Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 4, Issue. 4, April 2015, pg.413 – 420

RESEARCH ARTICLE

Analysis of Infant Cry through Weighted Linear Prediction Cepstral Coefficient and Probabilistic Neural Network

Roopa S M[1], Smt. Anitha G[2], Dr. Mohamed Rafi[3]

[1]Department of Computer Science, Visvesvaraya Technological University, India

[2]Department of Computer Science, Visvesvaraya Technological University, India

[3]Department of Computer Science, Visvesvaraya Technological University, India

[1] roopasmeharwade@gmail.com, [2] anithapoojar@yahoo.com, [3] rafimohamed17@gmail.com

Abstract— Acoustic analysis of infant cry signals can be a work in the area of automatic detection of pathological status of an infant. This work investigates the application for linear prediction cepstral coefficients (LPCCs) to provide the robust representation of infant cry signals. Three classes of infant cry signals were considered such as normal cry signals, cry signals from deaf babies and babies with asphyxia. A Probabilistic Neural Network (PNN) is suggested to classify the infant cry signals into normal and pathological cries. PNN is trained with different spread factors or smoothing parameter to obtain better classification accuracy. These methods are used to help medical professionals for diagnosing pathological status of an infant from cry signals.

Keywords— Acoustic Analysis, Infant Cry, Weighted LPCCs, Wireless Sensor Network

I. Introduction

Pathological status of babies using the conventional methods is commonly detected several months and sometimes even years after the infant is born. It is necessary to detect the pathological status earlier to avoid opportune treatments and therapies. Crying is the only way of communication for an infant. From the cry, a trained professional can understand the physical or psychological status of the baby. Infants cry are due to some possible reasons such as, hunger, pain, sleepiness, discomfort, feeling too hot or too cold, and too much noise or light. Acoustic analysis of infant cry signal is a non-invasive and has been proven to be a tool for the detection of certain pathological conditions. In the recent years, simple techniques have been proposed for analyzing the infant cry through linear prediction coding, Mel frequency cepstral coefficients and pitch information. In classical LPCC coefficients, the low-order cepstral coefficients are sensitive to overall spectral slope and the high- order cepstral coefficients are sensitive to noise and other forms of noise like variability. Hence, a standard technique is to weigh the cepstral coefficients by a tapered window so as to minimize these sensitivities. This paper presents the development of an intelligent classification system for classifying normal and pathological cries using weighted LPCCs and probabilistic neural network.

INFANT CRY

Baby cries are the result of the interaction between control of different areas in the brain, respiratory control and vocal fold vibrations. At early stage, it is believed that a cry is the result of respiratory action and the effect of air going through a pipe, causing the vocal folds to vibrate, resulting in a cry bout, this is the vocalization produced during one expiration. The more the neural system matures, the more laryngeal control can be exerted resulting in manipulation and modulation of the cry signal, but observations on maturational aspects of the cry signal are limited and conflicting. Decoupling is referred to as the absence of coupling between fundamental frequency variation and intensity contour during phonetic crying. In this cry model, cry production is controlled by a three level processor structure, in which the lower level can be separated into the sub glottal, glottal and supraglottal cry production areas. In the current study, we want to assess the relation between the energy envelope of the cry bout and frequency variation within the same bout using the method recently introduced by Huang, called Empirical Mode Decomposition.

DATABASE DISCRIPTION

The newborns database consists of audio signals of babies .We have collected 10 samples from each type of cry signals of Baby . The Chillanto Data Base is a property of the InstitutoNacional de AstrofisicaOptica y Electronica—CONACYT, Mexico.The database acquisition of newborns took one year to complete and thus it has minor variations in lightning conditions due to changes in weather conditions. The database of each subject is prepared in two different sessions. In the first session data is collected within four hours of birth of a child and the data of second session depends on the type of birth. For example, if there is a case of normal birth then data is collected after 20 hours otherwise in scissoring cases data is collected after 70 hours of the birth. In this data collection, the time for data acquisition is set according to the period that a newborn stays in the hospital after his/her birth.

II. Literature Survey

EXISTING SYSTEM

***Reyes-Galaviz, O, Cano-Ortiz, S, and Reyes-Garca, C*:** We are proposed Evolutionary-neural system to classify infant cry units for pathologies identification in recently born babies; this work presents an infant cry automatic recognizer development, with the objective of classifying two kinds of infant cries, normal and pathological, from recently born babies. Extraction of acoustic features is used such as MFCC, obtained from Infant Cry Units sound waves, and a genetic feature selection system combined with a feed forward input delay neural network, trained by adaptive learning rate back-propagation. For the experiments, recordings from Cuban and Mexican babies are used, classifying normal and pathological cry in three different experiments; Cuban babies, Mexican Babies, and Cuban & Mexican babies. In this paper the whole process is described; in which the acoustic features extraction is included, the hybrid system design, implementation, training and testing.

Advantages:

It is also shown a comparison between a simple traditional feed-forward neural network and another complemented with the proposed genetic feature selection system, to reduce the feature input vectors.

Disadvantages:

The genetic feature selection system combined with a feed forward input delay neural network, trained by adaptive learning rate back-propagation.

***Reyes-Galaviz. O, Verduzco. A, Arch-Tirado, E, and Reyes-García, C*:** We are proposed Analysis of an infant cry recognizer for the early identification of pathologies. This work presents the development and analysis of an automatic recognizer of infant cry, with the objective of classifying three classes, normal, hypo acoustics and asphyxia. We use acoustic feature extraction techniques like MFCC, for the acoustic processing of the cry's sound wave, and a Feed Forward Input Delay neural network with training based on Gradient Descent with Adaptive Back-Propagation for classification. The complete infant cry database is represented by plain text vector files, which allows the files to be easily processed in any programming environment. The paper describes the design, implementation as well as experimentation processes, and the analysis of results of each type of experiment performed.

Advantages:

The use principal component analysis in order to reduce vector's size and to improve training time.

Disadvantages:

The feed forward Input Delay neural network with training based on Gradient Descent with Adaptive Back-Propagation for classification.

Var*allyay.G. Jr, Benyo. Z, Illenyi. A, Farkas. Z, and Kovacs, L*: We are proposed Acoustic analysis of the infant cry: classical and new methods, many researches related to the infant cry analysis intent to estimate the context and/or obtain objective information concerning the physical and emotional condition of newborns. Using several techniques in signal processing, peculiar acoustics features, such as the fundamental frequency and formants are classically analyzed. However, the findings reveal the existence of some contests with respect to the conclusions. In this article a specific phonologic program was used to analyze the cry signal, aiming to investigate the real significance of some classical frequency domain parameters. The results point out that just two among four studied parameters seem to contribute in the analysis of the cry signal context. Beside the significance of the two parameters in such analysis.

Advantages:

Using several techniques in signal processing, peculiar acoustics features, objective information concerning the physical and emotional condition of newborns.

Disadvantages:

The findings reveal the existence of some contests with respect to the conclusions. In this article a specific phonologic program was used.

***Orozco.J and García.C*:** We are proposed Detecting pathologies from infant cry applying scaled conjugate gradient neural networks; this work presents the development of an automatic recognition system of infant cry, with the objective to classify two types of cry: normal and pathological cry from deaf babies. In this study, we used acoustic characteristics obtained by the Linear Prediction technique and as a classifier a neural network that was trained with the scaled conjugate gradient algorithm. Preliminary results are shown, which, up to the moment, are very encouraging.

Advantages:

The development of an automatic recognition system of infant cry, the crying signal is analyzed to extract the more important features in the time domain.

Disadvantages:

The crying wave is filtered to eliminate irrelevant or undesirable information like noise, channel distortion, and other particular signal's characteristics. The data are reduced when removing repetitive component.

***García.J and García. C*:** We are proposed Acoustic features analysis for recognition of normal and hypoacustic infant cry based on neural networks, Acoustic analysis of infant cry signals has been proven to be an excellent tool in the area of automatic detection of pathological status of an infant. This paper investigates the application of parameter weighting for linear prediction cepstral coefficients to provide the robust representation of infant cry signals. Three classes of infant cry signals were considered such as normal cry signals, cry signals from deaf babies and babies with asphyxia. A Probabilistic Neural Network is suggested to classify the infant cry signals into normal and pathological cries. The experimental results demonstrate that the suggested features and classification algorithms give very promising classification.

Advantages:

PNN is trained with different spread factor or smoothing parameter to obtain better classification accuracy.

Disadvantages:

A disadvantage here can be considered the parameter weighting for linear prediction cepstral coefficients to provide the robust representation of infant cry signals.

PROPOSED SYSTEM

In the recent years, simple techniques have been proposed for analyzing the infant cry through linear prediction coding, Mel frequency cepstral coefficients and pitch information. In classical LPCC coefficients, the low-order cepstral coefficients are sensitive to overall spectral slope and the high- order cepstral coefficients are sensitive to noise and other forms of noise like variability. Hence, a standard technique is to weigh the cepstral coefficients by a tapered window so as to minimize these sensitivities. This paper presents the development of an intelligent classification system for classifying normal and pathological cries using weighted LPCCs and probabilistic neural network.

III. System Design

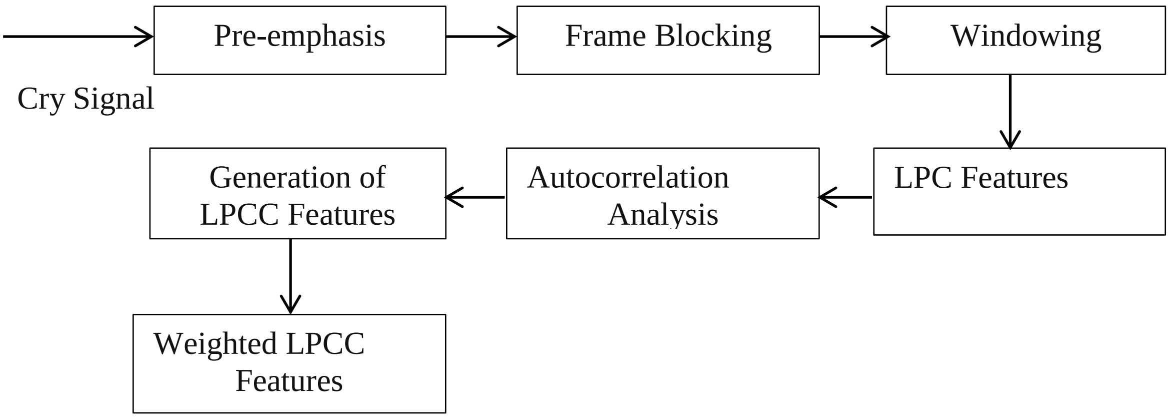


Fig.1. Block Diagram of the Feature Extraction

Fig.2. Block Diagram of the Feature Extraction and Classification Phase

A.FEATURE EXTRACTION

1. Pre-emphasis:

The basic block diagram of feature extraction process is first the cry signals are passed through a first order low pass filter. The purpose of this process is to spectrally flatten the signal and to make it less susceptible to finite precision effects later in the signal processing.

2. Frame Blocking:

The first order pre-emphasis filter is defined as, the pre emphasized cry signal is segmented into frames of N samples with an overlap of (1/3)*N samples. The different length of the frame is used in this work, such as 20 ms, 30 ms, 40 ms and 50 ms respectively.

3. Windowing:

The segmented cry signals are multiplied with window coefficients (Hamming) to avoid the discontinuities during segmentation and taper the signal at the beginning and the end of each frame.

4. LPC Features:

Linear Predictive Coding is one of the most powerful speech analysis techniques, and one of the most useful methods for encoding good quality speech at a low bit rate. It provides extremely accurate estimates of speech parameters, and is relatively efficient for computation. Based on these reasons, we are using LPC to represent the crying signals. Linear prediction is a mathematical operation where future values of a digital signal are estimated as a linear function of previous samples.

Each of the windowed signals is auto correlated according to the formula where p is the order of the LPC analysis and p is fixed as 8,10,12,14, and 16. The LPC analysis is performed to convert the autocorrelation coefficients into LPC coefficients. The LPC analysis is implemented using Durbin-Levinson recursive algorithm. The final solution for the LPC coefficients is given as order cepstral coefficients are sensitive to noise and other forms of noise like variability. Hence, a standard technique is to weigh the cepstral coefficients by a tapered window so as to minimize these sensitivities.

5. Autocorrelation Analysis:

The autocorrelation sequence of the signal is performed. Considering a sampling frequency 16 kHz, in order to achieve good time resolution, 256 points are required (16 milliseconds) which correspond to more than 3 pitch periods. The next step is multiplying the autocorrelation by a trapezoid window in the frequency range defined by fi, in order to restrict the search to a reasonable cry fundamental frequencies range. The result, peak picking is performed in order to search for a candidate F0, and a decision algorithm is applied in which the peak value is compared to a voiced threshold. A peak which exceeds the threshold value or a borderline peak with “voiced history” is defined as voiced. A peak which does not exceed the threshold value but has “voiced neighbors” is also defined as voiced.

6. Generation of LPCC Features:

The weighted LPCC features are extracted from every frame of a signal and formed as a feature matrix. The average of feature matrix is used to represent a cry signal. A feature database is created, after the computation of weighted LPCC features and they are used as input features for the classifiers to distinguish the cry signals as normal, deaf cries and asphyxia cries.

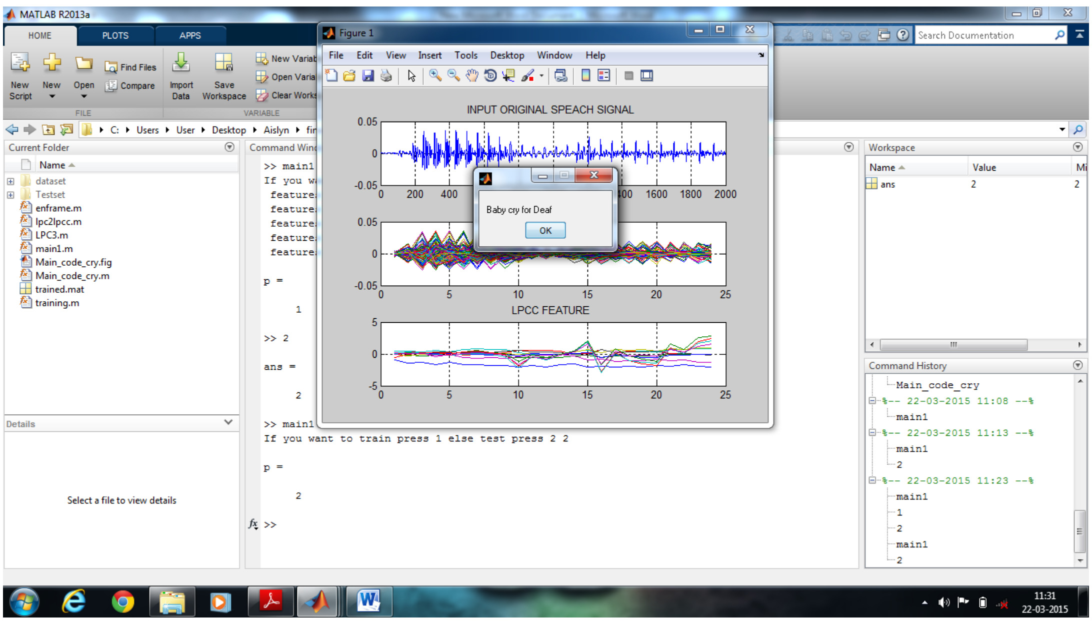
B. Classification Method:

Probabilistic neural network artificial neural networks are widely used in pattern recognition and classification problems by learning from examples. Different neural network paradigms are available for classifying the patterns. In this work, PNN structure is used for classifying normal and pathological cries. Donald F. Speech has proposed the probabilistic neural net based on Bayesian classification and classical estimators for probability density function. PNN comprises of four units, such as input units, pattern units, summation units and output units. All the units are fully interconnected and the pattern units are activated by exponential function, instead of sigmoidal activation function.

The pattern unit computes distances from the input vector to the training input vectors, when an input is presented, and produces a vector whose elements indicate how close the input is to a training input. The summation unit sums these contributions for each class of inputs and produces a net output which is a vector of probabilities. From the maximum of these probabilities, output units produce a 1 for that class and a 0 for the other classes using compete transfer function.

IV. Results

We have given input of deaf baby cry signal. Once input is given, we are training that to neural network.Here original speech signal of baby is given input, from that LPCC features are extracted, later these features are given as a input to the classifier which classifies the given signal as a deaf baby cry.



Final output

V. Conclusion

This paper presents the analysis of infant cry signals based on the weighted LPCC and PNN. The infant cry signals were segmented into samples of N frames. The weighted LPCC features were extracted from every frame of a signal and finally the average of the LPCCs were used as features. In this work, the frame length and order of LPC were varied and their effects on the performance of results were presented. Conventional validation was performed (one training set (70%) and one testing set (30%)), in order to test the reliability of the PNN classifier. The maximum classification accuracy of 99% was obtained for the frame length of 40 ms and for the LPC order 14. The classification results indicate that the suggested method could be used as a valuable tool for clinical diagnosis of the infant cry signals.

REFERENCES

[1] Reyes-Galaviz, O., Verduzco, A., Arch-Tirado, E., and Reyes- García, C., Analysis of an infant cry recognizer for the early identification of pathologies. Nonlinear Speech Modeling and Applications. 3445:404–409, 2005.

[2] Orozco, J., and García, C., Detecting pathologies from infant cry applying scaled conjugate gradient neural networks, in European Symposium on Artificial Neural Networks, Bruges (Belgium), 2003, pp. 349–354.

[3] García, J., and García, C., Acoustic features analysis for recognition of normal and hypoacustic infant cry based on neural networks. Lecture Notes in Computer Science, Artificial Neural Nets Problem Solving Methods, 2687:615–622, 2003. doi:10.1007/3-540-44869- 1_78

[4] Várallyay, G., Jr, Benyó, Z., Illényi, A., Farkas, Z., and Kovács, L., Acoustic analysis of the infant cry: classical and new methods, in 26th Annual Internation Conference of the IEEE EMBS, San Francisco,CA,USA, 2004, pp. 313–316.

[5] Reyes-Galaviz, O., Cano-Ortiz, S., and Reyes-Garca, C., Evolutionary-neural system to classify infant cry units for pathologies identification in recently born babies, in 8th Mexican International Conference on Artificial Intelligence, MICAI 2009, Guanajuato, Mexico, 2009, pp. 330–335.

