Removing Background Color from Blueprints

Martí María and Miguel A. López-Álvarez, Hewlett-Packard, Large Format Printing Division; Sant Cugat; Spain

Abstract

A blueprint is a type of paper-based reproduction. For almost a century blueprint was the only low cost process available for copying drawings. Despite today this technology has been largely superseded by digital printing, it is still being used in emerging countries. When making copies of those documents, the colored background often makes difficult reading the information in the scanned plot. Hence, having an accurate way to transform the background to white enhances the look of the plot, plus saves ink when printing. This paper introduces an algorithm for removing the background color from a blueprint scanned image, comprising processing the image in a colorimetric space by determining the colorimetric values of the background color of the image, computing a linear chromatic adaptation transform for transforming the background-color colorimetric values to values of a target white point, and applying the computed chromatic adaptation transformation to the image.

Problem statement

Large format scanners and printers are widely used in the architect-engineer-constructor technical market for printing, scanning, and copying all kind of drawing and CAD plots. These devices have to deal with a wide variety of substrates like plain, glossy, translucent, or tracing papers among others.

Hewlett-Packard Designjet T2300 was introduced in fall'11 as one of the very first totally integrated large format multifunction printers. Attached to a large format printer there is a 36" wide format scanner. Hence, the printer and the scanner are combined in the same device. With the advent of such hardware, we can deal with media not yet addressed in large format arena. One of those unsolved formats is blueprints.

A blueprint is a chemical type of paper-based reproduction usually of a technical drawing, architecture or engineering design. The photosensitive compound, a solution of ferric ammonium citrate and potassium ferricyanide, is coated onto paper. As a result the media get tinted with blue color. Blueprints are indeed a very ancient reproduction method: 1861 Alphonse Louis Poitevin, French chemist, found the reduction and conversion of ferric salts to a ferrous state when exposed to UV-light

We received a strong requirement from marketing about supporting blueprints in the Designjet T2300, mainly because they are still very useful in emerging countries, mostly China (where almost 100% of the small-medium civil construction projects are designed in blueprints, some of them even handmade ones), India, and South/middle America (mainly Brazil and Mexico). In the US, although blueprint is not used for new designs anymore, there is a huge amount of blueprints in the archives from 10 to15 years ago and beyond. Many of these plots could be accurately digitalized if necessary. The image above is a blueprint of the galleon La Belle commissioned by Louis XIV, King of France.

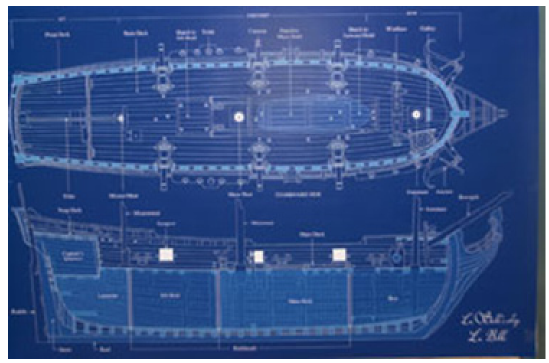

Figure 1. Blueprint sample

Prior solutions

Obtaining good scanned copies of blueprints is difficult because of the low contrast between the background and the printed information. Several methods are known for removing an image background color. Prior art is described in US patent 7,085,413 by Huang et al. The method disclosed in this patent involves identifying the background color using an RGB histogram of the image. After that, an adjustment is made to the background color entries in a mapping table used to convert RGB input to CYMK output; the adjustment made causes the RGB background color to be mapped to white in the CYMK output color space. This method works well for removing the background color, and since tri-dimensional interpolation is being used in the mapping, no contouring artifacts are generated. But since the changed colors are limited to background and its immediate neighbors, the rest of the blueprint -which is also blue-tinted- is unaffected; as a consequence, the overall image appearance is lost in the process.

Other algorithms clean the full image not only removing the background, but restoring all shades of color. Well-known solutions operating in this way are histogram normalization and equalization. The most frequent color in the histogram is assumed to be the background and the rest of colors are normalized to this value. This usually increases the global contrast of many images, especially when the important information in the image is represented by close contrast values. Those methods may give very good results, but have two important drawbacks. First, images are handled in their native color space, meaning that colors may be transformed in a way that is unexpected for our visual system; therefore, image appearance may be lost as well. Second, if the apparent gamma of the native color space is not close to linear, light detail loss may happen, especially in highlights and/or dark shadows. This is a well-know effect, and is the reason why black point compensation is performed in the XYZ colorimetric space [1]. This approach is also computationally expensive, as it requires previous histogram calculation and a scaling operation on each image pixel.

On commercially available devices, some large format scanners provide a set of image processing algorithms like “contrast”, “background cleaning” or “negative” that can help improving the quality of the final image. These settings are helpful for experienced users but they are also highly image dependent. Hence, for better results they have to be individually adjusted for each image. Alternatively, some scanning solutions only offer conversion to gray which often is useless as it only provides a smaller contrast than the original color image.

Our solution

In color science, the tern “chromatic adaptation” refers to the human visual system’s capability to adjust to widely varying colors of illumination in order to approximately preserve the appearance of object colors [2, 3]. For example, if an observer who is adapted to daylight moves to looking into a white (reflecting all the wavelengths equally) piece of paper under a slightly tinted blue light, the paper looks initially blue to the observer. After a certain period of time, the observer gets fully adapted to the illuminant. Chromatic adaptation works removing the blue component of the illuminant and makes the paper look approximately neutral white. This effect depends on the kind of illuminants involved and their chromatic difference, and it is limited to moderate amounts of chroma. It is quite uncommon for a real observer to be totally adapted to a highly tinted light.

Figure 2. Simulated CAT effect. Un-adapted on the left, adapted on the right.

Appearance models allow us to estimate the representation of an object under a different light source than the one on which it was recorded. A common application is to find a chromatic adaptation transform (CAT) that will make the recording of a neutral object appear neutral, while keeping other colors also looking realistic [4, 5]. On removing the chroma of the illuminant, a chromatic adaptation transform can effectively preserve the appearance of a scene and predict how the scene will look like when other illuminant is being used. Those properties seem to fit very well in solving the problem of appearance-preserving background removal: Since chromatic adaptation transform works great for illuminants, our idea is to leverage it for colored media like blue prints.

In order to develop this idea, we inspected modern appearance models in order to find a suitable chromatic adaptation transform that could be used to simulate adaptation to highly chromatic lights. Bradford was considered but discarded because a non-linearity of the blue channel. CIECAM02 [6-8], which is one of the most used modern appearance models, uses a modified CMCCAT2000 transform, CAT02 which is a linearized Von-Kries type of chromatic adaptation transform. This type of chromatic adaptation transform can be pushed to support full adaptation on very chromatic illuminants. As a part of a small observation experiment, we found that if we use a blue illuminant on a printed white paper and apply a CAT02 transform from the blue illuminant to any blackbody-locus neutral illuminant, then the appearance is mostly preserved and the results are pleasant to end users.

The math behind is very simple. The L, M, S cones responses to a given color stimulus can be calculated as:

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9831 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (1)$$

Chromatic adaptation can be calculated as the change in the cones responses due to a change in the illuminant the observer is adapted to:

$$\begin{bmatrix} L_c \\ M_c \\ S_c \end{bmatrix} = \begin{bmatrix} \alpha \frac{L_{wr}}{L_w} D + 1 - D & 0 & 0 \\ 0 & \alpha \frac{M_{wr}}{M_w} D + 1 - D & 0 \\ 0 & 0 & \alpha \frac{S_{wr}}{S_w} D + 1 - D \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (2)$$

In our case complete adaptation was being used and D factor turned to be 1. After some experimentation, we found this specific chromatic adaptation to perform very well for the task of removing background color in certain scanned images of blueprints.

Since chromatic adaptation works in the XYZ colorimetric space, we need some way to convert from the scanner space to XYZ, and from XYZ to the printer native space. The color management process, which already takes place in the multifunction large format printer, was using International Color Consortium (ICC) profiles [9, 10]. ICC profiles do use what is known as the PCS. This is a profile connection space which may be CIE L*a*b* or XYZ. The profiled color pipeline used in the multifunction printer was already using a conversion to the PCS, and since the color management is applied in hardware latter in the printing part, we could easily integrate the chromatic adaptation algorithm as an additional step in ICC the profile pipeline. Since the full profile transformation is embedded in a single tri-dimensional look-up table and applied at hardware level, we can apply our algorithm at zero cost in terms of computing time.

Algorithm

The complete algorithm can be described as follows:

1) Few first lines of the scanned image are captured and buffered. An RGB histogram on that data is created. The upper 95% part of histogram is averaged to figure out the mean RGB values of background.

2) XYZ tristimulus for this RGB triplet are obtained by using the input ICC profile corresponding to the currently scanned paper. This is the regular, colorimetric profile for the scanner. It has no blue-removal capabilities. If Y of background is lower than rest of colors, then the image should be inverted (negative). This is the case of the sample in figure 1, where the background is darker than the drawings, which is not usual in more modern blueprints.

3) A linearized CAT02 chromatic adaptation transform is calculated. It goes from the XYZ of blue background to D50. This illuminant is chosen because it is the reference white for ICC color management.

4) An abstract profile from L*a*b* color space to L*a*b* color space is then computed. The profile contains conversions from L*a*b* to XYZ, where the chromatic adaptation step is applied prior to a further conversion from XYZ back to L*a*b*. This profile may be computed in real time as the necessary computational overhead is very low.

5) Finally, the abstract profile is inserted in the color management workflow between the input scanner profile and the output printer profile. Designjet T2300 printers can apply the color management by means of application-specific integrated circuits (ASICs); therefore, with zero impact on throughput.

Results

The algorithm has been found to have some interesting properties;

1) No extra time required, since it is integrated in the color management logic. Most color managers create a single color map for all steps in the color transform pipeline, which is currently the case in our multifunction printers. Since this color map is being evaluated using hardware in the printing pipeline, there is no throughput penalty.

2) Colors other than the blue background are transformed and appearance of the image is preserved. This is a key advantage and makes a significant difference with other similar approaches.

3) It is not limited to blueprints, may handle any colored background. For example yellow paper and recycled or tracing paper as well. One interesting application of this feature is the ability of the algorithm to "clean" copies of colored newspaper [10, 11]. We keep discovering new use cases on areas that may initially seem unrelated.

4) It is harmless for drawings already printed on white; it can be always turned on for automatic background color removal. This feature suggests a multifunction copy mode that no longer would be deal only with blueprint but with any colored media. We are currently analyzing what are the limits the algorithm has.

5) No user interaction is required. There are no settings to play with.

In order to check the algorithm, a prototype was built using the Little CMS library [12]. The code is simple since it only computes an XYZ to XYZ 3-by-3 matrix and creates an ICC profile that encapsulates this matrix. It is up to the printer color manager to provide proper conversions for the scanner profile, which is operating in the CIE L*a*b* profile connection space and the output printer profile, operating in the same profile connection space. All profiles in the color pipeline are linked together in a single color map which is readily uploaded to the ASIC by the printer color firmware.

For testing our solution, we obtained a set of nearly 50 blueprint originals from China and Brazil, provided by Hewlett-Packard customers. Historic blueprints were also available at Free Library of Philadelphia. The result of our algorithm is quite evident, even for non-expert readers, by looking at figure 3. In this sample no image processing is being applied other that the chromatic adaptation. We can clearly see how most of the blue background has been effectively transformed to white, while the information of text and lines is maintained and the contrast is improved as a result of the transformation.

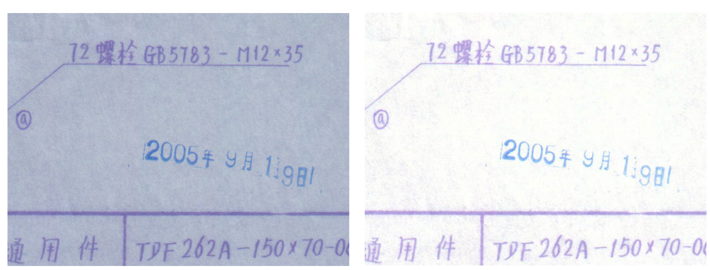

***Figure 3**. Original and result after chromatic adaptation background cleaning*

In order to properly evaluate the algorithm in a more systematic way, several well-trained engineers were asked to compare the performance of the transform versus the solution previously offered by some large format scanner vendors. At that time we only wished to compare the suitability of the algorithm. Since previous solutions typically involved a conversion to gray, many observers suggested a fairer comparative versus other more advanced algorithms, which would be capable to handle color as well. Currently a comparative versus histogram based methods is already planned, and will be conducted on depending on business needs.

Nevertheless, the original test, which was designed to check the acceptability of the algorithm and the preference of the observers, clearly showed that our approach was far superior to previously offered solutions based in gray conversion.

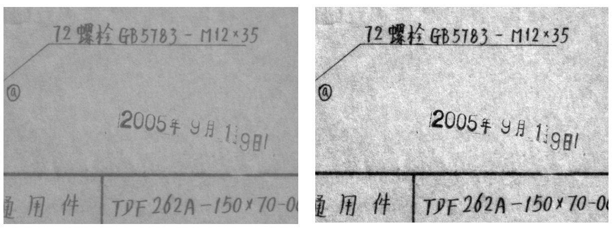

***Figure 4**. Result of scanning the original blueprint in Figure 3 with current scanner solutions that only offer transformation to grayscale (left). Same image after applying enhancement transformations (right).*

If we compare these results against the results obtained from previous scanning solutions, which consist on using different scan presets and let end user modify some image processing settings to improve the scanned image, our approach –in all cases- was chosen as superior in terms of image quality by all observers, with the added bonus that it needed no user interaction. In Figure 4 we show the results of the method suggested by the one current scanner vendor, including the best possible enhancement we could apply. The same blueprint is processed by our method in Figure 3.

The algorithm has proved to perform correctly, but some preliminary investigation indicates that can be improved by combining it with more traditional approaches. For example, after a chromatic-adaptation transform, a conventional background cleaning algorithm may be applied for obtaining even better results. Conversion to gray is also greatly improved if a chromatic adaptation transform is previously used.

Conclusions

This solution may be useful in the case of scanning original plots which have a colored (not white) background, like –typically- blueprints, tracing or recycled paper originals. The colored background often makes difficult to read the information in the scanned plot, since resulting contrast is usually low. Hence, having an accurate way to transform the background to white enhances the look of the scanned image, plus saves ink when printing copies of it.

References

[1] Lars Borg, Black Point Compensation from Adobe Systems. ICC White Paper Adobe 1 bpc, at www.color.org/whitepapers.html (2004).

[2] CIE Technical Report 159:2004 – A Colour Appearance Model for Color Management Systems: CIECAM02.

[3] Mark D. Fairchild. Color Appearance Models (Addison-Wesley, Reading Massachusetts, 1998).

[4] CIE TC 1-52 (2004). A Review of Chromatic Adaptation transforms. 160:2004. . s.l. : CIE. ISBN 978-3-901906-30-5.

[5] Changjun Li, M. Ronnier Luo, Bryan Rigg, Robert W. G. Hunt, "CMC 2000 chromatic adaptation transform: CMCCAT2000". Color Research & Application **27** 49–58, (2002).

[6] N. Moroney, M.D. Fairchild, R.W.G. Hunt, C. Li, M.R. Luo, T. Newman, The CIECAM02 Colour Appearance Model. Proceedings of the 10th IS&T/SID Color Imaging Conference, Scottsdale, USA, (2002).

[7] M.R. Luo and R.W.G. Hunt, "The Structure of the CIE 1997 Colour Appearance Model (CIECAM97s)", *Color Res. App.* **23** 138-146 (1998).

[8] N. Moroney, "A Hypothesis Regarding the Poor Blue Constancy of CIELAB", Color Research and Application, **28** 371-378, (2003).

[9] Specification *ICC.1:2001-04 File Format for Color Profiles*. International Color Consortium, available at http://www.color.org/

[10] P. Green. Color management: understanding and using ICC profiles (Wiley-IS&T, 2010).

[11] R. W. G. Hunt. Measuring Colou (Fountain Press, 3rd edition, 1998).

[12] M. Maria. Little CMS Engine – How to use the Engine in Your Application, 2004. Available from http://www.littlecms.com.

Author Biography

Marti Maria is a color engineer at the large format printer division of Hewlett-Packard. He worked previously at ICR; a company specialized in imaging and color. Marti is also the author of well-known open source color oriented packages, like the Little CMS open CMM and the LPROF profiler construction set. He has contributed to several color books and was session chair on Color & Imaging Conference 16.

Miguel A. López-Álvarez obtained a B.Sc. in Physics in 2003, a B.Sc. in Electrical Engineering, and a Ph.D. in multispectral color science in 2007 by the University of Granada (Spain). Since 2008 he works in Hewlett-Packard as a color and digital imaging engineer, focusing on color management and image quality of large format printers and scanners.