SIERRA - A Superimposed Application for Enhanced Image Description and Retrieval

Uma Murthy[1], Ricardo da S. Torres[2], and Edward A. Fox[1]

[1] Department of Computer Science, Virginia Tech,
Blacksburg, VA 24061, USA
{umurthy, fox}@vt.edu

[2] Institute of Computing, State University of Campinas,
Campinas, SP, Brazil, 13084-851
rtorres@ic.unicamp.br

Abstract. In this demo proposal, we describe our prototype application, SIERRA, which combines text-based and content-based image retrieval and allows users to link together image content of varying document granularity with related data like annotations. To achieve this, we use the concept of superimposed information (SI), which enables users to (a) deal with information of varying granularity (sub-document to complete document), and (b) select or work with information elements at sub-document level while retaining the original context.

Description

In many image-based applications, like biomedical teaching, research, and diagnosis, there is need to link (or integrate) image content with other multimedia information: text annotations, metadata (keywords or ontological terms), audio-visual presentations, etc. Not only does this contribute to richer image descriptions, it also helps in more effective retrieval of images and related information [10]. Further, for complex images (e.g., images with plenty of detail, or with specific hard-to find details), there may be a need to isolate and work with parts of images (meaningful objects within the image) without losing the original context (the actual image). For example, an ichthyologist may want to annotate a particular part of a fish after seeing annotations from other ichthyologists on the same type of fish. Yet, current image-based systems either focus on content-based [8] or text-based descriptions [3, 4]. Some systems [1, 2, 6, 7, 12], which combine both techniques to enhance the image annotation process, provide limited support for linking image content, at varying document granularity, to other multimedia content (like ontological terms, video descriptions, etc.).

We have developed SIERRA, an application which combines text- and content-based image retrieval so users can relate images, at varying document granularity, to other multimedia content, applying the concept of superimposed information (SI). SI refers to new information (or new interpretations) laid over existing information [9] (like bookmarks, annotations, etc.). Superimposed applications (SAs) allow users to

lay new interpretations over existing or base information. SAs employ "marks", which are references to selected regions within base information [11]. SAs enable users to (a) deal with information of varying granularity, and (b) select or work with information elements at sub-document level while retaining the original context.

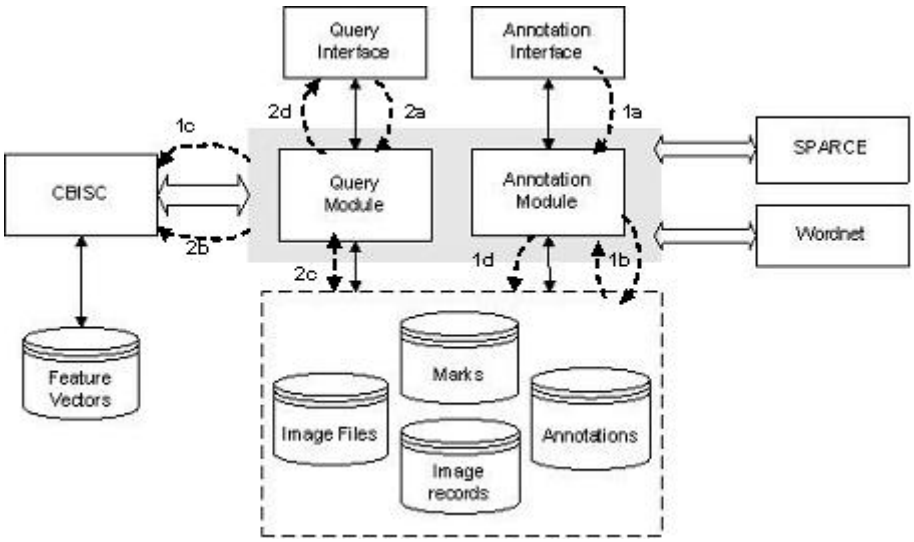

Figure 1. The high-level design of SIERRA and labeled steps involved in two major functions of SIERRA – 1) the marking and annotation process, and 2) the retrieval process

Figure 1 shows the high-level design of SIERRA. It consists of two main modules – the annotation module and the query module. The design is such that other existing modules may be plugged in to facilitate richer image description and retrieval. SIERRA makes use of the Content-Based Image Search Component (CBISC) [13], an OAI-compliant component that supports queries on image collections. It retrieves images similar to a user-defined pattern (e.g., color layout of an image, image sketch, etc.) based on content properties (e.g., shape, color, or texture), which are often encoded in terms of image descriptors. In addition, we foresee integration with other types of components like the ontology WordNet [5] (for suggesting annotation terms), and the Superimposed Pluggable Architecture for Contexts and Excerpts (SPARCE) – middleware for managing "marks" over text, audio, and video content [11]. Integration with SPARCE will enable associating image marks and annotations with marks in other content types.

Two major applications of SIERRA include: 1) the image marking and annotation process, and 2) the image/mark/annotation retrieval process. Figure 1 traces the high-level steps of scenarios involving each of these processes. 1a) the user identifies an image, marks a region of interest and annotates that region with keywords; 1b) a mark is created and all mark-relevant information is stored; 1c) the content of the sub-image referred to by the mark is stored in the CBISC; 1d) annotation information is

stored; 2a) the user identifies an image, then marks (selects) a region within the image and uses this mark to query SIERRA; 2b) SIERRA uses the sub-image referenced by the mark created by the user to query the CBISC and get a list of images or marks similar to the queried mark. 2c) all annotations associated with the result images/marks are retrieved; 2d) the user is able to view the result images/marks with associated annotations.

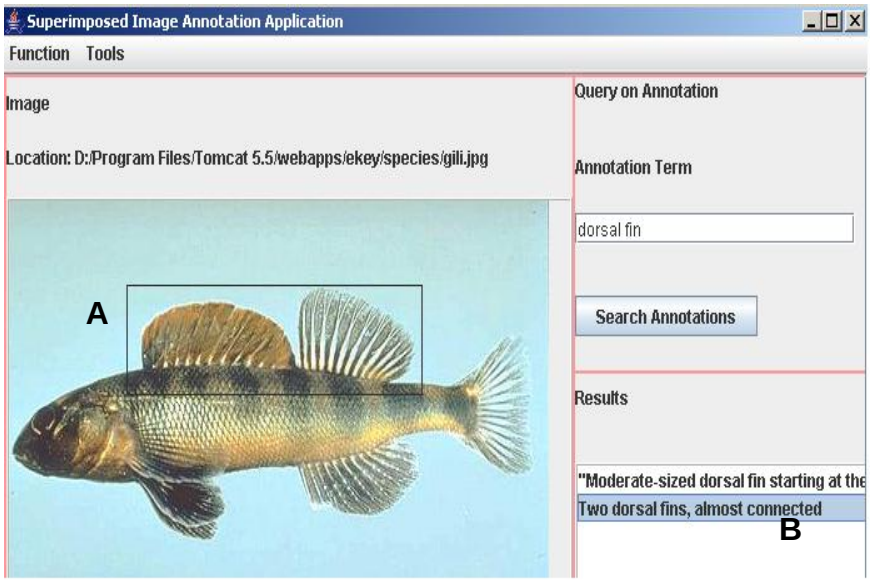

Figure 2. A snapshot of the initial prototype of SIERRA. All annotations with the phrase "dorsal fin" are listed. On selecting an annotation, the associated mark (and the containing image) is displayed. A) Mark associated with selected annotation; B) Selected annotation

The current prototype of SIERRA (see Figure 2) allows users to select parts of images and associate them with text annotations. Then, users can retrieve information as annotations and associated marks in two ways, either for (1) a specified image, or (2) annotations containing specified query terms. The first capability illustrates how this SA differs from a typical hypermedia application, in that important work can be done just with the marks, ignoring the base information.

This prototype has been developed in Java and makes use of the Java 2D API for image manipulation. Data is stored in a PostgreSQL database.

We are integrating this prototype with our content-based image search component [13] to extract content from complete images and marks. We will then undertake a formative usability evaluation on the prototype. Future work on this application includes integration with the ontology WordNet [5] and with SPARCE [11].

Acknowledgements

This work is funded in part by NSF DUE-0435059, under the NSDL research track. Ricardo Torres' work is funded in part by CAPES, FAPESP, FAEPEX, the CNPq WebMaps and AgroFlow project, and by a Microsoft eScience grant.

References

1. Barnard, K., Duygulu, P., Freitas, N.d., Forsyth, D., Blei, D. and Jordan, M.I. Matching Words and Pictures. *Journal of Machine Learning Research*, *3* (6): 1107-1135.
2. Benjamin, B.B., PhotoMesa: a zoomable image browser using quantum treemaps and bubblemaps. In *Proceedings of the 14th annual ACM symposium on user interface software and technology*, Orlando, Florida, 2001, ACM Press, 71 - 80.
3. Chen-Yu, L., Von-Wun, S. and Yi-Ting, F., How to annotate an image? the need of an image annotation guide agent. In *Proceedings of the 4th ACM/IEEE-CS joint conference on digital libraries*, Tuscon, AZ, USA, 2004, ACM Press, 394 - 394.
4. Elin, G., Rohlfing, M. and Parenti, M. Fotonotes.net - Image Annotation Standard and Scripts, 2004, http://www.fotonotes.net/.
5. Fellbaum, C. (ed.), *WordNet: An Electronic Lexical Database*. The MIT Press, Cambridge, MA, 2001.
6. Freitas, R. and Torres, R., OntoSAIA: Um Ambiente Baseado em Ontologias para Recuperação e Anotação Semi-Automática de Imagens. In *Proceedings of Workshop em Bibliotecas Digitais*, Uberlândia, MG, Brasil, 2005, 60-79.
7. Halaschek-Wiener, C., Schain, A., Golbeck, J., Grove, M., Parsia, B. and Hendler, J., A Flexible Approach for Managing Digital Images on the Semantic Web. *Presented at* the 5th International Workshop on Knowledge Markup and Semantic Annotation, Galway, Ireland, 2005.
8. Lieberman, H., Rosenzweig, E. and Singh, P. Aria: an agent for annotating and retrieving images. *Computer*, *34* (7): 57-62.
9. Maier, D. and Delcambre, L., Superimposed Information for the Internet. In *Proceedings of the WebDB Workshop*, Philadelphia, PA, USA, 1999, 1-9.
10. Muller, H., Michoux, N., Bandon, D. and Geissbuhler, A. A review of content-based image retrieval systems in medical applications - clinical benefits and future directions. *International Journal of Medical Informatics*, *73* (1): 1-23.
11. Murthy, S., Maier, D., Delcambre, L. and Bowers, S., Putting Integrated Information into Context: Superimposing Conceptual Models with SPARCE. In *Proceedings of the First Asia-Pacific Conference of Conceptual Modeling*, Denedin, New Zealand, 2004, 71-80.
12. Stein, A. COLLATE - Collaboratory for Annotation, Indexing and Retrieval of Digitized Historical Archive Material. IST-1999-20882, Fraunhofer IPSI, Dolivostrasse 15, D-64293, Place Published, 2004,
13. Torres, R., Medeiros, C.B., Goncalves, M.A. and Fox, E.A. A Digital Library Framework for Biodiversity Information System. *International Journal on Digital Libraries*, *6* (1): 3 - 17.