The Frequency-Domain Effects of Stochastic Image Foveation in Superpixelating Cameras

EE381K-14 Term Literature Survey

Thayne Coffman
coffman@ece.utexas.edu

Abstract

We are interested in the relative performance of various halftoning methods when they are used to translate real-valued desired resolution signals into binary control signals for variable acuity superpixel imager (VASI) cameras. This paper describes and compares three of the most popular halftoning techniques (classical screening, error diffusion, and dithering with blue noise), and we find that when they are considered in that order, the techniques form a spectrum moving from low complexity and low results quality to medium complexity and high results quality. We also briefly present the set of metrics we will use to measure performance and our plans for future work.

1. Introduction

The variable acuity superpixel imager (VASI) is a recent development in camera technology [1]. VASI cameras generate foveated imagery by sharing charges between pixels directly on the focal plane array (FPA). The sharing (or non-sharing) behavior of each pixel is specified at frame rate in the form of a binary vector. With this control signal, multiple foveae can be maintained, and they can be created, repositioned, or removed in each frame. Performing foveation on the focal plane also drastically reduces the bandwidth required to transfer images off the FPA, allowing effective frame rates up to and above 1000 Hz. This combination of a high field of view, high resolution on regions of interest, low bandwidth, and very high frame rates make the VASI camera an attractive sensor for automatic target recognition (ATR) applications [2].

To use these cameras, however, the user must specify the control signal that defines which pixels share charges. The varying resolution of foveated imagery is typically specified as a real-valued desired resolution image. We believe that the translation between the desired resolution and a binary control signal can be based on digital halftoning techniques with little additional implementation effort. This translation must be very efficient to avoid lowering the camera's effective frame rate. We also believe that stochastic approaches, where the translation is probabilistic, are particularly promising. The objective of our project is to compare how different halftoning approaches perform at translating desired resolutions to VASI control signals in the context of an ATR application. The objectives of this survey are to describe the halftoning methods we will use and to introduce the quality metrics we will use to measure performance.

2. Related Work

Digital halftoning methods (Section 2.1) and quality metrics (Section 2.2) are both well represented in the existing literature. The halftoning problem is analogous to our problem of converting the real-valued desired resolution function to a binary control signal for each VASI camera pixel, and quality metrics are required to measure the halftoning methods' performance.

2.1 Halftoning

Digital halftoning, also called spatial dithering, is the process of converting continuous-intensity images to binary images for display on media that can only produce two intensity levels [3]. For our discussion, the original images are assumed to be intensity images, normalized such that each pixel's value falls in the range [0, 1]. The values at each pixel in the resulting halftoned image are constrained to be exactly zero (black) or exactly one (white). There are a number of general approaches to halftoning, and many variations on each approach. Some approaches allow more than two intensity levels in the output image. Others operate on and produce color input images. Neither of those categories will be covered in this survey. Instead, we will focus our attention on three of the most common methods as applied to intensity images [4] – classical screening, error diffusion, and dithering with blue noise. There are many other approaches that are not covered in detail here. Those include but are not limited to:

- Dot diffusion [5], a hybrid between classical screening and error diffusion
- Direct binary search [6], a time-consuming iterative approach
- Look up table methods, which modify other approaches to improve efficiency [7].

2.1.1 Classical Screening

Classical screening, also called ordered dithering, is one of the oldest halftoning methods still in common use. Classical screening is based around the use of a *dithering matrix*, which has values in the range [0, 1]. At each pixel in the image, the pixel's gray level value is thresholded against a particular entry in the dithering matrix. The pixel is colored black if it falls below the threshold and white if it is above. The threshold is selected by periodically replicating the dithering matrix to cover the entire real-valued image. Most of the research in classical screening revolves around the engineering of the dithering matrix, and some of the most popular matrices in current use were introduced by Bayer in [8].

Dithering matrices can, at a high level, be classified into two categories – *clustered dot* matrices and *dispersed dot* matrices. Clustered dot matrices tend to put black pixels (or ink dots) near other black pixels, in a "cluster." Clustered dot matrices are commonly used on printing devices, where the tendency of ink dots to diffuse through the paper (or other medium) is mitigated by the clustering of the dots. In contrast, dispersed dot matrices tend to spread black pixels out more uniformly, and are commonly used when processing images for human consumption on displays. The dispersion of the dots reduces the perception of unpleasant visual effects by the human visual system.

Classical screening is very simple and efficient, but generally produces images of lower quality than other methods. It is a *point operation*, meaning that the value at a given pixel in the halftoned image can be computed using only the value (and location) of that same pixel in the original image – without requiring information about the value of neighboring pixels in the original or halftoned images. In general, point operations tend to be more efficient than *neighborhood operations*, which require information from neighboring pixels in the original or halftoned image to compute the value at a given pixel. Both clustered dot dithering and dispersed dot dithering suffer, however, from unpleasant visual artifacts introduced by the use of a periodic dithering matrix. As stated above, dispersed dot dithering suffers to a lesser extent than clustered dot dithering.

2.1.2 Error Diffusion

Error diffusion was introduced by Floyd and Steinberg in [9]. It attempts to spread the quantization error introduced at each pixel among neighboring pixels. Rather than varying the threshold as in classical screening, error diffusion quantizes every real-valued pixel against a constant threshold of 0.5. However, the quantization error (the difference between the quantized value and the original value) is then used to modify the pre-quantization values at neighboring pixels. The error is diffused among multiple neighboring pixels according to a weighting scheme. This feedback loop has a kind of anti-hysteretic effect, where a pixel that is quantized to 1.0 will increase the likelihood of neighboring pixels being quantized to zero.

Two aspects of the approach are commonly varied – the error diffusion weighting scheme and the scan order of the processing. A number of alternative weighting schemes have been proposed with varying justifications, but the original scheme proposed by Floyd and Steinberg gives performance that is comparable or better than others. The scan order of processing is often changed from raster scan to serpentine scan because error diffusion can produce anisotropic (directionally-biased) or periodic visual artifacts on areas of constant or near-constant intensity, and serpentine scans help to mitigate those visual effects.

Error diffusion produces visual results that are superior to classical screening, but at increased computational cost. Because the quantization of a given pixel is affected by neighboring pixels (it can potentially be affected by all previously-quantized pixels), error diffusion is categorized as a neighborhood operation (and bears the associated increased computational complexity). The process of diffusing the error across neighboring pixels does succeed in generally reducing the occurrence of unpleasant visual artifacts, although some anisotropy remains. The visual quality of error diffusion is generally superior to classical screening, but typically not as high as dithering with blue noise or direct binary search.

2.1.3 Dithering With Blue Noise

Dithering with blue noise (or just “blue noise”) is a method for extending error diffusion, and it was introduced by Ulichney in [10]. Blue noise introduces random perturbations into the diffusion weights and/or the relative locations to which the error is diffused. Different blue noise techniques vary in the set of perturbations they apply, but are conceptually very similar. While the application of the technique is relatively simple, the analysis supporting its use is somewhat involved.

Ulichney argues the superiority of the blue noise approach based on two metrics – anisotropy and radial frequency content. He argues that because of the characteristics of the human visual system, a dithering process will maximize perceived image quality when it introduces only *blue noise* artifacts, defined as isotropic noise that is constrained to high radial frequencies. He derives a metric for anisotropy, or the amount of variation in the frequency content at constant radial frequency but varying directions. Ulichney then compares halftoning approaches by comparing the anisotropy and radial frequency content of the artifacts introduced when the approaches are applied to constant-intensity images.

Using his metrics, Ulichney shows that blue noise halftoning yields characteristics closer to his defined ideal than either classic screening or error diffusion. His analysis is generally accepted and it accurately mirrors subjective evaluations – in terms of perceived quality, blue noise outperforms error diffusion, which outperforms classical screening. Ulichney argues that error diffusion performs closer to the ideal in part because the anti-hysteretic effect of the feedback loop shapes the noise towards higher frequencies. Blue noise introduces a relatively minor amount of additional computational cost over error diffusion, but this cost can be mitigated through efficient random number generators and/or precomputing the set of random numbers.

2.1.4 Comparison of Halftoning Approaches

Table 1 summarizes the computational complexity and results quality of the halftoning methods mentioned above, as well as other popular halftoning methods that were not covered in this paper.

Halftoning Approach	Computational Complexity	Results Quality
Clustered dot screening	Very low	Low. Better for printed media than dispersed dot screening.
Dispersed dot screening	Very low	Low. Better for human consumption than clustered dot screening.
Error diffusion	Medium	High
Dithering with blue noise	Medium	High
Dot diffusion	Low	High
Direct binary search	Very high	Very high
Look up tables	Very low	Varies with the type of approach it is based on.

Table 1: Complexity and quality (for human consumption) of various halftoning methods

2.2 Performance Assessment

We will now give a brief description of how we plan to evaluate the performance of the halftoning approaches when used to generate VASI control signals. Because we are targeting ATR applications specifically, the most direct way to measure performance would be to use actual ATR performance under different halftoning approaches. Typical ATR metrics include false positive rate (incorrect detection), false negative (missed detection) rate, classification accuracy, and area under the receiver operator characteristic (ROC) curve [11]. Time limitations prevent us from being able to implement this direct approach to performance measurement because we would also have to implement the ATR application itself.

Instead, we will measure performance through more objective image quality metrics. We will compute these metrics on the various halftoned images and then infer the relative expected ATR performance. We will start with simple metrics (taken in the spatial or frequency domains) like mean squared error (MSE), signal-to-noise ratio (SNR), and peak signal-to-noise ratio (PSNR). These have been shown to be inappropriate for modeling human perception of image quality, but they can be reasonable estimators of vision algorithm performance. Implementations of these metrics are widely available. We also plan to try adapting Ulichney's measures of anisotropy and frequency content to compare the power spectra of the non-foveated and foveated images. Finally, we plan to explore the Universal Quality Index developed by Wang et al [12] and available for download at [13] and [14].

3. Term Project Implementation Plan

As stated above, the goal of our term project is to compare how different halftoning methods perform when translating a desired resolution image into a share/no-share control signal at each pixel. Our performance metrics have been selected to measure the fidelity of the foveated images to the original full-resolution images in the context of an automatic target recognition application. We will select a set of test images, likely to include the popular “Lena” and “Mandrill” images, and generate our desired resolution functions by hand. We will download (from [15]) implementations of the halftoning approaches detailed in Section 2.1. If time permits we will also implement a stochastic approach of our own design. We will simulate in software the charge sharing behavior of a VASI camera that uses these halftones as control signals, and compare the quality of the resulting images with our selected metrics.

4. Conclusion

We have briefly presented our motivation for studying various halftoning approaches, some of the more popular methods, our approach to quantifying their performance in our application, and our plans for future work. We find that classical screening approaches are the least computationally complex but produce the lowest quality results. Error diffusion and dithering with blue noise each offer improvements in quality at the cost of increased complexity. We expect the “best” approach in our application to be a balance between computational complexity and performance against our selected metrics.

5. References

[1] P. McCarley, M. Massie, J.P. Curzan, “Large format variable spatial acuity superpixel imaging: visible and infrared systems applications,” *Proc. SPIE, Infrared Technology and Applications XXX*[sic.], vol. 5406, pp. 361-369, Aug 2002.

[2] D.J. Stack, C. Bandera, C. Wrigley, B. Pain, “Real-time reconfigurable foveal target acquisition and tracking system,” *Proc. SPIE*, vol. 3692, pp. 300-310, July 1999.

[3] M. Mese, P.P. Vaidyanathan, “Recent advances in digital halftoning and inverse halftoning methods,” *IEEE Trans. on Circuits and Systems*, vol. 49, no. 6, pp. 790-805, June 2002.

[4] B. Evans, *Image Halftoning Research*, Online tutorial available at http://www.ece.utexas.edu/~bevans/projects/halftoning/.

[5] D.E. Knuth, “Digital halftones by dot diffusion,” *ACM Trans. Graph.*, vol. 6, no. 10, pp. 245-273, 1987.

[6] M.A. Seldowitz, J.P. Allebach, D.E. Sweeny, “Synthesis of digital holograms by direct binary search,” *Appl. Opt.*, vol. 26, pp. 2788-2798, 1987.

[7] A.N. Netravali, E.G. Bowen, “Display of dithered images,” *Proc. SID’81*, vol. 22, pp. 185-190, 1981.

[8] B.E. Bayer, “An optimum method for two level rendition of continuous-tone pictures,” *Proc. IEEE Int. Conf. on Communications, Conf. Rec.*, pp. (26-11)-(26-15), 1973.

[9] R. Floyd, L. Steinberg, “An adaptive algorithm for spatial grayscale,” *Proc. SID’76*, pp. 75-77, 1976.

[10] R.A. Ulichney, “Dithering with blue noise,” *Proc. IEEE*, vol. 76, pp. 56-79, Jan 1988.

[11] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, Second Edition, John Wiley & Sons, 2001.

[12] Z. Wang, A.C. Bovik, L. Lu, “Wavelet-based foveated image quality measurement for region of interest image coding,” *Proc. IEEE Int. Conf. Image Proc.*, vol. 2, pp. 89-92, Oct 2001.

[13] N. Damera-Venkata, B. Evans, *Image Quality Assessment.* Matlab files. Version 1.0 released April 28, 2001. Available online at http://www.ece.utexas.edu/~bevans/projects/halftoning/.

[14] Z. Wang, *Demo Images and Free Software for a Universal Image Quality Index.* Available at: http://signal.ece.utexas.edu/~zwang/research/quality_index/demo.html.

[15] V. Monga, N. Damera-Venkata, B. Evans, *Halftoning Toolbox for Matlab.* Version 1.1 released November 7, 2002. Available online at http://www.ece.utexas.edu/~bevans/projects/halftoning/.