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

**Gerald Benoit**
*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 (Benoit & 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.

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:

- *RQ1:* What is the search behavior of users of a visual-only retrieval system? i.e. how do users



*ASIST 2012,* October 28-31, 2012, Baltimore, MD, USA.

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 motivation and literature review 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.

## MOTIVATION AND LITERATURE REVIEW

*RQ1:* While there is a considerably large body of literature about visualization techniques, there is little in terms of methods and evaluation of visual information retrieval (VIR) that is widely accepted. Gupta and Jain (1997) noted two approaches to VIR – 1) computational, based on pixels, and how this approach can be turned to specific needs, such as state change (aircraft location detection), and 2) ranking video frames based on pixel changes. Morse, Lewis, and Kai (2001) noted that VIR testing remains inconclusive with little benchmarking and proposed a case study to consider new ways of presenting data. Other work emphasizes the boundary-spanning necessary for applying visuals in information retrieval that concludes browsing, searching, user studies, affective computing, new types of queries, etc. All are needed for the future of retrieval and remain major challenges (Lew, Sebe, Djeraba, & Jain, 2006). Older work such as Rui, Huang, Mehrotra (1997), and Smith (1998) lament the lack of a common test-bed and "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 have been great advances in supporting VIR (Ward & Reinhard, 2000), there remain opportunities in semantic-level searches for testing VIR (Zhang, 2007), the inclusion of heterogeneous sources of data (Müller *et al.*, 2006; Benoit, 2008) and searching by appearance (e.g. Siggelkow, n.d.). 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 to search/manage VIR collections.

*RQ2:* 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.

## PROJECT DESCRIPTION

3000 images were digitized and first-level cataloguing records created using auroraDL, a multilingual digital assets management, search, and research platform (see Benoit, 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.

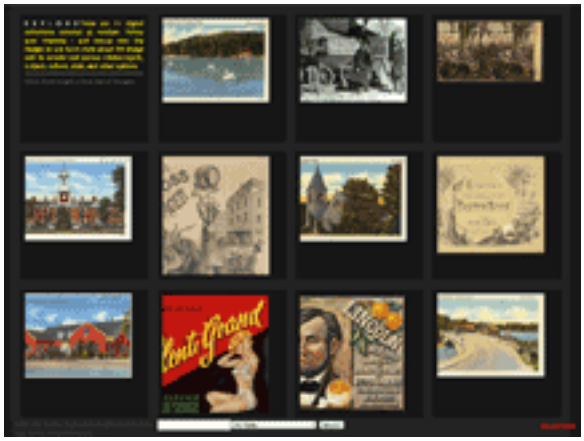

**Figure 1. Screen Image of Random Images.**

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 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.

### Representation as sets

The combining of traditional ($x$) and non-traditional attributes ($y$) is not a concern from an information retrieval point of view. The Collection representation C consists of proper subsets, $c_1$, $c_2$, $c_3$, $c_4$, $c_5$, $c_6$, $c_7$, each representing a domain and no member in $c_i = c_j$. The initial retrieval set is the set of elements $\{x: 1, 2, \ldots 11\}$ where $x \in \{c_{i1} \ldots c_{j11}\}$, $i, j \in C$. The system can be expressed as a rough set collection, letting $I = (U, A)$ consist of attributes and values, where $U$ is a non-empty set of finite objects in the collection and $A$ is a non-empty, finite state of attributes, such that a: $U \rightarrow V_a$ for every $a \in A$. Ua is the set of values that attribute a might take. The search engine retrieves a value $a(x)$ from $V_a$ to each attribute $a$ and object $x$ in the collection $C$'s universe $U$ of objects. Although the collection has non-traditional and traditional access points, they are indistinguishable attributes or properties from $P$: With any $P \subseteq A$, there is an associated equivalence relation IND($P$): $\text{IND}(P) = \{(x,y) \in U^2 \mid \forall a \in P, a(x) = a(y)\}$. The relation IND($P$) is a $P$-indiscernible relation. The partition of $U$ is a family of all equivalence classes of IND($P$). If $(x,y) \in \text{IND}(P)$, then $x$ and $y$ are indiscernible from attributes from $P$ (Hassanien, Suraj, Slezak and Lingras, 2008).

### 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).

| | |
|---|---|
| Contextualization of the original object | Brief 'wikipedia'-type text |
| Context of the agent | 'Why is [creator name] important?' |
| Database searches<br>1. What other images are there by this agent?<br>2. What other items are there on these subjects?<br>3. What other examples of this<br>(a) culture<br>(b) artistic style<br>How has this item been used before? | |

**Table 1. Contextualizing 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.

*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.

## 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.

## ACKNOWLEDGEMENT

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

## REFERENCES

Agarwal, N.K., Xu, Y. (C.), Poo, D.C.C. (2011). A Context-based Investigation into Source Use by Information Seekers. *JASIST*, 62(6), 1087-1104.

Argyrous, G. (2009). *Statistics for research with a guide to SPSS*. (2nd ed). London: Sage.

Benoit, G. & Hussey, L. (2011). Repurposing digital objects: Case studies across the publishing industry. *JASIST*, 62 (2), 363-374.

Benoit, G. (2008, Oct). Interactive 3D Visual Retrieval for Art History Education. *Proceedings of ASIS&T 2008*, 45(1), 1-13.

Benoit, G. (2012). auroraDL and responding to end-user digital library needs. *Proceedings of iConference 2012*, pp. 444-446, New York, NY: ACM.

Gupta, A., & Jain, R. (1997, May). Visual information retrieval. *Communications of the ACM, 40*(5), 71-79.

Gwizdka, J., & Lopatovska, I. (2009). The role of subjective factors in the information search process. *Journal of the American Society for Information Science and Technology.* 60(12).

Hassanien, A. E., Suraj, Z., Slezak, D. and Lingras, P. (Eds.) (2008). *Rough computing: Theories, technologies, and applications.* Information Science Reference. Hershey, PA: IGI Global.

Kuhlthau, C.C. (1991). Inside the Search Process: Information Seeking from the User's Perspective. *JASIST*, 42(5), 361-371.

Kuhlthau, C.C. (2004). *Seeking meaning: a process approach to library and information services* (2nd edition), Westport, CT: Libraries Unlimited.

Kuhlthau, C.C. (Jan 2005). Towards collaboration between information seeking and information retrieval. *Information Research*, 10(2).

Lew, M. S., Sebe, N., Djeraba, C., & Jain, R. (2006). Content-based multimedia information retrieval: state of the art and challenges. *ACM Transactions on multimedia computing, communications, and applications, 2*(1), 1-19.

Limberg, L. (1997). Information use for learning purposes. In P. Vakkari, R. Savolainen and B. Dervin (Eds.) *Information seeking in context* (Tampere, Finland: Aug 14-16, 1996), London: Taylor Graham, 275-289.

Metropolitan Museum of Art. (2006, Oct.). Statuette of a Hippopotamus [Egyptian; Middle Egypt, Meir] (17.9.1). In *Heilbrunn Timeline of Art History*. New York. Retrieved June 10, 2012 from http://www.metmuseum.org/toah/works-of-art/17.9.1.

Morse, E., Lewis, M., & Kai, A. O. (2001, Nov). Testing visual information retrieval methodologies. *JASIST*, 53(1), 28-40.

Müller, H., Clough, P., Hersh, W., Deselaers, T., Lehmann, T.M., Janvier, B. and Geissbuhler, A. (2006). Using heterogeneous annotation and visual information for the benchmarking of image retrieval systems. In S. Santini, R. Schettini and T. Gevers (Eds). *Internet Imaging VII. Proceedings of SPIE Volume: 6061*. Retrieved June 10, 2012, from citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.69.2557.pdf

Nahl, D. and Bilal, D. (2007). *Information and Emotion: The Emergent Affective Paradigm in Information Behavior Research and Theory*. ASIS&T Monograph Series. Medford, NJ: Information Today, Inc.

Rui, Y., Huang, T. S., & Mehrotra, S., (1997, Oct.) Content-based image retrieval with relevance feedback in MARS. International conference on Image Processing, 2, vol. 2., 815-818.

Siggelkow, S. (n.d.). SIMBA. Albert-Ludwigs-Universität-Freiburg. Retrieved June 10, 2012, from http://simba.informatik.uni-freiburg.de/

Smith, J. R. (1998, June). Image retrieval evaluation. *Proceedings of the IEEE Workshop on content-based access of images and video libraries*, 112-113.

Todd, R.J. (1997). Information utilisation: a cognitive analysis of how girls utilize drug information based on Brookes' fundamental equation [K[S] + Δ I = K[S + Δ S]. In P. Vakkari, R. Savolainen and B. Dervin (Eds.), *Information seeking in context* (Tampere, Finland, Aug 14-16, 1996), London: Taylor Graham, 351-370.

Ward, R., & Reinhardt, M. (2000, Sept.). *Oracle visual information retrieval*. Release 8.1.7, Part No. A85335-01. Redwood City, CA: Oracle Corp.

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