Halftoning-Inspired Methods for Foveation in Variable-Acuity Superpixel Imager* Cameras

Thayne R. Coffman[1,2], Brian L. Evans[1], and Alan C. Bovik[1]

[1]Center for Perceptual Systems, Dept. of Electrical and Computer Engineering,
The University of Texas at Austin, Austin, TX 78712-1084, USA
[2]21st Century Technologies, 4515 Seton Center Parkway, Suite 320, Austin, TX 78759, USA
(coffman@ece.utexas.edu, bevans@ece.utexas.edu, bovik@ece.utexas.edu)

***Abstract*- We compare approaches for generating binary control signals for variable acuity superpixel imager (VASI™) cameras. We foveate a set of images using control signals generated by various halftoning approaches and then measure their performance via figures of merit (FOMs). We find that two novel approaches provide superior FOM values but inferior bandwidth control, making them unsuitable for use with VASI™ cameras. Floyd-Steinberg error diffusion gives the best combination of FOM values and bandwidth control. Our contributions include a comparison of approaches, a lookup table method to improve bandwidth control, and two novel methods for binarizing VASI™ control signals.**

I. 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.

II. 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). Figure 1 illustrates this process. The sharing or non-sharing 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

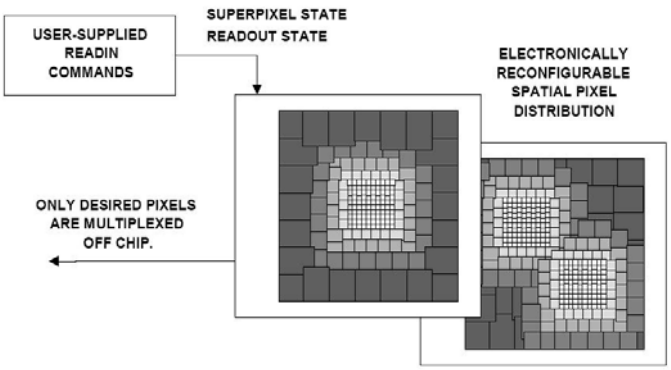

Figure 1: VASI™ camera foveation. Image from [1].

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, thereby 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 attempts to preserve the visual information in an image while reducing the word size used to represent each sample. This often (but not always) involves approximating continuous-valued images using only binary intensity levels [3]. Halftoning is well represented in the available literature and finds applications in grayscale and color printing, and display on low-cost displays (e.g., liquid crystal displays) [5]. 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

* VASI is a trademark of Nova Sensors, Inc.

Work performed at 21st Century Technologies was sponsored in part by the United States Air Force Research Labs under contract FA8851-05-C-0269.

quantization error introduced at each pixel among neighboring pixels [7]. More importantly, error diffusion shapes the quantization error into higher frequencies where the human visual system is less sensitive. For human observers, this results in more visually pleasing halftoned images. Serpentine scans are also sometimes used with error diffusion to further reduce visual artifacts by trading reduced horizontal and vertical artifacts for increased artifacts along diagonal directions. This can result in more pleasing images because the human visual system is less sensitive to high frequencies along diagonals than along horizontal and vertical directions. Error diffusion produces better visual results than classical screening (at an increased computational cost). Dithering with blue noise can yield even higher visual quality than error diffusion. Dithering with blue noise extends error diffusion by randomly perturbing the diffusion weights and/or directions. This again shapes the quantization noise into higher frequencies where the human visual system is less sensitive. It also makes the noise more isotropic (direction-independent), which can improve subjective visual quality [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 are interested in peak signal to noise ratio (PSNR), weighted signal to noise ratio (WSNR) [9], linear distortion measure (LDM) [9], and universal quality index (UQI) [10] for measuring the fidelity of our foveated images to the originals. 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%.

III. 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 the FOMs mentioned in Section II. 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]. Implementations for the standard halftoning approaches and all FOMs were taken from [12]. Other implementation was done in Matlab.

Not all FOMs are created equal. WSNR, LDM, and UQI were designed to predict subjective human perceptions of image quality. PSNR, in contrast, has been shown to be uncorrelated with subjective image quality. PSNR can, however, help predict the performance of some computer vision algorithms such as ATR. For WSNR and UQI, higher values indicate better visual quality. For LDM, lower values indicate better visual quality. For PSNR, higher values imply better performance for some algorithms, but they do not necessarily indicate better subjective visual quality.

Our desired resolution functions were defined to have one or two foveae superimposed over an image-wide minimum resolution. Each fovea is represented as a 2D Gaussian peak (with mean vector μ and covariance matrix Σ) 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" – see Section III.B), and white noise dithering ("wNoise").

A. *Inverse Function to Compensate for Sharing Geometry*

Halftoning the desired resolution signals directly yields very poor results. The halftoning algorithms all generally 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 is significantly affected by the geometry of exactly which pixels share charges.

Consider the variety of the control signals that we might choose. If we use a 'checkerboard' pattern (the control signal is one for those pixels whose row and column index sum to an even number), 50% of the pixels are one and the resulting PBW is nearly 0%. It is possible to engineer a tessellating control signal with 46% of the pixels set to one and a resulting PBW of 15%. In this signal, a lower percentage of the pixels are nonzero but the resulting PBW is significantly higher. At the extreme, if we set the control signal to one in the top half of the image and zero in the bottom half, 50% of the pixels are set to one and the resulting PBW is nearly 50%.

The effects of the particular geometries are illustrated in Figure 2 and Figure 3, which show the percentage of charge-sharing pixels and the achieved PBW as functions of desired resolution for a few of the approaches. 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 lookup tables improve our ability to achieve 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, which is reflected in our results below.

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

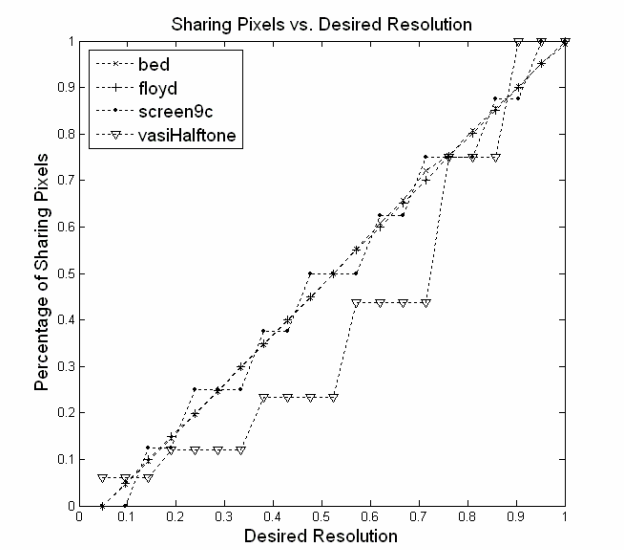

Figure 2: Percent of sharing pixels vs. desired resolution function.

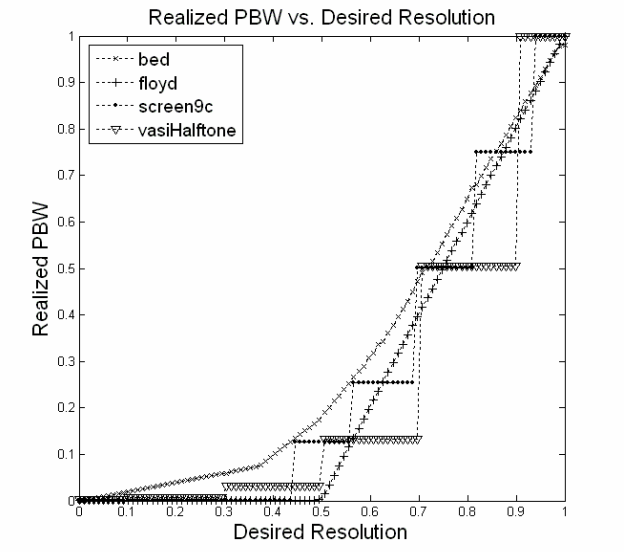

Figure 3: Realized PBW vs. desired resolution function

B. Specialized vasiHalftone and vasiHalftone2 Approaches

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 $q(r,c)$ 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, we use $m(r,c)=q(r,c)+1$ (typically with a larger choice for K than in vasiHalftone), which gives much more flexibility on superpixel sizes. Figure 4 shows how nontrivial superpixel shapes and sizes can occur if non-sharing pixels do not align to form a continuous boundary.

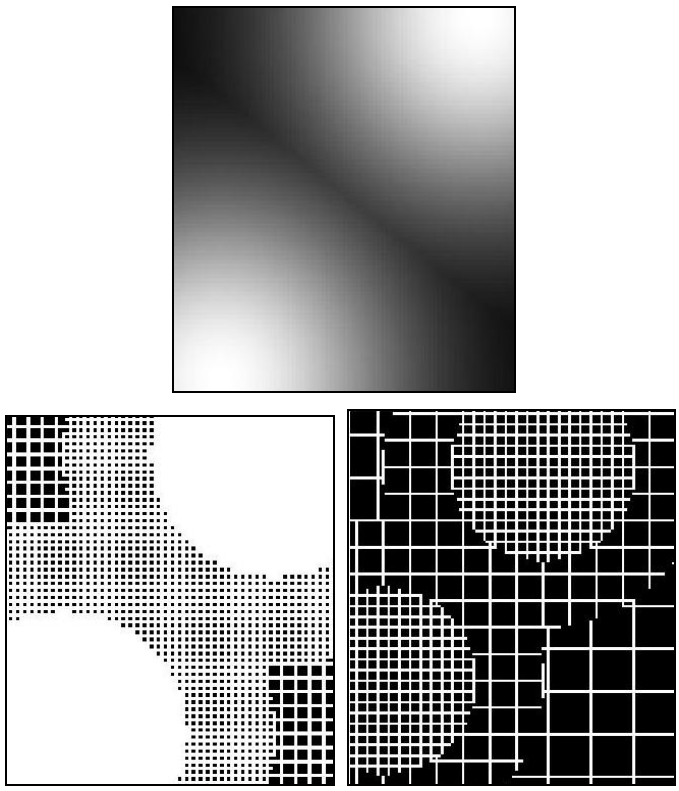

Figure 4: (top center) Desired resolution, (bottom left) Binary vasiHalftone control signal, (bottom right) Binary vasiHalftone2 control signal.

The vasiHalftone and vasiHalftone2 approaches are both point processes, meaning the control signal can be computed at each pixel independently or in parallel. Under our assumption that the vertical and horizontal control signals are equal, computing the binary control signals requires as few as one table lookup plus 5 operations per pixel. The two modulus comparisons can be accomplished by simple bit masking for vasiHalftone (but not vasiHalftone2).

C. Simulating VASI™ Charge Sharing Behavior

To simulate VASI™ foveation of the image $i(r,c)$ in software, we construct a binary image to represent the sharing behavior. The sharing behavior image $s(r',c')$ contains one white pixel surrounded by eight black neighbors for each pixel in $i(r,c)$. Thus pixel $s(2r,2c)$ corresponds to pixel $i(r,c)$,

and if image *i(r,c)* has size *[R,C]*, then *s(r',c')* has size *[2R+1,2C+1]*. For each pixel in *i(r,c)*, we switch *s(2r-1,2c)* to white if its VASI™ "share up" signal is set, and we switch *s(2r,2c-1)* to white if its "share left" signal is set. As a result, the pixels $s(2r_1,2c_1)$ and $s(2r_2,2c_2)$ will be in a connected region if and only if the pixels $i(r_1,c_1)$ and $i(r_2,c_2)$ are part of the same superpixel. We use binary region labeling on *s(r',c')* to build a many-to-one map from *i(r,c)* to superpixels and average the *i(r,c)* values for each superpixel to determine the superpixel value. While this approach lets us simulate VASI™ still images without a VASI™ camera (and it can be extended to simulate VASI™ video sequences), it consumes the majority of the processing time in the study.

IV. 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 foveal and non-foveal regions are given in Table 3 and Table 4, respectively. Figure 5 shows an example binary control signal and VASI™ image for the *floyd* method. Figure 6 shows similar examples for the *vasiHalftone* method.

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 ≅ 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 also consistently overshoot the desired PBW. Their lack of bandwidth control is undesirable because it artificially inflates FOM scores and will not allow adequate control of the VASI™ camera frame rate. All the other approaches have inferior FOM values, worse bandwidth control, or both.

V. 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 lookup table in the inverse function through a closed-form approximation to the inverse mappings of Section III.A would speed up control signal generation and reduce PBW inflation in some cases. Our assumption that the "share up" and "share left" control signals were equal is not a requirement in general, and may lower FOM performance or increase PBW inflation. Classical screening algorithms with more gray levels might give better PBW control than the 9-level approaches we tried. We would also like to explore ways to improve the PBW control on the vasiHalftone and vasiHalftone2 approaches. Mixed strategies that follow one halftoning 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.

TABLE 2
SUMMARIZED RESULTS FOR FULL IMAGES

	Figure of Merit				PBW
Approach	**PSNR**	**WSNR**	**LDM**	**UQI**	**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 FOVEAL REGIONS

	Figure of Merit				PBW
Approach	**PSNR**	**WSNR**	**LDM**	**UQI**	**Inflation**
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

TABLE 4
SUMMARIZED RESULTS FOR NON-FOVEAL REGIONS

	Figure of Merit				PBW
Approach	**PSNR**	**WSNR**	**LDM**	**UQI**	**Inflation**
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

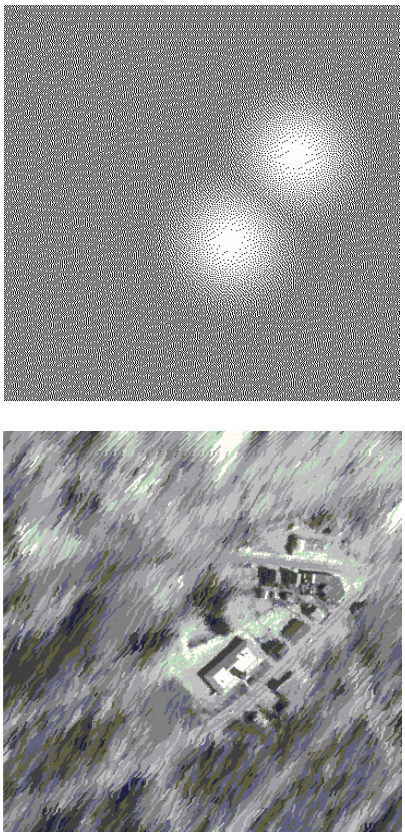

Figure 5: *Floyd* method:
(top) Binary control signal, (bottom) Resulting foveated image

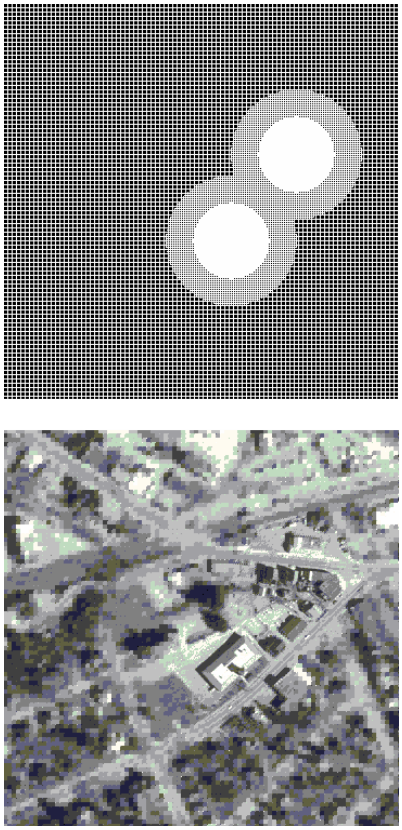

Figure 6: VasiHalftone method:
(top) Binary control signal, (bottom) Resulting foveated image

REFERENCES

[1] P. McCarley, M. Massie, and 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, and B. Pain, "Real-time reconfigurable foveal target acquisition and tracking system," *Proc. SPIE AeroSense Symp. on Acquisition, Tracking, and Pointing*, vol. 3692, pp. 300-310, July 1999.

[3] M. Mese and 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, and 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.L. 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 and L. Steinberg, "An adaptive algorithm for spatial grayscale," *Journal of the Society for Information Display*, vol. 17, no. 2, pp. 75-77, 1976.

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

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

[10] Z. Wang, A.C. Bovik, and 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, and P. McCarley, "Vision Chip for Navigating and Controlling Micro Unmanned Aerial Vehicles," *IEEE Int. Symp. on Circuits and Systems*, May 2003.

[12] V. Monga, N. Damera-Venkata, and B.L. Evans, *Halftoning Toolbox for Matlab.* Version 1.2 released July 25, 2005. Available online at http://www.ece.utexas.edu/~bevans/projects/halftoning/.

[13] H. Kollnig, H.-H. Nagel, and M. Otte, "Association of Motion Verbs with Vehicle Movements Extracted from Dense Optical Flow Fields," *Proc. European Conf. on Computer Vision*, vol. II, pp. 338-347, May 1994.