



A Hardcopy Backup and Reconstruction System for Digital Images

Eric Hoarau, Ingeborg Tastl, Nathan Moroney

HP Laboratories
HPL-2009-199

Keyword(s):

image reconstruction, auxiliary information, color transformation, color correction, fading

Abstract:

Large amounts of images are captured every day and digital storage solutions are the most prevalent approach to preserve them. This paper proposes to augment the existing solutions with a modernized hardcopy backup, which can serve as a fall-back solution if everything else fails and can be enjoyed at the same time. In particular, the collection of images is printed in form of an inexpensive photo-album, which is augmented with printed human readable meta-information designed to enable an easy reconstruction of the digital version of the original images. The methodology is illustrated using a set of digital sRGB images and printed simulations of faded versions of those images, which are scanned and reconstructed in an automatic and manual way and are compared to the results of a scanner reconstruction software and to the original images.

External Posting Date: August 21, 2009 [Fulltext]

Approved for External Publication

Internal Posting Date: August 21, 2009 [Fulltext]



To be presented at 2009 IEEE International Conference on Image Processing - Cairo, Egypt. November 7-11, 2009

© Copyright 2009 IEEE International Conference on Image Processing 2009

A HARDCOPY BACKUP AND RECONSTRUCTION SYSTEM FOR DIGITAL IMAGES

Eric Hoarau, Ingeborg Tastl, and Nathan Moroney

HP Laboratories, Palo Alto, CA, USA

Email: {Eric.Hoarau, Ingeborg.Tastl, Nathan.Moroney}@hp.com

ABSTRACT

Large amounts of images are captured every day and digital storage solutions are the most prevalent approach to preserve them. This paper proposes to augment the existing solutions with a modernized hardcopy backup, which can serve as a fall-back solution if everything else fails and can be enjoyed at the same time. In particular, the collection of images is printed in form of an inexpensive photo-album, which is augmented with printed human readable meta-information designed to enable an easy reconstruction of the digital version of the original images. The methodology is illustrated using a set of digital sRGB images and printed simulations of faded versions of those images, which are scanned and reconstructed in an automatic and manual way and are compared to the results of a scanner reconstruction software and to the original images.

Index Terms— image reconstruction, auxiliary information, color transformation, color correction, fading.

1. INTRODUCTION

With the advent of digital imaging, users have a multitude of options to store their images. The International Imaging Industry Association (I3A) has created a site which details the options and limitations of current technologies [1]. Digital storage solutions such as hard drives, CDs and online storage offer an exact copy of the original, but are vulnerable to hardware failure, system obsolescence, or simply user errors. Analog solutions such as single prints and photo albums are easily viewed and are resilient to damage, but they are a reduced representation of the original data.

The early research focused on professionally restoring faded Silver Halide images and old movies [2][3][4]. The steps involved either a skilled operator manually restoring individual images or building inversion models for the average fading of a specific media, time and storage conditions. The accuracy of these solutions depends on the correlation of the actual fading with the model and on how well the parameters are being estimated. Significant bodies of work look at the issue of color appearance changes of digital prints over time [5], but few publications look into anticipating an easy future image reconstruction at the time of printing [6]. Samadani proposes to encode information in printed media as side information and achieves promising results; however we are not aware of any solution that focuses on augmenting the printed image with visual features to track appearance changes and facilitate a reconstruction years or decades later. In this paper, we describe a simple yet robust hardcopy backup and reconstruction system for digital images which is not based on specialized ink and paper. We introduce an automated document creation layout algorithm, manual and automated reconstruction methods and present our experimental setup and results.

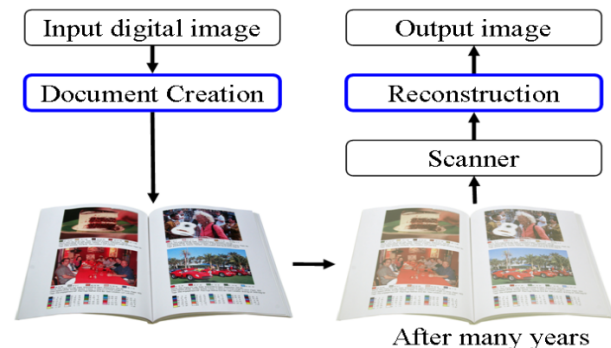


Figure 1. Reconstruction Framework.

2. SYSTEM

Our solution enables a user to automatically store their favorite digital images on a printed page or album that also contains human readable metadata and color tracking features. The resulting album can be enjoyed for years to come and in the event of loss of the original digital file used to reconstruct an approximate digital file as shown in Figure 1.

Printing the auxiliary information in a human readable format ensure that it can be used by current and future reconstruction methods. The embedded color patches track the modifications that are happening to the image during its life cycle. These include the perceptual color transformations at the printing and scanning stage, as well as the appearance changes due to light, temperature, humidity and pollutants. Reconstructing the original images simply means inverting those changes by using the original numeric color values printed next to the patches. Treating changes as black box makes our method suitable for any printing technology and it is independent on whether the actual changes fit a particular model or not. This robustness makes it well suited for real world applications.

3. DOCUMENT CREATION

We developed a web application to easily layout the images from an archive into a PDF formatted document. The tracking features and key metadata information are automatically embedded alongside the images. For speed, the core layout algorithm was developed in C++ coding language and utilizes various imaging libraries. The application GUI also enables the user to optionally modify or add to the metadata included in the final document. The resulting digital file can be printed and collected into various finishing formats such as folders, binders, and bound documents.

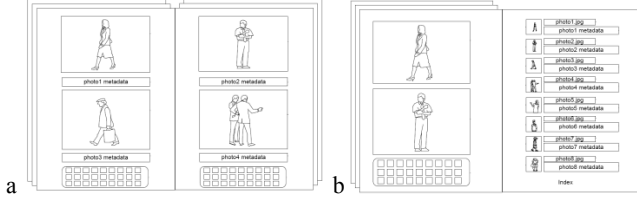


Figure 2. a) standard layout, b) index based layout.

The layout options for the document are endless and depend on the intended use. We present two possible layouts in Figure 2. The first layout has an archival focus and the images are printed using a 4x6" format with white space around them for both image protection and easy scanning. Larger images could be printed if desired, but this form factor maximizes the number of images and minimizes production cost by using standard paper. The resulting scanned image will be ~2MP. There is a wealth of information in the EXIF file of an image that is generally discarded during the printing stage. Such information might not seem valuable at this time, but in the future, might prove to be invaluable for automatic image recognition and understanding. In this layout, the image metadata information is printed below the image using a standard font such that it can easily be read or interpreted by OCR algorithms. It includes image-specific information such as file name, date, tags and GPS location, and camera information such as shutter speed, f-number, focal length, flash, and color space.

The second layout is more aesthetically pleasing and fosters everyday enjoyment by moving some of the tracking features and image metadata to an index section as shown in Figure 2b. This also provides an opportunity to include additional information into that section supporting the preferred method of reconstruction and to use a larger color target of for example 125 patches. A larger color target will generally result in more accurate reconstructions as shown in the result section of this paper.

We have developed and tested two color targets as our main tracking tool. The first one consist of 27 color patches evenly sampling the sRGB space (placing 3 samples along the Red, Green and Blue axis in a cube results in $3^3=27$ color patches). The second version has 30 colors patches based on high frequency, non-redundant lexical or named colors. This target is based on previous web-based color naming experimentation and was used as a non-uniform, image independent sampling of colors. The corresponding sRGB values have to be printed for reference and are used as goal values for the reconstruction process.

Finally we have created an optional smaller selection of image specific color patches to be added to the 30 lexical color patches to potentially help in the manual reconstruction steps. There are other solutions for creating image color palettes [7], but our solution leverages the geometry of lexical quantization [8]. Input images are center-weighted and then reduced to an additional color sampling based on minimum area requirements for the product of the corresponding chroma and regions areas. K-means is then used to reduce this initial sampling to a target five color patches per image.

4. RECONSTRUCTION OF SCANNED IMAGES

If a digital version of the hardcopy images is needed, the user can scan the prints and either manually adjust the images with the goal of bringing the color patches back to their reference values, or preferably he can automatically reconstruct the images.

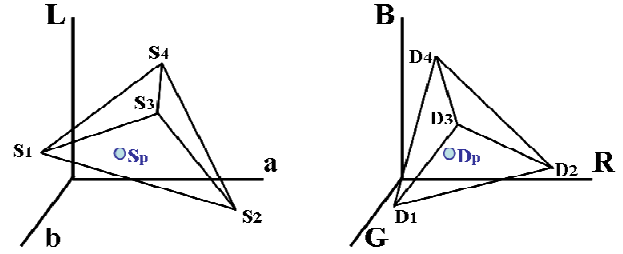


Figure 3. Tetrahedrons in Lab and sRGB space.

4.1. Basic concepts of the automatic reconstruction

The core idea is to use the scanner values (e.g. scanner RGB or Lab) of the faded color patches and the corresponding original recorded sRGB values to specify the color transformation for the reconstruction of the images. One could apply one of the many scattered 3D data interpolation methods described in [9].

One possibility would be a tetrahedral interpolation: As shown in Figure 3, let $\{S_1, S_2, S_3, S_4\}$ be four vertices of a tetrahedron in Lab space and $\{D_1, D_2, D_3, D_4\}$ be the corresponding tetrahedron in sRGB space. A point S_p inside the tetrahedron in the Lab space is transformed into a point D_p in sRGB space using the following formulas.

$$D_p = A_D A_S^{-1} (S_p - S_1) + D_1$$

where A_D and A_S are 3x3 matrices given by

$$A_S = [S_2 - S_1 \quad S_3 - S_1 \quad S_4 - S_1]$$

$$A_D = [D_2 - D_1 \quad D_3 - D_1 \quad D_4 - D_1]$$

All the color values of the faded image are interpolated using the tetrahedrons formed by the values of the color patches. The interesting questions are then in which color space to form the tessellation, what is the minimum amount of color patches necessary to enable a decent reconstruction and how to distribute them in the sRGB color space.

4.2. Our implementation of an automatic reconstruction

We previously came up with the idea to use ICC profiles commonly used for device color characterization to simulate how an image printed on a specific device and media would fade over time [10]. Instead of using original sRGB values of a large number of color patches and the corresponding measured Lab values of the color patches that have been exposed to accelerated fading in a fadometer and building ICC fade simulation profiles we use the original sRGB values of the color patches and the scanned Lab values of the faded patches to build an "ICC reconstruction profile" either for a specific image or for a whole document (in the case that a large chart has been placed at the end of the document). A printed, faded and scanned image/document can simply be reconstructed as visualized in Figure 4.

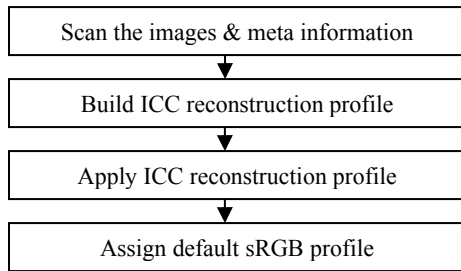


Figure 4. Steps of the automatic image reconstruction.

This process can be performed in Photoshop, inside a Matlab, C program or a scanner software. Instead of using a complete ICC reconstruction profile one could also just apply the color table (a 3D CLUT going from Lab to sRGB) for reconstruction purposes. One other possible implementation would be to offer the reconstruction as a web application, where users simply submit the scanned document and receive back the reconstructed images. The Lab values from the modified scanner pipeline that we used could be replaced by traditional scanner RGB values.

4.3. Manual reconstruction

The user skill level and time commitment will influence the result of the manual reconstruction. This is mainly due to the non linearity of the fading process, and to the wide range of current imaging software packages. A simple reconstruction would adjust the black and white point using the corresponding patches. A more advanced step involves removing a color cast by setting the grey point and adjusting the midtone level. The user can also locally color correct the image by segmenting the image and using the 5 image dependent color patches as reference. The availability of color patches greatly simplifies the task in comparison to a direct enhancement method.

5. TESTING

In order to test the feasibility of our approach, we assembled a set of test images, augmented them with different sets of sRGB color patches, simulated the effects of using 3rd party replacement inks for HP cartridges and 30 years of light exposure (using an ICC fade simulation profile), printed the simulated faded image, scanned the image using an HP scanner, retrieved a scanned image (in both scannerRGB and Lab color space) and restored the image first by using the “restore faded image” functionality of the scanner itself, second by using our new automatic method and third by using a manual method in Photoshop.

The color changes induced by the perceptual color transformation of the printer, the fading of the print, and finally the processing in the scanner are visualized in Figure 5a for 27 regular spaced color patches in the sRGB space. The vectors ending in blue squares visualize the changes that happen to the original colors. The overall gamut shrinks from the RGB cube (gray wire frame) to the transparent magenta gamut in the middle. The strongest changes occur in the dark Red, pure Red, dark Magenta and pure Green.

Figure 5b shows the relationship of the 27 original sRGB patches (mesh) in relationship with the (scanner)Lab values in Lab color space. Any attempted image reconstruction has to invert exactly those changes. The color transformation is challenging because the reconstruction is non-linear, and it has to perform a strong gamut expansion of the scanned data, which is often noisy.

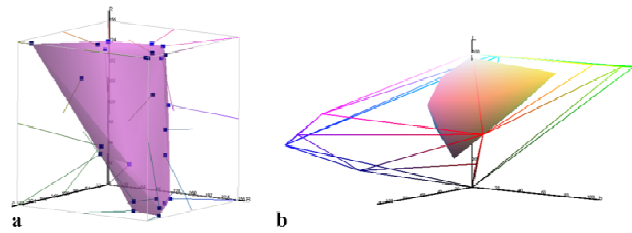


Figure 5. a) 27 patches from sRGB to scanner RGB (blue squares) in sRGB space, b) Gamuts formed by the 27 original (wire frame) and the scanned patches in Lab space.

6. RESULTS & DISCUSSION

The results for each solution are shown in Figure 6 and 7. We selected the “monk image” because it contains highly saturated content (robe of the monk), skin tones, and neutrals and the “flowers image” because of the natural objects (sunflowers, roses, leaves). Furthermore, these images also contain colors (reds & greens) that we know change quite strongly for the fading example that we are using. In addition, it is very easy to tell if the colors are off after the reconstruction of images containing natural objects.

We visually compared the results of the automatic and manual reconstruction against a reconstruction method implemented in an HP scanner and against the original digital image. We opted against a simple pixel-by-pixel numerical comparison of the original and reconstructed image as we think a complex human judgment is more appropriate. Any metric would need to consider that an image of a particular resolution has been halftoned, printed, might have faded quite dramatically over time and has been scanned through which noise has been added and finally restored as good as possible. Any metric would need to consider spatial aspect as well as color differences and could for example be based on S-CIELAB [12]. The ultimate evaluation would still be systematic psychophysical tests.

6.1. Results of the method deployed in HP Scanners

Looking at the results of the algorithm implemented in the HP scanner (Figure 6c and 7c), we can definitely see an enhancement of the scanned image. However, the images have a bluish cast and are still quite pale in comparison to the original images. This algorithm does not use any meta-information and has been developed for the restoration of faded Silver Halide prints.

6.2. Results of the Automatic Method

Figures 6f and 7f show the results of an automatic reconstruction using 125 color patches which are uniformly distributed in the sRGB color space. Considering that the fading is a non-linear process and that we are commonly using about 1000 patches to generate an ICC printer profile the results are actually quite good. Trying to reduce the number of color patches even further we found out that 27 (3x3x3) uniformly spaced color patches resulted in unsatisfying reconstructions. However, placing 30 color patches in more strategic color areas yielded to results presented in Figures 6e and 7e which are almost as good as the ones achieved with 125 patches. The skintones are not as well reproduced, neither are dark browns (e.g. center of the sunflowers) and near neutral colors (e.g. water container of the flowers).



Figure 6. "Monk" Image –a) Original sRGB b) Scanned c) Reconstructed by scanner software d) Manual reconst. e) Automatic reconst. 30 patches f) Automatic reconst. 125 patches.

Not surprisingly the reconstruction gets more and more challenging as the number of color patches are reduced. Ensuring an as accurate reconstruction of neutral colors and skintones is both challenging and can lead to visually disturbing artifacts. The human visual system is very sensitive in the areas around the neutral axis. Further testing and ensuring a truly robust solution for a set of different fading characteristics is still on our TODO list.

6.3. Results of the Manual Method

As previously mentioned, the result of the manual reconstruction is highly dependent on the user skill level and the software used. We focused on developing an easy reconstruction method as it is most likely the first correction step a user would undertake. Our results shows that while not achieving the quality of the automatic method, adjusting the black, white, and grey point using the patches and using curves to reduce the flatness of the image greatly improves the faded image. Since these are global changes and the fading is non linear, some colors such as the dark red roses in figure 7d is not reconstructed as well as with the automatic method.

7. CONCLUSION

We presented a simple hardcopy backup solution that can be used as a fall-back solution in the advent that a digital backup fails. This solution is based on augmenting a printed image with human readable tracking features. These features are based on standard color and imaging tools to give the user the most flexibility in his alternatives for reconstruction in the future. As it does not require a specific printing technology, it can also be inexpensive to produce. We also implemented a reconstruction system based on the color patches and compared it to current solutions. The manual and automated reconstruction showed good results and resiliency especially considering that fading is a destructive process. More user studies will need to be conducted to determine the ideal implementation.

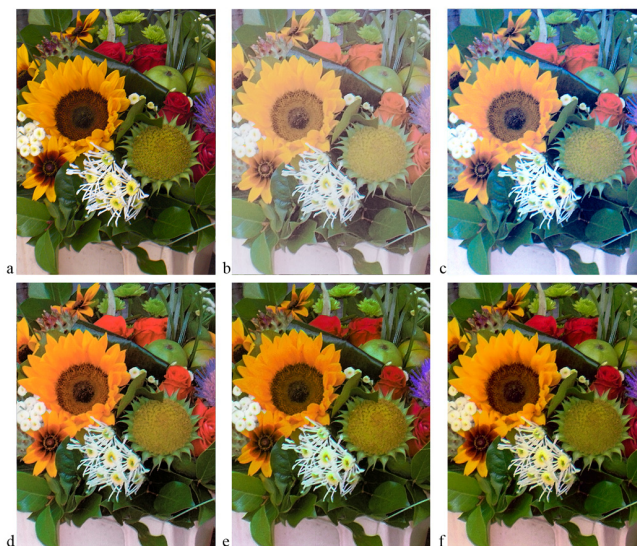


Figure 7. "Flowers" Image –a) Original sRGB b) Scanned c) Reconstructed by scanner software d) Manual reconst. e) Automatic reconst.. 30 patches f) Automatic reconst. 125 patches.

8. REFERENCES

- [1] I3A, International Imaging Industry Association <http://www.savemymemories.org/>
- [2] F.S. Frey, R. Gschwind and L. Rosenthaler: "Electronic imaging, a tool for the reconstruction of faded color photographs and motion pictures", IS&T/SID Color Imaging Conference, 1996.
- [3] Raymond S. Lee, "Innovative Technologies to solve imaging problems", Proc. 12th Int. Symp. on Photofinishing Technology.
- [4] Yujiro Ando, Aran Handsuesai, and Kanmanus Khantong: "Digital Restoration of Faded Color Images by Subjective Method", Journ. Imag. Sc. & Tech., Vol. 41, Nr. 3, May 1997.
- [5] www.wilhelm-research.com
- [6] R. Samadani and D. Mukherjee, "PhotoPlus: Auxiliary information for printed images based on distributed source coding", Proc. SPIE, Visual Comm. & Img. Proc., Jan. 2008.
- [7] B. Morse, D Thornton, Q. Xia, J. Uibel: "Image-Based Color Schemes", ICIP 2007.4379355
- [8] N. Moroney, P. Obrador, G. Beretta: " Lexical Image Processing", Proc. 16th. IS&T/SID Color Imaging Conference.
- [9] I. Amidror, "Scattered data interpolation methods for electronic imaging: a survey", Journ. of Elect. Imag., Vol. 11, 2002.
- [10] I. Tastl, K. Koh, D. Rossing, D. Berfanger, N. Moroney, "ICC Profile Based Defect Simulation", Proc of IS&T/SID CIC 2006.
- [11] G. Sharma: Digital Color Imaging Handbook, CRC Press, 2003.
- [12] X.M. Zhang and B.A. Wandell: "Color image fidelity metrics evaluated using image distortion maps", Signal Proc., Vol. 70, No 3, pp. 201-214, Nov. 1998.