Shutterfly's Printing Infrastructure

Dhiraj Kacker and Russell Muzzolini
Imaging Science
Shutterfly,Inc.
Redwood City, California

Abstract

Shutterfly is a digital photofinishing company providing Silver-Halide photographic prints, greeting cards and other products from digital images.* Shutterfly's printing infrastructure is comprised of multiple output devices with various different printing technologies. In this paper we describe the image processing and color management components of Shutterfly's printing infrastructure addressing issues related to hardcopy reproduction of consumer digital images. We also describe the architecture of the central image processing engine with emphasis on metrics of scalability, throughput, and cost.

1. Introduction

As digital cameras penetrate into the mass consumer market there is an ever increasing need to make prints from digital images. The choices for consumers are wide and varied. They include printing at home on inkjet photoprinters, printing at retail locations, and using one of many web based photofinishing companies for getting prints. Shutterfly is one such web-based photofinishing company. In this paper we describe the various components of the printing infrastructure in place at Shutterfly with key emphasis on the imaging processing aspects of the infrastructure.

A typical Shutterfly customer uploads his/her digital pictures to their account through Shutterfly's website. The website provides capabilities for organizing and sharing pictures in addition to ordering a number of different digital imaging products. These products include different sizes of silver-halide prints, customized photo greeting cards, customized calendars, and CD's of digital pictures. The silver-halide prints are printed on one of many digital minilabs while the customized photo greeting cards and calendars are printed on the Indigo UltraStream digital offset press. A key aspect of the ordering workflow is that it is entirely automated. There is no human intervention until the final printed product comes off one of the printers. This fundamental requirement for automation imposes a number of secondary requirements on the whole printing process. In particular, all image enhancements have to be done in a completely automated fashion and all printers, irrespective of their native printing technology, have to produce nearly identical colors at all times. This is quite different from the traditional film-based photofinishing world where the operator, at either the retail location or in a centralized printing facility, has the ability to make color adjustments to improve the quality of the prints.

From the point of view of the customer there are some significant differences in the digital world relative to the film-based world. With digital images customers have multiple opportunities to preview their images before requesting prints. These images are often viewed on uncalibrated output devices such as the LCD displays on cameras and on home monitors. This sets expectations which are often difficult to convey to the order fulfiller. In the film world the only opportunity to 'look' at the pictures is when the final prints come out.

In the next few sections we describe the architectural and process design choices relative to the aforementioned issues to build Shutterfly's digital photofinishing facility. In making these choices a very conscious attempt was made to build on well understood color imaging concepts. The key was to integrate these concepts in building a large, scalable, fully automated digital printing facility. It was, however, the actual integration of these pieces and scale of the operations that spawned the creation of new technologies and processes. We begin our discussion by describing the logical components of the imaging architecture in the next section. In Sec. 3 we describe the physical manifestation of this architecture. Finally, in Sec. 4 we close with some remarks.

2. ImagingWorkflow

It is well known in the photographic industry that an average consumer desires good reproduction of their pictures, not necessarily a picture that is true to the original scene in terms of color, contrast, brightness etc. In the traditional .lm world there are two points at which images are enhanced: in the .lm and during printing. In the digital world the "film enhancements" are replaced by the image processing in the firmware of a the capture device. Given the computational resource constraints on a typical consumer camera most cameras are limited in what they can do. In consumer studies conducted at Shutterfly, we found that average consumers overwhelming preferred images that had some post-processing performed on them. From a fulfillment point-of-view, we therefore have two distinct image processing components to be concerned about: automatic enhancements and physical reproduction. We call the Shutterfly's technology developed around the former VividPics™ and around the latter ColorSure™. These two logical components of the imaging workflow are shown in Fig. 1. Currently Shutterfly accepts JPG RGB images and assumes the input space to be sRGB.[1] As a first step, the images are converted to floating point YUV space and all VividPics™ image enhancements are performed in this device independent color space. Just before printing, the images are converted to device dependent (RGB for silver-halide printers and CMYK for the offset press) color space for accurate color reproduction. The final goal of this whole process is to provide pleasing prints to customers. To this end we have a mechanism for incorporating customer feedback received through various channels, including customer service, into enhancements to VividPics™. In the next few sections we discuss these components of the imaging workflow.

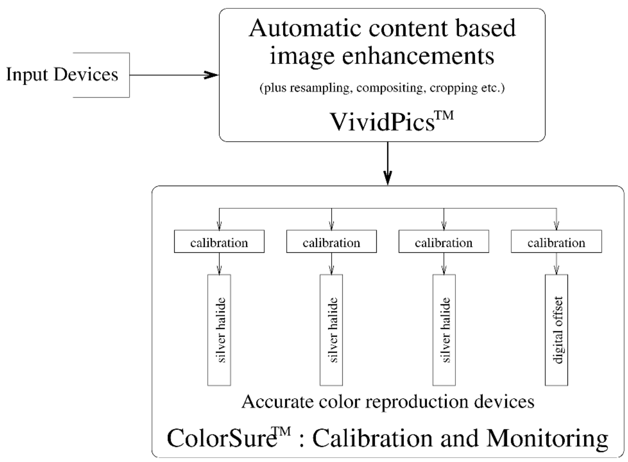

Figure 1. Conceptual imaging workflow

2.1. VividPics™

Every month we print millions of digital images that come from a wide variety of capture devices including digital cameras, film scanners, and flatbed scanners. The images are often edited using a wide variety of software packages. The images also come in a wide variety of resolutions. Given the large quantity of images processed, a key requirement in developing VividPics™ is to have a very low failure rate. In particular it is extremely important that the algorithms make very little change to images that already have good color balance, lighting etc. As VividPics™ is applied in a completely automated fashion, it limits the aggressiveness of the underlying algorithms; aggressive algorithms may greatly improve some images, but will in all probability have unacceptable failure rates. It is also non-trivial, and probably undesirable from a usability point-of-view, to provide multiple previews on the website. Most consumers just want a simple way to get prints from their images, most do not have calibrated monitors, and the computational requirements and available bandwidth are prohibitive for creating such previews in real-time. Given the wide variety of sources and scene types it was essential to develop VividPics™ as an image adaptive algorithm. The key image measurements used as inputs into this adaptive algorithm are colorspace location of image pixels relative to the final printer's gamut,[2] measures of luminance values in the image, and image metadata information. The operations performed on the images include brightness adjustment, contrast adjustment, and color balancing. The extent and aggressiveness of these operations is determined by the aforementioned image measurements. Given an ideal output device (one that can produce all colors with infinite color resolution) all image adjustments can be performed without any knowledge of the final printing device. However, as real printers have finite color gamuts and various printing artifacts such as contouring in highlights in printers using halftoning, the enhancements have to be cognizant of these while making adjustments to the image. In addition to all image enhancements done as part of VividPics™, we detect if an image either has low resolution, is underexposed, or has a non sRGB color space. If an image with any one of these is detected, we add icons on the header prints and insert information about the problem in the order. This helps us proactively set expectations.

2.2. ColorSure™

ColorSureTM is the second step of the imaging workflow. The goal of ColorSure™ is simple: ensure that all printing devices, irrespective of native printing technology, provide colorimetric reproduction at all times. In order to achieve this we use both a calibration and a monitoring process. To calibrate themachines we build ICC profiles[3] using Gretag's ProfileMaker software; RGB profiles are built for the silver halide machines and CMYK profiles are built for the offset presses. In order to monitor the printers we print MacBeth color charts that are automatically scheduled throughout the day on every printer. The charts are scanned on a spectrophotometer by operators to measure the color accuracy of the reproduction. The goal is to maintain all machines within 2.0-2.5 ΔE_{94}[4] of the target colorimetric values. If the scan produces colors greater than the deviations desired, we have production process in place to correct for the problem.

2.3. Production Benefits

Breaking up the imaging workflow into these two disjoint components has many advantages. A customer ordering prints from the same image on multiple days and/or different products of the same image, gets almost identical pictures at all times. This greatly reduces the number of complaints received by customer service. In addition, sophisticated customers who create images in a color managed workflow can disable VividPics™ on the website and get accurate color reproduction (see Ref. 5 for an independent evaluation) as all printing devices are managed to be colorimetric. This workflow also has some very significant benefits in operating the printing facility. Development and enhancements to either component of the imaging pipeline can take place without having an impact on the other. For example, if we learn something new about customer preferences or enhancements to some class of images, changes to VividPics™ will have an immediate effect on all products, irrespective of which printer type they are printed on. Due to the colorimetric reproduction of all printers, these changes will be accurately reflected on all images. In addition, as color accuracy can be mapped to a single _E number, we have been able to train all machine operators to monitor and keep printers within calibration. We have thus removed the need to make subjective evaluations which require additional training and are therefore less cost effective. Finally, this architecture also enables us to quickly add to our printing capacity without requiring any custom color related work on new printers.

3. Physical Architecture

The typical approach to processing images employed by most printer manufacturers is to provide software and hardware in addition to the printer and its peripheral control system. This "rendering" subsystem incurs extra cost and often becomes a bottleneck when trying to maximize the throughput of the printer. Modifications to the processing are almost impossible as the system is inevitably proprietary. Further, varied choices for hardware platform, OS and other software increase operational costs.

We realized that the above de.ciencies needed to be overcome for our printing infrastructure to be capable of high print throughput with different printing devices at a low cost. A centralized rendering system called the "Renderfarm" was implemented which meets the goals of throughput, scalability and cost effectiveness.

The Renderfarm architecture is illustrated in Figure 2. Thick lines represent image data flow while thin lines represent control data flow. The Renderfarm physically consists of a number of Rendernodes (servers) and their interconnect network, additional shared storage, as well as load balancing and monitoring components. All render requests are sent to the Renderfarm using the standard HTTP protocol and an XML interface and subsequently routed to an appropriate Rendernode via the Load Balancing component. Both simplistic (i.e. hardware load balancing) or priority and load based (software) policies are available. The Monitoring component provides status information and configuration capabilities for all the nodes in the Renderfarm. A render request contains the complete description for rendering one or more output images. This includes URLs for all source input images and a compositing script describing output format, processing effects, affine transformations, ICC Profile selection and instructions for compositing text, art borders and other images into the final output image.

An image is transformed from its initial source input(s) to its final printer dependent color space, resolution and file format in two "render passes". The first pass is the Device Independent pass in which all VividPics™ processing is performed. Since this pass is independent of any output speci.c device it can be performed well before (i.e. when the customer order is received by the Scheduler component) the images are required for printing. This significantly reduces the computational requirements for the nodes rendering this pass provided the render requests are scheduled appropriately in advance by the Scheduler component.

The second pass is the Device Dependent render pass in which all ColorSure™ processing is performed. ICC profiles are created for each printer by the ICC Profile Builder and sent to the Renderfarm. Thus, the Renderfarm keeps track of the reproduction characteristics of each printer at all times. The printer Linecontroller issues Device Dependent requests to the Renderfarm just prior to printing on its associated printer. The Renderfarm subsequently writes the rendered output images to the printer's local storage at which point printing is possible.

The Renderfarm achieves cost effectiveness with the use of commodity hardware. Each Rendernode is a headless, rackable, computer with (currently) a 1G Intel CPU, 1G memory, 40GB disk running Solaris 8 and costs $1000 each. Additional nodes are are purchased as needed to take advantage of the price/performance benefit with commodity hardware. This strategy reduces the performance requirements of the image processing software in order to obtain high throughput. Rather than writing specially tuned versions of the software that result in lengthy development with higher cost additional nodes are purchased. In this system which produces hundreds of thousands of images a day coarse grain parallelism at the image (or batch of images) level is adequate to achieve the desired throughput.

Scaling the Renderfarm by simply adding Rendernodes does imply that any shared resources such as network bandwidth and storage do not become bottlenecks. Storage is solved by simply using the local disk drive on each Rendernode (typically 40 GB or more on today's commodity hardware). The bandwidth issue is resolved in two ways; utilize less bandwidth and maximize the available bandwidth. The Renderfarm minimizes bandwidth by reducing output file size using standard JPG compression where possible. Further bandwidth saving are gained by caching output images and minimizing network reads/writes with the use of local storage on each Rendernode. Available bandwidth is maximized by using interconnects on the network which support 1GBit interfaces (although each Rendernode and printer only require a 100MBit interface to the network).

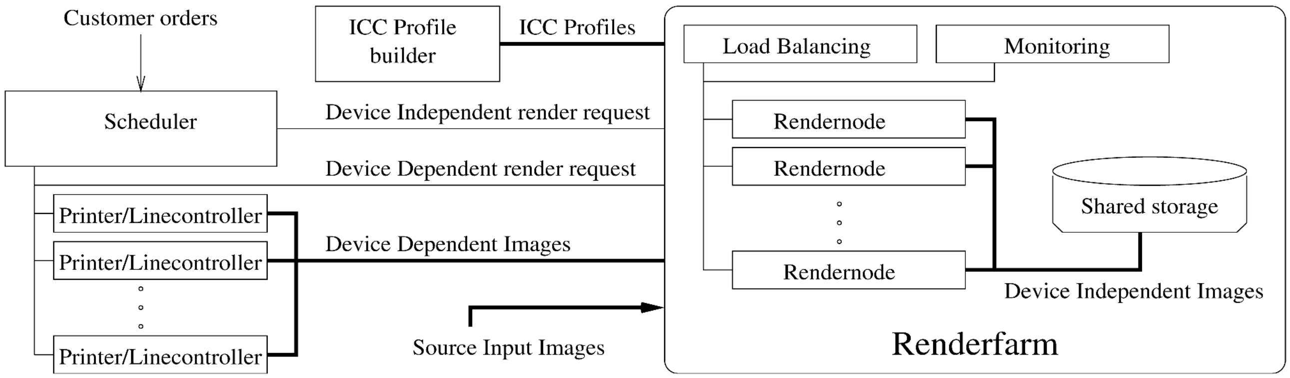

Figure 2. Imaging Architecture

4. Conclusion

In this paper we have described the printing infrastructure in place at Shutterfly's print facility. We have elucidated the key architectural components of the imaging workflow as well as the physical manifestation of the workflow. A number of commonly understood and well researched color imaging concepts such device independent and device dependent color processing, colorimetric reproduction, and image enhancements form the theme of this paper. One of the key innovations at Shutterfly is the integration of these concepts in building a scalable digital printing facility that has a very competitive cost structure associated with it. As much as possible we have integrated publicly available off the shelf software and algorithms. We have, however, developed key components of the printing infrastructure internally. These include the algorithms for automatic image enhancements that have good performance measures on an extremely large and varied set of images; production processes around colorimetric reproduction enabling a large number of line operators to keep printers in calibration; and finally creating an architecture for doing all the image processing that scales easily with increase in volume.

5. Acknowledgement

A number of people have contributed to the work reported in this paper. They include Greg Ward, Henry Li, Sumit Chawla, Jeff Boone, Jeannine Smith, Xin Wen, and Danny Loh.

References

* Certain aspects of the present paper have been disclosed in pending US patent applications.

1. IEC Standard 61966-2-1, “Multimedia systems and equipment - Colour measurement and management - part 2-1: Colour management - Default RGB colour space - sRGB,” 1999.
2. G. J. Braun and M.D. Fairchild, “Image lightness rescaling using sigmoidal contrast enhancement functions,” in *Color Imaging: Device Independent Color, Color Hardcopy, and Graphic Arts IV, Proc. SPIE 3648*, 1999, pp. 96–107.
3. *File Formats for Color Profiles (Version 4.0.0). Specification ICC.1:2001-12*, International Color Consortium.
4. CIE Publication 116, *Industrial Colour-Difference Evaluation (Vienna: CIE 1995)*.
5. *MacWorld*, Mac Publishing, LLC, Jun. 2002.