Chinese Web users’ daily image needs and seeking behavior in a Q&A community

Kun HUANG1†, Xi NIU2, Shanshan WANG1 & Kaifei WANG1

1School of Government, Beijing Normal University, Beijing 100875, China
2School of Informatics and Computing, Indiana University at Indianapolis, Indianapolis, IN 46202, USA

Abstract

Purpose: In order to further the understanding of Chinese Web users’ image-seeking behavior, this study explores the kinds of images that Chinese Web users seek online and how they express their requests.

Design/methodology/approach: We used five pairs of simulated keywords to collect 893 image-seeking questions from Baidu Zhidao. Then, we revised the subject category of questions to analyze popular image needs. In addition, we conducted content analysis and descriptive statistical analysis to identify image-seeking motivations and image features used in the requests in terms of the two theories of image feature classification and image use.

Findings: Among the 893 questions, the image searches for entertainment accounted for 47.59%, more than the searches for professional knowledge (37.40%) and personal daily activities (15.01%). With regard to motivation, over 60% of the questions were identified as used for learning, which is well over the proportion of questions used for illustrating. Thus, these questions requested images as sources of data rather than sources of objects. Non-visual features (47.58%) were used most frequently in question descriptions, slightly higher than semantic features (45.96%). Users who lacked domain knowledge tended to use general words rather than specific words to describe their requests. However, not many users used syntactic features when seeking images. Nevertheless, most of the users had a fairly clear idea about what the target image should look like.

Research limitations: We studied only one question and answer (Q&A) community using five pairs of simulated keywords.

Practical implications: The findings should be helpful in strengthening the functionality of Q&A systems, promoting the theories of image feature classification, and shedding light on information literacy training.

* This work is supported by Humanities and Social Science Fund from the Chinese Ministry of Education (Grant No.: 11YJC870010). The authors would like to thank Ms. Ya’nan BAI, Ms. Lulu ZHANG, Mr. Lingchao SONG, and Ms. Tingyu QIN for their assistance in collecting and coding the pilot data. Our sincere thanks also go to the anonymous reviewers and editors for their valuable comments and constructive suggestions.

† Corresponding author: Kun Huang (E-mail: huangkun@bnu.edu.cn).
Originality/value: This study is one of the first research efforts that discusses Chinese Web users’ daily image searches and querying behavior in natural language in a Q&A community, which should help to further the understanding of the principles of image-seeking behavior among Chinese Web users.

Keywords Image needs; Image seeking; Baidu Zhidao; Q&A communities; Image users

1 Introduction

With the widespread use of digital images in people’s everyday lives, the concept of image has become a significant part of people’s daily information needs[1]. Information systems designers have devoted a substantial amount of effort to solve technical problems, such as developing algorithms to extract meaningful visual features, finding ways to recognize semantic features, and matching image features and users’ queries to reach high levels of relevance and efficiency[2,3]. Over the past thirty years, research into image users has attracted more and more attention, focusing on the images that users request and the ways that users seek images and use and keep those images[4,5]. Before the year 2000, most image user studies focused on domain-specific collections and professional users, such as experts, workers, or users in specific domains[6,7]. However, after 2000, more studies have focused on non-professional users and ordinary daily needs. All of these studies have allowed researchers to combine the user studies and technical research studies, which has benefited the enhancements of image systems and image services.

In China, little research has been undertaken to examine image users. A small number of studies has involved surveys or analysis of search logs to detect users’ image requests and image-seeking behavior. Different from those works, this study examines image users in question and answer (Q&A) communities, where users are able to express their requests in everyday, non-technical language. The questions in Q&A communities tend to be more narrative than general keyword search logs. Therefore, they are an ideal data source to analyze users’ information requests and ways of describing those requests. However, although Q&A communities have been used to discuss such areas as medical information requests[8] and to explore the relationship between information searches and relevance evaluation criteria[9], they have not been used in the study of image information searches among the Chinese populace.

In this study, we collected image-seeking requests from a Chinese Q&A community to explore three research questions:

1) What kinds of images do users want in their daily lives?
2) For what reasons or purposes do they want to have these pictures or images? 
And 
3) How do they express or describe their image-seeking needs?

2 Related research on daily image needs

Savolainen defined everyday life information seeking as non-work-related information seeking\[^{10}\]. Yoon\[^{11}\] and Huang & Kelly\[^{12}\] regarded both work-related and non-work-related image searches as components of daily information requests. Likewise, this study does not distinguish between work-related and non-work-related information requests.

Several researchers have examined image queries found in large-scale search engines logs. For example, Goodrum & Spink\[^{13}\] analyzed 1,025,908 queries from Excite according to the unit of queries, terms, and search sessions. Their results revealed that on average one user would use 3.36 queries to conduct an image search, with 3.74 terms per query. Moreover, image queries with high occurrence were related to nude, sex, and girls, etc. Jansen et al.\[^{14}\] used Altavista’s search logs to conduct similar analysis. Later, Jansen\[^{15}\] also examined Excite’s logs. He employed three image feature classification theories, including classifications defined by Chen\[^{16}\], Enser and McGregor\[^{17,18}\], and Jörgensen\[^{19}\] to code the queries. From this work, Jansen\[^{15}\] discovered five new additional image attributes that were added to image feature classification theory. Similarly, Pu\[^{20,21}\] also used Enser and McGregor’s theory\[^{17}\] to code the top 1,000 queries in a Chinese image search engine, VisionNEXT. She proposed a general-purpose query analytic model for understanding image requests.

Several other researchers conducted survey studies. By surveying students at the University of North Texas, Ruiz & Chin\[^{22}\] found that the most common purposes of image searches were entertainment and professional. Yoon\[^{11}\] surveyed 58 students at a public American university and found that the image-seeking motivations were various, including preparing presentation materials, searching for computer wallpaper, shopping, etc. Huang & Kelly\[^{12}\] derived similar findings by collecting questionnaire responses from 519 Chinese undergraduate students. In addition, Zhu and his colleagues\[^{23}\] also conducted a survey study among ordinary Chinese citizens.

Other researchers collected data from Web Q&A communities. Cunningham et al.\[^{24}\] analyzed 404 image-seeking queries posted on Google Answers ‘Ask an Expert’. The study showed that bibliographic and content features were used most frequently. Chung & Yoon\[^{25}\] analyzed 192 requests posted on Yahoo! Q&A. They found that users looked for images mostly for the purpose of illustration.
In summary, previous research studies have explored the images that users requested, the motivations of image seeking, and the image features that users might use in queries. However, few Chinese researchers have used Q&A communities to study daily image searches. Therefore, this study examines a Chinese Q&A community to learn about Chinese Web users’ image requests and querying behavior.

3 Methodology

3.1 Data collection

We collected image-seeking questions from Baidu Zhidao, one of the most widely used Q&A websites in China. It is affiliated with Baidu, which is regarded as a counterpart to Google in the United States. The China Internet Network Information Center (CNNIC) has reported that 86.7% of the search engine users in China put Baidu as their first choice. Alexa Traffic Ranks also showed that Baidu.com ranked fifth for websites worldwide. Moreover, nearly one-fourth of the Baidu users went to Baidu Zhidao, ranking the second to the percentage of Baidu users who visited the Baidu homepage. All these data suggest that Baidu Zhidao could be a representative Q&A service in China.

Two basic strategies are typically used to collect data from Q&A communities. The first is to select the requests under specific categories that are related to images; for example, Cunningham et al. gathered data from the category of Arts in Google Answer. The second strategy is to use simulated image-searching keywords to ask questions; for example, Chung & Yoon used Boolean OR operator to find keywords such as ‘image’, ‘picture’, etc., to find image-seeking questions in Yahoo!Answer. We took the second approach for this study.

At the beginning of this research effort, we tried one pair of keywords ‘图(图)/图片(tupian)’ to search in Baidu Zhidao. Baidu Zhidao returned only 76 pages of search results at most for each query, with 10 questions on one page. We reviewed the first 50 pages and obtained only 91 questions that were related to image seeking. Due to the limitations of Baidu Zhidao’s function, we simulated another four pairs of keywords to obtain more questions. Finally, we reviewed 5,000 questions altogether and filtered 893 valid questions, as shown in Table 1. All 893 questions were submitted by 614 users with usernames and another 279 anonymous users.

http://zhidao.baidu.com
Table 1  Image-seeking questions collected within first 50 pages for each pairs of query

<table>
<thead>
<tr>
<th>Simulated queries</th>
<th>Valid questions</th>
<th>Reviewed questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>图(tu)/图片 (tupian, equivalence of image or picture)</td>
<td>91</td>
<td>1,000</td>
</tr>
<tr>
<td>要(yao)图 (tu)/要图 (yuotu, equivalence of want AND images/“want images”)</td>
<td>393</td>
<td>1,000</td>
</tr>
<tr>
<td>求(qiu)图(tu)/求图 (qiutu, equivalence of seek AND images/“seek images”)</td>
<td>237</td>
<td>1,000</td>
</tr>
<tr>
<td>找(zhao)图(tu)/找图 (zhaotu, equivalence of “look for” AND images/“look for images”)</td>
<td>132</td>
<td>1,000</td>
</tr>
<tr>
<td>有(you)图(tu)/有图 (youtu, equivalence of have AND images/“have images”)</td>
<td>40</td>
<td>1,000</td>
</tr>
<tr>
<td>Total</td>
<td>893</td>
<td>5,000</td>
</tr>
</tbody>
</table>

All these 893 questions were posted between 2005 and 2014, among which, over half were posted in the most recent two years (i.e., 2013 and 2014). Table 2 summarizes the yearly distribution of these questions.

Table 2  Distribution of questions' submission time

<table>
<thead>
<tr>
<th>Year</th>
<th>Frequency</th>
<th>Percentage (%)</th>
<th>Year</th>
<th>Frequency</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>328</td>
<td>36.73</td>
<td>2009</td>
<td>55</td>
<td>6.16</td>
</tr>
<tr>
<td>2013</td>
<td>148</td>
<td>16.57</td>
<td>2008</td>
<td>52</td>
<td>5.82</td>
</tr>
<tr>
<td>2012</td>
<td>91</td>
<td>10.19</td>
<td>2007</td>
<td>39</td>
<td>4.37</td>
</tr>
<tr>
<td>2011</td>
<td>98</td>
<td>10.97</td>
<td>2006</td>
<td>14</td>
<td>1.57</td>
</tr>
<tr>
<td>2010</td>
<td>66</td>
<td>7.39</td>
<td>2005</td>
<td>2</td>
<td>0.22</td>
</tr>
</tbody>
</table>

3.2 Data coding

3.2.1 Coding of subject categories of image searches

Theoretically speaking, it is practical to analyze the kinds of images that are sought based on the subject categories of the questions provided by the users. Because of different understandings with regard to the categories’ meanings, similar image-seeking questions may be classified into different categories. Moreover, the specificity of categories also would be different. Therefore, we made some revisions to reduce the potential for inconsistency.

- First, we set up three top categories, which included professional knowledge, daily activities, and entertainment.
- Second, we created subcategories for each top category depending on the questions. For instance, within the entertainment category, the subcategories
were classified according to image type. The subcategories for professional knowledge were created in terms of fields and disciplines. The subcategories for daily activities were created in terms of people’s basic needs, such as clothing, food, housing/accommodation, travel, health, etc.

- Third, we compressed 137 subcategories given by users to 51, and assigned new categories to each question. If disagreement existed among the three coders, we discussed together until an agreement was reached.

Table 3 presents three examples how the subject category has been revised in terms of each original questions in Baidu Zhidao. The average ratios of inter-coder agreement reached 0.97 for image uses and 0.92 for image features, which is acceptable reliability.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Original category in Baidu Zhidao</th>
<th>Revised category</th>
</tr>
</thead>
<tbody>
<tr>
<td>I need a subway line picture for Guangzhou. Simple is better. I want to use it as wallpaper.</td>
<td>Life &gt; Wallpaper</td>
<td>Entertainment</td>
</tr>
<tr>
<td>Pictures of military rank How does the military rank upgrade?</td>
<td>Computer/ network &gt; Military</td>
<td>Professional knowledge</td>
</tr>
<tr>
<td>Pictures of enlarged lymph gland on neck I want to know what is the matter with my lymph gland.</td>
<td>Internal medicine</td>
<td>Health care</td>
</tr>
</tbody>
</table>

### 3.2.2 Coding of image-seeking motivation

Normally, the motivation to seek an image has a close relationship with the purpose of the image. For this study, two different image use theories were used to code the motivations for the questions, the one is the theory of Conniss et al.[28], and the other is the framework developed by Fidel[29].

In the theoretical concept of Conniss et al.[28], there are seven types of image use, which include: 1) aesthetic value, 2) emotive/persuasive purpose, 3) illustration, 4) generation of idea, 5) information dissemination, 6) information processing, and 7) knowledge learning. We refer to this theory as the “7C theory” in our study.

As an image could be regarded either as a data or an object, Fidel[29] considered the use of image as a continuum between two poles: One is the data, and the other is the object. The data pole concerns the information such as medical slides and statistical graphics, while the object pole refers to actual images, such as the cover of magazines and pictures in a book. We called this theory as the “2P theory”. Table 4 lists some examples on how we coded the data with the “7C” and “2P” theories.
3.2.3 Coding image features

We employed the image feature classification theory summarized by Chung & Yoon[25] to analyze the use of image features in question descriptions. This comprehensive theory was summarized based on previous prominent theoretical and empirical work on image indexing and image-seeking behavior[19,30–32]. As shown in Table 5, we divided the image features into three categories: Non-visual, syntactic, and semantic.

4 Findings

4.1 What kind of images do users seek?

The findings indicate that nearly half of the questions (47.59%) were related to entertainment, followed by professional knowledge (37.40%) and daily activities (15.01%). Figure 1 illustrates the distribution of the top-ranked six subcategories within the three main categories.

Figure 1 shows that three mostly-requested images were comics and animation, mathematics, and photos of people. It is particularly interesting to note that images related to mathematics were requested frequently. When looking more closely into these questions, we found that many primary and middle school students took pictures of homework and posted them on Baidu Zhidao to seek help. They hoped that someone could show them the answers and solution steps with images.

Similarly, within the professional knowledge category, many questions were about basic concepts, correct methods, and procedures of problem solving. For the
Table 5 Examples of image classification according to Conniss et al. [28]

<table>
<thead>
<tr>
<th>Questions</th>
<th>Attribute component</th>
<th>Attribute</th>
<th>Top category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title: Images and detailed introduction of the Classic of Mountains and Rivers</td>
<td>Date, title, creator, rights</td>
<td>Bibliographical</td>
<td>Non-visual</td>
</tr>
<tr>
<td>File type: ...famous court painting and related introduction</td>
<td>Quality, size, file format/type</td>
<td>Physical</td>
<td></td>
</tr>
<tr>
<td>Similarity: I need a similar picture like this…</td>
<td>Associated information, comparison, Contextual similarity, external relation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global shape: I’d like to have a head portrait shaped like the character Z or several Zs.</td>
<td>Global color, global texture, global shape, global quality</td>
<td>Global distribution</td>
<td>Syntactic distribution</td>
</tr>
<tr>
<td>Local color: I need a picture with girls in black suits.</td>
<td>Local color, local texture, local shape</td>
<td>Local structure</td>
<td></td>
</tr>
<tr>
<td>Spatial relation of elements: I need a picture with gray font and a row of transparent lines passing through it.</td>
<td>Position, spatial relation of elements, visual relationships, focal point</td>
<td>Global composition</td>
<td></td>
</tr>
<tr>
<td>Activity: I need a fighting picture of the chief male actor in Guilty Crown.</td>
<td>Object, people, scene, time, place, activity</td>
<td>Generic</td>
<td>Semantic</td>
</tr>
<tr>
<td>People: I need a picture of Aria in the Scarlet Ammo.</td>
<td>Object, people, scene, time, place, activity</td>
<td>Specific</td>
<td></td>
</tr>
<tr>
<td>Concept: I need pictures that could symbolize the humanistic quality of Dongguan.</td>
<td>Object, people, scene, time, place, activity, concept, emotion</td>
<td>Abstract</td>
<td></td>
</tr>
</tbody>
</table>

daily activities category, image requests concerning health care were found most frequently. These questions always asked about positions of the internal organs and disease symptoms.

Afterwards, we further looked into the best answers to see how users were satisfied with the responses. Out of all 893 questions, over 80% of the questioners confirmed the best answer. Among the best answers, half provided the images directly, and less than one-fifth offered related links.

In addition, about 30% of the best answers provided the solution only in words. Thus, associated information for images, such as background information, ways to search, or what the images looked like, also seemed to be very helpful. For example, for the question “I want to know what gesture I should make if I win a basketball game. Please show me pictures,” three gestures introduced in the best answers were described in words.

4.2 For what purposes do Chinese Web users seek images for their daily lives?

In our study, every image-seeking question was coded both with Conniss et al. [28] and Fidel [29] image use theories to understand the reasons why users seek a particular
image. In fact, only 58.79% of the questions were found to describe their preferred reasons explicitly. Figure 2 indicates that learning and the data pole were the dominant uses of images.

Figure 2(a) indicates that the percentage of the first usage, learning, is much higher than that of the second highest, illustration (13.71%). This differs from some
previous findings in this regard. Chung & Yoon\cite{25} and McCay-Peet & Toms\cite{33} discovered that images were used for illustration more often than any other usages. Thus, images were more found to be used as an object source rather than a data source. However, it is not the inherent property of images that decides how an image ought to be used\cite{29}, since whether an image is used as a source of information or an object is normally determined by the users or their tasks. Furthermore, users visited websites with different motivations. Figure 3 depicts the cross-analysis between the two types of image uses in this study.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure3.png}
\caption{Cross-analysis of relationship between the two types of image uses.}
\end{figure}

As shown in Fig. 3, over three-fifths of the questions closer to the data pole were used for learning, which is the most popular use of images found in this study. When images were used for illustration and aesthetic value, they were much closer to the use of the object pole.

However, the frequency of questions related to illustration and aesthetic value was much less than the use of learning. This outcome led to the difference between our findings and those of previous studies. Among the questions related to learning, over 60\% of the questions asked for the way to solve a problem or the method used to make something. For example, one user wants to know how to make Chinese rice pudding (“I want to know the method to make traditional Chinese rice-pudding. I need the picture.”), another wants to inform more about the solar phenomenon (“How do the solar eclipse and lunar eclipse occur? Please show me images to explain that.”)

In our analysis, most of the users on Baidu Zhidao used images as a visual surrogate to present abstract or tacit knowledge, methods, or steps. In this way, tacit knowledge became explicit and vivid, which makes users much easier to understand. Moreover, users can hardly search tacit knowledge using general search engines, which is the reason why they seek help from Q&A communities.
4.3 How do users describe their image needs?

4.3.1 Overall characteristics

An image-seeking request usually contains both a title and content. If the title is sufficiently informative, the user may keep the content blank. For this reason, 219 (24.52%) out of the 893 questions had no content. With respect to the title, the average title length was 11.31 words (SD = 7.49), with the maximum length 56 words and the minimum 4 words. Three kinds of descriptions were identified in titles.

- First, over 50% of the titles were specific. These titles provided detailed and specific requirements, such as “I need images related to purebred golden retrievers”, “the graph of the relationship between the location of the naevus and people’s destiny”, and so on.
- Second, 28.78% of the titles were quite general, stating only, for example, “need picture” or “want picture”, without presenting any further useful clues.
- The third kind of titles was somewhere between the other two mentioned above, accounting for 18.03%. Although these titles did not specify the subject or type of image, they mentioned “big picture” and “big and high definition”.

In addition, three kinds of content description patterns were identified in rest 674 questions with content.

- First, 43.77% (295 of 674 questions) described requirements in pure text.
- Second, 34.71% (234 of 674 questions) merely provided example images without any text. These questions always stated the request in titles.
- Third, different from the above two, 21.51% (145 of 674 questions) provided image examples, related URLs, and/or narrative words.

4.3.2 Use of image features

According to image feature classification theory of Conniss et al.[28], we coded the image features that occurred in the questions. Except date, rights, external relation, and global quality (Refer to Table 2), all of these features occurred 2,028 times. Each question contained at least one feature and 15 at most, with an average number of features per question of 2.27 (SD = 1.65).

As shown in Fig. 4, non-visual features (Fig. 4(a)) occurred most frequently (47.58%), which was slightly higher than semantic features (45.96%, Fig. 4(c)) but significantly higher than the occurrences of syntactic features (4.46%, Fig. 4(b)).

As for the non-visual features, contextual information occurred 691 (34.07%) times, higher than the frequency of each semantic feature. Moreover, within the
category of contextual information, associated information was given for 599 times, accounting for 86.69% of contextual information.

In addition, 61.10% (366 of 599 times) associated information was offered with image examples. Users may use image examples to depict background information, for example, a picture of the floor plan was attached to the request “I just built a new house in the countryside. Could you please suggest me how to decorate my house?”

As for both the syntactic and semantic features, our findings indicate that ordinary users tend to describe their requests in simple ways, this has also been proved by Chung & Yoon[25], they found that image users used few syntactic features in their questions. It appears that it is easier for users to describe what should be contained in the image or what meaning the image symbolizes rather than to describe the specific visual features clearly. What’s more, users preferred to use generic features rather than specific features. It makes sense that more domain knowledge is needed if users hope to recognize the specific object or people in the images.

### 4.3.3 Relationship between image uses and image features

We used the 529 questions with explicit image usage either in pure text or only in images to conduct cross-analysis between image uses and image features.

#### (i) Relationship between seven kinds of image uses and image features

There are three distinctive features of image use, i.e., non-visual, syntactic, and semantic. Table 6 indicates that contextual features and generic features were used in all seven kinds of image uses. This finding shows that users preferred to provide more background information to make others better understand their requests and that they usually were aware of what should be in the images.

Besides, physical, specific, and abstract features occurred in six popular image uses, in which physical features were used most frequently in questions of aesthetic value use. Such questions often mentioned size, quality, and resolution.

Specific features were used most frequently in questions for information processing, whose occurrence was the highest of specific features, accounting for
55.88%; whereas, for the image use of information processing, no syntactic features were found. This means that image requests for the purpose of information processing were usually unambiguous. The more specific the given information, the more easily for users to obtain help from others.

Moreover, abstract features were used more frequently in emotive/persuasive purpose, but absolutely not in information processing use. An example of an emotive/persuasive question was “I’m working on a commemorative album for my friend. I need some retrospective and sentimental pictures.”

(ii) **Relationship between two poles of image uses and features.** Table 7 shows that the closer the questions are to the object pole, the more types of image features the questions may use, with more syntactic features and fewer specific features. By contrast, the questions that are closer to the data pole tend to use fewer visual features and depend more on contextual features. Moreover, specific features were used more for the questions that were closer to the data pole rather than to the object pole.

### 5 Discussion

#### 5.1 Daily image needs

Daily image needs that were discovered from the questions on Baidu Zhidao were diverse and covered 51 subject categories. The dominant image requests were associated with entertainment. Users sought comics and animation most frequently, which suggests that users seek images mostly for entertainment purposes[15].
Another major finding was that the proportion of professional needs was also high, just slightly less than for entertainment. Users sought images to solve professional problems, such as in areas of mathematics, biology, painting, etc. This result might be attributed to the function of Q&A communities that enables users to share ideas and help each other.

As reported by CNNIC, knowledge and experience accounted for over 40% of the total information that users were seeking, ranking the second after news[26]. These findings indicate knowledge and experience searching are very popular in people’s daily lives.

5.2 Image-seeking motivations

As to motivations, learning was found to be the most popular purpose for seeking images, even higher than illustration use. This result explains the large need for professional images.

Moreover, the knowledge gained from images could be further divided into ‘What’ knowledge and ‘How’ knowledge. ‘What’ knowledge is the kind of knowledge that consists of raw data and facts, whereas ‘How’ knowledge is manifested in information related to methods and work procedures. Furthermore, images used as ‘How’ knowledge were more than those as ‘What’ knowledge, although it is more difficult for general search engines to find ‘How’ knowledge than ‘What’ knowledge, which depends more on human beings’ intelligence. Thus, users moved to seeking help from Q&A communities. All these findings demonstrate the importance of interpreting abstract knowledge in visual form, which also explains the reason why visual literacy is highly emphasized in education[34].

### Table 7 Cross-analysis of relationship between image features and 2P theory

<table>
<thead>
<tr>
<th></th>
<th>Object pole (%)</th>
<th>In-between (%)</th>
<th>Data pole (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-visual</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biblio-</td>
<td>0.77</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>graphical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td>17.48</td>
<td>9.87</td>
<td>0.54</td>
</tr>
<tr>
<td>Contextual</td>
<td>21.59</td>
<td>21.05</td>
<td>47.76</td>
</tr>
<tr>
<td><strong>Syntaxic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global</td>
<td>2.83</td>
<td>1.32</td>
<td>0.18</td>
</tr>
<tr>
<td>distribution</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local structure</td>
<td>2.31</td>
<td>1.32</td>
<td>1.07</td>
</tr>
<tr>
<td>Global composition</td>
<td>5.14</td>
<td>3.29</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>Semantic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generic</td>
<td>32.39</td>
<td>36.18</td>
<td>25.76</td>
</tr>
<tr>
<td>Specific</td>
<td>7.71</td>
<td>15.13</td>
<td>20.93</td>
</tr>
<tr>
<td>Abstract</td>
<td>9.77</td>
<td>11.84</td>
<td>3.58</td>
</tr>
</tbody>
</table>
5.3 Descriptions of image needs

First, most users tended to express their needs clearly. Nearly 80% of the questions had meaningful titles that included requirements for image types or image subjects. Thus, the average lengths of the titles were much longer than the average search terms found in search engines[13]. Furthermore, users also preferred to provide informative background in question statements.

Second, users preferred to use both text and examples to describe their image needs. Once the examples were given, they were always used as background information or examples to find more similar, bigger images or the original versions with higher resolution. This finding shows that users who are seeking help from Q&A communities always have a fairly clear need.

Third, users preferred to use non-visual and semantic features to describe their needs. Among the non-visual features, contextual features occurred most frequently, especially in terms of associated information, and most associated information was provided by inserting an image example.

Semantic features also were used frequently, especially generic features. This supports previous studies’ findings of Choi & Rasmussen[4] and Ornager[6] that non-professional users prefer to use general words for searching[4,6]. Due to the lack of domain knowledge, users found it difficult to articulate specific names, titles, or detailed information about the images when they stated their needs.

Moreover, specific features were found to be important to information processing, since the purpose of these questions were specific. Furthermore, although the occurrence of abstract features was not very high, abstract features occurred in almost each image use. Thus, the importance of extracting and indexing images with abstract attributes should not be ignored.

6 Implications

6.1 Functions to strengthen Q&A systems

First, this study’s findings could be used to improve indexing and searching functions by strengthening the indexing of image features with high occurrences and offering multiple handy choices to navigate images[35]. Furthermore, because image features might depend on different image uses, the Q&A system could borrow from research into searching intent forecasting[36] to determine strategies for image indexing.

Second, this study’s findings could serve to strengthen the function of question automatic answering. Over half of the users gave examples or related URLs to seek images. Once it recognizes the image examples, the Q&A system should be able to search automatically and recommend the results first. This function has already been supported by most search engines.
Third, the Q&A system should provide more options and suggestions for users to express their questions more efficiently. When users submit questions, the Q&A system could encourage them to choose a desired answer type or the purpose of their search. The more attributes the questions have, the more ways the system should have to search and sort them, which would be helpful in developing new automatic answering algorithms\[37,38\]. All these measures would be effective in improving users’ experience.

6.2 Enhancements to existing image feature classification theories

A better understanding of the principles of image seeking will provide opportunities to enhance existing theories. We found that supplementary work should be undertaken to enhance image feature classification theory rather than image use theory. Although Chung & Yoon\[25\] have given a relatively comprehensive summary of image features, we suggest that the abstract features under the semantic category could be enhanced.

Current abstract concepts and emotions refer mainly to the meaning that the whole image symbolizes or the emotion it arouses. The number of objects (three sheep), the mood of a person (a sad woman), or other modifiers were not included in the theory. Thus, it is necessary to build a new category system for adjective and adverb modifiers.

For instance, Pu\[20\] divided modifiers into annotation, affect, and perception. Huang\[39\] divided the affective features into physics, style, emotion, and aesthetic. In this way, more abstract features could be recognized for future indexing and search use. To extract these features, user-generated content, such as tags and text surrounding the images, could be used to improve the technical feasibility.

6.3 Integration of information literacy training and image-seeking procedures

Our findings show that most users have a fairly clear need when they seek help from Q&A communities. Even when they have image examples, they were not aware that they could use a search engine to conduct example-based searching by themselves. Therefore, the goal of Q&A systems should not be a sharing platform. They could provide information literacy training at the same time. For instance, after the users submit their questions, Q&A systems could inform them about alternative search strategies, especially when no answer is returned, in order to reduce users’ frustrations.

7 Conclusions

By collecting image-seeking questions posted on Baidu Zhidao, we discovered the diversity of the needs and rich characteristics of descriptions among Chinese
Web users. On the one hand, our study shows that the theories of image uses and features classifications developed for western populations could help explain Chinese users’ image needs to a certain degree. On the other hand, the findings were useful for image service providers to develop new functions and to strengthen and improve service for Chinese users.

The shortcoming of this study is that it used only one Q&A website. Although Baidu Zhidao is used widely, if the findings were generalized, the sample sizes of both the Q&A website and questions would need to be enlarged.

Specifically, we found that Baidu Zhidao users tend to be very young. The questions related to learning use were posted mostly by primary and middle school students, yet only one-fifth of the citizens in China are aged between 10 and 19, and over half are aged between 20 and 39[26]. In addition, more theories could be applied to code the image-seeking questions, which would allow the principles of image needs representation to be more informative.

Author contributions

K. Huang (huangkun@bnu.edu.cn, corresponding author) was in charge of the overall research design, coded the questions, and drafted the first draft of the manuscript. X. Niu (xiniu@iupui.edu) revised the manuscript and helped with the data analysis and discussion of the findings. S.S. Wang (perfect_1992@126.com) was in charge of the data collection and image feature coding. K. F. Wang (942040581@qq.com) joined the discussion and did coding work.

References


Chinese Web users’ daily image needs and seeking behavior in a Q&A community

Kun HUANG et al.

Research Paper

[References]


Submission Guidelines

♦ Aims
Chinese Journal of Library and Information Science (CJLIS), being sponsored by the Chinese Academy of Sciences (CAS) and published quarterly by the National Science Library of CAS, is a scholarly journal in the field of library and information science (LIS). Its aim is to provide an international communication link between researchers, educators, administrators, and information professionals.

With the publication of the research results both from China and from other foreign countries, the Journal CJLIS strives to achieve an equilibrium between theory and practice. With its goal to provide an open forum for Chinese and international scholars in this field to exchange their research results, CJLIS also offers new possibilities in the advancement of Chinese library operations. The CJLIS tries to establish a platform for LIS students, researchers and staff all over the world to engage in intellectual dialog and also to improve library services so as to promote even more quickened and substantial development of LIS in China.

♦ Scope
Striving toward academic excellence, innovation, and practicality, the CJLIS mainly includes research papers both on the theoretical aspects as well as on the practical fronts in all aspects of the field. More specifically, it includes but not limited to informatics, library management, information technology application, knowledge organization system, knowledge management, archives, permanent preservation of library resources, LIS education, and so on.

♦ Refereeing Process
Articles and papers covering the topics or themes mentioned above will be refereed through a double-blind peer review process.

♦ Editorial Advisory Board
The Editorial Board is composed of the nationally and internationally well-known scholars and researchers in the LIS field and the high quality of this Journal is thus reasonably assured.

♦ Manuscripts Categories
As the first English-language academic journal on LIS published in Mainland China, the CJLIS will take a proactive attitude to trace and report the prevailing hot issues in the field around the globe as well as the more serious scholarly communications. As such, the submitted manuscripts are classified into constant categories and unfixed categories. In the former category, research papers, library practice and progress reports are the essential components. In the latter, book reviews, biographical sketches, anecdotes, reminiscence of prominent librarians and brief communications will appear occasionally.

Research papers represent original research work or a comprehensive and in-depth analysis of a topic. More than 3,000 words are considered as a proper length for such manuscripts, with a structured abstract ca. 200 words.

Library practice covers the latest development and application in any segment of library field work and information service. The length of the manuscript is preferred to be more than 3,000 words, with a structured abstract ca. 200 words.

Progress reports reflect the projects result or research progress on the key topics of the library and information science. Submissions of articles to this section are expected to be comprehensive and analytical, which may deepen the understanding of the discussed issue and stimulate further researches on the topics, or give a new perspective on future technological applications. The manuscript length should be within 5,000 words, with a structured abstract ca. 200 words.

♦ Manuscripts Requirements
All papers can be submitted either in English or in Chinese (or both) with a double-line space. For the assurance that all the materials of the to-be-submitted are included, please check the following:

Title. Please give a brief biographical introduction to all contributing authors and their research background on a separated paper. For a better organization of the paper, please use the headings and subheadings.

Authors and affiliations. Please do not forget to write down the mailing address of each and every article contributors.

References. Be sure all the references used should be cited properly in both in-text and in bibliography. Particular attention should be paid to the proceedings. Do not forget adding the name(s) of editors of the compilation, as well as the name of the publishers. For the detailed information, please request a copy of Reference Citation Format.

♦ Copyright
All submitted papers normally should not have been previously published nor be currently under consideration for publication elsewhere. For all the materials translated or obtained from other published resources, they should be properly acknowledged. All copyright problems should be cleared without any legal entanglements prior to the publication.

♦ Notes for Intending Submissions
A guide for authors and other relevant information, including submitting papers online, is available at the website of the Editorial Office of the CJLIS (http://www.chinalibraries.net). For any questions, you can e-mail the Office or directly to:

Prof. ZHANG Xiaolin
Editor-in-Chief of CJLIS
The CJLIS Editorial Office
National Science Library, Chinese Academy of Sciences
No.33 Beishihan Xiu, Zhongguancun, Haidian District, Beijing 100190, P.R. China
Tel: 86-10-82624454 or 86-10-82626611 ext. 6628
Fax: 86-10-82624454
E-mail: chinalibraries@mail.las.ac.cn
Website: http://www.chinalibraries.net

♦ Subscription
For single-copy subscription in China: RMB ¥ 200/Issue. For subscription outside of China, US $ 199/Volume yearly (including air shipping)
Chinese Journal of Library and Information Science (Quarterly)

Contents

- Research Papers

1 Chinese Web users’ daily image needs and seeking behavior in a Q&A community
Kun HUANG, Xi NIU, Shanshan WANG & Kuifei WANG

21 Collaborative information seeking: An investigation of community members’ behavior
Dan WU & Xue XIANG

37 Enhancement of students’ learning with mobile devices: Taking librarian-faculty collaboration in SJTU as an example
Qinling HUANG, Xie GAO, Wei PAN & Lin GU

53 Balance between education and game-play — A case study of design and optimization of the mini-game book classification
Jianhua WU, Yanan CHEN & Xiangtao MA

66 Factors influencing online information acquisition: The case of Chinese college students
Yihan ZHANG, Qinjian YUAN & Jabeen MISBAH

83 Author Index (2008–2014)

107 Reference Citation Format

108 Subscription

ISSN 1674 – 3393
CN 11-5670/G2