Einsteinian Neural Network for Spectrum Estimation

LEONID I. PERLOVSKY, CHARLES P. PLUM, PETER R. FRANCHI, ELIHU J. TICHOVOLSKY, DAVID CHOI AND BERTUS WEIJERS

Nichols Research Corporation and US Air Force Rome Laboratories, Hanscom Field AFB

(Received 3 March 1997; accepted 8 July 1997)

Abstract—A model-based neural network is developed for spectrum estimation. Its architecture and learning mechanism are founded on the Einsteinian interpretation of the spectrum as a probability distribution of photons. By considering a spectrum as an ensemble of photons, we derive the neural learning mechanism from the basic physical principle of entropy maximization of a canonical ensemble. This neural network is applied to characterizing a recently observed phenomenon known as equatorial ionospheric clutter that significantly affects operations of over-the-horizon (OTH) radars and communication links using high frequency radiowaves propagating through the ionosphere. We utilize a specific parameterization of the internal spectral model, which is derived from the physical principles of the propagation of electromagnetic waves through a turbulent ionosphere. A set of parameters characterizing equatorial ionospheric clutter is estimated. The developed technique may have a broad applicability in scientific data analysis. © 1997 Elsevier Science Ltd. All rights reserved.

Keywords—Model-based, Neural networks, Spectrum estimation, Entropy, Ionosphere, Radar, Propagation, Clutter.

1. INTRODUCTION

Model-based neural networks utilize internal models of signals, processes, or objects in order to combine adaptive learning with a priori knowledge. One of the first model-based neural network was Widrow’s Adaline (Widrow, 1959). Several neural networks based on more complicated models have been recently described for applications in classification, signal processing, tracking, and control (Perlovsky, 1987, 1994; Specht, 1990; Streit & Luginbuhl, 1990; Perlovsky & Jaskolski, 1994). In this paper, a neural network utilizing compositional, flexible model for spectrum estimation is developed and applied to characterization of ionospheric propagation effects in Doppler spectra observed by over-the-horizon (OTH) radars. The developed technique has a broad applicability for analyses of spectral data, including time–frequency spectra used in speech recognition as well as for other applications, where the model parameters are difficult to estimate, because of multiple interfering sources.

We present a new approach to model-based spectrum estimation founded on Einstein’s interpretation of the spectrum as a probability distribution function (pdf) of photon frequency. This leads to a “Einsteinian” likelihood function which is different from what usually is encountered in statistical estimation. We model the spectrum as a superposition of signals from several sources using physically based models for each signal source. And, we estimate parameters of this model by using the physical principle of the maximum entropy (ME) of a photon ensemble.

Section 2 discusses Einstein’s interpretation of photon spectra, introduces a physical model for signal spectra compatible with Einstein’s ideas, and presents the ME estimators for the model parameters. Section 3 briefly describes an architecture of Einsteinian neural network. In Section 4 the developed estimation technique is applied to radar data. The results are discussed in Section 5.

2. PHYSICAL MODEL FOR SPECTRUM ESTIMATION

Einstein interpreted the electromagnetic spectrum as a probability distribution function (pdf) of the photon energy (Einstein & Hopf, 1910). A similar interpretation is valid for phonons of acoustic spectra (speech, seismic
signals, etc.) and for any signal field obeying Bose-Einstein’s statistics (bosons). Statistical estimation theory usually considers a pdf of the given set of data as a function of model parameters, which is called likelihood. A new spectrum modeling and estimation technique, which relates statistical estimation of a spectrum to the equilibration of a physical ensemble of photons, is introduced in this section. It is inspired by the Einsteinian interpretation of the spectrum, which considers frequency (rather than spectral values) as a random variable. We define a spectrum model in terms of the number of physical photon states as follows. A spectrum \( S(\omega) \) is measured in units of energy and the energy of the individual photon is \( \hbar \); thus, the number of measured photons is

\[
N_\omega = S(\omega)/\hbar \omega, \quad (1)
\]

where \( \omega \) is frequency in radians per second. A pdf is proportional to an expected number of observations. For a single photon, this is proportional to the density of photon states as a function of frequency. Therefore according to Einstein’s interpretation, a spectrum model \( F(\omega) \) is interpreted as proportional to a number of physical states for a single photon at each frequency,

\[
NF_\omega = \text{const.} \cdot \hbar \omega \cdot NF_\omega. \quad (2)
\]

A computation of entropy of the physical ensemble, \( E \), can be found in statistical physics textbooks, for example (Sakurai, 1985); it results in

\[
E = \sum_\omega N_\omega \ln[NF_\omega/N_\omega] = \sum_\omega [S(\omega)/\hbar \omega] \ln [F(\omega)/(\text{const.} \cdot S(\omega))]. \quad (3)
\]

The physical equilibrium equations for a photon ensemble are obtained by maximizing entropy subject to the constraints of the conservation laws. Since we consider estimation of model parameters, given the data, the energy \( e \) and the number of photons \( N \) are constant during the estimation process (in physics it is called a canonical ensemble)

\[
\epsilon = \sum_\omega S(\omega) = N \sum_\omega F(\omega); \quad N = \sum_\omega S(\omega)/\hbar \omega = N \sum_\omega F(\omega)/\hbar \omega. \quad (4)
\]

A specific shape of the parametric model for \( F(\omega) \) is determined based on the physics of the electromagnetic wave propagation and scattering. Physical considerations discussed in Section 4 suggest that a spectrum should be modeled as a sum of contributions from four sources (which we will also call modes), one of which has a uniform distribution in photon energy, and the other three modes are Gaussian

\[
F(\omega) = \sum_m F(\omega|m), m = 1, \ldots 4. \quad (5)
\]

\[
F(\omega|m) = \hbar \omega A_m/(\omega_{\text{max}} - \omega_{\text{min}}), \quad F(\omega|m) = \hbar \omega A_m G(\omega|m), m = 2, 3, 4. \quad (6)
\]

\[
G(\omega|m) = (2\pi)^{-1/2}(\sigma_m)^{-1} \exp\{0.5(\omega - \omega_m)^2/\sigma_m^2\}. \quad (7)
\]

Here each Gaussian mode is specified by three parameters, the modal amplitude, \( A_m \), mean frequency, \( \omega_m \), and frequency standard deviation \( \sigma_m \). A multiplicative term \( \hbar \omega \) in eqn (6) is introduced so that \( [A_m G(\omega|m)] \) is measured in units of the number of photons, and \( G(\omega|m) \) can be interpreted as a pdf of a photon source \( m \). The ME estimation equations for the model parameters are derived by maximizing the entropy, \( E \), eqn (3), under constraints (eqn 4). By using the method of Lagrange multipliers, the following ME equations are derived.

\[
A_m = N_m/N, \quad N_m = \sum_\omega P(m|\omega) [S(\omega)/\hbar \omega], \quad E = \sum_\omega [S(\omega)/\hbar \omega] \ln [F(\omega)/(\text{const.} \cdot S(\omega))]. \quad (8)
\]

\[
\bar{\omega}_m = \sum_\omega P(m|\omega) [S(\omega)/\hbar \omega] \omega/N_m, \quad (9)
\]

\[
\sigma_m^2 = \sum_\omega P(m|\omega) [S(\omega)/\hbar \omega] (\omega - \bar{\omega}_m)^2/N_m. \quad (10)
\]

\[
P(m|\omega) = A_m F(\omega|m) / \left( \sum_m A_m F(\omega|m') \right), \quad \sum_m P(m|\omega) = 1. \quad (11)
\]

The term \( P(m|\omega) \) in eqn (11) has a meaning of a posteriori Bayes probability that a photon at frequency \( \omega \) has originated from the source (or mode) \( m \). Thus, \( N_m \) is the number of photons from the source \( m \), and \( N \) is the total number of photons. Equations (8) and (11) above are applicable to all four modes, and eqns (9) and (10) are applicable to modes \( m = 2, 3, 4 \).

3. ARCHITECTURE OF EINSTEINIAN NEURAL NETWORK

Equations (8)–(11) define an iterative dynamical system similar to the MLANS neural network considered previously (Perlovsky & McManus, 1991; Perlovsky, 1994). The parallel structure of these equations is similar to the original MLANS equations and we have used the MLANS architecture to implement this new neural network that we call MEANS for Maximum Entropy Adaptive Neural System or Model-based Einsteinian ANS. The MEANS architecture is illustrated in Figure 1a,b. It consists of two subsystems, the modeling subsystem, Figure 1a, that implements eqns (8)–(10) for estimating model parameters and the association subsystem, Figure 1b, that implements eqn (11) for associating spectrum values with various signal sources. The modeling subsystem input nodes contain Doppler spectrum values,
FIGURE 1. MEANS architecture (a,b) and spectra estimation results (c). A modeling subsystem (a) estimates model parameters, and an association subsystem (b) estimates weights probabilities associating spectrum values with modes corresponding to various signal sources. Comparison of measured and estimated spectra for B34, RB1 data set is shown in (c): (c1) data, (c2) estimated MEANS model, (c3) mode–mean frequencies superimposed on the data.
S(ω_\text{min}), \ldots S(ω_\text{max})$, which are provided by the radar instrumentation. The modeling subsystem includes three types of neurons for each mode, estimating parameters of the model and corresponding to eqns (8)–(10). The weights of these neurons, $P(\text{mle})$, can be interpreted as a probability of a photon at frequency $ω$ originating from source $m$, and the quantity, $P(\text{mle})[S(ω)/\text{hω}]$, can be interpreted as a number of photons at frequency $ω$ originated from source $m$. The output nodes of the modeling subsystem contain the parameters of the spectrum model, $A_m$, $\bar{ω}_m$, $σ_ω$, given by eqns (8)–(10). During learning, these parameters are used by the association subsystem, Figure 1b. The association subsystem input nodes contain the frequency values, $ω_\text{min}, \ldots, ω_\text{max}$, determined by the radar operating parameters. It is comprised of feedforward and competitive layers implementing eqn (11), and its output nodes contain the probabilities $P(\text{mle})$, which serve as weights in the modeling subsystem during learning. In this way, MEANS architecture, neural equations, and learning are determined by the physical model of the spectrum. MEANS learning proceeds via internal iterations; at each iteration, the modeling subsystem uses weights computed by the association subsystem at the previous iteration. Similarly, the association subsystem uses model parameters computed by the modeling subsystem at the previous iteration. Convergence is determined by requiring that model parameter changes between the iterations are below predetermined thresholds. Thus, MEANS learning is unsupervised, the a priori information is contained in its model. The MEANS input data are the Doppler frequency and spectral values and the output is the estimated values of the spectral model parameters.

It is instructive to describe the MEANS estimation process in comparison with classical paradigms in physics and statistics. The objective function in MEANS is the entropy of the photon ensemble, maximization of which is a fundamental principle of statistical physics. In statistical physics, a classical paradigm is to find the distribution of particles by maximizing entropy, given the physical model defining the number of states and their distribution (a forward problem). MEANS solves the inverse problem: given the observed distribution of particles (photons), find parameters of the model. The paradigm of an inverse problem is central in statistical estimation. A rich body of estimation theory is available for simple inverse problems, when there is a set of data. MEANS learning is very fast, because, like Adaline, its weights are functions of relatively few parameters determined by the physical model. MEANS is a stable, convergent system, which can be proved using an EM algorithm (Dempster et al., 1977).

4. CHARACTERIZATION OF IONOSPHERIC CLUTTER

This section describes an application of the developed technique to characterizing ionospheric clutter in over-the-horizon (OTH) radar Doppler spectra. OTH radars operate in the high frequency (HF) band between 5 and 30 MHz and are used to detect and track targets at distances up to 4000 km beyond the maximum line-of-sight range of conventional ground-based microwave radars. To achieve these distances the HF signals propagate, obliquely, reaching maximum altitudes ranging from 90 to 400 km in the ionosphere, and then reflect back to the earth in one bounce, or “hop”. Doppler processing allows one to separate the moving targets from the enormous, non-moving, ground clutter reflection. However, Doppler-spread clutter spreads over a number of Doppler cells and significantly degrade the OTH radar performance. For the basic studies of ionosphere, as well as for future development of clutter suppression techniques, development of clutter models and estimation of their parameters is of interest.

Example of typical OTH radar Doppler return spectra are shown in Figure 1c1, where the amplitude of returns is plotted using gray scale, as a function of Doppler frequency (horizontal axis) and time (or “spectrum number”, vertical axis). Every horizontal line in this plot contains a 64-point Doppler spectrum. There are 150 spectra shown, numbered 0 to 149, all corresponding to the same range of 2856 km and collected from 21:07:44 UT to 23:04:58 UT. Each spectrum is
determined by contributions from four major sources of signals. The main peak contains two contributions: first from energy directly reflected back from the ground (or ocean) through the ionosphere to the receive antenna in one “hop”, and second from a two-“hop” path, which has lower energy and is slightly offset. The vertical dark line corresponds to these ground return peaks near zero Doppler. The Doppler-spread clutter structure, which is of the main interest for our analysis, can be seen to the right of the ground return peak. At time 21:07:44 UT the equatorial (spread-Doppler) clutter is barely discernible from the ground clutter and over the next one and one-half hours (toward the top of the figure, spectrum number 125, recorded at 22:42:26 UT) it evolves into a large structure spread over many Doppler values. (This severely impairs radar functioning, in fact, the radar operator actually changed the operating parameters of the radar for time frames corresponding to spectrum numbers 58–65 and > 125 in an attempt to locate a better operating regime to improve the radar performance, however this resulted in a significant loss of received energy.) The clutter also changes in range and azimuth (not shown), depending upon the spatial variations in the ionosphere causing the spread-Doppler clutter.

For a while, the origin of clutter in OTH radar spectra, that spreads over significant Doppler interval, remained unexplained. Early studies attributing these effects to the interaction of HF signals with either meteor trails or auroral irregularities (Vandrak et al., 1977) did not account for all the instances of clutter observed. In 1989 Franchi and Tichovolsky developed the Phase Screen Ground Modulation (PSGM) model, demonstrating how turbulence could produce Gaussian phase fluctuations in HF signals. They related model parameters to a number of physically significant parameters of electron distribution in ionosphere. Accurate estimates of the temporal and spatial characteristics of clutter spectra are needed for two purposes, first, for improved clutter rejection in normal radar and communication operations, and second to establish the quantitative connection between the observed clutter and ionospheric properties, for the basic scientific studies of ionosphere.

Previous approaches to clutter characterization were of a rather approximate nature (Thomas, 1995): classical spectrum estimation methods did not allow to isolate spread-Doppler clutter from other signal sources and estimate model parameters. The technique developed here utilizes a spectral model composed of several submodels related to physical processes of ionospheric propagation of electromagnetic waves. During the estimation of model parameters, spectral values are partitioned in a fuzzy manner between the submodels. This condition is not uncommon in science, data often are affected by multiple phenomena with unknown characteristics that cannot be isolated in scientific experiments. Thus we believe that the technique described in this paper has broad applicability in scientific data analysis.

We use results of Franchi and Tichovolsky (1989) for modeling the four signal sources described above: spread-Doppler clutter, ground return, and two hop contributions are modeled using Gaussian shape models and the noise “floor” is modeled using a uniform distribution. Figure 1c2, c3 show the results of MEANS estimated spectral model for this data set. Each one of 150 horizontal lines in these plots is processed separately, so for every 64-point spectrum, MEANS iterated until convergence, resulting in a set of model parameters (150 sets of parameters for 150 spectra). These parameters have been used to compute the estimated spectra models according to eqns (5)–(7); these model estimates shown in Figure 1c2 are computed off-line, outside of MEANS, for the comparison purpose only. The model estimates are remarkably close to the data, considering a relatively simplistic form of the model used. The model is seen to correspond to the overall data structure as it changes over time for both the ground return and the spread clutter structures. A comparison of the model parameters, the mean frequencies for each mode \( \omega_m \) to the data is shown in Figure 1c3. The mean frequencies accurately correspond to the data for the most of the 150 spectra for the ground return structure \( (m = 2,3) \) and for the observed spread-Doppler clutter event \( (m = 4) \). Thus, an accurate estimation of the clutter parameters is achieved. These parameters can be related to ionospheric properties using the PSGM model (Franchi & Tichovolsky, 1989), and also, can be used for developing of clutter rejection techniques (Perlovsky, 1995).

5. CONCLUSION

This paper introduces a model-based neural network, whose internal model and learning algorithm implements a new type of spectrum model and a new estimation principle. The estimation principle is based on the maximization of the physical entropy of an ensemble of photons. The model is related to the Einsteinian interpretation of the spectrum as a probability distribution for the photon energy. While ionospheric propagation of electromagnetic waves can be considered as a classical phenomenon, the quantum nature of the electromagnetic field determines statistical properties of a photon ensemble and the estimation procedure, in the same way as the Plank distribution is determined by the quantum structure of the blackbody radiation. A specific parameterization of the spectrum model considered in the paper is determined by the PSGM model (Franchi & Tichovolsky, 1989). The proposed model is versatile and can be used to model any spectrum shape with a desired accuracy by adding modes. This follows from the fact that Gaussians form a complete set of functions. We call this neural network MEANS for Maximum Entropy Adaptive Neural System or Model-based Einsteinian ANS.
REFERENCES


