ELSEVIER

Computers in Biology and Medicine 34 (2004) 355–370

Computers in Biology and Medicine

www.elsevierhealth.com/locate/compbiomed

An information-theoretic approach to estimating ultrasound backscatter characteristics

Renata Smolíková[a,b], Mark P. Wachowiak[a], Jacek M. Zurada[b,*,1]

[a] *Department of Computer Engineering and Computer Science, University of Louisville, Louisville, KY 40292 USA*
[b] *Department of Electrical and Computer Engineering, University of Louisville, Louisville, KY 40292 USA*

Received 30 August 2002; received in revised form 30 May 2003; accepted 30 May 2003

Abstract

Analysis of backscatter in the ultrasound echo envelope, in conjunction with ultrasound B-scans, can provide important information for tissue characterization and pathology diagnosis. Statistical models have often proven useful in modeling backscatter. In this paper, an innovative approach to backscatter analysis based on generalized entropies and neural function approximation is presented. Entropy measures are shown to provide accurate estimates of scatterer density, regularity, and SNR of the amplitude distribution. Specific scattering distributions need not be assumed. Experimental results on ground truth envelopes show that generalized entropies can be used to accurately estimate backscatter properties.

Keywords: Ultrasound; Backscatter analysis; Speckle; Renyi entropy; Tsallis entropy

1. Introduction

Biomedical ultrasound (US) is a popular biomedical imaging modality because of its cost effectiveness, relative safety, and because ultrasound scanning can be performed in a variety of clinical settings. However, US brightness-mode scans, or B-scans, are characterized by a phenomenon known as speckle. Speckle has a negative effect on image interpretation. It contributes to the low signal-to-noise ratio (SNR) in B-scans, and makes human and computerized analysis very difficult. Speckle is present both in B-scans and in the radio-frequency (RF) signal and corresponding RF envelope.

* Corresponding author. Tel.: +1-502-852-6314; fax: +1-502-852-3940.
E-mail address: jmzura02@louisville.edu (J.M. Zurada).
[1] Sponsored in part by the Systems Research Institute, Polish Academy of Science, Newelska 6, 01-447 Warsaw, Poland.

0010-4825/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0010-4825(03)00072-6

Speckle is not additive noise. It results from the coherent accumulation of reflections from random scatterers, or small structures in tissue, such as cells, that reflect and scatter the incoming US beam. In fact, speckle is an inherent property of the US image formation process, and is highly correlated with the RF signal. Because of this correlation, speckle processes, henceforth known as backscatter, can also supply useful clinical information for tissue identification, tissue density, and whether a pathology is present.

Backscatter has both deterministic and stochastic characteristics, and therefore, the amplitude of the backscatter signal (envelope) can be modeled by statistical distributions. The parameters of these distributions indicate characteristics such as density (the number of scatterers within the resolution cell of the transducer), and scatterer amplitude. The latter is related to the size of the scatterers.

These distributions are physically-based models derived from considering the signal from a single resolution cell (the smallest unit of volume that the US transducer can isonify) [1]. Different backscatter models, such as the Rayleigh [2], Nakagami [3,4], Rician [4], K [1,3,5], generalized K [6], homodyned K [7,8], and the recently introduced generalized Nakagami distribution [9], have been proposed. The fit of the various distributions to the envelope under different scattering conditions has already been studied [3,10,11]. The parameters of these distributions have potential clinical importance in tissue characterization. For example, the shape parameter of the Nakagami distribution is monotonically related to the scatter density, and is therefore known as the effective density [9,12]. Additionally, the scale parameter of this distribution is the effective scatterer cross-section (size), and both parameters can be used to classify tissue [12]. The shape parameter of the K distribution is also a measure of effective density [1,3,5,13].

Parameter estimation for the Nakagami distribution is relatively easy and accurate [14–17], but is more complex for the K [17–21], generalized K [6], and homodyned K distributions [7]. The latter two models, specified by three parameters (the Nakagami and K distributions have two), encompass a wide range of scattering conditions. Furthermore, although many of these models are general, they are most appropriate for only specific density ranges [3]. Additionally, many of these parameters merely give an indication of backscatter characteristics, that are used subsequently as features in classification of tissue conditions [12].

In addition to density and scatterer amplitude, another important backscatter characteristic is scatterer spacing, or regularity. Regularity quantifies the randomness of the scatterers, and also indicates whether both randomly and regularly spaced scatterers are present. Regularity was previously studied with methods based on frequency and wavelet analysis [22–24].

In this paper, generalized entropies are introduced to characterize backscatter of the ultrasound echo envelope. Entropy is a functional of probability density, and is therefore related to distribution parameters. By using entropy, no specific distribution need be assumed. Entropy-based backscatter characterization also has an intuitive appeal, as entropy is a measure of dispersion, disorder, and complexity, all of which are affected by backscattering conditions [25]. In addition to the well-known Shannon–Boltzmann–Gibbs (or simply Shannon) entropy, many other entropy definitions have been introduced. This paper presents the efficacy of utilizing the Shannon, Renyi and Tsallis entropy measures for quantifying three scattering characteristics: density, regularity, and scatterer amplitudes. Analysis of regularity in the current study is limited to the presence of one type of scatterer spacing (i.e. clustered, random, and almost regular).

The paper is organized as follows. First, quantitative measures of scattering conditions are discussed. The basic theory of generalized entropy measures and entropy estimation follows. A neural

approach to estimating scattering measures from generalized entropies is presented in the next section. Experiments with a realistic US RF simulator, providing ground truth data, are described next, followed by results of using entropies to estimate the backscatter characteristics. Analysis, discussion, and concluding remarks end the paper.

2. Backscatter characteristics

In this study, three backscatter characteristics were considered: density (ρ), spacing (α), and scatterer amplitude (a). Density is a measure of the average number of scatterers in the resolution cell of the US transducer. Spacing (or placement) refers to the randomness or regularity of the distances between scatterers. The gamma distribution has been shown to accurately describe scatterer spacing [10]. By denoting the mean spacing between scatterers as $\bar{x}$, the spacing distribution is described by the gamma distribution, $\Gamma(\alpha, \beta)$, with shape parameter α and scale parameter $\beta = \bar{x}/\alpha$. For large α, the spacing is regular. For $\alpha = 1$ the scatterers are randomly spaced (diffuse). When $\alpha < 1$, the scatterers are clustered together [10].

The scattering amplitudes (henceforth termed amplitudes), or scattering cross-sections [4,26], were also characterized. The randomness of the amplitudes was modeled with the SNR of the scattering cross-section. In this case, the amplitudes are gamma-distributed with shape parameter a^2 and with unit scale parameter [4,26].

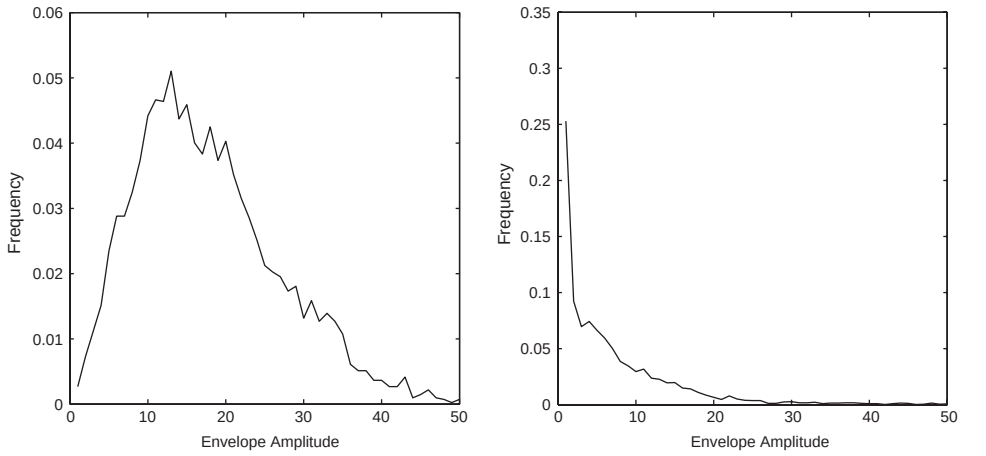
In this study, a realistic ultrasound RF simulator [4,26], described in Section 5.1, was used to generate echo envelopes with different scattering density, spacing and amplitude. The envelope histograms resemble specific probability density functions. Figs. 1a and b show the histograms for very high and very low density, spacing and amplitude. The former appears to have a Rayleigh form, while the latter has negative exponential characteristics. Fig. 1c shows a histogram for low density and high spacing and amplitude. The density function appears to have a general Rayleigh form, but with thinner tails, as does the density function in Fig. 1d. These observations about envelope distributions are consistent with those of other studies [1,3,4].

3. Generalized entropies and their estimation

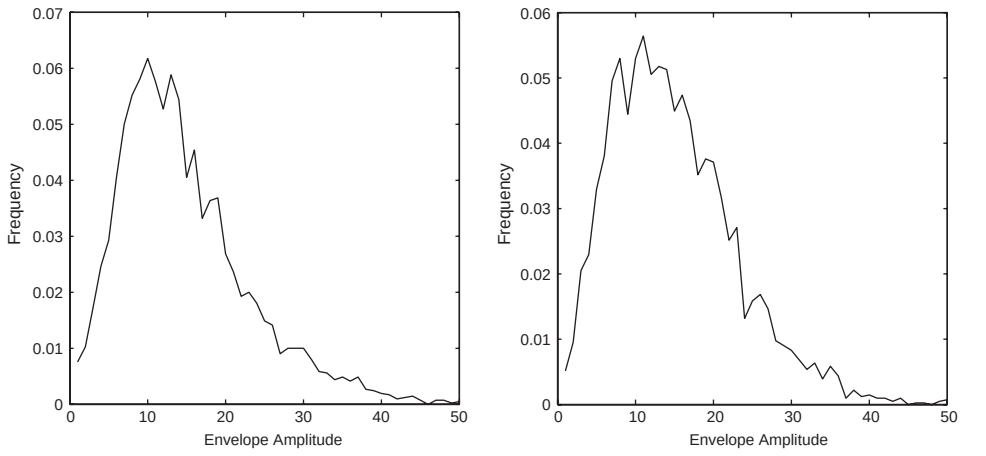
3.1. Background

Entropy, a fundamental concept in both physics and information theory, is a measure of disorder in a signal or data set. As it is a functional of the probability density of a data set, it can also represent different scattering conditions. In addition to the well-known Shannon entropy, several general families of entropy measures have been introduced [27]. Many of these measures have properties (e.g. smoothness, large dynamic range with respect to certain conditions) that make them desirable for specific applications [27–31].

The Shannon entropy is formally defined as an average value of logarithms of the probability density function. A generalization has been proposed to replace the average logarithms by an average of powers, leading to the Renyi and Tsallis entropy measures of order q [27–29]. Other entropy measures include trigonometric measures [32], polynomial β-entropy [33], and various other information



(a) High ρ (ρ = 50), high α (α = 8), and high a (a = 2) (b) Low ρ (ρ = 2.4), low α (α = 0.6), and low a (a = 0.9)



(c) High α (α = 8), high a (a = 2), and low ρ (ρ = 0.6) (d) High ρ (ρ = 50), high a (a = 2), and low α (α = 0.4)

Fig. 1. Envelope histograms for different scatterer densities (ρ), spacings (α) and amplitudes (a). (a) High density, spacing, and amplitude. (b) Low density, spacing, and amplitude. (c) Low density, high spacing and amplitude. (d) Low spacing, high density and amplitude.

measures [27]. However, in the current paper, the selection of entropies is limited to the Renyi and Tsallis measures, as they are widely used in many practical applications.

The Shannon entropy $H(X)$ of a continuous random variable X with density function $f(x)$ is defined as [27]:

$$H(X) = -\int_{-\infty}^{\infty} f(x) \ln f(x)\, \mathrm{d}x \equiv -E(\ln f(x)). \tag{1}$$

Renyi differential entropy of order $q > 0$, $q \neq 1$, is defined as follows [28]:

$$H_q^{\mathrm{R}}(X) = \frac{1}{1-q} \log \int_{-\infty}^{\infty} f^q(x)\, \mathrm{d}x. \tag{2}$$

It can be shown that

$$\lim_{q\to 1} H_q^{\mathrm{R}}(X) = H(X). \tag{3}$$

The Renyi entropy has been used in image registration [34,35] and feature evaluation [36]. In preliminary work, this measure has also been demonstrated to be promising in density estimation from US backscatter [37,38].

The Tsallis entropy measure is gaining importance in many applications, including optimization [31] and EEG signal analysis [30]. The Tsallis differential entropy of order $q > 0$, $q \neq 1$ is defined as [27,29]:

$$H_q^{\mathrm{T}}(X) = \frac{1}{q-1}\left(1 - \int_{-\infty}^{\infty} f^q(x)\,\mathrm{d}x\right). \tag{4}$$

Also, as in the case of Renyi entropy,

$$\lim_{q\to 1} H_q^{\mathrm{T}}(X) = H(X). \tag{5}$$

The Renyi and Tsallis measures are related by the following [39]:

$$H_q^{\mathrm{R}}(X) = \left(\frac{1}{1-q}\right)\log[1 + (1-q)H_q^{\mathrm{T}}(X)]. \tag{6}$$

Considering two independent random variables X and Y, $H(X,Y)$ and $H_q^{\mathrm{R}}(X,Y)$ are additive, whereas $H_q^{\mathrm{T}}(X,Y)$ is pseudo-additive [39]. However, it is precisely this property that makes it useful in the study statistical mechanics phenomena [29]. In the context of the current paper, however, only the properties of the generalized entropies with respect to backscatter characteristics are important.

3.2. Entropy estimation

Nonparametric kernel density entropy estimators has been frequently used in many applications and extensively studied [36,40]. The general form of the kernel density estimate of N random samples from a random variable X is given as [41]:

$$\hat{f}_N(x) = \frac{1}{Nh_N}\sum_{i=1}^{N} K\left(\frac{x - X_i}{h_N}\right), \tag{7}$$

where $h_N > 0$ is the window width, also called the smoothing parameter. Kernel functions which satisfy the conditions: $K(u) \geqslant 0$, $\int_{-\infty}^{\infty} K(u)\,\mathrm{d}u = 1$, $\int_{-\infty}^{\infty} uK(u)\,\mathrm{d}u = 0$, are usually symmetric probability density functions, e.g. Gaussian kernel $K(u) = (1/\sqrt{2\pi})\exp(-u^2/2)$.

By observing that the Shannon entropy is defined as: $H(X) = -E(\ln f(x))$, the entropy can be estimated by the quantity [41]:

$$H(\hat{f}_N) = -\frac{1}{N}\sum_{i=1}^{N} \ln \hat{f}_N(X_i). \tag{8}$$

The nonparametric Renyi entropy estimator, obtained in a similar manner to the Shannon entropy estimator, is given as [40]:

$$H_q^{\mathrm{R}}(\hat{f}_N) = \frac{1}{1-q} \ln \left\{ \frac{1}{N} \sum_{j=1}^{N} \hat{f}_N^{q-1}(X_j) \right\}. \tag{9}$$

For the Tsallis entropy, $H_q^{\mathrm{T}}(f)$, the nonparametric entropy estimate is expressed as [37]:

$$H_q^{\mathrm{T}}(\hat{f}_N) = \frac{1}{q-1} \left[1 - \frac{1}{N} \sum_{j=1}^{N} \hat{f}_N^{q-1}(X_j) \right]. \tag{10}$$

The choice of a smoothing parameter and kernel function for kernel-based entropy estimation methods affects the accuracy of the estimate. The mean integrated square error (MISE) [42] is widely used as a measure of the global accuracy of $\hat{f}$ as an estimator of f. The efficiency of any kernel can be compared with the Epanechnikov kernel, the optimal kernel minimizing the MISE, which has unit efficiency. It is given by [42]:

$$K_e(u) = \begin{cases} \frac{3}{4}(1 - \frac{1}{5}u^2)/\sqrt{5} & |u| < \sqrt{5}, \\ 0 & \text{otherwise.} \end{cases}$$

For example, the Gaussian kernel has an efficiency of 0.95 compared to the Epanechnikov kernel [42].

4. Function approximation with neural networks

Artificial neural networks (ANNs), especially multilayer perceptrons (MLPs) trained with the backpropagation (BP) learning algorithm [43], radial basis function neural networks (RBFNs) [43], general regression neural networks (GRNNs) [44], and Elman recurrent networks (ELMs) [45], are commonly used for function approximation and interpolation. They can approximate any linear or nonlinear function to an arbitrary degree of accuracy; i.e., they exhibit the universal approximation property [43]. Given a sufficient number of input data and network parameters, an approximation to an unknown mapping can be obtained. Neural networks also exhibit robust performance in the presence of noisy or inaccurate measurements, or data with outlier values [45]. However, in some cases, these networks generally converge very slowly for functions with strong nonlinearities.

Bayesian regularization in feedforward neural network training was proposed to improve network generalization [46,47]. Such regularization minimizes a weighted combination of squared errors and neuron weights by determining the combination weights.

In this study, entropies are considered as functions of backscatter characteristics. The neural network is trained to learn the mapping from the input space of the entropy values to the three-dimensional output space of the backscatter density ρ (the average number of scatterers in the resolution cell), spacing α (regularity of the distances between scatterers) and scatterer amplitude a (scattering cross-sections).

5. Methods

5.1. Data

An ultrasound RF simulator [4,26] was used to generate echo envelopes. The simulator realistically models the actual physical process in RF signal generation, and uses the density (ρ), spacing (α), and amplitudes (a) to describe the scattering process.

The parameters for the simulator are as follows: f_0 (center frequency) = 3.5 MHz, B (bandwidth) = 0.8 MHz, c (velocity of sound) = 1446 m/s, sampling window size = 3.7 cm. In the simulation, the RF backscattered signal consisting of 8 A lines (1-D RF signals) was sampled at 40 MHz. There were 2048 samples in each line. To ensure that samples were uncorrelated, every second sample in every second A line was used, for a total of 4096 ($= \frac{8}{2} \times \frac{2048}{2}$) data samples per envelope [4]. The received RF backscattered signal was demodulated at f_0, resulting in inphase and quadrature components X and Y, respectively. A 10th order 2 MHz Butterworth lowpass filter was applied to both components, and the echo envelope was computed as $\sqrt{X^2 + Y^2}$ [4].

Envelopes were generated for 20 densities ρ, log-spaced from 2 to 50 because, as will be demonstrated below, the entropy measures change much faster for low densities (< 10) than for higher values. Ten spacing parameters, α, were log-spaced from 0.3 (indicating a high degree of clustering) to 8 (indicating nearly regular spacing). The α were log-spaced because placement appears to be regular for $\alpha > 10$. Five SNR values for the amplitudes, a, were linearly spaced between 0.1 and 2. Therefore, each trial consisted of $20 \times 10 \times 5 = 1000$ training samples.

5.2. Entropy measures

Kernel-based estimation with the Epanechnikov kernel was used for the entropy estimates. Twenty trials were performed for each (ρ, α, a). For each envelope, the Renyi and Tsallis entropies of order $q \in \{0.1, 0.25, 0.5, 0.75, 2, 3, 4\}$, as well as the Shannon entropy, were computed.

Four generalized entropies were selected for approximating backscatter characteristics, based on the dynamic range and smoothness of the entropy estimates. These entropies, subsequently used as input to the neural networks, were: (1) Shannon entropy (H), (2, 3) Renyi entropy of orders $q = 0.25$ ($H^{\mathrm{R}}_{0.25}$), and $q = 0.75$ ($H^{\mathrm{R}}_{0.75}$), and (4) Tsallis entropy of order $q = 0.25$ ($H^{\mathrm{T}}_{0.25}$). The Shannon entropy estimate is shown in Fig. 2. Fig. 3 shows Renyi and Tsallis entropies for $q = 0.25$ and 0.75. In these figures, the SNR of the amplitude distribution was set to $a = 1.5$. Due to numerical inaccuracies in estimating some orders of Tsallis entropy (Eq. (10)), the $H^{\mathrm{T}}_{0.25}$ function is not as smooth as the other entropies in Fig. 3. However, one of the main advantages of neural networks is robust response in the presence of noisy and inaccurate measurements, and therefore, because of its high dynamic range, this entropy was selected as network input.

As seen in Fig. 2, the Shannon entropy does not markedly increase with increasing ρ. However, the low order Renyi entropy and, especially, Tsallis entropy have much higher dynamic range than the Shannon entropy. High order Renyi and Tsallis entropies are very large and negative for low densities and clustered scatterers, as demonstrated in Fig. 4. These properties reduce their usefulness over the full range of scattering conditions.

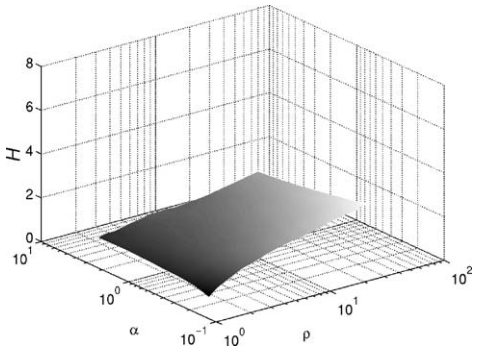


Fig. 2. Shannon (H) entropy estimate from envelopes. The backscatter coefficient is $a = 1.5$.

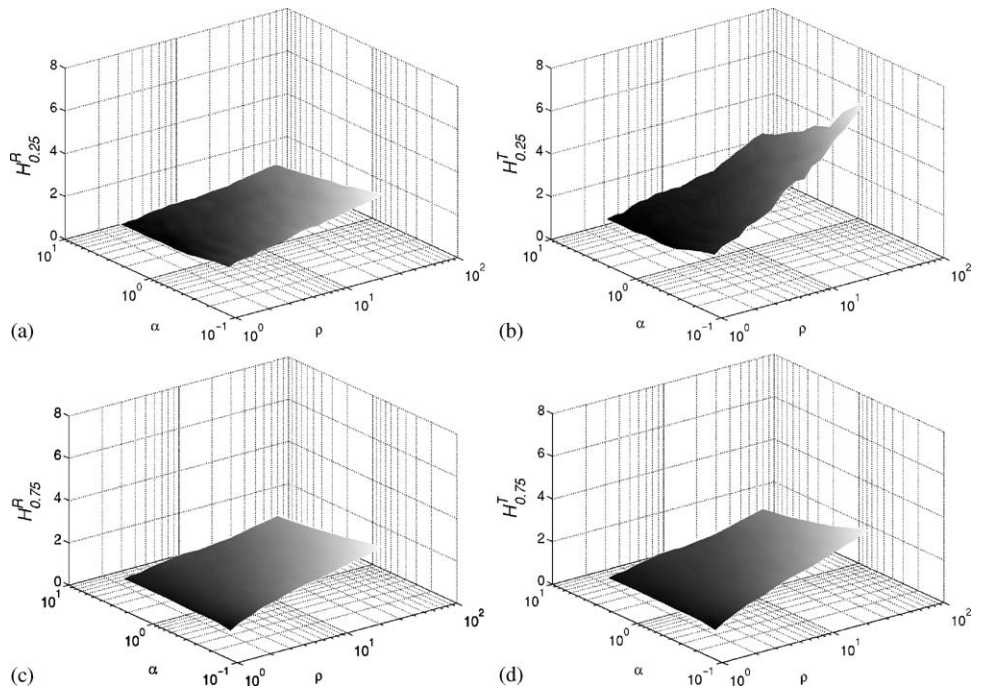


Fig. 3. Tsallis ($H^{\mathrm{T}}_{0.25}$, $H^{\mathrm{T}}_{0.75}$) and Renyi ($H^{\mathrm{R}}_{0.25}$, $H^{\mathrm{R}}_{0.75}$) entropy estimates from envelopes. The backscatter coefficient is $a = 1.5$.

5.3. Neural estimation of scattering characteristics

Neural network function approximators were used to estimate density, spacing, and SNR of the amplitude distribution. The networks were trained to learn the mapping from the 4-dimensional input space of 4 entropy measures to the three-dimensional output space (ρ, α, a). Several network architectures were employed: MLPs trained with the Bayesian regularization backpropagation learning algorithm, GRNNs and ELMs (radial basis function neural networks, which are special

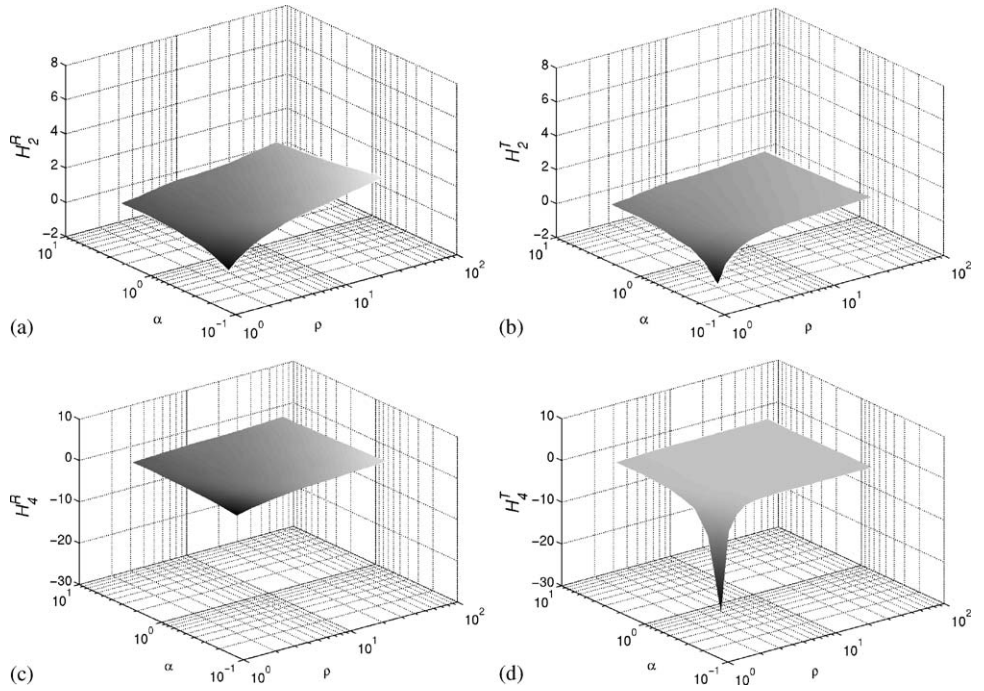


Fig. 4. Tsallis (H_2^{T}, H_4^{T}) and Renyi (H_2^{R}, H_4^{R}) entropy estimates from envelopes. The backscatter coefficient is $a = 1.5$.

cases of GRNNs, may also be used for this purpose). Regularization was implemented with the Gauss–Newton approximation to Bayesian regularization using Levenberg–Marquardt optimization [47].

In many cases, RF envelopes can be collected in a small amount of time window. As the envelope results from a stochastic process, statistics and other features obtained from averaged envelopes may be distorted. However, the features themselves can be averaged to obtain a smoother estimate. This was the approach taken in these experiments, and the envelope entropy estimates were averaged over the 10 trials. The mean entropies were presented as inputs to the neural network (1000 training samples) together with three targets. The training set was divided into two sets (10 trials for each set) so that cross-validation could be performed. The error metric was the root-mean square error (E_{RMS}) between the estimate from the neural network and the ground truth ρ, α and a values.

To test the network, 1000 simulated envelopes with randomly spaced densities in the range $[2, 50]$, spacings in $[0.3, 8]$, and amplitudes in $[0.1, 2]$ were generated. These data were generated to test the approximation and interpolation abilities of the networks. The generalized entropy measures from the resulting envelopes were then estimated and presented to the networks.

Although neural networks were used exclusively for backscatter characterization in this study, it should be noted that there are many other data modeling methods [48] that could also be used to estimate backscatter characteristics from generalized entropies.

Table 1
E_{RMS} for testing different ANN architectures with four entropy inputs

	Cross-validation				Testing			
Architecture	ρ	α	a	Total	ρ	α	a	Total
MLP (5)	0.1907	0.2148	0.1228	0.1804	0.2043	0.1601	0.1246	0.1662
MLP (10)	0.1780	0.1788	0.1230	0.1620	0.2128	0.1618	0.1213	0.1695
MLP (15)	0.1953	0.2291	0.1265	0.1885	0.2133	0.1642	0.1211	0.1704
ELM (5)	0.1780	0.1788	0.1230	0.1620	0.2055	0.1613	0.1227	0.1667
ELM (10)	0.1755	0.1847	0.1214	0.1630	0.2119	0.1619	0.1215	0.1692
ELM (15)	0.1864	0.2292	0.1210	0.1843	0.2098	0.1618	0.1218	0.1684
GRNN ($\sigma = 0.1$)	0.1613	0.1854	0.1177	0.1573	0.2309	0.1793	0.1270	0.1840
GRNN ($\sigma = 0.25$)	0.1763	0.2013	0.1208	0.1695	0.2446	0.1924	0.1273	0.1907
GRNN ($\sigma = 1.0$)	0.2001	0.2151	0.1274	0.1849	0.2279	0.2012	0.1278	0.1904

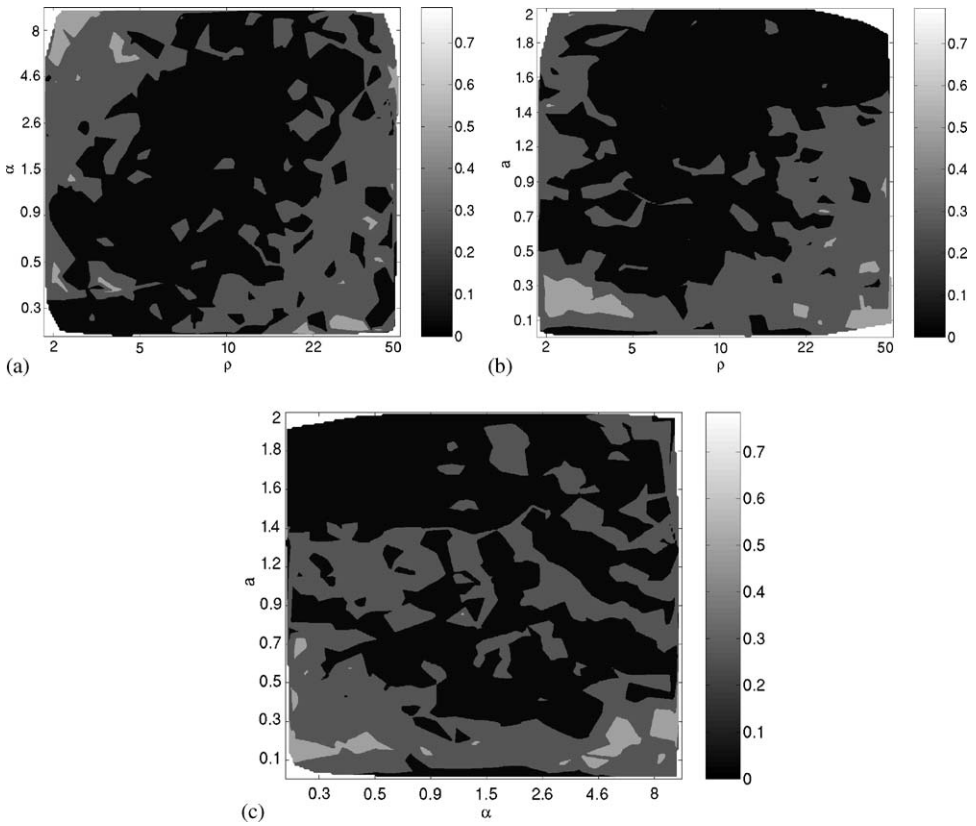


Fig. 5. The error (Euclidean norm) between (ρ, α, a) and the estimates $(\hat{\rho}, \hat{\alpha}, \hat{a})$. The MLP architecture with five hidden units was used for ANN training. Black indicates a very low error.

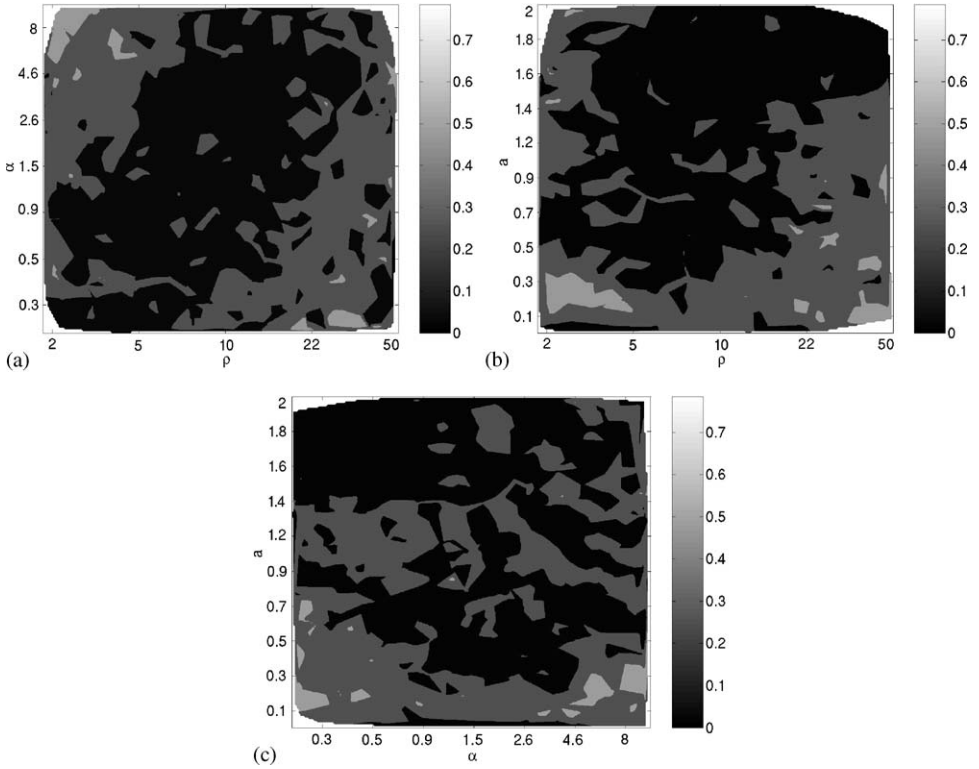


Fig. 6. The error (Euclidean norm) between (ρ, α, a) and the estimates $(\hat{\rho}, \hat{\alpha}, \hat{a})$. The ELM architecture with five hidden units was used for ANN training. Black indicates a very low error.

6. Results

The root-mean square error (E_{RMS}) results for cross-validation and testing of different neural network architectures are summarized in Table 1. As the densities and spacing training values were log-spaced, the three output values ρ, α, and a were linearly spaced: $\rho \in [\log(2), \log(50)]$, $\alpha \in [\log(0.3), \log(8)]$, $a \in [0.1, 2]$. These output values were scaled to the range $[0, 1]$. Different numbers of hidden units (5, 10 and 15) was used for the multilayer perceptron trained with Bayesian regularization BP learning algorithm and the Elman recurrent networks. All networks were trained for 200 epochs. The GRNNs were trained with spread $\sigma = 0.1$, 0.25 and 1. These networks are trained in a single pass [44]. The E_{RMS} was computed for each triple (ρ, α, a) for cross-validation and testing. The total E_{RMS} was also computed. It indicates the overall performance of a neural network.

All ANNs were most accurate for estimating the amplitudes. As seen in Table 1, the MLP trained with Bayesian regularization BP with 5 hidden units had the smallest total E_{RMS} for testing. The Elman recurrent network with 5 hidden units and GRNN with $\sigma = 0.1$ also performed well resulting in a slightly higher total E_{RMS} than the MLP.

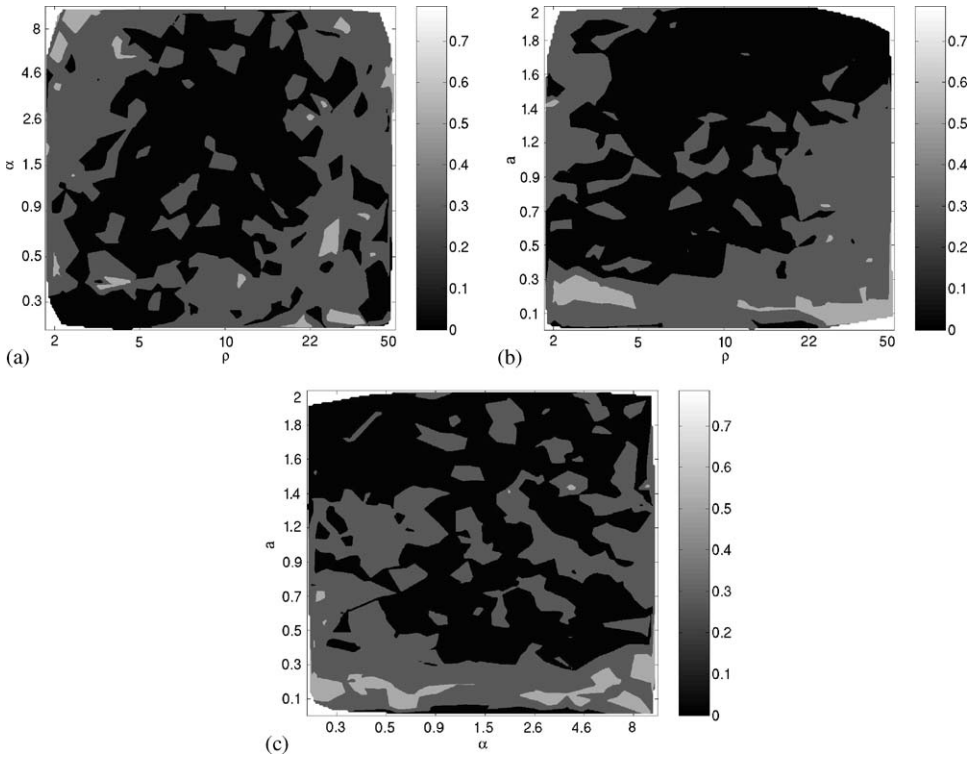


Fig. 7. The error (Euclidean norm) between (ρ, α, a) and the estimates $(\hat{\rho}, \hat{\alpha}, \hat{a})$. The GRNN architecture with the spread $\sigma = 0.1$ was used for ANN training. Black indicates a very low error.

To analyze the effect of different scattering conditions on the accuracy of the testing estimates, the Euclidean norm was computed between the ground truth values (ρ, α, a) and the estimates $(\hat{\rho}, \hat{\alpha}, \hat{a})$, and displayed as contour maps. Figs. 5–7 show the Euclidean norms (errors) for the best three architectures: MLP (5 hidden units), ELM (5 hidden units) and GRNN ($\sigma = 0.1$). The best results are observed for MLP and ELM networks. Because the testing parameters were uniformly distributed, the parameter and error values were interpolated for display purposes. In these figures, the brightness is proportional to error, with black indicating very low error.

The performance of the proposed method for backscatter characteristics estimation based on entropy was also compared with another estimation technique. The Nakagami *m* parameter, the effective scatterer density [4], was estimated from the envelopes with the Tolparev-Polyakov method, which has been shown to be very accurate [14]. The estimate was then used as a fourth input to the neural network along with three entropy values. The MLP architecture with 5, 10, and 15 hidden units was employed. The cross-validation was performed on a training set (1000 samples) divided into two sets (five trials each) and the network was tested on ten trials.

As seen in Table 2, the MLP with 10 hidden units was the best performer. These results indicate that the Nakagami parameter can be also used along with entropy estimates for backscatter characteristics estimation. However, it is most effective in estimating density, and, when used alone, cannot indicate spacing and/or amplitude.

Table 2
E_{RMS} for the MLP architecture testing with three entropy inputs and one Nakagami m parameter input

	Cross-validation				Testing			
Architecture	ρ	α	a	Total	ρ	α	a	Total
MLP (5)	0.1895	0.1932	0.1295	0.1722	0.1753	0.1933	0.1284	0.1679
MLP (10)	0.1769	0.1891	0.1267	0.1667	0.1674	0.1892	0.1278	0.1674
MLP (15)	0.1762	0.2056	0.1287	0.1733	0.1863	0.1992	0.1305	0.1746

7. Discussion

Entropy is a measure of disorder and complexity. It is also a functional of the probability density, and can therefore convey more information about a process than first-order statistics. In addition to the well-known Shannon definition, many other entropies exist, each having characteristics that make them useful for specific applications. The two generalized measures used in this study, the Renyi and Tsallis entropies, generalize the Shannon measure, and small orders of these entropies ($q = 0.25$ and 0.75) have a high dynamic range with respect to density and regularity. High order Tsallis and Renyi entropies ($q = 2.0$ and 4.0) are smoother than the low-order measures, but they have a lower dynamic range with respect to ρ, α, and a. Additionally, as seen in Fig. 4, H_2^T and H_4^T have a very sharp "tail" for low density, highly clustered scatterers, but are otherwise generally flat. Therefore, the high-order entropies are not useful for backscatter characterization.

Although entropy is nonlinear with respect to ρ, α, and a, the functions are smooth enough (see Fig. 3) so that accurate estimates can be obtained with relatively simple feedforward neural networks having only five hidden neurons. The results on testing data (with different density, spacing, and amplitude values than were used in training) also showed that the estimates were free of very large interpolation errors. Thus, the neural networks have been shown to be accurate estimators for the three backscatter characteristics. Although neural networks are universal approximators that exhibit robust behavior in the presence of inaccurate measurements, many non-neural numerical approximation techniques exist [48], and could also potentionally be used for backscatter characterization.

As seen in Figs. 5–7, most of the errors occurred for high densities ($\rho > 20$), for regularly-spaced scatterers ($\alpha > 5.0$), and for low amplitude SNR values ($a < 0.5$). The first two situations (density and spacing) can be explained by the fact that the entropy values for all measures increase very slowly as ρ increases above 20. They also change very slowly as the scatterer spacing becomes more regular. In the case of highly regular spacing, and the presence of both random and regularly spaced scatterers, frequency analysis or wavelet-based approaches may be appropriate [22–24].

A potential limitation of the proposed method is that RF characteristics may differ among US devices, and, consequently, entropy values may also differ. As a result, neural networks must be trained for a specific device, with entropy values computed from phantom or other ground-truth data.

Also, as stated earlier, it was assumed that all scatterers contributing to the RF envelope have one type of spacing (clustered, random, or some degree of regularity). Both types of spacing may

be present, leading to Rician statistics. Using generalized entropies to model these situations is the subject of future investigations.

8. Conclusions

In this paper, a new approach to ultrasonic backscatter characterization based on generalized entropies was presented. Using entropy, specific scattering distribution models need not be assumed. However, as shown in the results section, statistical parameters may be used in conjuction with generalized entropies. Because three characteristics are to be estimated, the Shannon entropy, whose values are not unique with respect to these properties, is not sufficient in itself. Low order generalized entropies, especially the Tsallis entropy measure, have a high dynamic range with respect to density, spacing and amplitude, and are therefore suitable for characterizing these properties.

Neural network estimators were used for backscatter characterization. Experimental results on ground truth simulated envelopes show that accurate estimates can be obtained with three orders of generalized entropy, in addition to the Shannon entropy. Accuracy of the estimates does not appear to be dependent on the specific network architecture. Furthermore, entropy-based characterization can be performed with other, non-neural, estimation methods.

Although further experiments on clinical RF envelopes are required, the controlled experiments presented in this paper strongly suggest that generalized entropies may be employed, along with other techniques, to accurately characterize ultrasonic backscatter.

References

[1] R.C. Molthen, P.M. Shankar, J.M. Reid, F. Forsberg, E.J. Halpern, Comparisons of the Rayleigh and K-distribution models using in vivo breast and liver tissue, Ultrasound Med. Biol. 24 (1) (1998) 93–100.

[2] J.W. Goodman, Statistical properties of laser speckle patterns, in: J.C. Dainty (Ed.), Laser Speckle and Related Phenomena, Vol. 9 of Topics in Applied Physics, 2nd Edition, Springer, Berlin, 1984, pp. 9–76.

[3] L. Clifford, P. Fitzgerald, D. James, Non-Rayleigh first-order statistics of ultrasonic backscatter from normal myocardium, Ultrasound Med. Biol. 19 (6) (1993) 487–495.

[4] P.M. Shankar, A general statistical model for ultrasonic backscattering from tissues, IEEE Trans. Ultrason. Ferr. Freq. Con. 47 (3) (2000) 727–736.

[5] P.M. Shankar, A model for ultrasonic scattering from tissues based on the K distribution, Phys. Med. Biol. 26 (9) (2000) 1633–1649.

[6] D.L. Liu, R.C. Waag, Harmonic amplitude distribution in a wideband ultrasonic wavefront after propagation through human abdominal wall and breast specimens, J. Acoust. Soc. Am. 2 (1997) 1172–1183.

[7] V. Dutt, J.F. Greenleaf, Ultrasound echo envelope analysis using a homodyned K distribution signal model, Ultrasonic Imag. 16 (1994) 265–287.

[8] X. Hao, C.J. Bruce, C. Pislaru, J.F. Greenleaf, Classification of normal and infarcted myocardium based on statistical analysis of high frequency intracardiac ultrasound RF signal, Proc. SPIE 4687 (2002) 139–146.

[9] P.M. Shankar, Ultrasonic tissue characterization using a generalized Nakagami model, IEEE Trans. Ultrasonic Ferr. Freq. Con. 48 (6) (2001) 1716–1720.

[10] R.M. Cramblitt, K.J. Parker, Generation of non-Rayleigh speckle distributions using marked regularity models, IEEE Trans. Ultrasonic Ferr. Freq. Con. 46 (1999) 867–874.

[11] M.P. Wachowiak, R. Smolíková, G.D. Tourassi, A.S. Elmaghraby, General ultrasound speckle models in determining scatterer density, Proc. SPIE 4687 (2002) 285–295.

[12] P.M. Shankar, V.A. Dumane, J.M. Reid, V. Genis, F. Forsberg, C.W. Piccoli, B.B. Goldberg, Classification of ultrasonic B-mode images of breast masses using Nakagami distribution, IEEE Trans. Ultrasonic Ferr. Freq. Con. 48 (2) (2001) 569–580.
[13] L. Weng, J.M. Reid, P.M. Shankar, K. Soetanto, Ultrasound speckle analysis based on the *K* distribution, J. Acoust. Soc. Am. 89 (6) (1991) 2992–2995.
[14] A. Abdi, M. Kaveh, Performance comparison of three different estimators for the Nakagami *m* parameter using Monte Carlo simulation, IEEE Commun. Lett. 4 (4) (2000) 119–121.
[15] J. Cheng, N.C. Beaulieu, Maximum-likelihood based estimation of the Nakagami *m* parameter, IEEE Commun. Lett. 5 (2001) 101–103.
[16] M.P. Wachowiak, R. Smolíková, M.G. Milanova, A.S. Elmaghraby, Statistical and neural approaches for estimating parameters of a speckle model based on the Nakagami distribution, Lecture Notes in Artificial Int. 2123, Springer, Berlin, 2001, pp. 196–205.
[17] R. Smolíková, M.P. Wachowiak, A.S. Elmaghraby, J.M. Zurada, Segmentation of ultrasound images with speckle modeling, Proc. of IEEE EMBS/EUROSIP Conf. (Biosignal 2002), Brno, Czech Republic, 2002, pp. 316–319.
[18] V. Dutt, J.F. Greenleaf, Speckle analysis using signal to noise ratios based on fractional order moments, Ultrasonic Imag. 17 (1995) 251–268.
[19] D.R. Iskander, A.M. Zoubir, B. Boashash, A method for estimating the parameters of the *K* distribution, IEEE Trans. Signal Proc. 47 (4) (1999) 1147–1151.
[20] D.R. Iskander, A.M. Zoubir, Estimating the parameters of *K*-distribution using higher order and fractional moments, IEEE Trans. Aerosp. Electron. Syst. 35 (4) (1999) 1453–1457.
[21] M.P. Wachowiak, R. Smolíková, J.M. Zurada, A.S. Elmaghraby, Estimation of *K* distribution parameters using neural networks, IEEE Trans. Biomed. Eng. 49 (6) (2002) 617–620.
[22] G. Georgiou, F.S. Cohen, Tissue characterization using the continuous wavelet transform, Part I: decomposition method, IEEE Trans. Ultrasonic Ferr. Freq. Con. 48 (2001) 355–363.
[23] G. Georgiou, F.S. Cohen, C.W. Piccoli, F. Forsberg, B.B. Goldberg, Tissue characterization using the continuous wavelet transform, Part II: application on breast RF data, IEEE Trans. Ultrasonic Ferr. Freq. Con. 48 (2001) 364–373.
[24] X. Tang, U.R. Abeyratne, Wavelet transforms in estimating scatterer spacing from ultrasound echoes, Ultrasonics 38 (2002) 688–692.
[25] Y. Zimmer, S. Akselrod, R. Tepper, The distribution of the local entropy in ultrasound images, Ultrasound Med. Biol. 22 (4) (1996) 431–439.
[26] V.M. Narayanan, P.M. Shankar, J.M. Reid, Non-Rayleigh statistics of ultrasonic backscattered signal, IEEE Trans. Ultrasonic Ferr. Freq. Con. 41 (6) (1994) 845–851.
[27] J.N. Kapur, H.K. Kesavan, Entropy Optimization Principles with Applications, Academic Press, New York, 1992.
[28] A. Renyi, Probability Theory, North-Holland, Amsterdam, 1970.
[29] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, J. Stat. Phys. 52 (1/2) (1988) 479–487.
[30] A. Capurro, L. Diambra, D. Lorenzo, O. Macadar, M.T. Martin, C. Mostaccio, A. Plastino, E. Rofman, M.E. Torres, J. Velluti, Tsallis entropy and cortical dynamics: the analysis of EEG signals, Physica A 257 (1998) 149–155.
[31] C. Tsallis, D. Stariolo, Generalized simulated annealing, Physica A 233 (1996) 395–406.
[32] A.P. Sant'anna, I.J. Taneja, Trigonometric entropies, Jensen difference divergence measures, and error bounds, Inf. Sci. 35 (1985) 145–156.
[33] B.D. Sharma, R. Autar, An inversion theorem and generalized entropies for continuous distributions, SIAM J. Appl. Math. 25 (1973) 125–132.
[34] B. Ma, A. Hero, J. Gorman, Image registration with minimimal spanning tree algorithm, Proc. of IEEE Int. Conf. on Imag. Proc. (ICIP), Vancouver, Canada, 2000, pp. 10–13.
[35] M.P. Wachowiak, R. Smolíková, T.M. Peters, Multiresolution biomedical image registration using generalised information measures, MICCAI 2003, Montréal, Canada, November 16–18, 2003.
[36] D. Xu, J.C. Principe, Feature evaluation using quadratic mutual information, In Proc. of INNS/IEEE Conf. on Neural Networks (IJCNN 2001), Washington, DC, 2001, pp. 459–463.
[37] R. Smolíková, M.P. Wachowiak, G.D. Tourassi, J.M. Zurada, Neural estimation of scatterer density in ultrasound, Proc. of INNS/IEEE Conf. on Neural Networks (IJCNN 2002), Honolulu, HI, 2002, pp. 1696–1701.

[38] R. Smolíková, M.P. Wachowiak, G.D. Tourassi, A.S. Elmaghraby, J.M. Zurada, Characterization of ultrasonic backscatter based on generalized entropy, In Proc. of EMBS/BMES Conference, Houston, TX, 2002, pp. 953–954.
[39] D. Holste, I. Grosse, H. Herzel, Bayes' estimators of generalized entropies, J. Phys. A 31 (1998) 2551–2556.
[40] D. Erdogmus, J.C. Principe, Entropy minimization algorithm for multilayer perceptrons, In Proc. of INNS/IEEE Conf. on Neural Networks (IJCNN 2001), Washington, DC, 2001, pp. 3003–3008.
[41] E.J. Dudewicz, E.C. Van der Meulen, The empiric entropy, a new approach to nonparametric entropy estimation. In M.L. Puri, J. Vilaplana, W. Westz (Eds.), New Perspectives in Theoretical and Applied Statistics, Wiley, New York, 1987, pp. 207–227.
[42] B.W. Silverman, Density Estimation for Statistics and Data Analysis, Chapman & Hall, London, 1986.
[43] R. Rojas, Neural Networks: A Systematic Introduction, Springer, Berlin, 1996.
[44] P.D. Wasserman, Advanced Methods in Neural Computing, Van Nostrand Reinhold, New York, 1993.
[45] C.M. Mehrotra, S. Ranka, Elements of Artificial Neural Networks, MIT Press, Cambridge, 1997.
[46] D.J.C. MacKay, Bayesian interpolation, Neural Comput. 44 (1992) 415–447.
[47] F.D. Foresse, M.T. Hagan, Gauss–Newton approximation to Bayesian learning, Proceedings of the INNS-IEEE Joint Conference on Neural Networks (IJCNN 1997), 1997.
[48] W. Duch, G.H.F. Diercksen, Neural networks as tools to solve problems in physics and chemistry, Comput. Phys. Commun. 82 (1994) 91–103.

Renata Smolíková received M.S. in Mathematics from Palacky University, Olomouc, Czech Republic in 1989 and Ph.D. in Mathematics from the same university in 1999. From 1992 to 1998 she was an assistant professor in the Department of Mathematics, University of Ostrava, Czech Republic and from 1996 to 1998, a research associate in the Institute for Research and Applications of Fuzzy Modeling, University of Ostrava. In 2002 Dr. Smolíková finished her second Ph.D. in Computer Science and Engineering and is currently working as a postdoctoral Fellow at Robarts Research Institute, London, ON, Canada. She is a member of SIAM, the IEEE, and the IEEE Engineering in Medicine and Biology Society. Her research interests include ultrasound backscatter analysis, medical image registration, and neural networks.

Mark P. Wachowiak earned his Ph.D. in Computer Science and Engineering from the University of Louisville in 2002. He completed the M.S. in Computer Science at the same institution. Dr. Wachowiak is a member of the IEEE, IEEE Engineering in Medicine and Biology Society, and SIAM. His research interests are in the areas of biomedical engineering and computation, including soft tissue modeling, biomedical imaging and visualization, image-guided therapy and surgery planning, signal analysis, high-performance computing, and neuroscience. He is currently a postdoctoral fellow in the Imaging Laboratories of Robarts Research Institute in London, ON, Canada.

Jacek M. Zurada is the S.T. Fife Alumni Professor of Electrical and Computer Engineering at the University of Louisville. He is the co-editor of the 2000 MIT Press volume *Knowledge-Based Neurocomputing*, and the author of the 1992 PWS text *Introduction to Artificial Neural Networks*, contributor to the 1994 and 1995 Ablex volumes *Progress in Neural Networks*, and co-editor of the 1994 IEEE Press volume *Computational Intelligence*: *Imitating Life*. He is the author or co-author of more than 180 journal and conference papers in the area of neural networks, and analog and digital VLSI circuits. Since 1998 Dr. Zurada has been the Editor-in-Chief of IEEE Transactions on Neural Networks. Dr. Zurada is also an Associate Editor of Neurocomputing, and was an Associate Editor of IEEE Transactions on Circuits and Systems, Part I and Part II. He is an IEEE Distinguished Speaker for the Circuits and Systems Society and the Neural Networks Society. Dr. Zurada serves as the President-Elect in 2003, and as the President of IEEE Neural Networks Society in 2004-05.