Decoding Baby Talk: A Novel Approach for Normal Infant Cry Signal Classification

Mrs. Sameena Bano, Asst. professor,
Dept. of Computer Science and Engg., Ghousia College of Engineering, Ramanagara, Affliated to VTU,India .
Sameena.parveez@yahoo.com

Dr. K.M. RaviKumar, professor & Head,
Dept. of Electronics & Communication Engg., S. J. C. Institute of Technology, Chickballapur,
Affliated to VTU ,India. kmravikumar@rediffmail.com

Abstract— This paper describes a novel approach to identify a baby physiological state and its needs. In this work normal infant cry signal of ages 1day to six months old is used. In particular there are fixed cry attributes for a healthy infant cry, which can be classified into five groups such as: Neh, Eh, Owh, Eairh and Heh. The infant cry signal is segmented by using Pitch frequency and the features like Short-time energy, Harmonicity Factor (HF) and Harmonic-to-Average Power (HAPR) are extracted and MFC (mel-frequency cepstrum) coefficients is computed over MATLAB. KNN classifier using Pitch, Short-time energy, Harmonicity Factor (HF) and Harmonic-to-Average Power (HAPR) are used to classify the normal infant cry signal. Percentages of results obtained are Neh 80%, Eh 90%, Owh 80%, Eairh 90%, and Heh 90% respectively. Decoding baby talk supports the mother's built-in intuition about knowing and responding to their baby's needs, and physician to treat infant early.

Keywords- Infant cry; Pitch frequency; Harmonicity Factor (HF); Harmonic-to-Average Power (HAPR); MFC

I. 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 crying signal conditions such as whether the baby cry is for Neh: Hungry, Eh: Pain/burp me (Pinching/Drawing blood), Owh: Sleepy, Eairh: Pain (Ear pain/tummy pain), or Heh: Discomfort. Fig.1 represents this classification.

Fig.1 Decoding Baby Talk [15]

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, HF and HAPR 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].

II. 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 figure2.

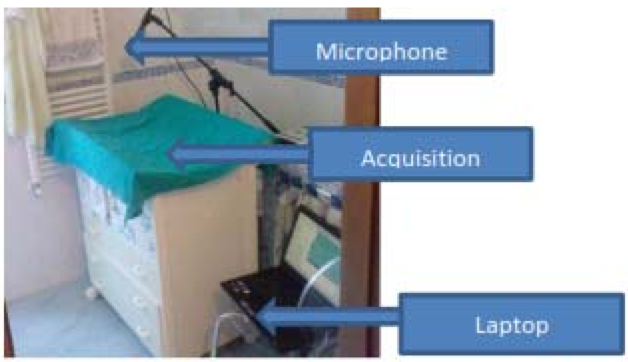

Fig.2 Acquisition system [5]

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. There are five different databases each of 50 samples. 1. Neh database, 2. Eh database, 3.Owh database, 4.Eairh database, 5.Heh database.

1. Neh database: Signal is recorded purposely delaying the feeding time.
2. Eh database: Signal is recorded while collecting blood for diagnosis or by pinching baby.
3. Owh database: Signal is recoded when the baby cry before sleeping.
4. Eairh database: Signal is recorded during tummy pain or ear pain for infant.
5. Heh database: Signal is recorded when pediatric perform health evolution.

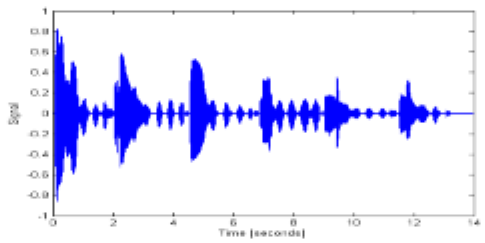

Fig.3 Signal waveform

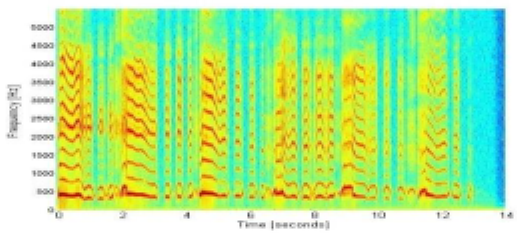

Fig.4 Spectrogram

B. *Feature Extraction*

The following features are extracted from the baby cry signals: 1. 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).

$$S(t) = s(t+T) \qquad (1)$$

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.

Step1: Load the normal infant cry signal
Step2: The sampling frequency is 44100 Hz.
Step3: Extract segment of 0.1 to 0.25 seconds for analysis.
Step4: Obtain the complex cepstrum for step3, is given in (2).

$$Ccep = ccep(x) \qquad (2)$$

Step5: 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.

2. 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).

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

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 f1,f2,…, fn.

$$HF = \sum_{i=1}^{n} fi \bmod f0 \qquad (4)$$

HF is zero for harmonic signals

4. 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 nth harmonic (5).

5. 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).

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

L=1,2, … M-1

$$Pcomp = |S(2\pi n f0, t)^2| \quad (5)$$

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

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

Step 4: Finally HAPR is computed (7).

$$HAPR(t,N) = \frac{1}{N}\sum_{n=2}^{N} 10\log 10\frac{|S(2\pi n f0,t)^2|}{|Sx(t)^2|} \quad (7)$$

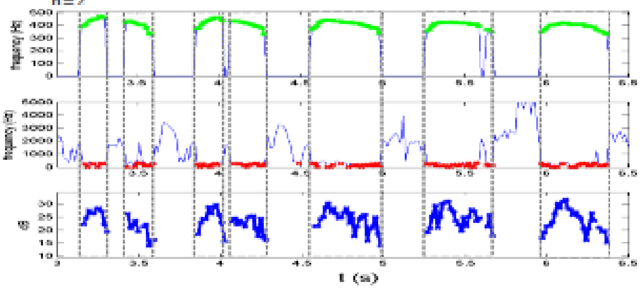

Fig.5 Pitch, Harmonicity Factor and HAPR of Normal Infant Cry Signal for analysis and detection of pitch

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

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

III. RESULTS

Fig. 6 and 7 represent the normal infant cry signal and sample segmentation of normal infant cry signal respectively. Table I below shows the different values of Pitch for segmentation: three different cry signals.

TABLE I. 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

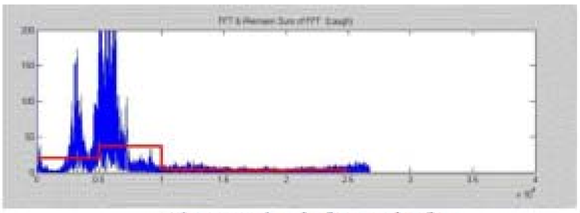

Fig.6 Cry signal of Normal Infant

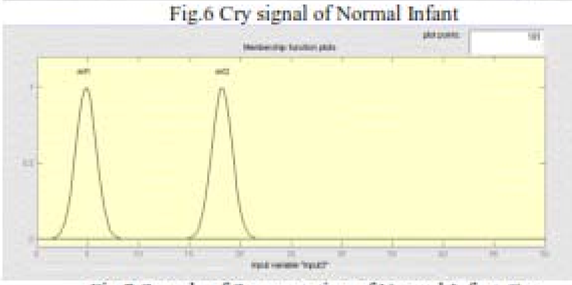

Fig.7 Sample of Segmentation of Normal Infant Cry Signal

TABLE II. Sample data of Statistical property

Type of Cry	Mean	Max	Min	SD	PSD
Neh	-2.2201	3.0907	-50.0873	4.9624	121.0349
Eh	-2.1500	3.6382	-60.3298	7.0185	120.7044
Owh	-2.4204	1.7076	-52.2498	5.5982	118.7385
Eairh	-.9026	1.7476	-6.1438	7.5617	86.3074
Heh	-1.6563	1.4171	-53.8510	5.1450	71.2117

Since the MFCC gives us a matrix of values, the data is compress by using statistical properties as shown in Table II. The signal is recorded for 20sec, in hospital environment. There are five different databases each of 50 samples. Neh database, Eh database, Owh database, Eairh database, Heh database, 40 samples used for training and 10 samples used for testing. The percentage detection is shown in Table.3 and fig.8.

TABLE III. Percentage of detection of type of cry

Type of Cry	% of Correct detection	% of Wrong detection
Neh	80	20
Eh	90	10
Owh	80	20
Eairh	90	10
Heh	90	10

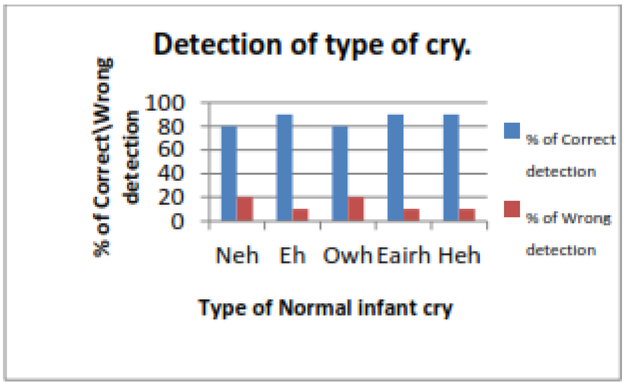

Fig.8 Percentage of detection of type of cry

IV. CONCLUSIONS

We present an efficient study of analyzing infant cry signals. The normal infant cry signal is classified into five types Neh: Hungry, Eh: Pain/burp me (Pinching/Drawing blood), Owh: Sleepy, Eairh: Pain (Ear pain/tummy pain) and Heh: Discomfort. This study is based on five different databases such as, Neh database, Eh database, Owh database, Eairh database, Heh database. Each has 50 samples of data 40 samples used for training and 10 samples used for testing. Percentages of results are Neh 80%, Eh 90%, Owh 80%, Eairh 90%, and Heh 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 help 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 a future research we plan to extend the evaluation of the proposed study, using a large set of signals.

References

[1] O. Wasz-Hockert, K. Michelsson, and J. Lind, Twenty-Five Years of Scandinavian Cry Research, New York, NY, USA, 1985. [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 Otorhinolaryngology, 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 Develop- ment 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.
[10] G. Varallyay. Crysamples, http://sirkan.iit.bme.hu/var- allyay/crysamples.htm, 2009.
[11] A. M. Noll. Cepstrum pitch determination. The Journal of the Acoustical Society of America, 41(2):293–309, 1967.
[12] D. Gerhard. Pitch extraction and fundamental frequency: History and current techniques. Technical report, University of Regina, Canada, 2003.
[13] A. Klautau. The MFCC. Technical report, Signal Processing Lab, UFPA, Brasil, 2005.
[14] 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.
[15] Dunstan Baby Ears Web Site Baby Ears by DBL Support http://www.dunstandbabay.com 2007.