An Interactive Image Retrieval Approach to Searching for Images on Social Media

Manali Gaikwad
Department of Computer Science
University of Regina
Regina, SK, Canada
gaikwadm@uregina.ca

Orland Hoeber
Department of Computer Science
University of Regina
Regina, SK, Canada
orland.hoeber@uregina.ca

ABSTRACT

Searching for images posted within social media services such as Twitter relies on matching textual queries to the contents of the posts that include the images. Unfortunately, social media posts may not always provide accurate or meaningful descriptions of the contents of the embedded images, making searching for images a challenging task. In this research, we augment the textual contents of the posts with new information extracted from the images using image processing and deep learning methods, and provide a visual interface to enable interactive image retrieval. A user study was conducted with 28 participants to collect evidence on how our approach was used in relation to Vakkari's three-stage model of information seeking. We also analyzed participants' perceptions of usefulness, ease of use, and satisfaction in comparison to a common grid-based image search interface. The results from this study highlight the value of providing visual and interactive features to enable searchers to discover images from social media sources.

KEYWORDS

Image search, interactive image retrieval, information seeking, search visualization

ACM Reference Format:
Manali Gaikwad and Orland Hoeber. 2019. An Interactive Image Retrieval Approach to Searching for Images on Social Media. In *2019 Conference on Human Information Interaction and Retrieval (CHIIR '19), March 10–14, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3295750.3298930

1 INTRODUCTION

Over the last decade, we have experienced a rapid growth of social media use due to the development and ongoing refinement of multimedia, networking, and mobile technologies. This technology has changed the way people collect, share, search for, and learn from the information available to them. Social media services such as Twitter have emerged as powerful tools for not only sharing personal opinions and staying informed about friends and family, but also as a source for the discovery of information [12].

CHIIR '19, March 10–14, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-6025-8/19/03…$15.00
https://doi.org/10.1145/3295750.3298930

While the textual aspects of social media continue to be important, embedded media (e.g., images) are becoming increasingly popular. The methods for searching within the image content are not well supported within social media services themselves, nor within other interactive search and exploration tools [7]. Most interfaces that do support image search (e.g., the photos tab of the Twitter search interface) provide a grid-based representation and an implied relevance ranking, supporting visual scanning and scrolling as the primary interactive mechanisms, with little support for interactive exploration and sensemaking.

Image search is an important special case of information seeking, with two specific aspects that make image search challenging: (1) it is common for the searcher to have an under-specified information need [16], and (2) searchers have difficulty textually describing what it is they are seeking [11]. As a result, the traditional targeted-search approach to search interface design (e.g., query box and search results list or grid) may not be sufficient to support the search tasks necessary to find relevant images. Our research approaches this problem from two directions: extracting useful information from the images to augment the social media posts; and developing a visual and interactive image retrieval interface that leverages this extracted information to organize the images and support exploratory search [23] and sensemaking [18] within the context of Vakkari's three-stage model of information seeking [20].

To evaluate our approach to supporting image search tasks, we have developed prototype software called ImgSEE. A controlled laboratory study was conducted to study how participants use ImgSEE in relation to the specific stages of information seeking specified in Vakkari's model. The study also had participants conduct similar search tasks using a common grid-based interface, allowing us to compare how they perceived the interfaces in terms of usability for the intended tasks. The results of this study provide evidence of the benefit of augmenting social media posts with supplemental information from the embedded images, and the value of using this information within a visual interface that supports interactive engagement with the search process.

The remainder of this paper is organized as follows. Section 2 provides an overview of the key literature that has informed this research. Section 3 presents the details of the design and implementation of ImgSEE. Section 4 explains the methodology and results of the user study. The paper concludes with a summary of the primary contributions of this research.

2 RELATED WORK

This work was informed by prior research in the domains of interactive information retrieval, information visualization, and image search. In this section, relevant work from these areas are presented and discussed in the context of interactive image retrieval from social media sources.

2.1 Interactive Information Retrieval

Information retrieval focuses on the process of obtaining relevant information objects to a user's information needs. The traditional approach has changed over time to support users in accomplishing the goal interactively, enabling them to engage in information seeking processes [2]. Information seeking is a problem-solving task that consists of recognizing and understanding the problem, creating a plan for the search, conducting the search, evaluating the results, and repeating the process as required. In order to develop an interactive search interface to support users' information seeking goals, it is important to understand the strategies people use while searching for information.

Vakkari's three-stage model of information seeking [20] is particularly relevant to image retrieval, given the underspecified nature of image search tasks. During the pre-focus stage, a searcher may undertake a series of broad searches in order to assess the breadth of images that are present in the collection. In the focus formulation stage, a searcher may refine the goal of the search activity, and seek relevant images. In the post-focus stage, a searcher may collect, organize, and make use of the found images. Exploratory search processes [23] may occur during the pre-focus and focus formulation stages, as the searcher moves from a broad exploration to a focused search. Sensemaking [18] may occur throughout this exploratory search process, and may also be present during the post-focus stage, as the searcher seeks to organize and manage the images found.

Information foraging theory [17] explains processes associated with people searching for information. An important part of this theory is the concept of information scent [5], which describes a mechanism by which people interpret the information available to them to decide where to direct their efforts. An interactive image retrieval interface may be designed to enable information scent, by providing visual cues that afford the searcher with information regarding the makeup of the image collection and the ability to predict the results of their interaction with the search interface.

2.2 Information Visualization

Information visualization deals with the presentation and communication of abstract data using graphical representations [22]. A visual representation can be processed more quickly compared to a similar amount of text, leveraging the human ability to perceive, interpret, and make sense of visual stimuli. While precision may be lost while visualizing information, what is gained is the ability to assess similarities and relationships among a large amount of data with minimal effort.

The visual encoding of information is a fundamental concern within information visualization. While there are a wide range of visual variables that can be manipulated to encode information (e.g., position, shape, size, brightness, hue, orientation, texture, motion [22]), of particular importance is how colour is used to convey information.

The Opponent Process Theory of Colour [9] explains how humans perceive colour stimuli. This theory suggests that humans can readily see differences in colour across three channels: red-green, yellow-blue, and luminance (black-white). This theory provides a logical organization of colour in a circle, with red and green opposite, yellow and blue opposite, the saturation as the radius from the centre, and the luminance (brightness) as an orthogonal dimension. This organization of colour features prominently in the HSB colour model [4].

Although images are inherently visual, their complexity makes them difficult process in the same way as well-designed information visualizations. With some careful design, it may be possible to leverage the elements of the human visual processing systems to convey meaningful information about images to a user.

2.3 Interactive Image Search

Although image search interfaces have been studied for many years, much of the focus has been on how to organize images in a meaningful order or structure. Common approaches are grid representations [14], hierarchical structures [10], and multi-layered clusters [19]. They often employ a combination of visual and semantic similarity when choosing how to group the images, and limit the interactivity to scrolling, pan & zoom, and the selection of images.

Others have discussed the importance of considering the intent of the image retrieval activity [11], and have identified unique characteristics of image retrieval beyond traditional search settings [1]. A noteworthy approach that is consistent with these works is CIDER [8], which uses automatic query expansion to promote diversification of the image search results, and then interactive visual exploration within the image space to allow searchers to focus on the aspect of the query relevant to their needs and intentions. Our work is inspired by this level of interactivity, but with a focus on enabling the searchers at each stage of Vakkari's three-stage model of information seeking.

3 IMGSEE

In order to address the challenges of searching for images within social media data, we have developed ImgSEE (Image SEarch and Exploration) with a goal of enabling interactivity during the image retrieval process. Necessary first steps are to collect the image data and augment it with supplemental information, which are described in the following two subsections. The ImgSEE interface is described in detail in the subsequent subsection. The final subsection provides a brief overview of an image search scenario, and a link to a video illustrating how ImgSEE may be used to support image seeking tasks.

3.1 Data Collection

Using a set of user-specified harvest queries, data is collected in real-time using the Twitter Streaming API (https://developer.twitter.com). The Apache Spark API provides the core stream processing framework (https://spark.apache.org). The incoming tweets are filtered to remove retweets, duplicate images, and tweets that do not contain image content. The remaining tweets are stored in an instance of Elasticsearch (https://www.elastic.co/products/elasticsearch), serving as the source for the data augmentation stage.

3.2 Data Augmentation

Although the raw tweet data provides a wealth of meta-data about the tweet and author, little information is included about the image embedded within the tweet. In this step, the goal is to augment the collected data with new information extracted from the image. The first step is to download the image and store it for use in these data augmentation steps, as well as for use by the search interface.

A simple image processing algorithm that uses the HSB colour model as its basis provides the average hue, saturation, and brightness over the entire image. In addition, two classes of textual descriptors are extracted from the image contents. Optical character recognition using the Tesseract OCR API (https://github.com/tesseract-ocr/tesseract) identifies explicit textual information from within the images themselves. The semantic contents of the images are identified using DeepDetect (https://deepdetect.com), which uses a deep learning approach trained to classify images based on 1000 generic tags.

This new information is added to the Elasticsearch index for each image on a batch basis, in order to ensure that it does not conflict with the real-time processing of the Twitter data collection. The augmented data records serve as the basis for the interactive image search interface.

3.3 ImgSEE Interface

Vakkari's three-stage model of information seeking [20] (pre-focus, focus formulation, post-focus) served as a guide in the design of the ImgSEE interface. Exploration of the image collection during the pre-focus stage is enabled by textual querying within the contents of the tweets and the new information added in the data augmentation step. In the focus formulation stage, further exploration is enabled through pan & zoom within the dynamic image space, filtering the images based on associated hashtags, and filtering the images based on the timestamp of the tweet. A workspace is provide to support sensemaking and interactive grouping of selected images, which is of use during the post-focus stage.

The core interface was implemented using the Angular library (https://angularjs.org); the visualization features were implemented using D3 (https://d3js.org). A screenshot of the core interface is provided in Figure 1, showing a search for "adventure" among the tweets collected over seven days in September 2017. Here, the searcher has filtered the search results to be those that use the hashtags #summer and #travel. The searcher has also saved some images in the workspace, and organized these into three categories. Each of the core features of ImgSEE are explained in more detail in the subsections below.

3.3.1 Pre-focus: Querying. After a specific image dataset has been selected (based on the harvest queries used to generate the collection), the searcher can issue textual queries to isolate subsets of images from the underlying collection. By default, this search is performed within the textual contents of the tweets. The user can also specify searching within the supplemental information added to the tweet: the text embedded in the images or the semantic descriptions of the images. Such pre-focus querying can allow searchers to quickly experiment with various topics of interest, and visually inspect the search interface (image space, hashtags, and timeline, as shown in Figure 1) to assess the general make-up of the image collection.

3.3.2 Focus Formulation: Image Space Exploration. While many image search interfaces developed within the research community use the visual features of the images to cluster or organize the search results (e.g., self-organizing maps), the visual structures in which the images are organized are dynamically determined by the search results sets. As a result, each search organizes the images differently, and the searcher has to re-learn these structures with each new query.

Our goal was to provide a static organizational structure that would remain constant regardless of the query and the search results set. Doing so promotes learnability and memorability. Leveraging knowledge of the Opponent Process Theory of Colour and the HSB colour space, we use a disc to represent the image space. Hue is represented in the angle in the disc, preserving opposite colours (red vs. green; yellow vs. blue). A combination of saturation and brightness is represented in the radius from the centre. Mapping images to their corresponding locations within this disc results in visually similar images (in terms of colour) being close to one another, and visually different images being distant from one another.

This image space provides a visual overview of the search results, which can be useful for quickly assessing the makeup of the search results as part of the focus formulation stage. Interactive exploration within this image space using pan & zoom operations allow regions that contain many visually similar images to be expanded and inspected, and potentially relevant images to be identified and selected.

Of note is that there may not be any direct correlation between the colour-based distribution of the images and the semantic meaning of the images. The semantics may be explored through interactive querying (previous section) and hashtag filtering (next section). Here, the goal is to provide a logical organization of the images that is easy to interpret (visual similarity of overall colour), and which may be used to isolate relevant images from irrelevant ones (e.g., zooming into a region of adventure images where the colour is predominantly blue can allow the searcher to find images of water-based adventures even if the terms that describe this concept do not appear in the original posts that included the images).

3.3.3 Focus Formulation: Hashtag Filtering. Given the underspecified nature of image queries, it is useful to provide mechanisms to help searchers to understand the makeup of the search results, and interactive features to support query refinement. ImgSEE extracts and presents the top hashtags associated with the search result set, and uses a small-multiples approach to show where the associated images for each of these hashtags exist within the image space framework. Given the small size of these micro-visualizations, each image is represented as a dot, rather than in detail. Selecting any of these hashtags automatically filters the search results to only show those images associated with any of the selected hashtags.

The micro-visualizations support searchers in making visual comparisons of the locations of the images within the image space. This, coupled with the detailed view of the image space, provide information scent, enabling the searcher to observe the potential outcome of selecting a specific hashtag as part of their exploratory search process. In addition, this feature of ImgSEE supports the

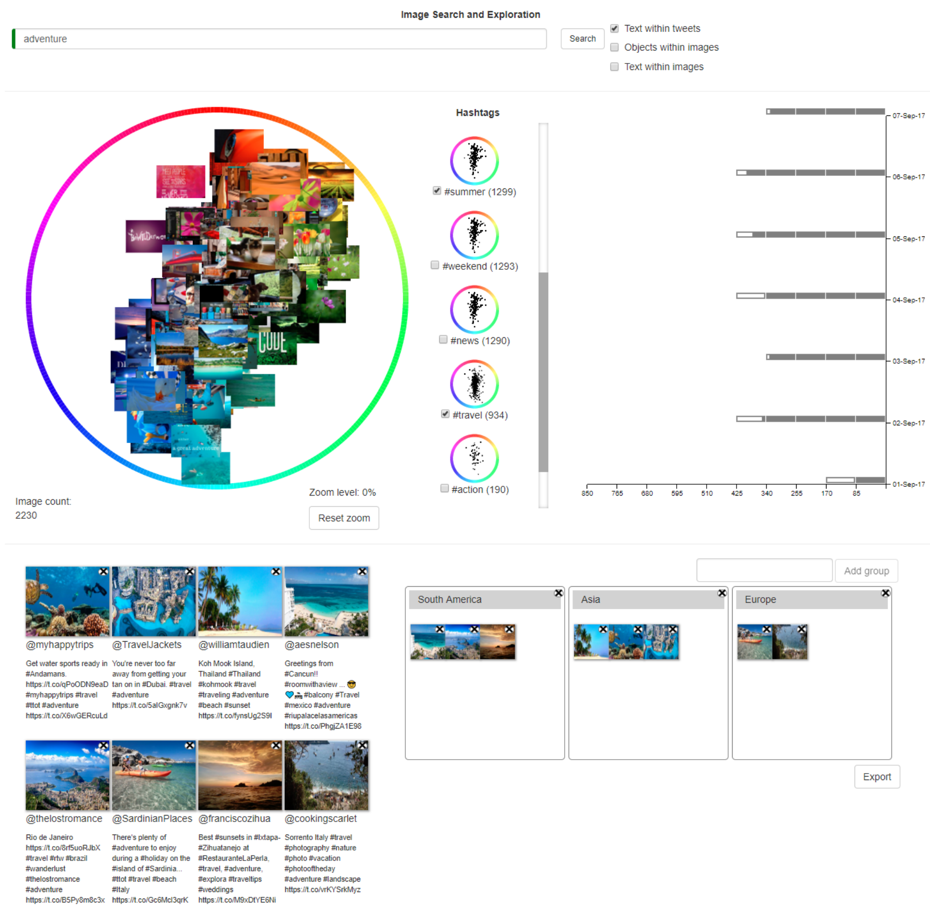


Figure 1: The core interface features of ImgSEE are the image space (left), top hashtags and micro-visualizations (middle), timeline (right), and workspace (bottom).

recognition of relevant information, rather than requiring that it be recalled, enabling searchers to filter the search results to more accurately reflect the information being sought.

3.3.4 Focus Formulation: Temporal Filtering. The timeline visualization provides an overview of temporal patterns of image-based posts, allowing searchers to quickly identify and compare these patterns. Since images are discrete data, a discrete timeline is used [3]. The size of the discrete bins is dynamically determined based on the temporal extent and size of the search results set. Partial bins are represented as partially-filled blocks. This binned approach makes it easy for the searcher to make comparisons and estimate the number of images on a given day.

The timeline supports interactive filtering of the search results, which is reflected in the other interface features (operating as multiple coordinated views). Selecting a specific day filters the search results to only include the images posted on that day. Similarly, selecting or unselecting individual bins further filters the search results. For example, a searcher may choose all images posted on September 5, 2017, and then further filter the timeline to only show those images taken earlier on that day.

Given the importance of temporal information in social media posts, providing searchers with an overview and interactive control over this element of their search activity is valuable. During the focus formulation stage, an overview of the temporal pattern can help the searcher make sense of the relationship between the images being posted and real-world events that have inspired people to post to social media.

3.3.5 Post-focus: Workspace Grouping. While many image search interfaces provide features that support exploratory search, and enable tasks that are useful for the pre-focus and focus formulation information seeking stages, few provide support for the post-focus stage. Given the dynamic nature of image search, it is beneficial to provide features that allow searchers to save, organize, and do something useful with the images discovered.

When a specific image is found to be relevant or useful, it can be selected from the image space, which will automatically copy it to the workspace. Images in the workspace are presented in a grid-based structure, with the most recent additions at the bottom of the grid. Searchers are able to dynamically create and name groups within the workspace, and images can be dragged into these groups. Since mutually exclusive groups would limit the utility of this feature, images may be dragged into multiple groups as needed. Clicking the export button automatically creates a directory structure based on the groups, adds the images to their respective directories, zips these directories into a single file, and initiates a download to the searcher's computer for further use as needed.

3.4 Image Search Scenario

In order to illustrate how ImgSEE can support image retrieval tasks, we have prepared a video [1] of an search scenario where the searcher is seeking images as inspiration for planning an adventure holiday. This scenario includes elements of what makes image search challenging: an under-specified information need and a difficulty in describing what is being sought. The video illustrates how the key interactive features of ImgSEE support the searcher in each stage of Vakkari's three-stage model of information seeking: conducting broad searches and assessing features of the image sets in the pre-focus stage; refining the search goal and seeking relevant images in the focus formulation stage; and collecting, organizing, and using the images in the post-focus stage.

4 EVALUATION

Although the features of ImgSEE were designed to support interactive image retrieval, conducting a controlled laboratory evaluation can provide insight regarding how the approach can be used in relation to Vakkari's three-stage model of information seeking, and the participants' perceptions of usefulness, ease of use, and satisfaction in comparison to a traditional grid-based approach. In order to support this evaluation, we implemented a common grid-based image search interface modelled on the photos tab provided by Twitter's search interface. We call this baseline comparison interface Grid.

4.1 Methodology

The specific methodology for this evaluation was designed to enable answering four research questions:

> **RQ1:** How does the search stage (pre-focus, focus formulation, post-focus) influence the use of the different features of ImgSEE?
> **RQ2:** In comparison to Grid, how useful did participants find ImgSEE?
> **RQ3:** In comparison to Grid, did the participants find ImgSEE easy to use?
> **RQ4:** In comparison to Grid, how satisfied were participants with using ImgSEE?

Two specific image search tasks were devised that each had a degree of under-specification and for which there may not be an obvious query that would lead to a precise set of relevant images. A starting query was used to collect and pre-process the images following the process outlined in Sections 3.1 and 3.2. For each task, the participants used ImgSEE as the image search interface. These tasks were interleaved with using the Grid image search interface on similar tasks. In order to address learning effects, the order of the tasks and the order of the search interface were varied using a Latin square. A fifth task was performed at the end only using ImgSEE on a topic provided by the participants in advance. Participants in the study were given brief instructions to guide them through Vakkari's three-stage model of information seeking. The specific details of these tasks are provided in Table 1.

The data for this study was collected using two primary instruments. To support answering RQ1, the ImgSEE interface was instrumented such that each interaction with the interface was logged. To support answering RQ2, RQ3, and RQ4, a post-study questionnaire based on the Technology Acceptance Model (TAM2) [21] was administered to collect perceptions of usefulness (three questions), ease of use (four questions), and satisfaction (one question). While some have suggested using usefulness as a key criteria when evaluating interactive information retrieval interfaces [6], including ease of use and satisfaction are important when the search interfaces under investigation are novel or somewhat complex to operate, as they provide a more comprehensive view of the participant's experiences during the study.

The study procedure followed the normal pattern for controlled laboratory evaluations. Each session began with an overview of the study process and verification of informed consent. A pre-study questionnaire was administered to measure demographic information and image search experience. A training video was shown for each of the two interfaces, illustrating how they can be used to support image seeking activities at the pre-focus, focus formulation, and post-focus stages. Two laptop computers were used in this study. One was used to manage the study process, administering the pre-study questionnaire, providing instructions to guide the participants through each of the tasks, and administer the post-study questionnaire; the other was pre-loaded with the ImgSEE interface and was used to conduct each of the search tasks.

Each study session took approximately 60 minutes. Participants (n=28) were recruited among our university's student body, and were entered into a draw for a 50% chance to win a $20 gift card.

4.2 Analysis

Because of differences in each of the image seeking tasks, the data on how ImgSEE was used were analyzed and are reported separately. The average number of interactions with each of the core ImgSEE features (RQ1) are presented in a graphical format using standard error of the mean as a measure of variability. These are organized by information seeking stage (pre-focus, focus formulation, post-focus). Using ANOVA, statistically significant differences in the use of the features between the three information seeking stages are noted in a table. In the case of statistically significant differences, post-hoc ANOVA of the two stages with the greatest number of interactions was conducted. If a statistically significant difference was found between these two, the one with the most frequent interaction is reported as the significant stage for the feature. If no statistically significant difference was found, both were reported as the significant stages for the feature. In either case, the mean number of interactions is also reported for these significant stages.

For the overall perceived usefulness (RQ2) and overall perceived ease of use (RQ3) questions, the data was collected using a set of five-point Likert scale measures (three and four questions, respectively). The average of the scores for each of these high-level constructs is provided in graphical format for each interface. Because the distances between these scale levels are not equal, applying statistical

[1] http://www.cs.uregina.ca/~hoeber/download/2018-chiir-imgsee-video.mp4

Introduction	Pre-Focus	Focus Formulation	Post-Focus	Tweets/ Images	Period
Adventure Task: You are in the process of making preliminary plans for an upcoming holiday with a group of friends. You have each decided to find sets of pictures of the places that you want to visit and activities that you want to undertake. You have decided that you want to focus on pictures that other people have posted to Twitter from their adventure vacations.	Using the provided application, start the search session using the keyword "adventure". Explore among the images to find those related to various locations and activities of interest to you.	Select one of the locations or activities you identified in the previous step and find more images relevant to this topic.	Among the images you found in the previous step, organize the images into some logical structure and then export them for sharing them with your friends.	4,883	7 days
Restaurant Task: Imagine you are a food blogger and you are looking to find images of food that other people are tweeting about right now. Your goal is to collect images that will give you ideas on the types of restaurants you might want to visit soon for your own food blog.	Using the provided application, start the search session using keyword "restaurant". Explore among the images to find what the latest trend is.	Select a specific type of food or restaurant that will be the focus of your next food blog and find more images related to this topic.	Among the images you found in the previous step, organize the images into some logical structure and then export them for use when you are writing your blog post.	3,781	15 days
Self-Selected Task: Your final task is to use ImgSEE to search for images related to one of the general topics you provided when you were recruited for this study. We have chosen the topic based on the one that had most images posted to Twitter in the last week. The topic along with the initial query is provided on the paper that the investigator has placed beside the computer.	Find images related to a small set of subtopics you find interesting.	Narrow this down to a single topic and find more images related to it.	Among the images you found in the previous step, organize the images into some logical structure and then export them to complete the task.	3,115 (average)	8 days (average)

Table 1: Tasks, phase instructions, and image collection properties.

measures such as ANOVA directly is not recommended. However, summing the set of Likert scale scores that measure the same underlying phenomenon can allow ANOVA to be used [13]. We follow this advice, using one-way ANOVA within-subjects analysis since all participants were exposed to both interfaces.

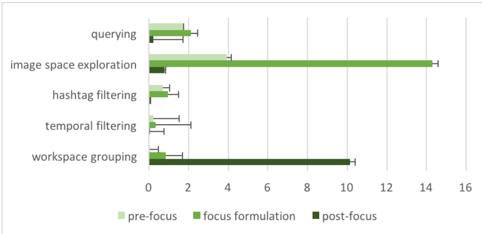
For the perceived satisfaction measure (RQ4), one question was asked and data was collected using a five-point Likert scale. As with the other data regarding the perceptions of the participants, the average of this score is presented in graphical format for each interface. Since just one question was asked about this high-level construct, a chi-squared test was used to analyze this measure.

4.3 Results

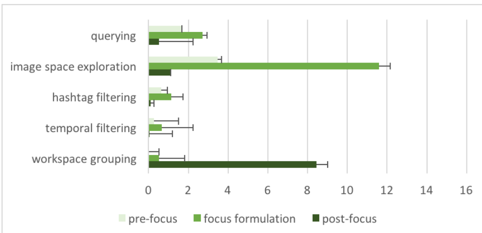
4.3.1 RQ1: Search Stage and Feature Use. For each of the three search tasks conducted using ImgSEE, the number of interactions with each of the core ImgSEE features were analyzed across each of the three information seeking stages. Figure 2 shows the mean number of such interactions. The general pattern across all tasks was for the querying to be used during the pre-focus and focus formulation stages; the image space exploration to be used extensively during focus formulation, but also somewhat during the pre-focus stage; the hashtag filtering to be used lightly during the pre-focus and focus formulation stages; and the workspace grouping to be used primarily during the post-focus stage. The temporal filtering features were used sporadically across the three stages, and with a high degree of variability among the participants (as shown with the relatively large standard error compared to the low interaction means).

Analysis of this data revealed statistically significant differences in which stage participants interacted with most of the ImgSEE features (see Table 2). The obvious features of Figure 2 were statistically significant (i.e., image space exploration is primarily used during the focus formulation stage and workspace grouping is primarily used during the post-focus stage). Less obvious are the statistically significant differences in which stages the querying and hashtag filtering features were used (both pre-focus and focus formulation stages, but not the post-focus stage). The temporal filtering feature was not used much, and there was no significant differences in how it was used across the stages.

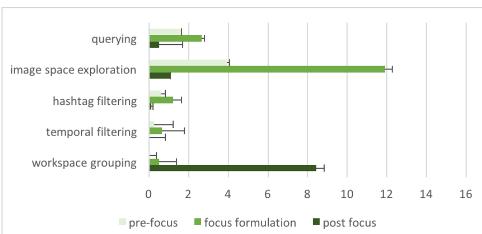
While there were very little differences between the tasks themselves, it is worth examining the participants' self-selected tasks in more detail. Given their personal interest and prior knowledge



(a) Adventure Task

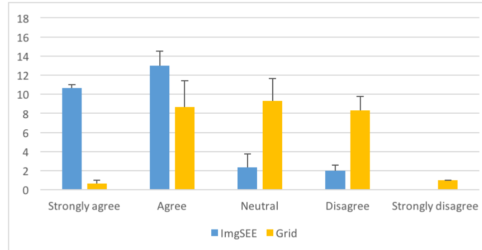


(b) Restaurant Task

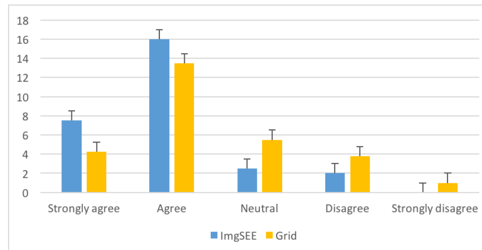


(c) Self-Selected Task

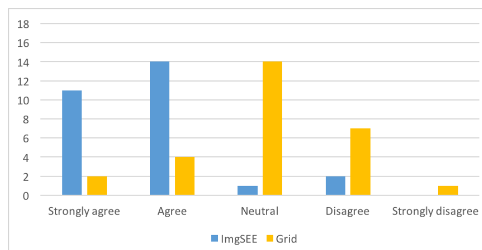
Figure 2: Frequency of interaction with each of the core ImgSEE features during each of the stages of information seeking. Error bars represent the standard error of the mean.



(a) Overall usefulness of the two interfaces.



(b) Overall ease of use of the two interfaces.



(c) Overall satisfaction with using the two interfaces.

Figure 3: Responses to questions regarding the overall usefulness, ease of use, and satisfaction with using ImgSEE and Grid. Error bars represent the standard error of the mean.

of this task, many participants moved quickly through the pre-focus stage, issuing one to two queries, using the hashtag filtering perhaps once, and interacting with the image space about four times. When in the focus formulation stage, the participants used each of these features more extensively, issuing up to three queries, using the hashtag filtering once, and interacting with the image space about twelve times. After entering the post-task phase, the workspace grouping became the primary interface element that was used, although some participants used the querying and image space exploration features as well. We view this self-selected task as providing the most realistic view of how ImgSEE might be used, since it was based on each participants' specific image seeking interests, and was performed last in the study (after the participants would have gained some experience in using the interface).

When designing ImgSEE, we used Vakkari's three-stage model of information seeking as a guide for developing visual and interactive features that support the searcher throughout the stages. More specifically, we hypothesized that the querying feature would primarily support pre-focus activities, that the image space exploration, hashtag filtering, and temporal filtering would support focus formulation activities, and that workspace grouping would support post-focus activities. While findings partially validate those hypotheses, it is clear that there is not a 1:1 correspondence between interface feature and information seeking stage. During the pre-focus stage, not only did participants issue queries, they also engaged with the exploratory search features to some degree (image space exploration and hashtag filtering). During the focus formulation stage, these same exploratory search features became the dominant mechanism for interaction, although there was also an increase in querying as a mechanism for focusing the search. The post-focus stage found participants primarily using the sensemaking features provided by the workspace grouping, although some querying and image space exploration was done, which indicates that some participants may have temporarily return to the focus formulation stage. This last point is noteworthy, since it provides evidence that Vakkari's model may be cyclical, rather than linear.

4.3.2 RQ2-4: Participant Perceptions of Usability. The validity of the results presented in the previous section can be enhanced by

Feature	Task	ANOVA Across Stages	Stage(s) of Significant Interaction
	adventure	$F(2, 81) = 6.27, \mathbf{p < 0.005}$	pre-focus (m=1.7) & focus formulation (m=2.1)
	restaurant	$F(2, 81) = 3.32, \mathbf{p < 0.05}$	pre-focus (m=1.7) & focus formulation (m=2.7)
querying	self-selected	$F(2, 81) = 5.02, \mathbf{p = 0.01}$	pre-focus (m=1.4) & focus formulation (m=3.2)
	adventure	$F(2, 81) = 55.45, \mathbf{p < 0.0001}$	focus formulation (m=14.1)
	restaurant	$F(2, 81) = 35.83, \mathbf{p < 0.0001}$	focus formulation (m=11.6)
image space exploration	self-selected	$F(2, 81) = 42.34, \mathbf{p < 0.0001}$	focus formulation (m=13.4)
	adventure	$F(2, 81) = 3.37, \mathbf{p < 0.05}$	pre-focus (m=0.7) & focus formulation (m=0.9)
	restaurant	$F(2, 81) = 3.58, \mathbf{p < 0.05}$	pre-focus (m=0.6) & focus formulation (m=1.1)
hashtag filtering	self-selected	$F(2, 81) = 4.15, \mathbf{p < 0.05}$	pre-focus (m=0.6) & focus formulation (m=1.0)
	adventure	$F(2, 81) = 0.95, p = 0.391$	
	restaurant	$F(2, 81) = 1.86, p = 0.162$	
temporal filtering	self-selected	$F(2, 81) = 1.74, p = 0.182$	
	adventure	$F(2, 81) = 75.45, \mathbf{p < 0.0001}$	post-focus (m=9.9)
	restaurant	$F(2, 81) = 46.46, \mathbf{p < 0.0001}$	post-focus (m=8.4)
workspace grouping	self-selected	$F(2, 81) = 76.00, \mathbf{p < 0.0001}$	post-focus (m=9.3)

Table 2: Statistical analysis (ANOVA) of the frequency of feature use across the three stages of information seeking. Statistically significant results are in bold, and indicate that feature was used more frequently in one or more stages than the others. For the stages with significant interaction using a given feature, the mean number of interactions are reported.

studying how the participants perceived the usability of ImgSEE in comparison to the more common grid-based approach to image retrieval interfaces. If the participants found ImgSEE useful, easy to use, and satisfying, this not only speaks to the quality of the implementation of the approach, but also enhances the trustworthiness of the findings on how the approach enabled the participants to undertake interactive image retrieval processes.

Overall, participants in this study reported more positive perceptions of the usefulness (RQ2), ease of use (RQ3), and satisfaction (RQ4) after using ImgSEE in comparison to Grid. The average number of responses at each level of the Likert scales for each of these high-level constructs are illustrated in Figure 3.

The participants' degree of agreement to statements regarding the usefulness of the Grid interface was evenly distributed between agree, neutral, and disagree, whereas for ImgSEE, the participants strongly agreed or agreed with these statements. For ease of use, there was a similar pattern of agreeing with the statements for both interfaces, but with a positive bias for ImgSEE. Participants were primarily neutral in terms of satisfaction with Grid, whereas most strongly agreed or agreed with the statement about satisfaction when considering their use of ImgSEE.

Analyses of these data revealed statistically significant differences across all three measures. Single factor ANOVA of the usefulness and ease of use data resulted in $F(1, 54) = 28.94, \mathbf{p < 0.0001}$ and $F(1, 54) = 5.78, \mathbf{p < 0.05}$, respectively. A chi-squared test of independence of the data for the satisfaction question resulted in $X^2(4, 56) = 26.83, \mathbf{p < 0.0001}$. In all cases, participants reported more positive responses regarding ImgSEE than Grid.

5 CONCLUSIONS

To the best of our knowledge, this is the first study that has used Vakkari's three-stage model of information seeking in the context of interactive image retrieval. We employed it as both a design guide, and as a mechanism for controlling the laboratory-based evaluation. This model is synergistic with exploratory search and sensemaking processes, which are especially useful for image search tasks where the information need may be under-specified and the searcher may be challenged in describing what they are seeking.

This paper provides examples of how information visualization techniques can enhance interactive information retrieval processes under the special constraints of image search. The study showed that the participants were able to follow the three stages of information seeking and make use of most of the features provide to support these stages. Of note is that the most frequently used features were those that provided the images themselves as the primary interaction elements (image space exploration and workspace grouping). As a result, for future image search interfaces, we suggest ensuring that the images themselves remain visible, rather than being abstracted away as in the hashtag filtering and temporal filtering features.

A final observation from this study was that some participants seemed to return to the focus formulation stage after entering the post-focus stage. For complex search tasks that have under-specified information needs like image search tasks, the search process may be more cyclical than linear. For future research, studying image search in the context of non-linear information seeking models such as Meho & Tibbo's model of digital library search [15] may provide further insight into how interactive information retrieval approaches can be designed to enhance image search tasks.

ACKNOWLEDGMENTS

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), through the Discovery Grant held by the second author (RGPIN-2017-06446).

REFERENCES

[1] Paul André, Edward Cutrell, Desney S. Tan, and Greg Smith. 2009. Designing Novel Image Search Interfaces by Understanding Unique Characteristics and Usage. In *Proceedings of the IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT)*. Springer, Berlin, Heidelberg, 340–353. https://doi.org/10.1007/978-3-642-03658-3_40

[2] Nicholas J. Belkin. 2016. People, interacting with information. *ACM SIGIR Forum* 49, 2 (2016), 13–27. https://doi.org/10.1145/2888422.2888424

[3] Matthew Brehmer, Bongshin Lee, Benjamin Bach, Nathalie Henry Riche, and Tamara Munzner. 2017. Timelines revisited: A design space and considerations for expressive storytelling. *IEEE Transactions on Visualization and Computer Graphics* 29, 9 (2017), 2151–2164. https://doi.org/10.1109/TVCG.2016.2614803

[4] Nilgün Camgöz, Cengiz Yener, and Dilek Güvenç. 2004. Effects of hue, saturation, and brightness: Part 2: Attention. *Color Research & Application* 29, 1 (2004), 20–28. https://doi.org/10.1002/col.10214

[5] Ed H. Chi, Peter Pirolli, Kim Chen, and James Pitkow. 2001. Using information scent to model user information needs and actions and the web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, USA, 490–497. https://doi.org/10.1145/365024.365325

[6] M. Cole, J. Liu, N. Belkin, R. Bierig, J. Gwizdka, C. Liu, J. Zhang, and X. Zhang. 2009. Usefulness as the Criterion for Evaluation of Interactive Information Retrieval. In *Proceedings of the Workshop on Human-Computer Interaction and Information Retrieval*. 1–4.

[7] Orland Hoeber, Larena Hoeber, Maha El Meseery, Kenneth Odoh, and Radhika Gopi. 2016. Visual Twitter analytics (Vista): Temporally changing sentiment and the discovery of emergent themes within sport event tweets. *Online Information Review* 40, 1 (2016), 25–41. https://doi.org/10.1108/OIR-02-2015-0067

[8] Enamul Hoque, Orland Hoeber, and Minglun Gong. 2013. CIDER: Concept-based image diversification, exploration, and retrieval. *Information Processing & Management* 49, 5 (2013), 1122–1138. https://doi.org/10.1016/j.ipm.2012.12.001

[9] Leo M. Hurvich and Dorothea Jameson. 1957. An opponent-process theory of color vision. *Psychological Review* 64, 6, Pt.1 (1957), 384–404. https://doi.org/10.1037/h0041403

[10] Yushi Jing, Henry Rowley, Jingbin Wang, David Tsai, Chuck Rosenberg, and Michele Covell. 2012. Google image swirl: A large-scale content-based image visualization system. In *Proceedings of the International Conference on World Wide Web*. 539–540. https://doi.org/10.1145/2187980.2188116

[11] Christoph Kofler, Martha Larson, and Alan Hanjalic. 2016. User Intent in Multimedia Search: A Survey of the State of the Art and Future Challenges. *Comput. Surveys* 49, 2 (2016), 36:1–36:37. https://doi.org/10.1145/2954930

[12] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. 2010. What is Twitter, a social network or a news media?. In *Proceedings of the International Conference on World Wide Web*. ACM, New York, USA, 519–600. https://doi.org/10.1145/1772690.1772751

[13] Bjorn Lantz. 2013. Equidistance of Likert-type scales and validation of inferential methods using experiments and simulations. *The Electronic Journal of Business Research Methods* 11, 1 (2013), 16–28.

[14] Mathias Lux, Michael Riegler, Pål Halvorsen, Konstantin Pogorelov, and Nektarios Anagnostopoulos. 2016. LIRE: Open source visual information retrieval. In *Proceedings of the International Conference on Multimedia Systems*. ACM, New York, USA, 30:1–30:4. https://doi.org/10.1145/2910017.2910630

[15] Lokman I. Meho and Helen R. Tibbo. 2003. Modeling the Information-Seeking Behavior of Social Scientists: Ellis's Study Revisited. *Journal of the American Society for Information Science and Technology* 54, 6 (2003), 570–587.

[16] Vicki L. O'Day and Robin Jeffries. 1993. Orienteering in an information landscape: How information seekers get from here to there. In *Proceedings of the INTERCHI Conference on Human Factors in Computing Systems*. ACM, New York, USA, 438–445. https://doi.org/10.1145/169059.169365

[17] Peter Pirolli and Stuart Card. 1999. Information foraging. *Psychological Review* 106, 4 (1999), 643–675. https://doi.org/10.1037/0033-295X.106.4.643

[18] Peter Pirolli and Daniel M. Russell. 2011. Introduction to this special issue on sensemaking. *Human-Computer Interaction* 26, 1-2 (2011), 1–8. https://doi.org/10.1080/07370024.2011.556557

[19] Grant Strong, Orland Hoeber, and Minglun Gong. 2010. Visual image browsing and exploration (Vibe): User evaluations of image search tasks. In *International Conference on Active Media Technology*. Springer, Berlin, Heidelberg, 424–435. https://doi.org/10.1007/978-3-642-15470-6_44

[20] Pertti Vakkari. 2003. Task-based information searching. *Annual Review of Information Science and Technology* 37, 1 (2003), 413–464. https://doi.org/10.1002/aris.1440370110

[21] Viswanath Venkatesh and Fred D. Davis. 2000. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science* 46, 2 (2000), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

[22] Matthew O. Ward, Georges Grinstein, and Daniel Keim. 2015. *Interactive Data Visualization: Foundations, Techniques, and Applications* (2nd ed.). A. K. Peters, Ltd., Natick, MA, USA.

[23] Ryen W. White and Resa A. Roth. 2009. *Exploratory Search: Beyond the Query-Response Paradigm*. Morgan & Claypool Publisher, San Rafael, CA. https://doi.org/10.2200/S00174ED1V01Y200901ICR003