Histogram Classification of Acoustic Breath Sounds

Bryce Hill B.S. NASA Fellow

***Abstract*—A novel approach to classifying sounds recorded at the trachea is introduced. The sounds in question are normal breath sounds and snore sounds. The sounds are classified by a the combination of a histogram classifier and a repeated pulse counting classifier. Individually the classifiers performed poorly, but when used together they performed quite well. This is an ongoing study into developing an acoustic respiratory monitor.**

I. Introduction

Patients undergoing sedation procedures often experience a cessation of breathing due to two types of respiratory distress. Respiratory depression is the first cause which happens as the subject receives enough sedation drug to remove drive to breath. The second cause is respiratory obstruction which happens when the subject's oropharynx collapses due to the lack of muscular tension. The rescue procedure for each of these cases is different and currently there is no one device that can alert the physician as to the cause of respiratory distress.

As a solution to the above mentioned problem an old technology is being considered not only for the classification of respiratory distress as obstruction or depression, but also a respiratory rate monitor. The old technology is that of listening to the tracheal breath sounds through a precordial stethoscope. The new technology is automating the analysis of these sounds. Previous work has already been done to calculate the breath rate from the acoustic signal and compare it to the breath rate of a measured flow signal with a good deal of accuracy [1]. An accurate breath rate gives an indication of when apnea occurs in general, but determining the type of apnea may be able to be determined by the characteristics of the sounds.

Currently the standard for monitoring during a sedation procedure is the pulse-oximeter which measures the blood oxygen saturation in a finger. The main problem with this monitor is the time delay from the cessation of breathing to the de-saturation of the blood at the finger. This delay can be several minutes[2]. Other monitors such as a flow meter are able to measure flow volume or proportional flow but are unable to determine the cause of apnea and are often cumbersome and awkward. Other monitors measure the chest and abdomen movements which can be proportional to flow volume but in cases of obstruction can continue to show valid breathing. Aside from missing periods of obstructive apnea these monitors only measure a relative flow volume and are cumbersome[3][4].

The two major challenges associated with building an acoustic respiratory monitor are detecting when apnea occurs by use of the respiratory rate measured and classifying the period of apnea by the sounds preceding the apnea.

II. Data Set

Data was collected from 24 subjects for an IRB approved study. Each subject was sedated using a combination of remifentinal and propofol in incrementing dosages. During the sedation procedure a precordial stethoscope equipped with a microphone was placed on the trachea and collected a single channel of acoustic data at 22050 Hz with 24 bit resolution. Flow data was also recorded with the use of a tight fitting facemask with the flow being measured by a differential pressure monitor (Cosmo +II Respironics) which was collected at 100 Hz. Each subject was also fitted with bands around their chest and abdomen to measure the change in circumference by means of the change of inductance in the bands, called Respiratory Inductance Plethysmography (RIP) which was collected at 100 Hz. Medical staff administered the drugs to the subjects and performed assessments on a regular basis.

The data collected at 100 Hz was later synchronized with the acoustic data by means of visual comparison of the flow data to the envelope of the audio data.

III. Methods

Snoring and obstruction are caused by the same mechanism [5] namely the collapse of the upper airway. The difference between partial airway obstruction and snoring is the volume of gas moved during each breath. Partial obstruction will limit the amount of volume to less than 200 ml. Snoring also often precedes obstruction as the subject gradually descends into apnea. [6] used the general rule that snoring is much louder than normal breath sounds to distinguish snores. Snoring sounds recorded at the trachea are audibly different from other breath sounds. Although amplitude is a simple indicator there is no threshold that can be defined to differentiate between snores and clear breaths due to change in background noise and placement of the precordial stethoscope.

In order to classify breath sounds features other than amplitude will need to be processed. Normal breath sounds are generated by turbulence in the trachea and upper airway. Sound generated by any kind of turbulence is literally caused by the sound of the fluid colliding with itself. Turbulence at the trachea is relatively faint until amplified by a stethoscope cup. It was also initially observed that normal breath sounds sound white and may have a Gaussian distribution. Later it was found that [7] used a Gaussian random noise generator to simulate this kind of sound. Initial observations have shown that most normal breath sounds do indeed have a Gaussian distribution, but this observation was not true across all normal breath sounds. Some normal breath sound distributions became heavy tailed and had a central peak greater than a Gaussian distribution as shown in figure 1. It was soon discovered that this disparity was caused by the non-stationary variance of the audio signal. In order to

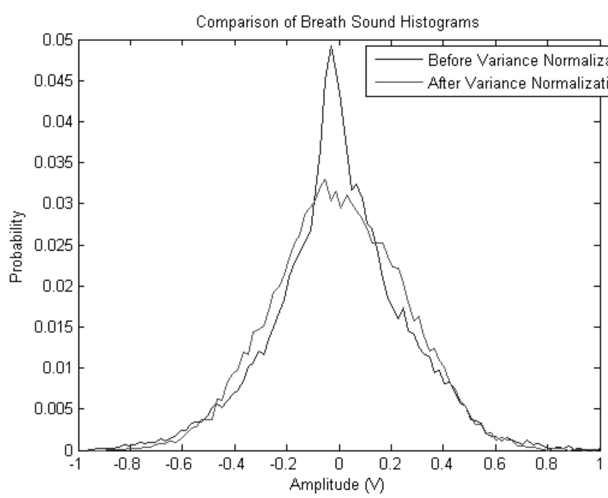

Figure 1: Comparison of the distribution of a normal breath sound before and after variance normalization.

further understand this a normal breath sound was modeled as a normal distribution with a changing variance over time $N(t,\sigma(t))$. The non-stationary variance of the normal distribution $N(t)$ was simulated by a single rectified sine wave $s(t)$ to create a simple match to the varying flow as shown in figure 2. The sound was further simulated as a Gaussian signal modulated by a sinusoid to create the signal $N(t,\sigma(t))=s(t)\cdot N(t)$. The standard deviation of the signal was then measured using transfer function $S(n)=\sqrt{(h(n)*(N(t,\sigma(t)))^2)}$ where

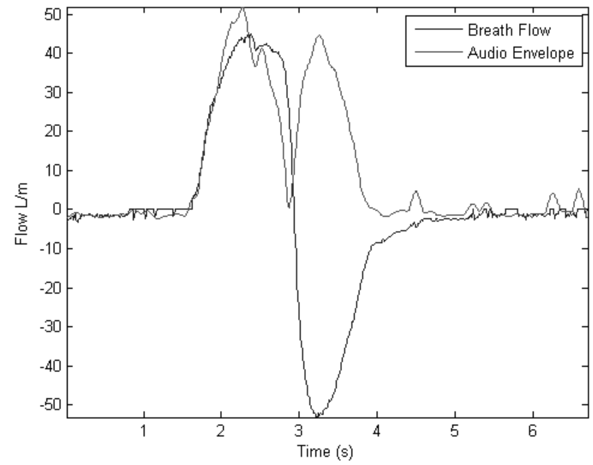

Figure 2: Normal breath to be modeled as a rectified sinusoid.

$h(n)=\frac{[1,1,1,1...1]}{N}$. The standard deviation was then divided from the modulated signal to attempt to create a uniform variance Gaussian signal which should have a normal distribution as shown in figure 3. The actual distribution for Gaussian signal modulated by a sinusoid is $f_Z(z)=\frac{1}{\pi\sqrt{(2\pi)}}[e^{\frac{-z^2}{4}}K_0(\frac{z^2}{4})]$ where K_0 is the order zero modified Bessel function of the second kind, which was derived in [8]. The theory that most normal breath sounds have a Gaussian distribution after variance normalization turns out to be true.

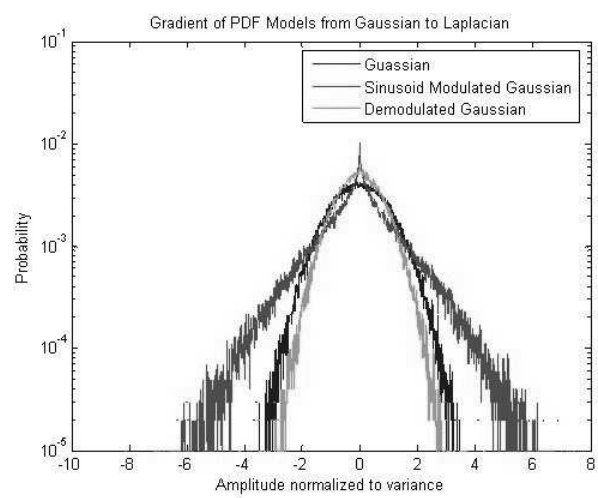

Figure 3: Distribution of a Gaussian signal before modulate, after modulation and after variance normalization

Snore sounds have been modeled based on the physical explanation of the sound produced. Snore sounds are produced as air passes through a partially collapsed pharynx.

This causes the loose tissue to slap against each other. Because of the relatively constant pressure the slapping occurs at a regular rate for the given pressure and tension of the pharynx. Each slap has a drum beat effect which appears as a exponentially decaying sinusoid with equation $s_n(t)=e^{\frac{-t}{b}}\cdot\sin(\omega t)$. After the variance normalization technique, the distribution of snores was observed to be approximately a Laplacian distribution with distribution $f_X(x)=\frac{1}{2b}e^{\frac{-|x|}{b}}$.

Given that the variance normalized distribution of a normal breath is Gaussian and for a snore is approximately Laplacian, a simple classifier can be built. In this case eleven distribution models were built with equation

$$f_X(x,p)=p\frac{1}{\sqrt{(2\mathrm{pi}\,\sigma^2)}}e^{\frac{-x^2}{2\sigma^2}}+(1-p)\frac{1}{2b}e^{\frac{-|x|}{b}}$$

where p is the probability of the distribution being Gaussian which was varied from 0 to 1 in increments of 0.1. A typical snore and normal breath are shown in figure 4 against the eleven models.

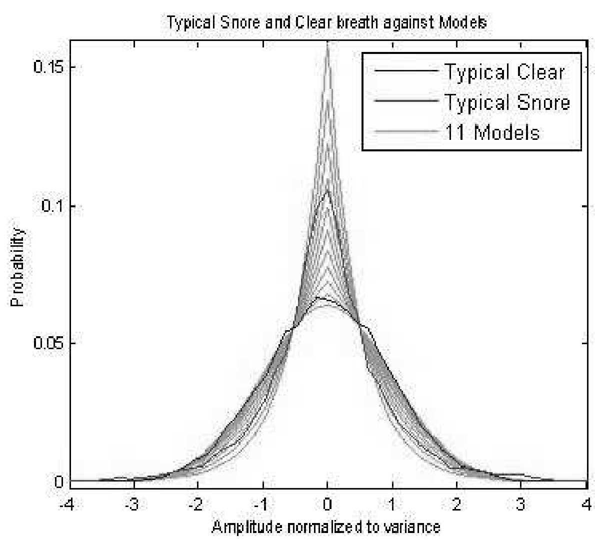

Figure 4: Distribution of a typical snore and normal breath against 11 Gaussian-Laplacian mixture models

An additional method for classifying breaths is to measure the time between each snore slap. The actual slapping sound is so low in frequency that it fails to register on an fft. If the sound is a snore the repeated impulse will have an observable spike in certain time intervals. To find the interval between each impulse a simple method was used which tracks each peak and exponentially decays when the audio signal dips below that of the tracking signal as shown in figure 5. The time between each tracked impulse is then stored for post processing. The constant in the exponential decay can be either heuristically found or tuned to a desired repetition rate

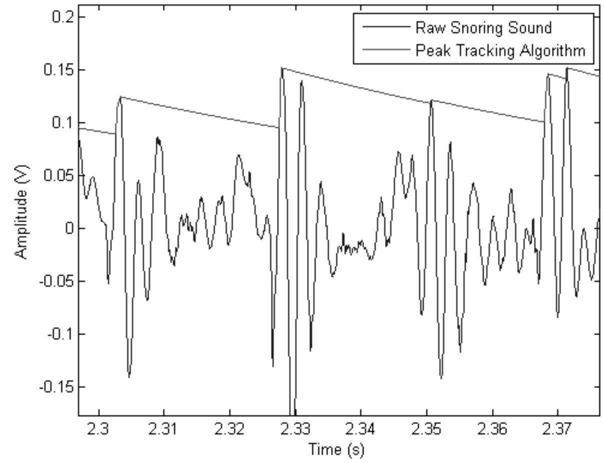

Figure 5: Raw snore sound with peak tracking algorithm

corresponding to the time constant of the exponential decay. In the case of a normal breath, the time between pulses should be small and erratic without a fundamental time period. This as a classifier simply compares the number of repetitions in a given repetition rate window. The optimal threshold can be derived experimentally.

IV. Results

Three ten minute audio clips were automatically segmented and then classified by ear to create a standard to classify against. The normalized histogram of each sound was then compared to the eleven models by summing the difference between each histogram and the distributions. The distribution with the smallest amount of difference was chosen as the best match. Figure 6 shows that the histogram classifier performed quite well with an error rate of about 0.0724.

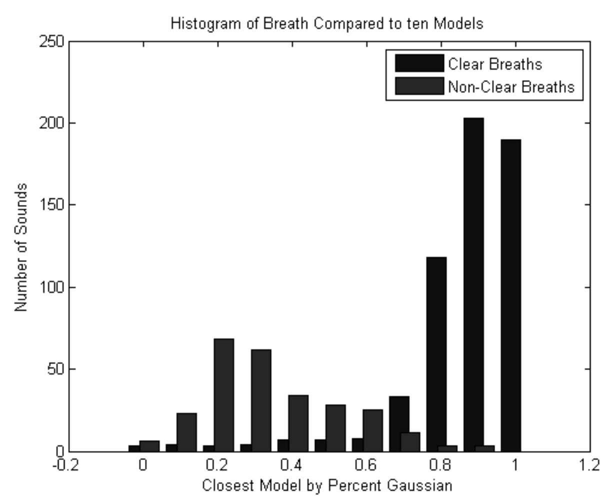

Figure 6: Histogram of classified snores vs. normal breaths when classified by histogram matching method

The ten minute sections were also classified using the repetition rate comparator and the results are shown in figure 7. This feature does not classify the set nearly as well with an error rate of about 0.169.

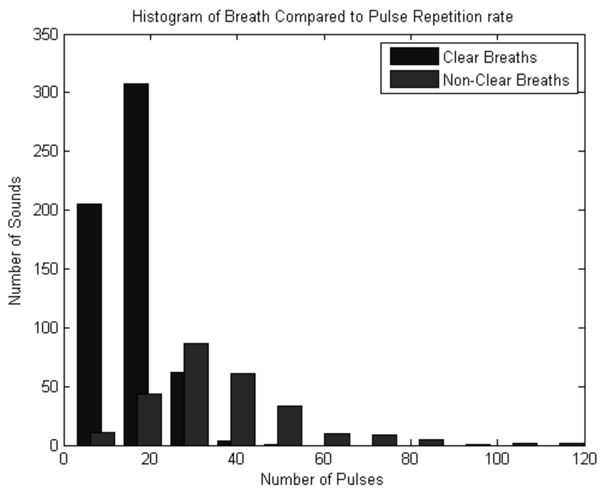

Figure 7: Histogram of pulse repetition classification of snores vs. clear breaths.

If both classifiers are used at the same time an improvement over using a single classifier emerges as shown in figure 8. A simple linear decision boundary can be seen and gives a probability of error of about 0.0483 with equal probability that a snore or a normal breath will be misclassified. Extraordinarily the pulse repetition rate dramatically improves the histogram classifier even though it individually does a very poor job.

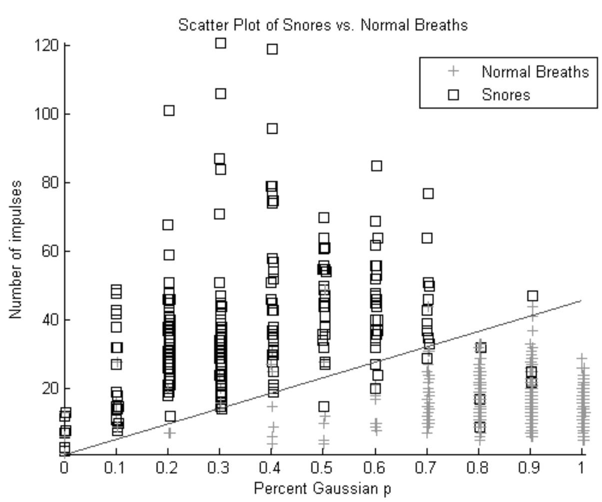

Figure 8: Dual feature classifier using histogram classification and pulse repitition classification

V. Conclusions

The histogram classifier is very simple and yet it does a very good job at classifying the snore vs. the normal breath sound. It may be robust enough on its own to be able to classify the breath sounds and be successful. The histogram method however throws away any useful data from the time domain. Perhaps the most common error that occurs is if there is a snap sound in a normal breath which can drastically change the distribution of the sound. It seems necessary to use some kind of time domain classifier in this case.

The time delay classifier by itself does not perform well. The time delay classifier uses very little of the data at all just as the histogram classifier does. The use of this classifier with the histogram classifier improves the decision error dramatically.

The combination of the two classifiers amplifies the accuracy of each of them making it less likely to error. The additional use of such technologies as Support Vector Machines [9] or Neural Networks [10] may further improve the reliability of these features to classify these sounds.

VI. Future Work

Although every obstruction sound could be considered to be a snore, not every snore is an obstruction sound. Rules will have to be made to classify periods of apnea as obstruction or depression based on the classified sounds. Since there is no gold standard to determine the difference between obstruction and depression, a method for classifying periods of apnea must be created. The method will be based on the physiology of the subject. During periods of normal breathing the RIP measurement correlates highly with the flow volume measurement. If the airway becomes occluded but the respiratory muscles continue to labor the effort can be seen on the RIP measurement but not on the flow measurement as

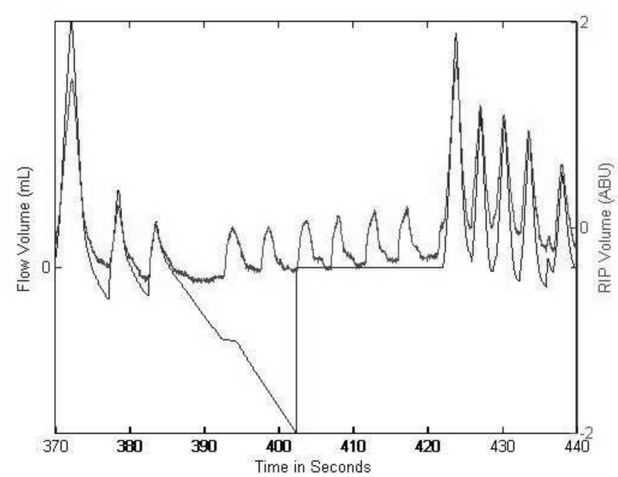

Figure 9: Respiratory obstruction. RIP bands show muscular movement and Flow volume zeros.

shown in figure 9. Similar to normal breathing is respiratory depression which can be identified by decreasing breath volumes until apnea occurs as shown in figure 10. Periods of apnea defined as breath volumes less than 200 ml for longer than 15 seconds will be manually classified as obstruction or depression.

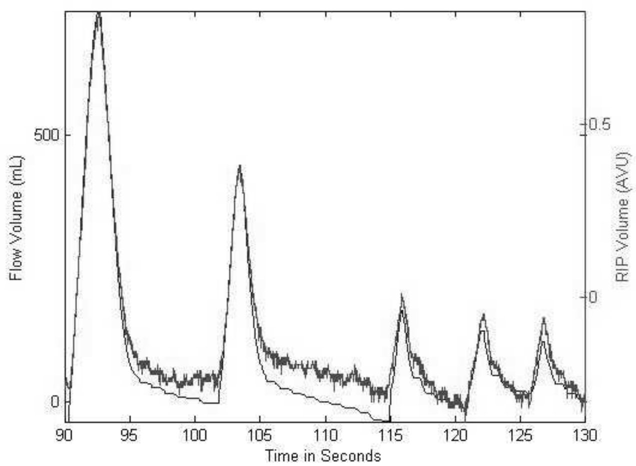

Figure 10: RIP and volume measurements as the subject descends to respiratory depression.

Next each period of apnea will need to be classified acoustically by looking at the sounds produced and classified prior to the apnea. The classifications will be compared and the ability of this algorithm to determine respiratory obstruction vs. respiratory depression will be determined

Acknowledgment

B.H Author thanks Dr. Dwayne Westenskow PhD, Douglas Christensen PhD, Joseph Orr Phd, Neil Cotter PhD, Neal Patwari PhD, and Ken Johnson MD for their continued support and help in collecting and analyzing the data.

References

[1] B. Hill, *Acoustic Breath Detection and Classification Modeling Respiratory Events*, 2008 Rocky Mountan NASA Space Grant Consortium.

[2] D. B. MacLeod, L. I. Cortinez, J. C. Keifer, D. Cameron, D. R. Wright, W. D. White, E. W. Moretti, L. R. Radulescu, and J. Somma. *The desaturation response time of finger pulse oximeters during mild hypothermia*. Anaesthesia, 60(1):65–71, Jan 2005.

[3] H. L. Watson, D. A. Poole, and M. A. Sackner. *Accuracy of respiratory inductive plethysmographic crosssectional areas*. J Appl Physiol, 65(1):306–308, Jul 1988.

[4] R. T. Brouillette, A. S. Morrow, D. E. Weese-Mayer, and C. E. Hunt. *Comparison of respiratory inductiveplethysmography and thoracic impedance for apnea monitoring*. J Pediatr, 111(3):377–383, Sep 1987

[5] D. W. Hudgel. The role of upper airway anatomy and physiology in obstructive sleep apnea. Clin Chest Med, 13(3):383–398, Sep 1992.

[6] Woodson BT Fairbanks DNF, Mckelson SA. *Snoring and Obstructive Sleep Apnea*. Lippincott Williams &Wilkins, 2002.

[7] M. Kompis, E.W. Russi, *Computer-based lung sound simulation*. Medical & Biological Engineering and Computing, May 1997, 35 321-238.

[8] K.V.Lever G.D. Cain and A.Yardim. *Probability density functions of amplitude-modulated random signals*.IEEE Electronics Letters, 34:1560–1561, 1998.

[9] G. Guo, S. Li, *Content-Based Audio Classification and Retrieval by Support Vector Machines*, IEEE Transactions on Neural Networks Vol 14 No. 1 January 2003.

[10] R.P. Lippman, *An Introduction to Computing with Neural Nets.* IEEE ASSP April 1987.