**All-visual retrieval: How people search and respond to an affect-driven visual information retrieval system**

**Gerald Benoît**  
*Associate Professor, Computer Science & Graduate School of Library & Information Science Simmons College, 300 The Fenway, Boston, MA benoit@simmons.edu, benoit@fas.harvard.edu*

**Naresh Agarwal**  
*Assistant Professor*  
Graduate School of Library & Information Science Simmons College, 300 The Fenway, Boston, MA agarwal@simmons.edu

**ABSTRACT**  
The design of information retrieval (IR) systems must respond to the goals, intentionality and prior domain knowledge of the users. During focus groups conducted, end users complained that instead of looking for specific items, they might be interested in a spectrum of concepts – all things related to something or all things of a particular color. To respond to these needs, an entirely visuals-driven information retrieval system project was developed using a test-bed of copyright-free images reflecting monographs, graphics, and work collections. In the absence of such IR systems, not much is known about how users will interact with a visuals-only retrieval system. This poster describes the project in general and its usage to explore (a) how users interact with graphic-only retrieval for exploring traditional and non-traditional access points and (b) how the affective component impacts the use of such systems. Findings based on the study will help shed light on research based on visual information systems and user behavior when interacting with such systems. The findings will be useful both in designing systems that respond to user needs, and add to prior research in information seeking and retrieval.

**Keywords**  
Visual information retrieval, affective computing.

**INTRODUCTION**  
Most research in information seeking and retrieval assumes a task or problem at hand, which gives rise to a need for information which motivates the seeker to interact with an information retrieval (IR) system. A number of researchers have looked at the concept of intentionality of the seeker (goals, purposes, motivation, etc.) when interacting with an IR system. Limberg (1997) looked at the influence of differing information goals, while Todd (1997) called these information intents. Similarly, Kuhlthau (2004, 2005) found that novices and experts had very different approaches to work tasks.

A number of issues came forth as a result of several focus groups in mixed information settings – combinations of libraries, archives, and museums (Benoît & Hussey, 2011) and student class projects in information retrieval and data interoperability. End users complained that they may not know about a topic well enough to search efficiently for resources (low domain knowledge; novices as per Kuhlthau, 2004) and that they were not always interested in specific items (non-specific need for information in the absence of well-defined tasks). Rather, these users were interested either in (a) exploring a spectrum of concepts, e.g., all things related to the “Barcelona chair” or (b) letting their curiosity drive their interaction with the retrieval system (differing level of goal, intent or intentionality as per Limberg, 1997 and Todd, 1977). One focus group member described vividly her motivation at seeing “a blue elephant” – the Egyptian Middle Kingdom Dynasty 12, ca. 1981-1885 B.C., Statuette of a Hippopotamus (Metropolitan Museum of Art, 2006) and so wanted just to click around and follow blue things in the collection. Traditional access points do not include the object’s color and it is unreasonable to add such data to existing collections.

To address these needs of end users based on the focus groups and class projects, an information retrieval system was modeled that is driven entirely by the user’s reactions to graphics. A test collection using copyright-free materials from Boston Public Library and Simmons College was created from five domains and rudimentary cataloguing added. In the absence of such IR systems, not much is known about how users will interact with a visuals-only retrieval system. When we claim visuals-only, it is to be noted that when a user takes over the mouse over or clicks on an image, the user would see text.

This poster describes the project in general and how it will be used to capture end-users’ responses when presented with a visuals-only retrieval system. We ask two questions:
Motivation and Literature Review

**RQ1:** What is the search behavior of users of a visual-only retrieval system? i.e. how do users interact with graphic-only retrieval for exploring traditional and non-traditional access points?

**RQ2:** How might visually-inspired affective states affect retrieval behaviors and use of such systems?

Findings based on the study will help shed light on research based on visual information systems and user behavior when interacting with such systems. The findings will be useful both in designing systems that respond to user needs, as well as in research in information seeking and retrieval.

In the next section, we will look at the literature review and motivation for the two research questions. This will be followed by a description of the project. We will then briefly discuss the methodology before concluding the poster. Let us now look at the literature review.

**PROJECT DESCRIPTION**

3000 images were digitized and first-level cataloguing records created using auroraDL, a multilingual digital assets management, search, and research platform (see Benoît, 2012). End-users are presented with 11 randomly-selected images and can refresh the images at will (see Figure 1). Users interact with the system by clicking on an image to reveal a “card,” consisting of a thumbnail, a brief text describing the creator of the object and an image of the creator, a text describing the object itself, and text fields enumerating the traditional and non-traditional access points, along with buttons to refine the retrieval set based on the user’s choice.

**Figure 1. Screen Image of Random Images.**

Each image maps to other records that share the same traditional access points and to records that share other kinds of properties. The traditional access points include a combination of VRA4 elements (creator, image-to-work relationships, culture format), Library of Congress subject tracings. The other properties are local use history and

---

**RQ1:** There is a considerable interest and work in the affective dimensions – emotions, feelings, attitude, etc. of the user during information retrieval (see Kuhlthau, 1991’s information search process model). Nahl and Bilal (2007), in their book, bring together research in information science that deals with the relationship between information and emotion. Along with other contextual factors, Agarwal, Xu and Poo (2011) look at the seeker’s inherent lack of comfort, learning orientation and task-self efficacy when interacting with an information source. Gwizdka & Lopatovska (2009) define quite a few aspects of the role of subjective factors in searching. They identify task characteristics and interface as independent factors with various constructs (behavior measures, subjective variables) as dependent variables to record how end-users’ sense of happiness, familiarity, interest, confidence, etc., impact their search behaviors. Gwizdka & Lopatovska (2009) also summarize past studies on the role of subjective behaviors either before search, or during and after search.

---

**RQ2:** Little effort in establishing a benchmark set of images and queries. Doing so would have many benefits in advancing the technology and utility of content-based image retrieval systems.” (Smith, 1998, p.112) In brief, then, there remains a need to understand how people actually interact with image-based systems. While there are advances in supporting VIR (Ward & Reinhard, 2000) and while Boolean searches remain a potential source for testing VIR at the semantic level (Zhang, 2007), the inclusion of heterogeneous sources of data (Müller et al., 2006; Benoît, 2008) and searching by appearance (e.g. Siggelkow, n.d.), projects or computational- or human-identified aspects remains a challenge. Creating profiles of behavior in a closed-set of images/resources might help to identify what motivations lead different people using the same collection to respond as they do, which, in turn, may help searching and managing VIR collections.
color.

The test bed consists of images of maps, daguerreotypes, soldiers and sailors of the American Civil War, anti-slavery manuscripts, sheet music covers, flowers, early chromolithography in product labels, early color off-set printed travel photographs and travel posters from railroad and steamship companies. Most of the images are from the 19th century and the entire collection ranges from the mid-1850s through 1950s. End-users, then, can pursue a traditional topic, such as the Civil War, progressing through a variety of formats and topics but all within the Civil War collection; or range across formats and topics by pursuing what interests them. For example, clicking on an image of the Civil War, the user can move towards a battle scene. There, noticing uniforms, the user can pursue other military uniforms that lead to the sheet music collection. In its turn, it might lead to images of Boston and then over to travel postcards of New England. From there, the user might move to advertisements of the New Haven Railroad or to the brewing company advertisements in the Chromolithography collection.

**User Interaction**

Users visit the site and are presented with eleven randomly-selected images, driven by a PHP script that integrates the collection from MySQL tables and text files that contextualize the creator of the object and the object itself. Taking the mouse over an image reveals a “flip card”, a window offering contextualizing data to provide the end-user with a sense of the accepted significance and values placed on the object. The database provides information about the organization’s holdings based on various properties: LCSH, AAT, and locally-created subject tracings: “why is this creator important”, “what other items are there by this agent?”, and offers searching by agent, subjects, culture, artistic style, use history and visual property related to the currently displayed item (Table 1).

<table>
<thead>
<tr>
<th>Contextualization of the original object</th>
<th>Brief “wikipedia”-type text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context of the agent</td>
<td>“Why is [creator name] important?”</td>
</tr>
</tbody>
</table>

**Database Searches:**

1) What other images are there by this agent?
2) What other images are there on these subjects?
3) What other examples are there of this
   a) culture
   b) artistic style
4) How has this digital object been used before?

<table>
<thead>
<tr>
<th>Database Searches:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) What other images are there by this agent?</td>
</tr>
<tr>
<td>2) What other images are there on these subjects?</td>
</tr>
<tr>
<td>3) What other examples are there of this</td>
</tr>
<tr>
<td>a) culture</td>
</tr>
<tr>
<td>b) artistic style</td>
</tr>
<tr>
<td>4) How has this digital object been used before?</td>
</tr>
</tbody>
</table>

**Table 1. Conceptualizing and default searches from user interaction pop-up window (“flip card”).**

**METHODOLOGY**

**RQ1:** Transaction log analysis and click-through analysis will be utilized to reveal the users’ responses to inputs at different points in their information seeking session. To gather data for the first research question, the user’s mouse action is recorded on the server, noting when and on what property the user reacts: viewing a card, fetching another randomly-selected set, and refining a search. The result is a record noting what properties motivated the user to explore in general and which prompted a further search. The transaction logs will be parsed and measures of association (crosstabs) determined (Argyrous, 2009, p. 112). The data, being not normally distributed with an unknown population, are likely to be asymmetric. The association will be measured using Gamma and Somers’ d.

We might consider screenflow movies to show and capture users visual searching (full image and action mousing, clicking, etc.). Another dimension to consider is “felt time pressure” in the task setting e.g. with a time limit on tasks. Time pressure is a normal feature of task search and is often correlated with frustration, anxiety, errors, lowered confidence, among other affective factors.

**RQ2:** To examine the relationship between subjective factors and retrieval behaviors, we will build on the work of Gwizdka & Lopatovska (2009). The subjective factors (pertaining to the seeker) will be operationalized as:

1) **happiness levels** (how happy are you right now?),
2) **satisfaction with** (how did you feel during the search? Please rate your overall satisfaction with this search experience.) and
3) **confidence in the search results** (I am confident that I found the desired information),
4) **feeling lost during search** (I was lost at some point during this search.),
5) **familiarity with** (how familiar are you with the topic of the search task?) and
6) **interest in the search topic** (how interesting do you find the topic of the search task?),
7) **estimation of task difficulty** (how easy do you think it will be to find information for this task? Did you find it easy or difficult to do this search task?).

Subjects will be recruited to participate in one of two groups: 1) task-based searching (Masters students from the Graduate School of Library and Information Science at Simmons College will be assigned to search) and 2) non-task-based (free) searching (general library population utilizing random search; a research assistant will monitor the searching). In the task-based group, participants will be assigned a mix of 1) fact-finding and 2) information gathering tasks (Gwizdka & Lopatovska, 2009). Following Gwizdka and Lopatovska, the participants will be introduced to the study, asked to sign the consent form, will get to practice the search task, will fill out a pre-session questionnaire, perform the search task, and then fill the...
post-session questionnaire. The questionnaires will gather subjective assessments of their feelings and task characteristics and feelings of lostness, confidence, satisfaction when using the VIR system. The relationship between these subjective factors and retrieval behaviors and system use will be examined. We might also consider using formative assessment in addition to pre- and post-testing to capture the process nature of affect. Emotion is dynamic and formative assessments might help capture some of the dynamic aspects. Online forms can be sent at intervals.

RESULTS, CODE, SET THEORY
A set theoretic description of the project, as well as the source code for capturing user decisions and inputs, and project results are available at pomme.simmons.edu/visir/

CONCLUSION
We have described the Visual IR system developed and its usage to explore (a) how users interact with graphic-only retrieval for exploring traditional and non-traditional access points and (b) how the affective component impacts the use of such systems. Future work will involve gathering data to answer these questions. The findings will be useful both in designing systems that respond to user needs, and add to prior research in information seeking and retrieval. Future work might also involve comparing with other live test beds such as ‘The Commons on Flickr’ and Google Image Search.

ACKNOWLEDGEMENT
This project is supported by the Emily Hollowell Research Fund, Graduate School of Library and Information Science, Simmons College, Boston, MA.

REFERENCES


Zhang, Y.-J. (2007). *Semantic-based visual information retrieval*. IGI.