EXTRACTING HALFTONES FROM PRINTED DOCUMENTS USING TEXTURE ANALYSIS

Dennis F. Dunn

Dept. of Computer Science and Eng.
The Pennsylvania State University
University Park, PA 16802
email: dunn@cse.psu.edu

Thomas P. Weldon

Dept. of Electrical Engineering
Univ. of North Carolina at Charlotte
Charlotte, NC 28223
email: tpweldon@uncc.edu

William E. Higgins

Dept. of Electrical Engineering
The Pennsylvania State University
University Park, PA 16802
email: weh@ruth.ece.psu.edu

ABSTRACT

Separating halftones from text is an important step in document analysis. We present an algorithm that accurately extracts halftones from other information in printed documents. We treat halftone extraction as a texture-segmentation problem. We show that commonly used halftones, consisting of a pattern of dots, can be viewed as a texture. This texture exhibits a distinct spectral component that can be detected using a properly-tuned Gabor filter. The Gabor filter essentially transforms halftones into high-contrast regions that can be isolated by thresholding. We propose a filter-design procedure and provide experimental results.

1. INTRODUCTION

To reproduce photographs and other grayscale images, a pattern of dots called a halftone is used to simulate shades of gray [1]. When documents are processed electronically, it is important to be able to separate halftones from text, since they typically have different processing requirements. *In this paper, we describe a method for extracting halftones from documents.* We show that halftones, because of their distinctive dot pattern, can be viewed as textures. It has been shown that textures can be segmented based on spectral differences by using bandpass filters (e.g., Gabor) [2, 3, 4, 5, 6, 7]. This suggests that bandpass filters can be used to isolate regions containing halftones. Other researchers have employed spectral techniques in document analysis. Jain and Bhattacharjee [8] used a bank of 28 Gabor filters to segment text from illustrations by treating lines of text as a texture. Randen and Husøy [9] adopted a similar, but more efficient, approach using perfect reconstruction filters. In these previous works, the authors focused on the textural properties of the text. In contrast, we concentrate on the textural characteristics of halftones. Although textures produced by the halftone process are not visible to the naked eye, these textures produce spectral characteristics that differ significantly from other information in the document. More importantly, halftones derived from different images (but produced using the same halftone screen) are effectively samples of the *same texture*, even though visually they appear quite different. Consequently, *all halftones* in any given document have similar spectral characteristics, and therefore a *single Gabor filter* can be used to segment halftones from printed documents. Our approach is simpler and produces more accurate segmentations (based on a visual comparison of results) than the Texture Co-occurrence Spectrum approach of Payne *et al.* [10] (although their approach addresses the more general problem of region classification).

2. SPECTRAL PROPERTIES OF HALFTONES

The subsequent analysis describes the spectral composition of an image when represented as a halftone. We show that the halftone dot pattern can introduce a high-frequency carrier that is essentially modulated by the baseband of the input image. The baseband signal corresponds to the grayscale variations in the original photograph. This effectively causes low-frequency phenomena in the original photograph to appear at much higher frequencies in the halftone. In particular, the spectral characteristics near DC appear to be replicated at multiples of the halftone dot frequency. For simplicity, we restrict our analysis to 1D cross-sections (*signals*) of 2D images, which allows us to perform the analysis in 1D. Section 4 provides experimental evidence suggesting that these principles also apply to 2D images.

We begin with an analytical expression for the halftone process derived by Kermish and Roetling [11]:

$$h(x) = s(x) + \frac{2}{\pi}\sum_{n=1}^{\infty} n^{-1}\sin(\pi n s(x))\cos(n\omega_0 x) \quad (1)$$

where $h(x)$ is the halftone signal, $s(x)$ is the baseband

signal, and ω_0 is the halftone dot frequency (the carrier). When contrast is limited, the spectrum of the halftone signal can be approximated by the spectrum of the original signal replicated at integer multiples of the halftone frequency. We expand the sine function in (1) as a Taylor series.

$$h(x) = s(x) + \frac{2}{\pi}\sum_{n=1}^{\infty} n^{-1} z(x;n)\cos(n\omega_0 x) \qquad (2)$$

where

$$z(x;n) = n\pi s(x) - \frac{(n\pi s(x))^3}{3!} + \frac{(n\pi s(x))^5}{5!} - \cdots \qquad (3)$$

Note that since $-1 \leq z(x;n) \leq 1, \forall n$, the quantity $z(x;n)/n$ in (2) decays as $1/n$. We approximate (2) by considering only the first two terms in the summation in (2), and assume that the contrast is sufficiently low such that $|2\pi s(x)| \ll 1$. This gives

$$h(x) \approx s(x) + 2s(x)\cos(\omega_0 x) + 2s(x)\cos(2\omega_0 x) \qquad (4)$$

From modulation theory, the spectrum of (4) is just the spectrum of $s(x)$ centered at DC plus the same spectrum at $\pm\omega_0$ and at $\pm 2\omega_0$ (appropriately scaled). Although contrast is seldom small enough to justify the assumption above, the result supports the notion of spectral replicas mentioned previously. Recall that our main concern is to isolate the halftone from the text. To do this, all we need to know are the basic characteristics of the halftone spectrum. Thus, an accurate approximation to $h(x)$ is not required.

The previous analysis suggests that spectral replicas can also occur in halftones of 2D images; i.e., approximate copies of the image spectra can occur at multiples of the halftone dot frequency. Note that for a page containing both text and images represented as halftones, the spectrum near DC will be a composite of text and image information. At multiples of the halftone dot frequency, though, the spectrum will consist primarily of image data associated with the halftones. Thus, a bandpass filter, tuned to the halftone dot frequency, can be used to extract the halftones. This is the foundation of our approach.

3. AN ALGORITHM FOR EXTRACTING HALFTONES

Although our approach to extracting halftones was presented using spectral arguments, halftone-extraction can also be viewed as a texture-segmentation problem. Since halftones consist of dots on a fixed lattice, halftones can be considered a strongly-ordered texture, based on Rao's taxonomy [12]. Strongly ordered textures are characterized by a regular pattern of geometric primitives called texels. In halftones, the texels correspond to the halftone dots that vary in diameter. Although the dot pattern is not visible to the naked eye, halftones exhibit all the properties characteristic of textures, just at a smaller scale.

Dunn and Higgins [6, 7] analyzed strongly-ordered textures in detail and showed how Gabor filters can be used for texture segmentation. By viewing halftones as textures, we can apply these same principles to isolate halftones. In particular, since the halftone dot spacing is typically the same for all halftones in a given document, *the halftones are effectively instances of the same texture.* Thus, *a single Gabor filter is sufficient to isolate them.* Below, we develop an algorithm for selecting an appropriate filter and demonstrate how the filter output can be used to segment halftones from other information in printed documents. We begin with a brief discussion of Gabor filters [6, 7], followed by a summary of the algorithm.

We define a Gabor filter O_h as:

$$m(x,y) = O_h(i(x,y)) \triangleq |i(x,y) * h(x,y)| \qquad (5)$$

where i is an image, m is the output, and $h(x,y)$ is a Gabor elementary function (GEF) [13]. $H(u,v)$, the frequency-domain representation of a GEF, is given by

$$H(u,v) = \exp\left\{-\frac{1}{2}\sigma^2((u-U)^2 + (v-V)^2)\right\} \qquad (6)$$

where σ characterizes the bandwidth of the filter, (u,v) denote frequency-domain rectilinear coordinates, and (U,V) represents a particular 2D frequency [3, 6, 14]. The 2D frequency can also be expressed in polar coordinates as a radial frequency $F = \sqrt{U^2+V^2}$, and angle $\phi \triangleq \tan^{-1}(V/U)$. The GEF's frequency response has the shape of a Gaussian, centered about the frequency (U,V). Thus, the GEF acts as a bandpass filter. Convolving an image with a GEF results in an output that is complex. The magnitude operation in (5) converts the complex output to a real-valued output and has other useful properties [6].

Our algorithm for extracting halftones begins with the initialization of a few constants. These constants are fixed for any given document or series of documents and can be determined off-line. First, we select one sample halftone and determine the halftone dot frequency (simply by counting the dots per inch under magnification). Only one example is needed since documents typically use the same screen for all halftones. The dot frequency determines the radial frequency F of the filter (taking into account any frequency folding due to low digitizer sampling rates). Next, we establish both the spatial extent and bandwidth of the filter by specifying the single parameter σ. Finally, we specify a segmentation threshold T as a percentage of the filter-output range. As we will see, performance is relatively insensitive to σ and T. Then, for each page in the document, the following steps are performed:

1. Digitize the page and compute its discrete Fourier transform.

2. Find the point of maximum energy in an annular region with radius equal to the Gabor-filter radial frequency F. This determines the angle ϕ, and completes the filter design.

3. Apply the filter to the page and threshold the filter output.

4. RESULTS

Fig. 1a is a clip from a popular journal [6] containing both text and a halftone illustration. Fig. 1b shows its Fourier spectrum. Note the four bright spots (circled for clarity) representing the halftone dot frequency. Applying a Gabor filter tuned to the halftone dot frequency produces the filter output shown in Fig. 1c. Note how much of the baseband signal is recovered by the single filter. This supports the notion of spectral replicas suggested in the 1D analysis. To segment the image, we simply threshold the filter output. The result after thresholding is shown in Fig. 1d. Note that thresholding produces a very good segmentation except where the original image is very dark. This can be expected, since as the grayscale approaches black, the dot size increases to the point where the individual dots become indistinguishable. In the limit, the local spectrum approaches DC. To illustrate the accuracy of the segmentation, we computed the minimum-area rectangular bounding box that includes all points above the threshold. The minimum is computed over all orientations between -45 to +45 degrees to account for possible image rotations. The result is shown in Fig. 1a, where the bounding box (shown in black) is superimposed over the original input image. Fig. 2a shows an example of a rotated image, with the bounding box superimposed.

Implementing our method is straightforward, requiring only four easily determined parameters: the filter center frequency, defined by F and ϕ, the filter extent/bandwidth σ, and the threshold T. The choice of filter center frequency is critical to performance, but since the halftone dot frequency is always significantly above the baseband (otherwise it would be visible to the eye), it is conspicuous. In contrast, our algorithm is insensitive to bandwidth and threshold. A large bandwidth (small σ) is desirable because the corresponding filter has a small spatial extent (more spatial resolution). If σ is too small, though, sampling the filter itself can cause aliasing. We used $\sigma = 4$ pixels for all examples. Segmentation results, however, were largely unaffected by differences in bandwidth. Regarding threshold sensitivity, we achieved similar segmentations for thresholds spanning 25% of the output range. For example, the segmentation in Fig. 1a was produced using a threshold $T = 64$ out of 255. Using a threshold of 128 produced a bounding box approximately 3% smaller.

Another example is shown in Fig. 2b, with the bounding box shown in white. This image was also used by [8] and [9] and is provided for comparison.

5. REFERENCES

[1] R. Ulichney, *Digital Halftoning*. Cambridge, MA: The MIT Press, 1987.

[2] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 23, no. 12, pp. 1167–1186, 1991.

[3] A. Bovik, M. Clark, and W. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Machine Intell*, vol. 12, pp. 55–73, January 1990.

[4] F. Farrokhnia and A. K. Jain, "A multi-channel filtering approach to texture segmentation," in *IEEE Conf. Comp. Vision Pattern Rec.*, pp. 364–370, 1991.

[5] J. Bigun and J. du Buf, "N-folded symmetries by complex moments in Gabor space and their application to unsupervised texture segmentation," *IEEE Trans. Pattern Anal. Machine Intell*, vol. 16, pp. 80–87, January 1994.

[6] D. Dunn, W. Higgins, and J. Wakeley, "Texture segmentation using 2-D Gabor elementary functions," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, pp. 130–149, February 1994.

[7] D. Dunn and W. Higgins, "Optimal Gabor filters for texture segmentation," *IEEE Trans. Image Proc.*, vol. 4, pp. 947–964, July 1995.

[8] A. K. Jain and S. Bhattacharjee, "Text segmentation using Gabor filters for automatic document processing," *Machine Vision and Applications*, vol. 5, pp. 169–184, 1992.

[9] T. Randen and J. H. Husøy, "Segmentation of text/image documents using texture approaches," in *Proc. NOBIM-konferansen-94*, (Asker, Norway), pp. 60–67, June 1994.

[10] J. S. Payne, T. J. Stonham, and D. Patel, "Document segmentation using texture analysis," in *Proc.* 12^{th} *IAPR, Jerusalem, Israel*, vol. II, pp. 380–382, October 1994.

[11] D. Kermish and P. G. Roetling, "Fourier spectrum of halftone images," *J. Opt. Soc. Am.*, vol. 65, pp. 716–723, June 1975.

[12] A. R. Rao, *A Taxonomy for Texture Description and Identification*. New York, NY: Springer-Verlag, 1990.

[13] D. Gabor, "Theory of communication," *J. IEE (London)*, vol. 93, pp. 429–457, 1946.

[14] J. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Am. A*, vol. 2, pp. 1160–1169, July 1985.

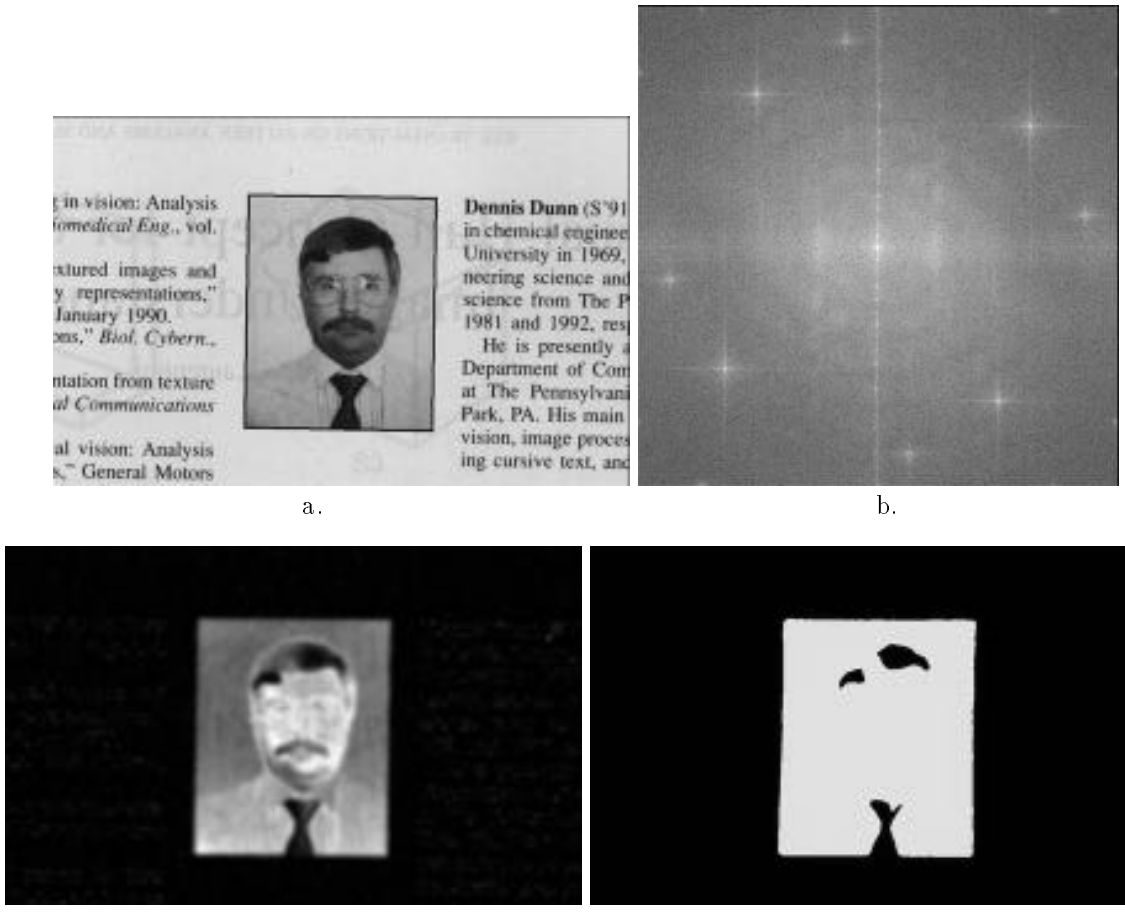

Figure 1: Halftone from [6; pp. 149]: (a) Input image with minimum-area bounding box superimposed in black; (b) Fourier transform magnitude plot (FFT log-magnitude) of the original image, digitized at 300 dpi. Spectral components of the halftone carrier are circled; (c) Output from a properly tuned Gabor filter. $\sigma = 4$ pixels, $F = 0.4078$ cycles/pixel, $\phi = 51.62°$; (d) Result from thresholding the filter output at 25% (graylevel 64 out of 255).

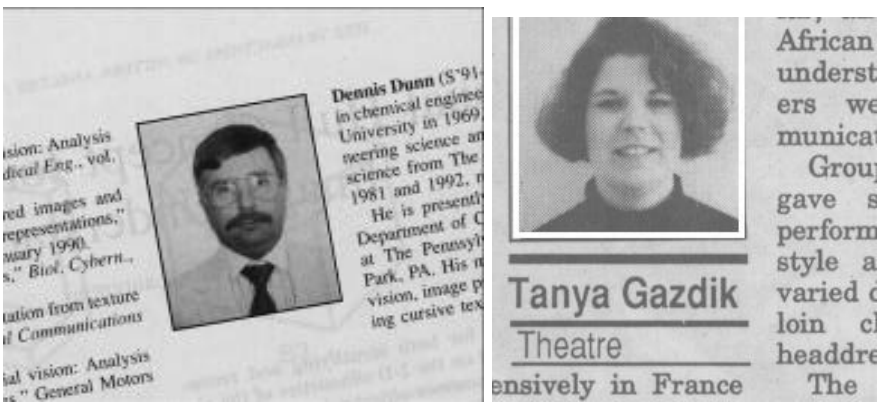

Figure 2: (a) Rotated version of the halftone in Fig. 1 with bounding box superimposed in black. Threshold = 64. Parameters the same as in Fig. 1. (b) "Gazdik" halftone from Jain and Bhattacharjee [8] with permission. Computed bounding box in white. Threshold = 64. Gabor filter parameters: $\sigma = 4$ pixels, $F = 0.561$ cycles/pixel, $\phi = 45°$.