Pervasive Digital Color Reprographics

Joseph S Czyszczewski *, Hong Li*, James T Smith *

Keywords: Reproduction, Digital, Color

Abstract: Work group digital color scanners are becoming affordable to the office environment, while color printers are already pervasive. There is an opportunity for a color reprographics solution that leverages these devices, avoiding the cost of dedicated color copiers. However, achieving color copier quality requires both color management techniques that are too complex for average users and image processing that is not generally available. Czyszczewski, Li and Smith will describe an alternative color management technique that delivers quality and consistency with superior simplicity.

Introduction

Digital color scanners and printers become pervasive in the office environment when they are affordable. Wide availability of these devices creates opportunity for new color applications. One common application is to leverage scanners and printers as a reprographics solution. A modular solution based on existing devices for casual copy requirements is compelling because it avoids the cost of new dedicated color copiers. However, achieving acceptable quality and performance in a distributed heterogeneous solution requires color management techniques which are too complex for office users. This open approach also requires automated image processing which is not generally available in the office environment.

Modular Solution

A modular color reprographics system is defined which integrates scan stations with printers where both are distributed, heterogeneous, and network attached. The scan station consists of a scanner connected to a personal computer. The personal computer includes a user interface, network interface, and copier solution software. The software provides a user interface, color management,

* IBM Printing Systems Division

image processing, and output job management. The user interface is graphical and resembles a traditional copier control panel to support general walk up office users. The simplicity of copier like user interface is required to avoid the need for prepress skills. The personal computer can be a dedicated end user machine or a shared work group machine which is hidden and secured.

The operation of the system starts with a color original placed on the scanner, a pre-configured local or remote color printer and copy options are selected. The document is scanned through image processing pipeline to optimize first page out time, converted to a printer format, spooled, and sent to the selected network attached printers as standard image print jobs. In addition to making color hard copies, this process can also be configured to distribute electronic copies as e-mail or via shared folders. Also, one scanned job can be simultaneously distributed to multiple destinations and multiple scan stations can target the same printer.

As a modular system, multiple and different types of scanners and printers must be logically integrated into an operationally seamless system. One key challenge in deploying such an open heterogeneous color copy solution is that different color scanner and printer models typically have significantly different color imaging characteristics. Also, alternate distribution options such as e-mail for viewing introduce yet another class of characteristics. To avoid handling $m \times n$ pairs of characteristics, our strategy is to individually characterize each scanner and output device. This characterization must be done with simple customization to enable on site configuration. Based on the configuration information, the system automatically and dynamically connects image processing and color management components based on the selected devices and job options.

Performance is another key challenge in such a system because general purpose devices and networks are not optimized to make color copies. The risk is that general office users will not be satisfied with a copier which has poor performance even if the quality is competitive. There are several performance considerations in the design of an open solution. One is to optimize first page out time. This requires a streaming architecture where the leading edge of the output page is sent to the printer before the trailing edge of the original page is scanned. A second consideration is throughput which is a problem due to the large file sizes and the bandwidth limitations of general scanners, printers, and networks. This requires a trade off between image quality and file size with compression and resolution tuning. Another option is to perform device binding in the personal computer for a reduction in print file size and processing requirements. A third consideration is the ability for an end user to interrupt or cancel a job that has been sent to the printer. In some cases, standard print jobs need to be interrupted when a walk up user needs to make a quick copy. Another case occurs when a user makes an error and need to cancel the current copy job.

These capabilities are dependent on the level of control provided by the printer for job management.

Challenges

The scanners and printers are the primary devices in the modular copier system. Recognizing and adapting to the different technologies used in each is essential to optimize end to end copy quality and performance. The overall strategy is to include algorithms which deal with each significant technology variation within a reasonable range of devices and to dynamically configure the ones required for each pair of devices selected by the end user when a copy request is made.

While some of the characteristics are best determined based on the device specifications and managed as configuration options, others are best detected with test targets. As a result, an automated process is defined for the set of recognized technology variations and the end user assists in determining the final configuration for each of the scanners and printers that must be supported by the solution. Typically, the test target is scanned or printed and scanned to complete the configuration process.

Secondary components such as the scan station personal computer and the office network are also considered, but are not as significant. Rapid improvements in affordable processors eliminates the scan station processor as a primary consideration for midrange office copies. Network bandwidth limitations are handled with a private network for the primary printer associated with the scan station for copies. Since remote copies are less common, their impact is comparable to standard print traffic and do not have a significant impact on the design.

The shared office environment requires the use of work group scanners which operate at much higher speeds than personal desktop scanners. In some cases, these higher speed scanners introduce a color fringe problem which is a significant issue when making copies of black text. This problem is amplified when printed on a printer that is not properly aligned. Another problem is "show through" of the back side image on a two sided document. A third problem is that some scanners with dual optical systems use different technologies on the front and back sides to reduce cost based on the fact that most document scanned are one sided.

In addition to image quality variations, performance variation must be factored into the design. For example, some scanner drivers support streaming while others are page oriented. Another consideration is the scanner communication bandwidth and higher speed attachment options may be required. A third performance consideration is the impact that resolution and bit depth have on the scanner throughput. For example, scanning at 300 dpi significantly increases scanner throughput while delivering sharp text and high quality images.

Different technologies must also be supported for the printers. Examples are toner, ink jet, or dry ink and different halftones are used with each of these different printing technologies. For example, ink jet and dry ink color printers generally use error diffusion or stochastic screening while color toner printers rely on cluster halftone or line screen techniques and different frequencies are supported. Another consideration is the use of bilevel or multilevel halftone technologies.

A third printer variation exists in printer color conversion with PostScript color management model where color conversions are done with a color space dictionary and a Color Rendering Dictionary (CRD). A CRD is developed for each printer by its manufacturer and different techniques are applied to different types of originals based on the limitations of the print engine and different techniques are used to convert CMY to CMYK. As a result, when equal amounts of RGB gray are sent to a variety of printers, the output color ranges from single black gray to process black with a color cast based on the CMY to CMYK conversion.

Color Management

Color management is more challenging in a modular multifunction color copier than in a traditional closed monolithic copier because interchange is required. In a traditional copier, the scanner is color corrected, followed by the printer, and finally additional color adjustments are made for poor originals and customer preferences.

However, in the case of a modular copier with a customer defined heterogeneous mix of input and output devices, each combination of scanner and printer pairs must be supported. Also, color corrected scan data may be simultaneously sent to different network printers or via e-mail to personal computers that have significantly different visual characteristics. Color interchange among different input and output devices is needed.

A device independent CIE-based color space is desired to support color interchange. CIE-based color space is defined relative to an international standard used in the graphic arts, television, and printing industries and it enables a page description to specify color values in a way that is related to human visual perception. The goal of this standard is for a given CIE-based color to produce consistent results on different output devices, within the limitations of each device.

The International Color Consortium (ICC) specification recognizes CIELAB and CIEXYZ color spaces as profile connection spaces. Although CIE-based color specifications are theoretically device independent, they are subject to the practical limitations in the color reproduction capabilities of the output device. Such limitations may require compromises to be made among various properties of a color specification when rendering colors for a given device. The ICC specification recognizes four rendering intents for different types of originals. The intent is to treat different image content optimally within the limitations of the output device.

The typical color management workflow to make copies is as follows: raw RGB scanned data is converted to CIE-based data through scanner color correction and CIE-based data is converted to printer CMYK data through printer color correction. End user color adjustments are done in a CIE-based space.

There are many color correction techniques developed for both scanners and printers. Scanners can be color corrected through a matrix or LUT and printers can be color corrected through an LUT. Since a CIE-based space is a device independent color space, CIE-based color can be sent to different output devices, such as networked printers and color displays, along with output device color correction to achieve consistent output quality.

To render device independent color on a PostScript printer, the PostScript interpreter converts from a specified CIE-based color space to the printer color space, taking into account the properties of the printer. Conversion from a CIE-based color into a printer color requires gamut mapping and development of a corresponding printer device color. PostScript implementations of such conversions are embodied in a CRD.

Calibrated RGB space is used as a color interchange in our system to optimize performance. Scanned data is color corrected with a simple 3x3 matrix to calibrated RGB space, the data is then sent to networked printers or displays. In the case of output to a PostScript color printer, two additional steps are taken to convert RGB color to printer CMYK color. In the first step, calibrated RGB color is transformed to CIEXYZ color through a PostScript color space dictionary. In the second step, CIEXYZ color is transformed to device CMYK color through a CRD. Compared to operating in the CIELAB space, color conversion is much faster in the calibrated RGB space. This is especially true for color display output since display devices use device RGB space and no color conversion is necessary assuming a calibrated RGB space for the display. End user color adjustments are also implemented in the calibrated RGB space. However, calibrated RGB is not a visually uniform space, and color adjustments implemented in RGB space are not as visually pleasing as in the CIELAB space.

Because the modular copier is intended for an office environment, this is a case where performance is optimized over color adjustment quality.

Reprographics Optimization

As a modular copier, quality reproduction requires a wide range of image processing algorithms. As mentioned above, scanners and printers have different characteristics and there are vastly different types of originals that need to be supported. For high quality reproduction, we need sharp single black text, without color fringe or moiré, and vivid color images without color shift. End users expect another "original" and, sometimes, a copy that is better than the original based on correction for poor quality in the original. In order to achieve this level of quality, we need to compensate for scanner and printer characteristics while correcting for original and printer halftone interactions.

In the case of a modular copier with a heterogeneous mix of input and output devices, the system need to take into consideration both quality and performance and a dynamically optimized system with selective pipeline image processing algorithms is required. Our system is designed to automatically manage the complexity of *m x n* device combinations with the simplicity of a copier like user interface. In such a system, we first detect the scanner characteristics. Then we detect the printer characteristics. Finally, the copy path workflow is dynamically configured based on the known scanner and printer characteristics to optimize quality and performance.

<u>First step: scanner characteristics detection</u>
During system configuration, the end user presses a "scanner characteristics detection" button and a standard IT8 target is scanned. This results in RGB three channel data being obtained for gray patches.

Gray balance is detected based on comparing RGB data in the center of each gray patch and if the difference exceeds a threshold, the scanner is not gray balanced. An end user message indicates that the scanner is not gray balanced and a gray balance correction curve is developed and stored in the system.

Color fringe is detected by evaluating RGB data in the outline area of each gray patch and if the difference exceeds a threshold the scanner has a color fringe problem. An end user message indicates that the scanner has color fringe problem and color fringe correction parameters are developed and stored in the system.

If no other scanner problems are detected, a simple 3x3 matrix color correction is developed using a regression method based on the scanned IT8 target data and the corrections developed above.

Second step: printer characteristics detection

During system configuration, the end users presses a "printer characteristic detection " button and two digital targets are printed. One target has a single step wedge with 16 steps of equal RGB values, the other has a four wedges representing C, M, Y, K and each with 16 steps.

Halftones are detected by scanning the CMYK step wedge and searching for a halftone frequency. If the printer uses a halftone, an end user message indicates the frequency of the halftone and it is stored in the system.

Gray balance is detected by scanning the RGB step wedge and comparing the three RGB values. If the difference exceeds a threshold, an end user message indicates that the printer is not gray balanced.

Third step: copy workflow configuration

When an end user makes a copy, the original is scanned, the scanned data goes through the required scanner characteristic correction filters in the pipeline and the 3x3 matrix color correction. The next step depends on the printer characteristics.

In the simple case where a printer is gray balanced and uses error diffusion, a complex original, with both text and halftone image, the scanner corrected RGB data is processed through background suppression, edge enhancement, color adjustment, and the compressed RGB data is sent directly to the printer. The PostScript interpreter takes care of RGB to CMYK conversion and the reproduction has sharp black text and moiré free vivid images. This approach also applies to high quality laser printers where the halftone frequency is above 220 lpi and multiple bits are supported. While there is some risk of a moiré problem, it is not significant enough to justify descreening.

In the more common and complex case, office color laser printers are used which have a halftone frequency below 170lpi. Some office color laser printers have poor gray balance resulting in equal RGB gray is printed with a color cast. Finally, some office color laser printers print equal RGB data using process black, prone to color fringe for scanned black text to be sharp and true black.

For these types of printers, the text and halftone image data needs to be segmented to provide acceptable copy quality. After segmentation, black text is sent to printer as compressed bilevel black and white data. This optimization also improves performance because binary data replaces gray scale data. Halftone images are descreened, color adjusted, and compressed RGB data is sent to the printer. Descreen filters are optimized to target the frequency of the printer halftone detected during configuration to minimize moiré while also minimizing copy degradation

Conclusion

This optimized open modular color copier solution delivers reprographics quality and consistency with simplicy. By leveraging existing office color scanners and printers, it improves color reprographics investment models. It also provides flexible document distribution and deployment options.

Acknowledgments

The authors thank our colleagues Ravi Rao ,Gerry Thompson and Chai Wah Wu in IBM Research at Yorktown Heights, New York for their insight and image processing algorithms. We also thank Israel Berger, Eugene Walach, Asaf Tzadok and Dan Chevion in IBM research at Haifa, Israel for their help and text segmentation algorithms.