Error Diffusion for Digital Image Halftoning

Lieutenant Colonel Barry L. Shoop

Abstract

A novel approach to digital image halftoning is proposed based on a symmetric error diffusion algorithm, a new form of artificial neural network, and a smart-pixel optoelectronic architecture. Using an *error diffusion neural network*, all pixel quantization decisions are computed in parallel and therefore fully-symmetric error diffusion can be accomplished. As a result, visual artifacts such as directional hysteresis and transient behavior near boundaries which have become characteristic of classical unidirectional error diffusion techniques are eliminated. This new artificial neural network computes the halftoned image asymptotically faster than classical error diffusion techniques and has demonstrated faster image convergence than more recent massively parallel halftoning algorithms.

Introduction

Optics has long held the promise of high-speed, high-throughput parallel information processing. The focus of its early applications was on analog signal processing techniques such as the optical Fourier transform, matrix-vector processors, and correlators. During this period, optics was used almost exclusively for front-end, pre-processing of wide-bandwidth, high-speed analog signals which were subsequently digitally processed using electronic techniques. Digital signal processing, however, provides higher resolution, improved flexibility and functionality, and increased noise immunity over its analog counterparts and therefore is the preferred method for accurate signal processing. Since the majority of signals encountered in nature are continuous in both time and amplitude, the analog-to-digital (A/D) interface is generally considered to be the most critical part of the overall signal acquisition and processing system. Because of the difficulty in achieving high-resolution and high-speed A/D converters, this A/D interface has been and continues to be a barrier to the realization of high-speed, high-throughput systems.

Digital image halftoning[1] is an important class of A/D conversion within the context of image processing. Halftoning can be thought of as an image compression technique whereby a continuous-tone, gray-scale image is printed or displayed using strictly binary-valued pixels. Common applications of this technique include laser printing and xerography and more recently, facsimile. Error diffusion[2] is one approach to digital halftoning in which the error associated with a nonlinear quantization process is diffused within a local region and subsequent filtering methods employed in an effort to improve some performance metric such as signal-to-noise ratio. Classical error diffusion is a one-dimensional, serial technique in that the algorithm raster scans the image from upper-left to lower-right and as a result, introduces visual artifacts directly attributable to the halftoning algorithm itself. A fully-parallel implementation of the error diffusion algorithm, however, provides the advantage that all pixel quantization decisions are computed in parallel and therefore the error diffusion process becomes two-dimensional and potentially symmetric. Visual artifacts attributable to the halftoning algorithm are eliminated and overall halftoned image quality is significantly improved[3]. Because of the massive interconnect requirement associated with the 2-D neural network approach, an optoelectronic implementation is the preferred method to implementing this fully-parallel architecture.

In the process of developing an optoelectronic implementation of the error diffusion algorithm we have developed a new artificial neural network architecture based on the mathematical formalism of the error diffusion algorithm called an *error diffusion neural network*[4]. One significant advantage of this new implementation is that arbitrary size and shape diffusion kernels can be implemented thereby reducing the visual artifacts which have become characteristic of halftoned images. This can be accomplished with no penalty in terms of computation speed. The error diffusion neural network computes the halftoned image asymptotically faster than a conventional Hopfield-type neural network. The error diffusion neural network also provides full-rank connectivity across the entire image whereas other error diffusion techniques only provide local error diffusion. We have quantitatively compared the resulting halftoned images with those produced using other halftoning techniques using performance metrics such as the radially averaged frequency spectrum and anisotropy of continuous-tone gray-scale images. The halftoned images that result from this

new halftoning process produce artifact-free halftone images and are currently among the best quality halftoned images available.

Digital Image Halftoning

Digital halftoning, sometimes referred to as spatial dithering, is the process by which a continuous-tone, gray-scale image is rendered using only binary-valued pixels. The underlying concept is to provide the viewer of the image the illusion of viewing a continuous-tone image when, in fact, only black and white pixel values are used in the rendering.

There are a number of different methods by which this digital image halftoning can be accomplished. Ordered dither, produces an output by comparing a single continuous-tone input value against a deterministic periodic array of threshold values. Dispersed-dot ordered dither[5] occurs when the halftone dots are of a fixed size, while clustered-dot ordered dither simulates the variable-sized dots of printer's halftone screens in the rendering. Among the advantages of ordered dither techniques are speed of implementation and simplicity. The primary disadvantage is that ordered dither produces locally periodic patterns in the halftoned image which are visually objectionable. In contrast, halftoning using the error diffusion algorithm, first introduced by Floyd and Steinberg[2], requires neighborhood operations and is currently the most popular neighborhood halftoning process. In this algorithm, the error of the quantization process is computed and spatially redistributed within a local neighborhood in an effort to influence pixel quantization decisions within that neighborhood and thereby improve the overall quality of the halftoned image. Classical approaches to error diffusion, however, suffer from implementation constraints. Here, the algorithm raster scans the image, and for each individual pixel, a binary quantization decision is made based on the intensity of the individual pixel and the weighted error from pixels within a predefined diffusion region of previously processed pixels. As a result, the diffusion filter is necessarily causal resulting in undesirable visual artifacts. In an effort to improve halftone image quality, others[6] have proposed the use of two-dimensional neural networks to provide a mechanism for symmetric error diffusion. Convergence and stability issues in conjunction with implementation constraints limited the usefulness of most of these approaches.

The Error Diffusion Neural Network

The popularity of the neural network-based approach lies in the ability to minimize a particular metric associated with a highly nonlinear system of equations. Specifically, the problem of creating a halftoned image can be cast in terms of a nonlinear quadratic optimization problem where the performance metric to be minimized is the difference between the original and the halftoned images.

The Hopfield-Type Neural Network

The dynamic behavior of an N-neuron Hopfield-type neural network can be described by the following system of N nonlinear differential equations:

$$c\frac{du_i(t)}{dt} = -u_i(t) + \sum_j T_{i,j}\Im\left[u_j(t)\right] + x_i, \tag{1}$$

where i=1,2,... N, $\Im[\cdot]$ is a monotonically increasing sigmoid function, x_i is an input N-vector, and c is a scaling factor. In equilibrium, Equation (1) implies

$$u_i = x_i + \sum_j T_{i,j}\Im\lfloor u_j \rfloor. \tag{2}$$

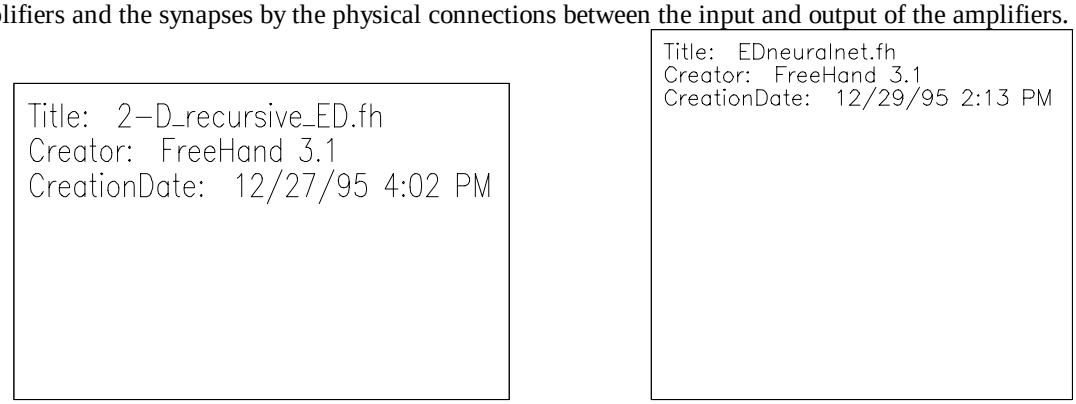
Hopfield showed that when the matrix of interconnection weights $\mathbf{T}$ is symmetric with zero diagonal elements and the high-gain limit of the sigmoid $\Im[\cdot]$ is used, the stable states of the N functions $y_i(t) = \Im\lfloor u_j \rfloor$ are the local minima of the energy function[7]

$$E = -\frac{1}{2}\mathbf{y}^T\mathbf{T}\mathbf{y} - \mathbf{x}^T\mathbf{y} \tag{3}$$

where $\mathbf{y} \in \{-1,1\}$ is an N-vector of quantized states.

The Error Diffusion Neural Network

Figure 1 shows the block diagram of a 2-D error diffusion architecture and an electronic implementation of a four-neuron error diffusion-type neural network. Here the individual neurons are represented as amplifiers and the synapses by the physical connections between the input and output of the amplifiers.



(a) (b)

Figure 1. (a) Block diagram of a 2-D error diffusion architecture, (b) Four-neuron electronic implementation.

In equilibrium, the error diffusion neural network satisfies

$$\mathbf{u} = \mathbf{W}(\mathbf{y} - \mathbf{u}) + \mathbf{x}. \tag{4}$$

For an $N \times N$ image, **W** is an $N^2 \times N^2$ sparse, circulant matrix derived from the original error diffusion weights $w_{i,j}$. If we define the coordinate system such that the central element of the error diffusion kernel is $(i, j) = (0,0)$, then the matrix **W** is defined as $W(i, j) = -w[(j-i) \operatorname{div} N, (j-i) \bmod N]$ where

$$x \operatorname{div} y = \left\lfloor \frac{x}{y} \right\rfloor \text{ if } xy \geq 0 \; or \left\lceil \frac{x}{y} \right\rceil \text{ if } xy < 0 \tag{5}$$

An equivalence to the Hopfield network can be described by

$$\mathbf{u} = \mathbf{A}(\mathbf{W}\mathbf{y} + \mathbf{x}), \tag{6}$$

where $\mathbf{A} = (\mathbf{I} + \mathbf{W})^{-1}$. Effectively, the error diffusion network includes a pre-filtering of the input image **x** by the matrix **A** while still filtering the output image **y** but now with a new matrix, **AW**. Recognizing that $\mathbf{AW} = \mathbf{I} - \mathbf{A}$ and adding the arbitrary constant $k = \mathbf{y}^T\mathbf{y} + \mathbf{x}^T\mathbf{A}\mathbf{x}$, we can write the energy function of the error diffusion neural network as

$$\hat{E}(x, y) = \mathbf{y}^T\mathbf{A}\mathbf{y} - 2\mathbf{y}^T\mathbf{A}\mathbf{x} + \mathbf{x}^T\mathbf{A}\mathbf{x} \;. \tag{7}$$

This energy function is a quadratic function which can be factored into

$$\hat{E}(x, y) = [\mathbf{B}\underbrace{(\mathbf{y} - \mathbf{x})}_{\text{error}}]^T [\mathbf{B}\underbrace{(\mathbf{y} - \mathbf{x})}_{\text{error}}] \tag{8}$$

where $\mathbf{A} = \mathbf{B}^T\mathbf{B}$. From Equation (8) we find that as the error diffusion neural network converges and the energy function is minimized, so too is the error between the output and input images. Computing the Fourier transform of this energy function we find that the error diffusion neural network minimizes the frequency weighted squared error between the output and input images.

If the neurons update independently, the convergence of the error diffusion network is guarenteed[8] if

$$\forall k : [\mathbf{AW}]_{k,k} \geq 0 \,. \tag{9}$$

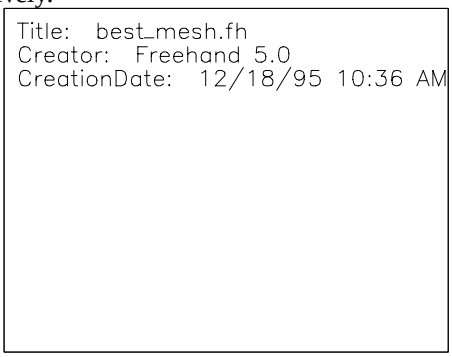
We find in practice, that even in a synchronous implementation, the halftoned images converge to a solution which results in significantly improved halftone image quality over other similar halftoning algorithms.

The Error Diffusion Filter

The purpose of the error diffusion filter is to spectrally shape the quantization noise in such a way that the error resulting from the quantization process is distributed to higher spatial frequencies which are less objectionable to the human visual system. In this application, the feedback filter for the error diffusion neural network was designed using conventional two-dimensional filter design techniques and resulted in the impulse response and frequency response shown in Fig. 2 (a) and (b), respectively.

0.0005	0.0020	0.0052	0.0069	0.0052	0.0020	0.0005
0.0020	0.0104	0.0249	0.0329	0.0249	0.0104	0.0020
0.0052	0.0249	0.0584	0.0767	0.0584	0.0249	0.0052
0.0069	0.0329	0.0767	•	0.0767	0.0329	0.0069
0.0052	0.0249	0.0584	0.0767	0.0584	0.0249	0.0052
0.0020	0.0104	0.0249	0.0329	0.0249	0.0104	0.0020
0.0005	0.0020	0.0052	0.0069	0.0052	0.0020	0.0005

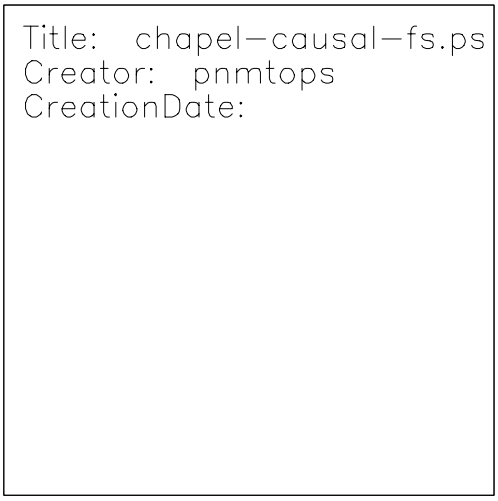
(a)



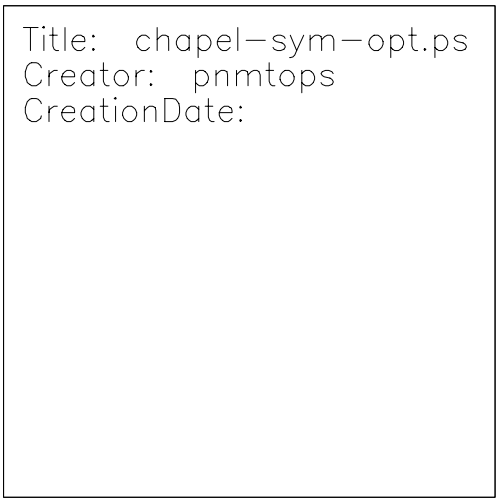
(b)

Figure 2. (a) Impulse response and (b) frequency response of one specific 7×7 error diffusion filter. ("•" represents the origin)

Figures 3 (a) and (b) show 348×348 halftoned images of the Cadet Chapel at West Point using classical halftoning techniques and our error diffusion neural network, respectively. Notice the uniformity of the pixel distributions in the upper-left of Figure 3(b) in cloud structure and also the improvement in the rendering of edges throughout the image. Figure 3(b) represents an *artifact-free* halftone image.



(a)

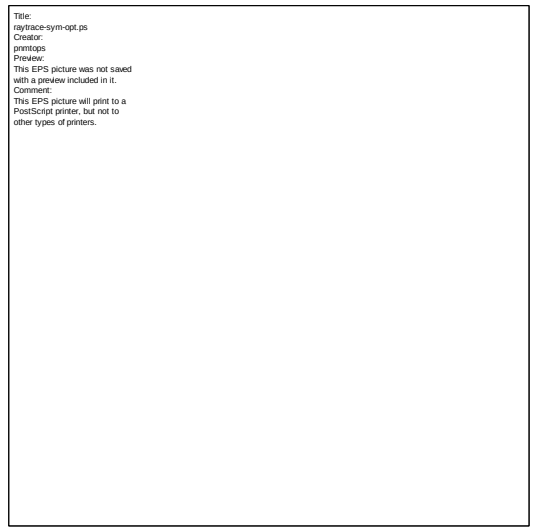


(b)

Figure 3. Halftoned images of the Cadet Chapel at West Point. (a) using classical error diffusion[3], and (b) using our error diffusion neural network.

Figure 4(a) shows a 512×512 halftoned image of a synthetic raytraced image[9] using a classical error diffusion halftoning technique and Figure 4(b) shows the same image after processing with our error diffusion neural network and the error diffusion filter coefficients shown in Figure 2(a). Figure 4(b) again results in a completely *artifact-free* halftoned image which accurately represents the original image. In this image, note particularly the detail in the reflection in the ball, in the black floor tiles, as well as in the shadows under the table.

(a)



(b)

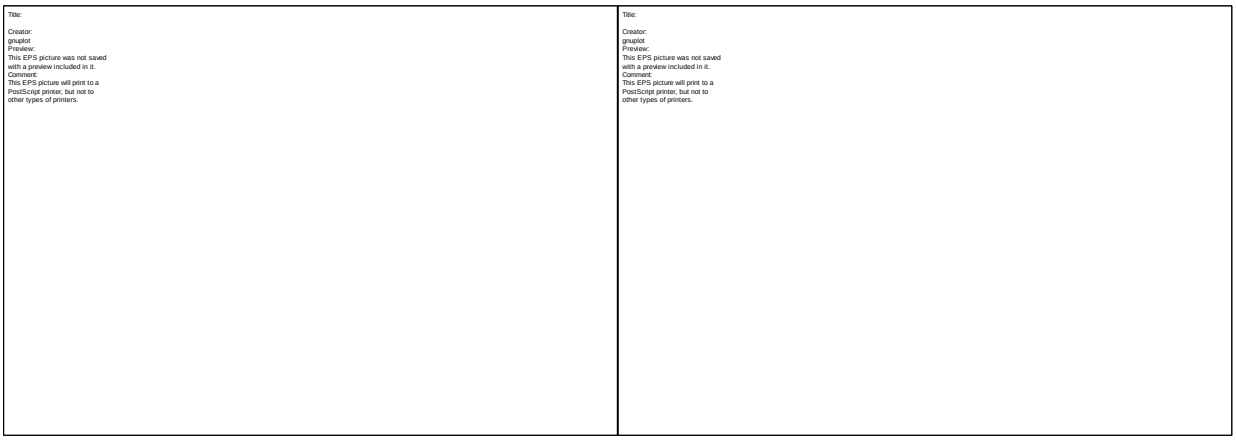
Figure 4. Halftoned images of a raytraced image. (a) using classical Floyd-Steinberg[2] error diffusion, and (b) using our error diffusion neural network.

A desirable attribute of a well-produced halftone of a fixed gray-scale image is radial symmetry of the dot pattern since the human visual system is extremely sensitive to directional artifacts. One quantitative metric for judging the quality of a halftoning algorithm is the anisotropy[1] which is a measure of the relative variance of frequency samples within a given annulus of the radially averaged power spectrum. Using Bartlett's Method[10] of averaging periodograms, ten periodograms of 256×256 output pixels from a halftone rendering of a constant gray-level input image are averaged to produce an estimate of the power spectrum $\hat{P}(\mathbf{f})$ of the halftone. This spectral estimate is then partitioned into annuli of width *Dw* and the radially averaged power spectrum is computed according to

$$P_t(f_t) = \frac{1}{N_r(f_r)} \sum_{i=1}^{N_r(f_r)} \hat{P}(\mathbf{f}) \tag{10}$$

where $N_r(f_r)$ is the number of discrete frequency samples in an annulus about the radial frequency f_r. The anisotropy is then defined as the ratio of the sample variance of the same frequency samples to the radially averaged power spectrum. If $\hat{P}(\mathbf{f})$ is perfectly radially symmetric, the anisotropy measure should be $\frac{1}{10}$ or -10 dB over the radial frequency range. When we computed the anisotropy on gray-scale images using our neural algorithm, the resulting anisotropy was consistently -10 dB with the plots being indistinguishable from the ideal curve.

Figure 5(a) shows one representative anisotropy curve from the error diffusion neural network with a constant input gray-level of 0.25. This anisotropy curve is consistent with the best anisotropy identified by Ulichney in Reference [1]. Figure 5(b) shows the radially-averaged frequency spectrum resulting from the same input. Spectral noise shaping to higher, blue frequencies, is clear. The cutoff frequency here is consistent with both the average input gray-level and the cutoff frequency of the error diffusion filter.



(a) (b)

Figure 5. (a) Anisotropy and (b) radially-averaged frequency spectrum the error diffusion neural network for a uniform gray-scale input of 0.25.

Extensions to Color

Within the past year, we have demonstrated the applicability of our error diffusion neural network to color halftoning, demonstrating excellent color image reproduction while simultaneously achieving a modest image compression ratio of 8:1. Figure 6 shows the results of this error diffusion network to color halftoning. Of particular interest is the detail in the upper-right of the image in the words: "We the People".

(a) (b)

Figure 6. Color halftoning using an error diffusion neural network. (a) Original color image (24 bits/pixel), and (b) halftoned color image (3 bits/pixel).

The results shown in Figure 6 were achieved by simply implementing three independent error diffusion networks in parallel, one for each of the primary colors in the image.

Conclusions

In our quest to develop an optical implementation of the error diffusion algorithm which provides improved halftoned image quality and performance, we have achieved several *firsts*. We have developed a new artificial neural network architecture based on the mathematical foundation of the error diffusion

algorithm which we call an *error diffusion neural network*. One significant advantage of this new implementation is that arbitrary size and shape diffusion kernels can be implemented thereby reducing the visual artifacts which have become characteristic of halftoned images. This can be accomplished with no penalty in terms of computation speed. The error diffusion neural network computes the halftoned image asymptotically faster than a conventional Hopfield-type neural network. This network also provides full-rank connectivity across the entire image in contrast to other error diffusion techniques which provide only local error diffusion. The halftoned images that result from this new halftoning process are currently the *best* halftoned images available. No other halftoning technique provides artifact-free images.

References

1. R. A. Ulichney, *Digital Halftoning*, Cambridge, MA: MIT Press, 1987.
2. R. Floyd and L. Steinberg, "An Adaptive Algorithm For Spatial Gray Scale," *SID 75 Digest* **36**, 35-36 (1975).
3. E. K. Ressler and B. L. Shoop, "High quality digital halftones from error diffusion networks," *Proc. of the Society for Information Display*, (San Diego, California) 506-509 (1996).

4. B. L. Shoop and E. K. Ressler, "An error diffusion neural network for digital image halftoning", *Proc. of the IEEE Workshop on Neural Networks*, (Institute of Electrical and Electronics Engineers, Signal Processing Society and Neural Networks Council, Boston, Massachusetts), 427-436 (1995).
5. B. E. Bayer, "An optimum method for two-level rendition of continuous-tone pictures," *Proc. IEEE Int. Conf. on Communication*, New York, NY, 26-11 - 26-15 (1973).
6. D. Anastassiou, "Error diffusion coding for A/D conversion," *IEEE Trans. Circuits Syst*., **36**, 1175-1186 (1989).
7. D.W Tank and J.J. Hopfield, "Simple neural optimization networks: A/D converter, signal detection circuit, and linear programming circuit," *IEEE Trans. Circuits Syst*., **33**, 533-541 (1986).
8. E. K. Ressler and B. L. Shoop, "Analysis and design of error diffusion networks for digital halftoning," *Proc. of the 20th Army Science Conference*, (Norfolk, Virginia), 191-195 (1996).
9. R. Geist, R. Reynolds, and D. Suggs, "A Markovian framework for digital halftoning," *ACM Trans. Graph*., **12**, 136-159 (1993).
10. M. S. Bartlett, *An Introduction to Stochastic Processes with Special Reference to Methods and Applications*, Cambridge University Press, New York, NY (1955).

Future Direction

One specific continuation of this research includes an investigation of an optimal approach to halftoning of color images, considering human color perception and the relationship and interconnectivity between the three individual networks. A theoretical analysis of this three-dimensional neural network is necessary to address such issues as convergence, stability, and network dynamics. Performance metrics must also be considered, providing quantitative results by which to compare these color halftoning results. A logical extension of this color halftoning which will also be investigated is an extension to more than three primary colors and the relationship to multi-spectral imagery and image compression.

Publications and Conference Presentations

B. L. Shoop, "Image processing algorithms for smart pixel architectures," Invited Paper, in *OSA Annual Meeting Technical Digest*, 1997 (Long Beach, CA) October 1997.

B. L. Shoop, J. Mait, and E. K. Ressler, "Optical error diffusion for analog-to-digital conversion," *Proceedings of the 20th Army Science Conference*, (Norfolk, Virginia), June 1996, *(Awarded Best Poster Paper in Microelectronics and Photonics Session).*

E. K. Ressler and B. L. Shoop, "Analysis and design of error diffusion networks for digital halftoning," *Proceedings of the 20th Army Science Conference*, (Norfolk, Virginia), June 1996.

E. K. Ressler and B. L. Shoop, "High quality digital halftones from error diffusion networks," *Proceedings of the Society for Information Display*, (San Diego, California), May 1996.

Cadet Publications and Presentations

J. Kettula, A. H. Sayles, and B. L. Shoop, "The neuron of an error diffusion neural network," *Proceedings of the Tenth National Conference on Undergraduate Research*, (Asheville, North Carolina), pp. 1124-1127, April 1996.

Cadet Involvement

John Kettula ('96)

Collaboration

Lieutenant Colonel Eugene K. Ressler	*Department of Electrical Engineering and Computer Science*
Professor Thomas D. Wagner	*Department of Electrical Engineering and Computer Science*
Dr. Diana Thomas	*Department of Mathematics*