Pattern Recognition Letters 15 (1994) 219–225
North-Holland

March 1994

PATREC 1173

A qualitative analysis of the resistive grid kernel estimator

Wendy L. Poston *[,a], George W. Rogers [b], Carey E. Priebe [b], Jeffrey L. Solka [b]

[a] *Naval Surface Warfare Center, Dahlgren Division, G33, Dahlgren, VA 22448-5000, USA*
[b] *Naval Surface Warfare Center, Dahlgren Division, B10, Dahlgren, VA 22448-5000, USA*

Received 14 May 1992; revised 12 July 1993

Abstract

Poston, W.L., et al., A qualitative analysis of the resistive grid kernel estimator, Pattern Recognition Letters 15 (1993) 219–225.

The ability to estimate a probability density function from random data has applications in discriminant analysis and pattern recognition problems. A resistive grid kernel estimator (RGKE) is described which is suitable for hardware implementation. The one-dimensional linear RGKE is compared to a kernel estimate using Gaussian kernels, and simulations are presented using both continuous and quantized data. The nonlinear form of the RGKE is shown to have desirable properties, such as the ability to detect discontinuities in the density function.

Keywords. Pattern recognition, kernel estimation, nonparametric probability density estimation.

Introduction

In this paper, the problem of estimating the probability density function $f(x)$ of a given sample of N real observations $X_1, \ldots, X_N$ will be considered. The examples presented here are for the one-dimensional case only. However, this concept can be extended to multi-dimensional data, subject to the constraints of circuit complexity and processing time. Related work which uses the resistive grid in a two-dimensional application can be found in Rogers et al. (1992). The approach studied here is the kernel estimator from Silverman (1986) using a kernel that can be implemented by a resistive grid network (Mead, 1989). This network is based on a nonlinear discrete dendritic process model. A discussion of kernel estimates and their application to neural networks can be found in Specht (1990) and a review of research concerning pattern classification is presented in Lippmann (1989). It will be shown qualitatively, that using the linear RGKE provides results comparable to an estimate formed using a Gaussian kernel. Results using the nonlinear RGKE will illustrate the ability of the resistive grid to detect discontinuities in the density. Finally, conclusions and possible applications will be discussed.

* Corresponding author.

Continuous resistive grid kernel estimator

The resistive network described in Mead (1989) is an implementable analog VLSI model of a biological dendritic process. This model provides a real-time analog means of computing the weighted average of many input signals or observations. The voltage at a node is determined by the weighted average of the

SSDI 0167-8655(93)E0036-N

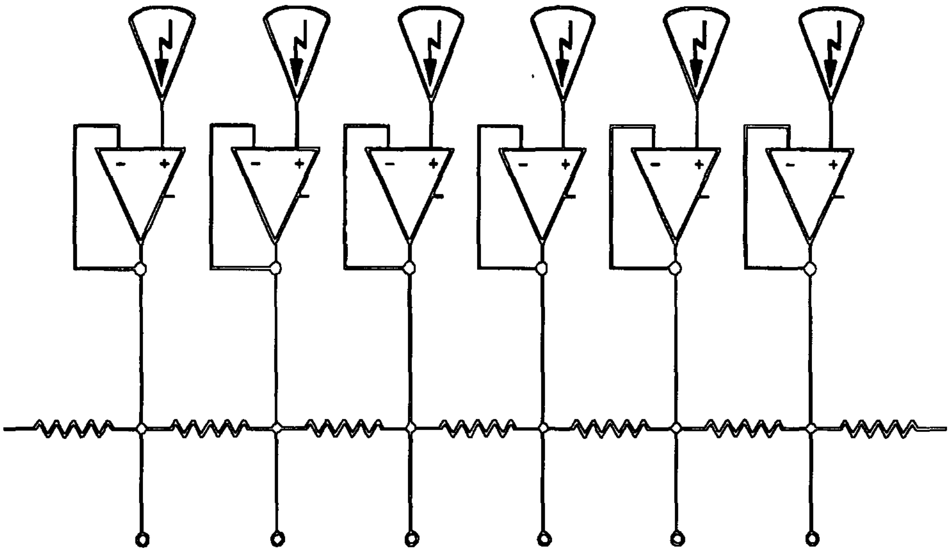

Figure 1. Schematic of the resistive grid network.

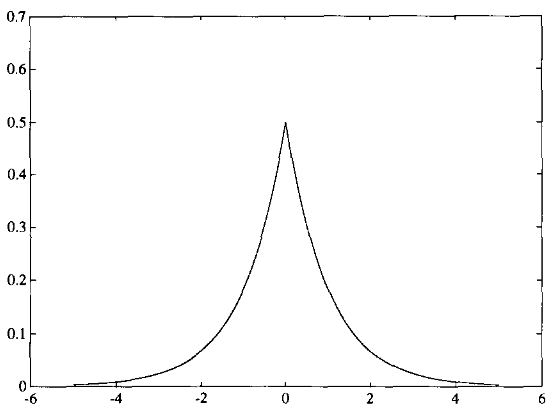

Figure 2. The continuous RGKE kernel is related to the Laplacian distribution.

inputs. Since the amplitude of the voltage due to a single input decreases exponentially with distance, signals that are farther away will carry less weight. A schematic of the one-dimensional resistive network is shown in Figure 1.

The equation for the general univariate kernel estimator of constant window width h is

$$\hat{f}(x)=\frac{1}{Nh}\sum_{i=1}^{N}K\left(\frac{x-X_i}{h}\right), \tag{1}$$

where N is the number of data points, h controls the degree of smoothness, and K denotes a kernel. For a one-dimensional continuous resistive network, the following equation can be used as the kernel centered at y

$$K_{RG}(y)=A\,\mathrm{e}^{-|x-y|/L}, \qquad L=\frac{1}{\sqrt{RG}} \tag{2}$$

where L is the characteristic length, R is the resistance per unit length, G is the conductance to ground per unit length and A is a normalization constant. The resistance R is related to the resistance in the cytoplasm of the neural process, and the conductance G corresponds to the leakage through the membrane to the surrounding fluid (Mead 1989). When relating this to density estimation, the characteristic length L can be identified with the window width h. The normalized voltage at each node of the resistive network determines the estimate of the density function. Note that the resistive grid kernel of Eq. (2) is just the Laplacian distribution centered at zero

$$f(x)=\frac{\lambda}{2}\mathrm{e}^{-\lambda|x-\mu|}, \quad -\infty<x<\infty \tag{3}$$

with $L=1$ and $A=1/2$. This density function is illustrated in Figure 2.

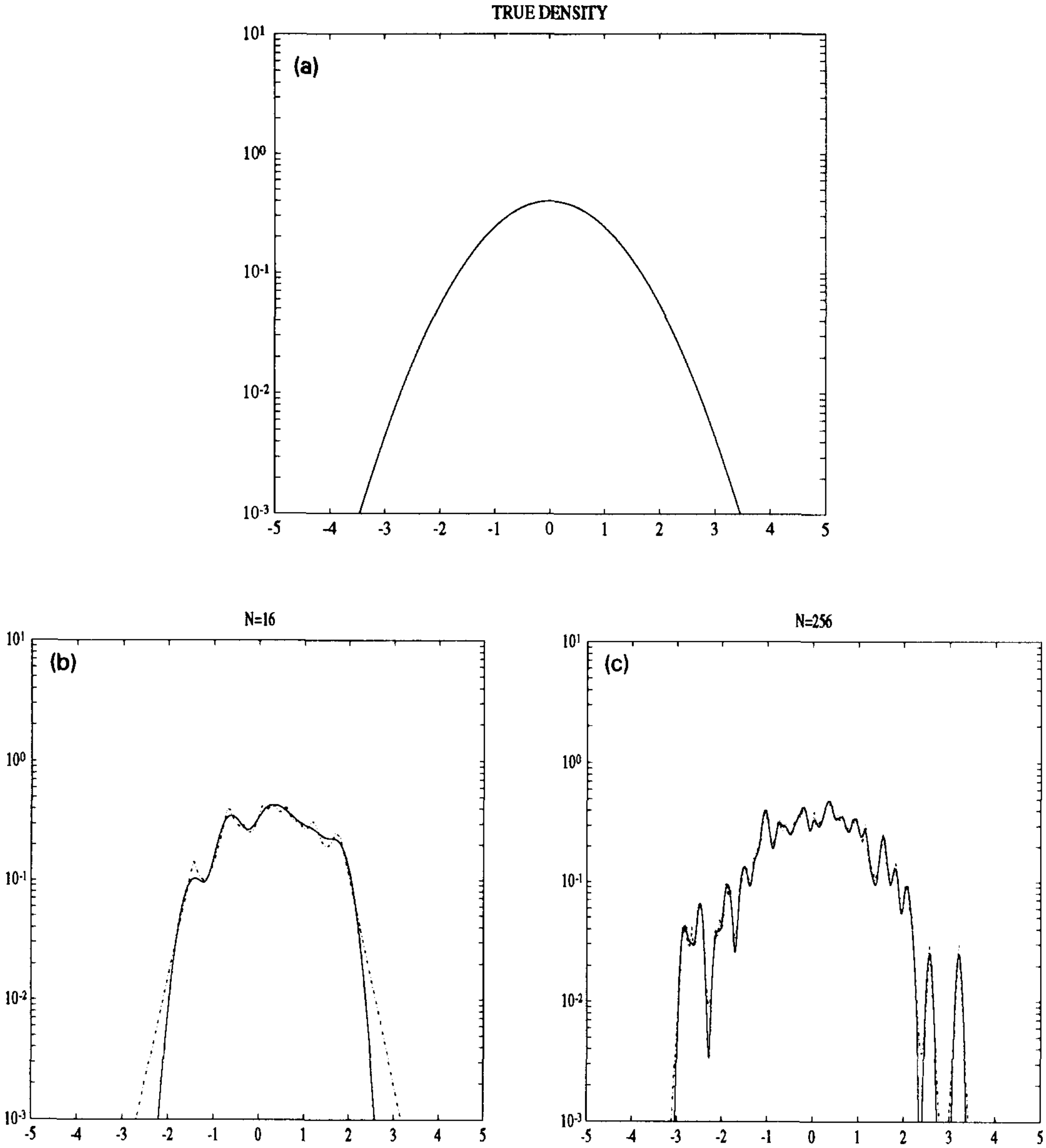

Figure 3. Comparison between an estimate using a Gaussian kernel (solid) and a Resistive Grid Kernel (dotted), with $h=1/\sqrt{N}$. The true density, depicted in 3a, is Gaussian.

Figure 3 shows a comparison of kernel estimates of a normal density using the continuous RGKE and Gaussian kernels for various window widths given by $h=1/\sqrt{N}$ (Duda and Hart, 1973). It can be seen from these plots that, as expected, the RGKE yields a density estimate similar to one using Gaussian kernels.

Since the continuous RGKE is a bounded Borel function satisfying the following conditions

$$\int |K_{RG}(x)|\, dx < \infty ,$$

$$\int K_{RG}(x)\, dx = 1 ,$$

$$|xK_{RG}(x)| \to 0, \quad |x| \to \infty \tag{4}$$

with $h \to 0$ and $Nh \to \infty$ as $N \to \infty$, then the estimate

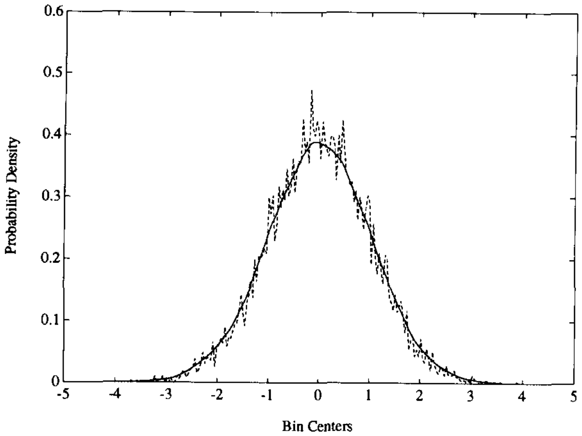

Figure 4. Comparison between an estimate using a histogram (dotted) and the discrete linear RGKE (solid) with $L=5$ and $N=10{,}000$. The true distribution is Gaussian.

$\hat{f}_{RG}(x) \to f(x)$ and the variance of this estimate goes to zero in probability as $N \to \infty$ (Silverman, 1986).

Discrete linear resistive grid kernel estimator

When implementing the RGKE, quantized inputs must be used (Mead, 1989). The probability density estimate for the kth node of an infinite discrete linear resistive network is given by

$$\hat{f}_{RGk} = A \sum_{i=-\infty}^{\infty} \frac{N_i \, e^{-|k-i| \ln \gamma}}{\sqrt{4L^2+1}}, \tag{5}$$

where N_i is the number of observations in bin i, A is a normalization constant and

$$\gamma = 1 + \frac{1}{2L^2} - \frac{1}{L}\sqrt{1+\frac{1}{4L^2}}. \tag{6}$$

This equation is valid under the assumption that linear superposition holds. This can be compared with a discrete or lattice version of Eq. (1).

$$\hat{f}(x) = \frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{(x-X_i)}{h}\right) = \frac{1}{Nh}\sum_{i=-\infty}^{\infty} N_i K\left(\frac{(x-X_i)}{h}\right), \tag{7}$$

where $N = \sum N_i$ with N_i giving the number of observations at the ith lattice site or node. If N_i of Eq. (7) is identified with the voltage at each node and normalization is neglected, the equivalence of Eqs. (5) and (7) can be seen.

To simulate the linear RGKE, $N=10{,}000$ data points are drawn from a quantized normal density with zero mean and variance one. The probability density function is then estimated only at the centers of the bins. Figure 4 illustrates the estimate from a discrete linear resistive grid with a characteristic length of 5.0. This shows that the RGKE can yield a smooth density estimate of the data.

A discontinuity is introduced in the same data to evaluate the performance of the linear discrete RGKE. This is illustrated in Figure 5a, which clearly shows the edge in the data. The estimates from the RGKE for different values of L are shown in Figures 5b–d. As these plots demonstrate, the discontinuity can be detected by decreasing the characteristic length L. However, smoothness in the estimate is lost when the characteristic length is made small enough to detect discontinuities in the density. This fundamental tradeoff in the choice of the smoothing parameter L (or h in the traditional formulation) is an inherent characteristic of linear kernel estimators.

Discrete nonlinear resistive grid kernel estimator

While the above equations for continuous and linear discrete RGKE allow for a development of the theory with respect to kernel estimators, including the nonlinearities of the network has some benefits. An application of Kirchoff's circuit laws to the ith node in Figure 1 yields the following equation

$$V_i = \frac{GV_i^{\text{in}} + \dfrac{V_{i-1}}{R_{i-1}} + \dfrac{V_{i+1}}{R_i}}{G + \dfrac{1}{R_{i-1}} + \dfrac{1}{R_i}}, \quad V_i^{\text{in}} = \beta N_i \tag{8}$$

which governs the voltage at the ith node of a resistive grid network when linearity is not assumed. The parameter β is a scaling factor controlling the degree of nonlinearity, V_i^{in} is the fixed input voltage to the ith node, R_i denotes the resistive element between nodes i and $i+1$, and G denotes the fixed conductance to ground at each node. The resistance used in Eq. (8) is derived from the current/voltage relation-

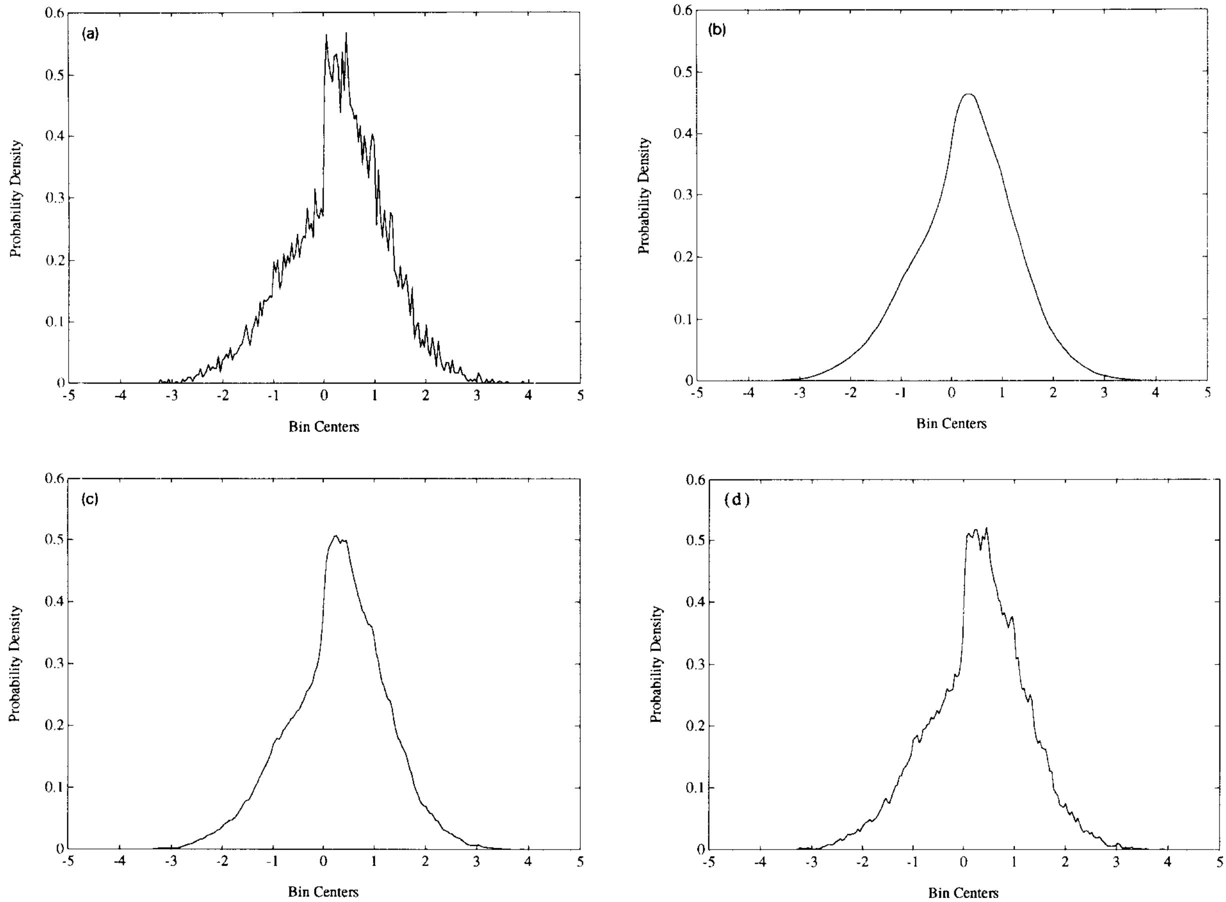

Figure 5. (a) Estimate of the density using a histogram with $N=10{,}000$. The true density if piecewise Gaussian. (b) Linear RGKE estimate with $L=5$. (c) Linear RGKE estimate with $L=2$. (d) Linear RGKE estimate with $L=1$.

ship for a nonlinear resistive element (Mead, 1989)

$$I \propto \tanh\left(\frac{V_i - V_{i+1}}{2}\right). \quad (9)$$

An application of Ohms law ($R=\Delta V/I$) leads to

$$R_i = R_0 \frac{V_{i+1} - V_i}{2} \Big/ \tanh \frac{V_{i+1} - V_i}{2}, \quad (10)$$

where R_0 is the zero signal resistance value. The estimate of the density function is given by the voltages at each node,

$$\hat{f}_{NLRGi} = \frac{V_i}{\sum_{i=1}^{N} V_i} = \frac{V_i}{\sum_{i=1}^{N} V_i^{\text{in}}} = \frac{V_i}{\beta N} \quad (11)$$

with equivalent normalization factors by virture of conservation of charge and the input scaling indicated in Eq. (8).

This set of coupled nonlinear equations is applied to the discontinuous data of the previous section. Figures 6a–c illustrate the probability density estimates for different degrees of nonlinearity. A value of 0.2 for β yields results similar to the linear case; the curve is very smooth and no discontinuity is detected in the density. However, as more nonlinearity is allowed, the discontinuity in the density becomes apparent. With $\beta=1.0$ the curve remains smooth while the discontinuity in the density function is clearly shown. Thus, the inherent nonlinearity in a resistive grid network can actually improve its probability density function estimator qualities.

For an additional illustrative example of the nonlinear RGKE, a piecewise uniform distribution on the interval (0,1) is used. The resistive grid estimates for $\beta=0.2$ and $\beta=1.0$ are shown in Figure 7. The results are similar to the previous case. As the scaling factor β is increased, the density becomes less smooth and the discontinuities are more evident. This example also demonstrates that the nonlinear RGKE will detect more than one discontinuity in the probability density function.

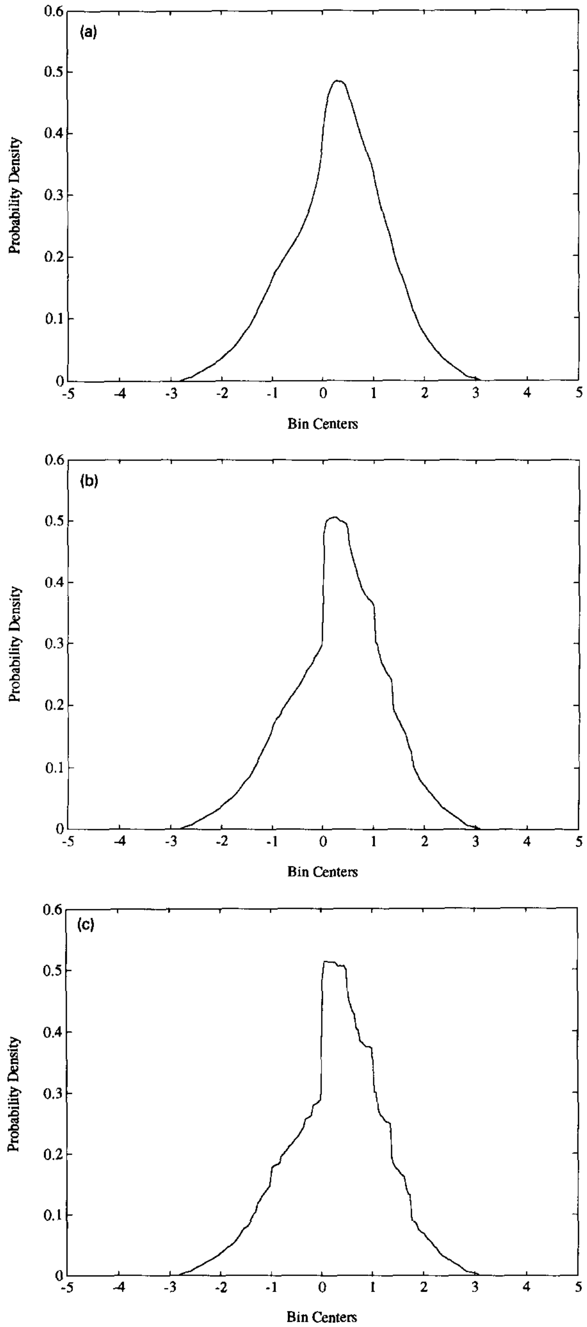

Figure 6. (a) Nonlinear RGKE estimate of the density shown in Fig. 5 with $\beta=0.2$ and $N=10{,}000$. (b) Nonlinear RGKE estimate with $\beta=1$. (c) Nonlinear RGKE estimate with $\beta=2$.

Summary

This study has shown qualitatively that the resistive grid kernel estimator yields results comparable to probability density estimates derived using Gaussian kernels. The ability of the nonlinear RGKE to detect discontinuities in the density, while continu-

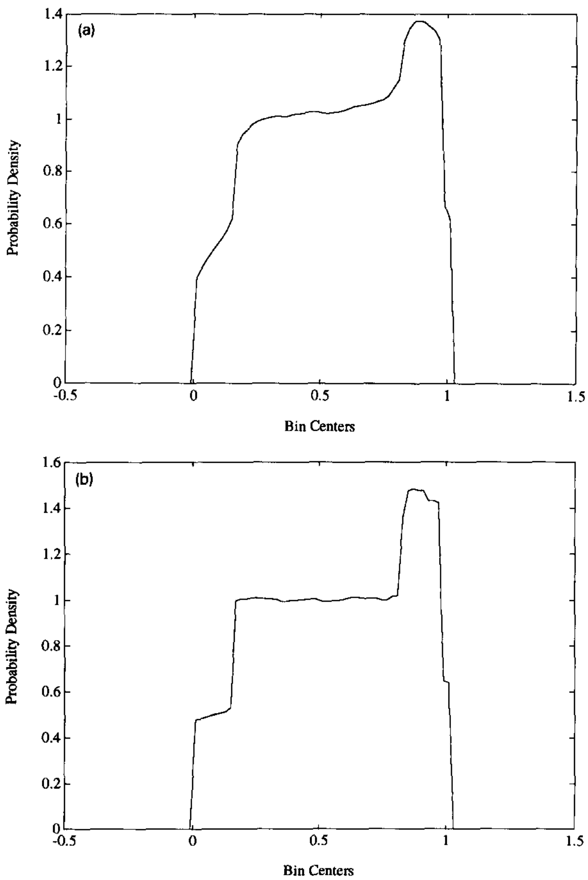

Figure 7. (a) Nonlinear RGKE estimate of a piecewise uniform density with $\beta=0.2$ and $N=10{,}000$. (b) Nonlinear RGKE estimate with $\beta=1$.

ing to produce a smooth function, illustrates its usefulness as an estimator. Since the RGKE is suitable for hardware implementation, it is possible to apply it to problems in discriminant analysis and pattern classification when real-time responses are required. Future research efforts include a quantitative statistical analysis of the approach, as well as the hardware implementation of the resistive grid network.

Acknowledgements

The authors gratefully acknowledge the comments on this algorithm by Prof. John Rice, UC Berkeley, and wish to acknowledge the help and interest of Dr. Jim Maar, both under the auspices of the Joint Services Advisory Group (JSAG) program. Also, the authors wish to acknowledge the support of the NSWCDD Independent Research Program.

References

Duda, R.O. and P. Hart (1973). *Pattern Classification and Scene Analysis.* Wiley, New York.

Lippmann, R.P. (1989). Pattern classification using neural networks. *IEEE Comm. Magazine*, Nov., 47–64.

Mead, C. (1989). *Analog VLSI and Neural Systems.* Addison-Wesley, Reading, MA.

Rogers, G.W., J.L. Solka, C.E. Priebe and H.H. Szu (1992). Optoelectronic computation of waveletlike-based features. *Optical Engineering* 31 (9), 1886–1892.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, London.

Specht, D.F. (1990) Probabilistic neural networks. *Neural Networks* 3, 109–118.