

## Special Section on Internet Imaging

### Giordano Beretta

Hewlett-Packard Company  
1501 Page Mill Road, 4U-6,  
Palo Alto, CA 94304 USA  
E-mail: beretta@hpl.hp.com

### Raimondo Schettini

ITIM CNR  
Via Ampere 56  
I-20131 Milano, Italy  
E-mail: centaura@itim.mi.cnr.it

Internet imaging differs from other forms of electronic imaging in that it employs an internet (network of networks) as a transmission vehicle. However, the internet is only one component (albeit a major one) in the total imaging system. The total system comprises client applications internet-worked with server applications, as well as offline authoring tools.

The internet is an evolving communication system. Its functionality, reliability, scaling properties, and performance limits are largely unknown. The transmission of images over the internet pushes the engineering envelope more than most applications. Consequently, the issues we are interested in exploring pertain to all aspects of the total system, not just images or imaging algorithms.

This emphasis on systems is what sets internet imaging apart from other electronic imaging fields. For a local imaging application, even when it is split between a client and a server linked by an Ethernet, a system can be designed by stringing algorithms in a pipeline. If performance is an issue, it is easy to identify the weak link and replace it with a better performing component.

On the internet, the servers are unknown, the clients are unknown, and the network is unknown. The system is not easily predictable and the result is that the most common problem today is scalability. To be successful one has to follow a top-down design strategy, where the first step is a detailed analysis of the problems to be solved. When

a solution is invented, algorithms are selected to produce a balanced system, instead of choosing algorithms of best absolute performance as is done in bottom-up approaches.

The paper on the Visible Human by Figueiredo and Hersch is a good example illustrating these fundamentals. Today, storing a 49-Gbyte 3-dimensional volume is not hard, and a RAID disk array can deliver fast access times. However, storage space and seek time are not the limiting factors for the extraction of ruled surfaces from large 3-dimensional medical images. The problem is one of load balancing, which requires detailed performance measurements for scalability. Eventually, a specialized parallel file striping system must be designed and optimized. Implementing and maintaining a system that must grow as more data becomes available and as surgeons require new staging techniques for tumors is practical only in a centralized solution served on the internet.

After e-mail, the most popular application on the internet is the World Wide Web, which is a hypertext system and as such is useful only when it can easily be navigated through a visual interface,<sup>1</sup> and search results are presented in a context,<sup>2</sup> as is illustrated for example by the KartOO search engine. Navigation requires structure,<sup>3</sup> and although techniques such as ontologies have been known for years, the particularities for decoupling and splicing ontologies are not yet sufficiently understood.<sup>4</sup>

In a recent paper, the Jörgensens

have described the challenges of developing an image indexing vocabulary,<sup>5</sup> and yet we know that taxonomies are not sufficiently powerful for efficiently finding related information through navigation.<sup>6</sup> Progress in bioinformatics has given us new computational tools that will allow the development of new collaborative structuring methodologies based on ontologies.

Another example of how wrong things can go when the fundamentals of internet imaging are not understood is content-based image retrieval (CBIR) systems. Today they are part of all the major search engines on the internet, and anyone who has tried to use them for real work has experienced how useless they are.

Although over the years a number of CBIR algorithms has been proposed, none has stood out as being particularly robust, despite the fact that each claims to perform best on some benchmark. Unfortunately there is no universally accepted benchmark for CBIR and the lack of a metric is probably one of the main causes for the poor quality of today's algorithms—without a performance metric is it impossible to diagnose the shortcomings of a particular algorithm and identify the critical control points.<sup>7</sup>

An international effort is underway to create a benchmark for CBIR,<sup>8</sup> similar to what was done in the past in the TREC effort for text retrieval. This requires an extensive collaboration to annotate a sufficiently large image corpus, which establishes the ground truth

against which performance can be measured. A tool has recently been developed for this purpose.<sup>9</sup>

One particularly nasty problem on the internet is that a preponderance of the available images is not normalized towards a standard rendering intent, as is done in conventional stock image collections. In fact, the subtleties of the various references for color encoding in the stages of a distributed workflow are only recently being described and standardized.<sup>10</sup>

A correct output-referred color encoding cannot be determined manually in the case of a large image corpus, as it is typically encountered in internet imaging. Contrary to silver halide photography, where contemporary films can largely compensate for illumination deviating from the intended illuminant, this is not the case in digital photography. This problem has led to the proposal of a number of automatic white balancing algorithms to compensate for these discrepancies by estimating the illuminant and applying a color appearance transformation.

To benchmark these algorithms it is necessary to develop a ground truth for combinations of illuminations and assumed illuminants. Tominaga's paper on a "Natural image database and its use for scene illuminant estimation" describes how such a database is created and how it is used in practice.

Digitalization, compression, and archiving of visual information has become popular, inexpensive and straightforward. Yet, the retrieval of this information on the World Wide Web—being highly distributed and minimally indexed—is far from being effective and efficient. A hot research topic is the definition of feasible strategies to minimize the semantic gap between the low-level features that can be automatically extracted from the visual contents of an image and the human interpretation of such contents. Two different approaches to this problem are described in the last two papers.

Lienhart and Hartmann present novel and effective algorithms for classifying images on the web. This type of algorithms will be a key element in the next generation of search engines, which will have to classify the web page media contents automatically. Experiments and results are reported and discussed about distinguishing photo-like images from graphical images, actual photos from only photo-like, but artificial images and presentation slides/scientific posters from comics.

The paper "Multimodal search in collections of images and text" by Santini introduces the intriguing issue of how to infer meaning of an image from both its pictorial content and its context. The author describes a data model and a query algebra for databases of images immersed in the World Wide Web. The author's model provides a semantic structure that, taking into account the connection with the text and pages containing them, enriches the information that can be recovered from the images themselves.

## References

1. F.J. Verbeek *et al.*, "Visualization of complex data sets over Internet: 2D and 3D visualization of the 3D digital atlas of zebra fish development," *Proc. SPIE* **4672**, 20–29 (January 2002).
2. G. Ciocca *et al.*, "A multimedia search engine with relevance feedback," *Proc. SPIE* **4672**, 243–251 (January 2002).
3. G. Beretta, "WWW+ Structure= Knowledge," *Technical Report HPL-96-99*, HP Laboratories, Palo Alto, June 1996, <http://www.hpl.hp.com/techreports/96/HPL-96-99.html>
4. J. Tillinghast *et al.*, "Structure and Navigation for Electronic Publishing," *Proc. SPIE* **3300**, 38–45 (January 1998).
5. C. Jørgensen *et al.*, "Testing a vocabulary for image indexing and ground truthing," *Proc. SPIE* **4672**, 207–215 (January 2002).
6. D.J. Watts *et al.*, "Identity and Search in Social Networks," *Science* **296**, 1302–1305 (May 2002).
7. N.J. Gunther *et al.*, "A Benchmark for Image Retrieval using Distributed Systems over the Internet: BIRDS-I," *Proc. SPIE* **4311**, 252–267 (January 2001).
8. <http://www.benchathlon.net/>
9. T. Pfund *et al.*, "A Dynamic Multimedia Annotation Tool," *Proc. SPIE* **4672**, 216–224 (January 2002).
10. S. Süsstrunk, "Color Encodings for Image Databases," *Proc. SPIE* **4672**, 174–180 (January 2002).



**Giordano Beretta**

is with the Imaging Systems Laboratory at Hewlett-Packard. He has been instrumental in bootstrapping the internet imaging community: in collaboration with Robert Buckley he

has developed a course on "Color Imaging on the Internet," which they have taught at several IS&T and SPIE conferences; with Raimondo Schettini he has started a series of Internet Imaging conferences at the IS&T/SPIE Electronic Imaging Symposium; and he has nursed the Benchathlon effort through its first two years. He is a Fellow of the IS&T and the SPIE.



**Raimondo Schettini**

is an associate professor at DISCO, University of Milano Bicocca. He has been associated with Italian National Research Council (CNR) since 1987. In 1994 he moved to

the Institute of Multimedia Information Technologies, where he is currently in charge of the Imaging and Vision Lab. He has been team leader in several research projects and published more than 130 refereed papers on image processing, analysis and reproduction, and image content-based indexing and retrieval. He is member of the CIE TC 8/3. He has been General Co-Chairman of the 1st Workshop on Image and Video Content-based Retrieval (1998), and general co-chair of the EI Internet Imaging Conferences (2000-2002). He was general co-chair of the First European Conference on Color in Graphics, Imaging and Vision (CGIV'2002).