Original Research Paper

Towards Semantic User Query: A Review

Hui-Hui Wang, Yin-Chai Wang and Soo-See Chai

Department of Computing and Software Engineering, Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia

Corresponding Author:
Hui-Hui Wang
Department of Computing and
Software Engineering, Faculty
of Computer Science and
Information Technology,
Universiti Malaysia Sarawak,
Kota Samarahan, Sarawak,
Malaysia
Email: hhwang@unimas.my

Abstract: This paper attempts to discuss the image query mechanisms and user needs for image retrieval. The explosive growth of image data leads to the need of research and development of Image retrieval. Image retrieval researches are moving from keyword, to low level features and to semantic features. Drive towards semantic features is due to the problem of the keywords which can be very subjective and time consuming while low level features cannot always describe high level concepts in the users' mind. This paper also highlights both the already addressed and outstanding issues.

Keywords: Image Retrieval, Semantic Gap, Semantic Based Image Retrieval

Introduction

An image can be interpreted and understood differently by different users (Yoon, 2011). It is important to understand the context of the users in an image seeking process for designing an image retrieval system (Choi, 2010). In some situations, users can imagine what they desired. However, they are unable to express their desire in precise wording (Datta *et al*., 2008; Zha *et al*., 2010). In such situations, the difficulties of the user are the query formulation problem, whereby they could not formulate and communicate the needed information effectively (Urban and Jose, 2006). Thus, understandings of the way the users express their search needs, which may lead to the provision of appropriate access points to visual materials for retrieval, remain important (Choi and Tasmussen, 2003).

Query mechanisms play an important role in bridging the semantic gap between users and image retrieval (Gudivada and Raghavan, 1995). The user query is used to express the user's intention and need to retrieve images from the collection of database that conform to human perception. The image search results and their accuracy are directly affected by the submitted queries (Zha *et al*., 2010). Researches have been focusing on reducing the semantic gap in image retrieval, however in order to define a semantic meaning and representation of the input query that can precisely describe the intent of the user queries as well as the domain coverage are the major challenges (Hu *et al*., 2009; Hua *et al*., 2013).

An image retrieval system should have an effective user interface that allows users to precisely express their need and intent for searching the images that are conformed to human perception (Hu *et al*., 1999). To help users find the desired images, image search has been intensively studied (Niblack *et al*., 1993; Flickner *et al*., 1995; Lew *et al*., 2006; Hua *et al*., 2013).

Evolution of User Query

Researchers have proposed a variety of ways to search images based on their intent from large image databases. Mostly, it can be classified into three levels (Eakins *et al*., 1996; Eakins, 2000). Each level represents different level of abstraction as illustrated in Fig. 1.

Level 1 comprises of retrieval process by primitive features, such as colour, texture, shape or the spatial location of image elements. For example, find pictures of an orange circle. At this level, image is retrieve directly derivable from the images themselves, without the need to refer to any external knowledge base.

Level 2 comprises of retrieval process by derived features, involving some degree of logical inference about the identity of the objects depicted in the image. For example, finding pictures of an object ('bus'). At this level, reference to some external knowledge base is normally required.

Level 3 comprises of retrieval process using abstract attributes and involving a significant amount of high-level reasoning about the meaning and purpose of the objects or scenes depicted. For Example, find pictures of a 'happy and cheerful girl'.
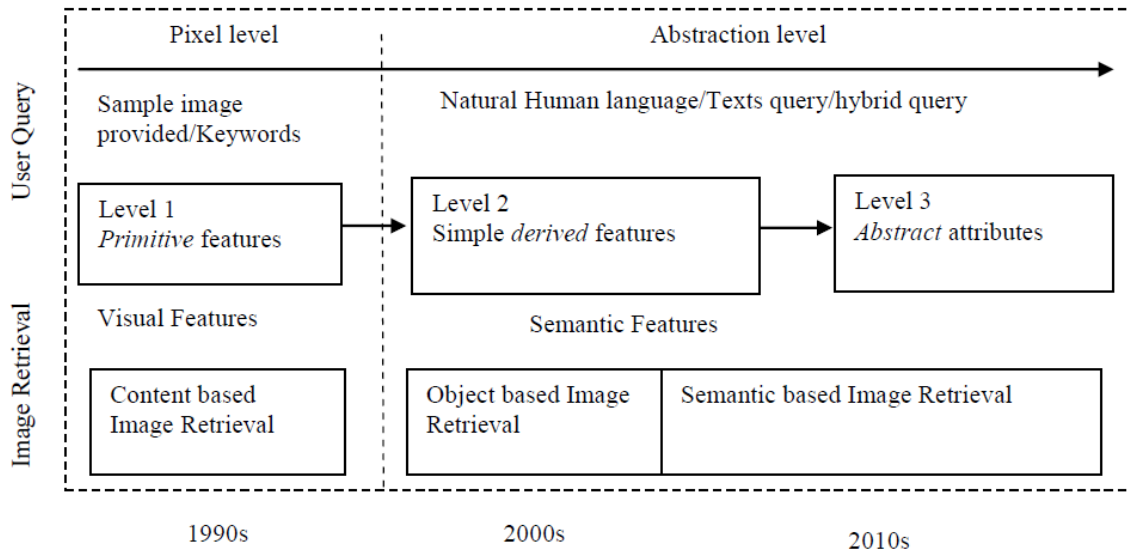
Fig. 1. Different level of abstractions

During the 1990s, Content-Based Image Retrieval (CBIR) has been introduced by providing user query by sample image besides using simple keywords. Initially, simple keywords have been used as image descriptors to index an image in initial CBIR. Texts are used to describe the content of the image which often causes ambiguity and inadequacy in performing an image database search and query processing. This problem is due to the difficulty in specifying exact terms and phrases in describing the content of images as the content of an image is much richer than what any set of keywords can express. Since the textual annotations are based on human language, variations in annotation pose challenges to image retrieval.

Besides, user query by providing the sample image and the visual features for CBIR system are extracted from the image itself. Visual features including color, texture, shape and spatial relations are used. Excellent survey on low-level image feature extraction in CBIR system can be found in (Wang *et al.*, 2006). Although there are many sophisticated algorithms to describe color, shape and texture features approaches, these algorithms do not satisfied and comfort to human perception. Main reason is due to the unavailability of low level image features in describing high level concepts in the users' mind. An example is the task of finding an image of a car in the middle of the road. The only way a machine is able to do this is by automatic extraction of the low level features that are represented in the low level features from the images with a good degree of efficiency.

Thus, semantic gap issue where the lack of correlation between the semantic categories that a user requires and the low level features that CBIR systems offer has been explored. The semantic gap issue has been investigated for years but still remains a big challenge. This is due to the fact that the visual image feature descriptors extracted from an image cannot (as yet) be automatically translated reliably into high-level semantics (Hu *et al.*, 2009).

In 2000s, semantic based image retrieval has been introduced. This is due to neither a single features nor a combination of multiple visual features could fully capture high level concept of images. Besides, the performance of image retrieval system based on low level features are not satisfactory, there is a need for the mainstream of the research converges to retrieve based on semantic meaning by trying to extract the cognitive concept of a human to map the low level image features to high level concept (semantic gap). In addition, representing the image content with semantic terms allows users to access images through text query which is more intuitive, easier and preferred by the front end users to express their mind compare with using images.

In order to improve the effectiveness and accuracy of CBIR systems, research direction has shifted from the designing of sophisticated low-level feature extraction algorithms to reducing the 'semantic gap' between the extracted visual features and the richness of human semantics (Vijay Kumar *et al.*, 2012). It's actually moving from the user query level 1 toward level 3.

User Intent

The concept of "intentions" is context-dependent, hence, hard to measure explicitly. A substitute of measuring user intentions is to study user search queries and browsing behavior for text-based search

(Broder, 2002; Jansen *et al*., 2008; Yin and Shah, 2010; Kumar and Tomkins, 2010) and image search (Kofler and Lux, 2009; Lux *et al*., 2010).

Intention Gap

The aim of CBIR systems is to provide maximum support in reducing the semantic gap between the simplicity of extracted visual features and the richness of the user semantics (Smeulders *et al*., 2000). The intention gap exists in CBIR with query-by-example between user's search intention and the query (Bian *et al*., 2012) which often leads to unsatisfactory search results (Zha *et al*., 2010). Intention gap may due to query image consists of many regions or components that are not part of user's interest or fails to capture important properties of the intended objects. Ability to select region-of-interest in a query image potentially solved the former (Yang *et al*., 2010) while query suggestion solved the later. User feedbacks are also widely exploited at inferring user search intent (Zhang *et al*., 2012; Yuan *et al*., 2011).

Towards Semantic User Query

Querying by visual example (Cox *et al*., 2000; Kelly *et al*., 1995; Tieu and Viola, 2000) is a paradigm that is used to express perceptual aspects of visual features of image content (Kelly *et al*., 1995) and it is refer to query level 1. The queries are based on colour, texture and shape features that can be formulated depending on the feature (s) extracted during image extraction stage. The user is required to select a candidate image through the query interface and the system converts the image into visual features for image representation. Finally, a list of similar images will be retrieved based on the image similarity criteria. Efforts have been made to extend the query by visual example to query by region selection (Chang *et al*., 1998; Carson *et al*., 1997) and sketch (Kato *et al*., 1992; Daoudi and Matusiak, 2000; Chans *et al*., 1997; Lai, 2000; Egenhofer, 1996). Users are allowed to select their "region of interest' in the image or draw their desired content. Query by sketch allows users to sketch the desired image with image editor software. The image properties of colour, texture, shape, sizes and spatial locations of the desired objects can be specified.

This search method often fails to capture similarity that could be inferred by humans, a phenomenon that is now commonly referred to as the semantic gap (Smeulders *et al*., 2000). Furthermore, in many cases locating a suitable example for a search may be a difficult task in itself (Rodden, 1999).

User usually prefers using keywords to indicate what they want (Khefi *et al*., 2004; Fu, 2007). Researchers believe textual queries usually provide more accurate description of users' information needs as it allows users to express their information needs at the semantic level and high level abstractions instead of limiting to the level of preliminary image features. Many popular search engines (e.g., Google, Yahoo) have developed technologies that allow users to search the Internet's images using keywords.

A retriever must represent the target image with verbal keywords in an image retrieval system which uses a keyword interface. Interpretation from an image to verbal keywords is difficult for humans and even if they can find the keywords, they have difficulty in finding an appropriate image when they have not been annotated or annotated with various keywords in the system. One example is that Google search engine offers the possibility to search for images using surrounding text and file names instead of image semantic meaning. So, database images that are not annotated with related keywords or not similar to the searching keyword will be not retrieved as similar images. That is why sometimes the search performed does not lead to satisfactory results.

This keywords image search also works well only when all the images are annotated with accurate textual information. Also, there are times and situations when users can imagine what they desire, but are unable to express this desire in precise wording (Niblack *et al*., 1993; Zha *et al*., 2010; Gerrig and Zimbardo, 2001). Moreover, keyword queries are usually ambiguous especially when they are short (one or two words) and thus cannot reflect users' intents precisely (Gerrig and Zimbardo, 2001).

User's Needs

An image can be interpreted and understood differently by different users (Yoon, 2011). It is important to understand the context of the users in an image seeking process for designing an image retrieval system (Choi, 2010). In some situations, users can imagine what they desire. However, they are unable to express their desire in precise wording (Datta *et al*., 2008; Zha *et al*., 2010). In such situations, the difficulties of the user are the query formulation problem, whereby they could not formulate and communicate the needed information effectively (Urban and Jose, 2006). Thus, understandings of the way the users express their search needs, which may lead to the provision of appropriate access points to visual materials for retrieval, remain important (Choi and Tasmussen, 2003).

Users often have a difficult time and unable to clearly articulate their information needs (Kelly and Fu, 2007; Belkin, 2000) and that users typically pose very short queries, usually between two and three words in length (Jansen *et al*., 2000). The short queries are due to the difficulty that users have in identifying and articulating their information needs and also conventional search interfaces encourage them to do so (Belkin *et al*., 2003).

Besides, with the different query language and the language associated with the images will also cause the inaccurate retrieval results (Ménard, 2011). The keyword queries are usually ambiguous especially when they are short and thus limited in reflecting users' intents precisely (Zha *et al*., 2010).

The Trend

Nowadays, the mainstream of image search techniques converges to retrieval that focuses on extracting and understanding user's search intent and the search query. The methods towards semantic user query are categorised into the following classes.

Relevance Feedback

Relevance Feedback (RF) is the key technique to reduce the intention gap in CBIR by exploiting user interactions (Zhou and Huang, 2003). The presence of the intention gap actually limits the understanding of user search intent by CBIR system, where the results from automatic retrieval system often do not satisfy users' information needs. Therefore, the practical way to identify the user search intent is to include the user into the retrieval process. Retrieved images are flagged as being either "relevant" or "irrelevant" or multi-relevance levels (Yang *et al*., 2002; Wu *et al*., 2004; Yuan *et al*., 2011). Users' feedbacks are then exploited by a relevance feedback algorithm to refine the search model. Through iterative interactions, relevance feedback attempts to capture users search intent and improve the search results. An interactive scheme, attribute feedback allows user to provide feedbacks on semantic features in order to capture user's search intent (Zhang *et al*., 2012; Cai *et al*., 2012).

Query Suggestion

One of the solutions to fill the intention gap is query suggestion, which allows interaction between users and search engines. A list of queries based on users' inputs is used to help users narrow down to precise queries. Most works focused on learning query suggestion model from user search queries logs or textual query suggestion (Baeza-yates *et al*., 2004; Mei *et al*., 2008; Carpineto and Romano, 2012). However, users usually perceive information in image more quickly than textual suggestions. Therefore, combination of visual and textual query suggestion can express the search intent better than using text suggestion alone.

Query suggestion for CBIR with query-by-example only has gained interest recently (Zha *et al*., 2010; Bian *et al*., 2012). The technique firstly provides a list of both representative images and keywords from user's initial textual query. User selected keyword-image from the list provided are used to refine the initial search results which helps user formulate a more precise search intents.

Interactive Interface

The Intention gap is closely related to user interface design of the image search to effectively express what is in the user's mind. Database researchers are concerned with indexing and querying images, whereas image processing experts tend to worry more on extracting the appropriate image descriptors. However, researchin providing user interface support for CBIR systems is scarce (Pečenovió *et al*., 2000; Santini and Jain, 2000; Nazakato *et al*., 2003; Heesch, 2008) until recently. Search engines designed to bridge the intention gap tend to integrate textual (keywords or tags) and visual features (Chen *et al*., 2013; Zhu *et al*., 2014; Craggs *et al*., 2014).

Friendly user interface not only means the interface should be as easy as possible to use, but also indicates that it can satisfy user's various searching needs. The easiest way to search similar images is different under different searching scenarios or users' searching needs; therefore, the corresponding user interfaces should be designed to fit users' various searching needs (Zhang, 2010). For image databases user, the ability to retrieve images based on their semantic content is important (Eakins *et al*., 2004). Besides, the user interface created for image searching may need to provide additional functionality to embrace the rich descriptions of the image, as opposed to the limited information provided by the usually short queries. The image attributes might be employed in user query. The fundamental idea of emergent image semantics is reflected in image searches. During the search, based on the previous result set, the user guides his/her queries towards the intended group of images. Besides this, the user could also judge the relevance of the images based on the features that differentiate these images (Westman, 2009).

An interactive and multi-modal mobile visual search application has been presented (Wang *et al*., 2011). A real-world image search prototype has been implemented where query refinement suggestion are derived from relevance feedback in an image retrieval system (Leiva *et al*., 2011).

Miscellaneous Techniques

Other techniques could have potential in inferring semantic user query. A contextual object retrieval model employing probabilistic reasoning has been proposed (Yang *et al*., 2010; 2011). The models are estimated using the visual words from both the region-of-interest and the visual context. These visual words are weighted using the search intent scores estimated using probabilistic techniques where the score indicates how likely the image region represented by the visual word reflects the search intent of the user. Significant performance improvement of QBE-based web image retrieval integrating global and local query context is reported (Yang *et al*., 2012).

An image knowledge base employing a graph representation that reflects the distribution of both images on the web and users' interests has been presented (Wang *et al.*, 2012). The image knowledge base is automatically constructed via both bottom-up and top-down approach that matches billion of web images onto an ontology of human knowledge. This image knowledge base can be used to order images of high quality and relevant to user interest and hence, more accurate images are returned for a user query. A similar system to build a large visual knowledge base automatically is reported (Zhang *et al.*, 2014).

Recently, exploratory search which combines browsing and searching are gaining interest (Lu *et al.*, 2014). The system developed allows users to specify object of interest by circling the visual objects while being browsed and the system automatically compiles the selected visual entities to represent users' underlying intent. The visual entity is scalable to textual context and visual attributes, hence, aid users gaining additional knowledge and form more accurate queries.

Issues

Success in answering queries at level 3 in Fig. 1 required some sophistication on the part of the searcher. Complex reasoning and often subjective judgment, can be required to make the link between image content and the abstract concepts it is required to illustrate. Queries at this level, though perhaps less common than level 2, are mostly encountered in both newspaper and art libraries. Recently researchers are focus on the user query at level 2 whereas it's only able to find and derived the semantic features of images instead of semantic concept. They ignore a further type of image query-retrieval by associated metadata such as, who created the image, where and when. This is not because such retrieval is unimportant. It is because (at least at present) such metadata is exclusively textual and its management is primarily a text retrieval issue. For example, find images of a happy and cheerful girl (level 3) instead of finding images of a girl (level 2). A user study confirms that users have difficulties in selecting search terms and finding images which illustrate abstract concepts (Besseling, 2011).

Even though the image retrieval research is moving towards semantic concept, much initial research in semantic image retrieval is focusing just on the simple semantic retrieval such as retrieval of objects of a given type. Little attention is on the retrieval by abstract attributes which involve a significant amount of high-level reasoning about the meaning and purpose of the objects or scenes depicted. In other words, the retrieval by abstract attributes was still not satisfied to human perception. Moreover, the retrieval involved human interference and is time consuming besides inconsistency. There is a need to further increase the confidence in image understanding and to effectively retrieve similar images that are conformed to human perception and without human interference.

There are still some spaces which need to be improved besides the challenges that are associated with mapping low level to high level concepts. Researchers are moving towards to intelligent image retrieval that are also supports more abstract in concept by understanding the image content in terms of high level concepts, which is closely related to the problem of computer vision and object recognition besides more intelligent system. The domain should be not specific but broad where all the extracted semantic features are applicable for any kind of images collection.

Conclusion

This paper provides a study of image retrieval work towards semantic based image retrieval. Recent works are mostly lack of semantic features extraction and user behaviour consideration. Therefore, there is a need of image retrieval system that is capable to interpret the user query and automatically extract the semantic feature that can make the retrieval more efficient and accurate to bridge the semantic gap problem in image retrieval.

Funding Information

Author's Contributions

All authors equally contributed in this work.

Ethics

The article is original and contains unpublished material. The corresponding author confirms that all the other authors have read the manuscript and approved it for publication.

References

Baeza-yates, R., C. Hurtado and M. Mendoza, 2004. Query recommendation using query logs in search engines. Proceedings of the International Conference on Current Trends in Database Technology, Mar. 14-18, Springer Berlin Heidelberg, pp: 588-596. DOI: 10.1007/978-3-540-30192-9_58

Belkin, N., 2000. Helping people find what they don't know. Commun. ACM, 43: 58-61. DOI: 10.1145/345124.345143

Belkin, N., D. Kelly, G. Kim, J.Y. Kim and H.J. Lee *et al.*, 2003. Query length in interactive information retrieval. Proceedings of 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Jul.28-Aug.1, ACM, Canada, pp: 205-212. DOI: 10.1145/860435.860474

Besseling, R., 2011. Designing an image retrieval interface for abstract concepts within the domain of journalism. CWI. Information Systems.

Bian, J., Z.J. Zha, H. Zhang, Q. Tian and T.S. Chua, 2012. Visual query attributes suggestion. Proceedings of the 20th ACM International Conference on Multimedia, Oct. 29-Nov. 2, ACM, Japan, pp: 869-872. DOI: 10.1145/2393347.2396334

Broder, A., 2002. A taxonomy of web search. ACM SIGIR Forum, 36: 3-10. DOI: 10.1145/792550.792552

Cai, J., Z.J. Zha, W. Zhou and Q. Tian, 2012. Attribute-assisted reranking for web image retrieval. Proceedings of the 20th ACM International Conference on Multimedia, Oct. 29-Nov. 2, ACM, Japan, pp: 873-876. DOI: 10.1145/2393347.2396335

Carpineto, C. and G. Romano, 2012. A survey of automatic query expansion in information retrieval. ACM Comput. Surveys, 44: 1-50. DOI: 10.1145/2071389.2071390

Carson, C., S. Belongie, H. Greenspan and J. Malik, 1997. Region-based image querying. Proceedings of IEEE Workshop on Content-Based Access of Image and Video Libraries, Jun. 20, IEEE Xplore, USA, pp: 42-49. DOI: 10.1109/IVL.1997.629719

Chang, S.F., W. Chen and H. Sundaram, 1998. Semantic visual templates: Linking visual features to semantics. Proceedings of International Conference on Image Processing, IEEE Xplore, USA, 3: 531-535. DOI: 0.1109/ICIP.1998.727321

Chans, Y., Z. Lei, D. Lopresti and S. Kung, 1997. A feature-based approach for image retrieval by sketch. Proceedings of the SPIE 3229, Multimedia Storage and Archiving Systems II, (SAS' 97), pp: 220-231. DOI: 10.1117/12.290343

Chen, Y., H. Sampathkumar, B. Luo and X. Chen, 2013. iLike: Bridging the semantic gap in vertical image search by integrating text and visual features. IEEE Trans. Knowl. Data Eng., 25: 2257-2270. DOI: 10.1109/TKDE.2012.192

Choi, Y., 2010. Effects of contextual factors on image searching on the Web. J. Am. Society Inform. Sci. Technol., 61: 2011-2028. DOI: 10.1002/asi.21386

Choi, Y. and E. Tasmussen, 2003. Searching for images: The analysis of users' queries for image retrieval in american history. J. Am. Society Inform. Sci. Technol., 54: 498-511. DOI: 10.1002/asi.10237

Cox, I., M. Miller, T. Minka, T. Papathomas and P. Yianilos, 2000. The Bayesian image retrieval system, PicHunter: Theory, implementation and psychophysical experiments. IEEE Trans. Image Process., 9: 1057-7149. DOI: 10.1109/83.817596

Craggs, B., M. Scott and J. Alexander, 2014. ThumbReels: Query sensitive web video previews based on temporal, crowdsourced, semantic tagging. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Apr. 26-May. 1, ACM, Canada, pp: 1217-1220. DOI: 10.1145/2556288.2557249

Daoudi, M. and S. Matusiak, 2000. Visual image retrieval by multiscale description of user sketches. J. Visual Langu. Comput., 11: 287-301. DOI: 10.1006/jvlc.2000.0159

Datta, R., D. Joshi, J. Li and J. Wang, 2008. Image retrieval: Ideas, influences and trends of the new age. ACM Comput. Surveys, 40: 1-60. DOI: 10.1145/1348246.1348248

Eakins, J., 2000. Retrieval of still images by content. Proceedings of the 3rd European Summer-School on Lectures on Information Retrieval, Sep. 11-15, Springer Berlin Heidelberg, pp: 111-138. DOI: 10.1007/3-540-45368-7_6

Eakins, J., P. Briggs and B. Burford, 2004. Image retrieval interfaces: A user perspective. Proceedings of 3rd Internation Conference in Image and Video Retreival, Jul. 21-23, Springer Berlin Heidelberg, pp: 628-637. DOI: 10.1007/978-3-540-27814-6_73

Eakins, J., K. Shields and J. Boardman, 1996. ARTISAN-a shape retrieval system based on boundary family indexing. Proceedings of SPIE 2670, Storage and Retrieval for Still Image and Video Databases IV, (IVD' 96) pp: 2670-2670. DOI: 10.1117/12.234792

Egenhofer, M., 1996. Spatial-query-by-sketch. Proceedings of IEEE Symposium on Visual Languages, Sep. 3-6, IEEE Xplore, USA, pp: 60-67. DOI: 10.1109/VL.1996.545269

Flickner, M., H. Saywhney, W. Niblack, J. Ashley and Q. Huang *et al.*, 1995. Query by image and video content: The QBIC system. Computer, 28: 23-32. DOI: 10.1109/2.410146

Fu, H., 2007. Attention-driven image interpretation, annotation and retrieval. Unpublished dissertation in partial fulfilment of the requirements for the degree of Doctor of Philosophy, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.

Gerrig, R. and P. Zimbardo, 2001. Psychology and Life. 16 Ed. Allyn and Bacon, Boston, ISBN-10: 0205344569, pp: 403.

Gudivada, V.N. and V.V. Raghavan, 1995. Content-based image retrieval systems. IEEE Comput., 28: 18-22. DOI: 10.1109/2.410145

Heesch, D., 2008. A survey of browsing models for content based image retrieval. Multimedia Tools Applic., 40: 261-284. DOI: 10.1007/s11042-008-0207-2

Hu, J., G. Wang, F. Lochovsky, J.T. Sun and Z. Chen, 2009. Understanding user's query intent with wikipedia. Proceedings of the 18th International Conference on World Wide Web, Apr. 20-24, ACM, Spain, pp: 471-480. DOI: 10.1145/1526709.1526773

Hu, P.H., P.C. Ma and P. Chau, 1999. Evaluation of user interface designs for information retrieval systems: A computer-based experiment. J. Decision Support Syst., 27: 125-143.
DOI: 10.1016/S0167-9236(99)00040-8

Hua, X.S., L. Yang, J. Wang, J. Wang and M. Ye *et al.*, 2013. Clickage: Towards bridging semantic and intent gaps via mining click logs of search engine. Proceedings of the 21st ACM International Conference on Multimedia, Oct. 21-25, ACM, Spain, pp: 243-252. DOI: 10.1145/2502081.2502283

Jansen, B., D. Booth and A. Spink, 2008. Determining the informational, navigational and transactional intent of web queries. Inform. Process. Manage., 44: 1251-1266. DOI: 10.1016/j.ipm.2007.07.015

Jansen, B., A. Spink and T. Saracevic, 2000. Real life, real users and real needs: A study and analysis of user queries on the web. Inform. Process. Manage., 36: 207-227. DOI: 10.1016/S0306-4573(99)00056-4

Kato, T., T. Kurita, N. Otsu and K. Hirata, 1992. A sketch retrieval method for full color image database-query by visual example. Proceedings of 11th IAPR International Conference on Pattern Recognition, Aug. 30-Sep.3, IEEE Xplore, The Hague, pp: 530-533.
DOI: 10.1109/ICPR.1992.201616

Kelly, D. and X. Fu, 2007. Eliciting better information need descriptions from users of information search systems. Inform. Process. Manage., 43: 30-46.
DOI: 10.1016/j.ipm.2006.03.006

Kelly, P., T. Cannon and D. Hush, 1995. Query by image example: The Comparison Algorithm for Navigating Digital Image Databases (CANDID) approach. Proceedings of SPIE 2420, Storage and Retrieval for Image and Video Databases III, (IVD' 95), pp: 238-248. DOI: 10.1117/12.205289

Khefi, M., D. Ziou and A. Bernardi, 2004. Image retrieval from the world wide web: Issues, techniques and systems. ACM Comput. Surveys, 36: 35-67. DOI: 10.1145/1013208.1013210

Kofler, C. and M. Lux, 2009. Dynamic presentation adaptation based on user intent classification. Proceedings of the 17th International Conference on Multimedia, Oct. 19-24, ACM, China, pp: 1117-1118. DOI: 10.1145/1631272.1631526

Kumar, R. and A. Tomkins, 2010. A characterization of online browsing behavior. Proceedings of the 19th International Conference on World Wide Web, Apr. 26-30, ACM, USA, pp: 561-570.
DOI: 10.1145/1772690.1772748

Lai, T., 2000. CHROMA: A photographic image retrieval system. PhD Thesis, University of Sunderland, Sunderland, UK.

Leiva, L., M. Villegas and R. Paredes, 2011. Query refinement suggestion in multimodal image retrieval with relevance feedback. Proceedings of the 13th International Conference on Multimodal Interfaces, Nov. 14-18, ACM, Spain, pp: 311-314.
DOI: 10.1145/2070481.2070539

Lew, M., N. Sebe, C. Djeraba and R. Jain, 2006. Content-based multimedia information retrieval: State of the art and challenges. ACM Trans. Multim. Comput. Commun. Applic., 2: 1-19.
DOI: 10.1145/1126004.1126005

Lu, S., T. Mei, J. Wang, J. Zhang and Z. Wang *et al.*, 2014. Browse-to-search: Interactive exploratory search with visual entities. ACM Trans. Inform. Syst., 32: 1-27. DOI: 10.1145/2630420

Lux, M., C. Kofler and O. Marques, 2010. A classification scheme for user intentions in image search. Proceedings of CHI Conference on Human Factors in Computing Systems, Apr. 10-15, ACM, USA, pp: 3913-3918.
DOI: 10.1145/1753846.1754078

Mei, Q., D. Zhou and K. Church, 2008. Query suggestion using hitting time. Proceedings of the 17th ACM Conference on Information and Knowledge Management, Oct. 26-30, ACM, USA, pp: 469-478. DOI: 10.1145/1458082.1458145

Ménard, E., 2011. Search behaviours of image users: A pilot study on museum objects. Partnership Canadian J. Library Inform. Practice Res., 6: 1-18.

Nazakato, M., L. Manola and T. Huang, 2003. Image grouper: A group-oriented user interface for content-based image retrieval and digital image arrangement. J. Visual Langu. Comput., 14: 363-386.
DOI: 10.1016/S1045-926X(03)00034-X

Niblack, W., R. Barber, W. Equitz, M. Flickner and E. Glasman *et al.*, 1993. The QBIC project: Querying images by content using color, texture and shape. Proceedings of SPIE 1908, Storage and Retrieval for Image and Video Databases, (IVD' 93) pp: 173-187.
DOI: 10.1117/12.143648

Pečenovió, Z., M.N. Do, M. Vetterli and P. Pu, 2000. Integrated browsing and searching of large image collections. Proceedings of 4th International Conference, VISUAL 2000, Nov. 2-4, Springer Berlin Heidelberg, pp: 279-289.
DOI: 10.1007/3-540-40053-2_25

Rodden, K., 1999. How do people organise their photographs? Proceedings of the 21st Annual BCS-IRSG conference on Information Retrieval Research, Apr. 19-20, eWiC, Glasgow, pp: 12-12.

Santini, S. and R. Jain, 2000. Integrated browsing and querying for image databases. IEEE Multim., 7: 26-39. DOI: 10.1109/93.879766

Smeulders, W., M. Worring, S. Santini, A. Gupta and R. Jain, 2000. Content-baed image retrieval at the end of the early years. IEEE Trans. Pattern Analysis Machine Intellig., 22: 1349-1380. DOI: 10.1109/34.895972

Tieu, K. and P. Viola, 2000. Boosting image retrieval. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Jun. 13-15, IEEE Xplore, USA, pp: 228-235. DOI: 10.1109/CVPR.2000.855824

Urban, J. and J. Jose, 2006. An explorative study of interface support for image searching. Proceedings of the 3rd International Workshop, Jul. 28-29, Springer Berlin Heidelberg, pp: 207-221. DOI: 10.1007/11670834_17

Vijay Kumar, K., R. Rajeswara Rao, V. Subba Ramaiah and J. Rani kaka, 2012. Content based image retrieval system consume semantic gap. Int. J. Comput. Sci. Inform. Technol., 3: 5231-5235.

Wang, X.J., Z. Xu, L. Zhang, C. Liu and Y. Rui, 2012. Towards indexing representative images on the web. Proceedings of the 20th ACM International Conference on Multimedia, Oct. 29-Nov.2, ACM, Japan, pp: 1229-1238. DOI: 10.1145/2393347.2396423

Wang, Y., T. Mei, J. Wang, H. Li and S. Li, 2011. JIGSAW: Interactive mobile visual search with multimodal queries. Proceedings of the 19th ACM International Conference on Multimedia, Nov.28-Dec.1, ACM, USA, pp: 73-82. DOI: 10.1145/2072298.2072310

Wang, Z., D. Feng, Z. Chi and T. Xia, 2006. Annotating image regions using spatial context. Proceedings of 8th IEEE International Symposium on Multimedia, Dec. 11-13, IEEE Xplore Press, San Diego, CA, pp: 55-61. DOI: 10.1109/ISM.2006.32

Westman, S., 2009. Image Users' Needs and Searching Behaviour. In: Information Retrieval: Searching in the 21st Century. Goker, A. and J. Davies (Eds.), John Wiley and Sons, Chichester, ISBN-10: 0470033630, pp: 63-83.

Wu, H., H. Lu and S. Ma, 2004. WillHunter: Interactive image retrieval with multilevel relevance. Proceedings of the 17th International Conference on Pattern Recognition, Aug. 23-26, IEEE Xplore Press, pp: 1009-1012. DOI: 10.1109/ICPR.2004.1334430

Yang, J., Q. Li and Y. Zhuang, 2002. Image retrieval and relevance feedback using peer indexing. Proceedings of IEEE International Conference on Multimedia and Expo, Aug. 26-29, IEEE Xplore Press, pp: 409-412. DOI: 10.1109/ICME.2002.1035624

Yang, L., B. Geng, Y. Cai, A. Hanjalic and X.S. Hua, 2011. Object retrieval using visual query context. IEEE Trans. Multimedia, 13: 1295-1307. DOI: 10.1109/TMM.2011.2162399

Yang, L., B. Geng, A. Hanjalic and X.S. Hua, 2010. Contextual image retrieval model. Proceedings of the ACM International Conference on Image and Video Retrieval, Jul. 5-7, ACM, China, pp: 406-413. DOI: 10.1145/1816041.1816100

Yang, L., B. Geng, A. Hanjalic and X.S. Hua, 2012. A unified context model for web image retrieval. ACM Trans. Multimedia Comput. Commun. Applic., 8: 1-19. DOI: 10.1145/2240136.2240141

Yin, X. and S. Shah, 2010. Building taxonomy of web search intents for name entity queries. Proceedings of the 19th International Conference on World Wide Web, Apr. 26-30, ACM, USA, pp: 1001-1010. DOI: 10.1145/1772690.1772792

Yoon, J., 2011. Searching images in daily life. Library Inform. Sci. Res., 33: 269-275. DOI: 10.1016/j.lisr.2011.02.003

Yuan, J., Z.J. Zha, Y.T. Zheng, M. Wang, X. Zhou and T.S. Chua, 2011. Utilizing related samples to enhance interactive concept-based video search. IEEE Trans. Multimedia, 13: 1343-1355. DOI: 10.1109/TMM.2011.2168813

Zha, Z., L. Yang, T. Mei, M. Wang and Z. Wang *et al.*, 2010. Visual query suggestion: Towards capturing user intent in internet image search. ACM Trans. Multimedia Comput. Commun. Applic., 6: 1-19. DOI: 10.1145/1823746.1823747

Zhang, D., 2010. A survey of user interfaces in content-based image search engines on the web. Proceedings of the 4th Workshop on Human-Computer Interaction and Information Retrieval, (IIR' 10), USA, pp: 129-132.

Zhang, H., Y. Yang, H. Luan, S. Yan and T.S. Chua, 2014. Start from scratch: Towards automatically identifying, modeling and naming visual attributes. Proceedings of the ACM International Conference on Multimedia, Nov. 3-7, Orlando, FL, USA, pp: 187-196. DOI: 10.1145/2647868.2654915

Zhang, H., Z.J. Zha, S. Yan, J. Bian and T.S. Chua, 2012. Attribute feedback. Proceedings of the 20th ACM International Conference on Multimedia, Oct. 29-Nov.2, ACM, Japan, pp: 79-88. DOI: 10.1145/2393347.2393365

Zhou, X. and T. Huang, 2003. Relevance feedback in image retrieval: A comprehensive review. Multimedia Syst., 8: 536-544. DOI: 10.1007/s00530-002-0070-3

Zhu, S., J. Cai, J. Zhang, Z. Li and J.C. Wang *et al.*, 2014. Bridging the user intention gap: An intelligent and interactive multidimensional music search engine. Proceedings of the 1st International Workshop on Internet-Scale Multimedia Management, Nov. 3-7, ACM, USA, pp: 59-64. DOI: 10.1145/2661714.2661720