Artistic Halftoning - Between Technology and Art

Victor Ostromoukhov
MIT Computer Graphics Group
victor@graphics.lcs.mit.edu

Abstract

Halftoning, a technique for producing an illusion of continuous tones on bi-level devices is widely used in today’s printing industry. Different variants of halftoning - clustered and dispersed dither, error-diffusion, blue-noise mask, stochastic clustered dithering - have been extensively studied during last thirty years.

In the present paper, we shall focus our attention on different artistic halftoning techniques developed by the author at EPFL (Swiss Federal Institute of Technology, Lausanne, Switzerland). First, we shall explore Artistic Screening, a library-based approach, which has been presented at SIGGRAPH95. Contour-based generation of halftone screens effectively provides a new layer of information. We show how this layer of information can be used to convey artistic and cultural elements related to the content of the reproduced images. Artistic Screening is basically black-and-white technique. Multicolor and Artistic Dithering, presented at SIGGRAPH99, extends it to multiple colors. This technique permits to print with non-standard colors such as opaque or semi-opaque inks, using traditional or artistic screens of arbitrary complexity.

Finally, we shall draw a way of extension of traditional dithering to purely artistic domain: Digital Facial Engraving presented at SIGGRAPH99. Inspired by traditional techniques, we first establish a set of basic rules thanks to which separate engraving layers are built on the top of the original photo. Separate layers are merged according to simple merging rules and according to range shift/scale masks specially introduced for this purpose. Technically, digital engraving is based on the analogy between physical copperplate preparation process and traditional dithering technique. The analogy is so perfect that the art of digital copperplate engraving may be resumed as the art of building appropriate threshold structures. We illustrate the introduced technique by a set of black/white engravings, showing different features such as engraving-specific image enhancements, mixing different regular engraving lines with mezzotint, irregular perturbations of engraving lines etc. We introduce the notion of engraving style which helps to port the look and feel of one engraving to another.

Keywords: Color Printing, Halftoning, Artistic Rendering, Non-Photorealistic Rendering (NPR)

1. Introduction

Halftoning and screening techniques are aimed at giving the impression of variable intensity levels by varying the respective surfaces of white and black within a small area. Traditional techniques use repetitive screen elements, which pave the plane and within which screen dot surfaces define either white or black parts [Ulichney87]. As long as the screen element period is small, or equivalently, the screen frequency is high, distinct screen elements cannot be perceived by the human eye from a normal viewing distance. In our SIGGRAPH’95 paper [SIGGRAPH95], we explored a different approach. Instead of looking at the halftoning layer as a pure functional layer producing undesired artifacts, we propose a new screening technique which enables the shape of screen dots to be tuned. By creating artistic screens which may take any desired shape, screening effects, which up to now were considered to be undesirable, are tuned to convey additional information for artistic purposes. The approach we follow is somewhat related to the pen and ink illustration techniques where pen strokes are used for sketching illustrations, at the same time creating texture and intensities.

For artistic screening, we extend the dynamics of screen dot shapes by using more sophisticated artistic shapes as screen dots. We would like to have full control over the evolution of the artistic screen dot shape and at the same time offer a halftoning method which is competitive with regard to conventional high-resolution clustered-dot screening. We have sought our inspiration in the work of medieval artists [Critchlow89], who after having tiled the
Artistic screening

State of the art techniques for generating screen dot shapes are based on dither threshold arrays which determine the dot growing behavior. Since the dither threshold levels associated with the dither cells of a dither threshold array specify at which intensity the corresponding binary screen element pixels are to be turned on, the so generated screen dot shapes have the property of overlapping one another. In order to generate screen dots of any shape, which need not overlap one another and which may have self-intersecting contours, we propose a new way of synthesizing screen dot shapes. We define the evolution of screen dot contours over the entire intensity range by inter-
polating over a set of predefined fixed dot contours which define the screen dot shape at a set of fixed intensity levels. Once the evolving shape of the halftone dot boundary is defined exactly for every discrete intensity level, the screen elements associated with each intensity level are rasterized by filling their associated screen dot contours. This halftoning process distinguishes itself from previous halftoning methods described in the literature by the fact that the screen elements associated with every intensity level are precomputed and that no comparisons between original gray levels and dither threshold levels are necessary at image generation time.

Classical clustered-dot halftoning techniques rely on ordered dither threshold arrays. A dither threshold array is conceived as a discrete tile paving the output pixel plane. A dither threshold level is associated with each elementary cell of the dither threshold array. The succession of dither threshold levels specifies the dot shape growing behavior at increasing intensity levels (see Fig 1). Dither threshold levels can either be specified manually or algorithmically [Ulrichney87]. Previous algorithmic approaches for generating discrete dither arrays are based on spot functions. A spot function \( z = S(x,y) \) defines the dither threshold levels for a dither element tile defined in a normalized coordinate space \(-1 < x, y < 1\).

By discretizing this spot function, i.e, by computing its elevation at the coordinates of the centers of individual screen cells, and by numbering successive intersection points according to their elevations, one obtains the dither threshold array used for the halftoning process. The comparisons between given source image pixel intensity levels \( z \) and dither threshold levels determine the surface of a screen dot. With a given dither threshold array, the classical halftoning process consists of scanning the output bitmap, for each output pixel, finding its corresponding locations both in the dither array and in the grayscale input image, comparing corresponding input image pixel intensity values to dither array threshold levels and accordingly writing pixels of one of two possible output intensity levels to the output image bitmap. Since artistic screening is not based on dither matrices, we precompute the screen elements (halftone patterns) representing each of the considered intensity levels. The halftoning process associated with artistic screening consists of scanning the output bitmap, and for each binary output pixel, finding its corresponding locations both in the grayscale input image and in the screen element tile. The input image intensity value determines which of the precomputed screen elements is to be accessed in order to copy its bit value into the current output bitmap location.

Spot functions \( S(x,y) \) generating simple screen dot shapes can be described easily. More complicated spot functions for generating shapes such as the dot shapes described in Fig 2 are impossible to generate, since they cannot be described as single valued functions. In order to generate complicated dot shapes capable of representing known subjects (birds, fishes) or objects (letter shapes), we define the evolving screen dot shape by a description of its
contours. For this purpose, we introduce fixed predefined screen dot contours which are associated with specific intensity levels. Shape blending techniques are used to interpolate between those predefined screen dot contours at all other intensity levels. The fixed predefined contours, defined in a screen element definition space, are designed by a graphist using a shape drafting software package such as Adobe Illustrator. The graphist defines his contours in the screen element definition coordinate space of his preference. Curved contour parts may be described by polynomial splines. For convenience, we use a cubic Bézier spline given by its control polygon to define each curved contour part. In order to simplify the interpolation process, we also assume that each straight line contour part is also defined by a Bézier control polygon having its vertices aligned on the given straight line segment. The arrangement of contour parts in each of the fixed predefined contours governs the interpolation process.

Once the fixed predefined contour parts have been transformed from screen element definition to rendition space, the discrete screen elements may be generated for each discrete intensity level. For reproducing 256 intensity levels, the intensity interval between \( z=0 \) and \( z=1 \) is divided by 255 and intermediate screen dot contours are successively generated at intensity levels \( z=0, z=1/255, \ldots, z=255/255 \). At each discrete intensity, the screen dot contours are rasterized by applying well known shape rasterization techniques [Foley90]. In the case of self-intersecting dot contours or dot contours having at a single intensity level multiple intersecting contours, care must be taken to use a scan-conversion and filling algorithm supporting the non-zero winding number rule and generating non-overlapping complementary discrete shapes. Furthermore, the filling algorithm must be able to fill shapes becoming smaller and smaller until they disappear. Fig 2 shows the result with an artistic screen dot shape inspired by Escher's drawing, reproduced on a grayscale wedge. Small details, such as the wings of the bird, progressively fade out as the bird's shape size decreases.

Since screen tiles can be as large as desired, they can be conceived so as to cover either the whole or a significant part of the surface of the destination halftoned image. Such large screen tiles are divided into elementary subscreen shapes which may contain either identical or different shapes. For microlettering applications, each elementary subscreen shape may contain a different letter shape. By defining the mapping from screen element definition space to screen element rendition space as a non-linear transformation, smooth, highly esthetic spatial variations of the subscreen shapes can be attained. For example, conformal mappings transform a rectangular grid of screen element sub-shapes into the subshapes of a deformed grid following electro-magnetic field lines (Fig 3a). In that example, the conformal mapping is \( w = k (1 + z + e^z) \), where \( k \) is a real scaling factor, \( z \) represents complex points \( z = x + i y \) lying in the original \((x,y)\) plane and \( w = u + i v \) the corresponding complex points lying in the destination \((u,v)\) plane. A possible transformation is one that keeps the angle and modifies the distance of points from the center of the circle (fisheye transformation).

In high-quality graphic applications, the shapes of artistic screen dots may be used as a vector for conveying addi-
tional information. This new layer of information may incorporate shapes which are related to the image. When reproduced in poster form, the screen elements of the screening layer will become sufficiently large to produce the desired visual effect. This screening layer adds a touch of islamic culture to the reproduced image. In the last example, we show that artistic screening can bring new solutions for avoiding desktop counterfeiting. In Fig 4, we show that by using artistic screening techniques, microletters can be incorporated into the grayscale image. Furthermore, due to the conformal mapping function $w = \tan(g(z))$ between screen element definition and rendition spaces, a non-repetitive screen is created which cannot be scanned easily without producing Moiré effects.

3. Multi-Color Dithering

Multi-color dithering algorithm converts a barycentric combination of color intensities into a multi-color non-overlapping surface coverage. Multi-color dithering is a generalization of standard bi-level dithering. Combined with tetrahedral color separation, multi-color dithering enables printing images made of a set of non-standard inks. In contrast to most previous color halftoning methods, multi-color dithering ensures by construction that the different selected basic colors are printed side by side. Multi-color dithering is applied to generate color images whose screen dots are made of artistic shapes (letters, symbols, ornaments, etc.). Two dither matrix postprocessing techniques are developed, one for enhancing the visibility of screen motives and one for the local equilibration of large dither matrices. The dither matrix equilibration process corrects disturbing local intensity variations by taking dot gain and the human visual system transfer function into account. Thanks to the combination of the presented techniques, high quality images can be produced, which incorporate at the micro level the desired artistic screens and at the macro level the full color image. Applications include designs for advertisements and posters as well as security printing. Multi-color dithering also offers new perspectives for printing with special inks, such as fluorescent and metallic inks.

The reproduction of color images requires in the general case (1) separating the image colors (for example red, green and blue) into the set of available printable colors (for example cyan, magenta, yellow and black), (2) halftoning each of the printable color layer and (3) possibly calibrating the system to ensure that the printed colors are close to the original image colors.

Most existing color halftoning techniques are based on the independent halftoning of each of the contributing color layers. In offset and in many electrographic printers, the color layers are generated independently at angles of 15, 45, 75 degrees in order to avoid interferences between the cyan, magenta and black layers [Yule67, chapter 13]. Ink-jet printers often use error-diffusion to halftone each of the color layers independently.

Existing approaches for color separation with non-standard inks, for example the segmentation of the CIE-XYZ color space into a set of tetrahedra having as a common edge the black-white axis [Ostromoukhov93], the creation of correspondence tables between input colors and combination of output inks [Boill94], or the selection of an optimal subset of inks according to an objective function [Stollnitz98] consider that color layers can be printed independently one from another. This is true as long as the inks are transparent and as long as the dots of each screen are randomly positioned with respect to the dots of the other screens, as assumed by the Neugebauer equations [Rogers98].

If these assumptions are not true, halftoning each of the color layers independently may generate color shifts depending on the amount of superposition between screen dots of individual halftone layers. As long as the inks are transparent, the color shifts are small and the reproduction fidelity can be ensured by calibrating the printing device [Kang97], for example, in the case of process colors, by establishing a 3D mapping between CIE-XYZ coordinates and the output Cyan, Magenta, Yellow and Black surface coverage values [ICC98].

Some applications however require that different inks are always printed side by side without overlapping. This is for example the case when printing with opaque inks [Kueppers85], [Kueppers89]. Manufacturers of valuable documents, such as banknotes, identity cards and checks make often use of the high registration accuracy of their original printing equipment to create a graphic design which is both visually pleasant and difficult to imitate using standard printing processes [VanRenesse98, chapter 6].

We propose a multi-color dithering method which automatically enforces side by side printing of several color layers. Combined with color separation by tetrahedral interpolation, multi-color dithering enables printing images made of non-standard inks. In addition, we explore ways of creating large dither matrices to produce artistically screened color images, i.e. color images whose screen elements are made of artistic color shapes (letters, symbols,
3.1 Multi-color dithering and color separation.

Error-diffusion in color space [Sullivan91] may be used for side by side printing with non-standard colors, i.e. for ensuring that different inks are always printed side by side. Error-diffusion in color space is a straightforward extension of standard error-diffusion [Ulichney87]: given an input color vector (for example in RGB space) the closest printable color is selected and printed on the current output pixel. The 3D difference vector (error) between input and selected color is distributed among the current pixel’s neighbors [Klassen94]. A disadvantage of color error diffusion are the classical error-diffusion artifacts and the fixed dot size which in case of dot gain may induce a strongly non-linear tone reproduction behavior [Knox94].

The new multi-color dithering method we introduce here has the same advantages as standard dithering. It enables the creation of dither matrices of the desired size and orientation and the generation of dither threshold levels according to the desired dot growth behavior. Furthermore, small individual dither matrices can be assembled to form large super-matrices incorporating many intensity levels [Kang97, chapter 9].

3.2 Standard black-white dithering

Before explaining multi-color dithering, let us describe the basics of standard dithering for black-white (two colors). To simplify the explanations we assume that an input grayscale image with normalized darkness values between 0 (white) and 1 (black) is dithered by comparing at each output location corresponding input darkness and dither threshold values. If the darkness \( b(x) \) is higher than the dither threshold value \( t(x) \), then the output location is marked as black, else it is marked as white (Fig 5a).

Standard dithering converts a darkness value into a surface coverage. Conceptually, we can look at a given darkness value \( b \) as a percentage \( b \) of black and as percentage of \( (1-b) \) of white. The dithering process converts an input signal of darkness \( b \) to a surface coverage \( b \) of black and \( (1-b) \) of white.

3.3 Multi-color dithering

Let us now extend dithering to color. Suppose that we would like to print with 4 different color inks \( C_1, C_2, C_3, C_4 \) (called basic colors). At each pixel of the output pixmap, the color separation we use gives us the relative percentages of each of the basic colors (Fig 5b), for example \( d_1 \) of color \( C_1 \), \( d_2 \) of color \( C_2 \), \( d_3 \) of color \( C_3 \) and \( d_4 \) of color \( C_4 \). One of the basic colors, for example \( C_4 \), may be white.

Extending dithering to multiple colors consists in intersecting the relative cumulative amounts of colors \( d_1, d_1+d_2, \) and \( d_1+d_2+d_3 \) with the dither function \( t \) (Fig 5c). In the interval where \( d_1(x) > t(x) \), the output location will be...
printed with basic color \( C_1 \) (Fig 5). In the interval where \( d_1(x) + d_2(x) > t(x) \) and \( d_1(x) \neq t(x) \), the output location will be printed with basic color \( C_2 \). In the interval where \( d_1(x) + d_2(x) + d_3(x) > t(x) \) and \( d_1(x) + d_2(x) \neq t(x) \), the output location will be printed with basic color \( C_3 \). In the remaining interval where \( d_1(x) + d_2(x) + d_3(x) + d_4(x) > t(x) \) and \( d_1(x) + d_2(x) + d_3(x) \neq t(x) \), the output locations are printed with basic color \( C_4 \). Multi-color dithering therefore converts the relative amounts \( d_1, d_2, d_3, d_4 \) of basic colors \( C_1, C_2, C_3, C_4 \) into relative coverage percentages and ensures by construction that the contributing colors are printed side by side.

Color dithering is generally applied with 4 basic colors, since a point in 3D color space within the printer’s gamut can be described by a barycentric combination of 4 colors.

### 3.4 Color Separation

When given a set of basic colors (inks), each specified by its tri-stimulus value in a given 3D color space (RGB or CIE-XYZ space), the 3D volume covered by these basic colors, called the printable gamut, can be segmented into a set of mutually adjacent tetrahedra [Hung93]. The vertices of each tetrahedron correspond to four neighboring basic colors. Many tetrahedrizations of a point set in 3D exist; there is however only a single tetrahedrization, which ensures that the enclosing sphere of a tetrahedron does not include another tetrahedron. Properties of tetrahedrizations and methods of construction are well described in the literature [Nielson97, chapter 20].

Color separation of an input tristimulus value (RGB or CIE-XYZ) is obtained by locating in the selected color space the tetrahedron enclosing the given tri-stimulus value and by finding the barycentric coefficients \( d_1, d_2, d_3, d_4 \) used to express the input tristimulus value as a linear combination of the tetrahedron's vertices. These barycentric coefficients give the relative amounts of basic colors \( C_1, C_2, C_3 \) and \( C_4 \) used to reproduce the input tristimulus value.

As an illustration for tetrahedral decomposition and interpolation, let us consider an RGB color cube whose vertices correspond to the basic colors black, red, green, blue, cyan, magenta, yellow and white (Fig 6a). A color wedge (Fig 6b) with vertices close to cyan, blue, red and yellow is reproduced by multi-color dithering using the available set of basic colors. This color wedge, a planar slice in RGB space (Fig 6a), intersects all tetrahedra into which the
RGB cube is decomposed. In each tetrahedron, the corresponding color wedge part is reproduced using the 4 basic colors associated with the tetrahedron’s vertices.

The dither matrix used for producing Fig 6b has been obtained by discretizing and renumbering an egg crate function [Kang97, section 9.3.1, Fig. 9.6a]. The resulting dithered color wedge incorporates ring shaped screen elements. All contributing basic colors are printed side by side.

If color reproduction fidelity is an issue, a calibrated input device (for example a scanner) is needed which provides a mapping from RGB input device values to CIE-XYZ device-independent values. Color separation in CIE-XYZ space is possible by tetrahedral decomposition of the volume formed by the measured CIE-XYZ values of the basic colors (inks + paper white). Colors close to the basic colors will be reproduced correctly. However, colors requiring the combination of several basic colors may deviate from their desired CIE-XYZ value, depending on various parameters such as printer registration accuracy, dot gain and ink density distribution of printed screen dots. The present contribution does not deal with printer color calibration. However, a possible printer calibration may be achieved by printing a large number of samples covering the printer’s gamut and by measuring their CIE-XYZ values in order to build a 3D calibration table providing a mapping between device independent input CIE-XYZ values and output space “predicted” CIE-XYZ values (CIE-XYZ values predicted by linear interpolation within each tetrahedron).

When printing with transparent inks, we may want to use the superposition of a pair of selected inks as an additional basic color. This is easily done by measuring its CIE-XYZ tristimulus values and incorporating the superposition of the two inks as a new color $C_j$ into the set of available basic colors.

In this paper, we generally assume that input color values are within the range of printable colors. If the input color values are not located within the range of printable colors, a gamut mapping method must be applied. Several gamut mapping methods are known to produce convenient results [Stollnitz98].

### 3.5 Synthesis of partially continuous, partially random dither matrices

For the purpose of artistic color screening, we need to generate dither matrices incorporating visually appealing symbols and ornamental motives. At low intensities however, one process color (for example black) may become very dominant and the corresponding screen motive may become too large to be recognized.

![Continuous dither matrix](image1)

(a) Continuous dither matrix

![Partially continuous, partially random dither matrix](image2)

(b) Partially continuous, partially random dither matrix

Fig 7 Grayscale wedge (a) with a standard dither matrix and (b) with a mixed dither matrix made of partially continuous, partially random threshold levels.

In order to solve this problem, we propose to generate dither matrices which produce partially ordered and partially random screen dots. To avoid enlarging the screen motive too much when strongly increasing the color surface coverage, we allow the background to contribute to the corresponding color. In the case of a black screen motive, from a certain darkness level, the background becomes successively darker. Thus instead of enlarging the motive shape, its contrast to the background becomes successively less pronounced, until it finally vanishes (Fig 7b). One
may observe that with a partially random screen, the desired screen motive remains visible over a larger intensity range than with a standard clustered screen. The final result however depends on dot gain: if dot gain is small, this effect is clearly visible. When dot gain is important, this effect is counterbalanced by the decreased contrast due to dot gain at high ink surface coverages.

![Fig 8 Producing a dither matrix comprising partially continuous, partially random threshold levels.](image)

It is relatively easy to create dither matrices generating partially clustered, partially random screens. One may first generate a continuous screen function representing the clustered part of the screen dot and sample it at each dither matrix cell (Fig 8a). Then a very small amplitude noise can be applied either to all dither matrix cells or to dither matrix cells having low dither values (Fig 8b). Dither matrix cells are successively numbered according to their respective dither values (histogram equalization). Their ordinal numbers represent, after normalization, their dither threshold levels (Fig 8c).

4. Digital Engraving

Engraving is among the most important traditional graphical techniques. It first appeared in the fifteenth century as illustrative support for budding book-printing, but very quickly became an art in its own right, thanks to its specific expressive power. Actually, four main classes of engraving are used by artists: letterpress or relief printing, intaglio or in-hollow printing, silk screen process and lithography, with several different techniques in each class. The history of printmaking was punctuated by prosperous periods of techniques which later declined for various reasons. Facial engraving is one such example. Extremely popular in seventeenth and eighteenth centuries, when photography did not exist, this wonderful art became almost unused, due to the extreme technical demands that it made on the engraver. Professional copperplate engravers are rare today, and the cost of true engravings is simply too prohibitive to be used in everyday printing. At the same time, traditional facial engraving has no doubt very specific appeal: its neat, sharp appearance distinguishes it advantageously from photos [Brunner84].

To appreciate the graphical impact of engravings it’s enough to compare the engraved portraits in the Wall Street Journal with portraits in other newspapers produced with traditional impersonal screening.

Does it mean that this enjoyable art is condemned to disappear for purely economical reasons? We don’t think so. We do believe that computer graphics can transform traditional engraving into a digital art. Already in the past, considerable effort has been made to transform traditional pen-and-ink illustrations into digital form [Winkenbach96], [Salisbury97], to make digital watercolors [Curtis97], various line art drawings [Dooley90a], [Dooley90b], [Elber95a], [Elber95b], [Elber98], [Lansdown95], woodcut imitations [Mizuno98], and expressive painting [Haeberli90], [Meier96].

In the present contribution, we establish the basis for digital engraving, and more specifically, for facial engraving. Our goal is very precise: starting from a digital photo of a person, to be able to reproduce it faithfully, relying on technical achievements and techniques that traditional engravers used in the past. The resulting digital engraving should be visually pleasant, and the person must be recognizable.
We tried to reuse the graphical techniques at the disposal of traditional artists who did not know anything about parametric surfaces. Visibly, the main rule for creating facial engravings was: the directions of the engraving lines should somehow follow facial features. Somehow means: just loosely related to large facial surfaces. Long curves traversing several features were very welcome [Ivins88]. This rough figure-hugging layered engraving, combined with cross-etching, forms the basis for our approach.

Fig 9 (a) A micro-photography of the graver’s tip cutting sharply into the copperplate showing a shavings lifted by the bevel.
(b) Virtual “universal” copperplate cut at different heights, producing furrows of different widths. Putting ink into the furrows imitates the true copperplate engraving printing process.

Technically, the micro-photography of the copperplate engraving process shown in Fig 9a gives us an important insight. The width of the furrows produced by the tip of the graver will influence the engraving line width. The furrows have a specific triangular shape clearly visible in Fig 9a (please note: they are concave, not convex). Based on this observation, we could build a virtual “universal” copperplate as shown in Fig 9b. This “universal” copperplate can be cut at different heights thus producing furrows of different widths. By putting ink into the furrows we imitate the true copperplate engraving printing process. But at the same time the process described here is nothing other than conventional dithering, well-known in computer graphics [Ulichney87]. Simply, the term of virtual “universal” copperplate stands for the threshold matrix (threshold levels corresponding to the height of our “universal” copperplate), while cutting and inking stands for comparison between the input signal level and the current threshold value, thus producing a black or white output signal. The analogy is so perfect that the art of digital copperplate engraving may be resumed as the art of building appropriate threshold structures looking like the “universal” copperplate in Fig 9b. Once this threshold structure is built, the rendering may be done using conventional dithering software.

4.1 Basic rules

Our goal is presently to develop a technique for building separate engraving layers, to transform them in order to follow the desired directions and finally to superimpose these layers, thus forming various cross-etching and smooth transitions between different parts of the artwork rendered by different engraving layers.

In this section we develop a simple and straightforward technique based on the most regular support for any line art: a sequence of equidistant straight lines defined on a given region. We call such a sequence the basic engraving layer. After appropriate morphing, the basic engraving layer roughly corresponds to the main etching technique used in traditional engraving. Later, we shall see how this relatively simple technique may be extended to imitate
more sophisticated etching techniques such as irregular lines or mezzotint.

4.2 Building separate layers

Let us suppose that the basic (non-transformed) engraving layer is defined on a unit square in the $uv$ coordinate system, as shown in Fig 10a. The layer is built of a sequence of threshold structures made up of uniformly spaced waves, as shown in Fig 10c. The cross-section of each wave has a simple saw shape, directly inspired by the shapes of the furrows in Fig 9a. The directions of the waves are not necessarily parallel to axes $u$ and $v$.

The basic engraving layer can easily be transformed into a warped layer (Fig 10d). There are many ways to perform such a transformation (see for example [Foley90, Chapter 11], or [Wolberg90]). We have chosen a very simple and intuitive way to define the transformation: it is given by four parametric border curves $U_j(u), U_2(u), V_j(v)$ and $V_2(v)$ in the image space $xy$ (see Fig 10b); the parameters $u$ and $v$ lie in the range $[0..1]$. The main advantage of our construction consists in the fact that the border curves $U_j(u), U_2(u), V_j(v)$ and $V_2(v)$ can easily be created on top of the source photo to be rendered by the engraving, using popular powerful tools such as Adobe Illustrator. As we want the engraving lines to somehow follow the directions of the features in the original image, we naturally build the border curves taking into account the borders of the features such as nose, eyes, cheeks, lips etc. The process of building the border curves becomes more clear by observing a concrete example of building an engraving style shown in the top row of Fig 13.

The border curves are built of an arbitrary number of straight line and/or Bézier curve segments (two segments for the curves $U_j(u), V_j(v), V_2(v)$, and three segments for the curve $U_2(u)$ in Fig 10b). Each curve has to be re-parametrized in order to preserve the uniformity of the curve’s Euclidean length when the parameter uniformly walks in the range $[0..1]$ (see for example [Farin90], section 9.4). The re-parametrization permits a smooth and uniform interpolation between the curves built of a different number of segments, of different lengths.

We suppose that our parametric curves $U_j(u), U_2(u), V_j(v)$ and $V_2(v)$ form the interior of the closed quadrilateral path $\mathcal{T}$, $U_j(0)=V_j(0), U_j(1)=V_2(0), U_2(0)=V_j(1)$ and $U_2(1)=V_2(1)$. Any point $P(u,v)=(x,y)$ inside the patch $\mathcal{T}$ can be defined as a function of parameters $u$ and $v$ (after re-parametrization), by applying the linear interpolation between the curves $U_j(u)$ and $U_2(u), V_j(v)$ and $V_2(v)$, taking into account the correction terms $L(v)$ and $R(v)$ due to deviations by left and right curves $V_j(v)$ and $V_2(v)$:

$$P(u,v) = (1-v) \cdot U_1(u) + v \cdot U_2(u) + (1-u) \cdot L(v) + u \cdot R(v)$$

$$L(v) = V_1(v) - (1-v) \cdot V_1(0) - v \cdot V_1(1)$$

$$R(v) = V_2(v) - (1-v) \cdot V_2(0) - v \cdot V_2(1)$$

where $0 \leq u, v \leq 1$

A uniform grid defined in the parametric space $uv$ is transformed using the transformation (1) into a warped grid, as shown in Fig 10a and Fig 10b. Any basic engraving layer defined in the parametric space is accordingly transformed into a warped engraving layer in the image space (Fig 10c and Fig 10d).

4.3 Superimposition of separate layers

Now, our task is to establish basic rules for the superimposition of several transformed engraving layers, each of which is a simple matrix of threshold values. For the sake of simplicity, we consider that the threshold values are in...
the range between 0 and 1; one entry of this threshold matrix corresponds to one pixel of the resulting image. We assume that there is a one-to-one correspondence between the matrix of threshold values representing the engraving layer and the resulting image. No geometrical transformation is performed at this stage, only the threshold values may be affected.

Table 11: Merging (superimposition) modes

<table>
<thead>
<tr>
<th>merging mode</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>copy CL</strong></td>
<td>[ T_{RL}(x,y) = T_{CL}(x,y) \cdot S(x,y) + D(x,y) ]</td>
</tr>
<tr>
<td><strong>smaller CL</strong></td>
<td>[ T_{RL}(x,y) = \min(T_{RL}(x,y), T_{CL}(x,y) \cdot S(x,y) + D(x,y)) ]</td>
</tr>
<tr>
<td><strong>bigger CL</strong></td>
<td>[ T_{RL}(x,y) = \max(T_{RL}(x,y), T_{CL}(x,y) \cdot S(x,y) + D(x,y)) ]</td>
</tr>
<tr>
<td><strong>multiply CL</strong></td>
<td>[ T_{RL}(x,y) = T_{RL}(x,y) \cdot (T_{CL}(x,y) \cdot S(x,y) + D(x,y)) ]</td>
</tr>
<tr>
<td><strong>add CL</strong></td>
<td>[ T_{RL}(x,y) = T_{RL}(x,y) + (T_{CL}(x,y) \cdot S(x,y) + D(x,y)) ]</td>
</tr>
</tbody>
</table>

The superimposition of several layers is performed sequentially, one layer after another. For this reason, it is important to define a set of basic rules for superimposing two layers, the extension to several layers being straightforward. Each engraving layer, before the superimposition, may undergo two range transformations: it may be scaled (range scale) and raised or lowered (range shift):

\[ T(x, y) = T(x, y) \cdot S(x, y) + D(x, y) \]

where range scale values \( S(x, y) \) and range shift values \( D(x, y) \) are two matrices of the same dimensions as the matrix of threshold values \( T(x, y) \) which forms the transformed engraving layer. We build the range scale and range shift matrices using popular image-manipulation tools such as Adobe Photoshop on top of the feature borders in the original image. A trimming operation may be needed when the resulting threshold value \( T(x, y) \) goes beyond the range \([0,1]\). Trimming may be performed either after every layer superimposition operation, or at the very end of the whole sequence of superimpositions.

As we see, superimposing engraving layers consists in consecutively merging the current engraving layer (CL) into the resulting one (RL). Once merged, the current engraving layer disappears as an independent entity. The merging is performed according to the merging mode. Table 11 enumerates some merging modes, among the most important ones. This list is not exhaustive: additional modes may be added if needed.

Fig 12 illustrates the use of merging modes. The sample image contains two parts: a uniform gray ramp and four flat patches whose respective intensities are 1/8, 3/8, 5/8 and 7/8. It may be noticed that the “copy” “smaller” and “bigger” modes are by far the most useful for engraving purposes. In fact, the “copy” mode serves to initialize the resulting engraving layer for the very first merging operation. The “smaller” mode produces cross-etching which is very close to traditional cross-etching known in the art. Please note that in the dark area this mode does not produce continuous lines. The “bigger” mode is complementary to the “smaller” mode: it produces continuous lines in the dark areas, and discontinuous ones in highlights. Finally, as can be observed in the bottom line of Fig 12, a judicious combination of appropriate merging mode with individual layer range shift may produce very interesting technical effects: one particular layer may become apparent only in a desired subrange of gray.

4.4 Equilibration

One may notice that the tone reproduction curve of two superimposed layers is no longer linear, even if both layers forming the superposition have a linear curve reproduction behavior. This means that when several layers are superimposed, the resulting engraving may appear locally darker or lighter than what is expected. In order to cope with this phenomenon, we have to equilibrate the resulting threshold structure: when the dithering process is applied, this modified threshold structure must produce visually uniform gray for a uniform input signal of any intensity, independently of the number of layers and the superposition rules that have been used. A simple histogram equalization will not work because of its global nature. Instead we need local model-based histogram equalization, taking into account the dot gain of the printing engine and the characteristics of the human visual system. This dramatically improves the quality of the result because the engraving lines are often very thin, and their visual
The threshold matrix to be equilibrated is corrected separately, for different input signals (flat uniform surfaces). For each input level, standard dithering is performed in order to obtain the bitmap to be sent to the printing device that we modelize. After applying dot-gain correction and low-pass filtering which simulates the human visual sys-

Fig 12 Merging modes. Left column: the threshold structure obtained by superimposing threshold structures L1 and L2 using different merging modes. Right column: a sample image (the topmost image) rendered using the threshold structures shown in the left column.

impact should be carefully taken into account. Let us outline the equilibration process in a few words.
tem, we get the model approximation of the printed and perceived surface corresponding to the uniform input signal of a given intensity. The response of the model is usually non-uniform. According to the local discrepancies between the input and model-based output, we locally modify the threshold matrix, then redo the whole cycle of the output model-based simulation described before. After a few iterations, we obtain the threshold matrix which produces a reasonably uniform output. Various parameters of the model are adjusted using the measurements of real test prints. Once all correction terms for the threshold matrix has been calculated for all input intensities, the resulting corrected threshold matrix may be calculated. For performance reasons, we calculate the correction terms only for a few (16) input intensity levels spread uniformly throughout the whole intensity range, the rest being linearly interpolated.

The results of our equilibration technique are very satisfactory for the digital engraving presented in this article, as well as for other dithering techniques.

4.5 Black and white facial engraving

Now, let us illustrate the techniques described in the previous section showing an example of black and white engraving of the head of Michelangelo’s Julien de Médicis. Five separate engraving layers for various parts of the face have been created (see Fig 13, upper row). The borders of the patches are arranged in such a way that the patches’ grids loosely follow the key features of the image: the nose line, the cheek profile, the eye shape etc.

The middle row of Fig 13 shows the range shift matrices associated with the corresponding engraving layers. Here we use the following convention: mid-gray corresponds to a zero range shift, white corresponds to a full-range raise ($D(x,y)=+1$) and black corresponds to a full-range lowering ($D(x,y)=-1$), and the gradations between these three states mean intermediate range shifts. It may be noticed that the pixels whose range shift values in a given layer are +1 (white in our convention) do not participate in building the resulting engraving layer. In such a way, the range shift matrices act as transfer masks between the current and the resulting engraving layers. Smoothness of gradation between the areas where shift $D(x,y)=0$ and $D(x,y)=1$ determines the nature of the fusion between several layers: abrupt boundaries define neat junctions between layers whereas smooth boundaries determine a very progressive fusion between the layers. In our example, the boundaries of the area where $D(x,y)=0$ in layer L2 (around the right eye) are relatively abrupt; the corresponding engraving shows a pretty neat junction between layers L1 and L2, as can be seen in Fig 14b. On the contrary, the junction between layers L1 and L3 in the area of the left eyebrow is more progressive - and the resulting engraving shows some overlapping between these layers (Fig 14b). It’s according to the artist’s taste that the degree of smoothness between the layers may be determined.

In the building process shown in Fig 13 we used only the “copy” and “smaller” merging rules between successive layers. For all layers, the scale values $S(x,y)=1$ (i.e. no scaling).

Fig 14 shows the resulting engraving achieved using the engraving style shown in Fig 13 (the only difference between the engraving layers shown in Fig 13 and Fig 14 consists in dividing by two all etching frequencies in the former, for the sake of visibility). It may appear quite surprising that a relatively simple technique described here produces such a decent result.

Furthermore, the visual quality of the engraving may be improved by applying more sophisticated techniques, and especially various cross-etchings, as shown in Fig 15. The principle of cross-etching shown in Fig 15a and enlarged in Fig 15b is very simple: from the same patch $\mathcal{T}$ we generate two engraving layers: in the first one, the engraving crests are oriented following the curvilinear coordinate $u$, and in the second one - following the curvilinear coordinate $v$.

Fig 15c shows that the “bigger” rule for superimposition of the layers may play an additional role.

4.6 Mixing etching and mezzotint

Fig 15 illustrates another expressive tool traditionally used by some engravers: mixing regular etching and mezzotint. In this example, the entire layer L5 has been replaced by a specially designed mezzotint engraving layer. Enlargement in Fig 15d shows that our simulated mezzotint is a relatively good match for the traditional aquatint texture shown in micro-photography in Fig 15e. We implemented our mezzotint following the description in [Ostromoukhov94]. Let us note the specifically warm appearance of such an engraving, as well as the acute contrast between regular and “stochastic” parts. Traditional engravers in the past often made the most of the additional expressive power of such juxtapositions.

4.7 Real photos

Compared to the photographs of sculptures, real-people photos may present some additional difficulties: the con-
The contrast of the key features such as the profile line, the nose, lips and eye contours may be insufficient. Real people may have some particular traits and features that one may wish to hide or stress. Typically, the wrinkles in female portraits often become almost invisible, whereas the vigorous lines on the male portraits are sometimes stressed. Also, special features such as glasses, moustaches, earrings etc. may require special technical attention. Real photos may need an additional pre-processing phase which would perform traditional cosmetic arrangements, contrast and edge enhancements, as well as engraving-specific enhancement techniques.

Engraving offers a large set of expressive tools for visual contrast enhancement. We have already mentioned the effect of mixing mezzotint and regular etching. Further examples are shown in Fig 16. The contrast between the glasses and the face is achieved by an abrupt change of the direction and frequency of the etching. As we explained before, such an effect can very easily be obtained by designing appropriate engraving layers, and by abrupt borders in the range shift masks.

Another tool for stressing particular features like a contour line or a small detail is illustrated in Fig 16c. Here, a small additional layer has been added in order to accentuate the nasolabial fold. It may be noticed that additional engraving strokes parallel to the feature did not modify the impression of the gray level of this particular area: our
equilibration process made the other lines locally thinner. On the contrary, the feature itself (nasolabial fold) appears much more contrasted.

4.8 Color engraving

Traditional engraving is essentially black and white art. This does not remove the temptation to experiment with colors. The simplest color scheme would be to use the same resulting engraving layer for all three RGB components: it is known as in-phase color printing. It certainly works but the achieved visual effect is not very different from black and white gravure.

One may expect a more advanced visual effect when the different color layers use different orientations. A simple but efficient scheme has been experimented: for the Red and Blue color planes, we use the engraving layers with line orientation along the parametric axis $u$, whereas for the Green color plane, we use the same engraving layers but with line orientation along parametric axis $v$. This technique may be called engraving by orthogonal (in parametric space) color lines. Yet another possibility is to mix normal regular engraving for some color planes with the mezzotint for the others. We hope that more advanced color mixing / texture mixing schemes may introduce unexpected, beautiful visual effects.

5. Conclusions

We have presented a family of halftoning techniques which extend traditional highly accurate methods of reproduction, in order to convey some artistic information.

Artistic Screening is a new halftoning technique, where screen elements are composed of artistic screen dot shapes, themselves created by skilled graphists. Fixed predefined dot contours associated with given intensity levels determine the screen dot shape’s growing behavior. Screen dot contours associated with each intensity level are obtained by interpolation between the fixed predefined dot contours. User-defined mappings transform screen elements from screen element definition space to screen element rendition space. These mappings can be tuned to produce
various effects such as dilations, contractions and non-linear deformations of the subscreen element grid. By choosing an appropriate mapping, images can be rendered while ensuring a highly esthetic behavior of their screening layer.

Multi-color dithering is a generalization of standard bi-level dithering. Combined with tetrahedral color separation, multi-color dithering enables printing images made of a set of non-standard inks. In contrast to most previous color halftoning methods, multi-color dithering ensures by construction that the different selected basic colors are printed side by side. We extended bi-level dithering to multi-color dithering and explored multi-color dithering in the context of artistic color screening. To generate high-quality dithered color images incorporating artistic screen motives, we developed two dither matrix postprocessing techniques, one for enhancing the visibility of screen motives and one for the local equilibration of large dither matrices. By combining within the same dither matrix continuous threshold levels for the screen motive and randomly distributed dither threshold levels for the background, we enhance the visibility of the generated screen shapes at high ink saturation levels. The dither matrix equilibration process we propose for avoiding disturbing local intensity variations takes both the physical behavior of the printer, i.e. the dot gain and the human visual system modulation transfer function into account.

Thanks to the combination of the presented techniques, high quality images can be produced, which incorporate at
the micro level the desired artistic screens and at the macro level the full color image. Possible applications include innovative designs for publicity and posters. Security printing may make use of multi-color dithering in order to print side by side with non-standard inks at a high registration accuracy. In that context, many new issues arise, such as the selection of the inks and the design of screen motives making the original very difficult to replicate, both by professional craftsmen and by simple color photocopying. Multi-color dithering also offers new perspectives for printing with special inks, such as fluorescent and metallic inks.

A very simple technique for producing digital engravings has been presented. The proposed system is based on the analogy between the “universal” copperplate which imitates the true copperplate engraving technique and conventional dithering. The art of digital copperplate engraving may be resumed as the art of building appropriate threshold structures. We have developed the basic technique for building separate engraving layers (threshold structures) which roughly follow the features of the original image, as well as the rules for merging them together. The resulting threshold structure is equilibrated in such a way that it generates a visually uniform output for a uniform input signal of any intensity. Applied on an input digital photo, using a standard dithering algorithm, such a threshold structure generates a reasonably faithful reproduction, which imitates traditional engraving. Several enhancement techniques, specific to engraving, have been proposed. The important notion of engraving style which comprises a set of separate engraving layers together with range shift and scale masks has been introduced. Engraving styles make it easier to adapt the look and feel of an engraving of person A to an engraving of person B. Finally, a simple color extension has been proposed.

References

Fig 16 This engraving shows various enhancement techniques: an abrupt change of the orientation and frequency of the etching (in glasses), or an additional layer for sharper appearance of the nasolabial fold in (c).