

IGroup: Web Image Search Results Clustering

Feng Jing¹, Changhu Wang², Yuhuan Yao³, Kefeng Deng¹, Lei Zhang¹, Wei-Ying Ma¹
Microsoft Research Asia¹ Department of EEIS, University of Academy of Art Design
49 Zhichun Road Science and Technology of China² Tsinghua University³
Beijing 100080, China Hefei 230027, China Beijing 100080, China
(86-10)58963609 (86)13581984028 (86-10)62777883

{fengjing, kefengd, leizhang,
wyma}@microsoft.com

wch@ustc.edu

crisyao@hotmail.com

ABSTRACT

In this paper, we propose, IGroup, an efficient and effective algorithm that organizes Web image search results into clusters. IGroup is different from all existing Web image search results clustering algorithms that only cluster the top few images using visual or textual features. Our proposed algorithm first identifies several query-related semantic clusters based on a key phrases extraction algorithm originally proposed for clustering general Web search results. Then, all the resulting images are separated and assigned to corresponding clusters. As a result, all the resulting images are organized into a clustering structure with semantic level. To make the best use of the clustering results, a new user interface (UI) is proposed. Different from existing Web image search interfaces, which show only a limited number of suggested query terms or representative image thumbnails of some clusters, the proposed interface displays both representative thumbnails and appropriate titles of semantically coherent image clusters. Comprehensive user studies have been completed to evaluate both the clustering algorithm and the new UI.

Categories and Subject Descriptors

H.3.3 [Information storage and retrieval]: Information Search and Retrieval – *clustering*. H.3.5 [Information storage and retrieval]: Online Information Services – *Web-based services*. H.5.3 [Information interfaces and presentation]: Group and Organization Interfaces – *organizational design, Web-based interaction*.

General Terms

Algorithms, Design, Experimentation.

Keywords

Web image search, search result clustering, user interface design.

1. INTRODUCTION

Image search is a long-standing research problem. Before the 1990s, images were manually labeled with texts based on which

they were further indexed and retrieved [18]. Since the early 1990s, because of the emergence of large-scale image collections, manual annotation became more and more impractical due to its tedious and time-consuming nature. To overcome the difficulty, content-based image retrieval (CBIR) was introduced ([23], [1]). In CBIR, images are indexed by their visual content, such as color or texture which could be extracted automatically. However, after over a decade's research, the performance of CBIR is still far from satisfaction and none of successful commercial image retrieval systems is currently based up on it. There are two main reasons for CBIR's lack of success as a technology in commercial image search systems. One is the so-called semantic gap between low-level visual features and high-level semantics. To reduce or bridge the gap, more semantic features should be utilized. Fortunately, such features are available in the Web image search scenario. For Web images, there are rich textual features, such as name, alt text, Uniform Resource Locator (URL) and the surrounding text [10]. These textual features are much closer to the semantics of the images than the visual features and have proved to be very helpful in many commercial Web image search engines ([27], [24], [29]). The other reason for disregarding visual features on a large scale image search, e.g. Web image search, is due to the low efficiency. Unlike textual features that can be indexed by inverted files in very efficient ways [15], visual features are generally indexed by tree structures [6]. However, since the dimensions of visual features are high, e.g. more than 100, tree-based indexing algorithms are very inefficient due to the curse of dimensionality [6]. Although in early research work ([5], [19], [12]), textual features and visual features were combined into an integrated system, because of the above problems visual features were eventually discarded in current practical commercial Web image search engines. Inspired by the huge success of PageRank [17] in Google's Web search [25], PicASHOW [16] used hyperlink analysis in Web image search. One basic assumption of PicASHOW is that images that are contained in authoritative pages, i.e. pages with high PageRank are good candidates to be high quality images [16]. However, this assumption is not sound. In fact, photos on some personal homepages¹ that have comparatively high PageRank are personal photos, and are usually of little interest to Web users. And there is also a large number of high quality photos contained in pages

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM'06, October 23-27, 2006, Santa Barbara, California, USA
Copyright 2006 ACM 1-59593-447-2/06/0010...\$5.00.

¹<http://www.columbia.edu/cu/sigmachi/Album/Weekend Party in February 2004.htm>

with rather low PageRank, e.g. photos of some photo forum sites.

2

1.1 Page Search Results Clustering (PSRC)

Existing search engines such as Google [25] and Yahoo [28] return ranked lists of Web pages in response to a user's query request. Web users have to sift through the list to locate pages of interest. This is a time-consuming task when multiple sub-topics of the given query are mixed together. A possible solution to this problem is to cluster search results into different groups and to enable users to identify their required group at a glance. Hearst and Pedersen [13] applied a traditional clustering algorithm. It first clustered documents into topically-coherent groups according to content similarity, and then generated descriptive summaries for the clusters. However, these summaries are often unreadable, which makes it difficult for Web users to identify relevant clusters. Zamir and Etzioni [14] presented a Suffix Tree Clustering algorithm that first identified sets of documents that share common phrases, and then created clusters according to these phrases. Zeng et al [8] also formulated the clustering process as a candidate phrase extraction process. The main characteristic of [8] is that it utilized learning methods to rank these salient phrases.

1.2 Image Search Results Clustering (ISRC)

Web image search shares some common features with Web page search. On the one hand, image queries also tend to be short and sometimes ambiguous. For example the average query length of the top (tail) 1,000 queries of Picsearch [27] is 1.368 (1.677) words. Moreover, 97% (92%) of the top (tail) queries contain one or two words. On the other hand, for most of the queries, tens of thousands of images will be returned. More specifically, on average each of the top 1,000 queries has 635,821 resulting images. 20% (55%, 87%) of the top queries have 1,000,000 (100,000, 10,000) resulting images. More important than the numbers is the fact that images on different topics are intermixed in the returned list, so the user has to sift through a long undifferentiated list to find images of interest. For example, images of apple fruit are intermixed with images of apple computer, apple ipod, apple logo, apple pie, Fiona Apple, and so on. To worsen the situation, for every query, most of the current image search engines ([24], [29]) only let the users see at most 1,000 images. All these facts motivate the clustering of image search results as the clustering of Web pages.

On the other hand, different from Web page search, Web image search has the following three unique properties. First, unlike a textual document, the content of an image can be taken in at a glance. For example, we can get the "gist" of an image in 110 ms or less, changing fixation roughly every 300 ms [20]. In that 110 ms, on average we read less than one word, or skim two words [2]. As a result, a user will possibly go through more files and pages in image search than in page search if both the quality and organization of the image search results are good. Second, different from Web pages -- the quality of which could be assessed by PageRank -- the quality of images is very hard to assess. Therefore, images in the first resulting page of some query are not necessarily better than those in the following pages in terms of both their relevance to the query and their quality. Last,

the query need of page search and image search are different. For page search, users often have specific page(s) in mind, e.g. the page(s) visited before. While for image search, most of the times, users don't have such clear expectations. All these properties make clustering more useful for image search results than the page counterpart.

Several image search result clustering (ISRC) algorithms have recently been proposed in the academic arena. In ([4], [9]), top result images were clustered based on visual features so that the images in the same cluster are visually similar. Considering that global image features do not describe individual objects in the images precisely, X.J. Wang [22] proposed to use region level image analysis. They formalized the problem as a salient image region pattern extraction problem. According to the region patterns, images were assigned to different clusters. Besides visual information, textual and link information has also been used recently. A reinforcement clustering algorithm and a bipartite graph co-partitioning algorithm are proposed to integrate visual and textual features in [21] and [3] respectively. The link information is also used in [7]. It first used block-level link analysis to construct an image graph. Then, spectral clustering techniques were adopted to hierarchically cluster the top image search results based on visual representation, textual representation, and graph representation. Leaving alone the clustering results, the efficiency of the algorithm, which was not mentioned, is very likely to be low due to the complicated preprocessing and heavy clustering algorithm.

The effectiveness of these visual feature-based approaches ([4], [7], [22], [9]) depends heavily on clustering performance and the quality of representative images of each cluster. Due to the continuous property of visual feature space, the boundary between different images is not very clear. It is not surprising that two images with some common attributes that users prefer will be grouped into different clusters. Under such circumstances, additional efforts are required to navigate around multiple clusters.

Besides the effectiveness issue, there are three severe limitations of the existing ISRC algorithms:

- Considering that ISRC is an online process, clustering of hundreds of images using high dimensional features is not efficient enough to be practical.
- Considering that for common queries, hundreds of thousands or even millions of images will be returned, 1,000 images are too few to be representative (less than 1% or even 0.1%). As a result, the helpfulness of the clustering results of these unrepresentative images will be rather limited.
- The clusters do not have semantic names as in PSRC. One reason for this is that the clusters are not semantically consistent.

For the above reasons, no commercial Web image search engine uses ISRC.

In this paper, we propose IGroup as an efficient and effective algorithm to organize Web image search results into semantic clusters. Different from all existing Web ISRC algorithms that only cluster the top images using visual or link features, the proposed algorithm first identifies several semantic clusters

² <http://www.photosig.com/go/photos/view?id=1605697&forward>

related to the query, then assigns all the resulting images to corresponding clusters. Our algorithm has the following three unique advantages over existing ISRC algorithms. First the most *representative* image groups can be found with *meaningful names*. Second, *all* resulting images are taken into account in the clustering process instead of only a small part. Last, our algorithm is *efficient* enough to be implemented in practical systems.

To make the best of the grouping results, a new user interface is proposed. Different from existing Web image search interfaces that show only limited number of suggested query terms ([27], [29]) or some representative image thumbnails of some clusters ([4], [7], [22], [3]), the proposed interface shows both the representative thumbnails and appropriate titles of the semantically coherent image groupings. Moreover, together with the thumbnail images, more informative description (cluster name) is shown instead filename. An online demo of our system (IGroup) is available at: <http://igroup.msra.cn>

The organization of the paper is as follows. Section 2 and 3 representatively describe the clustering algorithm and user interface. In Section 4, extensive user study results are presented and analyzed. We conclude in Section 5.

2. ISRC ALGORITHM

2.1 Learning Image Cluster Names

2.1.1 Learning candidate image cluster names

The candidate image cluster names are generated from two sources. One is the salient phrases extracted from the clustering results of Google’s web page search [25]. The other is from the suggested phrases of an image search engine, i.e. Picsearch [27]. For the former, we use the algorithm proposed in [8]. [8] reformalizes the clustering problem as a salient phrase ranking problem. Given a query and the ranked list of search results, it first parses the whole list of titles and snippets, extracts all possible phrases (n-grams) from the contents, and calculates several properties for each phrase such as phrase frequencies, document frequencies, phrase length, and more. A regression learning model from previous training data is then applied to combine these properties into a single salience score. The phrases are ranked according to the salience score, and the top-ranked phrases are taken as salient phrases. The salient phrases are further merged according to their corresponding documents. An online demo showing the algorithm of [8] is [26]. The resulting salient phrases are one source of the candidate image cluster names. On the other hand, Picsearch [27] will suggest up to five related phrases for each query. These suggested phrases are another source of the candidate image cluster names.

2.1.2 Merging and pruning cluster names

Given the candidate cluster names, a merging and pruning algorithm is utilized to obtain the final cluster names. First, we merged the same or very similar candidates from different sources. Second, the synonyms of “images,” e.g. “pictures” or “photos” are utilized to prune the candidate cluster names of possibly unhelpful clusters. Finally, the resulting candidate cluster names are used as queries to search an image search engine, e.g. Google image search [24] with the number of resulting images counted. The cluster names with too many or too few resulting images are

further pruned. Each of the remaining cluster names corresponds to a cluster that contains the images returned by the search engine using the cluster name as query. The reduced thumbnails of top ranked images are used as representative images of the clusters.

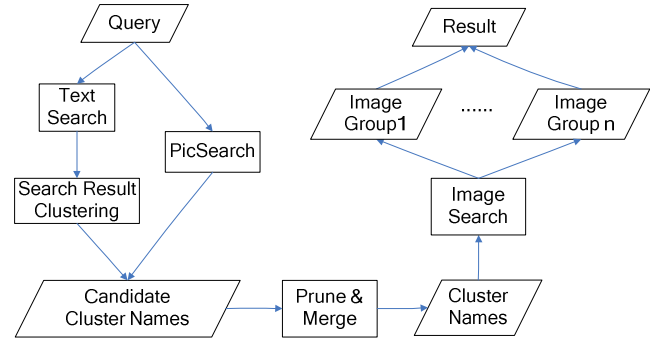


Figure 1. Flowchart of image search result clustering algorithm

The proposed ISRC algorithm is shown in Figure 1. Semantic clustering results for query “tiger” and query “moon” are shown in Figure 11. Note that the numbers within parentheses are the numbers of returned images using the phrases as queries.

As the byproducts, two problems of the existing Web image search engines are solved to some extent by our algorithm. One is that with the existing image search engines, one kind of images tends to dominant the search results. For example, for query “tiger”, most of the top result images are tiger animal. While with IGroup, other tiger related images, e.g. tiger tank and tiger shark, could be easily seen and accessed by the users. The other limitation is that for some general queries, especially those people name related queries, e.g. Linda, Eva and so on, many personal photos that will be uninterested to most users will be ranked very high. While with IGroup, since the most related key phrases could be found for each query, using the key phrases as supplement most of the personal images could be filtered out and more “interesting” images could come to the surface.

2.2 Efficiency Analysis

As we discussed earlier, efficiency issue is very critical for image search result clustering. Considering that ISRC is performed online on the server side, real time implementation is required. To analyze the efficiency of the proposed algorithm, we decompose the current implementation of the proposed ISRC algorithm into six components: 1. perform PSRC; 2. query Picsearch, download the first resulting page and parse the page to get the suggested phrases; 3. query Google image search with the candidate cluster names, download the resulting pages and parse them to get the resulting image numbers; 4. merge and prune the candidate image cluster names; 5. query Google image search engine with the cluster names; 6. download the first resulting page and parse it to get the top images. For each component, we calculate the average time cost. Since query length will affect the calculation of several components. We evaluate the time cost with one word queries, two words queries, and more than two words queries. Table 1 shows the time cost of each component and the whole process with different query lengths (the time value is an average of 20 queries from the top 1,000 queries of a commercial image search engine). As can be seen from the table, the top three time-consuming components are component 3, 6 and

2. However, imagining that we own an image search engine, such as Picsearch, component 3 will be simplified to fetch the size of the corresponding inverted files [15], component 6 will be simplified to select the thumbnails of the top images and component 2 will be simplified to fetch the pre-calculated related phrases. The time cost of all three components will be low enough to be neglected. As a result, the total time cost for single (two, more than two) word (s) queries will be around 3s (2.48s, 2.55s).

Table 1. Mean time in seconds of each component and entire process with different query length

	SRC	Pic	Count	Post	Search	D&P	Total
1	1.09	1.41	5.00	0.47	1.43	3.61	13.03
2	1.16	1.37	4.31	0.50	0.81	2.33	10.49
>2	1.24	1.38	4.40	0.55	0.75	2.75	11.08

3. ISRC USER INTERFACE

To make the best of the clustering results, a new Web image search UI is proposed with several variants. The UI basically consists of three views: search utility view, navigation view and list view. A snapshot of the UI of IGroup is shown in Figure 2.

3.1 Search Utility View

The search utility view embedded in the top frame as shown in Figure 2 provides two ways to formulate queries, i.e. typing query in the edit box and choosing query from the category lists. The search utility view also provides three search options which we will discuss later in the list view section.

3.2 Navigation View

In the navigation view, we show basic information of clusters, i.e. the cluster name and representative downscaled thumbnails. The clusters are sorted in descending order of the number of image they contain. When the cluster name is clicked, the corresponding images are shown in the list view. Moreover, to cover all resulting images, an “others” cluster is formed by using the “-” operator [24]. For example, assuming that the query is “A” and there are three clusters “A B,” “A C,” and “A D,” then “others” cluster will contain resulting images of query “A -B -C -D.” The “others” cluster is shown at the bottom.

3.3 List View

The list view has two variants. One is a cluster view that shows each cluster with four representative images and the corresponding cluster name (cluster ID). The other is very similar to the result view of existing Web image search engines ([24], [29], [27]). Thumbnails of resulting images are shown with related information, e.g. filename and file size. Like Google image search [24], 20 images are shown in one page. When original query is used to search, the images shown in the list view will depend on the current search option. When “List” is chosen, all resulting images will be displayed and ranked in the same order as in [24]. When “I feel lucky” is chosen, only images of the first (largest) cluster are shown. Finally, when “Mix” is chosen, all clusters are shown. In the list view, when a cluster is clicked, corresponding images within the cluster will be shown. In

the conventional view, when an image is clicked, the original image will be shown.

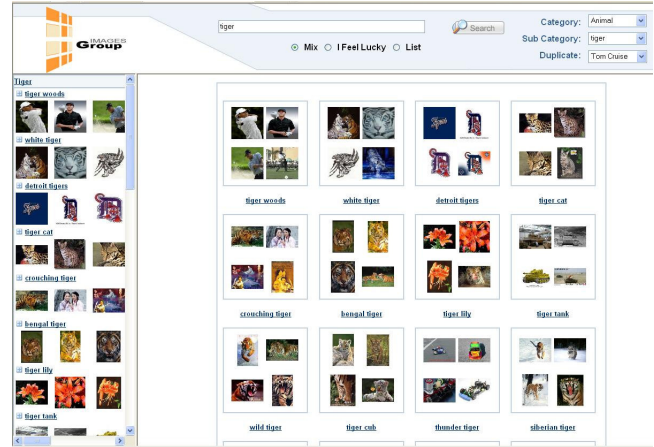


Figure 2. UI Snapshot of UI of IGroup.

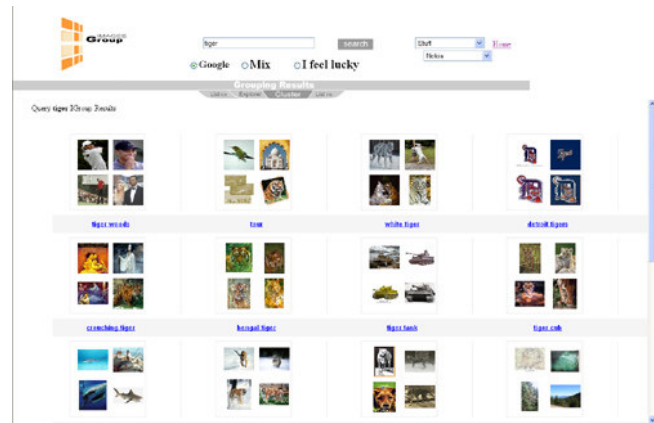


Figure 3. Cluster UI of IGroup

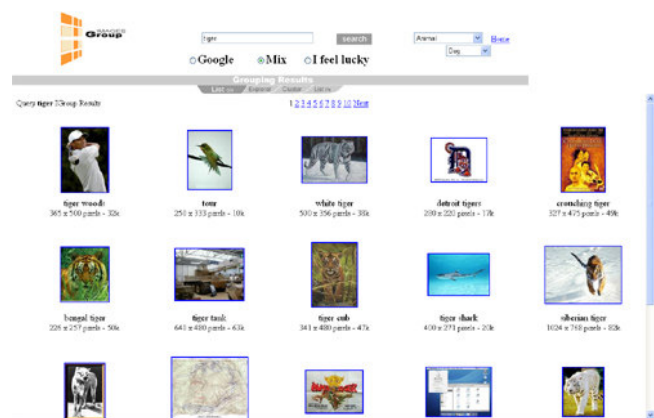


Figure 4. List UI of IGroup

4. USER STUDY

4.1 Participants

We asked 20 participants (5 female and 15 male) to be involved in all or part of the following user study. The participants said they were regular users of the Internet, searching for information either daily or a few times a week. They searched for images online less

frequently, with more than half of them searching for images several times a week and others less than once per week.

4.2 Apparatus

Participants received a mini-radiogram as a gift for participating in a session of about an hour. We partitioned the participants into four groups. Each group was tested on a machine using Internet Explorer 6 of Windows XP with a 17 inch LCD monitor set at 1280×1024 pixels in 24-bit color. Data was recorded with multiple methods: 1) server logs; 2) behavioral logs (time stamps) and 3) paper surveys after each interface, each task and at the end of the session. One experienced usability analyst conducted the sessions.

4.3 UI Evaluation

To facilitate and accelerate the evaluation process, we manually selected 305 queries. For these selected queries, cluster names were extracted offline and related pages were downloaded and parsed in advance. The selected queries are categorized into 14 categories: animal, cartoon, female celebrities, holiday, male celebrities, movie, nature, numbers, people, places, sports, stuff, vehicle and others. All the selected queries except those in the number category are from the top 1,000 queries of Picsearch [27]. Some representative queries are Britney Spears, Angelina Jolie, Brad Pitt, sex, tsunami, hurricane, and so on.

Three UIs were implemented and evaluated: the explorer UI (Explorer) in Figure 2, a cluster UI (Cluster) in Figure 3 and a conventional list UI (List) in Figure 4. For Explorer, the Mix mode is used. For Cluster, four representative images were shown with the same size as those of navigation view of Explorer. For List, images are shown in the same order as in the list view of Explorer. Different from the list view of Explorer, when images of List UIs are clicked, original images instead of the images in corresponding clusters are shown.

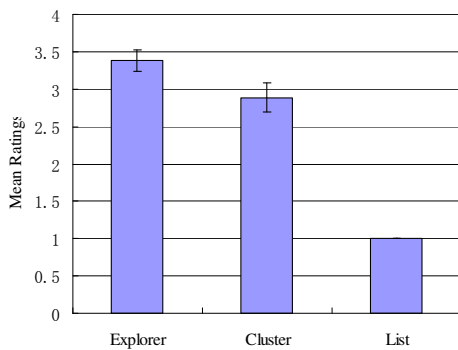


Figure 5. Mean ratings of three UIs.

Eighteen participants were involved in the evaluation. The participants were asked to evaluate all three UIs with arbitrary queries they liked. The UIs were presented to the participants in a random order. They were asked to rank the three UIs using “>”, “>>” or “=”. In this evaluation and the following evaluations of this paper, “>” means better, “>>” means much better and “=” means no difference. To quantitate the results, we convert the ranking into ratings. We assigned the numeral 1 to the worst one. If $A > (>>, =) B$, the rating of A will be the rating of B plus 1 (2, 0). The average ratings are shown in Figure 5. The figure indicates that Explorer > Cluster > List. We also performed the

ANOVA test. The test result is $F(2, 53) = 84.3$, $p \ll 0.00001$ which indicates the significant difference of the three UIs. Moreover, $F(17, 53) = 1.13$, $p > 0.36$ which means the difference among participants is statistically insignificant. Furthermore, ANOVA test of Cluster and List was performed. The result is $F(1, 35) = 92.7$, $p < 0.0000001$, which means Cluster >> List. We also collected the comments of users on the UIs. The main reason people who prefer Cluster most is that more information could be gained from one screen with Cluster.

4.4 Algorithm Evaluation

4.4.1 Extracting cluster names

Considering that the cluster names are from two sources: Picsearch and PSRC, we evaluated three cluster name extraction strategies: Picsearch only (Pic), PSRC only (SRC) and combination of them (P&S). The cluster names are shown to the participants with representative thumbnail images. The average ratings are shown in Figure 6. The figure indicates that P&S > SRC >> Pic. The ANOVA test result is $F(2, 53) = 71.17$, $p \ll 0.00001$. The reason that Pic is inferior to SRC is due to the too limited phrases (at most five phrases) Pic could suggest. For some longer queries, such as “lord of the rings,” no phrases were suggested in Picsearch, while 17 helpful phrases could be extracted from SRC, e.g. king, middle earth, Irr Tolkien, Frodo and so on. On the other hand, for some queries, Pic does suggest very useful queries that SRC could not produce. For example, for query “Nokia”, Pic will suggest “Nokia 6600,” “Nokia 6230,” “Nokia 3200,” and “Nokia 3220.” Therefore, the combination of Pic and SRC, i.e. P&S is the best choice.

4.4.2 Semantic clustering vs. visual clustering

To show the effectiveness of the proposed ISRC algorithm, we compared it with [4]. For [4], we randomly selected the following queries: snowboard, Christmas, sunset, Hawaii, Troy, South Park, David Beckham, Britney Spears, and Mercedes. For each query, the color correlogram [11] of top 1,000 result images were calculated. Based on the features, k-means algorithm was used to cluster the images into 20 clusters. The two algorithms were presented to the participants in a random order. All 18 participants considered our algorithm was significantly better (>>) than [4]. They gave two main reasons. One was that the clusters of [4] are semantically inconsistent. The other reason was that the clusters do not have meaningful names. As a result, it is very hard for the users to tell what the clusters really are.

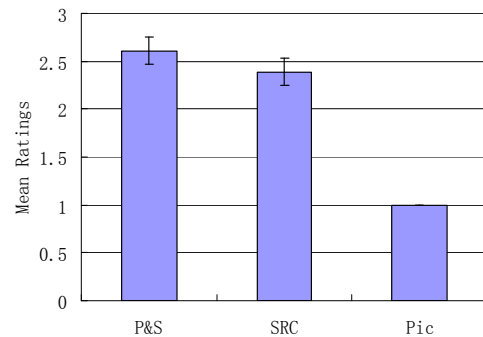


Figure 6. Mean ratings of three cluster name extraction algorithms.

4.4.3 IGroup vs. Google image search

To compare the proposed IGroup system with existing image search engines, e.g. Google image search [24], we defined three specific tasks. The first task is assuming that you will travel to London, search for 10 representative city images. The second task is assuming that you are a fan of Tom Cruise and want to build a homepage about him, search for 10 material images to use on the page. The third task is assuming that you will introduce dog to a little child, search for 10 dog images. All 20 participants were involved in the three tasks.

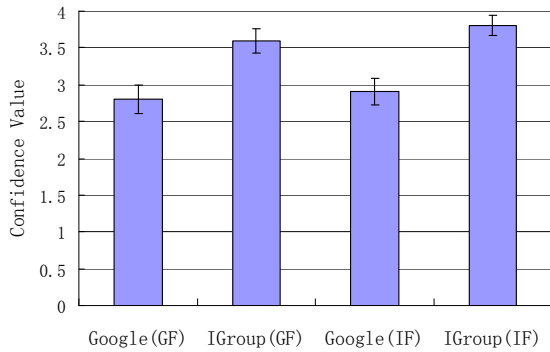


Figure 7. Confidence value comparison of Google and IGroup for task 1. GF (IF) indicates that Google (IGroup) is viewed first.

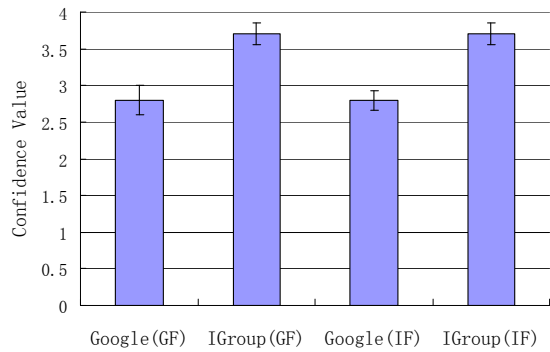


Figure 8. Confidence value comparison of Google and IGroup for Task 2.

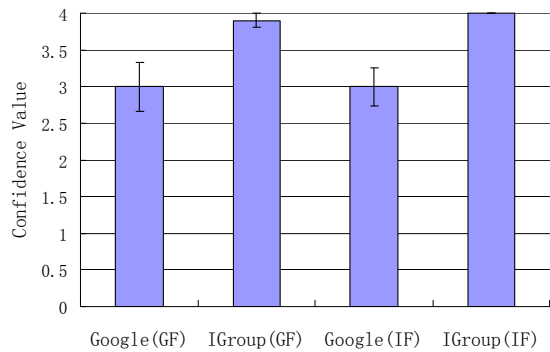


Figure 9. Confidence value comparison of Google and IGroup for task 3.

Since the order in which interfaces were viewed affected the ratings of the results, each interface was used as the starting view for half the participants.

After each task, participants completed a short questionnaire containing two questions. One was “Are you confident that the images you found are relevant images?” The other was “Are you satisfied with the results?” For the former, the participants were required to select from four options for Google or IGroup. The options were very confident, somewhat confident, unconfident, and very unconfident. For the latter, similar options were used: very satisfied, somewhat satisfied, unsatisfied, and very unsatisfied. As in the UI evaluation, the results are quantitated to facilitate further statistical analysis. Very confident (satisfied), somewhat confident (satisfied), unconfident (unsatisfied) and very unconfident (unsatisfied) correspond to rating 4, 3, 2, and 1. The average confidence (satisfaction) scores of task 1, 2, 3 are shown in Figure 7, 8, 9 (10, 11, 12). We also performed the ANOVA test. For all the comparisons, $F(1, 19) > 5.4$ and $p < 0.04$. Both confidence and satisfaction scores of IGroup are obviously higher than those of Google.

A good way to use IGroup is to first click most of the cluster names in the navigation view so that an overview of the cluster results could be gained. Then, clusters can be chosen to acquire additional details. Because users are so used to the existing image search UI, only eight of them used IGroup in the aforementioned way or very similar ways. However, for the users who did use IGroup appropriately, comments were highly positive. For example, all of the aforementioned eight users chose both very confident on the relevance of and very satisfied with the images found using IGroup for all three tasks.

4.5 Overall Comments

Seventeen participants were asked to complete a questionnaire to provide overall comments on the clustering algorithm and UI of IGroup.

First, the participants were asked to state negative impressions of the clustering results. According to the comments, two limitations existed. One was that for some queries, there are too many similar clusters. The other limitation was that the clustering process was slow for some queries.

Second, the participants were asked to state negative impressions of the UI of IGroup. Most were about the list view. According to users’ comments, the three ranking strategies have their own advantages. Therefore, we decided to support all three search options and let users choose favorite options. Another good suggestion is to use a draggable splitter to enable users to see more representative images of each cluster. We have implemented it in the current system.

Last, the participants were asked to answer the question of whether they would choose to use IGroup instead of Google image search. Four participants said that they would *always* use IGroup. Three participants said that IGroup will always be their first choice unless they can’t find desired images. Other participants gave three conditions that they would prefer in regard to using IGroup. The first is when they are not very familiar with the target images they want to find. The second condition is when they are not very clear about what they want. The third condition is they want to collect images about many aspects of a particular individual or subject area. Under these situations, it is hard to

formulate proper queries. Therefore, clustering results would be found to be advantageous.

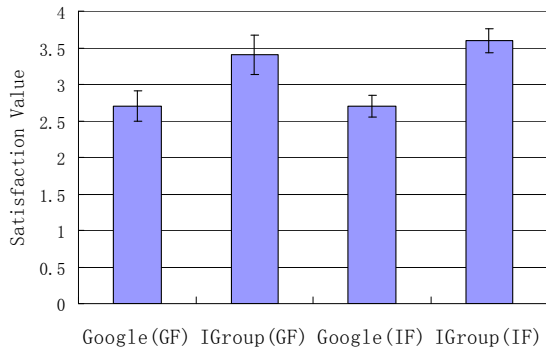


Figure 10. Satisfaction value comparison of Google and IGroup for task 1. GF (IF) indicates that Google (IGroup) is viewed first.

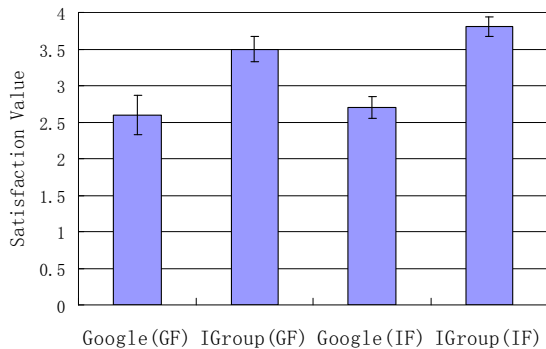


Figure 11. Satisfaction value comparison of Google and IGroup for task 2.

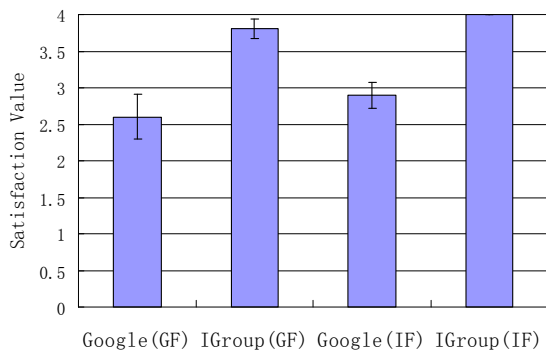


Figure 12. Satisfaction value comparison of Google and IGroup for task 3.

5. CONCLUSION

In this paper, we described IGroup an image search engine based on an effective and efficient image search result clustering algorithm. It is different from all existing image search result clustering algorithms that only cluster a small part of all resulting images using visual or link information. The proposed algorithm first identifies several semantic clusters related to the query, based on a search result clustering algorithm for general Web search. It

then assigns all the resulting images to the corresponding clusters. Our algorithm has the following three unique features. First, with our algorithm, the most representative image groups could be found with *meaningful names*. Second, *all* resulting images are taken into account in the clustering process instead of the small part. Lastly, our algorithm is *efficient* enough to be practical. To make the best use of the proposed clustering algorithm, a new UI was also proposed. Comprehensive user studies show the effectiveness of the clustering algorithm and the superiority of IGroup over existing image search engines, e.g. Google [24]. To get real experience of the system, please refer to our online demo: <http://igroup.msra.cn>

6. REFERENCES

- [1] A. W. M. Smeulders, et al. "Content-based image retrieval: the end of the early years". IEEE transaction on Pattern Analysis and Machine Intelligence, 22-12, 2000. pp. 1349-1380
- [2] A. Woodruff, A. Faulring, R. Rosenholtz, J. Morrison and P. Pirolli, "Using Thumbnails to Search the Web." Proceedings of the SIGCHI conference on Human factors in computing systems, pp: 198 – 205, 2001.
- [3] B. Gao et al. "Web image clustering by consistent utilization of visual features and surrounding texts." Proc. of ACM Multimedia 2005.
- [4] B. Luo, X. G. Wang, and X. O. Tang. "A World Wide Web Based Image Search Engine Using Text and Image Content Features." in Proc. of IS&T/SPIE Electronic Imaging 2003, Internet Imaging IV.
- [5] C. Frankel, M. Swain, and V. Athitsos, "WebSeer: An image search engine for the world wide Web", TR-96-14, Department of Computer Science, University of Chicago, 1996.
- [6] D. A. White, and R. Jain, "Algorithms and strategies for similarity retrieval," Storage and Retrieval in Image, and Video Databases, vol. 2,060, pp. 62-72, 1996
- [7] D. Cai, X. F. He, Z. W. Li, W. Y. Ma and J. R. Wen, "Hierarchical Clustering of WWW Image Search Results Using Visual, Textual and Link Analysis." Proceedings of the 12th annual ACM international conference on Multimedia, pp: 952-959.
- [8] H. J. Zeng, Q. C. He, Z. Chen, W. Y. Ma and J. W. Ma, "Learning to cluster Web search results", Proceedings of the 27th annual international ACM SIGIR conference, pp: 210 – 217.
- [9] H. Liu, X. Xie, X. O. Tang, Z. W. Li and W. Y. Ma, "Effective Browsing of Web Image Search Results." Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval, pp: 84-90.
- [10] H. T. Shen, et al, "Giving Meanings to WWW Images". Proc. ACM International Multimedia Conference, 2000. pp. 39-48.
- [11] J. Huang, et al, "Image indexing using color correlograms". In Proc. IEEE Comp. Soc. Conf. Comp. Vis. and Patt. Rec., pages 762--768, 1997.

[12] J. Smith and S.-F. Chang, "WebSEEK, a content-based image and video search and catalog tool for the Web", IEEE Multimedia, 1997.

[13] M.A. Hearst and J.O. Pedersen, "Reexamining the Cluster Hypothesis: Scatter/Gather on Retrieval Results," Proc. of the 19th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'96), pp. 76--84, 1996.

[14] O. Zamir and O. Etzioni. "Web document clustering: A feasibility demonstration." In Proceedings of SIGIR'98, pages 46-54

[15] R. Baeza-Yates, and B. Ribeiro-Neto, "Modern Information Retrieval". Addison-Wesley, June 1999.

[16] R. Lempel and A. Soffer, "PicASHOW: Pictorial authority search by hyperlinks on the Web", Proc. 10th Int. World Wide Web Conf., pp. 438-448, Hong Kong, China, 2001.

[17] S. Brin and L. Page, "The anatomy of a large-scale hypertextual (Web) search engine", In The Seventh International World Wide Web Conference, 1998.

[18] S. K. Chang, and A. Hsu, "Image information systems: Where do we go from here?" IEEE Transaction on Knowledge and Data Engineering, 4(5), Oct. 1992, pp. 431-442.

[19] S. Sclaroff, L. Taycher, and M. LaCascia, "ImageRover: a content-based image browser for the world wide Web", in IEEE workshop on content-based access of image and video libraries, pages 2-9, San Juan, Puerto Rico, June 1994.

[20] V. Coltheart, (ed.). "Fleeting Memories: Cognition of Brief Visual Stimuli." MIT Press: Cambridge, MA, 1999, pp. 32-70.

[21] X.J. Wang, W.Y. Ma, L. Zhang and X. Li, "Iteratively clustering Web images based on link and attribute reinforcements", Proc. of ACM Multimedia 2005.

[22] X. J. Wang, W. Y. Ma, Q. C. He and X. Li, "Grouping Web Image Search Result", Proceedings of the 12th annual ACM international conference on Multimedia, pp: 436-439.

[23] Y. Rui, T.S. Huang, and S.-F. Chang, "Image retrieval: current techniques, promising directions and open issues", Journal of Visual Communication and Image Representation, Vol. 10, 39-62, March, 1999.

[24] Google image search, <http://images.google.com>

[25] Google Web search, <http://www.google.com>

[26] MSRA clustering search, <http://wsm.directtaps.net/>

[27] Picsearch image search, <http://www.picsearch.com>

[28] Yahoo homepage, <http://www.yahoo.com/>

[29] Yahoo image search, <http://images.search.yahoo.com/>





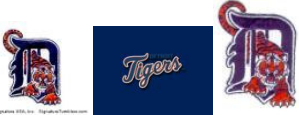

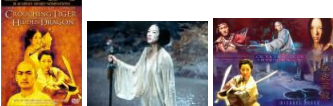



Clustering results of query "Tiger" (1,230,000)					
1.	Tiger Woods (61,300)		6.	Tiger Tank (11,100)	
2.	White Tiger (28,100)		7.	Tiger Club (9,650)	
3.	Detroit Tigers (17,100)		8.	Tiger Shark (9,290)	
4.	Crouching Tiger (13,500)		9.	Mac OS (4,430)	
5.	Bengal Tiger (11,400)		10.	Tiger Tool (865)	

Figure 11. Clustering results for query "tiger", the numbers within parentheses are the numbers of returned images.