CONF-Q51117--1

AN APPLICATION OF WAVELET TRANSFORMS AND NEURAL NETWORKS
FOR DECOMPOSITION OF MILLIMETER-WAVE SPECTROSCOPIC SIGNALS*

K. Gopalan†, N. Gopalsami, S. Bakhtiari, and A. C. Raptis

Energy Technology Division
Argonne National Laboratory
Argonne, IL 60439

July 1995

RECEIVED
JAN 30 1995
OSTI

The submitted manuscript has been authored by a contractor of the U. S. Government under contract No. W-31-109-ENG-38. Accordingly, the U. S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U. S. Government purposes.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Paper to be presented at the 21st Annual Conference of the IEEE Industrial Electronics Society, IECON '95, Orlando, Florida, November 6-11, 1995

†Purdue University Calumet, Hammond, IN 46323

*This work is supported by U.S. Department of Energy, Office of Nonproliferation and National Security, Office of Research and Development, Under Contract W-31-109-Eng-38

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED *dx*

An Application of Wavelet Transforms and Neural Networks for Decomposition of Millimeter-Wave Spectroscopic Signals

K. Gopalan ¹, N. Gopalsami ², S. Bakhtiari ² and A.C. Raptis ²

¹ Department of Engineering

Purdue University Calumet

Hammond, IN 46323

² Energy Technology Division

Argonne National Laboratory

Argonne, IL 60439

Abstract - This paper reports on wavelet-based decomposition methods and neural networks for remote monitoring of airborne chemicals using millimeter wave spectroscopy. Because of instrumentation noise and the presence of untargeted chemicals, direct decomposition of the spectra requires a large number of training data and yields low accuracy. A neural network trained with features obtained from a discrete wavelet transform is demonstrated to have better decomposition with faster training time. Results based on simulated and experimental spectra are presented to show the efficacy of the wavelet-based methods.

I. INTRODUCTION

A millimeter (mm) wave sensor is being developed for remote monitoring of airborne chemicals from industrial sites [1]. Detection of chemicals is based on measuring the molecular rotational energy transitions at mm-wave frequencies.. A mm wave source of radiation whose frequency is swept in the range of 225 GHz - 315 GHz, is transmitted to the plume of chemicals to be detected. Measurement of the received power as a function of frequency indicates the absorption of mm wave energy at

frequencies that are characteristic of the chemicals present in the plume. Under low-pressure conditions in which the analyte is sampled into a gas cell, individual chemicals will have resolvable spectral (absorption) peaks at distinct frequencies. In open-path monitoring situation, however, the absorption spectra of the received signal are complex because of the following: (a) Due to increase in pressure (above about 1 torr), the spectrum of a molecule becomes broader relative to that at low pressure; as a result, spectra of different molecules may overlap and merge together. (b) Presence of unknown, nontargeted chemicals may cause merging and broadening of spectra. (c) Instrumentation noise may contribute to the low-level absorption data and nonuniformly vary the spectral peaks. This paper reports the results of signal processing methods employed to obtain the composition of certain targeted chemicals from experimentally observed spectra.

II. ACQUISITION OF EXPERIMENTAL SPECTRA

The experimental setup used at Argonne for acquiring mm wave spectroscopic data employs an active mm wave source with a capability for sweeping the frequency in two ranges, 225 - 270 GHz and 270 - 315 GHz [2]. Using waveguides and an antenna the mm wave signal is transmitted into one end of a gas cell containing the chemical mixture to be analyzed. The transmitted signal available at the other end of the gas cell is detected by a hot-electron bolometer. Voltage output of the bolometer is proportional to the power of the received mm-wave radiation. A preamplifier filters high frequency noise and provides voltage gain. A 12-bit digitizer is used to acquire the data at the output of the preamplifier. The received signal power in the range of 225 - 315 GHz is

obtained from the combined data acquired for the two frequency ranges. Finally, the percent absorption spectrum is obtained as a ratio:

$$\% \text{ absorption} = 100(P_r - P_s)/P_r$$

where P_s = received signal power with chemical present

P_r = received signal power without chemical

III. FEATURE EXTRACTION AND CHEMICAL DECOMPOSITION

It is observed that the effect of high pressure on the mm wave spectrum of a chemical is to broaden the frequency of absorption while keeping the peak absorption at the same frequency. The broadening effect is given by the Lorentzian shape function as [3]

$$S(f) = K f_L / [(f - f_0)^2 + f_L^2]$$

where f_0 is the absorption (center) frequency

f_L is the Lorentz line width

K is a constant representing the total strength

In addition to the pressure-broadening effect, instrumentation errors contribute to the difficulty of decomposition. A major source of noise is the frequency jitter of the

mm wave source whose frequency is not calibrated on-line. Additionally, acquisition of data in two ranges of frequencies, because of the limit band width of the triplers, causes jitter in the spectrum around 225 GHz. Because of these and the other errors stated previously, it is essential that any chemical decomposition method reduces redundancies in the spectra and is robust with instrumentation noise.

The following sections describe the methods presently employed for chemical decomposition. These methods were initially tested using simulated mixture of data. Simulated pressure-broadened data were generated corresponding to 1 atm from low-pressure spectra available from a Jet Propulsion Laboratory data base [4]. Absorption spectra in the frequency range of 220 - 320 GHz for three chemicals, namely, $\text{CH}_3\text{Cl}^{35}$, CH_3CN and HC_3N , referred to as c_1 , c_2 and c_3 , were synthesized at atmospheric pressure by applying Lorentzian line shape function for each individual spectrum at low pressure. A random combination of these three chemicals were added to form a composite (synthetic mixture) at high (atmospheric) pressure as

$$p = c_1 n_1 + c_2 n_2 + c_3 n_3.$$

where n_1 , n_2 and n_3 are the individual concentrations of c_1 , c_2 and c_3 , respectively. It is to be noted that the concentration of each chemical is in the ratio of its pressure to the total pressure of all the chemicals in the mixture.

A total of nine such linear combinations, p_k , $k = 1, 2, \dots, 9$ were used as reference, or training, data base with known compositions, n_{1k}, n_{2k}, n_{3k} . Absorption data for each of the nine synthetic compositions had 941 samples covering the frequency range 220 - 320 GHz. A different mixture of the three chemicals was then used as test data set.

A two-layer backpropagation (BP) neural network was used in all tests to approximate the composite data as a combination of network weights and biases. Feature vectors of size $N \times 1$ corresponding to each raw spectral data vector were presented to the N source (input) nodes of the network. A hidden (first) layer of L neurons and an output layer of M nodes, where M is the number of chemicals in the composition, were used in the network. MATLAB Neural Network Toolbox from MathWorks, Inc. [5] was used to implement the two layer BP architecture with $L = 17$.

Initially, raw data vectors (simulated, pressure-broadened spectra) were used directly as reference and unknown inputs to the network. The network was trained using a random set of nine synthetic mixtures of data and their compositions. After 50,000 epochs (iterations) of training, an error of about 5.5×10^{-5} was reached. The resulting weight and bias values were used on a mixture of $[c_1 \ c_2 \ c_3]' = [0.7 \ 0.35 \ 0.5]'$. The network output for this data was observed as $[0.7004 \ 0.3508 \ 0.4997]'$.

IV. DECOMPOSITION USING WAVELET-BASED FEATURES

Clearly, the results of training the BP network are quite good, with a maximum error of below 0.3 % for the decomposition of individual chemicals. The large number of iterations, however, requires a long period of training for the network that uses raw spectral data. Training time is also increased for experimental spectra which are usually 2000 points or higher. By removing redundant information and presenting only the key features in the data to the network, training time can be reduced. The effectiveness of feature reduction, however, depends on the characteristics of the data to be represented.

Because of the nonstationary nature of the spectral data, a time-frequency analysis based on wavelets has been used for feature extraction. In these analyses the spectral data of a composite chemical are considered as time samples of a signal whose frequencies vary with time. By using a set of nonorthogonal wavelets in wavelet transform analysis, the one-dimensional spectral signal is decomposed into two dimensions of time and frequency [6,7]. The following sections describe the wavelet-based processing employed for the chemical decomposition.

A. Review of Wavelet Transform Analysis

Given an analyzing mother wavelet $g(t)$ (see below), the wavelet transform $Wg(a,b)$ of a signal $f(t)$ with respect to $g(t)$ is defined by [6]

$$Wg(a,b) = \langle f, g_{a,b} \rangle = (1/\sqrt{a}) \int f(t) g^*((t-b)/a) dt \quad (4.1)$$

where

$$g_{a,b}(t) = (1/\sqrt{a})g((t-b)/a) \quad (4.2)$$

and g^* is the complex conjugate of g . In the above equation a is the scaling (dilation) parameter, and b is the translation parameter. Note that if the Fourier transform of $f(t)$ is

$$\mathfrak{F}\{f(t)\} = F(w) ,$$

then, for $a > 1$,

$$\mathfrak{F}\{f(t/a)\} = aF(aw).$$

Hence, the signal $f(t/a)$ represents the frequency-compressed (or, time-dilated) waveform of $f(t)$.

A commonly used wavelet is the Morlet wavelet $g(t)$, which satisfies the admissibility conditions of integrability, square-integrability and low frequency behavior. This modulated gaussian function is given by

$$g(t) = \exp(-Kt^2) \exp(j\omega_0 t), \quad (4.3)$$

Note that $g(t)$ given above has zero mean, i.e., $\int g(t) dt = 0$ so that $G(0) \approx 0$ (for $\omega_0 > 5$) and

$$G(\omega) = (\sqrt{\pi/K}) \exp[-(\omega - \omega_0)^2 / (4K)]$$

is a smooth function in ω . More importantly, the Morlet wavelet $g(t)$ is concentrated in time and frequency, i.e., the time-bandwidth product is the smallest because of the gaussian function. It has been shown that this function is well suited for locating abrupt changes in signals. Note further that $g(t)$ is concentrated around $t = 0$ while $G(\omega)$ is centered at $\omega = \omega_0$. The scaled and translated function

$$g_{a,b}(t) = \exp(-K((t-b)/a)^2) \exp(j\omega_0((t-b)/a)) \quad (4.4)$$

is concentrated around $t = b$ while its Fourier transform, $G_{a,b}(\omega)$, has peak amplitude at $\omega = \omega_0/a$. Thus, increasing the scale a (> 1) narrows the time window and hence increases the analyzing range of frequency; conversely, as the dilation parameter a decreases, time window is widened and finer details of the signal become available.

B. Signal Decomposition using Morlet Wavelets

Based on the above observation of decreasing the dilation parameter a in Eq. (4.1) for obtaining coarser to finer details in $f(t)$, the raw spectral data are transformed using

Morlet wavelets. Fig. 1 shows a Morlet wavelet (for $K = 5$ and $\omega_0 = 12.5$ in Eq. (4.3)).

In Fig. 2 the synthesized spectral signals for c_1 ($\text{CH}_3\text{Cl}_{35}$) and c_2 (CH_3CN) at atmospheric pressure and their wavelet transforms (magnitudes of $Wg(a,b)$ given by Eq. (4.1)) for $a = 1$ and $b = 1, 2, \dots, N$, where N is length of signal in samples, are shown. From this figure it is clear that even at a single scale the wavelet transform brings out the key features of the signals while filtering out high frequency redundancies.

Fig. 1 Morlet wavelet ($K = 5$ and $\omega_0 = 12.5$ in Eq. (4.3))

Fig. 2 Spectral signals and their one-scale wavelet transform magnitudes (a) c_1 ($\text{CH}_3\text{Cl}_{35}$), (b) DWT of c_1 , (c) c_2 (CH_3CN), (d) DWT of c_2

Because of the discrete nature of the data, the continuous transform given in Eq. (4.1) is discretized for implementation; in addition, only discrete sets of scale $a = 2^i$ and translation $b = n2^i$ are generally used for efficient implementation. The resulting discrete wavelet transform (DWT), which enables extracting the features in $f(n)$ at higher scales of a in Eq. (4.1) using a particular $g(n)$, is given by Eq. (4.5).

$$W(a,b) = W(2^i, n2^i) = (1/\sqrt{2^i}) \sum f(k) g^*((k/2^i) - n), \quad (4.5)$$

$i = 0, 1, \dots$, and $n = 0, 1, 2, \dots N$

Each $i (= \log_2 a)$ is termed an octave of the transform. The scale a in the above equation increases successively by a factor of 2; hence, samples of discrete dyadic wavelets given by

$$\{g_{a,b}(t)\} = \{(1/\sqrt{2^i})g((k/2^i) - n)\} \quad (4.6)$$

are needed for the evaluation of Eq. (4.5). As i increases, i.e., at higher octaves, $g(t)$ in Eq. (4.6) must have larger number of samples so that the wavelet is adequately represented. To avoid generating a large number of samples of $g(t)$, $g(k/2^i)$ for noninteger arguments are obtained by interpolation using a low-pass filter [7,8]. This results in an efficient implementation as shown in Fig. 3 where the signal f is interpolated instead of g . In this figure, s is a low-pass (a trous) filter which, when convolved with $f(n)$, leaves the even-numbered values intact at the output while interpolating for odd-numbered points. DWT coefficients, which are the outputs of the decimated g -filter in each scale, form the feature vectors for a given $f(n)$.

Fig. 3 Computation of decimated DWT using Eq. (3.5)

V. RESULTS AND DISCUSSION

The above implementation of DWT with Morlet wavelets was used to obtain two-scale representations ($i = 0, 1$) for the same synthetic mixture of data discussed in Sec. II. Imaginary part of the DWT coefficients were used as feature vectors in a three layer neural network with 20 neurons in the hidden layer. Training error of below 10^{-5} was reached after approximately 2400 iterations. For a synthetic test mixture of $[c_1 \ c_2 \ c_3] = [0.7 \ 0.35 \ 0.5]$, the trained network yielded an output of $[0.7007 \ 0.3508 \ 0.4993]$. With random noise at 10 % of signal power added to the test mixture, the same network produced a concentration of $[0.7112 \ 0.3595 \ 0.5053]$. Clearly, the preprocessing by DWT reduced the redundancy in the data and contributed to fewer number of training steps. Although the DWT-based features of the spectral data are obtained offline, features for the unknown composition must be available for decomposition. DWT computation using the scheme shown in Fig. 3, however, is highly efficient requiring little added time for testing.

Individual spectral data for three chemicals, namely, CH_3Cl (c_1) CH_3CN (c_2) and CH_2Cl_2 (c_3), were obtained experimentally, each at a partial pressure of 1 torr. For testing, the three chemicals were mixed with partial pressures of 0.4 torr (c_1), 0.2 torr (c_2) and 0.6 torr (c_3) and the experimental spectra for the mixture was obtained. From the individual spectra of the three chemicals, nine random mixtures (linear combinations) were formed to constitute as training data set. The magnitudes of the DWT of the training spectra were formed as modified feature vectors and the neural network

discussed in III was trained using scales of $i = 0, 1$, and 2 . Fig. 4 shows the spectrum and wavelet domain feature vector of the test mixture. Results of the tests using two Morlet wavelets at different center frequencies are shown in Table 1.

Fig. 4 (a) Spectrum of experimental test mixture of three chemicals (b) Feature vector of the test mixture

TABLE 1
DECOMPOSITION RESULTS

(a) For a synthetic mixture of three chemicals, c_1 , c_2 and c_3

Synthetic mixture with composition	Network using Raw spectra	Network using DWT	Network using DWT ¹
$c_1 = 0.70$	0.7004	0.7007	0.7112
$c_2 = 0.35$	0.3508	0.3508	0.3595
$c_3 = 0.50$	0.4997	0.4993	0.5053
No. of epochs	50000	2400	2400

¹ Test spectra + random noise at 10 % of signal power, using the same network as without noise.

(b) For experimental mixture of three chemicals, c_1 , c_2 and c_3 using a set of individually measured spectra for each chemical

Unknown mixture with composition	Network using Raw spectra	Network using DWT magnitude ¹ Three scales	Network using DWT magnitude ² Two scales	Network using DWT magnitude ² Three scales
$c_1 = 0.4$	0.3526	0.3239	0.3539	0.3540
$c_2 = 0.2$	0.3034	0.2873	0.3557	0.3069
$c_3 = 0.6$	0.4286	0.5372	0.4482	0.4119
No. of epochs	1600	500	1010	1010

¹ $\omega_0 = 5$; ² $\omega_0 = 50$

Note that an analyzing wavelet with smaller the center frequency ω_0 ($= 5$) has narrower frequency resolution and hence more variations in the signal are brought out in the feature vector. To achieve the same frequency resolution using higher ω_0 ($= 50$), many more scales (with lower values of scale a) must be used in the construction of the feature vector. At lower number of scales, the burden of adequately representing the signals is placed on the neural network; hence, the network training time is increased. In general, therefore, the choice of the wavelet depends on the nature of the spectra of the signals that are available for training and decomposition.

VI. CONCLUSION

A method of representing mm wave spectroscopic signals using Morlet wavelets in discrete wavelet transform has been presented. When used with a neural network the representation has been shown to decompose accurately the concentrations of constituent chemicals in a synthetic mixture of chemicals. It is also demonstrated that the representation results in a reasonably accurate decomposition of experimental spectra of chemicals. Further studies in the decomposition of mm wave spectroscopic signals using wavelet domain features are in progress.

VII. REFERENCES

- [1] N. Gopalsami, S.L. Dieckman and A.C. Raptis, "Millimeter-Wave Remote Sensing for Detection of Effluent Species," Trans. Am. Nuclear Society, vol. 68, pp. 182-183, 1993.
- [2] N. Gopalsami, S. Bakhtiari, and A. C. Raptis, "Development of a Millimeter Wave Sensor for Environmental Monitoring," Proceedings of 40th SPIE Annual Meeting, San Diego, CA, July 9-14, 1995.
- [3] C. Elachi, *Introduction to the Physics and Techniques of Remote Sensing*, John Wiley & Sons, New York, 1987.

- [4] R.L. Poynter and H.M. Pickett, "Submillimeter, Millimeter, and Microwave spectral Line Catalog," JPL Publication 80-23, Revision 2, 1983.
- [5] MATLAB, *High Performance Numeric Computation and Visualization Software*, The Mathworks Inc., Natick, MA, 1993.
- [6] J.M. Combes, A Grossmann and Ph. Tchamitchian (Eds.), *Wavelets: Time-frequency Methods and Phase Space*, Springer-Verlag, Berlin, 1989.
- [7] M. Holschneider, et al., "A Real-time Algorithm for Signal Analysis with the help of the Wavelet Transform," *ibid*, pp. 286-297.
- [8] M.J. Shensa, "The Discrete Wavelet Transform: Wedding the A Trous and Mallat Algorithms," *IEEE Trans. Signal Proc.*, vol. 40, No. 10, pp. 2464-2482, Oct. 1992.

Fig. 1

Fig. 2 a

Fig. 2 b

Fig. 2 c

Fig. 2d

Fig. 3

Fig. 4a

Fig. 4b