From Concepts to Implementation and Visualization: Tools from a Team-Based Approach to IR

Uma Murthy¹, Ricardo da Silva Torres², Edward A. Fox¹, Logambigai Venkatachalam¹, Seungwon Yang¹ and Marcos A. Gonçalves³

¹Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, USA
²Institute of Computing, University of Campinas, SP, Brazil
³Computer Science Department, Federal University of Minas Gerais, Belo Horizonte, Brazil
{umurthy, fox, lokeya, seungwon} @vt.edu, rtorres@ic.unicamp.br, mgoncalv@dcc.ufmg.br

EXTENDED ABSTRACT
Researchers have been studying and developing teaching materials for information retrieval (IR), such as [3]. Toolkits also have been built that provide hands-on experience to students. For example, IR-Toolbox [4] is an effort to close the gap between the students’ understanding of IR concepts and real-life indexing and search systems. Such tools might be good for helping students in non-technical areas such as in the Library and Information Science field to develop their conceptual model of search engines. However, they do not cover emerging topics and skills, such as content-based image retrieval (CBIR) and fusion search. Although there is open source software (such as those in http://www.searchtools.com/tools/tools-opensource.html) that can be used to teach basic and advanced IR topics, they require a student to have high-level technical knowledge and to spend a long time to gain a practical understanding of these topics.

We present a new and rapid approach to teach basic and advanced IR topics, such as text retrieval, web-based IR, CBIR, and fusion search, to Computer Science (CS) graduate students. We designed projects that would help students grasp the abovementioned IR topics. Students, working in teams, were given a practical application to start with – the Superimposed Application for Image Description and Retrieval [5]. SAIDR (earlier, SIERRA) allows users to associate parts of images with multimedia information such as text annotations. Also, users may retrieve information in one of two ways: (1) Perform text-based retrieval on annotations; (2) Perform CBIR on images and parts of images that look like a query image (or part of a query image).

Each team was asked to build an enhancement for this application, involving text retrieval and/or CBIR, in three weeks time. The sub-projects are described in Table 1. The outcome of this activity was that students learned about IR concepts while being able to relate their applicability to a real world problem (Figure 1). Details of these projects may be found at http://collab.dlib.vt.edu/runwiki/wiki.pl?TabletPcImageRetrievalSuperimposedInformation. We will demonstrate the tools developed along with the IR concepts they illustrate (Table 1). We believe these tools may aid others to learn about basic and advanced topics in IR.

ACKNOWLEDGMENTS
Thanks go to NSF (DUE-0435059), Microsoft (tablet PC grant), CAPES, FAPESP, and CNPq to help support this work. Thanks to all students of CS5604, who worked on the sub-projects (listed at http://collab.dlib.vt.edu/runwiki/wiki.pl?SICBIRStudentsRoster).

REFERENCES