AIC Color 01, Proc. 9th International Congress of the AIC, SPIE Vol. 4421, 24–29 June 2001, Rochester, USA

Trends in color imaging on the Internet

Giordano B. Beretta,* Hewlett-Packard and Robert R. Buckley,** Xerox

ABSTRACT

We review three technological changes that will have an impact on color imaging on the Internet: bright LCD display replacing dim CRTs, compression based on foveation, and contents based image retrieval.

Keywords: color consistency, rendering, visual compression, image retrieval, LCD displays

1. INTRODUCTION

At the Color '97 AIC meeting in Kyoto one of us had given a paper[1] on this topic, in which we discussed color issues with the various Internet image formats and compared compression in the color domain to compression in the spatial domain. Our trend predictions at the time relied on the availability of cheap processing and on the observation that Internet access was becoming faster and cheaper. In fact, today's Internet is characterized by rich graphic content, created with one of the numerous software packages available for this purpose and browsed over fast connections.

After four years and with the speed at which the Internet and Web have progressed, we are at a juncture in Internet imaging where an update is necessary. The current trend is for people to not be any longer tethered to a desktop; instead they use the Internet wherever.

The industry is responding with small mobile Internet devices—palm-tops, personal digital assistants (PDA) and multi-function cellular phones—where the communications channels have lower bandwidth and the devices themselves have only a small screen real-estate. Combined with their low power requirements and low purchase price, these mobile devices allow a larger portion of the world's population to use the Internet in an even more personalized way than has been possible with the desktop PC. Multimedia in general and images in particular, are expected to play a significant role in defining such personalization.[5]

In a mobile wireless world of casual users it is necessary to develop new methods for serving color images on the Internet. In this paper we discuss three technological changes for consideration when developing research strategies for color imaging on the Internet: display technologies, compression for very low bit rates, and image retrieval. The technologies developed for the nomadic users will be applied also to the desktop, where LCDs will replace CRT displays.

2. LCD DISPLAYS AND COLOR INTEGRITY

2.1 LCD Display Technology

Today, LCD devices are mainstream. In 1997, a professional quality 19" CRT display for 1280×1024 pixels cost approximately $1500; today a 17" LCD display with 1280×1024 pixels costs $1000. It is true that CRTs are also becoming less expensive, but at a slower rate. According to an article in the March 2001 SID Information Display, the price-ratio difference will fall from 4.1× in 2000 to 2.0× in 2005. This is for desktop and laptop computers; in PDAs we will see the adoption of OLED displays, which from a color imaging perspective have the same implications.

An LCD monitor cannot be calibrated the same way as a CRT. A slightly more complex model has to be derived because LCDs often have sigmoidal TRCs. However, as has been shown by Gibson and Fairchild[4] and more recently by Okano and Shiotani,[10] a fairly accurate characterization of LCD displays is possible. Today's commercial software for ICC profile creation copes very well with LCD devices.

The traditional aim of computational color reproduction has been to achieve color fidelity. The first step was achieving a colorimetric match between a hardcopy original and a printed copy. The second step was to achieve an appearance match between a softcopy and a printed copy. These methods are based on colorimetry and efforts have been made to port the underlying methodology to the Internet. Hunt[7] reviewed the requirements for color reproduction on the Internet for applications such as Web shopping. Several commercial products are available to support color management on the web, using either client- or server-based techniques that require the end user to perform a simple monitor calibration.

Device characterization—e.g., in the form of ICC profiles—is only one component of color reproduction, the other important component being the viewing conditions. In homes and wireless applications, the viewing conditions can be expected to deviate significantly from those normally encountered in a controlled graphic arts studio or even an office. Moreover, the user cannot be expected to understand colorimetry and follow calibration procedures. The patent literature describes low-cost devices that can detect the ambient conditions and feed them into a color appearance model. However, although such devices can work in a fixed workstation, it is hard to imagine that a mobile user would carry one around at all times.

Fortunately there is a technological discontinuity. Although CRT manufacturers today claim 100 Cd/m^2 luminance for whites, this is not at the D_{65} white point required by standards such as sRGB. For a typical high end graphic arts CRT display we measured a luminance of 84 Cd/m^2 after calibrating it as close as was possible to D_{65} without removing the cover. After a further system calibration including the display controller and color management system to achieve 6500ºK correlated color temperature, we measured a luminance of 79 Cd/m^2, 30% less than the nominal specification.

In the case of LCD displays for workstations, the luminance is typically 200 Cd/m^2, three times that of a CRT, and the typical contrast ratio for a medium range LCD is 350:1. The LCD displays for portable computers are not yet up to such specifications due to battery limitations.

2.2 Visual Processes and LCD Displays

At a typical indoor workstation illuminance of 500 lux, a CRT with a white luminance of 80 Cd/m^2 is darker than its surround, therefore viewers will not adapt to the display and see it more like an object in their field of view. In the case of an LCD device, the reproduced image is brighter than the surroundings, a viewing condition similar to that of a film slide on an illuminator.

As Evans observed[3] (p. 599), under the illuminator condition color constancy mechanisms in the human visual system (HVS) correct for improper color balance. This means that, under typical viewing conditions, color fidelity is less of an issue for desktop LCD displays than for CRT displays. As Evans further notes[3] (p. 596), we tend to remember colors rather than to look at them closely; for the most part, he notes, careful observation of stimuli is made only by trained observers. Evans concludes that it is seldom necessary to obtain exact color reproduction of a scene to obtain a satisfying picture, although it is necessary that the reproduction shall not violate the principle that the scene could have thus appeared.

We can interpret Evans' consistency principle[3] (p. 600) as what is important is the relation among the colors in a reproduced image, not their absolute colorimetry. A color reproduction system must preserve the integrity of the relation among the colors in the palette. In practice this suggests that two conditions should be met. The first is that a color should not cross a name boundary, the second is that the field of reproductions error vectors of all colors should be divergence-free. The intuition for the divergence condition is that no virtual light source is introduced.

3. INTELLIGENT DATA COMPRESSION

3.1 Region of Interest

In our Color '97 paper[1] we discussed the JPEG encoding method. JPEG is not especially suitable for transmitting large digital images or image sequences over a slow wireless "last mile" link. A new standard known as JPEG2000[8] has been approved that addresses the issue of achieving acceptable image quality at very low bit rates. JPEG2000 also adds many features that allow users, and in particular Internet users, to interact with the compressed data in ways not supported by JPEG.

When Hunt[6] showed how to count and save bits when encoding color images, each bit was equally important. Extending Hunt's accounting, as is long known[3] (p. 589) when observers look at a scene their eyes are in almost constant motion (EM, eye movements). However, detailed vision occurs only when the eyes hesitate or dwell on (foveate) particular positions in the image—called human regions of interest (hROIs)—and the mind's eye (i.e., the brain) fills in the blanks. The hROIs are also known as areas of fixation and the EMs between ROIs are also known as saccades.

To achieve good image quality at very low bit rates, JPEG 2000 uses wavelets. More important, JPEG 2000 enables encoding ROIs at a higher bit rate than the remainder of the image. In practice, the ROIs will be compressed to a rate similar to JPEG, while the image areas in the periphery or outside the ROI are encoded at a very low bit rate. The number of foveation points is image dependent, so it is difficult to compile simple statistical data, but experience shows that for typical images the compression rate can be increased by an order of magnitude.[13]

Although in the HVS the hROIs are foveated relying on top-down processing, L. Stark and his collaborators[13] have succeeded in developing a set of bottom-up computer vision algorithms that can identify algorithmic regions of interest (aROI). In visual experiments, they have found very good correlation between aROIs and hROIs, thus providing guidelines for the deployment of JPEG 2000 ROI encoding.

These algorithms can also be used for intelligent clipping, dynamically reducing an image to fit a palm-top's small display without the need to scale it to a hard-to-recognize size or to scroll it. All an image server has to do is to compute the aROIs, compute their convex hull, and apply a clipping algorithm.

3.2 Compound Documents

Over the years a number of algorithms has been developed that perform well on particular types of images. Examples are JPEG for pictorial images, and JBIG2, JBIG and MMR for binary images. Unfortunately each algorithm behaves poorly if at all on image types for which they were not designed, which is a problem for compound documents that combine multiple types within a single image. As the digital world is becoming visually richer, such images are becoming much more common.

The Mixed Raster Content (MRC) imaging model has been approved in ITU-T Recommendation T.44.[9] With this model, an image is segmented according to the image types it contains and each type encoded with the most appropriate algorithm. This significantly improves the compression rate on typical compound documents, compared to single algorithms such as JPEG.

4. Content Based Image Retrieval

After the browser, the second most popular World Wide Web tool is the search engine. The prevalence of images has quickly spurred the need for locating them. The first image search engines searched the text in alternate image fields and around images, assuming them to be captions. In reality this text is often misleading and such search engines are not very useful.

Even if images would be manually tagged, this semantic information would not be universal, and the Internet is a global medium. For example, the Francophone surrealism movement has created a great number of visual artworks that convey a message based on the context in which an image is viewed. Also, the textual description of an image changes over time. For example, in 1889 the Eiffel Tower was an expressionist exhibition observation tower, in 1912 it was a futuristic object, while today it is a historical monument.

These problems are avoided by content based image retrieval (CBIR) algorithms. The first generation of such algorithms is based on metrics like the color histogram of images and other similar bottom-up approaches. However, the HVS matches images in a top-down process, even at the most rudimentary levels.[11] As a result, current CBIR algorithms are useful only in very restricted domains.[5]

The fallacy of metrics like color histograms is clear when one compares the images from a stock photo agency to those generally found on the Internet. While the former have been carefully rendered to a normalized intent, the latter are most often the raw output from digital cameras or scanners. A first step is to clearly specify for each image whether it is rendered for a particular output medium or unrendered. A family of large-gamut rendered/unrendered color encoding specifications has recently been proposed under the designation RIMM/ROMM RGB.[14] A second step is to develop bottom-up algorithms that can perform a canonical rendering operation, a process variously referred to as automatic enhancement[2] or intelligent enhancement.[12]

However, ultimately a useful CBIR algorithm must be based on top-down processes, because in humans image comparison is based on cognition. In addition, Internet imaging is different from conventional imaging in that the main issue is scalability; a methodology to support bulk indexing of large image collections still eludes the research community. Another peculiarity of Internet imaging is that client and server communicate according to standardized protocols. Today's CBIR implementations are closed systems, which means that they are specific to particular applications and cannot be universally deployed. Although a protocol has been proposed (see http://mrml.net/ for a description), it has not yet been widely recognized by the research community.

Finally, scientific progress is facilitated by the availability of performance metrics, which allow the critical comparison of different approaches. Establishing benchmarks for retrieval algorithms is a monumental task. The specific issues for image retrieval have been discussed in a recent paper on CBIR performance analysis.[5] Currently an international effort is underway for creating a benchmark (see http://www.benchathlon.net/ for more information).

REFERENCES

1. G. Beretta, V. Bhaskaran, K. Konstantinides, and B.K. Natarajan, "Encoding Color Images for the World Wide Web," AIC Color '97, *Proc. 8th International Congress of the AIC,* Vol. II, pp. 666–669, 25–30 May 1997, Kyoto, Japan
2. R. Eschbach, "Printing Images from the WEB", *Proc. Internet Imaging II,* SPIE Vol. 4311, pp. 1–6, 24–26 January 2001, San Jose, California
3. R.M. Evans, "Visual Processes and Color Photography", *JOSA,* Vol. 33, 11, pp. 579–614, November 1943
4. J.E. Gibson and M.D.Fairchild, *Colorimetric Characterization of Three Computer Displays (LCD and CRT),* Munsell Color Science Laboratory Technical Report, January 2000
5. N.J. Gunther and G.B. Beretta, "A Benchmark for Image Retrieval using Distributed Systems Over the Internet: BIRDS-I," *Proc. Internet Imaging II,* SPIE Vol. 4311, pp. 252–267, 24–26 January 2001, San Jose, California
6. R.W.G. Hunt, "Bits, Bytes, and Square Meals in Digital Imaging," *Proc. CIC 1997,* pp. 1–5, 17–20 November 1997, Scottsdale, Arizona
7. R.W.G. Hunt, "How to Shop on the Web Without Seeing Red," *Proc. CIC 2000,* pp. 2–7, 7–10 November 2000, Scottsdale, Arizona
8. JPEG2000: see http://www.jpeg.org/JPEG2000.htm for the latest information
9. ITU-T Recommendation T.44 — Mixed Raster Content (MRC), 1 November 1995
10. Y. Okano and N. Shiotani, "ICC Profiles Compensating the Color Reproductions of Liquid Crystal Display Panel," *Proc. IS&T PICS Conference,* pp. 403–406, 22–25 April 2001, Montréal, Quebec
11. A. Pascual-Leone and V. Walsh, "Fast Backprojections from the Motion to the Primary Visual Area Necessary for Visual Awareness," *Science,* Vol. 292, pp. 510–512, 20 April 2001
12. D. Pfeiffer, "Intellihance Client Side and Server Side Architectures for Photo Site Image Enhancement," *Proc. Internet Imaging II,* SPIE Vol. 4311, pp. 44–52, 24–26 January 2001, San Jose, California
13. C.M. Privitera, L.W. Stark, Y.F. Ho, A. Weinberger, M. Azzariti, and K. Siminou, "Vision Theory Guiding Web Communications," *Proc. Internet Imaging II,* SPIE Vol. 4311, pp. 53–61, 24–26 January 2001, San Jose, California
14. K.E. Spaulding, G.J. Woolfe, and E.J. Giorgianni, "Reference Input/Output Medium Metric RGB Color Encodings (RIMM/ROMM RGB)," *Proc. IS&T PICS Conference,* pp. 155–163, 26–29 March 2000, Portland, Oregon

---

* Correspondence: beretta@hpl.hp.com; phone 650-857-6713; http://www.inventoland.net/; Hewlett-Packard, 1501 Page Mill Road, 4U-6, Palo Alto, CA 94304, USA

** Correspondence: RBuckley@crt.xerox.com; phone 716-422-1282; fax 716-265-8871; Xerox Corporate Research & Technology, 800 Phillips Road, 0128-30E, Webster, NY 14580, USA