Recent Progress in Digital Halftoning for Color Reproduction
State of the Art Report

R.D. Hersch, V. Ostromoukhov

Ecole Polytechnique Fédérale, Lausanne, Switzerland
hersch@di.epfl.ch, victor@di.epfl.ch

ABSTRACT

Due to the proliferation of low-cost color printers during the last few years, converting full-color images to halftoned bi-level separations remains an important issue. A halftoning technique should faithfully reproduce the original color image, while satisfying several quality criteria such as absence of visible artifacts and Moires, providing reliable color reproduction as well as good small detail and texture rendition. Digital halftoning techniques are not limited to classical methods, such as clustered-dot dithering, dispersed dot dithering and error-diffusion methods. In order to obtain the best results, recent algorithms combine the respective advantages of dithering and error diffusion methods. One of the most spectacular progress in color reproduction technology has been made in the domain of ink-jet printers, where microdrops are used to achieve a larger dynamic range without increasing the spatial resolution. Colour rendition techniques specially adapted for these new printing devices are presented. A further advance in printing technology consists of incoporating into the printer more than the traditional 4 process colors. Advanced color separation techniques are required in order to print multicolor images with more than the Cyan Magenta Yellow and Black process colors. One of these advanced techniques combining color separation and halftone generation is error-diffusion in color space. Finally, we give a brief overview of the colour calibration standard proposed by the International Colour Consortium.

1. INTRODUCTION

Today's colour reproduction market is characterized by two main classes of applications: desktop reproduction and high-speed on demand printing. Desktop reproduction enables individual users to directly create and ouput their documents. For them, there is a large offer of low-cost, high-quality, low throughput printers (HP, Canon, Epson, Olivetti, etc.). High-speed on demand printing machines are generally operated by print shops who are used to work with offset machines. These professional print shops have high printing volumes and require therefore flexible high-throughput machines.

The main printing technologies are ink-jet and colour laser printing with either solid toners (Canon CLC, Tektronix Phaser 550) or special inks (Indigo). For the desktop, dye sublimation and wax transfer technologies are also used. These technologies, especially inkjet and laser printing, are being constantly improved. Higher quality and lower priced devices are proposed both for desktop reproduction and for high-speed on demand printing.

Higher quality output is obtained by increased resolution and by increasing the number of printable levels per pixel. A larger output colour gamut can be attained by printing with more than the traditional process colours Cyan, Magenta, Yellow and Black (hexachrome printing for example). Maintaining a high degree of colour fidelity throughout the colour reproduction process requires the appropriate calibration of input devices (scanners), display devices (monitors) and output devices (printers).

In this state of the art report, we will focus our attention on current approaches for colour halftoning and for calibrated colour reproduction.

2. COLOUR HALFTONING FOR VARIABLE DOT SIZE PRINTERS

In the last two years, 300 dpi to 600 dpi high-quality ink jet printers have been offered for desktop publishing at very low cost (below 300 dollars). New halftoning algorithms based on dispersed-dot dithering [Mitsa92], [Ostromoukhov94], on improved error-diffusion schemes [Eschbach93] or on combinations of error-diffusion and dithering techniques [Miller88] have provided the means to reproduce both grayscale and colour images.

Currently, ink-jet device manufacturers are making efforts to put on the market low-cost variable dot size ink-jet printers able to reproduce multiple intensity levels. Multi-level inkjet printers seem easier and cheaper to develop than devices having a significantly higher resolution. The main effort resides in ensuring a constant, repetitive small droplet diameter and at the same time a minimal dot gain by minimizing the ink spread on paper.

Multiple intensity levels per pixel are achieved by printing one, two or several droplets at the same position. Since a single droplet has a minimal diameter, say 50% of the diameter of the largest printable dot size, the first darkness level (or surface coverage level) an ink-jet printing device may print is at least 25%, the second darkness level is at least 45%, and the remaining levels cover the darker levels between 45% and 100% darkness. It is therefore of capital importance to use halftoning algorithms in order to obtain additional intermediate intensity levels (Fig. 1), e.g. levels between 0 and 25% darkness, levels between 25% darkness and 45% darkness, etc.

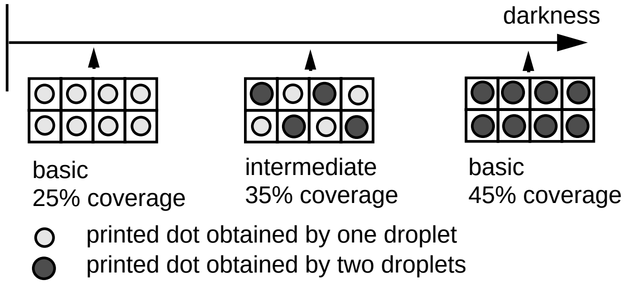

Figure 1. Intermediate darkness levels obtained by dithering between basic levels.

The quality criteria for judging and comparing halftoning algorithms are the following:

- visibility of individual dots or screen elements should be minimized

- the number of intensity levels should be large enough (> 40) to avoid banding effects
- structure artifacts, i.e. repetitive or semi-repetitive visible structures should be avoided
- false contours due to sharp halftone structure changes should be avoided

We present dither-based and error-diffusion methods for the halftoning of images on multi-level printing devices. The resulting visual effects are shown by simulating the printed dots of a multi-level inkjet printer.

3. DITHER TILE BASED MULTI-LEVEL HALFTONING[1]

Multi-level halftoning aims at generating additional intensity levels between the levels produced by printing successive ink drops on paper. When the set of available ink drops produces round dots, and when the colour inks are near to ideal inks or when colour layers can be placed without phase shifts one on top of the other, colour layers can be halftoned separately by a dispersed-dot dithering method such as Bayer dispersed-dot dithering [Bayer73], rotated dispersed dither [Ostromoukhov94] or with the help of any adequate dispersed-dot dither threshold array.

Dither-based halftoning methods [Ulichney87] are based on dither tiles paving the plane. Parallelogram or respectively hexagonal dither tiles for dispersed-dot dither can be generated recursively with two-fold (Fig. 2a) or respectively three-fold dispersion (Fig. 2b) of threshold levels [Ostromoukhov95].

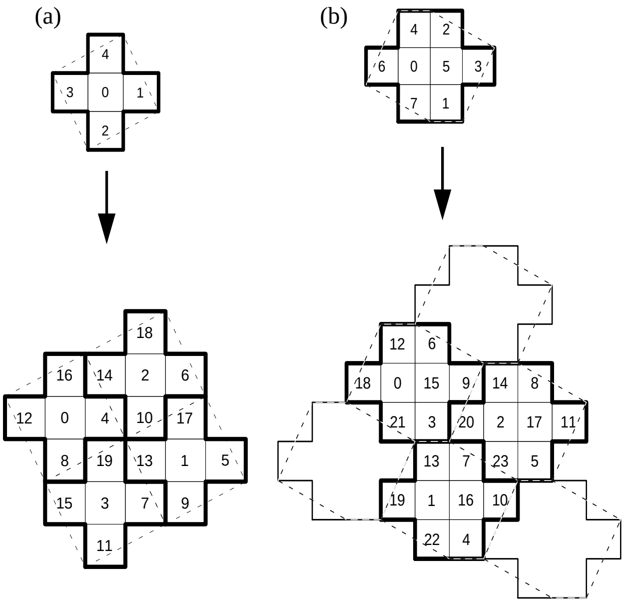

Figure 2. Recursive generation of a parallelogram (a) and hexagonal (b) dispersed dither tiles.

The number of additional intensity levels which may be produced between two levels given by *k* and *k*+1 droplets printed on a single pixel depends on the size of the dither tile. One has to be careful to select a small dither tile, since a large dither tile contains low frequency components and therefore generates more visible artifacts at low and middle resolution. Between the highest intensity level (no droplet printed) and the next lowest intensity level (one droplet printed), the dither array based multi-level dithering method behaves in the same way as bi-level dithering and produces similar artifacts, but at a reduced intensity.

In some cases, due to the displacement of the ink-jet head, successive droplets are not printed exactly at the same place, but in a slightly eccentric manner. The resulting printed dot has an elliptic shape. Since the elliptic dot touches neighbouring elliptic dots first in one direction and only after a certain number of intensity levels in the other direction, at certain levels, bands between elliptic dots become visible (Fig. 3a). In order to break these vertical or horizontal bands, the dither threshold levels should be arranged to produce an elliptic dot growing pattern breaking the continuity of the bands (Fig. 3b).

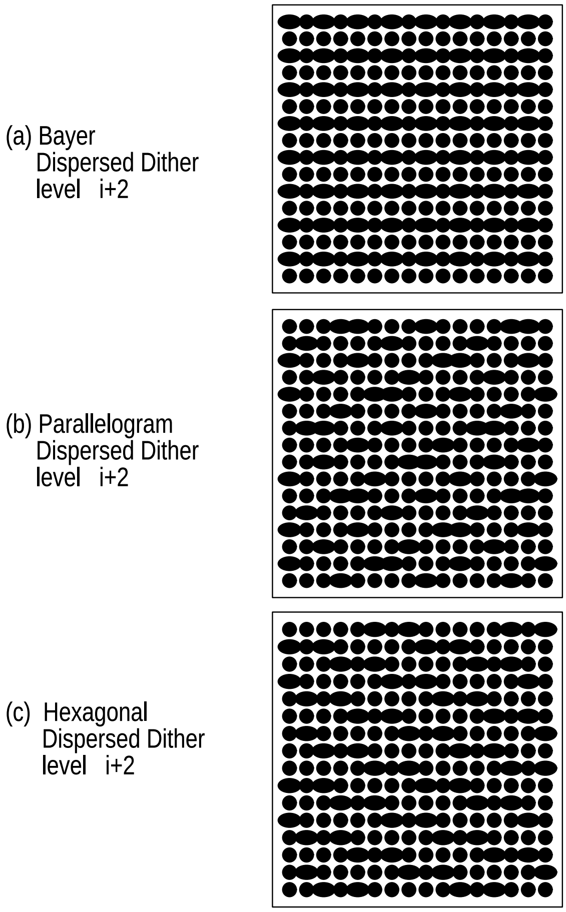

Figure 3. Breaking the horizontal bands produced by elliptic dot shapes with distributions of dither thresholds according to the tiles shown in Fig. 2a and 2b.

4. ERROR DIFFUSION BASED CONSTRAINED MULTI-LEVEL HALFTONING

Error-diffusion in colour space is proposed as a means of reducing artifacts which appear when error-diffusion is applied independently to each colour channel (red, green, blue or cyan, magenta, yellow). These artifacts, also called correlated noise or "worms" become strong when separatly error-diffused halftoned layers are superimposed.

---

[1] Parts of this section have been published in V. Ostromoukhov, et. al., “Dithering Algorithms for Variable Dot Size Printers,” Proc. IEEE Int. Conf. on Image Processing (ICIP’96), 1996, Vol 1 , 553-556.

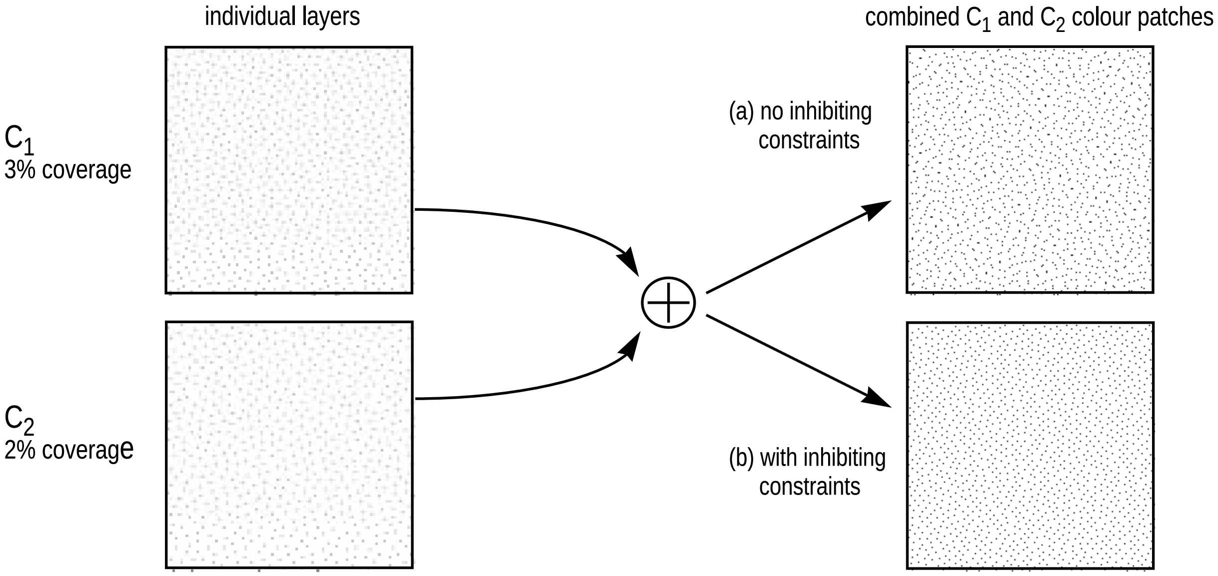

Figure 4. Colour error-diffusion (a) without and (b) with constraints inhibiting dot over dot printing.

Error diffusion in colour space relies on the idea of computing for each output pixel an error diffusion vector composed of the 3 colour components and to weight and distribute this error vector to neighbouring pixels. At each ouput device pixel, the choice of the ouput colour is made by computing the euclidian distance between the desired ouput colour plus the error colour vector diffused from neighbouring pixels and the available output colours. The ouput colour minimizing that euclidian distance is selected. Since the set of available basic colours can be colours lying anywhere in colour space, error-diffusion in colour space can be used both for colour separation when printing with more than four process inks and for halftoning purposes. In order to produce good results, the basic colors should span the same or a larger volume than the colours of the image which is to be reproduced.

As already pointed out by Sullivan, Miller and Wetzel [Sullivan89], when the chosen output colours lie on a parallelepiped whose sides are parallel to the chosen output space axes, separate error-diffusion in each of the colour layers and vector error diffusion in 3D colour space yield the same result. Sullivan, Miller and Wetzel propose to diffuse the error vector in the non-linear CIELUV colour space and to blur the output vectors in order to minimize error-diffusion artifacts.

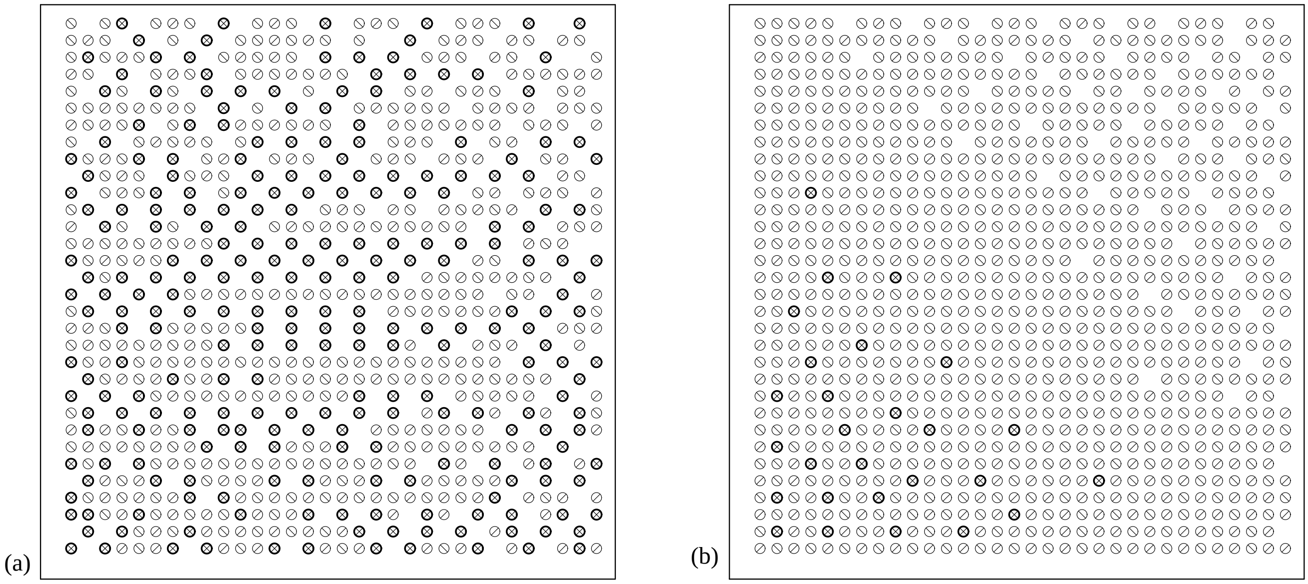

Figure 5. Bilevel colour error-diffused patch with surface coverages c1= 50%, c2=50% at the center and with ±12% coverage increase/ decrease towards the edges (a) without and (b) with constraints inhibiting dot over dot printing.

Klassen, Eschbach and Bharat [Klassen94] propose a colour error diffusion method for printing gray tone in colour images, which reduces the intensity of artifacts by distorting the colour space so as to induce the replacement of black pixels by side by side printing of cyan, magenta and yellow pixels. The approach we present is based on inhibiting constraints used for inhibiting the appearance of combined colour variable dot size pixels in highlight and mid-tone regions.

We apply color error diffusion in a linear RGB space which is obtained by a linear transformation of the parallelepiped whose vertices are given by the CIE-XYZ coordinates of solid printed Cyan, solid printed Magenta and solid printed Yellow. Worm-like artifacts are strong when colour layers are halftoned independently or equivalently, when colour error diffusion is applied with target colours C,M,Y,R,G,B,W forming a rectilinear parallelepiped in the RGB orthogonal output coordinate space. In highlights, when ink surface coverage percentages are low, one can inhibit the use of superimposed inks forming composed colour R,G,B,K. By removing these ouput colour candidates from the choice of output colours, highlight reds, greens and blues are generated by yellow-magenta, respectively cyan-yellow and cyan-magenta side by side dot printing. Fig. 4 shows the colour error-diffusion patterns when two colour layers (layer c1 at 3% and layer c2 at 2% surface coverage) are printed (a) without constraints, i.e. with the possibility of having overlapped c1 and c2 dots, and (b) with constraints inhibiting dot over dot printing.

Clearly, the solution with constrainted error-diffusion provides less worm-like visual artifacts. This visual result can also be explained by the fact that constrained error-diffusion generates more printed pixels (side by side printing instead of dot over dot printing) creating thereby higher frequency artifacts which are less perceptible to human vision.

In the next example (Fig. 5), a colour patch is error-diffused in colour space with two basic colours c1 and c2 having each at the center of the image a surface coverage of 50% which varies towards the borders of the image by ±12%. One can clearly see that the patch where the dot superposition of c1 and c2 is inhibited provides less disturbing artifacts. In the patch generated without inhibiting constraints, regions with superimposed c1 and c2 colours contain larger white areas which tend to create artificial boundaries.

Error-diffusion in colour space is also appropriate for variable-dot size printing. For the sake of simplicity and without loss of generality, we consider here a printer capable of printing dots either with one droplet at 50% surface coverage or with two droplets at 100% surface coverage. When printing two colours, for example cyan and magenta, the candidate printable colours for unconstrained error-diffusion are described in Table 1.

Figure 6a shows a two dimensional representation of the white-cyan- magenta-blue colour plane which is a part of the printable colour gamut in the RGB ouput colour space (Fig. 6b).

From Fig. 6, one can see that for example Blue at 50% can be rendered in two different ways: either by printing 1 droplet of cyan (cyan50%) overlaid with 1 droplet of magenta (magenta50%), which gives a substactive blue or by printing side by side dots formed by 2 droplets of cyan (cyan 100%) and 2 droplets of magenta (magenta100%), which gives a weighted additive blue.

In order to reduce halftone artifacts, intensity-dependent inhibiting constraints must be introduced. At some surface coverage levels, the colour resulting from overlaid c1 and c2 colours can be inhibited and at other surface coverage levels, it must be allowed. For example, for bi-level colour printing, blue between 50% and 100% can only be achieved if a minimal amount of superimposed cyan and magenta inks are allowed. This is also true for variable dot-size printing.

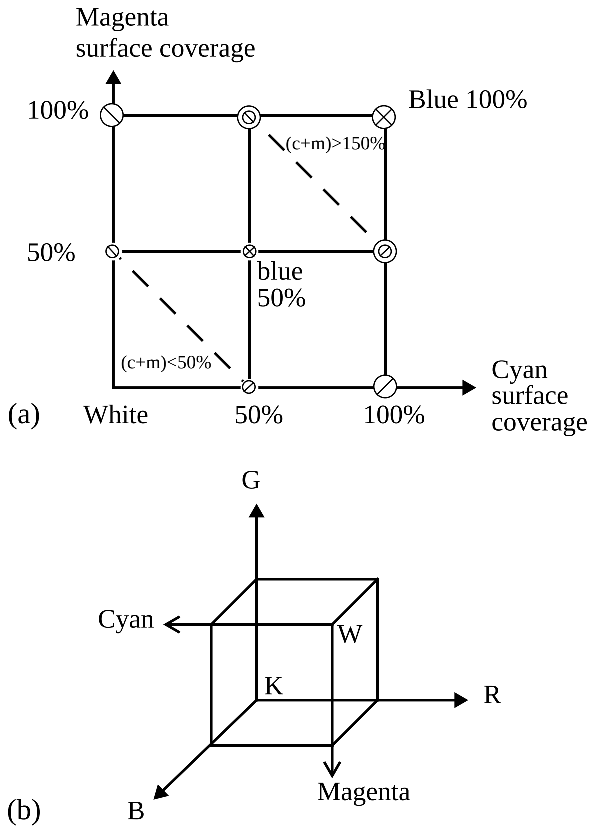

Figure 6. Two-dimensional representation of the white cyan magenta blue gamut and of corresponding error-diffusion output colours.

Table 1: Printable colours for variable dot size error-diffusion.

Colour	Label	Droplets	Symbol
White	White	no droplet	
Cyan 50%	c50%	one C-droplet	⊘
Cyan 100%	c100%	two C-droplets	⊘
Magenta 50%	m50%	one M-droplet	⦸
Magenta 100%	m100%	two M-droplets	⦸
Blue 50%	c50%m50%	overlapped one C & one M droplet	⊗
CyanBlue	c100%m50%	overlapped two C & one M droplet	◎
MagentaBlue	c50%m100%	overlapped one C & two M droplets	◎
Blue 100%	c100%m100%	overlapped two C & two M droplets	⊗

Constraints which vary according to the current colour intensity can be introduced for example by enlarging the distance between the current colour to be printed and the combined colour (blue) with an intensity dependent penalty factor.

Fig. 11 shows the feasibility of constained error-diffusion in colour space for variable dot size printing. One can see that at combined surface coverages below 50% (cyan + magenta < 50%) only few superimposed half- size colour dots (blue50%) appear. Similarly, at combined surface coverage levels below 150% (cyan + magenta < 150%), only few superimposed full size colour dots appear (blue 100%).

Fig. 12 shows a different solution, where the single drop superimposed 50% cyan and 50% magenta (blue 50%) colour has been completely discarded from the choice of printable colours. Furthermore, the intensity dependent penalty function is applied to all remaining printable colours, whose euclidian distance from the current colour in RGB space is larger than 1/2, assuming colour coordinates ranging between 0 and 1. Fig. 12 shows less artifacts and provides smoother intensity transitions than Fig. 11. The intensity dependent penalty function has a heavy impact on the resulting error-diffusion halftone quality. Further research is needed to optimize the penalty functions used for the inhibiting constraints.

5. CALIBRATED COLOUR REPRODUCTION

The International Colour Consortium, founded by computer and peripheral equipment manufacturers Adobe, Agfa,-Gevaert, Apple, Kodak, Microsoft, Silicon Graphics, Sun Microsystems, Taligent and by FOGRA, the German graphic arts research institute, has created a standard for describing the device profiles of input, display (preview) and output devices [ICC95]. The input devices profiles for scanners and digital cameras have the function of relating scanned R,G,B data to a device-independent profile connection space, either CIE-XYZ or CIE-LAB. The input model (Fig. 7) consists of non-linear tone reproduction curves (TRC) for mapping device red green blue values to linear red green blue values and of a linear 3×3 transformation for mapping the linear red green blue values to the connection space CIE-XYZ values.

Display devices are considered to be either colour CRT or colour liquid cristal displays. RGB display profiles are characterized by each of the RGB channels tone reproduction curves (redTRC, greenTRC, blueTRC) and the CIE-XYZ values of their respective phosphors (or filters). Fig. 8a illustrates the transformation between device RGB space and the CIE-XYZ connection space. For the visualization of colorimetric connection space referenced data, the inverse transformation needs to be used (Fig. 8b).

Colour printing on paper is a strongly non-linear process. Professional output device calibration systems are generally based on 3-dimensional mapping tables. A 3D mapping table is constructed by printing combinations of all 3 or 4 output colours (cyan, magenta, yellow and black) at given output intensity intervals and by measuring the ouput samples as CIE-XYZ tri-stimulus values. The set of measured samples enables establishing a correspondance between CIE-XYZ and output CMYK values. The set of measured samples in CIE-XYZ can be tetrahedrized. Within each tetraheder, linear interpolation can be applied in order to compute for a new CIE-XYZ value the corresponding CMYK value (Fig. 9).

(a)

$$
\begin{aligned}
linR &= redTRC\,[deviceR] \\
linG &= greenTRC\,[deviceG] \\
linB &= blueTRC\,[deviceB]
\end{aligned}
$$

$$
\begin{bmatrix} conX \\ conY \\ conZ \end{bmatrix} = \begin{bmatrix} redColX & greenColX & blueColX \\ redColY & greenColY & blueColY \\ redColZ & greenColZ & blueColZ \end{bmatrix} \begin{bmatrix} linR \\ linG \\ linB \end{bmatrix}
$$

(b)

$$
\begin{bmatrix} linR \\ linG \\ linB \end{bmatrix} = \begin{bmatrix} redColX & greenColX & blueColX \\ redColY & greenColY & blueColY \\ redColZ & greenColZ & blueColZ \end{bmatrix}^{-1} \begin{bmatrix} conX \\ conY \\ conZ \end{bmatrix}
$$

$$
\begin{aligned}
deviceR &= redTRC^{-1}\,[linR] \\
deviceG &= greenTRC^{-1}\,[linG] \\
deviceB &= blueTRC^{-1}\,[linB]
\end{aligned}
$$

Figure 8. Display device calibration.

The ICC output device profiles characterizing the transformation between device-independent connection space (CIE-XYZ) to output space (CMYK or output RGB) is composed (Fig. 10) by a linear transformation (3×3 matrix), a 1D look-up table associated to each channel, a multidimensional look-up table and finally a 1D look-up table associated to each output channel. The dimensionality of the multidimensional look-up table is given by the number of input channels (generally 3). Each multidimensional look-up table entry contains as many values as output channels. Therefore, the profile characterizing a CIE-XYZ to CMYK conversion will contain a 3D look-up table, where the X',Y',Z' values are indices to access the table and where each table entry contains either a Null value (out of gamut colour) or a valid Cyan Magenta Yellow and Black value. One way of generating the 3D look-up table consists of computing the CMYK values of corresponding CIE-XYZ table entries (uniform 3D grid in CIE-XYZ) by interpolating them in the tetrahedrized sample space described above.

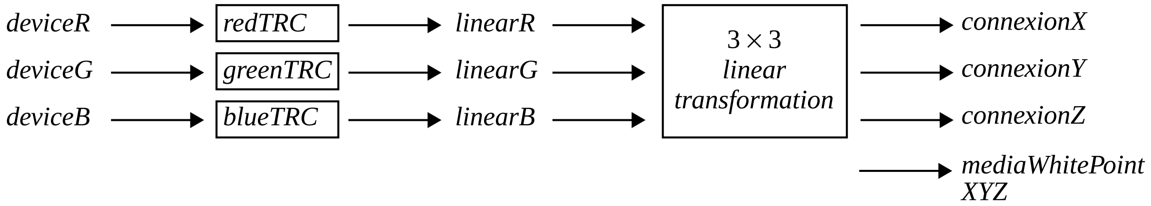

Figure 7. Input device calibration.

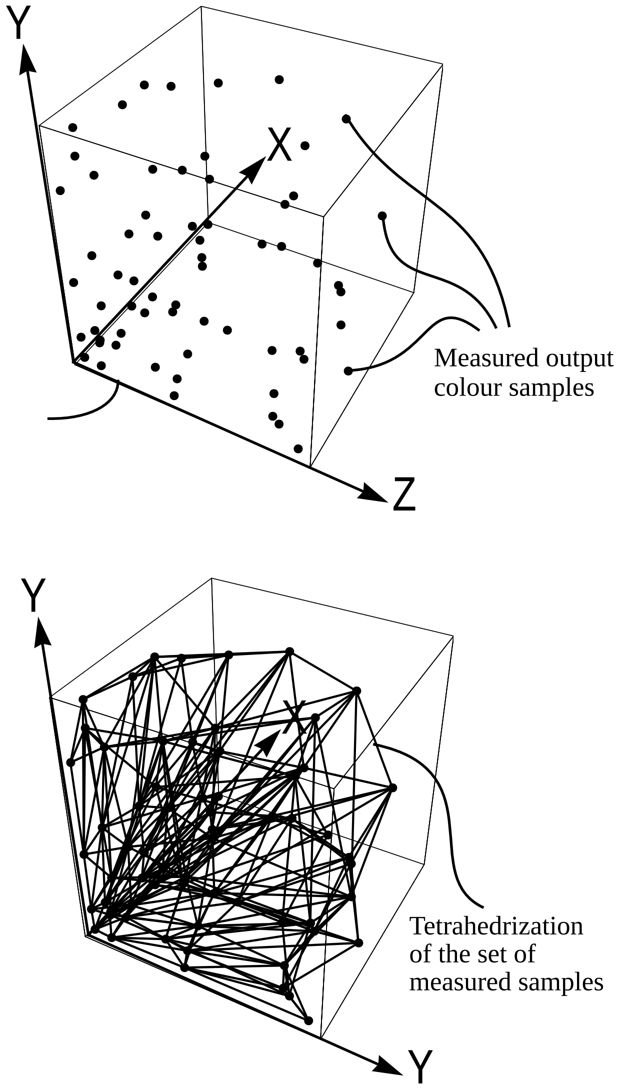

Figure 9. Tetrahedrization of the CIE-XYZ space by the measured output samples.

In addition to profiles for input, display and output devices, ICC profiles incorporate information about how measurement were made (0/45 degrees illumination geometry or diffuse illumination) and what the standard illuminants are (D50, D65, A, etc.). It also incorporates the viewing conditions (absolute XYZ of illuminant and surround in cd/m^2) and printing conditions (screen angle, frequency and spot shape).

Generally, the values given in the profile connection space (CIE-XYZ) are based on relative colorimetry. They are measured relative to a given illuminant, for example D50. When such values have to be displayed, respectively printed, the reference white CIE-XYZ$_i$ is transformed to media white (display white, respectively paper white). In that case, in order to obtain the printable tri-stimulus values CIE-XYZ$_a$, the connection space CIE-XYZ$_r$ values have to be rescaled so as to match the new media white (XYZ$_{mw}$).

$$X_a = \left(\frac{X_{mw}}{X_i}\right) X_r$$

$$Y_a = \left(\frac{Y_{mw}}{Y_i}\right) Y_r$$

$$Z_a = \left(\frac{Z_{mw}}{Z_i}\right) Z_r$$

In general, input devices and display devices have a larger gamut than output devices. Therefore, the ICC profile format supports the following rendering intents: perceptual, relative colorimetric, saturation and absolute colorimetric. Perceptual rendering means rendering on the output device so that input picture and output picture are perceptually identical, even if the measured colours vary considerably. Rendering by maintaining the relative colorimetry requires the adaptation of the colours to the new reference white as shown above. Saturation rendering consists in producing well saturated output images, keeping the hue as constant as possible. Absolute colorimetric rendering requires that all colours in the connection space are referenced in respect to an illuminant with known luminance (cd/m^2) and that exactly the same colours are reproduced on the output device, independently of the media white point. Out of gamut colours are not reproduced.

The ICC standard profile format therefore provides the framework, within which device manufacturers and application program producers may support input, visualization and output device profiles. It is up to the colour management systems integrated into operating systems (for example the ColorSync 2.0 for the MacIntosh operating system) to make use of the device profiles and reproduce the colour images accordingly.

The ICC standard however does not specify how to generate input, display and output device profile parameters. Generation of profile parameters appropriate for the different devices (scanners, displays, colour printers) remains part of the know-how of each of the contributing industries.

6. CONCLUSIONS

Variable dot size inkjet printers at moderate cost are being commercialized. Due to the relatively large size of single droplets, halftoning algorithms are still needed. However, since halfton-

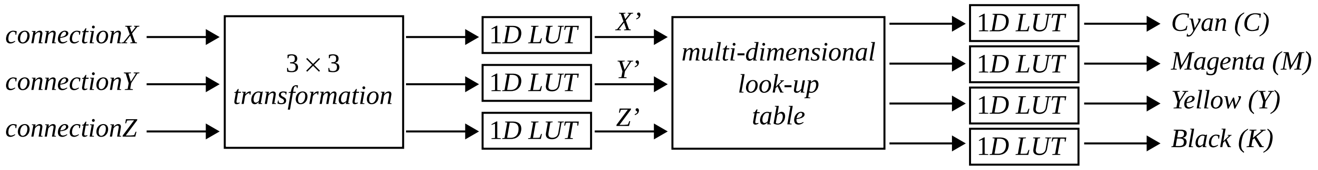

Figure 10. ICC output device profiles.

ing occurs between the basic levels attainable by printing one, two or several droplets at the same position, artefacts are less visible than in equal resolution bilevel printers. When dithering algorithms are used for the halftoning task, the dither threshold tiles should have oblique orientations so as to make the halftoning artifacts less visible. They should be designed so as to break up the inherent artifacts of variable dot size printers, such as for example continuous lines made up of elongated elliptic dots. In the case of error-diffusion in colour space, the introduction of dot over dot colour inhibiting constraints considerably reduces visual artifacts.

Improved resolution, introduction of dot-size modulation, printing with more than 4 process colors, advanced halftoning algorithms and calibrated scanning, previewing and printing paves the way to high-quality colour reproduction. Nevertheless, due to variations in the behavior of the scanners, displays and printers (inks, paper) over time, frequent recalibration is necessary. Dynamic recalibration will become possible when special additional colorimetric sensors will be included in scanners, monitors and output devices.

7. BIBLIOGRAPHY

[Bayer73] B.E. Bayer, "An Optimum Method for Two-Level Rendition of Continuous-Tone Pictures", *IEEE 1973 International Conf. on Communications,* Vol. 1, 26.11 – 26.15.

[Eschbach93] R. Eschbach, "Reduction of artifacts in error diffusion by means of input-dependent weights", *Journal of Electronic Imaging,* Vol. 2 , No. 4, Oct. 1993, 352–358.

[ICC95] ICC Profile Format Specification, Version 3.2, International Color Consortium, Nov. 20, 1995 (see http://www.color.org/).

[Klassen94] R.V. Klassen, R. Eschbach, K. Bharat, "Vector Error Diffusion in a Distorted Colour Space", Proc. of IS&T 47th Annual Conference, 1994, Reprinted in Recent Progress in Digital Halftoning, (Ed. R. Eschbach), IS&T, 1994, 63–65.

[Miller88] R.L. Miller, R.A. Morton (inventors), Image Processor with Smooth Transitioning between Dither and Diffusion Processes, US Patent 5'014'333, issued May 7, 1991, filed Jul. 21, 1988.

[Mitsa92] T. Mitsa, K.J. Parker, "Digital halftoning technique using a blue-noise mask, *J.Opt.Soc.Am.* A 9(11), 1992, 1920–1929.

[Ostromoukhov94] V. Ostromoukhov, R.D. Hersch, I. Amidror, "Rotated Dispersed Dither: a New Technique for Digital Halftoning", *ACM Siggraph'94,* Conference Graphics Proceedings, Annual Conference Series, 1994, 123–130.

[Ostromoukhov95] V. Ostromoukhov, Reproduction couleur par trames irrégulières et semi-régulières, PhD thesis no 1330, EPFL, Lausanne, 1995.

[Sullivan89] J.R. Sullivan, R.L. Miller, T.J. Wetzel (inventors), Color Digital Halftoning With Vector Error Diffusion, US Patent 5 070 413, issued Dec. 3rd, 1991, filed Oct. 10, 1989.

[Ulichney87] R. Ulichney, *Digital Halftoning,* The MIT Press, Cambridge, Mass., 1987.

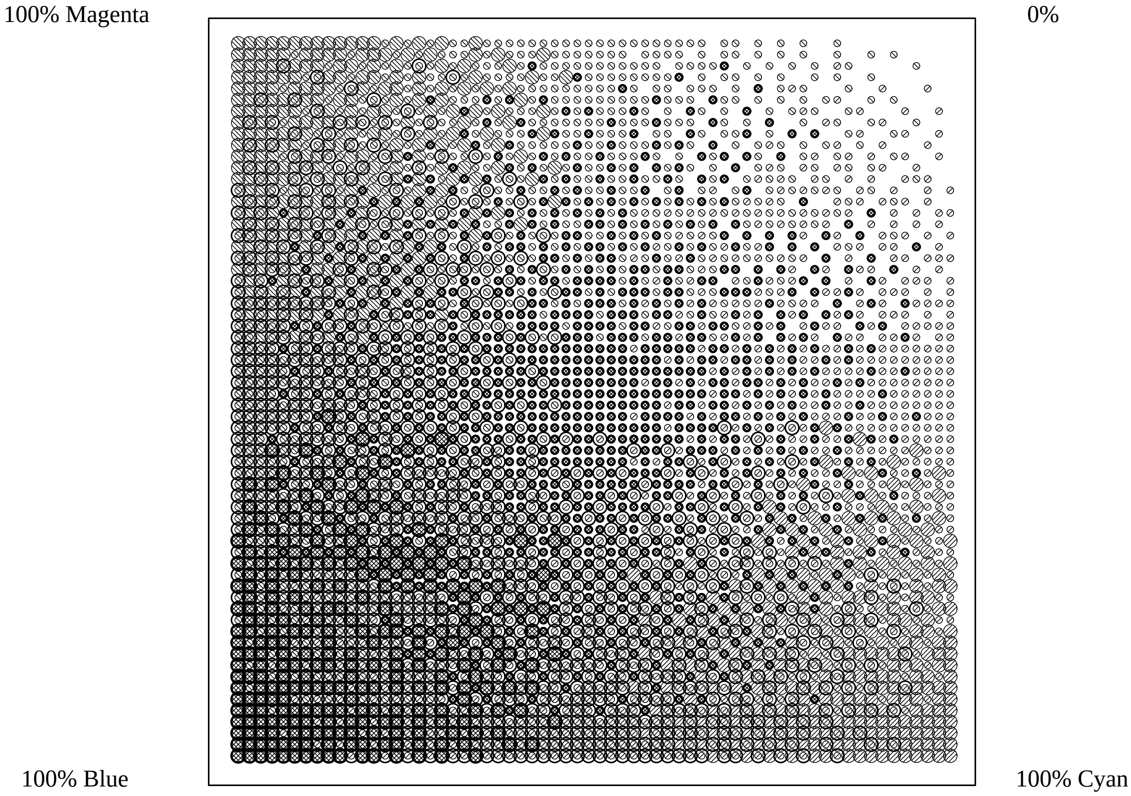

Figure 11. Variable dot size colour error-diffused cyan magenta wedge, with intensity-dependent inhibiting constaints (blue 50% allowed).

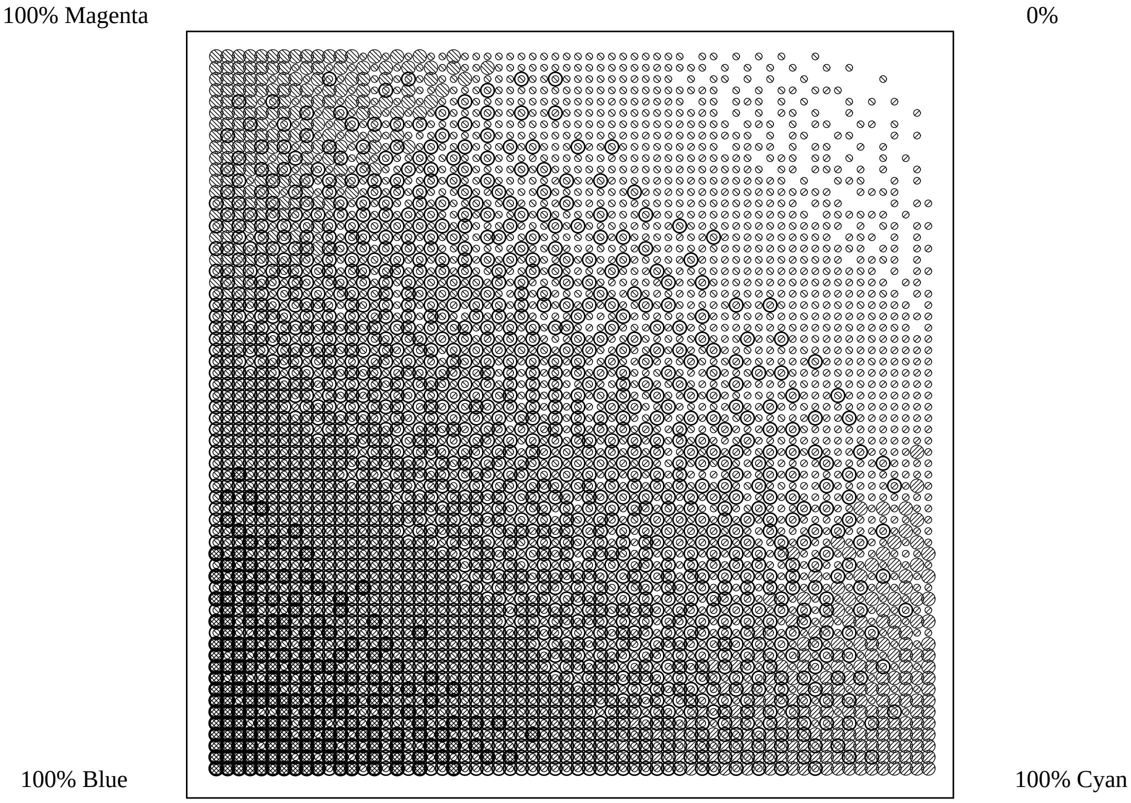

Figure 12. Variable dot size colour error-diffused cyan magenta wedge, with intensity-dependent inhibiting constaints (blue 50% discarded).