Decoding Baby Talk: Basic Approach for Normal Classification of Infant Cry Signal

Sameena Bano,
Asst. Professor,
Dept. of Computer Science and Engg.,
Ghousia College of Engineering,
Ramanagara, India .

K.M. RaviKumar, Ph.D
Professor & Head,
Dept. of Electronics & Communication Engg.,
S. J. C. Institute of Technology,
Chickballapur-562101, India

ABSTRACT
The analysis of infant cry has become more prevalent due to advances in areas such as digital signal processing, pattern recognition and soft computing. The analysis of infant cry has changed the diagnostic ability of physicians to correctly diagnose new-born. This work presents an approach to decode baby talk by classifying infant cry signal. We use normal infant cry signal of ages 1day to six months old. 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 features are extracted using MFC (mel-frequency cepstrum) coefficients over MATLAB. Statistical properties are calculated for the extracted features of MFCC and KNN classifier is used to classify the cry signal. KNN is the most successful classifiers used for audio data when their temporal structure is not important. This study is based on five different databases such as, Neh, Eh, Owh, Eairh, and Heh databases. 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. 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.

General Terms
Digital signal processing, pattern recognition, soft computing

Keywords
Infant cry, Pitch frequency, KNN, 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 same purpose that adults talk, that is, to let others identify about their needs or problems. Since crying is the one mean of communication of baby to express any discomfort. In earlier studies of the infant cry analysis, the structure of infant crying was analyzed to describe the diseases [1–3]. The information in normal cry signal could be used to classify the infant present condition. This study analyze infant cry signal and classify healthy infant cry signal to help mother to know the baby need and physician to early treatment.

This paper analyze 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), or Heh(Discomfort). Fig.1 represents this classification of normal infant cry signal.

Fig.1 Decoding Baby Talk [15]

As there exist a large number of approaches to do the modeling and the classification tasks. The focus is on MFCC to extract the features in infant cry signal, Statistical properties are calculated for the extracted features of MFCC and KNN 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
2.1 Data Acquisition
The cry signals used in this paper were obtained, by using diagnosis table, laptop which is connected to a microphone. Acquisition system is considered for being used in the hospital, minimizing the discomfort of the involved subjects and the impact of the external environment on children habits. Hence, the basic requirement is the comfort 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

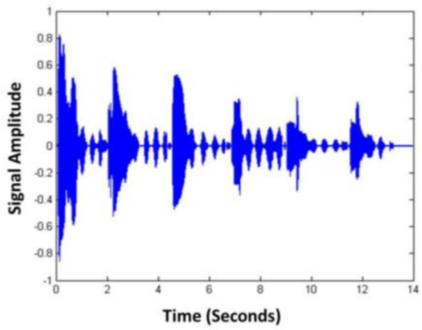

Fig.2 Signal waveform

2.2 Feature Extraction

The following features were extracted from the baby cry signals:

1) Pitch frequency: The fundamental frequency f0 is important for classification purposes. The pitch detection algorithm is based on a combination of the Cepstrum method [9] for coarse-pitch period detection. The results are distinguished by using cross-correlation method [10].

2) Short-time energy: The short-time energy (STE) of a signal x [n], using an analysis frame of N -samples length (beginning at n = N0), is defined as:

$$E[N0] = \frac{1}{N\sum x2[n]}$$

3) 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 the 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:

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

L=1,2, … M-1

2.3 Classifier

KNN (K-nearest neighbor) Classifier is used as it is the most successful classifiers used for audio data when their temporal structure is not important. KNN_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.

3. RESULTS

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	0.53	0.8	0.8	0.71
error	0.47	0.2	0.2	0.29
Correct	0.58	0.66	0.48	0.51
error	0.42	0.54	0.52	0.49

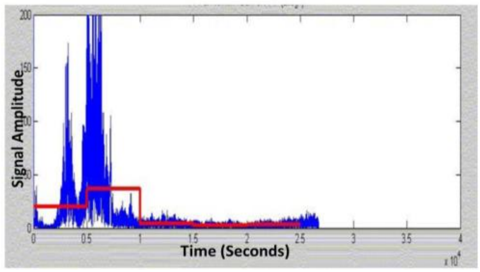

Fig.4 Cry signal of Normal Infant

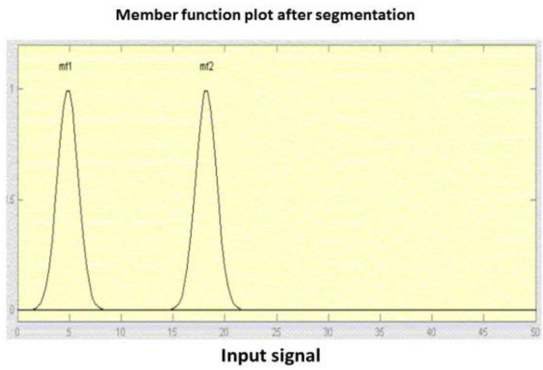

Fig.5 Sample of Segmentation of Normal Infant Cry Signal

Table 2. Sample data of Statistical properties

Type of Cry	Mean	Max	Min	SD	PSD
Neh	-3.2201	3.0907	-4.0873	3.9624	110.0349
Eh	-2.1500	3.6382	-6.3298	7.0185	110.7044
Owh	-1.4204	1.7076	-2.2498	5.5982	89.7385
Eairh	-.5026	1.7476	-4.1438	5.5617	76.3074
Heh	-1.6563	1.4171	-3.8510	4.1450	51.2117

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.6.

Table 3. Percentage of detection of type of cry

Type of Cry	% of Correct detection	% of Wrong detection
Neh	60	40
Eh	70	30
Owh	60	40
Eairh	70	30
Heh	70	30

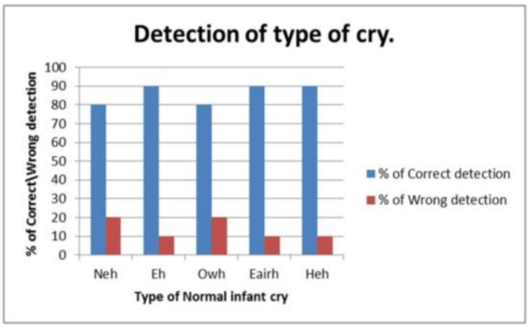

Fig.6 Percentage of detection of type of cry

4. CONCLUSIONS

The study presents analysis of 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 and Heh: Discomfort. The 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 60%, Eh 70%, Owh 60%, Eairh 70%, and Heh 70% respectively.

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.

Future research can extend the evaluation of the proposed study using a large set of signals.

5. 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/ varallyay/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.