

Inferring Random Cry Signal: A Neural Network Approach

¹Yogesh Kumar Mishra² Dr. Jai Prakash Saini

¹Associate Professor, ²Director, NSIT, New Delhi.

¹Electronics Engineering Department,

¹Kamla Nehru Institute of Technology, Sultanpur, India.

Abstract: Detection of the infant's cry signal is important because of the fact that infants convey the information only through cry signal. There are two major class of cry signal identified as pain and hunger cry. The detection of the pain cry signal is crucially important as cries are used for clinical purposes to diagnose several pathological states such as chromosomal abnormalities, brain damage, and sudden infant death syndrome (SID). Here neural has been used for the detection of the pain cry signal to success rate worth mentioning accuracy of 96.4%.

Index Terms - Infant's Cry, Neural Network, LPC, Supervised Learning Algorithm.

I. INTRODUCTION

Primitive stage of speech related to human being corresponds to cry of infant either due to pain or hunger. It is therefore desirable to recognize pain and cry signal associated with infants that are whether they correspond to the hunger or pain in certain part of the body. In new born the cry is of great importance in the sense that it serves some physiological needs by reorganizing the cardio respiratory system at the time of birth. Spectrographic study of cry signal has revealed by Arnon Cohen and Ehud Zmora, 1984 that it can be used for the diagnostic purpose. The cry analysis performed by listener from spectrograph is entirely subjective and hard to be recommended for clinical purposes. From the spectrographic method following types of cry signals have been identified as (1) Birth Cry (2) Pain Cry (3) Hunger / Thirst Cry (4) Pleasure / Emotional Need Cry. Out of these four type of cry signal it is of utmost important to detect the cry signal because this signal may be pertaining to some health related emergency need. Using sound spectrogram and audio perception special characteristics of the cries are identified for clinical purposes to diagnose several pathological states such as chromosomal abnormalities, brain damage, and sudden infant death syndrome (SID). Infant's cry in pathological states is reported to have distinct features such as high pitched cry in cases of brain damage. The latency of cry is also

an important feature in analysis. The analysis of infant's cry signal as reported in literature, is performed mainly by the following three methods, (1) Trained observer (2) Manual analysis of spectrograph (3) Statistical methods.

For application of cry analysis in clinical use, it is necessary to develop a sound objective method or a well-established tool for automatic monitoring and evaluation. Several such methods have been reported in the literatures. S. E. Barajas-Montiel and C. A. Reyes-Garcia 2005 has compared the classification capability of systems based on neural network and support vector machine and concluded that the neural network based system has better classification accuracy than the support vector machine

II. SYSTEM IMPLEMENTATION

The schematic block diagram of implementation of system that can discriminate the hunger and pain cry is shown in following Figure:1.



Figure:1 Block diagram of Cry detection System

The system handles the cry signal in Time domain. The cry signal source is constrained one in terms of sampling rate bits per sample and pitch of the cry signal. In the second block the LPC coefficients are computed. The third and fourth blocks are devoted to feature extraction and classification respectively. The autocorrelation LPC coefficients are computed using model of vocal tract described by Sadaoki Furui (1999) as system function

$$H(z) = \frac{G}{(1 - \sum_{k=1}^p a_k z^{-k})}, \quad (1)$$

where G is the gain., a_k are the linear predictor coefficient sand k being the order of prediction . The linear difference equation of order p pertaining to speech signal model is described by

$$S(n) = \sum_{k=1}^p a_k S(n-k) + Gu_n. \quad (2)$$

In quest of minimizing the predictor error

$$e(n) = S(n) - \sum_{k=1}^p a_k S(n-k) \quad (3)$$

Considering stationary random noise as input, thus obtained a_k are the linear predictor coefficients. Tacking $S(n)$ lies in the interval $0 \leq n \leq N - 1$, set of linear difference equation is required to be solved.

$$R(i) = \sum_{k=1}^p a_k R(|i - k|) \quad 1 \leq i \leq p \tag{4}$$

and $R(k) = \sum_{m=1}^{N-1-k} a_k S(m)S(m+k)$. (5)

Where $R(i)$ is the autocorrelation function. There are other type of the functions are also used in the minimization of the error and accordingly named as in case of Levinson – Durbin algorithm.

As the order of the prediction increases the number of predictor coefficients increase. In order to reduce the number of coefficients reduces by clustering the linear predictor coefficients. The clustering algorithm to handle such type of data is given below as suggested by Ghanim 1997.

Initialize the initial weight vector b_{ji} with the condition that the b_{ji} are unique for $j = 1, 2, \dots, N$, where N is the number of neurons and keep small weights for practical reasons as suggested by Pao, 1989, Amjed AL-Ghanim, 1997:

1. Apply the pattern $\bar{x} \equiv x_i$ to all neurons of input layer and apply bottom-up processing to calculate a weighted sum given by

$$y_j = \sum b_{ji}x_i \tag{6}$$

2. Use the MAXNET procedure to find the neuron in the output with the largest y_j value.
3. Verify that \bar{x} genuinely belongs to the j th cluster by finding output by top-down calculation; i.e. form the weighted sum

$$\sum_i t_{ji}x_i \tag{7}$$

to satisfy the condition that \bar{x} belongs to j th cluster if $\frac{\sum_i t_{ji}x_i}{\|\bar{x}\|} > \rho$ where ρ is vigilance parameter. If this equation is satisfied the proceed to step 5 otherwise to step 6.

4. Update bottom up and top down weights respectively b_{ji} and t_{ji} for that specific j and all i .
5. If \bar{x} does not belong to the most likely node, then deactivate that node and perform step 2 for initiating another cluster center. The initialization and update procedures are:

$$\begin{aligned} t_{ji}(0) &= 1 \\ b_{ji}(0) &= \frac{1}{1+N} \\ t_{ji}(n+1) &= t_{ji}(n)x_i \\ b_{ji}(n+1) &= \frac{t_{ji}(n)x_i}{0.5 + \sum_{i=1}^N t_{ji}(n)x_i} \end{aligned}$$

Once the data is segregated linear predictor coefficients classification can be done by machine learning tools. Here the supervised learning paradigm of artificial neural network that uses back propagation algorithm is employed. The various steps of the algorithm is given below with the notation that i, j and k belongs to input, hidden and output layer respectively (R.P.Srivastava, 1991):

1. Initialize the weights and apply the sample of input vector $x_{p1}, x_{p2}, \dots, x_{pn}$ to the input layer.
2. Find out the value at every node of hidden layer $net_{pj}^h = \sum_i w_{ji}^h x_i + \theta_j^h$, where w_{ji} weight vector of node i and j and θ_j^h is the bias unit of j th node of hidden layer.
3. Obtain output of hidden layer using equation $i_{pj} = f_j^h(net_{pj}^h)$.
4. Calculate the net input output layer neuron by $net_{pk}^o = \sum_{j=1}^i i_{pj} + \theta_k^o$.
5. Calculate the output of output layer neuron using equation $o_{pk} = f_k^o(net_{pk}^o)$.
6. Obtain error for output layer neuron by $\partial_{pj}^o = y_{pk} - o_{pk}$.
7. Hidden layer error is calculated as $\partial_{pj}^h = f_j^h(net_{pj}^h) \sum_k \partial_{pk}^o w_{kj}^o$.
8. Update the weights of the output layer neurons using equation $w_{ji}^o = w_{ji}^o(i) + \eta \partial_{pj}^o i$.
9. Update the weights of the hidden layer neurons $w_{ji}^h = w_{ji}^h(i) + \eta \partial_{pj}^h x_i$.

The steps 2 to 9 are repeated till the error is minimizes to tolerable limit. This total system error is calculated by using formula

$$E_p = \frac{1}{2} \sum_{k=1}^M \partial_{pk}^2 \tag{8}$$

III Simulation Result

The linear predictor coefficients obtained the time domain sampling of cry signal. The sampling frequency is 8 KHz. The hunger and pain cry signal in time domain are shown in Fig. (1) and Fig. (2) respectively. Frame length is kept at 50 msec. with a overlap period of 16 msec. Number of samples in a frame is kept at 256. An all pole model with four pair of poles has been considered to compute the 10th order linear predictor coefficients with the help of signal processing software MATLAB R1016b. Staring with 66 linear predictor coefficients, the number of cluster generated to 50. The classification is optimized to train with only 50 coefficients which take into account to various features of the input cry signal. The normalized system error has been minimized to a tolerable limit.

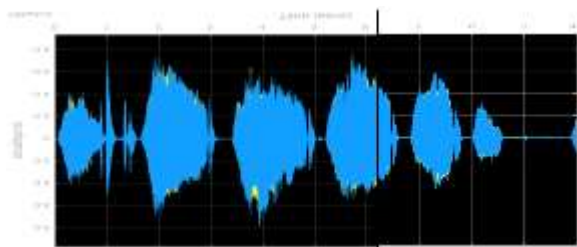


Fig. (1): Hunger Cry Signal (Time-Domain)

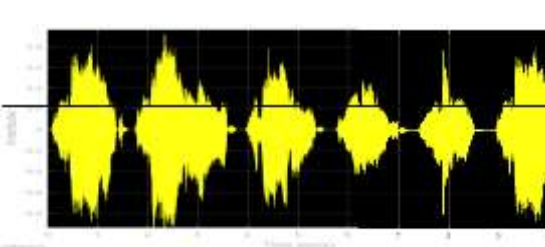


Fig. (2): Pain Cry Signal (Time-Domain)

Various training network and training parameters are listed in Table 2.1 and Mean Square Error (MSE) versus number of iterations graph is shown in Fig. 2.10.

Parameter	Value
Network	50: 32:2
L.R. (η)	0.7
M.R. (α)	0.03
Total Error	10^{-3}
Max. Individual Error	10^{-4}
Number of Iteration	200

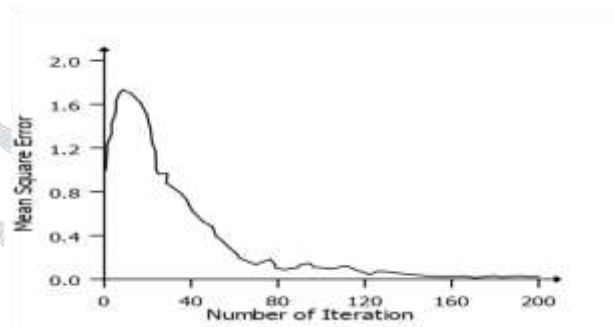


Table 1: Neural Network Parameters

Fig. (3) MSE Vs Number of Iteration plot

The knowledge presentation for classification of the cry signal was made by deriving two types training data file. First Training samples two outputs as hunger cry signal and pain cry signal while second data file contains four out puts as cry signal, not a cry signal, pain cry signal and not a pain cry signal. Because prior to data classification during it is not known that whether the cry signal is due to pain or hunger or some other reason. In the first case the correct classification rate is 94.6% which becomes 97.3% in the second type of the network.

REFERENCES

- [1] Arnon Cohen and Ehud Zmora, 1984. Automatic classification of infant's hunger and pain cry, Proceeding of the International Conference, Florence, Italy, pp. 667-672.
- [2] S. E. Barajas-Montiel; C. A. Reyes-Garcia, 2005. Identifying Pain and Hunger in Infant Cry with Classifiers Ensembles, International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, vol. 2, pp. 770 – 775.
- [3] Sadaoki Furui, 1999. Speech Analysis, Wiley Encyclopaedia of Electrical and Electronics Engineering.
- [4] Amjed AL-Ghanim, 1997. An Unsupervised Learning Neural Algorithm for Identifying Process Behavior on Control Charts and a Comparison with Supervised Learning Approaches, Computers & Industrial Engineering.
- [5] Y. H. Pao, 1989. Adaptive Pattern Recognition and Neural Networks, Addison - Wesley Publishing Company.
- [6] R.P.Srivastava, 1991. Transformations and Distortions Tolerant Recognition of Numerals Using Neural Networks, Proceedings of the 19th annual conference on Computer Science, San Antonio, Texas, USA, pp. 402-408.