Improving user interaction with image collections

Christian Hartvedt
Dept. of Information science and Media studies, University of Bergen

Abstract

To date, research in image retrieval has focused on the underlying retrieval algorithms, while little research has been focused on the many aspects pertaining to user expectations, user actions and user experiences. It is here postulated that ultimately, it is the user who decides the quality of an image retrieval system based on a match between his/her expectations and system performance. This paper presents a study of user expectations and experiences from query formulation in image retrieval. The study used an image retrieval prototype that supports text-based and content-based image retrieval, as well as a combination of the two. Results indicate that users are able to make use of a range of image retrieval tools, that content-based query formulation is not well understood, and that the different approaches to image retrieval may be useful in different situations depending on if the users know what they are looking for.

Introduction

The number of searchable images available are now in the hundreds of millions and the amount is constantly growing [1]. Important image collections made available on the Internet include Museum collections, city/tourist information, personal albums (e.g., Flickr, photobucket, piczo, Picasa web album and photo.net) and more general-purpose image collections (e.g., Google Images and Yahoo! Images).

One of the most profound challenges in image retrieval is the problem referred to as the semantic gap [2]. The semantic gap represents the mismatch between the semantic content visualized by user requests and the capabilities of the image retrieval systems. It is the belief of this author that two important factors contribute to the semantic gap:

1. Users often cannot formulate their queries in a manner that may be used effectively by the system, and
2. Image retrieval systems commonly lack an understanding of the user intentions behind a query. The latter challenge has been referred to as an *intention gap* [3, 4]

As an example of these two challenges, consider the structure depicted in figure 1. This sculpture is placed in a park in the city center. However, no information about the artwork can be found on or around it.

Figure 1 - The Saeverud monument

This paper was presented at the NIK 2010 conference. For more information see http://www.nik.no/

If having no information to work with in searching the web for information about the structure using a text-based query, possible search words could be: "*rusty artwork Bergen*". However, these words will not return any images of the structure if using for example Google images or Flickr[1].

This sculpture is commonly referred to as "*The Saeverud monument*", in memory of the famous Norwegian composer, but is formally named "*Three Sculptures of four Arcs of 233,5' each*". The artwork and information about the sculpture and the sculptor, Bernar Venet, can be found on a site run by the city of Bergen[2]. With this information, the sculpture can be found both in Google images and Flickr.

From the example above, it becomes clear that finding images is relatively easy if you know what you are looking for *and* if you are able to specify your information need using text. However, if this is not the case, then search engines relying on the use of text queries are not particularly useful.

The content of an image may be useful in its own right as a query term in situations where an information need cannot easily be formulated using text, but in order to get a retrieval system to perform satisfactorily, the system must be able to recognize and understand the depicted content of an image. In addition, systems must also be able to deduce what kind of information may be of interest to the person submitting a query.

Addressing user requirements

In terms of focus on functionality, it has been a common trend that new image retrieval systems have been proposed as improvements on previously proposed techniques and algorithms [5]. As such, much of the development within the field of image retrieval has been what Jörgensen [6] refers to as system-driven rather than user-driven. A system-driven approach tends to overlook some key aspects:

- Different users could be interested in different aspects of the same images. These differences could manifest themselves in the words used to describe/query for the items. It is thus valuable to have a diversity of annotations and to establish an effective mechanism to collect and store these keywords for subsequent use through user interaction, and
- It is important to create functionality to be used to process the returned result set through user interaction, hence facilitate support for active use of available information to filter the results according to users information need in order to find relevant images.

Related Work

Before discussing how the use of available information and user interaction can be combined with traditional approaches to image retrieval to help improve user interaction with image collections, some of the work on which this approach is built is briefly presented.

Image Retrieval

The two most common approaches to retrieving images are Text Based Image Retrieval (TBIR) and Content Based Image Retrieval (CBIR). TBIR, in which text annotations and descriptions are used for image retrieval, appeared in the mid 1980s [7]. CBIR, in which the colors, textures and shapes of the search image are compared with the same

[1] Searches performed October 1st, 2010
[2] http://www.skulpturibergen.no/

features of stored images, appeared in the mid 1990's [8].In image retrieval from the Internet, TBIR accounts for most image searches [9].

Two potential problems with text annotations are that they may provide only a limited coverage of the possible semantic interpretations of the image content, and that they are commonly biased because of human perception and subjectivity [10, 11]. This is not to say that it is not being worked on solutions. Amongst the efforts to tackle this challenge are so-called image-tagging games [e.g. 12, 13] and folksonomies [e.g. 14]. See [15] for an overview.

Also, a prerequisite for good TBIR is that users are able to formulate their information need using text. However, in many situations, this may be a quite challenging task due to the limited expressiveness of keyword-based query formulations [16]. In fact, in order to achieve a satisfactory result, textual queries may need a large number of modifications even by professional users [17]. This may be symptomatic of users' lack of understanding of query formulation and query strategies [18].

Much research effort has also been put into CBIR during the last two decades [1, 11, 19]. The CBIR approach may be convenient in situations where textual descriptions are hard or impossible to create, but CBIR lacks support for image retrieval based on high-level semantic concepts. To date, not much is known about the actual use searchers make of content- based functionalities in image searches [18].

Combining TBIR and CBIR

TBIR may return images reflecting the semantic information need but that do not necessarily resemble the desired outcome visibly. On the other hand, if using a CBIR approach alone, the user may get images that are visually similar to an image used as a query but that have little semantic similarity [20].

Research results from the field of image retrieval indicate that a combination of text and image contents may be a viable approach to improve the quality of image result sets [20-25]. This line of thought is also apparent in [26, 27], who propose creating image descriptions by combining low-level image features and high-level semantic information. It has also been found that even though CBIR methods may have limited utility on their own, having multiple access strategies to visual content is beneficial [28].

It has been demonstrated that people are able to formulate visual image queries that combine visual and textual conditions [18], but differences in the willingness of experts and novices to engage in visual searches have also been found [29]. McDonald and Tait [30] found that users that used visual similarity searches were mostly employed in visually cued tasks, while textual similarity was used in abstract tasks. From this, it seems as if text and image content cover different aspects of an information need.

Context

In applications where the management and retrieval of images are based on using low-level features, the value of context has been recognised for a long time. For example, in Computer Vision, context played a central part in image management and understanding already in the early 1990's [31].

Numerous definitions of the concept of context are discussed in Dey et al. [32], who find them to be too specific. They then propose an extension that has been used extensively. This definition states that context is *any information that can be used to characterize the situation of an entity*. Here, an entity may be a person, a place, or an object that is considered relevant to the interaction between a user and an application, including the user and applications themselves [32].

A different definition of context is presented in Elgesem and Nordbotten [33]. Although the authors generally agree with the definitions given in Dey et al., their approach differs in an important way. Elgesem and Nordbotten attempt to move the focus away from being solely on characterizing situations of entities important to the interaction between users and applications. Instead, the authors focus on the users and applications themselves, and especially on what information must be available to both in order for *communication* between them to succeed [33].

These two lines of thought are shared here, and in the following, context information is defined as:

the information available to user and system that can help to identify images of interest to the user, and that can be communicated back and forth to help satisfy the user's information need.

Examples of the kinds of context that are available to the system and can be useful include: the geographic location of the source image [see for example 34], the season and time of day and annotations of stored images that are deemed relevant by the user to his/her information need.

Context-focused interaction is a form of man-machine interaction in which users make use of context information both as query terms and as a tool for filtering results. Some of this information is made available to the user by the system as a response to actions performed by the user.

Focus on the User in Image Retrieval

A number of researchers have noted that user-centered evaluation techniques in image and information retrieval may be advantageous [35-38]. Draper [35] suggested that one advantage associated with taking a user-centered approach to system evaluation was that users could be observed as they used a system instead of having researchers relying on their own expectations on how users would use a given system.

In the field of image retrieval, the user has not received a lot of attention [39]. User interaction is commonly limited to submitting queries and giving feedback if the system has relevance feedback functionality. However, it has been found that in image retrieval in general, and CBIR in particular, interaction is a complex interplay between the user, the images, and their semantic interpretations [40].

Observational studies indicate that most searchers making use of text-based search engines on the Internet rarely take advantage of the available interactive relevance feedback opportunity [41, 42]. If this is correct, it poses somewhat of a paradox as the relevance feedback process has been thought of as being very useful in providing users with more relevant results. Research by Crucianu et al. [43] and Lew et al. [39] indicate that a reluctance to make use of relevance feedback functionality also applies to search engines developed for CBIR.

User interaction in image retrieval

It is the belief of this author that the problems associated with the less than optimal performance of image retrieval systems are a complex mix of human and technical factors. From this, it is also believed that image retrieval algorithms should be combined with other kinds of tools in order to improve user interaction with image collections. Active use of context information combined with interaction and communication between a system and its users are tools that may help improve the situation.

The user study presented here was carried out in order to gather information that may be important when designing a user-centered image retrieval system supporting context focused interaction. The first objective was to determine how well an existing image retrieval system that supports combined TBIR/CBIR performed with regard to result set precision. The second objective was to investigate how users experienced such a system. In order to assess the performance and perceived usability of TBIR/CBIR, the study also included tasks using only TBIR or CBIR.

Experiment design

Choice of Retrieval system

The image retrieval system (IRS) needed for this study required support for multiple query forms. As the main purpose of the experiment was to record user expectations and experiences from query formulation and also to be able to monitor the IRS' performance, the VISI (VORTEX Image Search Interface) system[3], developed by the CAIM project[4] was chosen for the experiment.

The version of the VISI system used in this study consists of an Adobe ColdFusion application running on top of an Oracle 10g database system. VISI supports TBIR in a traditional manner in that users submit their search terms using text in the form of keywords. Support for CBIR in VISI is in the form of query by example (QBE). Users performed CBIR searches by uploading example images to be used by the retrieval engine. Combined TBIR/CBIR in VISI is based on the approach presented in [25] and is performed as two separate queries. The two separate result sets (from TBIR and CBIR respectively) are ranked according to their respective relevance scores. A new combined similarity score is calculated from the results returned from each query, and images are re-ranked according to this score and then presented to the user.

Participants, Tasks and Questionnaire

Fifteen student volunteers participated in the study, seven females and eight males. Their ages ranged between 20 and 30 years, averaging 23.4. All participants used all the three approaches to image retrieval described above. The participants in the study were all students from the city of Bergen, and the majority of them were bachelor students from the department of information science and media studies in the University of Bergen.

Prior to the actual experiment, each participant was given a general description of image retrieval, covering both TBIR and CBIR, as well as a general description of the test system.

The three experimental tasks used in the study were:

- Use of text to search for a specific object whose name presumably was not known to all of the students participating in the study
- Use of CBIR to search for one specific object out of two in an example image, and
- Use of TBIR to support the CBIR from the second task

Since the tasks have increasing complexity, they were given in the above order.

One purpose of these tasks was to make participants think about the degree they felt that the functionality implemented in the test system supported them in solving specific

[3] http://bulmeurt.uib.no:8500/caim/VISI2/

[4] CAIM is a collaboration project between the Universities of Bergen, Tromsø and Trondheim, and Telenor R&I.http://caim.uib.no/index.shtml

tasks where it was hard to express their information need. The tasks were presented to the participants in a target specific form in that the participants were given instructions on what the target images from each search should look like. This was done in an effort to help create a concrete information need for each task in order to hopefully avoid some of the problems commonly associated with query based approaches. These problems are especially related to the query formulation in situations where users only have a vague idea of what they are looking for [44].

The image collection used in this study consisted of a test collection currently containing about 500 images displaying various buildings and structures in the city of Bergen. Other materials used in the experiment were a desktop computer and a questionnaire.

The questionnaire used to collect user responses started with general questions about the participants' previous experience with image retrieval. Prior to each image retrieval task, the participants were asked to rate to what extent they believed TBIR, CBIR and TBIR/CBIR respectively could aid in finding relevant images for the task at hand. They were also asked to rate how difficult they anticipated it would be to formulate a query to solve the task at hand. After completing each task, each participant was asked to rate to what extent the results met their expectations. The participants were also asked to rate how hard they felt it was to specify their information needs using text, image content, and a combination of the two respectively. The participants gave their responses using a 10-point scale.

During the experiment, the participants were encouraged to talk out aloud and ask any questions they might have during the experiment. The questionnaire also allowed for the participants to write down their comments and/or questions as they completed the different tasks.

Results

Previous Experience with Image Retrieval

All but one of the participants reported having used TBIR more than fifteen times prior to participating in the study. The last participant had never used search engine technology to retrieve images. Eleven of the participants claimed to have previous knowledge about CBIR, but only two of the participants had previously used a search engine developed for CBIR (one had tried an earlier version of the test system, while another participant had used the TinEye[5] search engine).

User expectations and experiences using TBIR

In the first task, the participants used TBIR to find more images of an image presented to them. The image used in the first task was the Saeverud monument as depicted above. Only two out of the fifteen participants knew the name of the monument. Of the remaining thirteen, one participant guessed the object to be a monument, but did not know the name of the object. The rest of the students did not have any idea of what the name of the object was. User expectations and experiences with TBIR are shown in table 1.

[5] http://tineye.com/

	TBIR expectations	Anticipated difficulties with TBIR queries	Met TBIR expectations	Experienced difficulties with TBIR queries
Mean value of responses on a 10-point scale	6,1	4,8	4,4	5,2
Std. Deviation	1,19	1,97	3,54	3,10

Table 1 - User expectations and experiences with TBIR

The average Precision value for all participants was 20 % for the first ten images in the result set. However, only three participants were able to retrieve relevant images (10 each).

One interesting result from this task is that when the participants who did not know the name of the object they were looking for, they instead specified their information need referring to what they perceived the object to *be*, e.g. sculpture (4), statue (4), Art/Work of art (3), Circle (1) or monument (1). However, as indicated by the responses on expectations, it seems that the participants still had high expectations that the system would retrieve images of the monument.

The results indicate that the participants expected the system to retrieve images of the monument even if the query terms were vague.

User expectations and experiences using CBIR

In the second task, the participants were asked to submit two queries with images selected from a set of example images. The participants could choose between the four different example images displayed in figure 2, each depicting at least two distinct structures and/or objects in the city of Bergen. The participants recorded their chosen images and objects in the questionnaire, but they had no means of marking/selecting the object they were looking for when submitting the image to the image retrieval system.

Figure 2 - Images used in the second and third tasks in the experiment

Thus, the participants used CBIR to try to find images depicting only one of the objects depicted in each of the two selected example images without being able to indicate to the system which of the objects they were actually interested in. The main purpose of the CBIR task was to investigate if the participants intuitively would understand that it would be very hard for the system to accommodate their request, the reason being that CBIR is based on retrieving similar images and thus would retrieve the images most similar to the example image, i.e. containing both objects. User expectations and experiences from the CBIR task are shown in table 2.

	CBIR expectations	Anticipated difficulties with CBIR queries	Met CBIR expectations Image1	Met CBIR expectations Image2	Experienced difficulties with CBIR queries
Mean value of responses on a 10-point scale	5,7	6,0	3,3	3,5	5,6
Std. Deviation	1,79	2,56	1,87	2,66	2,82

Table 2 - User expectations and experiences with CBIR

The results indicate that the participants were less confident in the capabilities of CBIR compared to TBIR. However, the results also indicate that the participants' expectations with regard to how hard it would be to formulate a query using and image were only a little higher than when using text.

The average Precision value for all participants with regard to finding images depicting only the object chosen by the participant was 6 % and 7 % for query one and two respectively. This fall in Precision was expected as attempting to retrieve only one object from an example image depicting several objects is disadvantageous for the CBIR approach. However, only two of the participants noted this in their responses, pointing out that: "I would crop the query image to include only the object I chose", and "it's hard with regards to explaining to the system which part of the image I'm actually interested in". The responses from the rest of the participants indicate that many of them appeared to expect that the system would be able to *auto-magically* interpret their vague query formulation.

This tendency is visible in the mean values with regard to what degree participants felt that the results provided by the test system met their expectations. Hence, the participants actually expected much more than the system could deliver. Another indication of unrealistic expectations to CBIR functionality in general was that most of the participants reported having little trouble formulating queries, even if the task required finding only a subset of the example image.

With regard to images used in the second task, it was assumed that the participants would choose image objects at random as this was an artificially created information need. However, ten of the fifteen participants actually reported a specific reason for choosing the particular objects they chose. The perhaps most interesting aspect here is that nine out of the ten reasons converged around three categories of reasons. The ten participants reported that they chose the particular objects because:

- The objects caught their interest or were perceived as prominent in the image (4)
- Because of the objects' specific colors and shapes (3)
- The objects were seen as "good"/charming, and well known (2)
- The objects were perceived as being the most challenging for the system to find (1)

User expectations and experiences using TBIR/CBIR

In the third task, the participants used the same images as they used in the second task. However, in this task, the participants used both TBIR and CBIR to find images depicting one of the objects depicted in the selected example image. User expectations and experiences with TBIR/CBIR are shown in table 3.

	TBIR/CBIR expectations	Anticipated difficulties with TBIR/CBIR queries	Met TBIR/CBIR expectations Image1	Met TBIR/CBIR expectations Image2	Experienced difficulties with TBIR/CBIR queries
Mean value of responses on a 10-point scale	6,9	3,5	5,1	4,9	2,2
Std. Deviation	1,40	1,69	2,69	2,93	2,10

Table 3 - User expectations and experiences with TBIR/CBIR

The results indicate that the participants were more confident in combining TBIR with CBIR compared to CBIR alone, but less confident in this approach than TBIR alone. However, the results also indicate that the participants' expected it to be easier to formulate a query using TBIR combined with CBIR than each of the two other approaches respectively.

The Precision value of number of images retrieved depicting only the object chosen by the participants when combining TBIR and CBIR was 40 % for the first query and 30 % for the second query. Although still a low value, this is by far the highest degree of precision recorded in the study.

Interestingly, participants were able to retrieve images depicting objects of which they did not know the name. However, the tendency observed in the first task (participants that did not know the name of the object specified their information need referring to what they perceived the object to be) was also apparent here.

The TBIR/CBIR approach to image retrieval also achieved the highest score with regard to meeting user expectations. The participants also reported that formulating queries was less difficult using TBIR/CBIR compared to TBIR and CBIR respectively. However, because of the way the experiment was set up, it cannot be ruled out that some of this result was the consequence of a learning effect.

Results from the third task of the experiment show that out of the fifteen participants, only one of them actively made use of information presented in the result set in the second task as a query term in the third task. This came as a surprise since the interface of the test system displays an image caption together with each image in the result set. When asked about this, most participants answered that they focused mainly on the images in the result set and hence overlooked the text-based information.

Conclusions from the Pilot Study

Overall, the results from this study indicate that an image retrieval system that combine TBIR with CBIR performs better than the same system's TBIR or CBIR algorithms respectively in terms of Precision. In addition, responses from the participants indicate that the combination approach may be particularly useful in situations where a user has difficulties with formulating a text-based query, e.g. does not know the name of the object of interest. However, responses from the task where the participants searched for images using CBIR also suggest that the way in which this technology actually works is indeed not well understood.

The responses given by participants on their experiences from using the test system suggest that users found the image retrieval tools easy to use. This is consistent with results presented in [18]. However, in both of the tasks involving TBIR, the participants formulated their queries using very general terms. This is a problematic aspect in that retrieving specific images using general terms is quite difficult. For example, use of the words “rounded sculpture metal”[6] as search the criterion to find the Saeverud monument in Google images generates about 800 000 images in the result set.

Results from this study indicate that what sets the participants challenges with image retrieval apart from challenges associated with information retrieval in general is closely related to problems associated with the creation of good representations of images. These problems occur because neither a text-based nor a binary representation can represent a real world object in a satisfactorily manner [45]. As such, there is a real difference between traditional information retrieval where the searches commonly are directed at retrieving a given document from a collection based on the documents actual text, and image retrieval, where the searches are directed against retrieving a given image based on an external representation of it (textual or binary). Also, since the query terms (i.e. example image or keywords) used for image retrieval actually are representations of an information need, the process of retrieving images is even more challenging than traditional information retrieval. Following from this is that in order to narrow the semantic gap, the results from this study indicate that there is a need to focus research on narrowing the intention gap.

Further Research

From the lessons learned in this study, the focus in future work will be on studying the ways in which CBIR may complement traditional TBIR when users interact with image collections. It is believed that a context-focused approach will be especially helpful in improving user interaction with image collections when users are looking for specific images depicting the content in a specific context.

Three important assumptions underlie the context-focused approach:

- Utilizing TBIR and CBIR in combination will enable users to specify their information need in a simpler, yet more complete manner since they may use both visual data and text.
- Combining two distinct communication channels with active use of both query and stored image context information will enable the image retrieval system to provide more relevant result sets. And
- User interaction, with the possibility for giving feedback, will provide the system with a better understanding of user intentions.

It has been noted that users of image retrieval systems should be presented with a diverse result set [16]. This view is shared in this work, and it is believed that presenting users with a richer result set that includes both the images matching the query terms given by the user, as well as associated context information available in the system will improve the user’s ability to locate relevant images for his/her information need. Future work will thus focus on investigating user expectations and user experiences from using a system that provides users with information to choose from in the various phases of the image retrieval process.

[6] These keywords were used by one participant in the search for images of the Saeverud monument

Acknowledgements:

Research for this paper has been supported by the National Research Council, Norway, under the NFR project #176858/S10, CAIM - Context Aware Image Management. I thank the student respondents for their willingness to participate in the study. I greatly acknowledge Associate Professor Joan C. Nordbotten for her encouragement and helpful suggestions. I would also like to thank the reviewers for their helpful comments on how to improve the paper.

References

1. Datta, R., et al., *Image Retrieval: Ideas, Influences, and Trends of the New Age.* ACM Computing Surveys, 2008. **40**(2): p. 1-60.
2. Arnold, W., et al., *Content-Based Image Retrieval at the End of the Early Years.* IEEE Transactions on Pattern Analysis and Machine Intelligence 2000. **22**(12): p. 1349 - 1380
3. Hua, X.-S., M. Worring, and T.-S. Chua, *Internet Multimedia Search and Mining*, in *Internet Multimedia Search and Mining*. 2010, Bentham Science Publishers.
4. Zhang, H.J. *Multimedia content analysis and search: new perspectives and approaches*. in *seventeen ACM international conference on Multimedia*. 2009. Beijing, China.
5. Jaimes, A. *Human Factors in Automatic Image Retrieval System Design and Evaluation*. in *IS&T/SPIE Internet Imaging*. 2006. San José, CA.
6. Jörgensen, C., *Image Retrieval Theory and Research*. 2003, Laham, Maryland: Scarecrow Press, Inc. 360.
7. Prasad, B.E., et al., *A Microcomputer-Based Image Database Management System.* IEEE Transactions on Industrial Electronics 1987. **IE-34**(1): p. 83 - 88.
8. Rui, Y., T.S. Huang, and S.F. Chang, *Image Retrieval: Current Techniques, Promising Directions And Open Issues.* Journal of Visual Communication and Image Representation, 1999. **10**(4): p. 39 - 62.
9. Hua, G. and Q. Tian, *What can visual content analysis do for text based image search?*, in *Proceedings of the 2009 IEEE international conference on Multimedia and Expo*. 2009, IEEE Press: New York, NY, USA.
10. Marques, O. and B. Furht, *Content-Based Image and Video Retrieval*. First ed, ed. B. Furth. 2002: Kluwer Academic Publishers 200.
11. Rui, Y., T.S. Huang, and S. Mehrotra. *Relevance feedback techniques in interactive content-based image retrieval*. in *SPIE/IS&T Conf. on Storage and Retrieval for Image and Video Databases* 1998. San Jose, CA.
12. Diakopoulos, N. and P. Chiu. *Photoplay: A collocated collaborative photo tagging game on a horizontal display*. in *UIST*. 2007.
13. Golder, S. and B.A. Huberman, *The structure of collaborative tagging systems.* Journal of Information Sciences, 2006. **32**(2): p. 198-208.
14. Mathes, A., *Folksonomies - cooperative classification and communication through shared metadata.* Computer Mediated Communication, 2004.
15. Gupta, M., et al., *Survey on Social Tagging Techniques*, in *SIGKDD Explorations*, O.R. Zaiane, C. Aggarwal, and B. Goethals, Editors. 2010. p. 58-72.
16. Zwol, R.v., et al., *Diversifying image search with user generated content*, in *Proceeding of the 1st ACM international conference on Multimedia information retrieval*. 2008, ACM: Vancouver, British Columbia, Canada.
17. Jörgensen, C. and P. Jörgensen, *Image querying by image professionals: Research Articles.* J. Am. Soc. Inf. Sci. Technol., 2005. **56**(12): p. 1346-1359.
18. Westman, S., A. Lustila, and P. Oittinen, *Search strategies in multimodal image retrieval*, in *Proceedings of the second international symposium on Information interaction in context*. 2008, ACM: London, United Kingdom.
19. Flickner, M., et al., *Query by Image and Video Content: the QBIC System.* IEEE Computer, 1995. **28**(9): p. 23–32.
20. Zhou, X.S. and T.S. Huang, *Unifying keywords and visual contents in image retrieval.* Ieee Multimedia, 2002. **9**(2): p. 23-32.
21. Lu, G., *Multimedia database management systems*. Artech House computing library. 1999, Boston: Artech House. 373.

22. Müller, H., P. Ruch, and A. Geissbuhler, *Enriching content-based image retrieval with multi-lingual search terms.* Swiss Medical Informatics, 2005. **54**: p. 6-11.
23. Westerveld, T. *Image Retrieval: Content versus Context*. in *RIAO 2000 Conference Proceedings*. 2000. Paris.
24. Kowalski, G. and M.T. Maybury, *Information Storage and Retrieval Systems - Theory and Implementation*. 2000: Kluwer Academic Publ.
25. Hartvedt, C. *Utilizing Context in Ranking Results from Distributed CBIR*. in *Norwegian Informatics Conference, NIK*. 2007. Oslo, Norway.
26. Grosky, W.I., R. Agrawal, and F. Fotouhi, *Mind the Gaps - Finding Appropriate Dimensional Representation for Semantic Retrieval of Multimedia Assets*, in *Semantic Multimedia and Ontologies*, I. Kompatsiaris and P. Hobson, Editors. 2008, Springer: London. p. 229-252.
27. Möller, R. and B. Neumann, *Ontology-Based Reasoning Techniques for Multimedia Interpretation and Retrieval*, in *Semantic Multimedia and Ontologies*, I. Kompatsiaris and P. Hobson, Editors. 2008, Springer: London.
28. Christel, M. *Examining User Interactions with Video Retrieval Systems*. in *SPIE*. 2007.
29. Christel, M. and R. Conescu. *Mining Novice User Activity with TRECVID Interactive Retrieval Tasks*. in *CIVR 2006*. 2006.
30. McDonald, S. and J. Tait, *Search strategies in content-based image retrieval*, in *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*. 2003, ACM: Toronto, Canada.
31. Strat, T.M. *Employing Contextual Information in Computer Vision*. in *DARPA93*. 1993.
32. Dey, A.K., et al., *Towards a better understanding of context and context-awareness.* Handheld and Ubiquitous Computing, Proceedings, 1999. **1707**: p. 304-307.
33. Elgesem, D. and J. Nordbotten. *The Role of Context in Image Interpretation*. in *Context-based Information Retrieval, CIR'07 at CONTEXT'07*. 2007. Roskilde University, Denmark.
34. Sandnes, F.E., *Where was that photo taken? Deriving geographical information from image collections based on temporal exposure attributes.* Multimedia Systems, 2010. **16**(4-5): p. 309-318.
35. Draper, S. *Evaluation in HCI and IR.* in *Second Mira Workshop*. 1996. Glasgow University.
36. Dunlop, M., *Reflections on Mira: Interactive evaluation in information retrieval.* American Society for Information Science, 2000. **51**(14): p. 1269-1274.
37. McDonald, S., T.-S. Lai, and J. Tait. *Evaluating a Content Based Image Retrieval System* in *24th ACM SIGIR Conference on Research and Development in Information Retreival* 2001. New Orleans, Louisiana: ACM Press
38. Jose, J., J. Furner, and D. Harper. *Spatial Querying For Images Retrieval: A User-Oriented Evaluation*. in *21st annual international ACM SIGIR conference on Research and development in information retrieval*. 1998. Melbourne, Australia: ACM New York, NY, USA
39. Lew, M.S., et al., *Content-based Multimedia Information Retrieval: State of the Art and Challenges.* ACM Transactions on Multimedia Computing, Communications, and Applications, 2006: p. 1-19.
40. Smeulders, A.W.M., et al., *Content-Based Image Retrieval at the End of the Early Years*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000. **22**(12): p. 1349–1380.
41. Anick, P. *Using terminological feedback for web search refinement—a log –based study*. in *SIGIR '03*. 2003. Toronto, Canada.
42. Debraski, P. *The Limited Use of Relevance Feedback.* in *Special Libraries Association 2005 Annual Conference*. 2005. Toronto, Ontario CANADA.
43. Crucianu, M., M. Ferecatu, and N. Boujemaa, *Relevance feedback for image retrieval: a short survey*, in *State of the Art in Audiovisual Content-Based Retrieval, Information Universal Access and Interaction, Including Datamodels and Languages*. 2004, DELOS2 European Network of Excellence.
44. Lai, T.S., J.I. Tait, and S. McDonald, *Image Browsing and Navigation Using Hierarchical Classification*, in *The Challenge of Image Retrieval 1999 conference, CIR'99*. 1999.
45. Mitchell, W.J.T., *Picture Theory: Essays on Verbal and Visual Representation*. 1994, Chicago: University of Chicago Press. 445.