Exploratory Product Image Search With Circle-to-Search Interaction

Dr.C.Sathiyakumar[1], K.Kumar[2], G.Sathish[3], V.Vinitha[4]
K.S.Rangasamy College Of Technology, Tiruchengode, Tamil Nadu, India[2.3.4]
Professor, B.E-CSE, K.S.Rangasamy College Of Technology, Tiruchengode, Tamil Nadu, India[1]

Abstract- Exploratory search is emerging as a new form of information-seeking activity in the research community, which generally combines browsing and searching content together to help users gain additional knowledge and form accurate queries, thereby assisting the users with their seeking and investigation activities. However, there have been few attempts at addressing integrated exploratory search solutions when image browsing is incorporated into the exploring loop. In this paper, we inves-tigate the challenges of understanding users' search interests from the product images being browsed and inferring their actual search intentions. We propose 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 interests in product 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 same or similar product images obtained by large-scale image search technology. Users can then utilize the recommended queries to complete intent-specific exploratory tasks. The proposed solution is one of the first attempts to understand users' interests for a visual-based exploratory product search task by integrating the browse and search activities. We have evaluated our system performance based on five million product images. The evaluation study demonstrates that the proposed system provides accurate intent-driven search results and fast response to exploratory search demands compared with the conventional image search methods, and also, provides users with robust results to satisfy their exploring experience. Exploratory search is emerging as a new form of information-seeking activity in the research community, which generally combines browsing and searching content together to help users gain additional knowledge and form accurate queries, thereby assisting the users with their seeking and investigation activities. However, there have been few attempts at addressing integrated exploratory search solutions when image browsing is incorporated into the exploring loop. In this paper, we inves-tigate the challenges of understanding users' search interests from the product images being browsed and inferring their actual search intentions. We propose 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 interests in product 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 same or similar product images obtained by large-scale image search technology. Users can then utilize the recommended queries to complete intent-specific exploratory tasks. The proposed solution is one of the first attempts to understand users' interests for a visual-based exploratory product search task by integrating the browse and search activities. We have evaluated our system performance based on five million product images. The evaluation study demonstrates

that the proposed system provides accurate intent-driven search results and fast response to exploratory search demands compared with the conventional image search methods, and also, provides users with robust results to satisfy their exploring experience.

1.INTRODUCTION

Analysis the challenges of understanding users search interests from the product images being browsed and inferring their actual search intentions. A novel interactive images 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 interests in product images by circling any visual objects in web pages, and then the system automatically infers users underlying intents by analyzing the browsing context and by analyzing the same or similar product images obtained by large-scale image search technology. Users can then utilize the recommended queries to complete intent specific exploratory tasks. The existing approach to facilities users selecting the interesting visual objects, and then automatically infer the multiple aspects of the interest via large-scale images retrieval, constructing a query entity enriched by the semantic meaning is the eventual purpose.The proposed solution is one of the attempts to understand users interests for a visual-based exploratory product search task by integrating the browse and search activities. Have evaluation our system performance based on product images. The evaluation study demonstrates that the proposed system provides accurate intent-driven search results and fast response to exploratory search demands compared with the conventional image search methods, and also provides users with robust results to satisfy their exploratory experience.The proposed system implements all the existing system aspects. Also presents a flexible and effective re-ranking method called CR Re-Ranking, to improve the retrieval effectiveness. To offer high accuracy on the top-ranked results, CR Re-Ranking employs a cross-reference (CR) strategy. Search quality, especially on the top-ranked results, is improved significantly.

2.EXPERIMENTAL RESEARCH

2.1. 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 the semantic meaning is the eventual purpose. 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.

In this stage, there are two analysis modules.

- 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.
- Large-scale image search aims to find the associated attributes annotated with identical images and similar images in the repository to help enrich the

semantic meanings embedded in images.

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.

Context validation

In the client side, the nearby texts of the circled image are captured for enriching semantic meaning. The surrounding textual information is rich but commonly less informative due to the presence of noise. To effectively detect the representative context from the surrounding texts, hence it proposes a lexicon-based method to validate the meaningful information from the context. In our system, three elements constitute the surrounding textual information:

- Title of the web page being browsed;
- Name of the captured image
- Text from the same div web element nearby the image element.

It removes the redundant texts from the three-part constitution and utilize the entity extraction methods to only recognize the general entity words..

Large-Scale Image Search

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. It performs two content-based visual search approaches in a parallel way

1) Partial-duplicate image search
2) Similar image search

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-theart 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. The proposed partial-duplicate image search approach.

1) Vocabulary tree and inverted index
2) Compact orientation geometric verification
3) Visual keyword extraction

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.

$$S_{Q/x} = \mathrm{Wx}(\mathrm{Hq} - \mathrm{Hx})$$

where

Hx→ the transposed histogram of the xth image;

Wx→term frequency–inverse document frequency (TF-IDF) weighting value

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 2π/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.

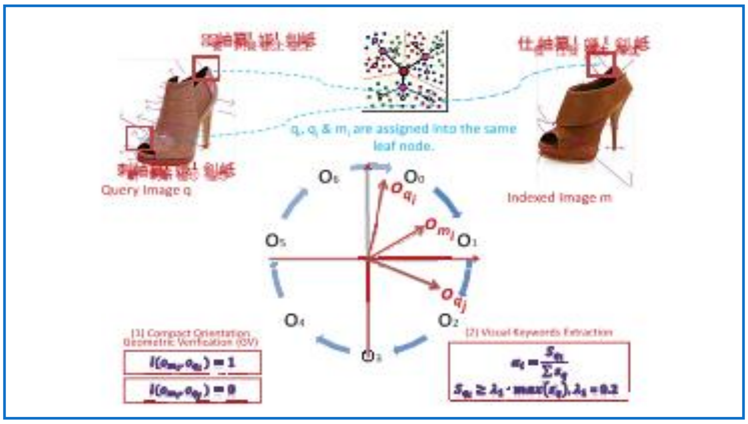
$$W_x^{GV} = \left[\frac{n_l^x}{n^x} \times log\left(\frac{(\bar{C}v_1).n}{nv_1}\right), \ldots, \frac{n_m^x}{n^x} \times log\left(\frac{(\bar{C}v_m).n}{nv_m}\right)\right]$$

$$\text{s.t. } \bar{C}v_{(i)} = \sum_{mi,qi \in V(i)} l(O_{mi}, O_{qi})$$

where

$mi, qi \in V(i)$ → local feature point visit the node V_i,q_i,m_i

$\bar{C}v_{(i)}$→no. of paired local points between query image and matching images and ℓ(.)



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 form the visual keywords that are used to represent an image for the search.

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. The search stage consists of two steps.

1) The first step is to find candidate images with searching the complementary hash tables.

2) The second step is to rerank those candidates and use the 10 most similar images as the visual expansions to represent the query image.

2.2. 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 provides users with visual cues and helps them express the search interests more precisely. On the client side, these suggested multimodality queries provide users many alternatives and help them select an intent-specific query. Keyword suggestion was initially proposed. Their goal is to find a set of keywords to resolve the ambiguity of the initial textual query. In our work, the initial textual query is not applied at the beginning. Hence, it creates a new strategy to seek a set of keywords to construct the queries, which can simultaneously reflect the dominant semantic knowledge of the visual content, context, and different aspects of the semantic knowledge. The recommended attributes in our system should satisfy the following two criteria.

1) **Representativeness**: It is jointly determined by the relatedness and the dominance.

i)Relatedness: Each attribute is inherently relevant to the visual appearance of prospected visual content and the surrounding context.

ii) Dominance: The selected attributes are with high co occurrence among all the labeled attributes.

2) **Diversity:** The selected keywords are able to reflect different semantic aspects of the initial search demand

2.3. 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. . During the similar image search process, the reranking step is modified and takes the selected attributes into the search loop.

- **Attribute-Embedded Image Search**

Attribute-embedded partial-duplicate image search: Attributes are commonly utilized for labeling images at a global level or referring to the partial objects inside the image. To get a reasonable retrieval precision, we further exploit the semantic clues among visual words at a local level and propose a novel partial-duplicate search scheme to encode the attribute value among local features in an image. The attribute can be reintroduced as a post processing step to re rank the retrieved images. Since post verification techniques are computationally expensive, in this paper, we adopt the same strategy as introduced in the GV approach, and we embed the attributes into the VT scheme.

$$W_x^{GV} = \left[\frac{n_l^x}{n^x} \times log\left(\frac{(\bar{C}v_1 + \bar{C}v_1).n}{nv_1}\right), \ldots, \frac{n_m^x}{n^x} \times log\left(\frac{(\bar{C}v_m + \bar{C}v_m).n}{nv_m}\right)\right]$$

s.t. $\bar{C}v_{(i)=} \sum_{mi,qi\in V(i)} l(O_{mi}, O_{qi})$

s.t. $\bar{C}v_{(i)=} \sum_{mi,qi\in V(i)} l(a_{mi}, a_{qi})$

where

$\bar{l}(.)$ → indicator function

$r_i = a_0 s_i + \sum_{\in=1}^{\gamma} l(a_\in) a_\in s_i$

$a_0 + \sum_{\in=1}^{\gamma} a_\in = 1$

where

γ → no. of the selected attributes

- **Search Results with Multiple Aspects**

In the results derived from our system, the top results are generated by the partial-duplicate image search and the remaining results are from the similar image search. In this paper, is set as two. This hybrid combination is driven by the facts that users are more likely to find both partial-duplicate images and visually similar images to enrich their interests. If the partial-duplicate images are not sufficient in the repository, the similar resulting images will replenish the final result list.

2.4.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. The size of the image cluster selected as visual query expansion and its similarity to the query image indicate the confidence that the expansion captures the user's search intention. If they are below certain thresholds, expansion is not used in image re-ranking.

2.5. VISUAL FEATURE DESIGN

Visual Feature Design is used to design and adopt a set of features that are both effective in describing the visual content of images from different aspects, and efficient in their computational and storage complexity. Some of them are existing features proposed in recent years. Some new features are first proposed by us or extensions of existing features. It takes an average of 0.01 ms to compute the similarity between two features on a machine of 3.0 GHz CPU. The total space to store all features for an image is 12 KB. More advanced visual features developed in recent years or in the future can also be incorporated into this framework.

2.6. MULTI MODEL FUSION

Analysis on click-through data from a very large search engine log shows that users are usually interested in the top-ranked portion of returned search results. Therefore, it is crucial for search engines to achieve high accuracy on the top-ranked images. While many methods exist for boosting video search performance, they either pay less attention to the above factor or encounter difficulties in practical applications.

It has been presented with a flexible and effective re-ranking method, called CR-Re-ranking, to improve the retrieval effectiveness. To offer high accuracy on the top-ranked results, CR-Re-ranking employs a cross-reference (CR) strategy to fuse multimodal cues. Specifically, multimodal features are first utilized separately to re-rank the initial returned results at the cluster level, and then all the ranked clusters from different modalities are cooperatively used to infer the shots with high relevance. Experimental results show that the search quality, especially on the top-ranked results, is improved significantly.

2.7. CONTENT-BASED IMAGE RETRIEVAL

"Content-based" means that the search analyzes the contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. The term "content" in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself. CBIR is desirable because most web-based image search engines rely purely on metadata and this produces a lot of garbage in the results.

It also has method of manually entering keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes the image. Thus a system that can filter images based on their content would provide better indexing and return more accurate results. There is a growing interest in CBIR because of the limitations inherent in metadata-based systems, as well as the large range of possible uses for efficient image retrieval. Textual information about images can be easily searched using existing technology, but this requires humans to manually describe each image in the database.

2.8. BASIC IDEA OF CROSS REFERENCE RE-RANKING

In this 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.

2.9. 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

2.10. ENHANCED K-MEANS CLUSTERING ALGORITHM

In form the number of specified clusters. The k-means method first selects a set of n points called cluster seeds as a first guess of the means of the clusters. Each observation is assigned to the nearest seed to form a set of temporary clusters. The seeds are then replaced by the cluster means, the points are reassigned, and the process continues until no further changes occur in the clusters.

- Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
- Assign each object to the group that has the closest centroid.
- When all objects have been assigned, recalculate the positions of the K centroids.
- Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.
- The dataset is partitioned into K clusters and the data points are

randomly assigned to the clusters resulting in clusters that have roughly the same number of data points.

- For each data point: Calculate the distance from the data point to each cluster.
- If the data point is closest to its own cluster, leave it where it is. If the data point is not closest to its own cluster, move it into the closest cluster.
- Repeat the above step until a complete pass through all the data points results in no data point moving from one cluster to another. At this point the clusters are stable and the clustering process ends.
- The choice of initial partition can greatly affect the final clusters that result, in terms of inter-cluster and intracluster distances and cohesion.

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.

A trial run of the system has been made and is giving good results the procedures for processing is simple and regular order. The process of preparing plans been missed out which might be considered for further modification of the application.