

Space Filling Curve Dither with Adaptive Clustering

LUIZ VELHO^{1,2}
JONAS DE MIRANDA GOMES²

¹IMPA-Instituto de Matemática Pura e Aplicada

Estrada Dona Castorina, 110
22460 Rio de Janeiro, RJ, Brasil
lvelho|jonas@visgraf.impa.br

²University of Toronto
Department of Computer Science
Toronto, Ontario, Canada
lvelho@vis.toronto.edu

Abstract. In this paper we investigate the problem of varying the size of dots in digital halftoning processes. We show that the dot size is the main factor determining the trade-off between tonal and spatial resolution in the reproduction of gray scale images. We discuss the control of pixel aggregation in the context of dithering algorithms. In particular, we extend the dither with space filling curves method, introduced in (Velho & Gomes, 1991), to change the size of pixel clusters according to local characteristics of the image. This makes possible to achieve optimal rendition of gray shades while preserving image details.

Keywords: Electronic Printing, Digital Halftoning, Dithering Algorithms, Space Filling Curves, Adaptive Clustering.

1 Introduction

The reproduction of gray scale images in bilevel graphic display devices is achieved through a process called *halftoning*. Given a gray scale image it consists in generating a binary image which causes the impression of continuous-tone variation. This process can be analog or digital. Analog halftoning is done using photographic means to generate high contrast halftone screens. Digital halftoning is done using a technique called *dithering* to determine the binary state (black or white) of the elements of the output image. Dithering algorithms try to distribute the black and white areas in such a way that the input and output images are perceptually as close as possible, within the physical limitations of the display equipment.

The dithering methods can be subdivided into two main groups according to the type of images they produce (Ulichney, 1987). *Dispersed dot* dithering methods generate images in which black and white pixels are evenly distributed throughout the image area. *Clustered dot* dithering methods generate images in which black and white pixels are concentrated together forming clusters. These two approaches serve different purposes and are suitable for different classes of display devices. Dispersed dot dithering is

used with low resolution devices such as bitmap displays while clustered dot dithering is used with high resolution devices such as digital phototypesetters.

The early work in this area focused mainly in dispersed dot dithering algorithms. This was motivated at first by the introduction of graphic display devices and subsequently by the popularity of bitmap workstations. The classical dithering algorithms are the *ordered dither* algorithm (Limb, 1969) and the *Floyd-Steinberg* algorithm (Floyd and Steinberg, 1975). The ordered dither employs a spatial look-up table to perform the quantization of intensity values, while the Floyd-Steinberg method modulates the quantization threshold to distribute the local discretization error. Donald Knuth combined these two approaches in his *dot-diffusion* algorithm (Knuth, 1987). A good survey of the initial research in dithering can be found in (Jarvis, Judice and Ninke, 1976).

The advent of high resolution hardcopy devices and applications to digital printing created a need for clustered dot dithering algorithms. The standard solution is a variation of the ordered dither with clustering which approximates analog halftone screens (Adobe Systems, 1985). An alternative solution is based on space filling curves to create pixels aggregates similar in structure to film grain (Velho and

Gomes, 1991). The main differences between these methods are related to the nature of the patterns they create and the degree of control over the cluster size allowed by them. The *clustered-dot ordered dither* generates periodic patterns with dot sizes determined by a square integer sublattice of the image. The *space filling curve dither* generates aperiodic patterns with dot sizes that can vary in one pixel increments. This method is more flexible than the clustered-dot dither specially with respect to changes in the dot size of the halftone screen.

A common characteristic to all current dithering methods is that the dot size is determined a priori and remains constant for a given image. Although this allows the resolution of the halftone screen to conform to the requirements of the imaging system, it does not give the best possible rendition of image detail.

In this paper we investigate the effect of variable sized dots in dithering algorithms. By changing the clustering of pixels based on local characteristics of the image we are able to achieve the best compromise between rendition of gray shades and image detail. This idea is implemented extending the space filling curve dither to handle space varying pixel clusters. It is the first dithering algorithm capable of generating halftone screens adapted to the characteristics of the image. It is suitable for displays with a variable dot sizes.

The remaining of the paper is organized as follows: Section 2 gives an analysis of the halftoning process. Section 3 reviews the space filling curve dither algorithm. Section 4 discusses the dither with adaptive clustering. Section 5 describes the implementation of the method extending the space filling curve algorithm. Section 6 shows results of the application of the method. Section 7 summarizes with concluding remarks and prospects of future research.

2 The Halftoning Process

The display of continuous-tone images in bilevel devices implies in the quantization of its intensities to one of two levels (black or white). This is an extreme case of discretization in which continuous intensity values must be converted into a discrete set of values (in this case, only two). This operation may cause a loss of information which is estimated by the *quantization error*. The quantization error is defined as the difference between the continuous and discrete values. It can be measured at a particular image element, region or even the whole image.

The halftoning process compensates the quantization error inherent in the bilevel discretization using a *dithering technique*. In the regular quantiza-

tion operation continuous input values are compared with a fixed quantization threshold and the output is set to 0 or 1 depending on whether the continuous value is respectively less or greater than the threshold value. It is clear that this introduces an artificial contour separating the regions of intensities less and greater than the threshold. The dithering technique makes this threshold value *spatially variable*. Using such strategy, the quantization error incurred in one pixel of the image can be compensated in the quantization of the pixels nearby. The result is that the average quantization error over small regions of the image can be made very small and, as a consequence, contouring artifacts are minimized.

This technique implies in a trade-off between spatial and tonal resolution. As the quantization error is spread over larger areas of the image more tones can be represented. Gray levels are rendered in this way as patterns of black and white pixels. On the other hand, if dithering avoids false contours it eliminates at the same time high frequency information contained in the image. The process transforms also true intensity boundaries into patterning features. In summary, intensity variation is displayed at the cost of poorer rendition of fine details. At first it may seem that this problem is inherent to the halftoning process and totally unavoidable. As will be shown next, this is not really the case.

The usual expedient to minimize the loss of image detail is to perform edge enhancement, either as a preprocessing step, or incorporated in the dithering algorithm prior to quantization. Although this alleviates the problem, it is an ad-hoc solution and the results are far from being optimal.

Much better results can be obtained by a careful application of dithering where it is needed. In image areas where the intensity changes slowly there is only shading information. Conversely, in image areas with abrupt changes of intensity there is also shape information that is often manifested in the form of edges. So, when it is applied to image areas of low contrast the dither generates patterns of dots conveying the impression of gray tones with no loss of information. But, when it is applied to image areas of high contrast the dither eliminates edges destroying spatial information.

In order to preserve spatial detail it is necessary to constrain the contours created by transitions between black and white areas to align as much as possible with the real edges of the original image. This must be done without changing the overall image contrast. In the case of dispersed-dot dither these goals can be achieved by various methods related to constrained optimization. An implementation us-

ing neural networks has been tried with limited success by (Geist and Reynolds, 1990). In the case of clustered-dot dither details can only be faithfully reproduced if adaptive clustering is used. Otherwise, with a fixed cluster size it will not be possible to capture features smaller than the size of the halftone screen dots. The best strategy is to make the size of clusters vary according to rate of change in intensity over regions of the image. This is the key idea of the method introduced in this paper.

3 Space Filling Curve Dither

In this section we briefly review the dithering with space filling curves introduced in (Velho & Gomes, 1990). The method takes advantage of the characteristics of space filling curves to perform neighborhood operations essential to the spatial dithering process. The trace of a space filling curve approximation is used to scan the image generating a parametrization of the image elements with many desirable properties. This approach effectively reduces a 2-dimensional problem to a 1-dimensional problem.

The method consists of the following steps:

- Subdivision of the image into regions;
- Computation of the average intensities of each region;
- Generation of corresponding dot patterns for each region;

The subdivision of the image is performed by following the path of the space filling curve until the number of elements visited is equal to the cluster size. This is illustrated in Figure 1-a.

For each region, the computation of the accumulated intensity is performed as each one of its elements is visited. Then, the corresponding dot pattern is generated by selecting a group of contiguous elements proportional in number to the total intensity. In this way, the region is subdivided into two subregions of respectively black and white pixels such that its average intensity is the same as the original image. Figure 1-b shows the configuration of dots corresponding to intensity levels 15/16 to 0 for a cluster of 4×4 pixels.

Because of the discrete nature of the representation there is a quantization error associated with each region. It is the difference between the average intensity of the continuous and quantized values. Note that this residue decreases as the size of the region increases. The quantization error in a region is compensated in the algorithm by propagating it to neighbor regions along the path of the space filling curve.

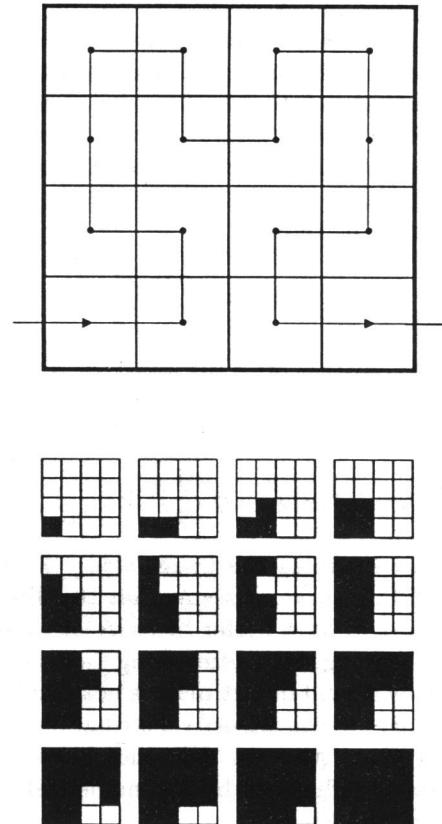


Figure 1: Dithering with Hilbert space filling curve – cluster of 4×4 pixels – (a) path of the curve approximation (b) configuration of dots corresponding to intensity levels 15/16 to 0

4 SFC Dither with Adaptive Clustering

The space filling curve algorithm subdivides the image in cluster blocks, and at each block it approximates the image function $f(x, y)$ by some bi-level image function $\bar{f}(x, y)$. The approximation criteria is a perceptual one, based on the pixels intensities. The adaptive clustering dithering consists of changing the size of each cluster, based on some adaptive criteria, in order to get a better approximation \bar{f} of the image function f .

The space filling curve dithering algorithm lends itself very nicely to adaptive clustering. In fact, it is the only dithering algorithm which allows fine control of the halftone screen resolution.

The adaptive criteria to compute the cluster size depends on the desired effect to be obtained by the halftone screen. In our case, the goal is to achieve the

best rendition of image detail without compromising tonal reproduction. According to our previous discussions in Section 2 we should use an adaptive criteria that varies the cluster size according to the rate of change in image intensity. In order to accomplish for this, we need to measure the variation of image intensities as we scan the image along the path of the space filling curve.

The gradient vector

$$\text{grad } f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right),$$

of the image function gives the direction of maximum rate of change of the image intensity. The rate of change can be measured by the squared norm

$$\|\text{grad } f(x, y)\|^2 = \left(\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right).$$

Since we are scanning the image along the path of the space filling curve, it seems that the directional derivative would be a better measure of the rate of change of the image intensities along the scanning direction. This should in fact make a noticeable difference for images whose intensity distribution has some directional bias. Also the difference between these two criteria is influenced by the numerical method used to compute them.

In our implementation we use the gradient adaptation criteria, and we approximate the gradient by a first order finite difference

$$\begin{aligned} \text{grad } f(x, y) &= (f(x+1, y) - f(x, y)) \\ &\quad , (f(x, y+1) - f(x, y)). \end{aligned}$$

After deciding that the gradient will take care of the adaptiveness criteria, it remains to obtain the correct relationship between the cluster size and the gradient intensity. As the magnitude of the gradient gets bigger, image intensities change faster and, therefore, the cluster size should get smaller. We need to find the correct relationship between the cluster size and the gradient norm. For this, we first observe that the intensities distribution in a dithered image must follow a perceptual criteria; also, the eye response to intensity changes is a logarithmic one (see (?)). Based on the two previous remarks, we conclude that we should vary the cluster size exponentially with the gradient magnitude. This rule maintains a linear relationship between the perceptual intensity inside each cluster and the slope of the image intensity.

5 Implementation

This section describes the implementation of the space filling curve dither algorithm with adaptive

clustering. The algorithm was programmed using the C language in the Unix operating system.

In our implementation we extended the space filling curve dither to work with adaptive clustering in the following way:

- Handle clusters of variable size within the same image;
- Change the cluster size according to a predefined function.

The separation of these two mechanisms makes the algorithm more flexible and allows for experimentation with different sets of criteria in the specification of the cluster size. The latter procedure gives input to the former establishing a clean interface between them.

Although in this work we have experimented mainly with an adaptive criterium based on the variation of the gradient norm of the image intensity function, as explained in the previous section; there are other types of criteria that we believe are worth exploring. One example is the cluster adaptation based on the physical reproduction function of the imaging system in order to compensate its deficiencies. Other example is an adaptation criterium based on a function of the image plane in order to create graphic effects. This opens up many possibilities of use for the method.

Note that if the cluster size is kept constant throughout the image, the algorithm reduces to the original space filling curve dithering.

The cluster size control is done in the first pass through the cluster region. As mentioned before, at each point of the image the maximum allowable cluster size is determined by the adaptation criteria. While the image elements in the region are processed (to accumulate intensity), the current cluster size is compared with the maximum allowable cluster size at that point. If this maximum allowable size is smaller than the current size, it becomes the current cluster size. The algorithm terminates the first pass when the number of elements in the region exceeds the cluster size. Then, it moves on to the second pass as in the original space filling curve dithering (to generate the dot pattern configuration).

In the current implementation the function that defines the cluster size is obtained from the norm of the gradient of the image intensity function at each point of the image.

To make the algorithm more versatile, the adaptation criteria can be passed as a parameter to the program. The default criterium is the rule described above using an exponential of base 2. The other

options provided include: linear variation with the norm of the gradient intensity function, and using a table supplied by the user.

The main structure of the algorithm is given below in pseudo-code:

```

Read input image;
Compute the gradient magnitude;
Create the cluster size table;

Set image_pointer to first region;
while (more pixels) {
    do {
        Accumulate intensity;
        Get max_size at point;
        if (max_size < region_size)
            region_size = max_size;
    } until (pixels_visited >= region_size)

    foreach (pixel in region) {
        if (accumulator >= 255) {
            accumulator -= 255;
            output_pixel = 1;
        } else {
            output_pixel = 0;
        }
    }
    Advance image_pointer to next region;
}

Write output image;

```

6 Results

In this section we show the results of applying the algorithm to various types of images. These test images reflect the main characteristics of images encountered in digital printing situations. All input images are gray-scale with 8 bits of intensity resolution and width of 256 pixels. The output images are printed using a 300 dpi laser printer.

The figures bellow compare the output of the space filling curve dithering algorithm with and without adaptive clustering. The images are scaled by a factor of 4 with pixel replication in order to show more clearly the halftone screen dots. The maximum cluster size is approximately 30 pixels.

Figure 2 is a test pattern of intensity gradations using linear ramps. The gradient increases in steps such that the rate of change in intensity is almost constant for each rectangle. As the slope of the ramp doubles, the cluster size also decreases by a factor of two. Observe that the space filling curve dither without adaptive clustering cannot resolve the last

two columns and rows, because the level of detail there is too high for the dot size.

Figure 3 is a test pattern using a radial cosine function. The cluster size changes continuously according to the magnitude of the gradient. Note that as the frequency increases the cluster size decreases proportionally.

Figure 4 is a synthetic image of a 3D scene generated using ray tracing. It exhibits low detail areas with small intensity variations as well as highly detailed areas with large changes in intensity. This can be seen respectively in the soft shadows and in the reflections of the environment. It can be seen that the adaptive algorithm captures perfectly both the shape and shade information. This is particularly evident in the smooth transitions of the shadow boundaries and in the sharp transitions of the silhouettes of the spheres.

Figure 5 is a cartoon image consisting of line drawings and areas of constant gray level. The shadow on the wall in the background is rendered as regular pattern of dots simulating a standard halftone screen. The effect of the adaptive method is striking. The improvement obtained is mainly due to the high contrast of the image. It needs dithering only in the areas of intermediate intensity. In this case, the algorithm was capable of matching exactly the edges of the drawings and at the same time reproducing with uniform dot patterns the different gray shades. We remark that essentially only two cluster sizes were used by the adaptive method: a cluster size of 1 pixel for the edges and a cluster size of 27 pixels for the areas of constant shade.

Figure 6 is a scanned in photography. This portrait of a woman exhibits a wide range of features: from the smoothness of the skin to the details of the eye and hair. It is interesting to note that the dithering with fixed cluster size did not perform very badly. Nonetheless, the adaptive clustering method shows far superior results, clearly close to an optimal solution. It is worth noting the rendering of the ear rings.

Figure 7 is an enlargement of the left eye area in Figure 6. It is interesting because it reveals in a small area the full range of dot sizes used for the previous image. It also shows the configuration of the dot patterns.

7 Conclusions

In conclusion, an adaptive digital halftoning method with variable size clustering was presented. It extends the space filling curve dithering algorithm to change the size of pixel clusters according to local characteristics of the image. This makes possible to

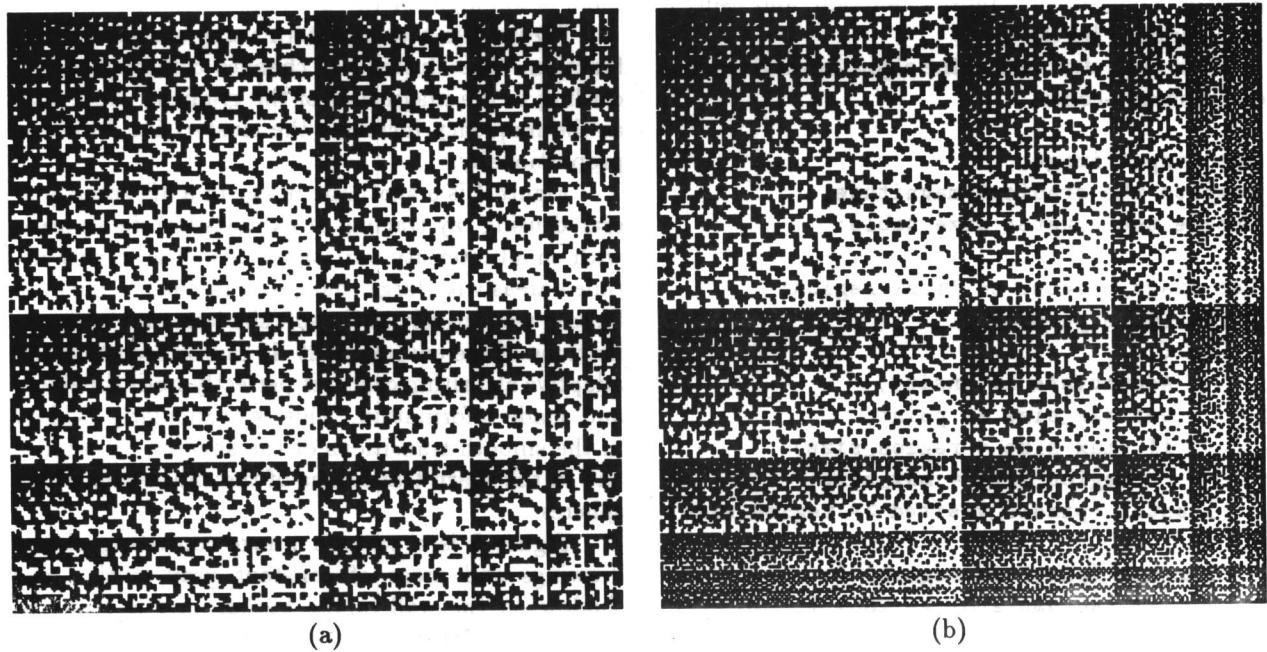


Figure 2: Linear intensity gradations (a) constant cluster size of 31 pixels (b) variable cluster size

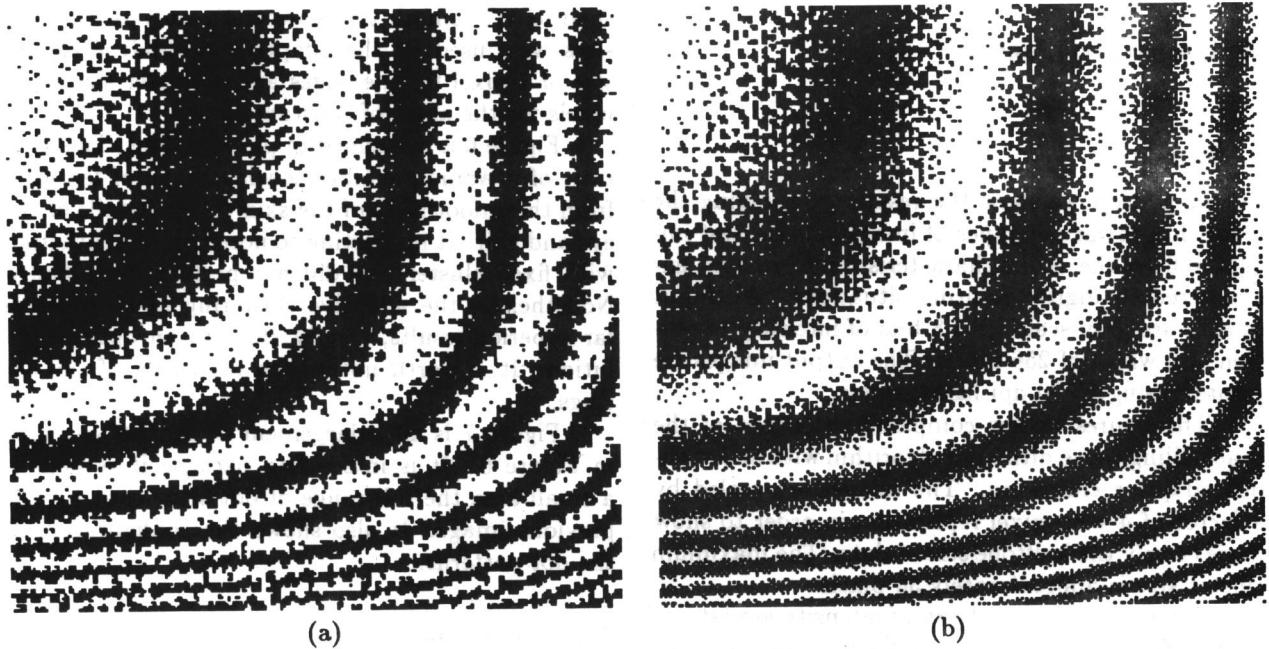
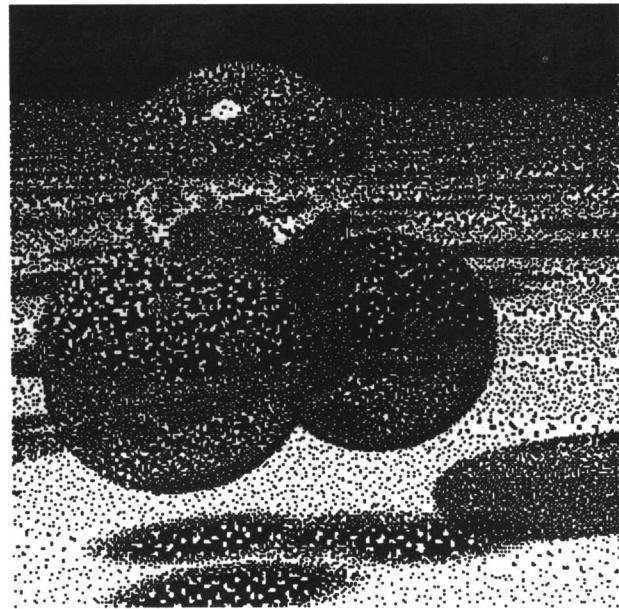


Figure 3: Radial cosine function (a) constant cluster size of 31 pixels (b) variable cluster size

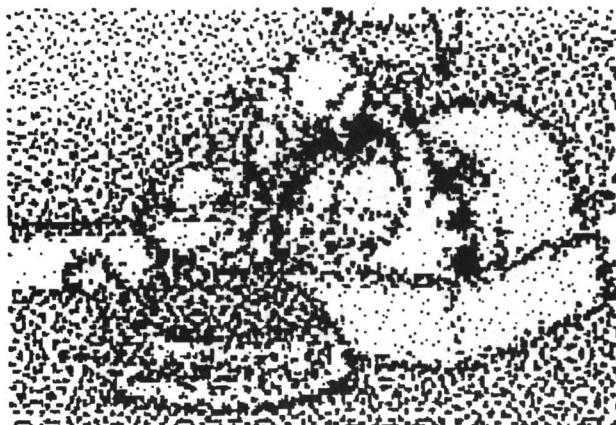


(a)



(b)

Figure 4: Computer generated image (a) constant cluster size of 27 pixels (b) variable cluster size



(a)



(b)

Figure 5: Line drawing cartoon (a) constant cluster size of 27 pixels (b) variable cluster size

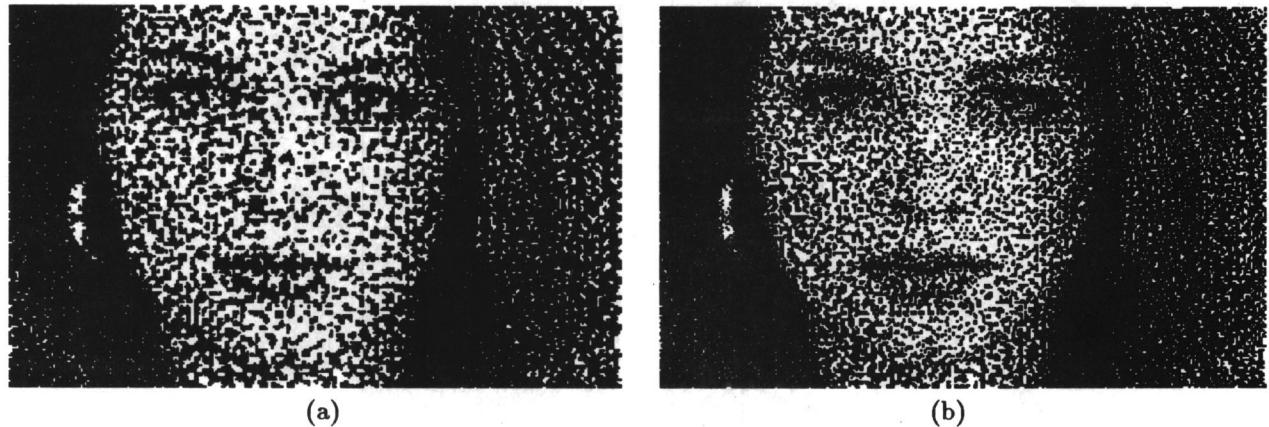


Figure 6: Scanned photograph (a) constant cluster size of 29 pixels (b) variable cluster size



Figure 7: Enlargement of the eye in Figure 6 (a) detail of Figure 6-a (b) detail of Figure 6-b

achieve an optimal rendition of gray shades while preserving image detail.

The dot size variation in the halftone screen can also be used for other purposes besides the improvement of quality in an imaging system. This method could be used to create textures in a paint system or to generate special effects in digital printing and video. Experimentation in this direction is being done. This is facilitated by the structure of the algorithm that separates the cluster size control from the specification of cluster sizes.

8 Acknowledgements

This research was done under the VisGraf project in the computer graphics laboratory at IMPA. The project is partially supported by CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico, FAPERJ - Fundação de Amparo à Pesquisa do Estado Rio de Janeiro and IBM Brasil.

The authors would like to thank André Silva for kindly providing some of the images used in the examples.

9 References

- Adobe Systems (1985). *Postscript Language Reference Manual*. Addison-Wesley, Reading Massachusetts.
- Floyd, R. and Steinberg, L. (1975). An adaptive algorithm for spatial gray scale. *SID Symposium*, pages 36-37.
- Geist, R. and Reynolds, R. (1990). Colored noise inversion in digital halftoning. *Proceedings of Graphics Interface*, pages 31-38.
- Jarvis, J., Judice, C., and Ninke, W. (1976). A survey of techniques for the display of continuous tone pictures on bilevel displays. *Computer Graphics and Image Processing*, (5):13-40.
- Knuth, D. (1987). Digital halftones by dot diffusion. *ACM Transactions on Graphics*, 6(4):245-273.
- Limb, J. (1969). Design of dither waveforms for quantized visual signals. *Bell Systems Technical Journal*, 48(7):2555-2582.
- Ulichney, R. (1987). *Digital Halftoning*. MIT Press, Cambridge, Ma.
- Velho, L. and Gomes, J. (1991). Digital halftoning with space filling curves. *Computer Graphics*, 25(4):81-90.