MIT Media Laboratory Perceptual Computing Group Technical Report #288.
Also appears in: Proc. of the First IEEE Int. Conf. on Image Processing, Austin, TX, Nov. 1994.

ORIENTATION-SENSITIVE IMAGE PROCESSING WITH M-LATTICE – A NOVEL NON-LINEAR DYNAMICAL SYSTEM

Alex Sherstinsky and Rosalind W. Picard

Media Laboratory, E15-383
Massachusetts Institute of Technology, Cambridge, MA 02139
shers@media.mit.edu and picard@media.mit.edu

ABSTRACT

Researchers in image processing have long recognized the importance of modeling the human observer. Although a full human vision model remains elusive, orientation detection, one of the key components of human vision, can be directly incorporated into a variety of image processing algorithms. Orientation detection also provides cues that allow an algorithm to adapt to inhomogeneities in images. In this paper we show how the M-lattice system, a new non-linear dynamical system, can easily incorporate orientation sensitivity for two different types of problems. First, simultaneous adaptive filtering and non-linear restoration is illustrated for fingerprint enhancement. Second, constrained non-linear optimization is illustrated for halftoning in a "hand-drawn" style.

1. INTRODUCTION

The physiological evidence for orientation detectors in the human visual system system [1] has led to widespread recognition of the importance of orientation for both human vision and perceptually-based image processing. Not only is orientation perceptually significant, but image processing algorithms that sense local orientation are better equipped to adapt to inhomogeneous data.

In this paper we show how local orientation information can be smoothly incorporated into two different types of image processing problems, while simultaneously accomplishing a desired image processing goal such as image halftoning.

The computational vehicle for both experiments is the M-lattice, a non-linear dynamical system recently introduced into the signal processing community for a variety of applications formulated as constrained non-linear optimization and as pattern-extraction [2]. The M-lattice system was derived from the reaction-diffusion model, first proposed by Turing in 1952 in order to explain mammal coat patterns [3]. Although primarily explored for biological pattern formation [4], reaction-diffusion systems can perform image processing operations that emphasize oriented patterns [5]. Recently, directional image processing was performed using anisotropic non-linear diffusion [6]. Yet, the M-lattice system is more flexible than both the reaction-diffusion system and the anisotropic non-linear diffusion, thereby providing a common framework for a variety of image processing problems.

2. BACKGROUND: M-LATTICE SYSTEM

We briefly review the essentials of the M-lattice system [2]. Let $\psi_i(t) \in \Re$ be a state variable as a function of time at each lattice point i, where $i = 1, \ldots, N$. Let $\chi_i(t)$ be an output variable, obtained from $\psi_i(t)$ via $\chi_i(t) = g(\psi_i(t))$. The "warping" function, $g(u)$, is a saturating piece-wise linear non-linearity with an arbitrarily large number of segments. The values of $\chi_i(t)$ will correspond to the intensities of the pixels in the output image at the time when the system has converged. Construct $\vec{\psi}(t)$ and $\vec{\chi}(t)$ by concatenating $\psi_1(t), \ldots, \psi_N(t)$ and $\chi_1(t), \ldots, \chi_N(t)$, respectively into column vectors.

Definition 2.1 *Suppose that a given function, $\Phi(\vec{\chi}(t))$, is continuous, twice-differentiable, and bounded above. Let the matrix $\mathbf{A}$ be real, symmetric, and negative-definite:* $\mathbf{A} \in \Re^{N \times N}$, $\mathbf{A} = [a_{ij}]$, $\mathbf{A} = \mathbf{A}^T$, *and* $\forall\ i\ \lambda_i[\mathbf{A}] < 0$. *Then the M-lattice system* [1] *is the following non-linear dynamical system:*

$$\frac{d\vec{\psi}(t)}{dt} = \mathbf{A}\vec{\psi}(t) + \vec{\nabla}_{\vec{\chi}}\Phi(\vec{\chi}(t)). \tag{1}$$

Notice the right-hand side contains two components – a linear function of the state variables and the gradient of a typically non-linear function of the warped state variables. Compared to reaction-diffusion systems, the linear function, $\mathbf{A}\vec{\psi}(t)$, generalizes diffusion to any linear filter, while the gradient term allows many types of the non-linear reaction.

The convergence and stability of non-linear dynamical systems is an area of much research. We have shown that a subclass of the M-lattice system possesses asymptotic convergence properties, regardless of the initial conditions [2]. Thus far, no proof of total stability exists for the system in (1). However, fixed points of the form $\vec{\chi} \in \{-1, 1\}^N$ are asymptotically stable [7]. In the applications we discuss below, the general M-lattice system exhibited convergence in computer simulation.

In non-linear optimization, $\Phi(\vec{\chi})$ is the objective function to be maximized. Using the results of Vidyasagar [8],

This research was sponsored by Hewlett-Packard Laboratories, Palo Alto, California.

[1]This is the definition adapted for the present paper. The general M-lattice system is defined in [7].

it can be shown that for certain types of objective functions, the M-lattice system converges to the (appropriately defined) local maxima of $\Phi(\vec{\chi})$ with respect to $\vec{\chi}$ [7]. Thus, in many situations it is advantageous to use the M-lattice system for non-linear optimization. Of the two examples described in Sections 4 and 5, the first uses a 1-lattice system for non-linear filtering; the second uses another 1-lattice system for non-linear optimization. Both employ orientation sensitivity.

3. ESTIMATING LOCAL ORIENTATION

Rather than compute orientation at all possible angles and then decide which angle dominates, we employ the computation-saving "steerable" set of basis filters described in [9]. Steerable filters have been shown to give a good match to orientation perception by humans [10]. The output of the steerable filters at each pixel i gives the angle, $\theta_i \in [-\pi, \pi]$, and relative strength (or magnitude), $m_i \in [0,1]$, of the dominant orientation present at that pixel.

For the applications that follow, we will use the orientation to guide the action of the M-lattice system. For example, to design a low-pass adaptive filter that rotates to the dominant orientation, denote the diagonal matrix of variances by $\mathbf{V}_i$ and the rotation matrix by $\mathbf{\Theta}_i$:

$$\mathbf{V}_i = \begin{bmatrix} \sigma_{i,x}^2 & 0 \\ 0 & \sigma_{i,y}^2 \end{bmatrix}, \quad \mathbf{\Theta}_i = \begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix}. \tag{2}$$

The relative sizes of $\sigma_{i,x}^2$ and $\sigma_{i,y}^2$ depend on m_i and determine the skewness of filters with respect to the dominant orientation:

$$\sigma_{i,y}^2 = \frac{L}{2}(1-m_i), \quad \sigma_{i,x}^2 = L - \sigma_{i,y}^2, \tag{3}$$

where $L \times L$ is the size of the filter mask in pixels. Let $\vec{n} \in \mathcal{Z}^2$ be the pixel position. Then the (unnormalized) oriented low-pass filter is given by:

$$h_i^u(\vec{n}) \quad = \quad \exp\left\{-\vec{n}^T \mathbf{\Theta}_i^T \mathbf{V}_i \mathbf{\Theta}_i \vec{n}\right\}. \tag{4}$$

4. NON-LINEAR RESTORATION OF FINGERPRINTS

A typical fingerprint identification system contains a pre-processing step for binarizing the original scanned fingerprint image. A binary fingerprint image is more amenable to classification than a gray-scale fingerprint image [11].

For fingerprint images, the binary output can be obtained by a standard halftoning algorithm, but this can aggravate any noise or scratches that are present. We propose a pre-processing scheme that not only binarizes the original fingerprint image, but also removes artifacts that can hinder the identification process. The method uses the ability of the M-lattice system to excite growing spatial waves, which are the signature of reaction-diffusion systems. The motivation for using the reaction-diffusion aspect of the M-lattice system is that fingerprint images have distinct patterns of thin curves with bifurcations resembling zebra stripes, and reaction-diffusion systems have been shown to be capable of some fingerprint enhancement [5], [12]. Reinforcing the harmonics that create fingerprint curves will emphasize the essentials of the fingerprint, while orientation-sensitive filtering will suppress the artifacts due to the image acquisition.

Let $\vec{n} \in \mathcal{Z}^2$, $a(\vec{n})$ and $h(\vec{n})$ be FIR filters , and $s(\vec{n}) \in [-1,1]$ be the original finely-quantized input image signal [2]. Consider the following version of the M-lattice system, defined in (1):

$$\frac{d\psi(\vec{n},t)}{dt} \quad = \quad a(\vec{n}) * \psi(\vec{n},t) + s(\vec{n}) - h(\vec{n}) * \chi(\vec{n},t). \tag{5}$$

For notational convenience, we have dropped the subscript i. Nevertheless, it is implied that each pixel i has its own $a(\vec{n})$ and $h(\vec{n})$. Using (5) for the restoration and enhancement of fingerprints is convenient, because it can be designed to produce binary outputs and, as we now show, possesses the desired pattern-formation properties [7], [13].

Choose $a(\vec{n})$ and $h(\vec{n})$ such that the unique interior fixed point, $\psi(\vec{n}) \in (-1,1)$, of (5) is at $s(\vec{n})$. The procedure for doing this will be given below. Denote the DFT of the filters by $A(\vec{k})$ and $H(\vec{k})$. By Definition 2.1, it is necessary for the matrix $\mathbf{A}$, representing the linear term of the M-lattice system, to be made negative-definite. For the system in (5), $\mathbf{A}$ is block-circulant and symmetric, containing $a(\vec{n})$ in the first row. Then all the $A(\vec{k})$ coefficients must be negative, because they are proportional to the eigenvalues of this $\mathbf{A}$ matrix. Before $\psi(\vec{n},t)$ reaches the clipping levels, (5) simplifies to:

$$\frac{d\psi(\vec{n},t)}{dt} \quad = \quad s(\vec{n}) + \left(a(\vec{n}) - \frac{1}{T}h(\vec{n})\right) * \psi(\vec{n},t). \tag{6}$$

Taking the DFT of both sides of (6) yields:

$$\frac{d\psi(\vec{k},t)}{dt} \quad = \quad S(\vec{k}) + \left(A(\vec{k}) - \frac{1}{T}H(\vec{k})\right)\psi(\vec{k},t), \tag{7}$$

whose solution for each $\vec{k}$ is:

$$\psi(\vec{k},t) \quad = \quad \left[S(\vec{k}) + \frac{S(\vec{k})}{F(\vec{k})}\right]\exp\left\{F(\vec{k})t\right\} - \frac{S(\vec{k})}{F(\vec{k})}, \tag{8}$$

where $F(\vec{k}) \stackrel{\text{def}}{=} A(\vec{k}) - \frac{1}{T}H(\vec{k})$, and the initial condition is set to $S(\vec{k})$, the DFT of the original image. Making $F(\vec{k})$ positive for a set of spatial frequencies creates the onset of growing spatially-stationary waves [7], [13].

For fingerprint restoration, filtering should delineate the ridges, while canceling fluctuations in the DC level and suppressing noise. In fingerprints, the ridges are not collinear, making it difficult to achieve this enhancement. The problem is solved here by using the adaptive filters, $H(\vec{k})$ and $A(\vec{k})$, which incorporate the orientation information of the original image, in a way that assigns negative DFT coefficients to the frequency bands corresponding to the ridges.

[2] Actually, $s(\vec{n})$ is obtained by scaling and shifting the original 256-gray-level image, $I(\vec{n})$: $s(\vec{n}) = \left(\frac{I(\vec{n})}{128}\right) - 1$.

Starting with (4), the filters are designed as follows. First, $h^u(\vec{n})$ is normalized to produce $h^n(\vec{n})$. Then the intermediate forms of $a(\vec{n})$ and $h(\vec{n})$ are computed:

$$h^f(\vec{n}) = \delta(\vec{n}) - \frac{1}{1-h^n(\vec{0})}\left(\delta(\vec{n}) - h^n(\vec{n})\right), \tag{9}$$

$$\begin{aligned} a^f(\vec{n}) &= -\left(\delta(\vec{n}) - h^f(\vec{n})\right) = h^f(\vec{n}) - \delta(\vec{n}) \\ &= -\frac{1}{1-h^n(\vec{0})}\left(\delta(\vec{n}) - h^n(\vec{n})\right) \\ &= \frac{1}{1-h^n(\vec{0})}\left(h^n(\vec{n}) - \delta(\vec{n})\right), \end{aligned} \tag{10}$$

where $\delta(\vec{n})$ is a unit sample in 2-D.

Based on (9) and (10), $h(\vec{n})$ is designed to have negative DFT coefficients at the spatial harmonics corresponding to the ridges. Since $0 < h^n(\vec{n}) < 1 \;\forall\; \vec{n}$, (10) guarantees that $A^f(\vec{k}) \leq 0$. Then by slightly perturbing $a^f(\vec{0})$, $A^f(\vec{k})$ is transformed into $A(\vec{k}) < 0 \;\forall\; \vec{k}$, thereby assuring that $\mathbf{A}$ is negative-definite. Substituting the resulting $a(\vec{n})$ and $h(\vec{n}) = \delta(\vec{n}) + a(\vec{n})$ into (5) with $T = 1$ confirms that the fixed point of (5) is indeed $\psi(\vec{n}) = s(\vec{n})$ as required.

The system goes unstable along the ridges, essentially giving them an infinite boost. The $\frac{1}{T}$ factor amplifies this effect as T is gradually decreased during the computer simulation from its initial value of $T = 1$. The warping keeps the system variables within the allowable range.

Figure 1(a) contains a typical scanned and finely quantized fingerprint image from the NIST database [14]. The original image is 512×512 pixels and was low-pass filtered and down-sampled by a factor of 2 in each dimension in order to speed up the computation. From the figure, it can be seen that the original fingerprint is corrupted by a number of scratches, and several regions are obscured by uneven illumination.

As shown in Figure 1(b), the common fingerprint halftoning method, based on adaptive median filtering and thresholding [11], only makes these artifacts more apparent, because it increases the image's contrast [7]. The adaptive threshold is set to the average of the minimum and the maximum gray levels within some neighborhood surrounding each pixel of the original fingerprint image. The optimal size of the window was determined to be 5×5 pixels by trial and error. Other standard halftoning methods, such as ordered dither or error diffusion, will perform poorly also because they have no built-in restoration mechanism and will halftone both signal and noise alike.

Figure 1(c) displays the fingerprint, processed by the M-lattice. The scratches have been removed and the unevennesses in the DC levels throughout the image have been eliminated. Essential detail such as ridges and bifurcations appear as continuous black curves, distinctly enhanced against a noise-free white background. Moreover, ridges and bifurcations have been extended even into the regions where they are barely detectable in the original image. This illustrates the celebrated synergetic property of reaction-diffusion: the emergence from noise of a spatial pattern, whose qualitative characteristics are pre-determined by the system's parameters [3], [4]. Using a Connection Machine (CM-2), the final image is produced in 25 iterations at the time step of 0.1 sec for the total time of less than 3 seconds or 1 minute including the system time and the I/O.

The use of orientation within the M-lattice system differs from existing approaches in that it accomplishes restoration and halftoning simultaneously. Alternatively, one might use a two-stage system consisting of some conventional image restoration algorithm, followed by adaptive-thresholding. However, there are two arguments in favor of using the M-lattice system-based approach. First, the attainable signal-to-noise ratios can be very large. This is due to the fact that the M-lattice system applies the filters, $a(\vec{n})$ and $h(\vec{n})$, a large (infinite in the limit) number of times by the virtue of being a continuous-time system. Second, no separate halftoning step is needed, since the M-lattice system binarizes the image.

5. ORIENTATION-DEPENDENT HALFTONING AS NON-LINEAR PROGRAM

In this section, we consider the problem of synthesizing a binary caricature that brings out the directional content of an image. The resulting halftoning method must yield an image that appears similar to the original gray-scale image in some indirect sense. A least-squares halftoning approach is appropriate for this task, because it can employ an explicit model of perception as the measure of performance [15]. Here we show how to implement such an approach using the M-lattice system.

Suppose $\vec{n} \in \mathcal{Z}^2$; $s(\vec{n}) \in [-1, 1]$ is the finely quantized original input image signal; $y(\vec{n}) \in \{-1, 1\}$ is the output halftone image; and $h(\vec{n})$ is a 2-D filter (not necessarily the same as $h(\vec{n})$ in the previous section). Let $\mathbf{B} = \mathbf{H}^T\mathbf{H}$, where $\mathbf{H}$ is a circulant matrix with $h(\vec{n})$ in the first row. The problem of halftoning can be stated as a non-linear program:

$$\min_{\vec{y}} \frac{1}{2}\vec{y}^T\mathbf{B}\vec{y} - (\mathbf{B}\vec{s})^T\vec{y} \tag{11}$$

$$\text{subject to constraints: } y_i^2 - 1 \quad \geq \quad 0, \tag{12}$$

where the vectors are the standard concatenations of the corresponding sequences. The particular form of constraints, (12), forces each pixel to assume binary values.

In order to solve this problem using the M-lattice system, we combine the objective function to be minimized, (11), with the N constraints, (12), into the Lagrangian cost functional with the help of the Karush-Kuhn-Tucker conditions [16]:

$$\min_{\vec{y}} \mathcal{L}(\vec{y}), \quad \text{where}$$

$$\mathcal{L}(\vec{y}) = \frac{1}{2}\vec{y}^T\mathbf{B}\vec{y} - (\mathbf{B}\vec{s})^T\vec{y} + \frac{1}{2}\sum_i p_i(y_i^2 - 1), \tag{13}$$

$$p_i \leq 0, \quad p_i(y_i^2 - 1) = 0. \tag{14}$$

The Lagrange multipliers, p_i, are the varying penalty terms that enforce the constraints according to (14). As a result, the unconstrained minimization of $\mathcal{L}(\vec{y})$ in (13) produces the optimal halftone image.

The optimization problem, (13), is "programmed" onto the M-lattice system, (1), by setting $\vec{y}$ equal to $\vec{\chi}$, $\Phi(\vec{\chi})$ to

$-\mathcal{L}(\vec{y})$, and taking partial derivatives. This yields:

$$\frac{d\vec{\psi}(t)}{dt} = \mathbf{A}\vec{\psi}(t) + \mathbf{B}\vec{s} - \mathbf{B}\vec{\chi}(t) - \mathbf{P}\vec{\chi}(t), \qquad (15)$$

where $\mathbf{P} = \text{Diag } \{p_1, \ldots, p_N\}$. The elements of $a(\vec{n})$ are chosen so as to guide the system towards an optimum corresponding to a perceptually-pleasant halftone. It has been shown that $\mathbf{A} = \mathbf{B} - \mathbf{I}$ is a good choice, because it filters out objectionable correlated spatial patterns [2].

Halftoning with the Hopfield network [17] would be similar, but requires setting $b_{ii} \geq 0$. Otherwise, the optimal values of y_i will not be binary [7]. Since no effort is made here to design $\mathbf{H}$ in a way that would result in $b_{ii} \geq 0$, the non-linear constraints provide the only mechanism for driving the output pixels to the limits of the gray scale.

Treating halftoning as a non-linear programming problem and solving it with the M-lattice system offers considerable flexibility in the choice of the quality metric and in the functional form of constraints. In order to demonstrate this flexibility, we incorporated orientation detection into the halftoning quality metric. The adaptive filter matrix, $\mathbf{H}$, was designed using (4) so as to include the information about the dominant orientation at each pixel of the original image, shown in Figure 2(a). Figure 2(b) displays the result, which exhibits more of the line and curve features found in hand-drawn "halftones" (such as the Wall Street Journal portraits) [7]. According to the poster provided by the Wall Street Journal Classroom Edition program, the entire process is done by hand and takes an artist from three to five hours [18]. In contrast, the simulation of the M-lattice system implementation on the CM-2 takes 3000 iterations at the time step of 0.01 sec for the total time of approximately 3 minutes including the system time and the I/O.

6. SUMMARY

Explicit use of local orientation information in image processing can lead to better adaptive algorithms for enhancement as well as to novel image processing such as "hand-drawn" style halftoning. We have described the incorporation of orientation information into two different image processing applications, and shown how the M-lattice system, a new non-linear dynamical system, can be used for both implementations. Additionally, the two applications show how the M-lattice system is well-suited to image processing problems which require either simultaneous binarization and enhancement, or which perform non-linear constrained optimization.

Acknowledgements

The authors would like to thank Tom Minka for implementing the M-lattice system simulator on the CM-2.

7. REFERENCES

[1] H. D. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *J. Physiology*, vol. 195, pp. 215–243, 1968.

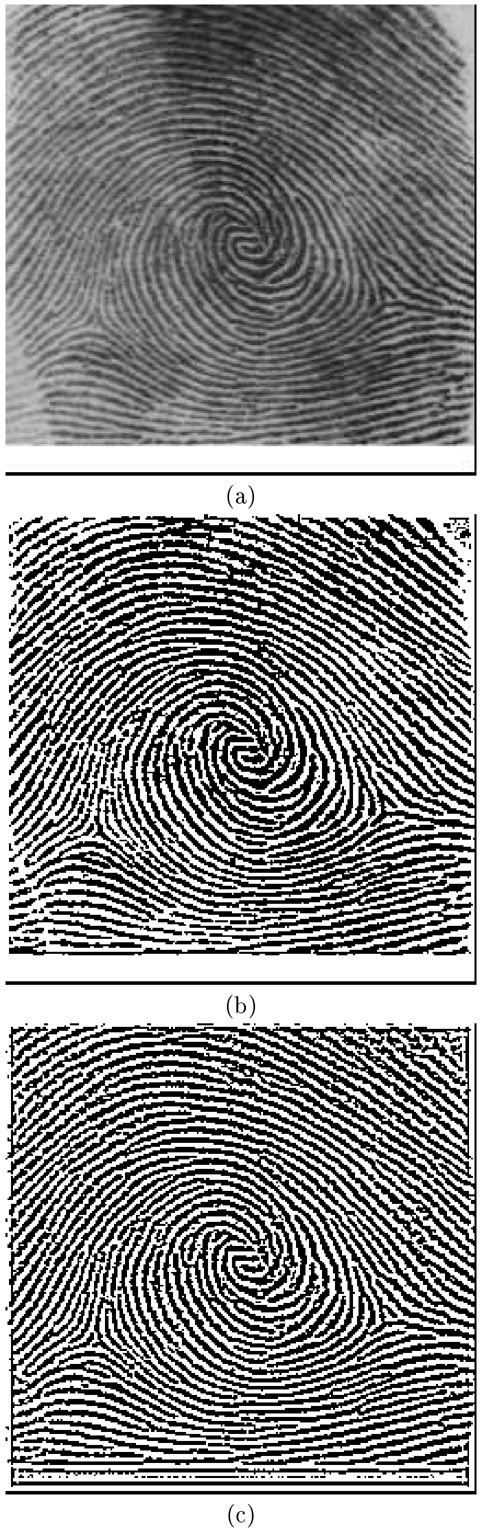

Figure 1: Restoration and halftoning of fingerprints. (a) the original "fingerprint" image; (b) the "fingerprint" image halftoned by a standard adaptive-threshold method; (c) the "fingerprint" image restored and halftoned by the M-lattice system operating in the reaction-diffusion mode utilizing orientation information at each pixel of the original.

(a)

(b)

Figure 2: Orientation-sensitive halftoning. (a) the original "Reagan" image; (b) the "Reagan" image adaptively halftoned using orientation information at each pixel of the original.

[2] A. Sherstinsky and R. W. Picard, "M-lattice: A novel non-linear dynamical system and its application to halftoning," in *Proc. IEEE Int. Conf. Acous., Speech and Sig. Proc.*, (Adelaide, Australia), Apr. 1994.

[3] A. M. Turing, "The chemical basis of morphogenesis," *Phil. Trans. Royal Soc. London*, vol. 237, no. B, pp. 37–72, 1952.

[4] J. D. Murray, *Mathematical Biology.* Berlin: Springer-Verlag, 1993.

[5] C. B. Price, P. Wambacq, and A. Oosterlinck, "Applications of reaction-diffusion equations to image processing," in *Third Int. Conf. Image Proc. and App.*, (Warwick, UK), July 1989.

[6] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pat. Anal. and Mach. Intell.*, vol. 12, pp. 629–639, July 1990.

[7] A. S. Sherstinsky, *M-Lattice: A System For Signal Synthesis And Processing Based On Reaction-Diffusion.* ScD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1994.

[8] M. Vidyasagar, "Minimum-seeking properties of analog neural networks with multilinear objective functions," Tech. Rep. (Unpublished), Bangalore, India, 1993.

[9] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pat. Anal. and Mach. Intell.*, vol. PAMI-13, pp. 891–906, Sept. 1991.

[10] R. W. Picard and M. Gorkani, "Finding perceptually dominant orientations in natural textures," *Spatial Vision, Spec. Julesz Issue*, 1994. To Appear; also avail. as Percep. Comp. TR #229, M.I.T. Media Lab.

[11] A. P. Russo, "An automatic system for fingerprint analysis," Master's thesis, Rensselaer Polytechnic Institute, Troy, NY, 1986.

[12] G. Zorpette, "A horse of a different color," *IEEE Spectrum*, pp. 17–18, July 1992.

[13] A. Sherstinsky and R. W. Picard, "Restoration and enhancement of fingerprint images using M-Lattice – a novel non-linear dynamical system," in *Proceedings of International Conference on Pattern Recognition*, (Jerusalem, Israel), Oct. 1994.

[14] National Institute of Standards and Technology, "Special database 4: 8-bit gray scale images of fingerprint image groups," July 1993.

[15] T. N. Pappas and D. L. Neuhoff, "Least-squares model-based halftoning," in *Proc. SPIE Conf. Elect. Imaging Sci. and Tech.*, (San Jose, CA), Feb. 1992.

[16] M. S. Bazaraa and C. M. Shetty, *Nonlinear Programming.* New York: John Wiley and Sons, 1979.

[17] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," *Proc. Nat. Acad. Sci., USA*, vol. 81, pp. 3088–3092, May 1984.

[18] "Private communications with Richard Kelman, Dow Jones," May 1994.