ULTRASONIC CLASSIFICATION OF GASES USING PRINCIPAL COMPONENT ANALYSIS

P.-E. Martinsson[*,†]**, J. E. Carlson**[†]**, and R. E. Carlson**[‡]

[†]EISLAB, Dept. of Computer Science and Electrical Engineering, Luleå University of Technology, Sweden.
[‡]Department of Chemistry, Tromsø University, Tromsø, Norway
[*]Email: Par-Erik.Martinsson@sm.luth.se

Abstract

On-line measurement of the energy content of natural gas is of interest for both industry and customers, since the energy content determines its monetary value.

Experiments with pulsed ultrasound show that, in addition to changes in speed of sound and acoustic attenuation, there is also a change in the shape of the sound waveform. In this paper we investigate the feasibility of using *Principal Component Analysis* (PCA) to quantify this change in pulse shape.

The principle is evaluated for Oxygen and Ethane, for different pressures and different temperatures. The results show that by using PCA, it is possible to distinguish between pulses that have propagated through Oxygen from pulses in Ethane.

Introduction

Natural gas contains a mixture of several gases, like Ethane, Methane, Nitrogen, etc. Sometimes also small, but highly undesired fractions of Oxygen. It is of interest to both provider and customer to know the composition of such gas mixtures, since this determines the energy content (calorimetric value) and the combustion properties and thus, the monetary value of the gas.

With the use of different sources of gas (i.e. different gas fields) the energy content of the gas delivered to customers may vary considerably. Variations up to 20% are not uncommon and can be on scale of seconds [1]. For both customers and gas companies it therefore becomes of importance to measure the energy content in order to get a fair invoice. This research is motivated by the lack of commercial methods that are both cheap and reliable.

Today, the energy content of gases is measured using either gas chromatography or calorimetery. Both methods are accurate, but requires samples of the gas to be removed and analyzed separately. They are also relatively slow. Because of this, the existing techniques are not suitable for on-line measurement at the customer side of the distribution line.

Typically, both the speed of sound and the absorption of sound within a gas vary with temperature, frequency and pressure etc. Hence both these parameters can be used to classify and/or detect a pure gas. The hypothesis we wanted to investigate with this work was weather the ultrasonic pulses transmitted through a gas contain additional information. We investigate the feasibility to use *Principal Component Analysis* (PCA) [2] to extract parameters from the measured ultrasound pulses, that can be used to classify which gases are present.

Before the analysis, the effects of speed of sound and attenuation were removed from the pulses. For narrowband pulses the effect of dispersion (i.e. the frequency dependence of the speed of sound) is minimal. Hence, we assume that a speed of sound change is visible mainly as a change in propagation time of the entire pulse. Since we are interested primarily in the shape of the received pulse (and not the arrival time) the pulses were aligned in time before the analysis. The effect of attenuation can easily be removed by normalizing each pulse to unit energy. Once this is done, we are left with a set of pulses with the same energy, starting at the same time. In this paper we show how the remaining variations can be studied with the use of PCA, and how this can be used to discriminate between pulses that have propagated through different gases.

Pure samples of Oxygen (O_2) and Ethane (C_2H_6) where used in order to perform the experiments required to test and validate the hypothesis.

The results show that in addition to the speed of sound and acoustic attenuation, the shape of the pulse can be used to extract parameters necessary to identify a gas. The results in this paper are preliminary and should be interpreted as an example of the analysis strategy rather than a complete analysis method.

Experimental Setup

A pulse echo measurement scheme was used to measure the attenuation and speed of sound for Oxygen and Ethane. The acoustic properties of interest vary with frequency, f, and pressure, P. The temperature and frequency dependence on sound velocity and acoustic attenuation, has previously been investigated by, for example, Martinsson [3], Lueptow [4], [5], and Bhatia [6]. Changes in frequency and pressure are normally studied as the ratio f/P. There are two ways to vary f/P in a pulse echo system. Transducers of different center frequency can be used for a fixed pressure, or one transducer can be used while the pressure is changed. The later of the two principles was chosen for the work in this paper. A 1 MHz air transducer was used while the static pressure was varied.

A custom-built pressure chamber (see Fig. 1) was
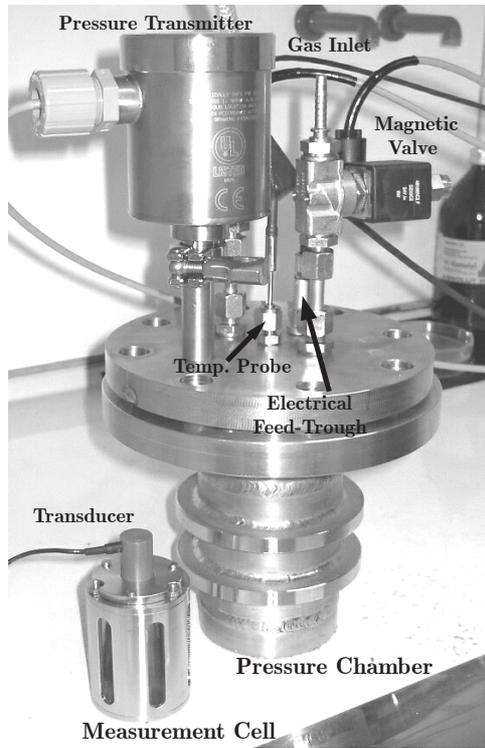
Figure 1 : The measurement equipment. All measurements were performed in a pressure chamber. The pressure was varied between 0 bar and 6.4 bar. The air transducer had a center frequency of 1 MHz.

used to achieve different static pressures. For each gas, the pressure was varied from 0 to 6.4 bar in 19 steps. Since the attenuation in Ethane is extremely high at low pressures and high frequencies, we were not able to make measurements for that particular gas for pressures above 1.86 bar.

The transducer was mounted on a stainless steel measurement cell, as seen in the lower left corner of Fig. 1. The measurement cell was then immersed into the pressure chamber. The whole setup was then placed in a temperature controlled chamber (*Heraeus Vötsch HT4010*), keeping the temperature constant at either 15 °C, 20 °C, and 25 °C.

The pressure in the chamber was measured with an *ANDERSON TPP Pressure Transmitter* with a range of up to 13.6 bar above atmospheric zero. The transmitter have an accuracy of approximately 30 mbar. This includes the combined effects of linearity, hysteresis and repeatability.

To excite and receive acoustic pulses from the transducer, a *Panametrics Pulser/Receiver Model 5072* was used. For the transmitting mode, the pulser/receiver was set to deliver maximum energy to the transducer, which corresponds to a short voltage peak with 380 V amplitude with an energy of 104 μJ. In receive-mode, a built-in amplifier allows the signals to be amplified

up to 59 dB. For the experiment described in this paper the gain was typically set to 20 dB throughout the measurements, but for some of the pulses recorded at low pressure, a higher gain was used. The output of the amplifier is linear up to ± 1 V.

All pulses were sampled with a *Tektronix TDS 724*, 1 GHz oscilloscope. For each experimental setting, 50 pulses were recorded and transferred to a PC, where they were averaged and further processed.

For each measurement, the temperature was recorded using an encapsulated PT100 sensor mounted trough the wall of the pressure chamber.

The transducer used was originally designed for operation in air which have an acoustical impedance $Z_{air} = 4154$ Pa·s/m (at T=20 °C and P=1 bar). Since the acoustic impedance of the gases used are different from air, the transducer will not operate at its optimal performance, and consequently not transmit as much power as desirable.

Theory

Aligning and normalizing

In order to be able to compare what effect the gases have on the shape of the ultrasound pulse, the effect of speed of sound and acoustic attenuation have to be removed. This was done by normalizing each pulse to unit energy, and aligning the start of the pulses in time. Of course, the values of attenuation and speed of sound can be stored for future use.

Principal component analysis (PCA)

With the *Singular Value Decomposition* (SVD) [7], any matrix $\mathbf{X}$ with rank r, can be factored as

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T. \qquad (1)$$

The SVD has some important properties, summarized below:

- First r columns of $\mathbf{U}$ form an orthonormal basis for the column space of $\mathbf{X}$.
- First r columns of $\mathbf{V}$ form an orthonormal basis for the row space of $\mathbf{X}$.

The r non-zero diagonal elements of $\mathbf{S}$ are the square-roots of the eigenvalues of $\mathbf{X}^T\mathbf{X}$ and $\mathbf{X}\mathbf{X}^T$.

To examine how much each of the columns in $\mathbf{U}$ contributes to the total experimental variation, the singular values, σ_i, are be scaled according to

$$\tilde{\sigma}_i^2 = \frac{\sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2}, \qquad (2)$$

Fig. 2 shows this for the experiments in this paper.

PCA [2] is normally described as factoring the matrix $\mathbf{X}$ into *scores*, $\mathbf{T}$, and *loadings*, $\mathbf{P}$, where

$$\mathbf{P} = \mathbf{V} \tag{3}$$
$$\mathbf{T} = \mathbf{US}. \tag{4}$$

In this paper, the the columns of $\mathbf{X}$ are given by the samples of the pulse waveform, and we wish to find a model for the shape of the pulse. To examine how different experimental conditions affect the pulse, the columns of $\mathbf{X}$ are scaled to unit energy, that is

$$\mathbf{x}_n = \frac{\tilde{\mathbf{x}}_n}{\tilde{\mathbf{x}}_n^T \tilde{\mathbf{x}}_n}, \tag{5}$$

where $\tilde{\mathbf{x}}_n$ denotes the n:th original pulse. The reason for this scaling is to avoid the risk of *a priori* assigning any of the pulse waveforms greater significance than the other. In other applications the pulses are scaled to unit variance or to unit maximum amplitude.
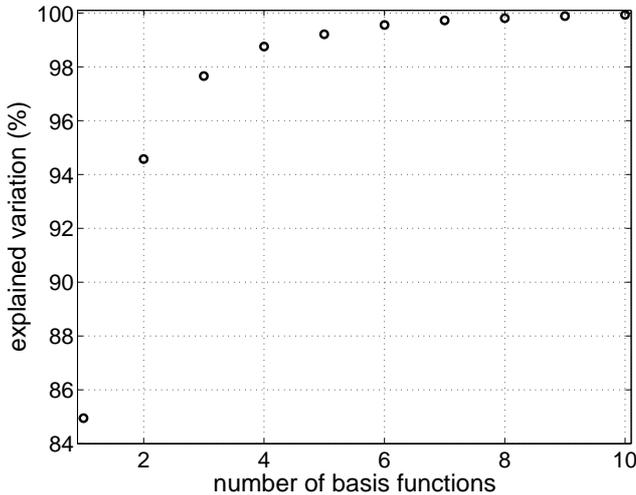
Figure 2 : Contribution of the basis vectors (columns in $\mathbf{U}$) to explaining the total variation in the measurement set. Note that the first 6 describe over 99% of the total experimental variation.

Results

Experiments were made for Ethane and Oxygen for frequency/pressure ratios (f/P) in the range of 0.0–6.4 bar, and for temperatures 15 °C, 20 °C, and 25 °C. For each experiment, 50 pulses were measured and averaged. These average pulses were then normalized in energy and aligned in time. The resulting pulses were stored as columns of the matrix $\mathbf{X}$.

The matrix $\mathbf{X}$ was then decomposed into its principal components, as described in the previous section. Fig. 2 shows how much of the total experimental variation is explained as we append the model with basis functions

(columns of $\mathbf{U}$). From Fig. 2 we see that using the first 6 basis functions is enough to explain more than 99% of the total variation. This means that, once attenuation and speed of sound has been accounted for, all the pulses can be accurately modeled as a linear combination of the first 6 columns in $\mathbf{U}$, i.e.

$$\mathbf{x}_n \approx v_{1n}\sigma_1\mathbf{u}_1 + v_{2n}\sigma_2\mathbf{u}_2 + v_{3n}\sigma_3\mathbf{u}_3 +$$
$$+ v_{4n}\sigma_4\mathbf{u}_4 + v_{5n}\sigma_5\mathbf{u}_5 + v_{6n}\sigma_6\mathbf{u}_6. \tag{6}$$
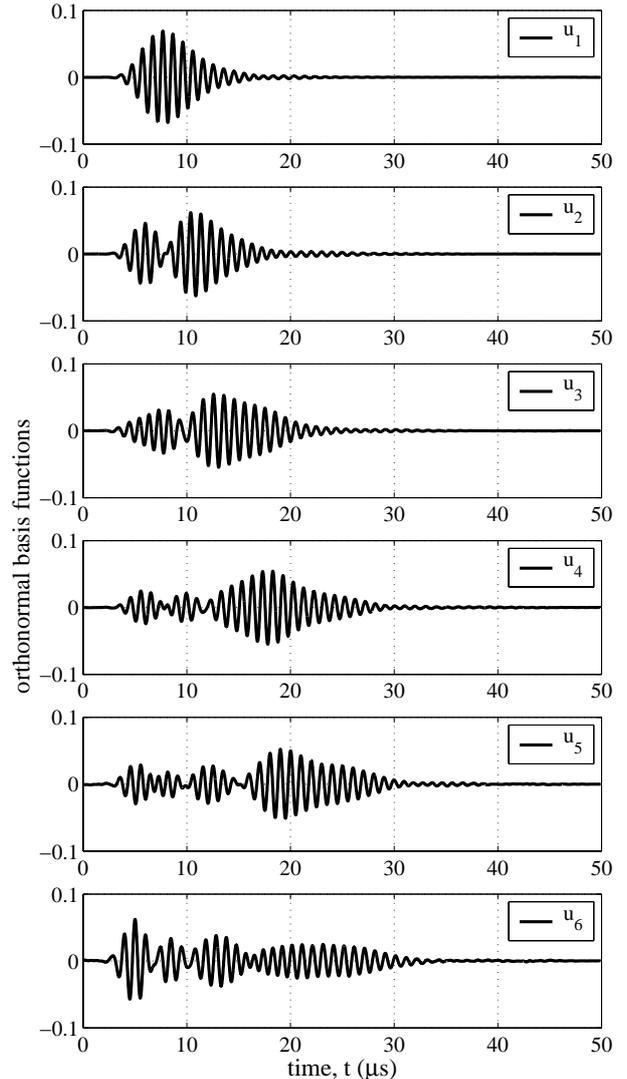
Figure 3 : Basis functions ($\mathbf{u}_1 \ldots \mathbf{u}_6$) corresponding to the six largest singular values.

The column vectors of $\mathbf{U}$ are orthonormal, which means they represent linearly independent components. Fig. 3 shows the 6 most significant components.

In Fig. 4 we clearly see that there are some properties of the ultrasound waveform that differ from Ethane and Oxygen. Most of the coefficients for Oxygen are located in the upper half of the plot, while those for
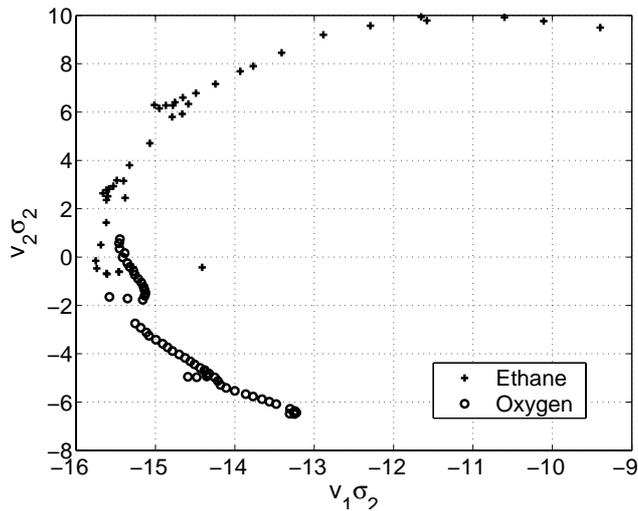
Figure 4 : Coefficients of the first basis function plotted versus the coefficients of the second basis vector.

Ethane are located at the bottom half. In Fig. 4 the y-axis is the numerical value of the second coordinate corresponding the second basis function in Fig. 3. The x-axis is the first coordinate corresponding to the first basis function in Fig. 3. Both axes have been scaled with the corresponding singular value.

Discussion

In this paper we have shown that a PCA can be used to extract information from the shape of the ultrasound pulse, that can be used to distinguish between Oxygen and Ethane. For example, the value of v_2 (see Fig. 4) differs significantly between Oxygen and Ethane.

Once we know which gas is present, more information can be extracted easily. If pressure, temperature, speed of sound, and acoustic attenuation are measured at the same time, other properties, like acoustic impedance and density, can be calculated [8].

It should be pointed out that some points of the Ethane sample occurs among the Oxygen points, and vice versa. These points correspond to samples measured at low pressure, and hence the noise level is higher for those particular measurements. This problem can be eliminated by, for example, using a more powerful transmitting equipment or by decreasing the travel distance in the measurement cell.

If the principles discussed in this paper would be used in a sensor application, we recommend using a high pressure and/or a low frequency. That would minimize the frequency to pressure ratio and give good signal-to-noise ratio.

As discussed in the theory section, the pulses are aligned in time, to remove the effect of changes in speed of sound before the PCA. However, there is still some time jitter remaining between the different measurement. This is difficult to compensate for, since the shape of the pulse changes. A successful jitter correction would most likely further reduce the dimensionality of the problem. This will be considered in future work.

Conclusions

In this paper we have shown that, except for the traditional attenuation and speed of sound values, it is possible to extract additional information from the pulse shape. This was done using a *Singular Value Decomposition* (SVD). Experiments show that the basis functions obtained by the SVD can be used to quantify changes in pulse shape stemming from changes in temperature f/P-value, and gas composition. Preliminary experimental results indicate that it should be possible to distinguish between Ethane and Oxygen, even during variation of pressure and temperature.

References

[1] J. Delsing and I. Blom, "On-Line Measuements of Energy Content of Bio Gas and Natural Gas Mixtures," Tech. Rep. ISSN: 0282-3772, Lund Institute of Technology, 1995.

[2] I. T. Jolliffe, Principal Component Analysis. Springer Verlag, New York, 1986.

[3] P.-E. Martinsson and J. Delsing, "Ultrasonic Measurements of Molecular Relaxation in Ethane and Carbon Monoxide," in Proc. IEEE Int. Ultrason. Symp., Munich, Germany, October 8–11 2002, pp. 494–499.

[4] Y. Dain and R. M. Lueptow, "Acoustic Attenuation in Three-Component Gas Mixtures - Theory," J. Acoust. Soc. Am., vol. 109, pp. 1955–1964, 2001.

[5] Y. Dain and R. M. Lueptow, "Acoustic Attenuation in Three-Component Gas Mixtures - Result," J. Acoust. Soc. Am., vol. 109, pp. 1974–1979, 2001.

[6] A. B. Bhatia, Ultrasonic Absorption - An Introduction to the Theory of Sound Absorption and Dispersion in Gases, Liquids and Solids. Dover Publications, New York, 1985.

[7] G. Strang, Linear Algebra and its Applications, 3rd ed. Harcourt Brace Jovanovich, San Diego, 1986.

[8] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, Fundamentals of Acoustics, 3rd ed. Wiley, New York, 1982.