ISSN:2348-2079

Volume-6 Issue-2

International Journal of Intellectual Advancements and Research in Engineering Computations

Exploratory Product Image Search with Cross Re-Ranking Model

[1]Ms.N.Zahira Jahan, M.C.A.,M.Phil., Associate Professor/MCA,
[2]Mr.S. Naveen, III MCA,
Department of MCA,NandhaEngineering College (Autonomous), Erode-52.
E-MailID: zahirajahan1977@gmail.com,nvncte@gmail.com

***Abstract*- In this paper analysis the challenges of understanding users' search interests from the product images being browsed and inferring their actual search intentions. In this paper a novel interactive image exploring system for allowing users to lightly switch between browse and search processes, and naturally complete visual-based exploratory search tasks in an effective and efficient way. This system enables users to specify their visual search interestsinproduct images by circling any visual objects in web pages, and then the system automatically infers users' underlying intent by analyzing the browsing context and by analyzing the sameor similar product images obtained by large-scaleimage search technology.**

***Index Terms*-- content-based image retrieval, exploratory search, gesture, interactive visual search, user interaction.**

I.INTRODUCTION

A few research attempts address the information-seeking problem in terms of improving the browsing experience. Several other research efforts focus on typical keyword-based search scenarios, primarily tackling the challenges of query understanding problems in a search system, such as query classification, query disambiguation, and query suggestion. A number of works leverage contextual cues for facilitating search intent expression and query augmentation. In the last few years, exploratory search has been regarded by many in the research community as an alternative for improving the standard search. Exploratory search is defined as a specialization of information-seeking behavior, which is commonly depicted by three kinds of activities lookup learn and investigate, especially pertinent to the learn and investigate activities. As the web has become the first choice for information seekers, people expect it to serve other kinds of information needs and search engines must strive to provide services beyond lookup. Consequently, research has focused on defining a broader set of information behaviors to learn about the situations when a user is, or feels, limited by only having the ability to perform a keyword search, especially when visual information is involved.However, discovering information and understanding interests in the images in the exploratory search.

II. RELATED WORKS

Kota Yamaguchi et al demonstrate an effective method for parsing clothing in fashion photographs, an extremely challenging problem due to the large number of possible garment items, variations in configuration, garment appearance, layering, and occlusion. In addition, provide a large novel dataset and tools for labeling garment items, to enable future research on clothing estimation. Finally, present intriguing initial results on using clothing estimates to improve purpose identification and demonstrate a prototype application for pose-independent visual garment retrieval Consider the upper east consider in address and pearls, the banker in his tailored suit and wingtips, or the hipster inhis flannel shirt, tight jeans, and black framed glasses.

D.Twidale, M.B. et al describe a novel system for user-driven integration of name variants when interacting with web-based information in particular digital library systems. We approach these issues via a client-side JavaScript browser extension that can reorganize web content and also integrate remote data sources. Designed to be agnostic towards the web sites it is applied to, illustrate the developed proof-of-concept system through worked examples using three different digital libraries. We discuss the extensibility of the approach in the context of other user-driven information systems and the growth of the Semantic Web.

Yang Cao etaldescribe an interactive sketch-based image search engine. Different from existing work, most of which is limited to a small-scale database or only enables single modality input, Mind Finder is a sketch-based multi model search engine for million-level database. It enables users to sketch major curves of the target image in their mind, and also supports tagging and coloring operations to better express their search intentions. Owning to a friendly interface, our

system supports multiple actions, which help users to flexibly design their queries.

After each operation, top returned images are up dated in real time, based on which users could interactively refine their initial thoughts until ideal images are returned. The novelty of the Mind Finder system includes the following two aspects anmulti model searching scheme is proposed to retrieve images which meet users' requirements not only in structure, but also in semantic meaning and color tone. An indexing framework is designed to make Mind Finder scalable in terms of database size, memory cost, and response time. By scaling up the database to more than two million images, Mind Finder not only helps users to easily present whatever they are imagining, but also has the potential to retrieve the most desired images in their mind.

Gabriel Take's et al **e**stablishing visual correspondences is an essential component of many computer vision problems, and local feature-descriptors. Transmission and storage of these descriptors are of critical importance in the context of mobile distributed camera networks and large indexing problems. To propose system a framework for computing low bit-rate feature descriptors with a 20× reduction in bit rate.The framework is low complexity and has significant speed-up in the matching stage. Represent gradient histograms as tree structures which can be efficiently compressed. The efficiently compute distances between descriptors in their compressed representation eliminating the need for decoding and perform a comprehensive performance comparison with SIFT, SURF, and other low bit-rate descriptors and show that proposed CHoG descriptor outperforms existing schemes.

Wai-Tat Fu, et al describe an extension of a computational cognitive model of social tagging and exploratory search called the semantic imitation model. The model assumes a probabilistic representation of semantics for both internal and external knowledge and utilizes social tags as navigational cues during exploratory search. Proposed system used to generate a measure of information scent that controls exploratory search behavior and simulated the effects of multiple presentations of navigational cues on both simple information retrieval and exploratory search performance based on a previous model called SNIF-ACT.

In this paper found that search performance can be significantly improved by these model-based presentations of navigational cues for both experts and novices. The result suggested that exploratory search performance depends critically on the match between internal knowledge and external knowledge structures.

III.SYSTEM METHODOLOGY

In Spite of the emergence of user-friendly computing devices and the evolution of user-centric exploring behaviors, under today's web interfaces, the two most ordinary online activities, browse and search, are still performed in isolation, owing to typical search systems rarely serving people to carry a search as means to seek information.However, discovering information and understanding interests in the images in the exploratory search loop still pose significant challenges.
In conventional image search systems, people commonly find it difficult to define or describe their seeking demands with respect to images even when general search interests appear, owing to a lack of specific knowledge of how to express intent by forming search queries. It is often difficult to precisely formulate queries to express users' specific intent. Even when the user formulates the intent-specific query, due to the well-known semantic gap between low-level image features and high-level semantic concepts, the resulting images might be still irrelevant to the query. That is, users seek semantic similarity, but the systems can only return the results with the similar visual appearances. The user has to submit other image queries to conduct iterative exploratory search.The search intent might be multifaceted, but the resulting images of each search iteration can only reflect one fact. In the proposed exploratory search system, unlike conventional image search engines, returning the replicas of the query image is not the goal of exploratory search demands.

A. ATTRIBUTE MINING BY LARGE-SCALE IMAGE SEARCH

In this section, it proposes an approach to facilitate users selecting the interesting visual objects, and then automatically infer the multiple aspects of the interest via large-scale image search. Rather than simply performing a query-by-example content-based image retrieval, constructing a query entity enriched by semantic meaning.Except utilizing the visual features to understand the meaning embedded in the selected image, the semantic meaning is also enriched by the surrounding context and the high-concurrent attributes from the annotated similar images in the repository. To address the multiple aspects of users' interest, it utilizes the browsing context and the associated attributes to enhance the semantic meaning of the original visual objects. The attributes in this scope refer to the textual labels (e.g., gender of the clothes, brand names, and category names) annotated on the visually similar images in the repository, which are discovered through performing the image search. In the client side, a user selects a region that contains a visual object of interest by a natural circling gesture, named a lasso gesture. Context validation aims to extract the valid domain specific attribute from the context for understanding the underlying intent. Due to the present noise in the surrounding texts, it is not guaranteed that all the text in the delivered context can help the server understand the search interest.

B. SELECT IMAGE OF INTEREST BY LASSO GESTURE

It introduces a lasso gesture, which is triggered by a technique called the Lasso Menu that combines selection, command invocation, and parameter adjustment in one fluid stroke. Gesture-based interaction systems are also reported. In our system, it is designed to circle a region to indicate the object of interest inside the image. Users press the lower-left corner of the touch screen using the thumb of the left hand to trigger the session, and then use the right hand to select items

by the typical lasso fashion of drawing a path that encloses them. Once the lasso stroke has begun, a semitransparent orange circle is placed at the beginning of the lasso stroke.

After the drawing of the stroke is completed, it uses the upper-left and lower-right positions of the circled area to form a rectangular bounding area to determine the selected area in the image of interest. The server will download the image by the delivered URL and crop the image by the coordinates of the bounding area. This operation is very natural and convenient for users to specify the search interest. Compared with other gestures like tapping gestures on selecting the whole image, the lasso gesture can help discriminate region of the visual object of interest from the other objects in the same image. Only the circled area represents the user's interest. In addition, it provides robust interaction and does not require drawing an exact silhouette or a closed circle. This lasso gesture requires no training and can naturally enable common users to get used to. Moreover, the gesture-based utility provides a unique user experience with two aspects.

C. CONTEXT DETECTION

In the client side, the nearby texts of the circled image are captured for enriching semantic meaning.In our system, three elements constitute the surrounding textual information title of the web page being browsed. It removes the redundant texts from the three-part constitution and utilize the entity extraction methods to only recognize the general entity words. For example, after combining the nearby captured text, the obtained text is as Fred Perry Men Bleach Washed Chambray Shirt Super Saver Shipping Free Returns. Although the general entity extraction is able to discard the vague and general words or phrases, the extracted entities are still not informative enough for our case. It builds a domain-specific attribute lexicon, and only the entities stored in the lexicon will be extracted and regarded as a valid context. For building a lexicon for the commercial product images, it collects all the possible attributes that appeared in the domain-specific categories.

D.LARGE-SCALE IMAGE SEARCH

To enrich the semantic meaning embedded in the circled image, it performs a large-scale image search to find identical or similar images in the repository and further infer the search interest from the highly concurrent attributes annotated on the resulting images. The interest may be present in the images that share the same local visual patches of the circled region, e.g., logos of the products, background of the scenes, or the similar global visual appearance. The search is performed over an image repository that contains five million images.

To make our system robust on responding the diverse search interests, it collects five million images from Amazon products in the category of clothing, which currently supports 83 clothing product categories. For each image, the product domain-specific attributes such as brand names, prices, and major colors are stored. Since the successive explorations are closely dependent on the obtained results in each iteration, the slow response for performing large-scale image search is not acceptable. In this section, it proposes a solution for ensuring the relevant results are returned in a timely manner.

E. PARTIAL-DUPLICATE IMAGE SEARCH

The identical or near duplicate images to the given circled query images are rarely present in the repository. Commonly, there are discrepancies between two images with similar visual appearance. For example, the background of the product image stored in the repository is relatively clean, but the background of web images is usually cluttered. For performing a large-scale partial-duplicate image search, it follows the state-of-the art bag-of-word (Bow) framework and use the hierarchical vocabulary tree (VT) to construct the vocabulary and inverted index files. Considering that a query image from the bounded lasso region may be a cropped raw image, e.g., the lasso stroke falsely passes through the region of interest and an incomplete object might be included in the bounded area; to overcome this difficulty, it incorporates the GV step to improve the accuracy. To make a robust search, it proposes a visual keyword extraction method to improve the search performance.

i) Vocabulary tree and inverted index

State-of-the-art large-scale image retrieval systems have relied on quantizing local descriptors into visual words, and then applying scalable textual indexing like scheme to construct a VT, which can be used to compare images in a large database in a timely manner. In this paper, VT is constructed by performing hierarchical K-means to group local descriptors into quantized visual code words.

ii) Compact orientation geometric verification

GV becomes an important post processing step for getting a reasonable retrieval precision, but full GV is computationally expensive. To address this dilemma, it proposes a novel scheme by embedding the compact orientation of the SIFT local descriptors into the VT structure and ensure the geometric constraints can be efficiently enforced. It divides the orientation space into $2\pi/r$ parts (subspace), which are indexed as a list as Os = [o1, o2.. , or]. r is defined as seven in this paper.Each local descriptor will be mapped to an orientation subspace and be assigned an orientation index. Consequently, it can only store the index, rather than the orientation information or location information of the features.

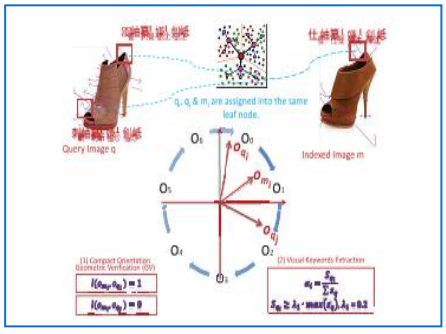

Fig 1. Partial-duplicate image search

iii) Visual keyword extraction

In our experiment, it observes that local descriptors have two different characteristics.Saliency: The larger patches in an image are generally more salient than the smaller ones. Hence, the local patches with the largest scales are selected to identify the visual keywords. Repetitiveness: When projecting the local features of an image into the VT, several assigned nodes will be visited more frequently than other nodes. Based on these observations, it proposes a method to filter out a set of less important nodes with small-scale descriptors or the nodes that are infrequently visited. The remaining one forms the visual keywords that are used to represent an image for the search.

F. SIMILAR IMAGE SEARCH

Partial-duplicate search approaches are utilized to search repository images that may contain similar local patches of the query image. To broadly infer multiple aspects of underlying intent, it also employs a similar image search to find those repository images that share similar global visual appearances (e.g., color, contour, texture, etc.) regardless of the presence of the dissimilarities among local details. Similar image search is also complementary to partial-duplicate image search and enforces diversity of the search results. To feed the feature into the similar image search, it extracts a 576-D global feature vector concatenated by the GIST feature with the dimension of 384 and the gradient feature with the dimension of 192 to represent each image. To avoid increasing system latency, it performs a similar image search running parallel with the partial-duplicate image search. To perform an effective and efficient visual similar search, it applies approximate nearest neighbor (ANN) search approach using multiple complementary hash tables. It can also perform an ANN search over the reference database and discover the 10 most similar images for each reference image, which is regarded as a visual expansion for the reference image.

G. MULTIMODAL QUERY SUGGESTION

In this section, it proposes a method to automatically discover the most representative attributes to represent multifaceted intent. It can use the associated metadata from the top W returned images obtained by large-scale image search and the extracted context by context detection to form an attribute corpus. Then, the system predicts the most likely attribute keywords, suggests image examples to users, and enables them to express a well-formulated intent-specific query. On the server side, it designs a multimodal query suggestion module aiming to provide textual query candidates and the corresponding image examples to users. The textual attributes are determined by the contexts and the attributes obtained from large-scale image search. In this stage, attributes associated with top-ranked images will be recommended to the users.It is jointly determined by the relatedness and the dominance. Relatedness: Each attribute is inherently relevant to the visual appearance of prospected visual content and the surrounding context.Dominance: The selected attributes are with high co-occurrence among all the labeled attributes.*Diversity:*The selected keywords are able to reflect different semantic aspects of the initial search demand

H. JOINT TEXTUAL-IMAGE SEARCH

This section proposes a new strategy to perform a joint keyword-image search approach, given the selected exploratory-search-specific queries of users in the format of the textual keywords and image examples. Since the underlying interest is commonly multifaceted, it is believed that the partial-duplicate images can enrich the exploratory search purpose by providing local visual details, and the similar images can provide the results with the similar global appearance. Hence, it is still performed partial-duplicate image search and similar image search for returning the relevant images, but unlike strategy in, it mainly emphasizesthe differences on how to utilize both visual content and the associated textual attributes simultaneously to perform intent-driven image search. Unlike the partial-duplicate image search described in, the existing content-based VT search scheme is modified and a new attribute-embedded indexing and retrieval method is presented. During the similar image search process, the re-ranking step is modified and takes the selected attributes into the search loop.

I. ADAPTIVE WEIGHT SCHEMA

Humans can easily categorize images into high-level semantic classes such as scene, people or object. The observed images inside these categories are usually agreed on the relative importance of features for similarity calculations. Inspired by this observation, it assigns the query images into several typical categories, and adaptively adjust feature weights within each category. The user first submits query keywords q and a pool of images is retrieved by text-based search. Then the user is asked to select a query image from the image pool, it is classified as one of the predefined adaptive weight categories. Images in the pool are re-ranked based on their visual similarities to the query images, the similarities are computed using the weight specified by the category to combine visual features. In the keyword expansion step words are extracted from the textual descriptions (such as image file names and surrounding texts in the html pages) of the top k images most similar to the query image, and the method is used to rank these words. To save computational cost, only the top m words are reserved as candidates for further processing.

However, because the initial image re-ranking result is still ambiguous and noisy, the top k images may have a large diversity of semantic meanings and cannot be used as visual query expansion. The word with the highest score computed from the top k images is not reliable to be chosen as keyword expansion either. In our approach, reliable keyword expansions are found through further image clustering. For each candidate word w_i, they find all the images containing w_e and group them into different clusters $\{c_{i,1,} c_{i,2\ldots}c_{i,ti}\}$ based on visual content. Images with the same candidate word may have a large diversity in visual content. captures the user's search intention. If they are below certain thresholds, expansion is not used in image re-ranking.

J. BASIC IDEA OF CROSS REFERENCE RE-RANKING

In this research work a new method called CR-Re-ranking method has been introduced, which combines multimodal features in the manner of cross reference. The fundamental idea of CR-Re-ranking lies in the fact that the semantic understanding of image content from different modalities can reach an agreement. Actually, this idea is derived from the multi-view learning strategy, a semi supervised method in machine learning.

In Multi-view learning, first partitions available attributes into disjointed subsets (or views), and then cooperatively uses the information from various views to learn the target model. Its theoretical foundation depends on the assumption that different views are compatible and uncorrelated. In this context, the assumption means that various modalities should be comparable in effectiveness and independent of each other. Multiview strategy has been successfully applied to various research fields, such as concept detection.

However, this strategy, here, is utilized for inferring the most relevant shots in the initial search results, which is different from its original role. CR-Re-ranking method contains three main stages: clustering the initial search results separately in diverse feature spaces, ranking the clusters by their relevance to the query, and hierarchically fusing all the ranked clusters using a cross-reference strategy.

K. IMAGE SEARCH RE-RANKING

Search engine results are often biased towards a certain aspect of a query or towards a certain meaning for ambiguous query terms. Diversification of search results offers a way to supply the user with a better-balanced result set increasing the probability that a user finds at least one document suiting her information need. In this dissertation, to present a re-ranking approach based on minimizing variance of Web search results to improve topic coverage in the top-k results

Web search engines frequently show the same result repeatedly for different queries within the same search session, in essence forgetting when the same documents were already shown to users. Depending on previous user interaction with the repeated results, and the details of the session, to show that sometimes the repeated results should be promoted, while some other times they should be demoted. The three key contributions are made to the image search re-ranking. The first contribution is that multiple modalities are considered individually during clustering and cluster ranking processes.

It means that re-ranking at the cluster level is conducted separately in distinct feature spaces, which provides a possibility for offering higher accuracy on the top-ranked documents. Multimodal features are first concatenated into a unique feature, and the subsequent clustering and cluster ranking are then implemented once in the above unique feature space. The second contribution is defining a strategy for selecting some query-relevant shots to convey users' query intent. Instead of directly treating the top-ranked results as relevant examples like PRF, they further filter out some irrelevant shots using some properties existing in the initial rankings. Reliably selecting a query-relevant shot set has a beneficial effect on cluster ranking.

The third contribution is presenting a cross-reference strategy to hierarchically combine all the ranked clusters from various modalities. They assume that the shot with high relevance should be the one that simultaneously exists in multiple high-ranked clusters from different modalities. Based on this assumption, the shots with high relevance can be inferred cooperatively using the cross-reference strategy and then be brought up to the top of the result list. As a result, the accuracy on the top-ranked documents is given more consideration. Because the "unequal overlap property" is employed implicitly, this fusion strategy is similar to the metasearch methods to a certain extent. However, our cross-reference strategy differs in two ways from meta search. The first difference is that, instead of combining multiple ranked lists from different search engines, they integrate multiple reordered variants of the same result list obtained from only one text-based video search engine. The second one is that, instead of fusing multiple lists at the shot level, to first coarsely rank each list at the cluster level, and then integrate all the resulting clusters hierarchically.

IV.CONCLUSION

The new system eliminates the difficulties in the existing system. It is developed in a user-friendly manner. The system is very fast and any transaction can be viewed or retaken at any level. Error messages are given at each level of input of individual stages. This research work is very particular in reducing the work and achieving the accuracy. It will reduce time be avoids redundancy of data. The user can easily understand the details available from the report. This work will support for the future development.The research work is menu driven. Image can be uploaded and processed very easily. Speed and accuracy is maintained in image processing.Data is entered in formatted manner.The related images can be searched with additional input. Modification and maintenance can be made to web site very easily.

V.FUTURE ENHANCEMENTS

The research work has covered almost all the requirement. Further requirements and improvements can easily be done since the coding in mainly structured or modular in nature. Improvements can be appended by changing the existing modules or adding new modules. Several areas to be developed in future, so the application must be upgraded for the new ones required and it is possible to modifications according to new requirements and specifications. The Future Analysis of this research work as follows research work only developed for the purpose of only image editing. In future, they will plan to add another process of video concept. Facilities fast data backup and restoration facility in case of data loss situations. The images search details can be cached

in server memory so that they can be used in future searches within the near time interval of previous search.

REFERENCES

[1] D. Bainbridge, M. B. Twidale, and D. M. Nichols, "Interactive contextaware user-driven metadata correction in digital libraries," *Int. J. Digital Libraries*, vol. 13, no. 1, pp. 17–32, 2012.

[2] H. Cao *et al.*, "Context-aware query classification," in *Proc. 32nd ACM SIGIR*, 2009, pp. 3–10.

[3] Y. Cao, H. Wang, C. Wang, Z. Li, L. Zhang, and L. Zhang, "MindFinder: Interactive sketch-based image search on millions of images," in *Proc. ACM Multimedia*, 2010, pp. 1605–1608.

[4] V. Chandrasekhar, G. Takacs, D. Chen, S. Tsai, R. Grzeszczuk, and B. Girod, "CHoG: Compressed histogram of gradients a low bit-rate feature descriptor," in *Proc. IEEE CVPR*, Jun. 2009, pp. 2504–2511.

[5] E. Cheng, F. Jing, and L. Zhang, "A unified relevance feedback framework for web image retrieval," *IEEE Trans. Image Process.*, vol. 18, no. 6, pp. 1350–1357, Jun. 2009.

[6] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE CVPR*, Jun. 2009, pp. 248–255.

[7] W.-T. Fu, T. G. Kannampallil, and R. Kang, "Facilitating exploratory search by model-based navigational cues," in Proc. 15th Int. Conf. Intell. User Inter. (ICIUI), 2010, pp. 199–208.