ImageFlow: Streaming Image Search

Varun Jampani [1], Gonzalo Ramos [2], & Steven Drucker [2]
[(1)] IIIT-Hyderabad, [(2)] Microsoft Corporation

ABSTRACT
Traditional grid and list representations of image search results are the dominant interaction paradigms that users face on a daily basis, yet it is unclear that such paradigms are well-suited for experiences where the user's task is to browse images for leisure, to discover new information or to seek particular images to represent ideas. We introduce ImageFlow, a novel image search user interface that explores a different alternative to the traditional presentation of image search results. ImageFlow presents image results on a canvas where we map semantic features (e.g., relevance, related queries) to the canvas' spatial dimensions (e.g., x, y, z) in a way that allows for several levels of engagement – from passively viewing a stream of images, to seamlessly navigating through the semantic space and actively collecting images for sharing and reuse. We have implemented our system as a fully functioning prototype, and we report on promising, preliminary usage results.

ACM Classification: H.5.2 [Information Interfaces and Presentation]: User Interfaces – graphical user interfaces

General Terms: Design, Human Factors.

Keywords: Image search and browsing, new interfaces.

INTRODUCTION

The amount of available digital images and videos available on the web has been increasing at an exponential rate. As a result, several commercial search engines have techniques and systems that allow users to search for images given a particular text query. These systems' interaction metaphor is close to that of plain text search interfaces and only depart from this model in the form they display search results: in a standard grid view. Researchers have found [1] that image based search differs from text based search in three primary ways: greater use of search for entertainment purposes, more emphasis on resource gathering, and a greater use of query reformulation. It is reasonable to then expect that the design requirements for image search interfaces are different.

We present and implement ImageFlow, a novel image search interface that explores an alternate approach to satisfying core user activities around image search tasks. ImageFlow streams images towards the user in a 3D-like environment and supports both the passive and active exploration of a search result set. A user can type an initial query into a search box and then passively observe images as they flow towards the user. A user can also interact with the system by steering through the flow of images with the mouse. ImageFlow also introduces a new way to explore different semantic and image attributes by mapping them onto its canvas's spatial dimensions.

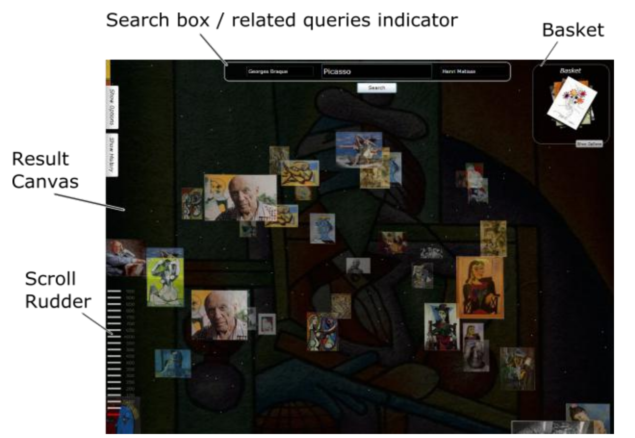

Figure 2: ImageFlow Main Screen

RELATED WORK

The visual nature of image sets has inspired many novel layouts presentations, some of which have been explored in previous work. Liu et al. [3] proposed using an image-similarity organization strategy to organize and present search results. Rodden et al.[5] found that sometimes, this approach actually prevents finding distinct images if all visually similar images are clustered together. Zhang et al. [7] expanded on the notion of using fisheye views for laying out images by using a so called 'Force Transfer' layout which allows a more uniform distribution of images, though all the images are still organized in a grid.

Porta [4] explored new visualization for users that explored an entire database of images. Our interface has similarities to his 'Shot' display, where images are like bullet's fired by a virtual gun towards the screen. Similarities end there as his system has no notion of semantic layout.

Google's swirl [2] shows clusters of related images around a central image and allows the user to navigate from one cluster to the next. This is a promising approach, but the number of images shown at any one time is limited, as it is dependent on the amount of screen real-state.

André et al [1] proposed principles for image search and browsing systems. We follow several of these principles to inform our own design. Their exploration of this design space is a significant accomplishment that opens the door to research such as ours.

DESIGN MOTIVATIONS FOR IMAGE SEARCH

An important difference between text- and image-based searches is that image results can be visually scanned without having to follows the orientation defined by the reading conventions of text. Also, relevance of image search results is dependent to a person's subjectivity as the meaning of an image might not as explicit as plain text. Presentation of image results in the form of an ordered list might be inadequate, especially if a user wants to perform comparisons between different results [3].

Ranking considerations aside, user intent may differ between image and text queries. André et al. [1] summarize some of the characteristics of image search:

- It can be 'exploratory' and 'goal-directed'. But, it tends to be more exploratory than web search.
- It is often used purely for entertainment to play or explore in visual space with no end goal.
- It encourages tangents. If a person is browsing for entertainment, this kind of tangent is actually desired.
- Users view more pages of search results, they spend more time looking at those pages of search results, and they click on more results than web searches.
- Image search tasks tend to explore a single line of questioning or topicality, iterating on related queries.

From the above discussion, we identify three main features that an image search interface should support:

- Rapid browsing. It should be able to present as many images as possible as clearly as possible.
- Easy collection of results for later reuse or sharing.
- Easy and direct query refinement.

IMAGEFLOW SEARCH PROTOTYPE

Guided by the previous section, we have developed a novel interface for searching images over the web, we call this interface *ImageFlow*. We developed our system using the Microsoft Silverlight SDK and the Bing Search API.

The UI of the system is divided in two regions. The top region of the screen contains a traditional search box. The bottom region displays the image search results (Figure 1). We call this area the result canvas.

RESULTS CANVAS

After users enter a search term into the system's search box, results appear in the results canvas as image thumbnails flying towards the user. The desired effect is that of users continuously "flying" though a star field of images, where the depth (z-index) of an image is directly mapped to its relevance, i.e., an image far away has been assigned less relevance by the image search engine. As the system presents the results of a search, the system also renders small dots that as a whole aim to reinforce the illusion of flying through a star field.

The layout of image results in a 3D-space configuration has a number of interesting properties. In addition to being eye-catching, this layout permits us to fit more images on a screen than with a grid view. The "flying through space" metaphor invites exploration, can support loosely defined searches, and foster serendipitous discoveries. Furthermore, such layout lets us explore interesting mappings of meaning to physical x and y dimensions, in the same way the z-axis maps to relevance.

Users can navigate through the result canvas in a passive or active way. While passive, users just let the system run and observe how images from a result set slowly fly towards them.

Alternatively, users can use the mouse to directly interact with the result of as query. The mouse wheel lets users fly forwards or backwards in the direction of the Z axis, i.e., scrolling through results from most to least relevant and vice versa. In the absence of a mouse wheel, users can drag or click on a particular spot in the depth rudder to achieve a similar effect. Users can reveal occluding thumbnails by moving up, down, left or right in the space thus effectively changing the point-of-view into the 3D field. Users do this action by dragging the image result canvas.

When a user holds the mouse's cursor over, or clicks on an image's thumbnail the system reveals information such as title, size and source. Clicking on an image enlarges it and brings it into the foreground. One exits this mode by clicking on the image again.

MAPPING IMAGE ATTRIBUTES

Internet image search sessions often involve query refinement or reformulation. In addition to explicitly altering a search term, current Internet search engines provide users with functionality that allows users to filter content by size, color, type etc., and to access search result sets from "similar" or "related" content.

This functionality is accessible through links or selectors in the UI and it often takes users away from the result to a new page or dynamically re-arrange search results. In contrast to these types of refinement that replaces one view with another, our system explores the idea of mapping different image features to 3D space. The underlying idea is that users can access regions of the search result set by navigating space (e.g., go up to see black and white images, go left to see older pictures), rather than pivot or filter by clicking links. This type of spatially stable representation of the result set has its trade-off when compared to the standard way results are presented in current search engines, yet we believe it has number of interesting properties worth considering. For example, for certain features, a semantically-organized canvas allows for reversible interactions. Also, because interactions occur by navigating through 3D space, one does not need to take real-state away from the canvas (e.g. a page's left panel) for pivoting UI elements such as menus, list of links, etc. Finally, some features might not easily described with words, yet still comprehended in the context of a spatial organizational principle.

In the particular case of our system, we chose to map the canvas 3D space as follows: z, y, and x to relevance, color / black and white and related searches respectively. These choices represent a possible set of valid mappings and for the purposes and scope of this paper they will be the ones tacitly in place, unless noted otherwise.

We believe that mapping related queries to the horizontal dimension is an attractive proposition as it generates a potentially infinite quilt of results where canvas sections connect at their seams with related content in a way that makes visual sense. E.g., a search for "Michael Jackson" will produce a canvas with images of Michael Jackson and as one moves to the right, the result images will show images of Michael and of Janet Jackson until all of the images correspond to the search "Janet Jackson". Figure illustrates the concept using colors as search and related terms.

To reinforce the effect of laterally dragging the canvas we use the search box as indicator of the result dominating the result canvas. In particular, we show to each side of the search box disabled search boxes containing the top two related queries to the center box (Figure 3). As a user drags the canvas left or right, these search boxes move accordingly. When a search box reaches the center of the screen it becomes the active one, and the image set from that query dominates the result canvas.

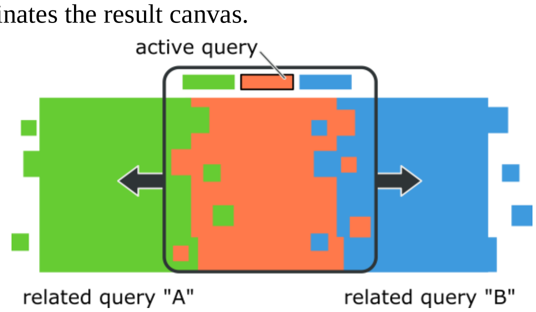

Figure 3: Related results are mapped to the x axis. Dragging the canvas to the left/right "selects" the related query.

As a related search query becomes the central query, we compute its top two related queries and populate the appropriate boxes at the left and right. Our current implementation does not check for cycles in the graph defined by related queries, and as a result, some queries can result in a small set of related queries. We plan to perform a better analysis of these graphs so as to avoid cycles and to provide users with a richer set of related query choices.

SUPPORTING USERS' TASKS

Searching and collecting images are two intimately related tasks. It is common to find the action of searching for an image as a component of the larger task of collecting a set of images, e.g., images for a particular presentation on a topic. Conversely, a loosely defined searching for an image often results in a person selecting and saving a shortlist of candidates from a large set of results. Loosely defined searches or serendipitous searches also often result in the user collecting a particular set of results, e.g., celebrity photo searches.

Imageflow lets users have access to a image basket to collect images. The basket is on the top-right corner of the result canvas (Figure 1) and in its minimized state it displays a pile of the collected images. Users collect images by clicking on the basket icon that appears result thumbnails when the user's mouse is over a thumbnail.

Users expand and minimize the basket by clicking on its "resize" toggle button. While expanded, the basket is divided in two main sections: one on the left containing its contents in a grid, and the other on the right presenting an enlarged preview area. Users can change what image is previewed by clicking on a particular thumbnail in the grid.

Besides previews, the basket allows a number of browsing support operations such as sharing, storing, importing, slideshows, etc. We implemented some of these, but their detailed description escapes the scope of this paper.

Figure 4: ImageFlow's basket. A basket icon appears over an image when the mouse is over it (left). Clicking the icon adds the image into the basket (right).

INFORMAL SYSTEM EVALUATION

In most cases usefulness of as system correlates to traditional metrics of performance such as time and error rates. However in the case of image browsing systems (where task are often open-ended or exploratory) it is not completely determined by them.

We designed our evaluation around explicit and implicit search tasks, which aim to uncover usability issues around explicit and implicit information gathering scenarios. As users perform the same tasks with a new and a baseline condition they have great experience with, we expect to hear them articulate key design features from both interfaces that otherwise would go unnoticed.

At the end of the search tasks, we presented participants with a questionnaire capturing basic demographic information, preferences and additional feedback.

We implemented the study as a Silverlight application and deployed it within our company's Intranet.

PARTICIPANTS

We recruited twelve participants (8 male, 4 female), ages 20-40 from our company via e-mail announcements. They were familiar with traditional web image search interfaces.

RESULTS

Our informal evaluation aimed to reveal user's impressions of the Imageflow and classic interfaces in the context of casual browsing and matching images to specific ideas.

Participants unanimously reported their familiarity with the *grid* interface allowed them to perform search tasks rapidly. A participant expressed that the grid layout combined with a flat scroll surface made it easy to get a sense of where he was when browsing the result set, how much data has been seen and how much data remained to be seen. A few participants indicated that the grid was not fun for general browsing tasks. In particular, one participant indicated that "it remains awkward, hard on the eyes and difficult to use when attempting to process information from a large number of images".

The majority of participants commented on how "different", "cool" and "fun" the ImageFlow interface was. Many of these users found that some features (in particular, the animations) that made ImageFlow "cool" also made the experience "distracting".

DISCUSSION

It is challenging to fairly compare a new UI against one people have used for years and for which they have developed effective strategies of use. Still, the feedback from our study reveals several points were ImageFlow can improve:

Explicit Animations. Making constant movement through the image stream an explicit action by the user, e.g., by measuring the speed of the scroll wheel and using an inertia effect to keep the motion going.

Occlusion. Reducing thumbnail occlusion algorithmically and not expecting a user changing to change her viewpoint.

Location awareness. Providing a full sense of location awareness for a very large information space is a hard problem. It is more tractable to give users an idea of what images they had seen before, and which ones they had missed. A history browsing mechanism could assist with this problem.

Grouping similar images. Result sets often contained re results that were different entities, yet looked identical and confused the participants. Grouping visually similar results can reduce both clutter and confusion.

CONCLUSION AND FUTURE WORK

Besides ImageFlow's contributions as a whole, there are a number of its design elements can be applied to other systems and interfaces. E.g., using the horizontal dimension to navigate towards results from related queries is a concept that fits well with the emerging use of infinite scrolling canvases, where the y dimension is mapped to relevance. Having the ability to spatially browse related queries without the need of menus or panels is a concept that can also benefit interactive surfaces with limited input real estate, such as small displays and mobile devices. Also, we are curious after reading through users' feedback as to how do the Grid and the ImageFlow interfaces affect the eye-scanning patterns (and eye strain) users exhibit during different tasks. These issues remain the topic of future work.

REFERENCES

[1] P. André, E. Cutrell, D. Tan, and G. Smith, "Designing Novel Image Search Interfaces by Understanding Unique Characteristics and Usage," Human-Computer Interaction – INTERACT 2009, 2009, pp. 340-353.

[2] Google Swirl. http://image-swirl.googlelabs.com/

[3] Liu, H., Xie, X., Tang, X., Li, Z., and Ma, W. 2004. Effective browsing of web image search results. In *Proceedings of the 6th ACM SIGMM international Workshop on Multimedia information Retrieval* (New York, NY, USA, October 15 - 16, 2004). MIR '04. ACM, New York, NY, 84-90.

[4] Porta, M. 2006. Browsing large collections of images through unconventional visualization techniques. In *Proceedings of the Working Conference on Advanced Visual interfaces* (Venezia, Italy, May 23 - 26, 2006). AVI '06. ACM, New York, NY, 440-444.

[5] Rodden, K., Basalaj, W., Sinclair, D., and Wood, K. 2001. Does organisation by similarity assist image browsing?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, United States). CHI '01. ACM, New York, NY, 190-197.

[6] von Ahn, L. and Dabbish, L. 2008. Designing games with a purpose. *Commun. ACM* 51, 8 (Aug. 2008), 58-67.

[7] Zhang, L., Chen, L., Jing, F., Deng, K., and Ma, W. 2006. EnjoyPhoto: a vertical image search engine for enjoying high-quality photos. In Proceedings of the 14th Annual ACM international Conference on Multimedia (Santa Barbara, CA, USA, October 23 - 27, 2006). MULTIMEDIA '06. ACM, New York, NY, 367-376.