for a given query. However, it is unclear to what degree such diversification should be promoted during the search process. In this evaluation, the goal is to study the inherent trade-off between precision and diversity in detail. In particular, for a set of queries, the changes in precision are measured as diversity in the image search results is increased. Using this information, a simple approach is proposed to automatically determining the degree of diversification based on features of the user-supplied query.

6.2.1 Methodology

For these experiments, 12 query topics were chosen, divided between those that were deemed to be highly ambiguous (having four senses), moderately ambiguous (three senses), slightly ambiguous (two senses), and non-ambiguous (one sense). This distribution of different degrees of ambiguity helped to examine the effect of the experimental condition (i.e., the degree of diversification) in the context of ambiguity. For each of these different degrees of ambiguity two queries were selected, except for the moderate ambiguity, for which six queries were chosen. The moderate degree of ambiguity was examined more carefully since it represents the most common case of ambiguity.

To evaluate the effect of diversification on precision, the top 60 search results from Google Image Search were retrieved using the concept-based query expansion method with ten different values of N, ranging from 0 to 40 ($N = 0, 2, 4, 6, 8, 10, 15, 20, 30, 40$). Here, $N = 0$ implies that no query expansion has occurred (i.e., the search results are not diversified, and are simply the results provided by the underlying image search engine). At the other extreme, $N = 40$ causes the system to return a highly diversified set of image search results from 40 different associated concepts chosen in the query expansion procedure. Data was collected more frequently in the low end of this range in order to more closely observe the effect of a low degree of diversification. Preliminary experiments illustrated that the effect of the degree of diversification at the higher range became rather stable; as such, the values of N ware sampled at increments of 10 at the high range.

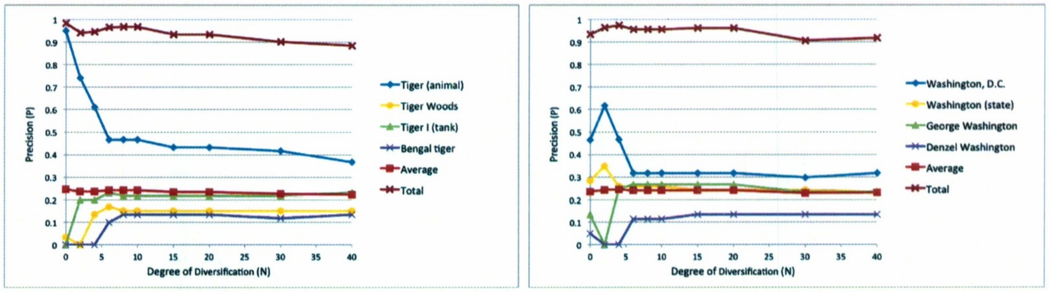
For each of the different senses of the query, assessors were asked to judge the relevance of each image. This assessment of relevance provided the ground truth information in the calculation of the precision scores (the ratio of relevant images to the total number of images retrieved). Since there were ten trails (i.e., ten different values of N) and 60 images retrieved with each trial, this resulted in the evaluation of a total of 600 images for each test query.

6.2.2 Results

In these experiments the precision for each of the test queries were measured as the diversification parameter N was varied from 0 to 40. The hypothesis was that as N increased, the distribution of the senses would become more balanced across all of the meanings of the query. This would result in a reduction in the precision for the most common senses of the query, and an increase in the precision for the less common senses. This feature can be readily identified in the graphs in Figures 6.1, 6.2, and 6.3. To further understand this effect, the average precision (the red lines with the square markers) and the total precision (the dark red lines with the x marker) were plotted across all of the senses.

Figure 6.1 shows the results from the highly ambiguous queries. Generally, the precision for the most common sense (i.e., the blue line in each graph) was automatically reduced as a result of the diversification. At the same time, the precision for all other senses increased. In both cases, images from some of the less common senses of the query were not represented in the search results, when the value of N was set rather low ($N < 6$).

For both queries, the precision of the images over each of the senses of the query



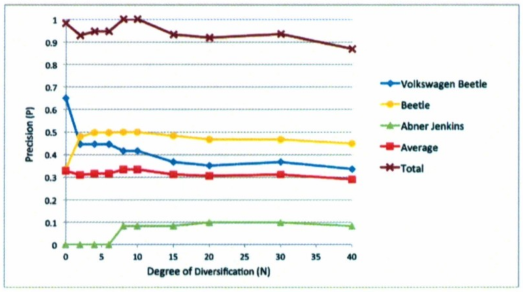
(a) Tiger (b) Washington

Figure 6.1: The effect of varying the degree of diversification (N) on precision (P) for highly ambiguous queries with *four* different senses.

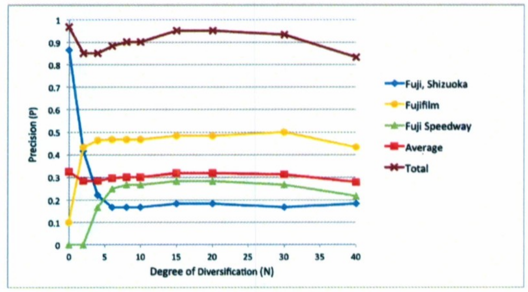
did not change significantly once the value of N was set beyond six to ten, depending on the specific query. Even when the value of N was large (i.e., 30 or 40), the average and total precision did not decrease drastically. This suggests that for highly ambiguous query even a higher degree of diversification does not have much negative impact on precision.

The expectation when designing these experiments was that for queries that have a higher degree of ambiguity, it would be necessary to set the diversification parameter rather high in order to capture enough information on all of the different senses. However, it is clear that even with a diversification parameter set at $N = 10$, the desired effect appears.

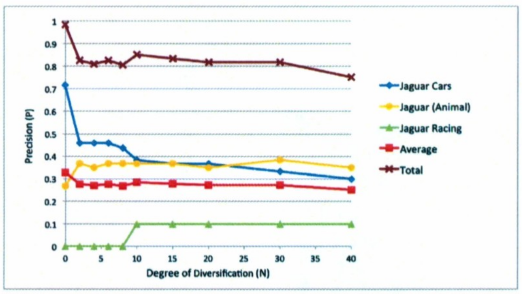
Figure 6.2 shows the results from the moderately ambiguous queries. Similar to the highly ambiguous queries, the precision for the most common sense (i.e., the blue line in each graph) was automatically reduced as a result of the diversification. In most cases, this occurred in a more or less smooth fashion even with very low values of N. Simultaneously, the precision for all other senses increased that indicates the



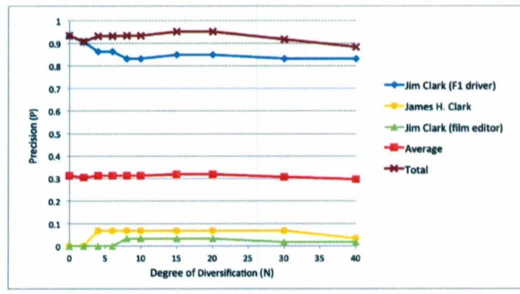
(a) Beetle



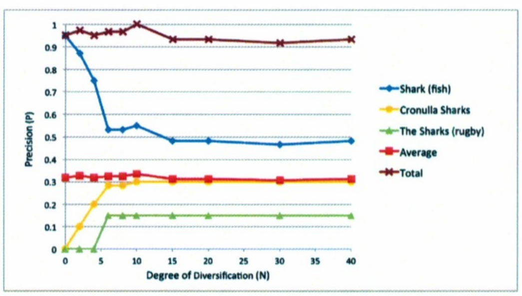
(b) Fuji



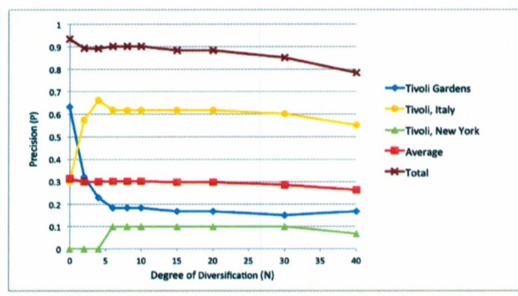
(c) Jaguar



(d) Jim Clark



(e) Sharks



(f) Tivoli

Figure 6.2: The effect of varying the degree of diversification (N) on precision (P) for moderately ambiguous queries with *three* different senses.

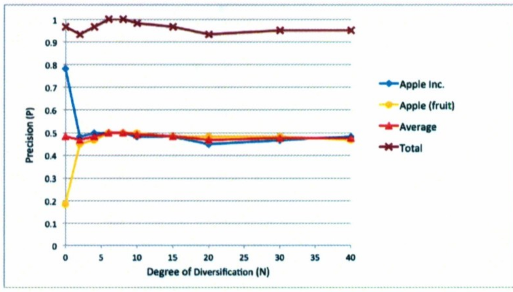
enhancement in diversity. In some cases, it was necessary for the value of N to be set higher than six to ensure that images from the less common senses of the query are represented in the search results.

Although there was, in some cases, a minor reduction in the total precision for very low values of N, this was often accompanied by a subsequent increase in the total precision as more diversification occurred. This effect is a result of the method by which the number of images retrieved for each expanded query is dynamically determined. With very few expanded queries, more images are retrieved for each (which may result in the inclusion of some less relevant images deeper in the search results list). As N is increased and more expanded queries are generated, the images that are retrieved have higher rankings with respect to their source query.

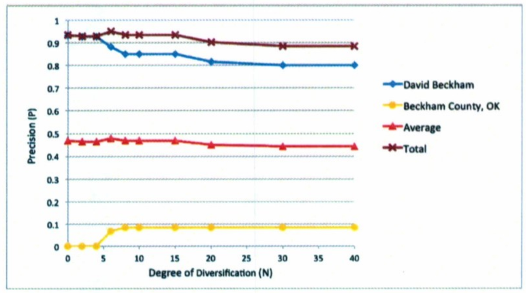
For all of these queries, the precision of the images over each of the senses of the query did not change significantly once the value of N was set beyond six to ten, depending on the specific query. Furthermore, if the value of N was set too large (i.e., 30 or 40), the average and total precision started to decrease, indicating that some non-relevant concepts and their associated images were being included in the search results set due to over-diversification.

In some cases, the most common sense remains the most common regardless of the level of diversification (e.g., Figure 6.2c, d, and e). In the other cases, senses that were less common in the original search results become dominant. The reason for this change is that there may be disagreement between what the underlying search engine assumes is the desired information need (i.e., the most common sense of the query), and the amount of information that can be extracted from Wikipedia on the other senses and their associated concepts.

Figure 6.3 shows the results from the slightly ambiguous queries. As with the highly and moderately ambiguous queries, even with a low degree of diversification (from two to six), the outcome of the diversification is a balancing of the precision between the two senses of the query. In addition, the effect of a dropping then



(a) Apple



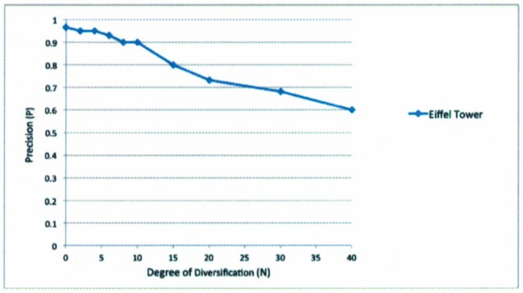
(b) Beckham

Figure 6.3: The effect of varying the degree of diversification (N) on precision (P) for slightly ambiguous queries with *two* different senses.

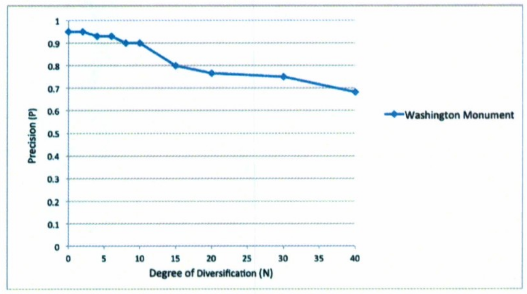
increasing total precision over low values of N is present, as is the dropping average and total precision when N is set too large. However, the later effect occurs with lower values of N than with the highly or moderately ambiguous queries (15 - 20 for the slightly ambiguous queries).

For the queries where there was only one sense (Figure 6.4), it is clear that diversifying the search results can very quickly have a negative effect on the precision. However, for small values of N (e.g, 2 - 4) this effect is negligible. This indicates that even for very specific queries, a small degree of diversification might be tolerable and perhaps even beneficial as highly related images are drawn into the search results set.

As a result of this analysis, it can be concluded that diversifying the image search results can be very useful for addressing the situation where an ambiguous query has multiple senses. Rather than relying on the search engine to choose the most common sense, it is beneficial to diversify the image search results and let the user focus on those images that match their needs. The more senses that can be inferred from a query, the more diversification is necessary to sufficiently balance all of these



(a) Eiffel Tower



(b) Washington Monument

Figure 6.4: The effect of varying the degree of diversification (N) on precision (P) for non-ambiguous queries with only *one* sense.

senses in the search results. However, when there are few different senses, the degree of diversification should be limited to avoid including irrelevant concepts and their associated images in the search results.

6.2.3 Automatically Determining the Degree of Diversification

One of the the major goals in this study is to automatically determine the degree of diversification needed to enhance the searcher's ability to find relevant documents. As illustrated in the experiments, the degree of ambiguity of the query has an impact on the value of diversification. That is, a highly ambiguous query can benefit from more diversification, whereas a very specific query may be harmed by diversifying the search results too much.

Based on this, a linear scaling of the degree of diversification is proposed based on the number of senses of the query (as determined by Wikipedia). One such automatic function for determining the how aggressively to diversify the search results is:

$$N = a \times q \tag{6.1}$$

Here, q is the number of senses for query Q (following the terminology from Chapter 3).

Based on the experiments reported in this section, setting $a = 3$ will produce reasonable results. That is, if a query has one sense, the diversification parameter will be set to 3, performing a relatively low degree of diversification that does not harm the precision of the one sense of the query. For the slightly, moderately, and highly ambiguous queries, this results in the diversification parameter being set to 6, 9, and 12, respectively. By inspecting the graphs in Figures 6.1 - 6.4, it can be seen that this is a near-optimal value for N.

6.2.4 Discussion

Without diversification, the most common sense of an ambiguous query can dominate the image search results. Images that match other senses of the query may not be very common or even represented at all within the top search results. This is a preferential outcome if the searcher's information needs match the underlying search engine's interpretation of the query. However, when this is not the case, it can be very difficult for the searcher to find images that match their needs. Their only recourse is to attempt to reformulate the query into something more specific. However, studies on Web image search behaviour indicate that this may be a rather difficult task for searchers to perform [2].

The key benefit of performing image search diversification is that instead of assuming a single interpretation of the query, images are retrieved from different senses discovered using Wikipedia, providing a more balanced distribution of the senses within the set of search results. As a result, this diversified set of search results may be suitable to a wider range of users with a wider range of information needs.

A common problem with query expansion in general, and search result diversification in particular, is that it can reduce the precision by including documents (images) that might not be relevant to all searchers. This remains the case for this work on image search diversification. In particular, this study has shown how the precision for the most common sense will invariably be reduced. This is because images from the less common senses are being included in the search results as a result of the diversification process. However, even amid this reduction in the precision of the most common sense, and the increase in the precision of the less common senses, the average precision is not chaining noticeably when the degree of diversification is not set unreasonably large. On the contrary, in some cases the total precision actually increases, when calculated over all senses of the query. This indicates that the proposed approach is effective in keeping the expanded queries focused on the intended senses of the query, even when there are multiple such senses.

Because the process of diversification will commonly result in a mix of relevant and irrelevant images for a given interpretation of an ambiguous query, it is important that an interface to the image search results be used that makes it easy for the searcher to ignore the senses of the query that are not relevant to their needs, and focus on those that are. Such an interface was described in Chapters 4 and 5. Using this visual interface to explore the search results, the precision for a particular sense of the query may be improved as irrelevant images are moved out of the field of view or are filtered based concepts that are uninteresting to the searcher. An evaluation of this interface from the perspective of improving the precision of the search results through these interactive operations is discussed in the next section.

It is also worth noting that the benefit that this approach to image search diversification can provide depends greatly on the completeness and interconnected nature of information in Wikipedia. If there is a sense of a given query that is not well represented in Wikipedia, then it will not be well represented within the diversified image search results set. As such, as Wikipedia continues to grow and be enhanced, it will become a better tool for providing knowledge for information retrieval systems such as what has been discussed in this research.

This simple approach to determining the diversification parameter N could benefit from further research in fine-tuning this formula and its parameters for automatically determining the value of N based on some measure of query ambiguity.

6.3 Evaluating the Potential Benefits of Search Results Exploration Techniques

In Chapter 4, the approach for the organization and exploration of image search results used in CBIS was discussed in detail. Within this approach, a set of interactive operations were presented to support image search results explorations: concept-based focusing, concept-based filtering, and visual zooming. In order to evaluate the potential benefits of these operations, the changes in image retrieval performance are studied as the image search results space is explored.

Within this analysis, it was assumed that searchers were able to make correct judgements regarding which concepts were relevant to their search needs, and were able to identify regions within the visual organization of the images that appeared to include relevant images. This assumption can be justified by the fact that image search is more visual in nature than document search. That is, a searcher generally has a mental picture of the desired images prior to searching, and when an area of images having visual similarity to that image needs is present in the display, it can be easily identified even by a moderately skilled searcher. Within this context, the goal of this evaluation was to measure the ability of the interactive features to support focusing and filtering of the image search results, but not the performance of individual searchers.

6.3.1 Methodology

The performance of CBIS was studied using a set of six ambiguous queries. Each of these queries were chosen such that they had a similar level of ambiguity. In particular, each query produced a strong match to three different interpretations within Wikipedia. Using the concept-based query expansion process described in Chapter 3, 300 images were retrieved for each query. A pair of assessors were asked to judge the relevance of each image in the search results with respect to each of the different senses of the query. These relevance scores were determined by examining each image and their corresponding text snippets. The combined relevance scores from the assessors were used as ground truth scores in the calculation of the precision.

Since both visible and hidden images may be included within the viewport, two different variants of the precision metric were defined for this analysis. Pv is the precision considering only the visible images within the viewport, and Pa is the precision calculated over both the visible and hidden images within the viewport. Pa and Pv are calculated according to the following equations:

$$Pv = \frac{|I_{visible} \cap I_{relevant}|}{|I_{visible}|} \tag{6.2}$$

$$Pa = \frac{|(I_{visible} \cup I_{hidden}) \cap I_{relevant}|}{|I_{visible} \cup I_{hidden}|} \tag{6.3}$$

Here, $I_{visible}$ and I_{hidden} represent the sets of visible images and hidden images within the viewport. Also, $I_{relevant}$ represents the set of all the relevant images within the image space.

To measure the potential benefit of the concept-based focusing and filtering operations, the change in precision was analyzed as these features were used. For the focusing operation, Pv was measured using the default settings (with no focusing, filtering, or zooming), and when the best concept for each of the three different interpretations of the query were selected. Note that a measurement of Pa does not make sense here since the focusing operation brings all of the images linked to the selected concept into view.

For the concept-based filtering operation, the middle frequency interpretation of the query was chosen, and only the top ten relevant concepts of this interpretation were selected using the checkboxes within the concept hierarchy. Both Pa and Pv were measured before and after this filtering operation was performed. Observing the effect of the filtering operation is difficult for the least common interpretation of the query, since it does not relate to sufficient number of concepts that can be filtered out. On the other extreme, for the most common interpretation of the query, filtering operation would obviously produce more positive impact on precision comparing to the middle frequency interpretation, since more concepts could be filtered out. Therefore, the experiments of filtering operation on the most common interpretations and least common interpretations were not performed.

To study the effect of the visual zooming operations, the precision measurements Pa and Pv were taken at three different levels of zoom for each query. At the top-level zoom (defined as zoom = 0), the viewport was completely zoomed out so that the entire image space was shown, but with many images hidden. At the bottom-level zoom (defined as zoom = 2), the viewport was zoomed in completely so that there were no hidden images, but with many images outside of the viewport. The mid-level zoom (defined as $zoom = 1$) was set at half-way between the top-level and bottom-level zoom settings, such that there were both images outside of the viewport and images hidden. The focal points for these zoom operations were chosen separately for each of the three different interpretations of each query.

6.3.2 Results

6.3.2.1 Concept-Based Focusing

The effect of the concept-based focusing operation on the precision the visible images (Pv) is illustrated in Figure 6.5. For each of the interpretations of the test queries, the focus operation resulted in many relevant images being brought into view. At the same time, many of the images retrieved as a result of the other non-focused concepts

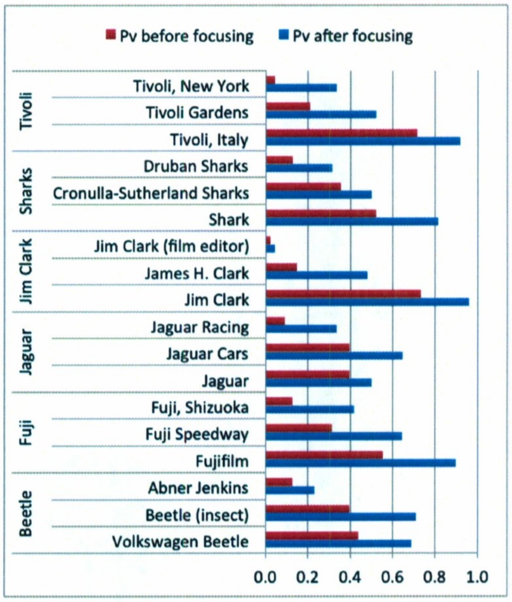


Figure 6.5: Precision calculated over the visible images (Pv) before and after using the concept-based focusing operation. In all cases, more relevant images are pulled into view by focusing on concepts that are related to each of the interpretations of the queries.

became hidden. As a result, Pv increased in every case.

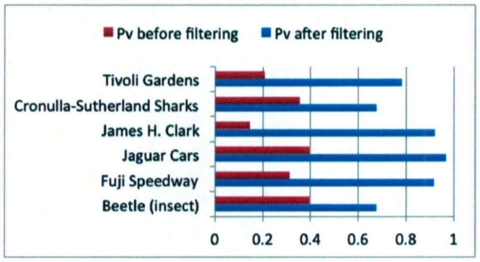
The goal of the query expansion process was to diversify the image search results so that they were not all from the most common interpretation of the query. Even so, the precision of the visible images prior to any user interaction are not evenly distributed across all of the senses of the queries. This is due to the method by which the query expansion is performed, wherein the number of concepts selected, as well as the number of images retrieved for each concept, vary depending on how closely the concept matches the home article in Wikipedia. However, as a result of the concept-based focusing operation, it is possible for many more relevant images to be shown, even for interpretations of the query that are relatively uncommon.

6.3.2.2 Concept-Based Filtering

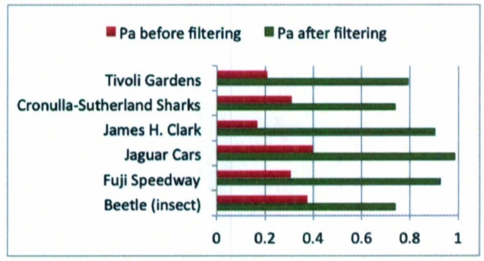
The results of the concept-based filtering operation performed for the middle frequency interpretation of the queries are shown in Figure 6.6, measured over both the precision of the visible images (Pv) and the precision of all the images within the viewport (Pa). Here a substantial increase in the precision can be seen as a result of only showing the images that were retrieved as a result of the ten best concepts for these interpretations of the queries. That is, by de-selecting the non-relevant concepts, many of the non-relevant images were also filtered out, resulting in increases in precision for each of the test queries.

The similarity between the Pa and Pv values here illustrates that the images hidden due to space constraints are not any more or less relevant than those that are shown. This outcome provides evidence of the value of the multi-resolution SOM method for organizing the images and defining the display priority to determine which images to show when the display space is constrained. That is, the visual feature vectors of the most relevant images generally closely match with the average feature vector of a region in SOM (as they have visual similarity with other relevant images), that causes them to have higher display priority.

The goal of this experiment was to demonstrate that even though a number of potentially irrelevant concepts could be included within the query expansion process, eliminating them through the concept-based filtering operation improves precision (Pa and Pv). However, the amount of improvement in precision may depend on the number of concepts filtered as well as the individual features of the query and the



(a) Precision calculated over the visible images (Pv)



(b) Precision calculated over all of the images (Pa)

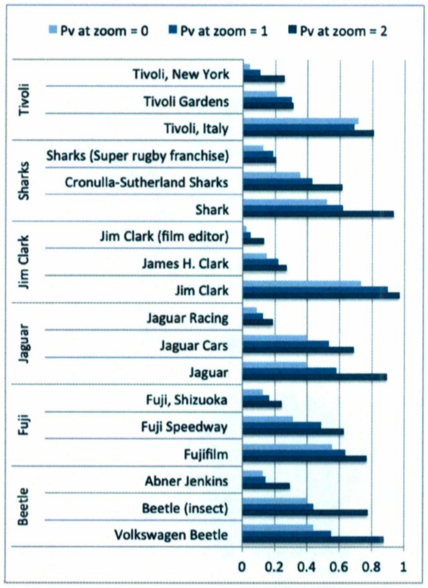
Figure 6.6: Filtering out all but the ten most relevant concepts associated with the middle frequency interpretation of the queries results in a significant improvement in precision, both among the visible images and over all of the images within the viewport.

corresponding search results set.

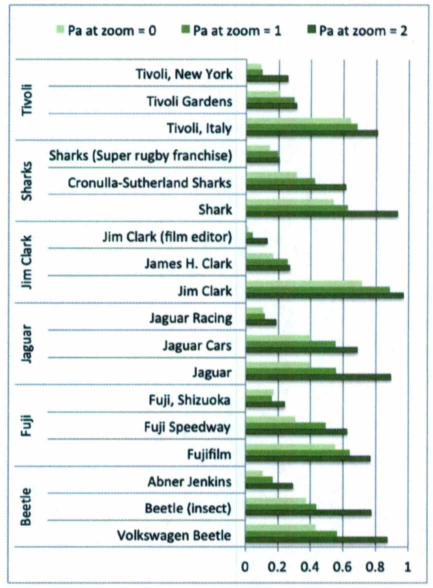
6.3.2.3 Visual Zooming

The evaluation results for the visual zooming feature are depicted in Figure 6.7. For nearly every interpretation of each of the test queries, as the image space was zoomed into a region that appeared to contain relevant images, both the precision of the visible images and the precision of all the images within the viewport increased. The zoom operation had two simultaneous effects on the image search results. Not only were images moved out of the viewport (effectively providing a filtering operation), but as more space was created, more images that were previously hidden became visible (effectively providing a focusing operation).

Note that at $zoom = 0$ (i.e., top-level zoom), the Pv values are the same as those reported for the concept-based focusing operation in Figure 6.5, and the Pv



(a) Precision calculated over the visible images (Pv)



(b) Precision calculated over all of the images (Pa)

Figure 6.7: In almost all cases, zooming into regions that appear to contain images relevant to a given interpretation of a query results in an increased precision among the images, both for those that are visible and for all the images within the viewport. Here, zoom = 0 is the top-level zoom; zoom = 1 is the mid-level zoom; and zoom = 2 is the bottom-level zoom.

and Pa values for the middle interpretation are the same as those reported for the concept-based filtering operation in Figure 6.6. Furthermore, the similarity between the Pa and Pv at the different zoom levels is consistent with the findings from the concept-based filtering operation. Indeed, at $zoom = 2$ (i.e., the bottom-level zoom), there are no more hidden images, resulting in $Pv = Pa$.

The goal of this experiment was to demonstrate that zooming into a suitable area in image space (where a large number of images are available with respect to the user's meaning) could potentially improve precision (Pa and Pv). As with the concept-based focusing operation, the degree to which the zoom operation had the ability to increase the precision depends on the particular makeup of expanded set of image search results. For most interpretations of the queries, there were many visually similar relevant images that were grouped together, making the zoom operation very effective (e.g., Fuji Speedway). For others, the images were visually dissimilar, resulting in their distribution throughout the image space, limiting the ability of the zoom operation to increase the precision (e.g., Tivoli, Italy). Also, the ability of the searcher to find an area of interest in the image space has an important effect on precision.

6.3.3 Discussion

The evaluation of the interactive features of CBIS for image search results exploration provides empirical evidence regarding the potential benefits of the approach described in Chapter 4. The conceptual focusing operation was shown to increase the precision of the images that are visible. The conceptual filtering operation was shown to be highly effective in removing non-relevant images among both those that are visible and those that are hidden. Zooming into a region of interest of the image space also resulted in improving the precision among both the visible and all the images within the viewport.

Within this analysis, all of the experiments were conducted based on the assumption that the searcher would be able to make intelligent choices regarding which concepts were relevant and which regions of the image display space appeared to contain relevant images. That is, the searcher does not make mistakes during the search process. The ability of the real users to achieve this level of performance is studied in the next section.

6.4 User Study

6.4.1 Purpose

For evaluating Web search interfaces, conducting user studies is not only regarded as an accepted practice but it is one of the more valuable ways in which researchers can evaluate the potential benefits that their work might have for real users. It provides researchers with a means to verify and validate design assumptions, confirm or reject hypotheses, and make comparisons between different systems and techniques [27]. As such, to obtain empirical evidence regarding the performance of the developed image search system (CBIS) in comparison to a baseline system (Google Image Search), a user study was conducted in a controlled laboratory setting.

Designing a user study for image search interfaces is inherently difficult because of the exploratory nature of the search activities, where the users' image needs are ill-defined. The main challenge here is to clearly and precisely define the image search task and the conditions under which the task is considered completed. While a common procedure for conducting user studies is to assign tasks to the participants and measure the time it takes to complete the tasks and the number of errors made, such measures alone may not provide a clear picture of how much knowledge is gained through the search activities.

In these situations, participants' subjective opinions can also be collected in order to measure the knowledge gain through the search process, overall usefulness and satisfaction in completing the task, and the ease of use of the software. Participants' judgements can also be recorded in terms of the usefulness of specific features of the systems being studied.

6.4.2 Methodology

The study was designed as a 2×2 ($interface \times$ $search\ task\ complexity$) between-subjects design. Each participant used two interfaces (CBIS and Google Image Search). Task complexities were of two types: simple tasks (S) and moderately complex tasks (C). For each interface, the participant performed two different tasks (a simple task followed by a moderately complex task) for a total of four tasks. To alleviate potential learning effects, the order of the tasks assigned to the participants was varied (see Table 6.1).

Since the participants are expected to be more familiar with list-based representations of image search results, they were exposed to Google Image Search first for performing two different search tasks, and then performed another two different search tasks using CBIS. Prior to performing any of the tasks, participants were given a brief introduction to the features of each of the two interfaces.

To reflect the exploratory nature of the image search activities, tasks were designed based on a set of four ambiguous queries (Table 6.2). To avoid the effect of variation in ambiguity of the query, each of these queries was chosen such that they had a similar level of ambiguity. In particular, each query produced a strong match to three

Table 6.1: Order of tasks assigned to participants.

Order 1	Order 2	Order 3	Order 4
S1	S2	S1	S2
C1	C1	C2	C2
S2	S1	S2	S1
C2	C2	C1	C1

different interpretations within Wikipedia. However, task complexity was varied for different queries. For the query "Beetle" and "Tivoli", the tasks were relevant to the dominant interpretations of the query, and desired images were found within the first few top ranked images provided by Google Image Search. As such, these two tasks were considered as simple tasks. However, for the other two queries ("Fuji" and "Jaguar"), the tasks were related to the less common interpretations of the queries, and the relevant images were scattered throughout the search results.

Table 6.2: Ambiguous queries and assigned tasks

Task ID	**Query**	**Possible Interpretations**	**Assigned Task**	**Task Complexity**
S1	Beetle	Beetle (insect) Volkswagen Beetle Abner Jenkins	Find five images of the Volkswagen cabriolet car	Simple
S2	Tivoli	Tivoli, Italy Tivoli Gardens Tivoli, New York	Find five images of Villa d'Este	Simple
C1	Jaguar	Jaguar (animal) Jaguar Cars Jaguar Racing	Find five images of the Jaguar S-Type model car	Moderately complex
C2	Fuji	Fuji, Shizuoka Fujifilm Fuji Speedway	Find five images of the Fuji Speedway	Moderately complex

For CBIS, the concept-based query expansion process was performed, 300 images were retrieved for each query, and the diversified image set was organized and presented to the searcher. The diversity within the search results allowed the participants to explore the results in some detail in order to find the relevant images. For the baseline interface, 300 top ranked images returned by Google Image Search were presented in a traditional list-based interface.

At the beginning of the user study, pre-study questionnaires were administered to determine prior experience with image search, educational background, and computer use characteristics. Then, the tasks were assigned according to one of the possible orders provided in Table 6.1. For each task, participants were given a scenario in which they were asked to find five images that are relevant to the described information need. During each task, measurements of time to task completion, accuracy, and subjective measures (such as perceived difficulty of the tasks, perceived satisfaction and perceived knowledge gain) were made. In addition, post-study questionnaires designed by the guidelines of the Technology Acceptance Model (TAM) [14] were filled out by the participant, capturing the subjective reactions about usefulness and ease of use of the both interfaces. As well, questions were asked regarding the usefulness of specific interface features. Finally, the participants were asked to provide an indication of preference for an image search interface.

6.4.3 Hypotheses

Based on the observations and knowledge about the two image search interfaces under investigation, a set of hypotheses were formulated for the user study. These hypotheses are as follows:

H1: *For simple tasks, participants will find relevant images faster with Google Image Search than CBIS. However, for moderately complex tasks, participants will find relevant images faster with CBIS.*

This hypothesis was provided based on the time taken by the participants to find five relevant images with respect to the assigned tasks. The reasoning is that for simple tasks, sufficient relevant images are available within the first or second page of Google Image Search results, which, it is expected, would allow participants to find images rather quickly. For CBIS, searchers would need to interact with the system by selecting the interpretations and concepts (focusing), filtering the irrelevant concepts, and zooming into an area of interest. This overhead of interaction consumes time regardless of the task complexity. As such, our expectation was that for simple tasks Google Image Search might take less time than CBIS.

For the moderately complex tasks, when Google Image Search interface is used, relevant images are distributed throughout the search results, requiring the participant to scan almost 300 images and perhaps read their captions, which is very time consuming. On the contrary, since CBIS groups the images based on conceptual and visual similarities, guided by a few interactions, the related ones could easily be found in a particular area of the image space. As a result, searchers could take less time for moderately complex tasks with CBIS than Google Image Search.

H2: *Using Google Image Search participants will take more time to complete moderately complex tasks than simple tasks. However, using CBIS there will not be any significant difference in the time they take between simple tasks and moderately complex tasks.*

This hypothesis was also formulated based on the time to task completion measure. The anticipation is that moderately complex tasks would take more time to complete than simple tasks for Google Image Search, since the images are more scattered throughout different pages.

However, for CBIS the overhead of interaction is expected to be similar regardless of the complexity of the search activities. Since in CBIS, no interpretation is highly dominated within the search results, the image space does not contain large amount of images of any particular interpretation in the initial view. This forces the participants to use some minimum level of interactions (such as concept-based focusing and zooming) to find the five images relevant to the given tasks. As such, it is expected that CBIS would require the searcher to take similar time to complete the tasks.

H3: *There will not be any significant differences between CBIS and Google Image Search in terms of accuracy.*

During each task, participants were asked to find five relevant images according to the information needs that were assigned to them. The accuracy was calculated based on the participants' answers (the ratio of number of relevant images to the number of total images provided by the participant). The expectation is that the CBIS system neither helped nor hindered the participants in deciding the relevance of individual images to the search task. As such, the accuracy might not be significantly varied between the two systems.

H4: *For simple tasks, the perceived difficulty level of a given task will be higher after using CBIS compared to Google Image Search. For moderately complex tasks, the perceived difficulty level of a given task will be higher after using Google Image Search compared to CBIS.*

To validate this hypothesis, after each task participants were asked to rate how difficult they thought the given task was. The observation is that for simple tasks, since the searcher can easily find desired images within a few top ranked results in Google Image Search, the perceived difficulty level may become less.

For moderately complex tasks, the relevant images are distributed throughout different pages. Going through all the pages sequentially and finding the images among a large amount of irrelevant images might be cumbersome and tedious. However, using CBIS the searcher can easily perform a few interactions to find all the relevant images grouped together, which may make them feel that the tasks are easier.

H5: *Participants will be more satisfied with the provided results after using CBIS than after using Google Image Search.*

To validate this hypothesis, after each task participants were asked to rate how satisfied they were in using the interface for searching images. The anticipation is that in Google Image Search, the relevant images with respect to the searchers' interpretations are not necessarily grouped together. Very often, it could be possible that once the searchers find a relevant image and move to next page to find more, they can forget the image and its location in the previous page. In this way, although they may find the required images, the scattering of images in different pages may give them the impression of dissatisfaction with the provided results. However, using CBIS, generally searchers can see the relevant images placed nearer to each other, which can provide them with more satisfaction.

H6: *Participants will gain more knowledge about the search topic after using CBIS than after using Google Image Search.*

The amount of knowledge the participants are gaining as they perform their assigned search activities is very difficult to estimate [27]. While it may be possible to quiz the participants at the end of the session, separating the ability of the participant to learn about the topic from the support the Web search interface provides may not be possible. As such, a simple approach was followed to measure perceived knowledge gain, where after each task, the participants rated how much knowledge they gained about the given search topic. The observation is that the concept-based query expansion technique used within the CBIS system can automatically discover different interpretations and underlying concepts for the given query, and present them in the concept hierarchy. As such, when the searcher explores the images corresponding to different interpretations and concepts, they can gain more knowledge about the query topic, which may not be the case for Google Image Search.

H7: *Participants will report that CBIS is easier to use than the Google Image Search.*

To validate this hypothesis, participants were asked to rate how easy they found it to use the interface for searching images (collected via a post-task questionnaire). The expectation is that once the searchers become familiar with CBIS, they can be more skillful in using different interaction techniques. This may result in more positive impacts on their impression on the ease of use of CBIS than Google Image Search. Although Google Image Search has a simple interface, the scrolling iteration for browsing images might not be seen as an easy way to examine them, particularly when the relevant images are distributed throughout different pages.

H8: *Participants will report that CBIS is more useful for finding images than the Google Image Search.*

To validate this hypothesis, participants were asked to rate how useful they found the interface for searching images (collected via a post-task questionnaire). The anticipation is that using CBIS the searchers' overall performance will be better (such as enhancing their productivity and effectiveness) compared to Google Image Search. As such, they may provide more positive feedback regarding the usefulness of CBIS.

H9:*Participants will prefer CBIS over Google Image Search.*

The post-task questionnaire included a question asking the participants to rank their preference for a search results interface. The expectation is that due to the overall positive impressions on CBIS, searchers might prefer this image search interface over Google Image Search.

6.4.4 Participant Demographics

Sixteen participants were recruited from undergraduate computer science and engineering courses to participate in this study. All of them were registered in first and second year courses. Further, pre-study questionnaires were provided consisting of a series of selection-based background questions on computer and image browser usage to determine their level of expertise.

From their responses, it was estimated that the participants constituted a relatively coherent group having a moderate level of expertise in using the image search software. Also it was apparent that all of them were quite familiar with the traditional grid-based image search interfaces (e.g., Google Image Search). Most of them reported that generally they searched for named entities (e.g., people, location/landmarks, car, etc.). Some others also expressed their interest about searching for general objects and themes. As such, according to the discussion on Section 2.1 regarding the Web image search behaviour, it can be assumed that the participants were representative of real-world image retrieval users.

6.4.5 Results

6.4.5.1 Time To Task Completion

The average time required to complete the four tasks using the two interfaces are illustrated in Figure 6.8. For the moderately complex tasks ("Jaguar" and "Fuji"), participants performed better using CBIS. For the simple task "Beetle", the participants took less time to find relevant images with Google Image Search, as more of the results were available on the very first page. For the other simple task "Tivoli", CBIS was superior in finding images quickly. In order to determine whether these differences in the results were statistically significant, Analysis of Variance (ANOVA) [65] tests were performed. Among these results, three were found to be statistically significant (see Table 6.3). For "Beetle" the results were not statistically significant.

The statistical results suggest that for moderately complex tasks, CBIS performed better than Google. For simple tasks, superiority of one interface over the other could not be confirmed. As a result of this study, it can be concluded that CBIS improves the searcher's ability to find images faster, particularly for complex tasks (hypothesis H1). These results are actually even better than what has been stated in hypothesis H1. That is, even with extra interactions, CBIS is not necessarily always slower than Google for simple tasks.

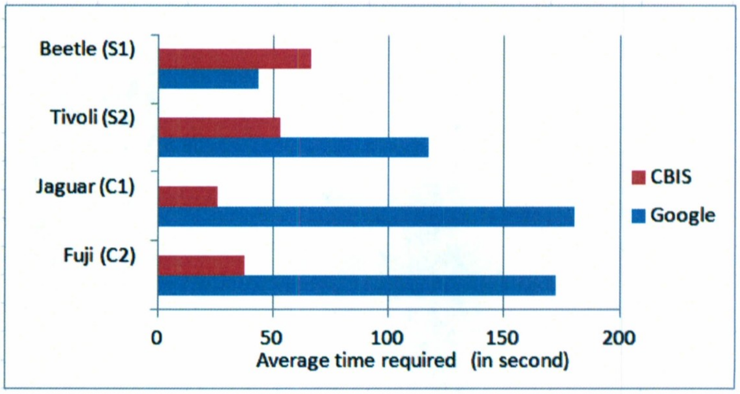


Figure 6.8: Average time to task completion measurements for the different tasks.

Table 6.3: Statistical analysis (ANOVA) of the responses for time to task completion.

Task	CBIS vs. Google
Beetle (S1)	$F(1, 14) = 1.84, p = 0.196$
Tivoli (S2)	$F(1, 14) = 13.61, p < 0.05$
Jaguar (C1)	$F(1, 14) = 39.94, p < 0.001$
Fuji (C2)	$F(1, 14) = 23.41, p < 0.001$

6.4.5.2 Simple vs. Complex Tasks

With Google Image Search, the average time required to complete the simple tasks took much less time than moderately complex tasks (80.37 and 148.87 accordingly). For CBIS, the performance with moderately complex tasks were even better than simple tasks (59.56 and 31.62 respectively). From ANOVA tests, both results were found to be statistically significant (for CBIS: $F(1, 30) = 6.044$, $p < 0.05$; for Google Image Search: $F(1, 30) = 12.059, p < 0.05$).

Based on the empirical evidence it can be concluded that using CBIS the participants were able to find images even more quickly for the moderately complex tasks than the simple ones, while for Google Image Search it was the opposite. These positive results achieved for CBIS were beyond the expectation of hypothesis H2 (no significant differences were expected between the different tasks performed with CBIS). The possible reason could be the fact that searchers were using the CBIS interface for the first time when they were assigned simple tasks. As they started the complex tasks with CBIS, they were already familiar with the interface controls and the interaction features, which made them more productive later. This illustrates a positive finding about CBIS: even though the participants were not familiar with CBIS before, they were able to learn the interactive features of this interface quickly and effectively.

6.4.5.3 Accuracy

After the participants completed the tasks, the images selected by the participants were carefully inspected to verify their relevance to the information need. Two participants abandoned the tasks when they were using Google Image Search, as they were not interested in moving to the next pages once they could not find relevant images on the top few results. As such, these two records were treated as outliers and removed in the statistical calculation. ANOVA tests across all four tasks indicate that there are no statistically significant differences in accuracy when using the different interfaces (see Table 6.4).

These results suggest that CBIS does not have much impact on the searchers' ability to decide the relevance of individual images (hypothesis H3). However, those participants who abandoned the tasks using Google Image Search indicated that they

Table 6.4: Results of the accuracy measurements and corresponding statistical analysis (ANOVA).

Task	Accuracy (Google)	Accuracy (CBIS)	ANOVA Results
Beetle (S1)	92.50 %	97.50 %	$F(1,14) = 0.737, p = 0.405$
Tivoli (S2)	82.50 %	97.50 %	$F(1,13) = 4.629, p = 0.051$
Jaguar (C1)	77.14 %	95.00 %	$F(1,14) = 2.291, p = 0.152$
Fuji (C2)	73.33 %	95.00 %	$F(1,12) = 3.760, p = 0.076$

were frustrated and unable to find relevant images. Thus, if these two outliers are treated as incorrect search results, the accuracy of Google Image Search would be lower than that of the CBIS.

6.4.5.4 Perceived Difficulty

After each task was completed, participants were asked to indicate the perceived difficulty of the tasks. The goal was to compare how this subjective measure changes for the very same tasks but using different interfaces. The average responses to this question are reported in Figure 6.9. For simple tasks, the results were mixed. For the query "Beetle", the task was perceived to be easier with Google Image Search. However, for "Tivoli" the result was the opposite. For moderately complex tasks "Fuji", and"Jaguar", it is clear that participants perceived the tasks to be easier to perform when using CBIS compared to Google Image Search. The statistical significance of these results were evaluated using pair-wise Wilcoxon-Mann-Whitney [10] tests (see Table 6.5). Significance was found for all the queries except "Beetle".

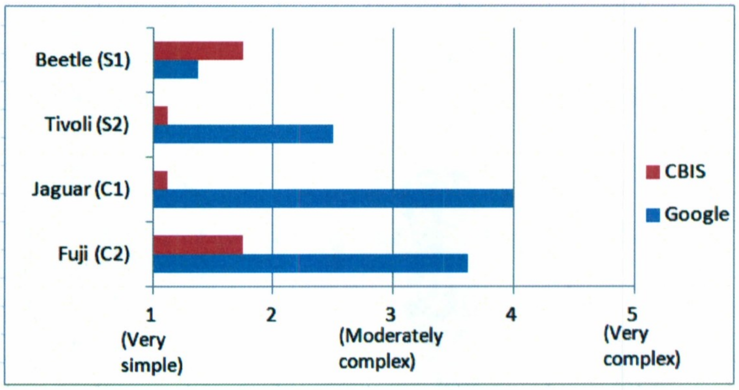


Figure 6.9: Average response to perceived difficulty for different the search tasks.

Table 6.5: Statistical analysis (Wilcoxon-Mann-Whitney tests) of the responses for perceived difficulty.

Task	CBIS vs. Google
Beetle (S1)	$Z = -.973$, $p = 0.442$
Tivoli (S2)	$Z = -3.303$, $p < 0.01$
Jaguar (C1)	$Z = -3.651$, $p < 0.001$
Fuji (C2)	$Z = -2.486$, $p < 0.05$

From these results, it can be said that for queries where the given tasks are designed based on the uncommon interpretation of the queries, the participants find it difficult to identify the relevant images that were distributed throughout different pages of Google Image Search. However, for the same queries, the amount of interaction required for CBIS was much less, for which the participants felt that the tasks are rather easier (hypothesis 4). This indicates a correlation between the perceived difficulty of the task and the complexity of the task.

6.4.5.5 Perceived Satisfaction

The *perceived satisfaction* measures compare how participants' satisfaction changes with the provided results from different interfaces for the given task. The average responses to this question are reported in Figure 6.10. It can be seen that for the simple task "Beetle", similar responses were produced for both interfaces. For the other simple task ("Tivoli"), CBIS was superior. For moderately complex tasks, the participants were more satisfied with CBIS. From Wilcoxon-Mann-Whitney tests, significance was found for all the queries except for "Beetle" (see Table 6.6).

The statistical results suggest that the subjective measures on perceived satisfaction produce almost opposite results to the perceived difficulty (i.e., low difficulty results in high satisfaction). That is, if participants can find desired images and become highly satisfied with the search results, they feel that the given tasks is easier to perform (hypothesis 5).

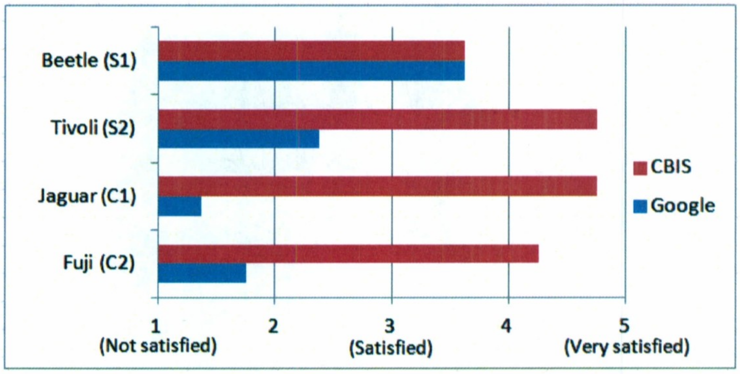


Figure 6.10: Average response to perceived satisfaction for different the search tasks.

Table 6.6: Statistical analysis (Wilcoxon-Mann-Whitney tests) of the responses for perceived satisfaction.

Task	CBIS vs. Google
Beetle (S1)	$Z = 0.000$, $p = 1.00$
Tivoli (S2)	$Z = -3.382$, $p < 0.001$
Jaguar (C1)	$Z = -3.520$, $p < 0.001$
Fuji (C2)	$Z = -2.750$, $p < 0.05$

6.4.5.6 Perceived Knowledge Gain

The *perceived knowledge gain* measures how much knowledge the participants gained about the given search topic. As can be seen from Figure 6.11, the results were in favour of CBIS for both simple and moderately complex tasks. From Wilcoxon-Mann-Whitney tests, significance was found for all the queries (see Table 6.7).

From these results, it has been found that using CBIS the participants' perceived knowledge gain about the query was improved significantly (hypothesis H6). These results support the fundamental idea of this thesis. That is, by discovering of Wikipedia concepts for a given query and utilizing them to present the search results in an organized way allow the participants to learn different perspectives about the query, which they may not aware of before. As such, in general they gain deeper knowledge of the query.

6.4.5.7 Ease of Use and Usefulness

In addition to collecting subjective reactions in the context of a particular task, overall perceptions of the usefulness and ease of use of the interface were collected in the post-

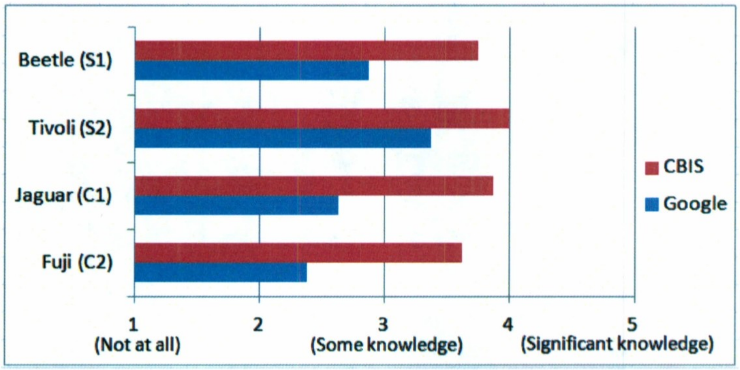


Figure 6.11: Average response to perceived knowledge gain for different the search tasks.

Table 6.7: Statistical analysis (Wilcoxon-Mann-Whitney tests) of the responses for perceived knowledge gain.

Task	CBIS vs. Google
Beetle (S1)	$Z = -1.819$, $p < 0.05$
Tivoli (S2)	$Z = -1.313$, $p < 0.05$
Jaguar (C1)	$Z = -2.674$, $p < 0.01$
Fuji (C2)	$Z = -1.976$, $p < 0.05$

study questionnaire, using the TAM instrument. Results of these questionnaires are shown in Figure 6.12.

Wilcoxon-Mann-Whitney tests were performed on the responses using a pair-wise grouping of the interfaces. Statistical significance was found for the usefulness measure but not for the ease of use measure (Usefulness: $Z = -8.874$, $p < 0.001$; Ease of Use: $Z = -0.071$, $p = 0.09$).

From these results, it has been found that for the usefulness measure, CBIS not

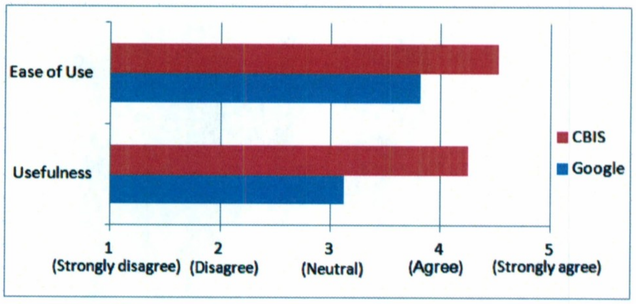


Figure 6.12: Average response to statements regarding to the perceived usefulness and ease of use of the interface.

only received a higher rating from the participants, but also the results were statistically significant (hypothesis 7). Also the participants reported a higher degree of ease of use with CBIS, although this difference did not prove to be statistically significant (hypothesis 8). It is promising that even though the participants were required to perform some interactive operations, they still found the CBIS interface at least as easy to use as Google Image Search.

6.4.5.8 Preference

At the end of the study, participants were asked to indicate their preference for an image search interface. Fifteen participants indicated their preference for CBIS (93%) over Google Image Search. A Wilcoxon signed rank test found statistical significance ($Z = -3.5$, $p < 0.001$) in the preference of CBIS over Google Image Search. This result clearly indicates the dominance of CBIS in getting the preference from the participants (hypothesis 9).

6.4.6 Discussion

In this evaluation, a user study was conducted among a group of real users in a controlled setting to validate the potential value of CBIS compared to a baseline system (Google Image Search). As a result of this study, it can be concluded that CBIS improves the searcher's ability to find images faster, particularly for the complex tasks. Another interesting finding about the time to task completion measure was the fact that using CBIS the participants were able to find images even more quickly for the moderately complex tasks than the simple ones, while for Google Image Search it was the opposite.

The possible reason for the better performance of CBIS for complex tasks is the potential learning effect. That is, the searchers were not familiar with CBIS before as opposed to the fact that they used Google Image Search quite frequently (as they mentioned in the pre-study questionnaires). As such, for the first task, they generally took more time to understand the appropriate use of the interactive features of the CBIS interface. Once they became more acquainted with CBIS, they performed significantly faster in the second task. As a result, it has been surprisingly found that with CBIS the participants took even less time for complex tasks than for simple tasks. In this context, even though it was not possible to conclude whether CBIS can perform faster than Google Image Search for every single task, an indication can be obtained that once the searcher can learn more about CBIS, their performance regarding the simple tasks can be even more improved.

From the in-task questionnaires, the searchers' impressions ware recorded with the provided image search interfaces for each given task. The perceived difficulty of the tasks may be correlated with their complexity. That is for the easier tasks where the relevant images are presented within the first page of Google Image Search, the searcher may find the task less difficult to perform, as opposed to CBIS and vice versa. Similar results have been found for the *perceived satisfaction* measure. Furthermore, it has been found that using CBIS the participants' perceived knowledge gain about the query was improved significantly.

Participants' reactions about overall usefulness, and the ease of use of the software were recorded via post-study questionnaires. From these results it has been found that participants found CBIS more useful comparing to Google Image Search. However, the results on ease of use measure were not statistically significant. One can expect that as searchers become more accustomed to the features of CBIS, this ease of use measure will increase further.

Even though the results of this study were very positive, and almost all the participants indicate their preference in favor of CBIS over Google Image Search, real-world use may result in even better results. The user study was conducted in laboratory settings, where the participants had little chance to get familiar with CBIS within the controlled environment, limiting their ability to become more skilful with the software. In real-world use, searchers can take whatever time they need to become familiar with the interface features, performing the real-world tasks according to their own needs, which can provide more valuable insights into the learnability, usability, and utility of the interface.

Chapter 7

Conclusions and Future Work

The goal of this thesis has been to address fundamental issues related to the shortcomings of current Web image search methods. To fulfil this goal, an approach for performing concept-based query expansion was introduced, for the purpose of generating a diversified set of image search results (Chapter 3). The image sets were organized based on their conceptual and visual similarities. To support the searcher in exploring the organized images a set of visualization and interaction techniques were designed (Chapter 4). An approach for the interactive query refinement was also proposed to match the searcher's intent more accurately (Chapter 5). Finally, three different evaluations were conducted to answer the different research questions related to this work (Chapter 6). The remainder of this chapter summarizes the contributions of the research work presented in this thesis, and potential future research directions.

7.1 Research Contributions

7.1.1 Concept-Based Query Expansion

The approach to query expansion proposed in this thesis uses knowledge contained within Wikipedia to extract different interpretations and related concepts to the given query, and uses these concepts to expand the query and retrieve a diverse range of images. The primarily benefit of this approach is that instead of only satisfying the information needs for the most common interpretation of the query, it provides a more balanced view of the different interpretations of the query.

A fundamental research question raised about this method was, *what is the impact of the proposed query expansion method on diversity and precision for short and ambiguous queries?* The anticipation was that broadening the query too much may result in a potentially significant drop in precision and that it may be necessary to somehow control the degree to which a query is expanded by the system. This trade-off between diversification and precision resulting from the query expansion method was evaluated, using a set of test queries of varying ambiguity (Section 6.1). The evaluation was designed such that the degree of diversification can be controlled through the diversification parameter N.

From these experiments, it has been found that the degree to which diversification is valuable depends on the level of ambiguity of the query. That is, a highly ambiguous query can benefit from a high degree of diversification more so than a very specific query. Based on this, a simple linear equation for determining the diversification parameter N was proposed based on the number of interpretations of the query. In this way, the system can dynamically control the number of concepts chosen within the query expansion process, and ensure that the trade-off between precision and diversity is well-balanced.

7.1.2 Visualization and Interaction Techniques for Exploring the Image Space

Query expansion has the side effect of introducing potentially irrelevant images within the search results, especially when a query is ambiguous and can be interpreted in multiple ways. To address this problem, a visual interface was developed that supports the searcher in exploring the image space using a set of novel interactive features. The benefits of this approach is that the images themselves are organized using a multi-resolution grid layout derived from a SOM, which not only groups conceptually and visually similar images, but also provides a solution to the problem when there are more images to show than display space allows. It allows searchers to perform concept-based filtering and focusing operations using a hierarchical representation of the concepts. Zooming into the image space results in a visual focusing and filtering operation, displaying more images as space between the images is created, and moving images that are distant from the focal point out of the viewport.

The first research questions emerging for this approach was, *Does the approach for organization and exploration of image search results improve the user performance for image search tasks?* To answer these questions, two different evaluations were performed (Sections 6.2 and 6.3). The first evaluation studied whether the interactive exploration features of the CBIS interface help to improve the precision of the search results. These experiments show that the precision calculated over both the visible and hidden images within the viewport increases as a result of concept-based focusing and filtering, as well as visual zooming operations, even for uncommon interpretations of ambiguous queries.

To further examine the system in a controlled laboratory settings, a user study was conducted. This study shows that the interactive features of CBIS make a positive impact on the searcher's ability to to find images faster particularly for the complex tasks. Also, it has been found that using this interface the participants were able to find images even more quickly for the moderately complex tasks than the simple ones, while for Google Image Search it was the opposite.

Regarding the subjective measures of the developed image search interface the research question was *Does the approach for organization and exploration of image search results improve the user's perceptions of usefulness, ease-of-use, and satisfaction while performing image search tasks?* This question was analyzed through the user study. This study shows that the interactive features of CBIS improve the user's perceptions of usefulness and the search tasks being easier when compared to a traditional image search interface. In addition, it has been found that using CBIS the participants' perceived knowledge gain about the query was improved significantly.

7.1.3 Interactive Query Refinement

For ambiguous queries, query expansion followed by image search results organization provides an effective way to support the searcher in discovering the concepts and images that they are interested in. To further support the searcher in finding images, an interactive query refinement approach was developed (Chapter 5) based on concepts or example images. Query refinement by expanding a concept allows searchers to enhance the original query by treating the selected concept as a new query, resulting in a diversification of the search results in the same way as described in Chapter 3.

Query refinement by images operates more like a narrowing operation, rather than a diversification operation. If the searcher selects a set of images that are near matches to their interests, the system will retrieve more images that are like these by using their source concepts as the queries. Since no query expansion is performed, more images from each of the selected images' concepts can be retrieved, resulting in a more focused set of search results.

Interactive query refinement was not evaluated via a user study, since it is very difficult to measure the benefits of this feature in a controlled laboratory setting. The main difficulty is that the ability to effectively refine a query is largely dependent on the individual's prior knowledge on the query topic and their ability to learn about the topic during the search task [27]. As such, rather than conducting a user evaluation of this approach, detailed illustrative examples were provided in Chapter 5, to demonstrate the potential benefits of the interactive query refinement.

7.2 Future Directions

7.2.1 Further Enriching Conceptual Information in Image Retrieval

In this thesis work, Wikipedia was used to extract different interpretations and corresponding related concepts for use in the query expansion process. While such an approach can be effective in promoting semantic diversity in the image search results, further room for enhancing semantic information within image retrieval could be possible. At the moment it is possible to discover the meaningful concepts for a given query, but we do not know how each of these concepts is related to the given query.

If we can extract additional information about the given query, it can be effectively used in the query expansion process and in visualizing the search results. For example, if the query is sensitive to location and time, and a spatio-temporal tag can be assigned to each image in the search results, this could provide valuable clues to the user in exploring and refining the query. Recently, a link-analysis based method was proposed for extracting the main locations and periods associated with all Wikipedia concepts [57]. Extracting such semantic information can be used in the process of retrieval and visualizing the image search results. An enhanced image search interface that takes advantage of this information to support the searcher in interactively exploring and filtering the search results based on the space and time dimension, can provide even more support in finding the desired images.

7.2.2 Application of Conceptual Information in Other Domains

In this thesis, the process of extracting conceptual information for a given query and utilizing it in the retrieval and presentation of image search results has been found to be effective. Although within this method there were some specific aspects that were designed for the image search application, the primary idea can still be applied to the domain of text-based retrieval applications (such as news, blogs, and video searching). This may provide a way to improve the representation of the search results, and to support exploration activities within the result set. Although recently there have been a number of research works conducted on how to support the searcher in exploring the search results and refining the queries [29], generally the underlying methods are based on a bag-of-words approach that represent the query and refining terms as individual words, which often lead to ambiguity [21].

In this context, it might be a promising direction to replace this bag-of-words approach with a concept-based approach. In a previous work, a query space was generated that represented the query terms in relation to the concepts they described and the terms that are related to these concepts. A visual representation of this query space allowed the users to interpret the relationships between their query terms and the query space, supporting their ability to refine the query [30]. Incorporating such concept-based approach in retrieval and visualization of search results might be promising in different text-based searching applications.

7.2.3 Studying Alternative Visualization Techniques for Image Search Results Exploration

Even though strong evidence has been found in favour of the the visualization techniques described in this thesis, there are still rooms for improvements. At the moment, the related concepts and their images are presented using two different visual components (the concept hierarchy and the image space). As a result, searchers may not be able to easily understand the association between each individual image and the corresponding concept. They need to scan each individual concept to find the images that are associated with the selected concepts.

To address this issue, it might be promising to examine how the concept information and images can be presented in an unified interface rather than using multiple visual components, so that the searcher can instantaneously find the images belonging to a particular concept, and then further explore the image space. The benefits and drawbacks of such an alternative visualization method could be evaluated through user evaluations.

Another potential issue is that the approach for image organization is not particularly fast. It was sufficient for the laboratory studies since the images could be downloaded in advance and the organization pre-calculated in an off-line manner. However, in order to develop a live image search system (for public use), the image organization process would need to occur in near-real time. As such, further study for speeding up the SOM-based methods is needed. However, it is presumable that this speeding up process might affect the ability of the organization method to arrange images accurately. In this context, it may be necessary to study the trade-offs in organizational accuracy vs. speed. Such a study could be done within the context of a user study, to measure the searchers' willingness to accept sub-optimal organizations of the images in exchange for returning the image search results quickly.

7.2.4 Further Evaluations

In this thesis, the benefits of the interactive exploration features were evaluated by a user study in a controlled laboratory setting. Although, laboratory studies are the most common method for evaluating information visualization systems, the ability of this method to capture the subjective measures through exploration of search results and query refinement activities are limited by the time constraints and task designs. In this context, field trials can be appropriate, wherein the evaluation is in real-world use, but with the direct participation and guidance of the researchers to assist the participants in using the system. Longitudinal studies may also help by capturing activities and impressions of the system during real-world use over extended periods of time [27]. Since image search can occur in various times, locations, and contexts, a longitudinal study may provide a better indication of the value of the approach over the breadth of web image search activities.

In this direction, a more logical way would be to follow a stepped model of evaluation [27], by developing the current refined prototype into a fully operational system. The data in longitudinal studies can be collected by log analysis and administering questionnaires. These longitudinal studies would reveal the true value (and potential problems) of the approach. Based on the findings from longitudinal studies, the system can be further developed and refined. Considering the potential for commercial value in this research, this may lead to the development of commercial-grade Web image search tools.

Bibliography

[1] ACM. ACM computing classification system. `http://www.acm.org/class/`, 2011.

[2] P. André, E. Cutrell, D. S. Tan, and G. Smith. Designing novel image search interfaces by understanding unique characteristics and usage. In *Proceedings of the IFIP Conference on Human-Computer Interaction*, pages 340–353, 2009.

[3] P. André, J. Teevan, and S. T. Dumais. From x-rays to silly putty via Uranus: serendipity and its role in Web search. In *Proceedings of the International Conference on Human Factors in Computing Systems*, pages 2033–2036, 2009.

[4] P. Anick and S. Tipirneni. Interactive document retrieval using faceted terminological feedback. In *Proceedings of the Annual Hawaii International Conference on System Sciences*, volume 2, pages 2036–2047, 1999.

[5] B. B. Bederson. PhotoMesa: a zoomable image browser using quantum treemaps and bubblemaps. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, pages 71–80, 2001.

[6] J. L. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509–517, 1975.

[7] J. Bhogal, A. Macfarlane, and P. Smith. A review of ontology based query expansion. *Information Processing and Management*, 43(4):866–886, 2007.

[8] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann. DBpedia - a crystallization point for the web of data. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 7(3):154–165, 2009.

[9] I. Borg and P. Groenen. *Modern Multidimensional Scaling: Theory and Applications.* Springer, 2005.

[10] S. Boslaugh and P. A. Watters. *Statistics in a nutshell.* O'Reilly, Sebastopol, CA, 2008.

[11] D. Cai, X. He, Z. Li, W.-Y. Ma, and J.-R. Wen. Hierarchical clustering of WWW image search results using visual, textual and link information. In *Proceedings of the ACM International Conference on Multimedia*, pages 952–959, 2004.

[12] J. Carbonell and J. Goldstein. The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 335–336, 1998.

[13] R. Datta, D. Joshi, J. Li, and J. Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys*, 40(2):1–60, 2008.

[14] F. D. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Management Information Systems Quarterly*, 13(3):319–340, 1989.

[15] T. Deselaers, T. Gass, P. Dreuw, and H. Ney. Jointy optimising relevance and diversity in image retrieval. In *Proceedings of the ACM International Conference on Image and Video Retrieval*, pages 1–8, 2009.

[16] H. Ding, J. Liu, and H. Lu. Hierarchical clustering-based navigation of image search results. In *Proceedings of the ACM International Conference on Multimedia*, pages 741–744, 2008.

[17] A. Diriye, A. Blandford, and A. Tombros. A polyrepresentational approach to interactive query expansion. In *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, pages 217–220, 2009.

[18] E. N. Efthimiadis. Query expansion. *Annual Review of Information Systems and Technology*, 31:121–187, 1996.

[19] J. Fan, D. A. Keim, Y. Gao, H. Luo, and Z. Li. JustClick: personalized image recommendation via exploratory search from large-scale Flickr images. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(2):273–288, 2009.

[20] B. M. Fonseca, P. Golgher, B. Pôssas, B. Ribeiro-Neto, and N. Ziviani. Concept-based interactive query expansion. In *Proceedings of the ACM International Conference on Information and Knowledge Management*, pages 696–703, 2005.

[21] E. Gabrilovich and S. Markovitch. Computing semantic relatedness using Wikipedia-based explicit semantic analysis. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1606–1611, 2007.

[22] B. Gao, T.-Y. Liu, T. Qin, X. Zheng, Q.-S. Cheng, and W.-Y. Ma. Web image clustering by consistent utilization of visual features and surrounding texts. In *Proceedings of the ACM International Conference on Multimedia*, pages 112–121, 2005.

[23] A. Goodrum and A. Spink. Image searching on the Excite Web search engine. *Information Processing and Management*, 37(2):295–311, 2001.

[24] D. Heesch. A survey of browsing models for content based image retrieval. *Multimedia Tools and Applications*, 42(2):261–284, 2008.

[25] D. Heesch and S. Rüger. Image browsing: A semantic analysis of NNk networks. In *Proceedings of the International Conference Image and Video Retrieval*, pages 609–618, 2005.

[26] O. Hoeber. Web information retrieval support systems: The future of Web search. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence - Workshops (International Workshop on Web Information Retrieval Support Systems)*, pages 29–32, 2008.

[27] O. Hoeber. User evaluation methods for visual Web search interfaces. In *Proceedings of the International Conference on Information Visualization*, pages 139–145, 2009.

[28] O. Hoeber and X. D. Yang. Evaluating WordBars in exploratory Web search scenarios. *Information Processing and Management*, 44(2):485 – 510, 2008.

[29] O. Hoeber and X. D. Yang. Supporting Web search with visualization. In J. Yao, editor, *Web-based Support Systems*, pages 183–214. Springer, 2010.

[30] O. Hoeber, X. D. Yang, and Y. Yao. VisiQ: Supporting visual and interactive query refinement. *Web Intelligence and Agent Systems*, 5(3):311–329, 2007.

[31] E. Hoque, O. Hoeber, and M. Gong. Evaluating the trade-offs between diversity and precision for Web image search using concept-based query expansion. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence - Workshops (International Workshop on Intelligent Web Interaction)*, pages 130–133, 2011.

[32] E. Hoque, O. Hoeber, G. Strong, and M. Gong. Combining conceptual query expansion and visual search results exploration for Web image retrieval. *Journal of Ambient Intelligence and Humanized Computing, 2011 (In Press)*, 2011.

[33] E. Hoque, G. Strong, O. Hoeber, and M. Gong. A concept-based interactive visualization approach to web image search (interactive poster). In *Conference Compendium of the IEEE Information Visualization Conference*, 2011.

[34] E. Hoque, G. Strong, O. Hoeber, and M. Gong. Conceptual query expansion and visual search results exploration for web image retrieval. In *Proceedings of the Atlantic Web Intelligence Conference*, pages 73–82, 2011.

[35] V. Jampani, G. Ramos, and S. Drucker. ImageFlow: Streaming Image Search. Technical Report MSR-TR-2010-148, Microsoft Corporation, November 2010.

[36] T. Janjusevic and E. Izquierdo. Visualising the query space of the image collection. In *Proceedings of the International Conference Information Visualisation*, pages 86–91, 2009.

[37] B. J. Jansen, A. Spink, and J. Pedersen. An analysis of multimedia searching on AltaVista. In *Proceedings of the ACM SIGMM International Workshop on Multimedia Information Retrieval*, pages 186–192, 2003.

[38] Y. Jing and S. Baluja. VisualRank: Applying PageRank to large-scale image search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(11):1877–1890, 2008.

[39] Y. Jing, H. A. Rowley, C. Rosenberg, J. Wang, M. Zhao, and M. Covell. Google image swirl, a large-scale content-based image browsing system. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, pages 267 –267, 2010.

[40] H. Joho, C. Coverson, M. Sanderson, and M. Beaulieu. Hierarchical presentation of expansion terms. In *Proceedings of the ACM Symposium on Applied Computing*, pages 645–649, 2002.

[41] D. Joshi, R. Datta, Z. Zhuang, W. P. Weiss, M. Friedenberg, J. Li, and J. Z. Wang. PARAgrab: A comprehensive architecture for Web image management and multimodal querying. In *Proceedings of the International Conference on Very Large Databases*, pages 1163–1166, 2006.

[42] M. L. Kherfi, D. Ziou, and A. Bernardi. Image retrieval from the World Wide Web: Issues, techniques, and systems. *ACM Computing Surveys*, 36(1):35–67, 2004.

[43] S. Kiranyaz and M. Gabbouj. Hierarchical cellular tree: An efficient indexing scheme for content-based retrieval on multimedia databases. *IEEE Transactions on Multimedia*, 9(1):102 –119, 2007.

[44] T. Kohonen. *Self-Organization Maps*. Springer-Verlag, Berlin, 1995.

[45] J. Kustanowitz and B. Shneiderman. Meaningful presentations of photo libraries: rationale and applications of bi-level radial quantum layouts. In *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, pages 188–196, 2005.

[46] J. Laaksonen, M. Koskela, and E. Oja. PicSOM-self-organizing image retrieval with MPEG-7 content descriptors. *IEEE Transactions on Neural Networks*, 13(4):841–853, 2002.

[47] H. Liu, X. Xie, X. Tang, Z.-W. Li, and W.-Y. Ma. Effective browsing of Web image search results. In *Proceedings of the ACM SIGMM International Workshop on Multimedia Information Retrieval*, pages 84–90, 2004.

[48] G. A. Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.

[49] D. Milne and I. H. Witten. An effective, low-cost measure of semantic relatedness obtained from Wikipedia links. In *Proceedings of the AAAI Workshop on Wikipedia and Artificial Intelligence*, pages 25–30, 2008.

[50] D. Milne and I. H. Witten. Learning to link with Wikipedia. In *Proceedings of the ACM Conference on Information and Knowledge Management*, pages 509–518, 2008.

[51] D. Milne, I. H. Witten, and D. M. Nichols. A knowledge-based search engine powered by Wikipedia. In *Proceedings of the ACM Conference on Information and Knowledge Management*, pages 445–454, 2007.

[52] P.-A. Moëllic, J.-E. Haugeard, and G. Pitel. Image clustering based on a shared nearest neighbors approach for tagged collections. In *Proceedings of the International Conference on Content-Based Image and Video Retrieval*, pages 269–278, 2008.

[53] D. Myoupo, A. Popescu, H. L. Borgne, and P.-A. Moëllic. Multimodal image retrieval over a large database. In *Proceedings of the International Conference on Cross-Language Evaluation Forum: Multimedia Experiments*, pages 177–184, 2009.

[54] G. P. Nguyen and M. Worring. Interactive access to large image collections using similarity-based visualization. *Journal of Visual Languages and Computing*, 19(2):203–224, 2008.

[55] J. Nielsen. Enhancing the explanatory power of usability heuristics. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 152–158, 1994.

[56] Z. Pecenovic, M. Do, M. Vetterli, and P. Pu. Integrated browsing and searching of large image collections. In *Proceedings of the International Conference on Advances in Visual Information Systems*, pages 279–289, 2000.

[57] A. Popescu and G. Grefenstette. Spatiotemporal mapping of wikipedia concepts. In *Proceedings of the Annual Joint Conference on Digital Libraries*, pages 129–138, 2010.

[58] M. Porta. Browsing large collections of images through unconventional visualization techniques. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pages 440–444, 2006.

[59] P. Pu and P. Janecek. Visual interfaces for opportunistic information seeking. In *Proceedings of the International Conference on Human - Computer Interaction*, pages 1131–1135, 2003.

[60] K. Rodden, W. Basalaj, D. Sinclair, and K. Wood. Does organisation by similarity assist image browsing? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 190–197, 2001.

[61] M. Sanderson. Ambiguous queries: test collections need more sense. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 499–506, 2008.

[62] R. Santos, J. Peng, C. Macdonald, and I. Ounis. Explicit search result diversification through sub-queries. In *Proceedings of the European Conference on Information Retrieval*, pages 87–99, 2010.

[63] R. L. Santos, C. Macdonald, and I. Ounis. Exploiting query reformulations for Web search result diversification. In *Proceedings of the International Conference on the World Wide Web*, pages 881–890, 2010.

[64] G. Schaefer. A next generation browsing environment for large image repositories. *Multimedia Tools and Applications*, 47(1):105–120, 2010.

[65] H. Scheffé. *The Analysis of Variance.* Wiley, New York, NY, 1959.

[66] P. Schonhofen. Identifying document topics using the wikipedia category network. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, pages 456–462, 2006.

[67] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.

[68] K. Song, Y. Tian, W. Gao, and T. Huang. Diversifying the image retrieval results. In *Proceedings of the ACM International Conference on Multimedia*, pages 707–710, 2006.

[69] R. Song, Z. Luo, J.-Y. Nie, Y. Yu, and H.-W. Hon. Identification of ambiguous queries in Web search. *Information Processing and Management*, 45(2):216–229, 2009.

[70] G. Strong and M. Gong. Browsing a large collection of community photos based on similarity on GPU. In *Proceedings of the International Symposium on Advances in Visual Computing*, pages 390–399, 2008.

[71] G. Strong and M. Gong. Organizing and browsing photos using different feature vectors and their evaluations. In *Proceedings of the International Conference on Image and Video Retrieval*, pages 1–8, 2009.

[72] G. Strong and M. Gong. Data organization and visualization using self-sorting map. In *Proceedings of the Graphics Interface*, pages 199–206, 2011.

[73] G. Strong, O. Hoeber, and M. Gong. Visual image browsing and exploration (Vibe): User evaluations of image search tasks. In *Proceedings of the International Conference on Active Media Technology*, pages 424–435, 2010.

[74] G. Strong, E. Hoque, M. Gong, and O. Hoeber. Organizing and browsing image search results based on conceptual and visual similarities. In *Proceedings of the International Symposium on Visual Computing*, pages 481–490, 2010.

[75] M. Strube and S. P. Ponzetto. WikiRelate! computing semantic relatedness using Wikipedia. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1419–1424, 2006.

[76] F. M. Suchanek, G. Kasneci, and G. Weikum. Yago: a core of semantic knowledge. In *Proceedings of the International Conference on World Wide Web*, pages 697–706, 2007.

[77] B. Taneva, M. Kacimi, and G. Weikum. Gathering and ranking photos of named entities with high precision, high recall, and diversity. In *Proceedings of the ACM International Conference on Web Search and Data Mining*, pages 431–440, 2010.

[78] R. H. van Leuken, L. Garcia, X. Olivares, and R. van Zwol. Visual diversification of image search results. In *Proceedings of the International Conference on the World Wide Web*, pages 341–350, 2009.

[79] R. van Zwol, V. Murdock, L. Garcia Pueyo, and G. Ramirez. Diversifying image search with user generated content. In *Proceeding of the ACM International Conference on Multimedia Information Retrieval*, pages 67–74, 2008.

[80] R. van Zwol, B. Sigurbjornsson, R. Adapala, L. Garcia Pueyo, A. Katiyar, K. Kurapati, M. Muralidharan, S. Muthu, V. Murdock, P. Ng, A. Ramani, A. Sahai, S. T. Sathish, H. Vasudev, and U. Vuyyuru. Faceted exploration of image search results. In *Proceedings of the International Conference on World Wide Web*, pages 961–970, 2010.

[81] B. Vélez, R. Weiss, M. A. Sheldon, and D. K. Gifford. Fast and effective query refinement. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 6–15, 1997.

[82] S. Wang, F. Jing, J. He, Q. Du, and L. Zhang. IGroup: presenting Web image search results in semantic clusters. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 587–596, 2007.

[83] C. Ware. *Information Visualization: Perception for Design.* Morgan Kaufmann, 2004.

[84] Wikimedia. English Wikipedia dumps in SQL and XML. `http://dumps.wikimedia.org/enwiki/`, 2010.

[85] J. Yao and Y. Yao. Web-based information retrieval support systems: building research tools for scientists in the new information age. In *Proceedings of the IEEE/WIC International Conference on Web Intelligence*, pages 570–573, 2003.

[86] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 401–408, 2003.

[87] C. X. Zhai, W. W. Cohen, and J. Lafferty. Beyond independent relevance: methods and evaluation metrics for subtopic retrieval. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 10–17, 2003.

[88] L. Zhang, L. Chen, F. Jing, K. Deng, and W.-Y. Ma. EnjoyPhoto: a vertical image search engine for enjoying high-quality photos. In *Proceedings of the ACM International Conference on Multimedia*, pages 367–376, 2006.

Appendix A

User Study

This appendix includes the formal approval received from the Interdisciplinary Committee on Ethics in Human Research (ICEHR) for the study. All the evaluation documents are provided as well.

MEMORIAL UNIVERSITY

Interdisciplinary Committee on Ethics in Human Research (ICEHR)

Office of Research · IIC2010C
St. John's, NL Canada A1C 5S7
Tel: 709 864-2561 Fax: 709 864-4612
www.mun.ca/research

ICEHR Number:	**2012-194-SC**
Approval Period:	**September 8, 2011 – September 30, 2012**
Sponsor:	-
Responsible Faculty:	Dr. Orland Hoeber Department of Computer Science
Title of Project:	***Concept-based image search***

September 8, 2011

Mr. Enamul Hoque
Department of Computer Science
Memorial University of Newfoundland

Dear Mr. Hoque:

Thank you for your email correspondence of September 7, 2011 addressing the issues raised by the Interdisciplinary Committee on Ethics in Human Research (ICEHR) concerning the above-named research project.

The ICEHR has re-examined the proposal with the clarification and revisions submitted and is satisfied that concerns raised by the Committee have been adequately addressed. In accordance with the *Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans (TCPS2)*, the project has been granted *full ethics clearance* for one year from the date of this letter.

If you intend to make changes during the course of the project which may give rise to ethical concerns, please forward a description of these changes to Mrs. Brenda Lye at blye@mun.ca for the Committee's consideration.

The TCPS requires that you submit an annual status report on your project to the ICEHR, should the research carry on beyond September 30, 2012. Also to comply with the TCPS2, please notify us upon completion on your project.

We wish you success with your research.

Yours sincerely,

M Shute

Michael Shute, Th.D.
Chair, Interdisciplinary Committee on
Ethics in Human Research

LF/bl

copy: Supervisor – Dr. Orland Hoeber, Department of Computer Science

MEMORIAL UNIVERSITY St. John's Campus

Department of Computer Science

St. John's, NL Canada A1B 3X5
Tel: 709 737 8627 Fax: 709 737 2009
cs.compsci@mun.ca www.mun.ca/computerscience

Informed Consent by Subjects to Participate in
User Evaluations of Concept-Based Image Search (CBIS)

I understand that this form and the information it contains are given to me for my protection and full understanding of the procedures of this research. My signature on this form signifies that I have received a copy of this consent form, that I understand the procedures to be used in this study, and the personal risks and benefits to me in taking part. I voluntarily agree to participate in this project. I understand that I may withdraw my participation in this study at any time, and that my decision to participate in this study, and my subsequent involvement in it, will have absolutely no bearing on any other dealings I have with Mr. Hoque, Dr. Gong, or Dr. Hoeber.

Knowledge of my identity is not required. I will not be required to write my name or any identifying information on the research questionnaires. My activities will be video and audio recorded for analysis purposes and the comments I make relevant to the assigned tasks or to the use of the software will be transcribed. The analysis of the collected data will be used in research publications (thesis, journal articles, conference papers, conference presentation, etc.) by the investigators. However, the original raw data will only be accessed by the investigators. All research materials will be held confidential by the Principle Investigator and kept in a secure on-campus location and on password-protected computers for a period of five years.

I agree to participate in this user evaluation by completing four image search tasks using two different image search interfaces. I will be given a description of the types of images I am to find, and will use the assigned interface to explore the search results. My activities will be recorded, and I will be asked to explain what I am doing and what I am thinking while performing the tasks. Three different sets of questionnaires will be administrated in this study: a pre-study questionnaire to get insight about my image search and browsing behaviour, in-study questionnaires following each task to evaluate what new knowledge I have gained in the process of performing each search, and a post-study questionnaire to capture my feelings and experiences with using the interfaces. I understand that these activities will require approximately 40 minutes, and will be conducted in the User Experience Lab in the Department of Computer Science at Memorial University. I understand that I will be compensated $10 for participation in this study, regardless of my performance or ability to complete the tasks.

The proposal for this research has been reviewed by the Interdisciplinary Committee on Ethics in Human Research and found to be in compliance with Memorial University's ethics policy. If you have ethical concerns about the research (such as the way you have been treated or your rights as a participant), you may contact the Chairperson of the ICEHR at icehr@mun.ca or by telephone at 864-2861.

I may obtain copies of the results in this study, upon completion, by contacting Dr. Hoeber, in care of the Department of Computer Science, Memorial University.

NAME (please print legibly): ______________________________

SIGNATURE : ______________________________

INVESTIGTOR: ______________________________

DATE: ______________________________

Investigators:

Mr. Enamul Hoque
M.Sc. Student
Department of Computer Science
Memorial University
enamulp@mun.ca

Dr. Minglun Gong
Associate Professor
Department of Computer Science
Memorial University
gong@mun.ca

Dr. Orland Hoeber
Assistant Professor
Department of Computer Science
Memorial University
hoeber@mun.ca

PRE-STUDY QUESTIONNAIRE

Please answer the following questions in regards to your background. Circle the answer the best describes you or your opinion.

1. How many hours a week do you use a computer?

0 1-5 6-10 11-20 21+

2. How do you prefer to interact with your computer?

Mouse (e.g. Click and Drag)	Keyboard (e.g. Shortcuts and Commands)	Other (Specify: ____________)

3. How many times a week do you search images on the Web (e.g., Google, Bing, Flickr, Facebook, etc.)?

0 1-5 6-10 11-15 15+

4. What type of image browser do you use most?

Thumbnails (e.g. Google)	Slideshow (e.g. Facebook)	Other (Specify:____________________)

5. When the results do not fulfil your information need, how often do you modify the original query?

Seldom Sometimes Often Always

6. What are kind of images you are generally searching for in Web?

Named entity (people, location/landmarks, car, etc.)	General Objects/ Themes	Other (Specify:_____________)

Task Scenario 1

Imagine that you are a big fan of Volkswagen Beetle car. And, your particular interest is in the Volkswagen cabriolet (convertible) car. Please find any **five** images of the **Volkswagen cabriolet car** by browsing in the given search interface.

When you find a relevant image, please indicate it to the corresponding investigator, who will note it on the form below.

____

____

____

____

____

In-Task Questionnaires

1. How much difficult do you think the search task was?

1	2	3	4	5
(Very simple)		(Moderately complex)		(Very complex)

2. How much are you satisfied with the provided search results?

1	2	3	4	5
(Not satisfied)		(Satisfied)		(Very satisfied)

3. How much did you gain more knowledge about the given search query after evaluating the search results:

1	2	3	4	5
(Not at all)		(Some knowledge)		(Significant knowledge)

Task Scenario 2

You are interested in knowing different models of Jaguar Cars. One of luxury model of Jaguar Car is S-Type. Find any **five** images of the **Jaguar S-Type model car** by browsing in the given search interface.

When you find a relevant image, please indicate it to the corresponding investigator, who will note it on the form below.

____

____

____

____

____

In-Task Questionnaires

1. How much difficult do you think the search task was?

1	2	3	4	5
(Very simple)		(Moderately complex)		(Very complex)

2. How much are you satisfied with the provided search results?

1	2	3	4	5
(Not satisfied)		(Satisfied)		(Very satisfied)

3. How much did you gain more knowledge about the given search query after evaluating the search results:

1	2	3	4	5
(Not at all)		(Some knowledge)		(Significant knowledge)

Task Scenario 3

Imagine that you are a tourist and you would like to visit an ancient Italian town named **Tivoli**. Before you tour there you want to know about some attractive landmarks to visit. One of such interesting landmark is Villa d'Este. Your task here is to find **five** highly relevant images of **Villa d'Este** by browsing in the given search interface.

When you find a relevant image, please indicate it to the corresponding investigator, who will note it on the form below.

____

____

____

____

____

In-Task Questionnaires

1. How much difficult do you think the search task was?

1	2	3	4	5
(Very simple)		(Moderately complex)		(Very complex)

2. How much are you satisfied with the provided search results?

1	2	3	4	5
(Not satisfied)		(Satisfied)		(Very satisfied)

3. How much did you gain more knowledge about the given search query after evaluating the search results:

1	2	3	4	5
(Not at all)		(Some knowledge)		(Significant knowledge)

Task Scenario 4

Imagine that you are a tourist and you would like to visit Fuji, a city situated in Japan. There is a famous race track situated in this city named Fuji Speedway. **Find any** five **images of** Fuji Speedway **by browsing in the given search interface.**

When you find a relevant image, please indicate it to the corresponding investigator, who will note it on the form below.

____

____

____

____

____

In-Task Questionnaires

1. How much difficult do you think the search task was?

1	2	3	4	5
(Very simple)		(Moderately complex)		(Very complex)

2. How much are you satisfied with the provided search results?

1	2	3	4	5
(Not satisfied)		(Satisfied)		(Very satisfied)

3. How much did you gain more knowledge about the given search query after evaluating the search results:

1	2	3	4	5
(Not at all)		(Some knowledge)		(Significant knowledge)

POST-TASK QUESTIONNAIRE

The following questions relate to your experiences using **Google** as your Web image search engine. Your answer to these questions will allow for a more accurate analysis of the data collected during the study.

INSTRUCTIONS: Please rate how strongly you agree or disagree with each of the following statements by circling the appropriate number.

n/a	strongly disagree	Disagreed	neutral	agree	strongly agree	
0	1	2	3	4	5	Using **Google** in my Web image searching enabled me to accomplish tasks more quickly.
0	1	2	3	4	5	Using **Google** improved my Web image search performance.
0	1	2	3	4	5	Using **Google** in my Web image searching increased my productivity.
0	1	2	3	4	5	Using **Google** enhanced my effectiveness when searching images the Web.
0	1	2	3	4	5	Using **Google** made it easier to search images.
0	1	2	3	4	5	I found **Google** useful for searching images.

n/a	strongly disagree	Disagreed	neutral	agree	strongly agree	
0	1	2	3	4	5	Learning to operate **Google** was easy for me.
0	1	2	3	4	5	I found it easy to get **Google** to do what I wanted it to do.
0	1	2	3	4	5	My interaction with **Google** was clear and understandable.
0	1	2	3	4	5	I found **Google** to be flexible to interact with.
0	1	2	3	4	5	It was easy for me to become skillful at using **Google**.
0	1	2	3	4	5	I found **Google** easy to use.

n/a	never	seldom	some-times	often	always	
0	1	2	3	4	5	The interface controls in **Google** are easy to use.
0	1	2	3	4	5	Organization of search results in paged grids is useful for finding the images.
0	1	2	3	4	5	Going through the images sequentially is useful in finding the desired images.
0	1	2	3	4	5	Text snippets about the images are useful in finding the desired images
0	1	2	3	4	5	For ambiguous query, **Google** helps me in finding my desired images.

POST-TASK QUESTIONNAIRE (CONTINUE)

The following questions relate to your experiences using **Concept-Based Image Search (CBIS)** as your Web image search engine. Your answer to these questions will allow for a more accurate analysis of the data collected during the study.

INSTRUCTIONS: Please rate how strongly you agree or disagree with each of the following statements by circling the appropriate number.

n/a	strongly disagree	Disagreed	neutral	agree	strongly agree	
0	1	2	3	4	5	Using **CBIS** in my Web image searching enabled me to accomplish tasks more quickly.
0	1	2	3	4	5	Using **CBIS** improved my Web image search performance.
0	1	2	3	4	5	Using **CBIS** in my Web image searching increased my productivity.
0	1	2	3	4	5	Using **CBIS** enhanced my effectiveness when searching images the Web.
0	1	2	3	4	5	Using **CBIS** made it easier to search images.
0	1	2	3	4	5	I found **CBIS** useful for searching images.

n/a	strongly disagree	Disagreed	neutral	agree	strongly agree	
0	1	2	3	4	5	Learning to operate **CBIS** was easy for me.
0	1	2	3	4	5	I found it easy to get **CBIS** to do what I wanted it to do.
0	1	2	3	4	5	My interaction with **CBIS** was clear and understandable.
0	1	2	3	4	5	I found **CBIS** to be flexible to interact with.
0	1	2	3	4	5	It was easy for me to become skillful at using **CBIS**.
0	1	2	3	4	5	I found **CBIS** easy to use.

n/a	never	seldom	some-times	often	always	
0	1	2	3	4	5	The interface controls in **CBIS** are easy to use.
0	1	2	3	4	5	Organizing conceptually similar images together in **CBIS** was useful in finding the desired images.
0	1	2	3	4	5	*Concept hierarchy* in **CBIS** was useful in gaining more knowledge about the query topic
0	1	2	3	4	5	Highlighting (focusing) images by concepts in was useful in exploring image search results
0	1	2	3	4	5	Zooming and panning into an area of interest was useful in finding image search results

POST-TASK QUESTIONNAIRE (CONTINUE)

Please rank your search interface preference (order from 1 to 2):
___________ Google
___________ Concept-Based Image Search Interface

Please make any other comments about the image search interface and/or the search tasks.

____

____

____

____

____

____

____

____

____