

COMPLEXITY ANALYSIS OF NORMAL AND DEAF INFANT CRY ACOUSTIC WAVES

Kathiresan Manickam, Haizhou Li,

Institute of Infocomm Research (I²R), 21 Heng Mui Keng Terrace, Singapore 119613

Abstract: This work describes the complexity found in the normal and the deaf crying acoustic waves. Using approximate entropy, in a single figure, the complexity of the auto-covariance of the signal is computed. Thus, using this complexity value, we are able to discriminate the normal and the deaf infants crying domains with ($P < 0.01$).

Keywords: Infant Cry, Complexity, deaf infant

1. INTRODUCTION

The infant crying waves, seemingly chaotic, carry useful and nevertheless essential information to establish its culture. Such entity vividly clarifies an infant's physiological anatomy and psychological condition. Physiological quantities from the laryngeal configuration, i.e. the length of the vocal tract, in return exemplify the resonance and formant effects. Psychologically, the infant's mental stamina correlates with the origins of a cry type. Modes of cries are unequivocally classified as normal, pathological, pain, hunger, etc. Hitherto, cry from a "normal" and "deaf" infant has been notably studied in literature [1]. Curiosity may result in the selection of the deaf population as the target group. Deaf infants cry exhibit contrasting acoustical characteristics compared to their rivals, the healthy infant population. Following investigation, it became apparent that the caseload for the deaf population is ever increasing in most institutions. It also came to light that both deaf and healthy infants commonly share attributes like pain or hunger cry. But, the hearing impaired infants acquired anatomical deficiency. Despite this setback, paediatricians have expressed that infant's cry is reciprocal to adult's speech.

Modes of infant's cry can be qualitatively characterised with commonly employed acoustics cues. Accordingly, normal crying features constitute of raising-falling pitch pattern, ascending-descending melody and high intensity seen from the spectrum [2-3]. Pathological infant crying correlate well with some normal infant's acoustics features. Spectral intensity will be lower than normal, rapid pitch shifts, generally glottal plosives, weak phonations and silences during the crying. Parameters, incorporating the pitch and formant descriptions, have been well utilised in the infant cry analysis up till now. Clement et al have substantiated that variations in the pitch

between hearing impaired and normal groups become meaningful from 8.5 months onwards [1]. In reality, this time gap is a result of lack of auditory feedback on the speech. Thus, evidence proves that a deaf group tends to voice louder since they want to hear themselves.

Pitch, a famously captured feature, aids in distinguishing cry types, as well as for a diagnostic tool, etc. La Gasse states, "The cry has enormous potential diagnostic". His quotation suggests that extremely high pitched cry indicates the pathological status of the infant which needs urgent medical attention [4]. A typical example is when infants exposed to drugs tend to have high pitch with variation at lower amplitude. The consequent effect of these drugs is the instability of the neural control of the vocal tract. Consequently, the vocal tract configuration determines the structure of the formant. Nevertheless, estimating the exact formant frequencies is complex [5]. A Fort et al has incorporated a parametric model using poles and zeros method to estimate fundamental frequency and formants from an infant cry, the conventional method involving the glottal pressure and tract configurations.

Regardless of these scientific features, experts in this field are able to distinguish the modes of cries. Garcia has mentioned that parents are specialists who were able to differentiate modes of cry solely using their instincts and comparing different types of cries [2-3]. However, uncertainty in the therapeutic service has brought irrefutable questions. A professional has quoted, "A deaf infant's characteristic varies from one another based on three factors: degree of loss; type and period of rehabilitation and the age of pathology identification". Consequently, concrete answers are unavailable scientifically (energy, pitch, duration, etc) regarding cry prosodic information which has forced us to lead this research in order to create an expert system to verify the cry status. The expert system should be able to analyse and classify modes of cry signals and probably, being realistic, at a later stage, diagnostic applications. Because the cry signals are noisy and evidences are showing their chaotic features, currently, analysing such signals itself has become problematic.

Deaf infants revealed more variations in their phonation using false vocal cords producing falsetto

waveforms [1]. Emergence of these irregular signals is the source for this paper. Since cry is an early form of adult speech, it is acceptable to use the conventional speech processing techniques to quantify these complex signals [6]. In this paper, our aim is to analyse if there is any complexity difference between the deaf and normal infant crying populations. Thus this initial step might help us in the diagnostic process.

II. CRY WAVE COMPLEXITY

Fig 1 below shows an example waveform of deaf and normal infant. Cry breaks and amplitude variations are often seen in the deaf infants. This propagates us to our initial comment regarding irregularity, chaos and complexity. Quantifying non-linear biological signals, due to their complex structure, require a reliable approach. Since we are aware of variability statistics like median, mean, standard deviation etc, it is observed that such methods are insufficient to quantify an erratic waveform. A suitable candidate, using regularity statistics, approximate entropy (ApEn), might rescue us from this problem [7]. One of our initial studies discriminates the phonation voice quality changes seen in healthy and radiotherapy larynx cancer patients using approximate entropy [8-9]. A segment from the waveform will be used as reference to identify a similar segment across the entire data.

(that means more complicated pattern) or a small value (with more determinable pattern). The single parameter is so robust that the detailed medical characteristics can be displayed in a simplistic scientific means. Investigating time domain signals requires suitable normalisation in order to compare across all individual infants.

Auto-covariance function of a frame, 1024 points, might ease this criterion. By doing this, white and other non-structured noises will result in low lags. Non-cry or cry breaks may produce low lag which is often seen in the deaf more than the normal infant as in Fig 1. However, alternating burst of cries and non-cries for the deaf groups will result in high complexity estimate. Low or zero lags will produce nil or low complexity. Fig 2 shows examples of auto-covariance signals for both disciplines (normal and deaf). More determinable features are normally observed from the healthy infant with a low complexity (0.148) as in Fig 2. A portion of the auto-covariance function is a classic example of the low complexity healthy infant's crying since it retains a sinusoidal waveform. Fig 3, demonstrates a typical example of deaf cry which earns a higher complexity estimate (0.719). The undeterminable irregular auto-covariance waveform is the cause for high computation.

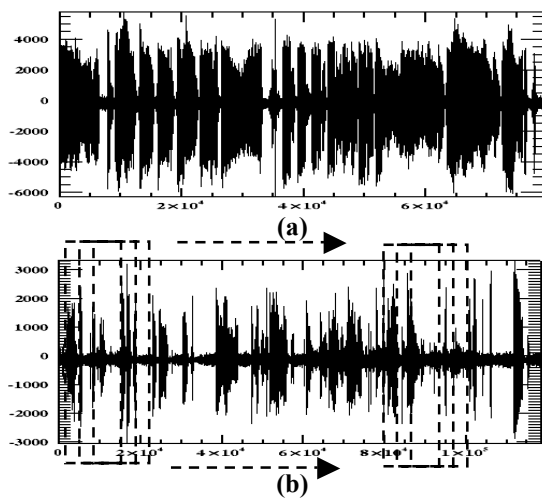


Figure 1
Infant Cry Waveform

(a): Normal Infant Cry Waveform

(b): Deaf Infant Cry Waveform

Abscissa: Time Ordinate: Amplitude

Because of the temporal sliding window across the desired signal of analysis, Fig 1 (dashed line); approximate entropy has a possible potential for characterising the complexity of a similar pattern. The complexity itself is expressed in a single featured Fig that branched to either a large value

III. COMPLEXITY ANALYSIS

The present corpus is a collection of infant crying samples, from early born up to 7 months. A total of 31 normal and 103 deaf infant cries were analysed in this research. Cry signals were recorded and sampled at 8 KHz. The data files were stored, visualised and analysed using software written in scientific language IDL from Research Systems. The recorded signal is divided into frames of 1024 data points each. To normalise the data frames, each frame was performed with auto-covariance function. Complexity value is calculated for each auto-covariance frame based on $N=1024$, $m=3$ and $r=0.2*\sigma$. Each frame produces a complexity value and these complexities do not always conform to a normal distribution. Subsequently, the median from the collection of the complexities was calculated and used for further analysis.

Fig 4 shows the distribution of the median complexity for an individual infant. The ratio of a normal infant below 0.6 and above 0.6 is nearly 5:1. A reversed scenario is echoed in the deaf population. The ratio of a deaf infant with below 0.6 to above 0.6 is nearly 1:4.

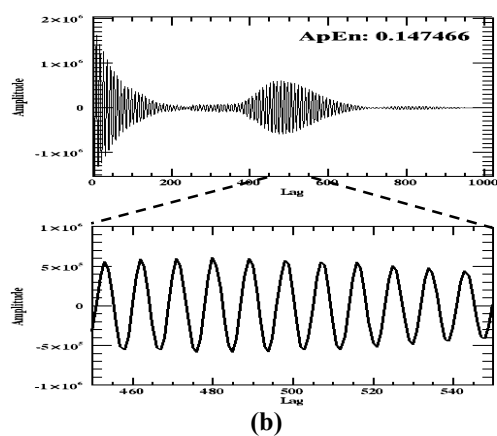
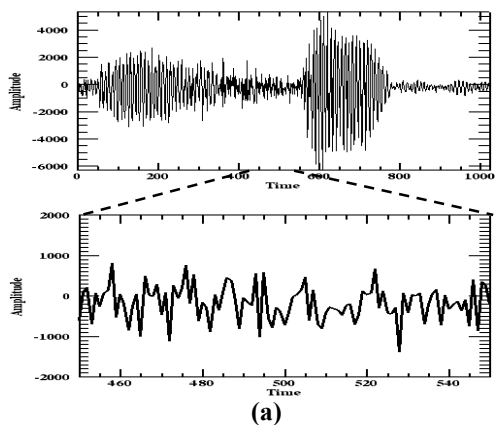


Figure 2
Normal Infant Cry Waveform

(a): Cry Waveform
(b): Auto-Covariance Cry Waveform
Top: Entire Frame Signal
Bottom: Portion of Frame Signal
Abscissa (a): Time **Abscissa (b):** Lag **Ordinate:** Amplitude

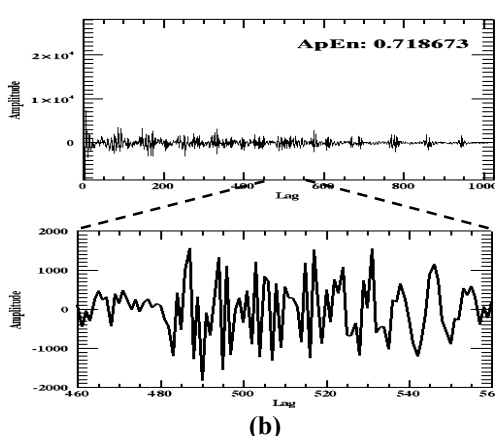
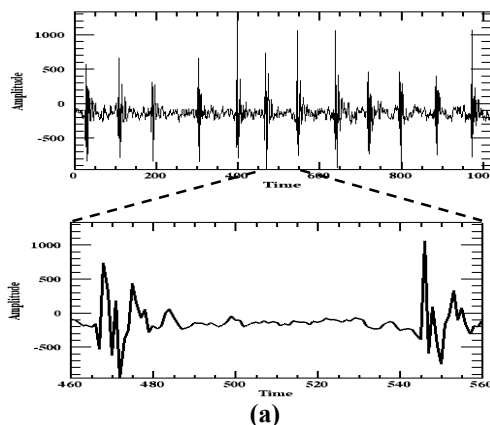


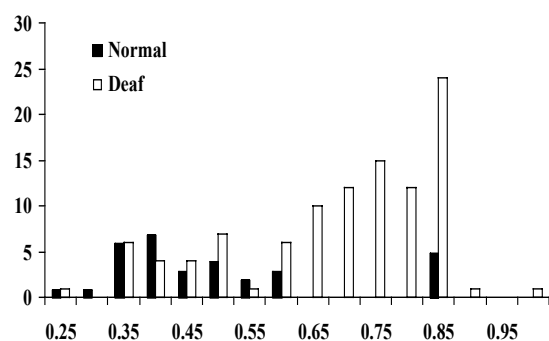
Figure 3
Deaf Infant Cry Waveform

(a): Cry Waveform
(b): Auto-Covariance Cry Waveform
Top: Entire Frame Signal
Bottom: Portion of Frame Signal
Abscissa (a): Time **Abscissa (b):** Lag **Ordinate:** Amplitude

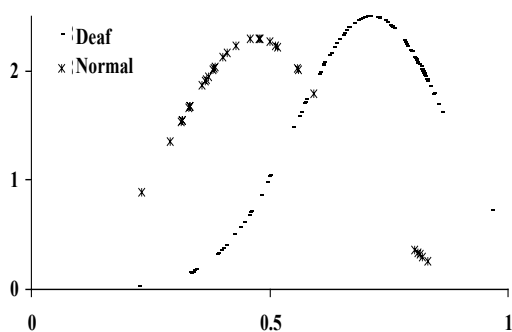
The distribution clearly shows a distinct separation between the deaf and normal infant population. Assuming that these two populations (deaf and normal) are independent, Wilcoxon Rank Sum Test showed that these two populations were indeed significantly different with ($P < 0.01$). Deafness is the most common of all forms of permanent damage following meningitis. Early detection and therapy might reduce the effect or severity following such disease.

IV. CONCLUSION

Despite successfully discriminating the normal and deaf infant crying modes using complexity analysis (approximate entropy), further research has to be carried out in order to reduce the over-lapping portion between the two domains. Nevertheless, this initial study on the infant cry wave analysis is encouraging but more features need to be incorporated like pitch, formant, and energy to enhance the findings.



(a)



(b)

Figure 4

(a): Histogram of Normal & Deaf Infant Cry Complexity

Abscissa: Median Complexity

Ordinate: Frequency of the Complexity

(b): Probability Density Function of Normal & Deaf Infant Cry Complexity

Abscissa: Median Complexity

Ordinate: Probability Density

REFERENCES

- [1] Clement, Chris J. / Koopmans-van Beinum, Florian J. / Pols, Louis C. W. (1996): "Acoustical characteristics of sound production of deaf and normally hearing infants", *ICSLP*, 1549-1552.
- [2] Jose Orozco Garcia and Carlos A Reyes Garcia, (2003) "Acoustic Features Analysis for Recognition of Normal and Hipoacusic Infant Cry Based on Neural Networks" Lecture Notes in Computer Science, 2687:615 – 622, ISSN: 0302-9743
- [3] Jose Orozco Garcia and Carlos A Reyes Garcia, (2003) "A Study on the Recognition of Patterns of Infant Cry for the Identification of Deafness in Just Born Babies with Neural Networks." 8th Iberoamerican Congress on Pattern Recognition, CIARP 2003, Havana, Cuba, 342-349, ISBN 3-540-20590.
- [4] Medical News Today: Babies cry linked to their neurological and medical status, 16 May 2005. www.medicalnewstoday.com
- [5] K Wermke, W Mende, C Manfredi, P Bruscaiglioni, (2002) "Developmental aspects of infant's cry melody and formants", *Medical Engineering Physics* 24:501-514.
- [6] Qiabing Xie, Rabab K Ward and Charles A Laszlo, (1996) "Automatic Assessment of Infants' Levels-of-Distress from the Cry Signals", *IEEE Trans Speech and Audio Processing* 4:4:253-265.
- [7] Pincus S M, (2001), "Assessing Serial Irregularity and its Implications for Health", *Ann NY Acad Sci*, 954:245-267.
- [8] Moore C J, Manickam K, Willard T, Jones S, Slevin N, Shalet S, (2004) 'Spectral Pattern Complexity Analysis and the Quantification of Voice Normality in Healthy & Radiotherapy Patient Groups', *Medicine Engineering Physics*, 2004,26(4), 291:301
- [9] Manickam K, Moore C J, Willard T and Slevin N, 'Quantifying Aberrant Phonation Using Approximate Entropy in Electro- Laryngography' (*Journal of Speech Communications-In press Corrected Proof*)