International Journal of Research in Engineering, Science and Management
Volume-2, Issue-10, October-2019
www.ijresm.com | ISSN (Online): 2581-5792

Identifying Wheezing Disorder in Infant Cry Signal

Sameena Bano[1], Mohammed Kaleem[2], Salma Jabeen[3]

[1]*Assistant Professor, Dept. of Computer Science and Engg., Ghousia College of Engg., Ramanagaram, India*
[2]*Lecturer, Department of Computer Science and Engineering, Middle East College, Muscut, Oman*
[3]*Associate Professor, Dept. of Computer Science and Engg., Ghousia College of Engg., Ramanagaram, India*

***Abstract*: This paper describes basic approach to identify wheezing disorder in the infant cry signal. In this work the infant cry signal of ages one day to two years old is used. In particular infants with prolonged cough cases wheeze which may lead to acute diseases like pneumonia. The infant cry signal is segmented by using Pitch frequency and Short-time energy and the features of segments are extracted using MFC (mel-frequency cepstrum) coefficients over MATLAB. KNN classifier is used to classify whether the infant cry signal has wheezing disorder. Percentages of results obtained are 80%, 70% and 90% based on variation of input samples. Identifying wheezing disorder in infant supports physician to treat infant in early stage.**

***Keywords*: Infant cries; Pitch frequency; Short-time energy; MFCC.**

1. Introduction

Crying is the first sound the baby makes when he enters the world, which is a very positive sign of a new healthy life. Infants cry for the similar reason that adults talk, lets other to know about their needs or problems. Since crying is all a baby can do to convey any discomfort, it seems that this multi-model signal carries a lot of information. In earlier studies of the infant cry analysis, the structure of infant crying was analyzed to describe the diseases [1–3]. The concealed information in normal cry signal could be used to classify the infant present condition. This approach analyze infant cry signal and classify healthy infant cry signal to help mother to know the baby need and physician to early treatment.

In this paper we study infant's cry signal to identify wheezing disorder. As there exist a large number of approaches to do the modeling and the classification tasks. We will focus on Pitch frequency for segmentation and Short-time energy for feature extraction and MFCC is calculated for the extracted features, Statistical properties are calculated for the MFCC and k-nn classifier is used to classify the cry signal, which are the most successful classifiers in use for audio data when their temporal structure is not important [4].

2. Methods

A. Data Acquisition

The cry signals used in this paper were obtained, by using diagnosis table and a laptop which is connected to a microphone [5] as shown in the figure1.

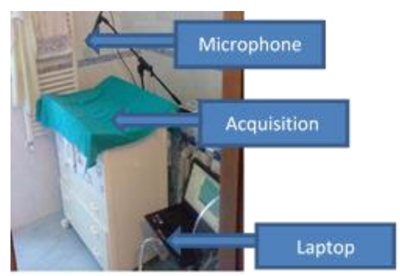

Fig. 1. Acquisition system [5]

Acquisition system was designed for being used in the hospital, minimizing the discomfort for the involved subjects and the impact of the external environment on children habits. Hence, the basic requirement is the ease in transporting and assembling the system.

The signal is recorded for 20sec, in hospital environment. Databases of 100 samples are used in this study of which 50 samples are normal infant and 50 with wheezing disorder.
Figure2 represents the sample signal waveform, and Figure3 represent the corresponding spectrogram.

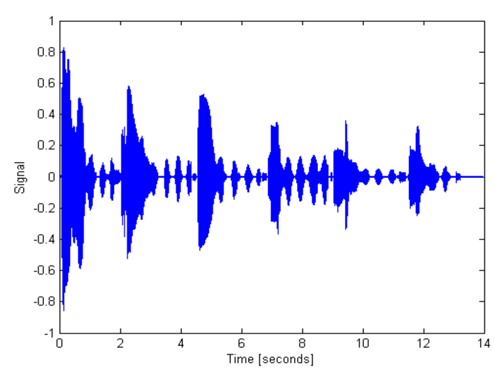

Fig. 2. Signal waveform

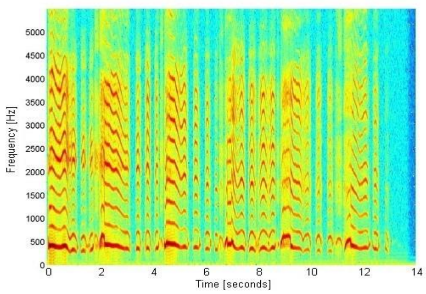

Fig. 3. Spectrogram

B. Feature Extraction

The following features are extracted from the baby cry signals:

Pitch frequency: Cry bursts are produced by quasi-periodic excitations of the vocal tract. The cry waveform is pseudo-periodic: at each point T, given in (1).

$$s(t) = s(t + T) \tag{1}$$

Typical pitch periods are 1.5 - 4ms

Pitch detection is based on Complex Cepstrum Fundamental Frequency Estimation.

Algorithm: Pitch frequency estimation based on complex cepstrum.

Step 1: Load the normal infant cry signal

Step 2: The sampling frequency is 44100 Hz.

Step 3: Extract segment of 0.1 to 0.25 seconds for analysis.

Step 4: Obtain the complex cepstrum for step3, given in (2).

$$ccep = ccep(x) \tag{2}$$

Step 5: Plot the cepstrum for times ranging from 2 to 10 msec corresponding to a frequency range of approximately 100 to 500 Hz.

Step6: Identify the peak in the cepstrum and find the corresponding frequency to the peak. Use this peak as fundamental frequency f0.

Short-time energy: The short-time energy (STE) of a signal s[n], using an analysis frame of N-samples length (beginning at n = N0), given in (3).

$$E[N0] = \frac{1}{N\sum s2[n]} \tag{3}$$

Harmonicity factor (HF): It is an estimation of the presence of harmonics in each frame analysis. HF is computed in (4), find the n highest peaks in DFT, their corresponding frequencies are f, f2,…, fn.

$$HF = \sum_{i=1}^{n} fi \bmod f0 \tag{4}$$

HF is zero for harmonic signals

Harmonic-to-Average Power Ratio (HAPR):

HAPR is a basic spectral feature, which determines the ratio of the harmonic component power and the average spectral power.

Algorithm: HAPR

Step 1: Identifying the highest peaks around the HF in the DFT magnitude.

Step 2: Find the power component around the n^{th} harmonic (5).

$$Pcomp = |S(2\pi nf0, t)^2| \tag{5}$$

Step 3: Calculating the average spectral power (6).

$$ASPs(t) = \frac{1}{N}\sum_{K=0}^{N}|S(Wk, t)^2| \tag{6}$$

Step 4: Finally HAPR is computed (7).

$$HAPR(t, N) = \frac{1}{N}\sum_{n=2}^{N} 10\log 10 \frac{|S(2\pi nf0, t)^2|}{|Sx(t)^2|} \tag{7}$$

Mel-Frequency Cepstrum Coefficients (MFCC): MFCC [11] provide a representation of short-term power spectrum of a signal. These coefficients are obtained by multiplying the short-time Fourier Transform (STFT) of each analysis frame by a series of M triangularly-shaped ideal band-pass filters, with their central frequencies and widths arranged according to a mel-frequency scale. The total spectral energy E[i] contained in each filter is computed and a Discrete Cosine Transform (DCT) is performed to obtain the MFCC sequence (8).

$$MFCC(L) = \frac{1}{M\sum logE[i]} \times \cos\left(\left(\frac{2\pi}{M}\right) \times \left(i + \frac{1}{2}\right) \times L\right) \tag{8}$$

L=1,2, … M-1

C. Classifier

k-nn (K-nearest neighbor) Classifier is used as it is the most successful classifiers used for audio data when their temporal structure is not important, K-nn Classification using an instance-based classifier can be a simple matter of locating the nearest neighbor in instance space and labeling the unknown instance with the same class label as that of the located (known) neighbor. This approach is often referred to as a nearest neighbor classifier.

Euclidean distance is used to get the Pair wise distance between two sets of observations like Neh and Eh and is repeated for all the remaining three sets and finally it classifies the category it fall.

The Euclidean distance between the n-dimensional vectors a and b is calculated (9).

$$Eucdist(x, y) = \sqrt{\sum_{i=1}^{n}(ai - bi)^2} \tag{9}$$

3. Results

The Table below shows the different values of Pitch for segmentation: three different cry signals.

Table 1
Values of different parameters

	cry1	cry2	cry3	Pitch
Correct	0.63	0.9	0.9	0.81
error	0.37	0.1	0.1	0.19
Correct	0.68	0.56	0.58	0.606667
error	0.32	0.44	0.42	0.393333

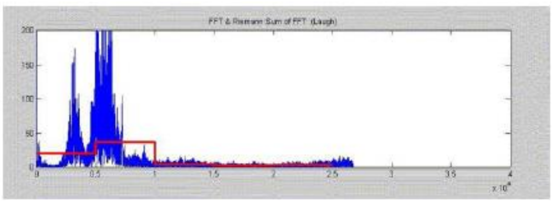

Fig. 4. Cry signal of normal infant

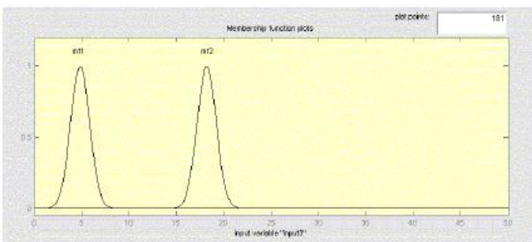

Fig. 5. Sample of segmentation of normal infant cry signal

Table 2
Sample data of Statistical property

Type of Cry	Mean	Max	Min	SD	PSD
Cry1	-2.2201	3.0907	-50.0873	4.9624	121.0349
Cry2	-2.1500	3.6382	-60.3298	7.0185	120.7044
Cry3	-2.4204	1.7076	-52.2498	5.5982	118.7385
Cry4	-.9026	1.7476	-6.1438	7.5617	86.3074
Cry5	-1.6563	1.4171	-53.8510	5.1450	71.2117

The signal is recorded for 20sec, in hospital environment. The database has 100 samples in which 50 samples are with wheezing disorder and 50 samples without wheezing disorder. 80 samples used for training and 20 samples used for testing. The percentage detection is shown in Table 3 and fig. 4.

Table 3
Percentage of detection of type of cry

Type of Cry	% of Correct detection	% of Wrong detection
Trail1	80	20
Trail 2	70	30
Trail 3	90	10

Trail 1 uses 50 samples of wheezing and 50 samples of normal infant without any disorder and the result is 80%. Trail 2 uses 50 samples of wheezing and 50 samples of infant with other pathological disorder like tummy pain, fever and the result is 70%. Trail 3 uses 50 samples of wheezing and 50 samples of infant without disorder like pinching baby, Giving Injection or by doctor diagnosis and the result is 90%.

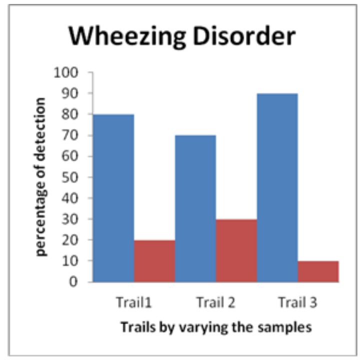

Fig. 6. Percentage of detection of type of cry

4. Conclusions

We present an efficient study of analyzing infant cry signals for wheezing disorder. This study is based on a database of 100 samples in which 50 samples are with wheezing disorder and 50 samples without wheezing disorder. 80 samples used for training and 20 samples used for testing. Percentages of results are 80%, 70%, and 90% respectively.

The multiple parameter analysis is aimed at providing a classifier with very high rate, while keeping a low rate negatives. The advantage of the study is its simplicity. It is based on a small number of features, which are relatively simple to implement. This study helps to decode the baby cry which supports the mother's built-in intuition about knowing and responding to their baby's needs, which empower every mother & father to feel more relaxed, more capable, more confident in caring for their new baby. This also help physician to treat infant early. In future research we plan to extend the evaluation of the proposed study, using a large set of signals.

References

[1] T. van Waterschoot and M. Moonen. Fifty years of acoustic feedback control: State of the art and future challenges. Proceedings of the IEEE, 99(2):288 –327, Feb. 2011.
[2] J. Benson, Social and Emotional Development in Infancy and Early Childhood, Elsevier, 2009.
[3] O. Wasz-Hockert, The Infant Cry: A Spectrographic and Auditory Analysis, Lippincott, Philadelphia, Pa, USA, 1968.
[4] A. Divakaran, Multimedia Content Analysis: Theory and Applications (Signals and Communication Technology), Springer, 2009.
[5] S. Orlandi, L. Bocchi, C. Manfredi, M. Puopolo, A. Guzzetta S. Vicari and M.L. Scattoni, "Study of cry patterns in infants at high risk for AUTISM" at Models and analysis of vocal emissions for biomedical applications at 7th international workshop on August 25-27, 2011 firenze, Italy.
[6] J.O. Garcia and C.A. Reyes Garcia. Mel-frequency cepstrum coefficients extraction from infant cry for classification of normal and pathological cry with feed-forward neural networks. In Proceedings of the International Joint Conference on Neural Networks, volume 4, pages 3140–3145, July 2003.
[7] G. Varallyay. The melody of crying. International Journal of Pediatric torhinolaryngology,71(11):1699 – 1708, 2007.
[8] P. Ruvolo and J. Movellan. Automatic cry detection in early childhood education settings. In 7th IEEE International Conference on Development and Learning (ICDL), pages 204–208, Aug. 2008.
[9] A. Messaoud and C. Tadj. A cry-based babies identification system. In Proceedings of the 4th international conference on Image and signal processing, ICISP'10, pages 192–199, 2010.