A Comparison of Halftoning-Inspired Methods for Foveation in Variable-Acuity Cameras

EE381K-14 Term Project Final Report
Thayne Coffman
coffman@ece.utexas.edu

Abstract

Our objective is to compare how different halftoning-inspired approaches perform at generating the binary control signal used by Variable Acuity Superpixel Imager (VASI) foveating cameras, in the context of automatic target acquisition and recognition (ATA/ATR) applications. To compare performance, we foveate a small set of test images with a variety of halftoning-inspired approaches and then measure five objective figures of merit (FOMs). We explore seven standard halftoning variations and two specialized methods of our own design (also presented) named vasiHalftone and vasiHalftone2. Our control signal generation approaches require an inverse function operation to counteract the effects of the pixel geometries produced by each particular halftoning approach, as explained in the paper. We find that the vasiHalftone and vasiHalftone2 approaches provide superior FOM values but are unable to accurately produce a target bandwidth reduction, making them unsuitable for use with VASI cameras. Floyd and Steinberg error diffusion is found to give the best FOM values while allowing precise bandwidth control. Other halftoning approaches are characterized by inferior FOM values, inferior bandwidth control, or both.

1. Introduction

Our objective is to compare how different halftoning-inspired approaches perform at generating the binary control signals used by Variable Acuity Superpixel Imager (VASI) foveating cameras, in the context of an automatic target acquisition and recognition (ATA/ATR) application. To accomplish this, we measure five objective figures of merit (FOMs) after foveating a small set of test images with a variety of halftoning-inspired approaches.

VASI cameras have a number of characteristics that are attractive for the ATR application, but to use them the user must specify a binary control signal that defines the camera's pixel-by-pixel behavior. The translation from a continuous-valued desired resolution signal to a binary VASI control signal can be based on halftoning. It must be efficient to avoid lowering the camera's effective frame rate. It must also accurately achieve the target resolution – or equivalently a target bandwidth reduction, which we express as percentage of original bandwidth **(PBW)** – because this drives the frame rate achieved by the VASI camera.

2. Background

VASI cameras generate foveated imagery by varying the camera's spatial acuity [1]. They do this by sharing charges between pixels directly on the focal plane array (FPA). The behavior at each pixel is specified as a binary vector supplied for each pixel in each frame. Thus multiple foveae can be maintained, and they can be created, repositioned, or removed at frame rate. VASI cameras use two bits per image pixel to control sharing with the pixel above and sharing with the pixel to the left. *In this study we made the significant constraint that the "share up" and "share left" controls are identical for each pixel.* Foveating directly on the FPA reduces the bandwidth needed to transfer pixel data off the FPA, allowing frame rates of over 1000 Hz. The combination of wide fields of view, high resolution on targets, low bandwidth, and high frame rates make VASI cameras attractive sensors for ATR applications [2].

Digital halftoning is the process of approximating continuous-valued images using only two (binary) intensity levels [3]. Halftoning is well represented in the available literature. A comprehensive survey of approaches is beyond the scope of this paper – other surveys already exist, including [3] and [4]. Three of the most common and popular methods are classical screening, error diffusion, and dithering with blue noise [5]. The choice of halftoning approach is often a tradeoff between visual quality and computational complexity. Classical screening is an efficient point operation that thresholds intensity values against a periodic dithering matrix [6]. It can be broadly divided into *clustered dot* and *dispersed dot* screening. It is one of the earliest methods and generally produces results that are inferior to other approaches. Error diffusion is a neighborhood operation that spreads the quantization error introduced at each pixel among neighboring pixels [7]. Serpentine scans are sometimes used in error diffusion to reduce visual artifacts. Error diffusion produces visual results that are superior to classical screening (at an increased computational cost) but not as high as dithering with blue noise. Dithering with blue noise extends error diffusion by randomly perturbing the diffusion weights and/or directions, in order to shape noise into higher frequencies and make it more isotropic [8]. Blue noise introduces minimal computational cost beyond basic error diffusion.

Image quality assessment can be more accurately called “image fidelity assessment” because the goal is to determine how well a distorted image reflects an original. We choose peak signal to noise ratio (PSNR), weighted signal to noise ratio (WSNR) [9], linear distortion measure (LDM) [9], and universal quality index (UQI) [10] to measure the fidelity of our foveated images to the originals. Some FOMs have been shown to be uncorrelated with subjective human perceptions of image quality, but they can accurately predict the performance of computer vision algorithms like ATR. Implementations of these metrics are widely available.

While VASI camera characteristics, halftoning, and image quality assessment are all relevant, relatively little prior work has been done with the specific goal of engineering the VASI control signal. To our knowledge, only the developers of the VASI camera and their associates have addressed this problem at all, and based on [1], [2], and [11], their strategy appears to support only three resolution choices – 1x1 pixels, 2x2 superpixels, and 4x4 superpixels. This would seem to grant only limited control over bandwidth reduction and VASI camera frame rate, and the lowest PBW it could achieve (with no foveae at all) would be 6.25%.

3. Approach

We selected a set of test images and manually defined the desired resolution functions and foveae. We then performed a transformation on the desired resolution signal and used a variety of standard and specialized halftoning approaches to convert the desired resolution to the VASI control signal. The VASI charge-sharing behavior was simulated in software and the fidelity of the resulting images was evaluated based on our FOMs. Implementations for the standard halftoning approaches and all FOMs were taken from [12]. Other implementation was done in Matlab. Our five test images (converted from color to grayscale where necessary) were *board.tif* ("board"), *westconcordorthophoto.png* ("concord"), and *eight.tif* ("eight") from Matlab's Image Processing Toolbox, *lena.tif* ("lena") from [12], and an image of an intersection ("traffic") from [13].

The desired resolution functions were defined to have one or two foveae superimposed over an image-wide minimum resolution. Each fovea is a 2D Gaussian peak normalized to a maximum value of 1.0. The generation parameters are given in Table 1, along with the resulting average desired resolution (which was used as the target PBW in our tests).

The halftoning approaches we used were block error diffusion ("bed"), dithering with blue noise ("bnoise"), Floyd & Steinberg error diffusion with raster ("floyd") and serpentine

("floydSerp") scans, clustered dot ("screen9c") and dispersed dot ("screen9u") classical screening with 9 gray levels, two specialized approaches of our own design ("vasiHalftone" and "vasiHalftone2"), and white noise dithering ("wNoise").

Table 1: Desired Resolution Function Generation Parameters

Image	Min. Res.	Fovea μ	Fovea Σ	Target PBW
board	0.05	[220 180]	[1200 0; 0 1200]	0.0956
		[440 160]	[600 0; 0 600]	
concord	0.05	[140 270]	[900 0; 0 900]	0.1164
		[220 210]	[900 0; 0 900]	
eight	0.05	[145 60]	[1200 0; 0 1200]	0.2086
		[75 250]	[1200 0; 0 1200]	
lena	0.05	[128 128]	[900 0; 0 900]	0.1190
traffic	0.02	[150 200]	[2500 0; 0 3600]	0.0584

Halftoning the desired resolution signals directly yields very poor results. The various standard halftoning algorithms accurately translate a desired resolution of X% into a control signal with X% of the pixels set to 1.0, but the PBW of the foveated image can be significantly affected by the geometry of which specific pixels share charges as determined by the halftoning approach. These effects are illustrated in Figure 1, which shows for each halftoning approach the percentage of charge-sharing pixels and the achieved PBW as functions of desired resolution. In order to counteract the approach-specific effects of geometry, we perform a function inverse via a lookup table (LUT) to recover the desired resolution value that will cause a given halftoning approach to generate a given target PBW. Larger tables improve our ability to hit the target PBW, but they also cause significant increases in the runtime of the translation. Any discontinuous jumps in the relationship between desired resolution and realized PBW limits our ability to achieve the target PBW, as seen in our results.

We also developed two specialized approaches, vasiHalftone and vasiHalftone2. These are not intended to be general halftoning algorithms, although we believe it might be possible to formulate one or both as variations on classical screening. They are designed to generate

rectangular superpixels in the foveated image, where the size of the rectangle is inversely related to the desired resolution. The desired resolution is first scaled and quantized to take integer values in the range $q(r,c) \in [0, K-1]$, and then $q(r,c)$ is then used to compute a modulus $m(r,c)$ at each pixel. The control signal is set to "no sharing" for pixels whose row or column index is divisible by their modulus. In vasiHalftone, the modulus is $m(r,c) = 2^{q(r,c)}$. For vasiHalftone2, the modulus is $m(r,c) = q(r,c)+1$ (typically with a different choice of *K*), giving much more flexibility on superpixel sizes. Figure 2 shows how nontrivial superpixel shapes and sizes can occur if non-sharing pixels do not align to form a continuous boundary.

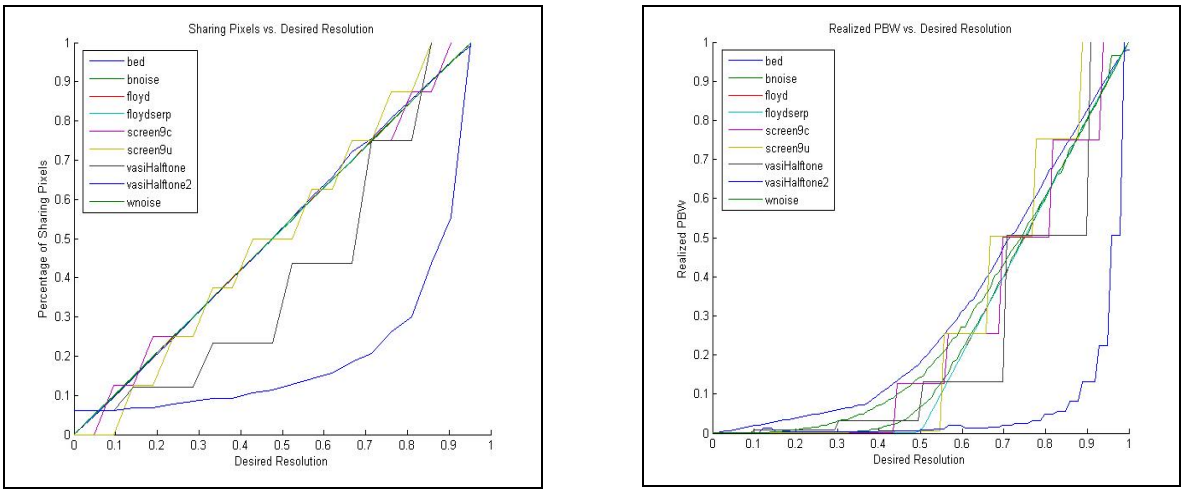

Figure 1: Percent of sharing pixels (left) and PBW (right) vs. desired resolution function

To simulate VASI foveation in software, we first construct a binary image to represent sharing behavior with one white pixel surrounded by eight black neighbors for each pixel in the original image. For each pixel, we switch the top neighbor to white if its VASI "share up" signal is set, and switch its left neighbor to white if its "share left" signal is set. We then use binary region labeling to build a many-to-one map from full-resolution pixels to superpixels and average the original pixel values in each region to determine the superpixel value. While this approach works, it consumes the majority of the processing time in the study.

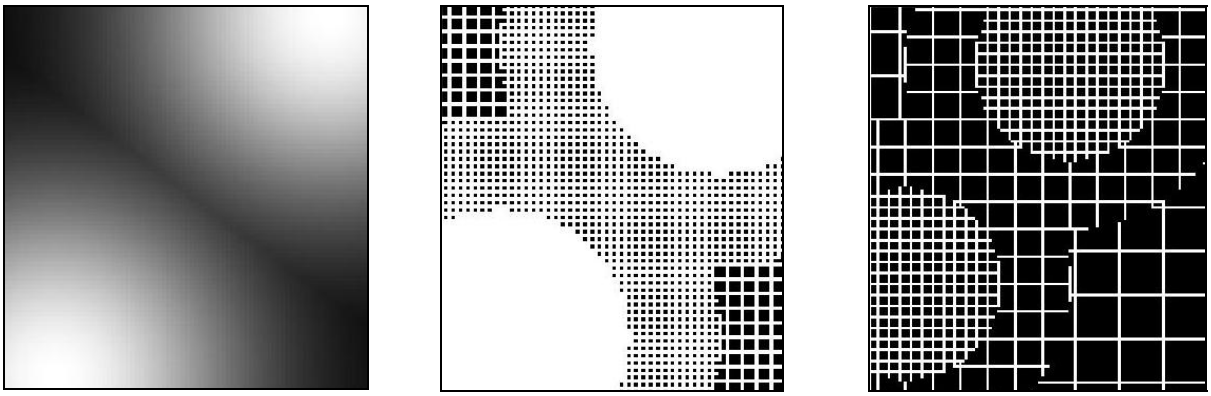

Figure 2: (left) Desired resolution, (center) vasiHalftone signal, (right) vasiHalftone2 signal.

4. Results

We normalize all FOM values on a given image by the maximum achieved by any approach on that image, and then average the normalized FOM performance for a given approach over all test images. This let us combine our results when FOM values varied between test images (e.g., the maximum PSNR value on *lena* was 25.2 dB, but the maximum on *board* was only 13.3 dB). Results are shown in Table 2. The last column in the table gives the "PBW Inflation Factor", which is the actual bandwidth divided by the target PBW (averaged over all test images). It measures the approach's ability to achieve a target PBW. We also computed the FOM values on two disjoint subregions of each image – the foveae and non-fovea regions. Good performance in foveae is important for accurate ATR, and good performance in non-fovea regions is important for effective ATA. Results on subregions are given in Table 3.

Our results show that Floyd & Steinberg error diffusion on a raster scan (*floyd*) offers the best FOM values while retaining precise bandwidth control (PBW inflation $\cong 1.0$). The *floyd* approach achieves good results in both fovea and non-fovea regions, implying that it will perform well for both ATR and ATA. While not the fastest halftoning approach, the *floyd* approach still has a low enough complexity to make it viable. The vasiHalftone and vasiHalftone2 approaches give consistently better FOM performance, but they consistently overshoot the desired PBW. They are essentially cheating by using extra bandwidth, and they do

not allow adequate control of the VASI camera frame rate. All the other approaches have lower FOM values, worse bandwidth control, or both.

Table 2: Summarized results for full images

	Figure of Merit				
Approach	PSNR	WSNR	LDM	UQI	PBW Inflation
bed	0.69	0.48	1.00	0.25	1.01
bnoise	0.72	0.51	1.00	0.34	1.02
floyd	0.85	0.71	1.00	0.51	1.06
floydSerp	0.83	0.67	1.00	0.46	1.06
screen9c	0.75	0.56	1.00	0.53	1.87
screen9u	0.83	0.73	1.00	0.81	3.01
vasiHT	1.00	1.00	1.00	0.91	1.79
vasiHT2	0.93	0.84	1.00	0.68	1.30
wNoise	0.69	0.48	1.00	0.29	1.03

Table 3: Summarized results for foveae (left) and non-fovea (right) regions

	Figure of Merit				
Approach	PSNR	WSNR	LDM	UQI	PBW Infl.
Bed	0.72	0.44	1.00	0.95	1.01
Bnoise	0.82	0.64	1.00	0.98	1.02
Floyd	0.86	0.78	1.00	0.98	1.06
floydSerp	0.80	0.72	1.00	0.98	1.06
screen9c	0.90	0.87	1.00	0.99	1.87
screen9u	0.84	0.64	1.00	0.99	3.01
vasiHT	1.00	1.00	1.00	1.00	1.79
vasiHT2	0.98	0.94	1.00	0.99	1.30
wNoise	0.79	0.55	1.00	0.97	1.03

	Figure of Merit				
Approach	PSNR	WSNR	LDM	UQI	PBW Infl.
bed	0.69	0.47	1.00	0.38	1.01
bnoise	0.72	0.50	1.00	0.41	1.02
floyd	0.85	0.70	1.00	0.56	1.06
floydSerp	0.83	0.66	1.00	0.52	1.06
screen9c	0.75	0.54	1.00	0.56	1.87
screen9u	0.84	0.72	1.00	0.85	3.01
vasiHT	1.00	1.00	1.00	0.91	1.79
vasiHT2	0.93	0.84	1.00	0.70	1.30
wNoise	0.69	0.47	1.00	0.37	1.03

5. Conclusions and Future Directions

Floyd & Steinberg error diffusion using a raster scan provides the best results overall. The vasiHalftone and vasiHalftone2 approaches give good FOM values but have poor bandwidth control. The other methods are inferior approaches for constructing the VASI control signal.

There are a number of interesting directions for future work. The elimination of the LUT-based inverse function through a closed-form approximation to the inverse mappings would significantly speed up control signal generation and reduce PBW inflation. Here we assumed that the “share up” and “share left” control signals are equal, which is not a requirement. Classical screening algorithms with more gray levels might give better PBW

control. We would also like to explore ways to improve the PBW control on the vasiHalftone and vasiHalftone2 approaches. Mixed strategies that follow one approach in foveal regions and another in non-foveae are also possible. Finally, we would prefer to measure the performance of ATR applications directly, instead of inferring their performance from our figures of merit.

6. 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*, 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.L. Evans, V. Monga, N. Damera-Venkata, "Variations on Error Diffusion: Retrospectives and Future Trends," *Proc. SPIE/IS&T Conf. on Color Imaging: Porcessing, Hardcopy, and Applications*, vol. 5008, Santa Clara, CA, Jan 20-24, 2003.

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

[6] 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.

[7] R. Floyd, L. Steinberg, "An adaptive algorithm for spatial grayscale," *Proc. SID'76*, pp. 75-77, 1976.

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

[9] J. Mannos, D. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 4, pp. 525-535, July 1974.

[10] 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.

[11] M. Massie, C. Baxter, J.P. Curzan, R. Etienne-Cummings, P. McCarley, "Vision Chip for Navigating and Controlling Micro Unmanned Aerial Vehicles," *IEEE ISCAS03*, May 2003.

[12] 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/.

[13] H. Kollnig, H.-H. Nagel, M. Otte, "Association of Motion Verbs with Vehicle Movements Extracted from Dense Optical Flow Fields," *Proc. ECCV'94*, vol. II, pp. 338-347, May 1994.