Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, 2017

Infant Cry Analysis of Cry Signal Segments Towards Identifying the Cry-Cause Factors

Shivam Sharma[1] and Vinay Kumar Mittal[2]
[1]Indian Institute of Information Technology Chittoor, Sri City, A. P., India
[2]CEO, Ritwik Software Technologies Pvt. Ltd., Hyderabad, Telangana, India
shivam.sharma@iiits.in, DrVinayKrMittal@gmail.com

***Abstract*—Infants communicate their physiological state and emotions mostly by crying. Identifying the cause of crying is natural and easy for human beings, but challenging for machines. Automatic identification of cry-cause factors has vast applications in assistive healthcare and timely remedial measures in critical situations. In this paper, Infant cry signals analysis is carried out to identify four cry-causes, namely, *Environmental Change*, *Discomfort*, *Stranger's Anxiety* and *Pain*. Different causes of infant crying are characterized by examining changes in the acoustic features that are extracted from acoustic signals in the cry segments. Mainly, the instantaneous fundamental frequency (F_0) is extracted using auto-correlation of cry signal. Changes in F_0 contours for different cry-causes are examined by deriving parameters mean, standard deviation and normalized standard deviation. A *Infant Cry Signals Database (IIITS ICSD 2)* is used, which was especially collected for this study. Results are validated by observing changes in the F_0 and its harmonics in spectrograms, derived using short-time Fourier transform. Few distinct patterns of changes in the F_0 contours are observed for different cry causes. *Pain* cry is characterized by higher standard deviation in F_0, than for *Discomfort*. The standard deviation in F_0 for *Environmental Change* is more than for *Stranger's Anxiety*. Normalized standard deviation is highest for *Pain*, than for *Environmental Change*, than for *Discomfort* and lowest for *Stranger's Anxiety*. These insights should be helpful towards automated identification of infant cry-causes and related numerous applications.**

***Keywords*—infant cry analysis, *ICSD2*, pain, discomfort, environmental change, stranger's anxiety, F_0 contour, cry segments**

I. Introduction

Crying forms an important clue towards ensuring good health of the new born baby and up-to a certain age of a neonate, there can be a lot that can be inferred about the psychological and physical state of an infant. This knowledge can be leveraged towards timely health-care. Crying is coordinated by the neural centres [1]. It can be understood as an amalgamation of different forms of vocalization, constrictive silence, coughing, choking and breaks as stated in [2].

Physiological behaviour of an infant were studied in [3] to understand his/her needs. Source excitation features were analysed using Spectrographic analysis [2]. F_0 analysis was done by implementing the Welch's method and FFT in [4]. F_0 and first three formants were determined using Cepstrum analysis in [5]. System characteristics were studied as part of [7], where short time analysis and MFCC were obtained, followed by k-NN based classification. The same author in [7] proposed a robust automatic detection of the infant cries in a natural conditions using the features like MFCCs, pitch and Formants and finally evaluated a deep learning based Convolutional Neural Network. A 3D crosscorrelogram is implemented to extract F_0 in [8]. F_0 analysis was done by implementing the Welch's method and FFT for hunger, sleepiness and discomfort [9].

The initial investigation of infant cry sounds was reported in detail in [10]. This was followed by comparative study of the characteristics using the features like F_0, *Strength of Excitation* (SoE) and *Signal Energy*, for *Discomfort* vs. *Pain* cry sounds in [11] and [13]. Study on expressive voices using sub-harmonic regions was done for Noh voices in [13]. Non-verbal paralinguistic are studied in [16] and using *modZFF* in [14] and [15]. Still there is need to examine other important cry causes and the corresponding acoustic behaviour. With our initial results being reported in [17], our extended dataset *ICSD2* was introduced in [21]. The next step is to quantify the characteristic behaviour of these features.

The aim of this paper is to quantify the acoustic characteristics using statistical parameters, of the cries, that are previously observed in our earlier work [21]. F_0 related characteristics are compared for *Environmental Change, Stranger's Anxiety, Pain and Discomfort*. Also, an additional feature *Cry Duration*, important from its historical implementation's perspective, is observed for the above mentioned categories. There are some striking observations made for some cases, whereas few further challenges are elucidated regarding important distinctions.

This paper is organized as follows. Details of *ICSD2* database are discussed in Section II. In Section III, the methods and features evaluated are described. Details of the experiments are included in Section IV. Detailed description of the observations from acoustic analysis are given in Section V. Section VI constitutes the key results and observations. The paper is summarized and concluded with the scope of further work in Section VII.

II. Infant Cry Signals Database [ICSD 2]

The dataset for the current work is collected from a paediatrics clinic. Zoom H4n, a handy recorder, with built in X/Y stereo mics is used to record the perturbation free cry voices from a distance of 10-20 cm. The recordings are made at 48 KHz sampling rate with bit coding rate of 16 bit. Care is taken that only the cry parts clean enough for a reliable analysis are considered for study.

TABLE I *Age and gender-wise (M/F) summary of Infant Cry Signals Database (IIITS-ICSD 2)*

Age (Months)	*File Count*			*Cry Count*			*Duration (Min.)*
	M	*F*	*Total*	*M*	*F*	*Total*	
< 1	2	7	9	11	79	90	3
1-3	9	10	19	107	156	263	11
4-6	12	10	22	70	58	128	22
7-9	11	12	23	98	68	166	8
10-12	21	14	35	149	99	248	11
13-15	6	17	23	58	180	238	10
16-18	6	9	15	34	99	133	6
19-24	11	11	22	96	80	176	7
25-30	3	2	5	17	26	43	2
31-36	0	4	4	0	18	18	1
> 36	2	2	4	20	8	28	2
Total	*83*	*98*	*181*	*660*	*871*	*1531*	*83*

The cry sounds are recorded for a total of over 100 subjects (50 Male and 50 Female) for a period of 5 days. Although, the infants age between 2 days and 6 years in the collected dataset, but majority fall in the range 6-15 months. The detailed summary of the total cases collected along with the number of cries in the sound files with respect to the age and gender is given in Table I. Primary crying cause for the inpatients, as attributed by the doctors is majorly due to vaccination pain, or Stranger's Anxiety. Other interesting categories that are noted are Discomfort and Environmental Changes. Overall, there are 7 distinct categories which are enlisted in the next section. The relevant cry sample details including their ground truth cause labels are collated in the form of metadata.

The cry recordings made during the check-up are considered and categorised based upon doctor's consultation. Fever and physical wound like anal fissure are clubbed under Ailment, whereas cries while yearning to go back to parents are brought under Emotional Need. On-going and historical medical conditions, parent's inputs and present health status, as adjudicated by the doctor formed the basis of ground truth categorization.

III. Signal Processing Methods And Features

A. Signal Processing Methods

(1) *Auto-correlation Method:* Autocorrelation provides a measure of the self-similarity given by,

$$r_x(m) = E[x(n)x(n+m)] \tag{1}$$

between the waveforms of the time functions, i.e., the source signal $x(n)$ and the delayed version of itself $x(n+m)$ [18]. F_0 contour for cry melody contour analysis is derived using this technique.

(2) *Short-time Fourier transform (STFT):* This is obtained by processing the segments of the cry signals in the frequency domain, which is given by

$$X(\tau,\omega) = \sum_{n=-\infty}^{n=\infty} x[n]w[n-m]e^{-j\omega n} \tag{2}$$

with $x[n]$ being the signal and $w[n]$ the window function [19]. The log magnitude of this gives a spectrogram and energy, used for primary spectral detailing of an acoustic cry signal.

B. Features

Acoustic analysis of cry signals involves consideration of relevant features, whereas quantitative analysis of the characteristic trends need some selective parameters as,

(1) F_0 *contour:* Considered important for analysis [8], the contours computed are plotted along-with the F_0 estimates, by a popular SOA algorithm YIN [20], which served as a reference here. The F_0 contour estimates are computed using Autocorrelation with window length 30 ms and a shift of 10 ms.

(2) *Short-time Spectrogram:* Short-time narrow band spectrogram obtained using Frame-size 30 ms and shift of 10 ms. The FFT size is set to 4096.

The F_0 estimates for the contour are analysed for individual cry segments, thereby giving an idea of the contour behaviour at the cry segment level. Various statistical parameters considered here for analysis are enlisted in Table II ahead.

IV. Experiments

The input cry signals (Fig. 1 and 2, Sub-plots (a)) are observed post application of the voice activity detection, marked in red in observation plots ahead, after silencing the regions with unwanted background noises. At first, spectrograms are used for observing any characteristic spectral behaviour. This is a 512 point FFT'd short-time spectrogram with Hamming window, and pre-emphasis factor of 0.97.

Further spectrogram analysis and comparison is done using Matlab generated spectrograms (Fig. 1 and 2), sub-plots (b)), for which the details are given in Section III-B. The cry cause categories examined are: *Ailment, Discomfort, Emotional Need, Environmental Factors, Hunger/Thirst, Pain and Strangers Anxiety*. In each category, 3, 15, 2, 15, 2, 16 and 20 cases, respectively, are analyzed for this study.

The variations in the cry signals are too much, primarily because of the age factor, which shows intense changes from 9 months onwards. Besides, the presence of different sound types that make up a cry signal induce lot of fluctuations in its characteristic behaviour. The experiments conducted are upon the infants lying mostly between the age group of 1-20 months. Also, at times the presence of different sound types can be crucial for analysis, which is duly considered and noted in [22].

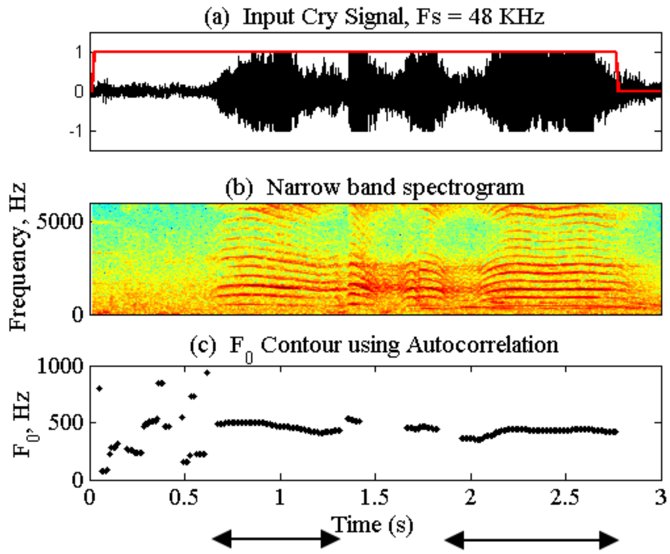

Fig. 1 Acoustic features for infant cry due to *Environmental Changes* [Spkr 60, *Male*, 19 months]. (a) signal, (b) Narrow-band Spectrogram, (c) F_0 contour (using autocorrelation) [Notice flat contours in the arrow-marked regions].

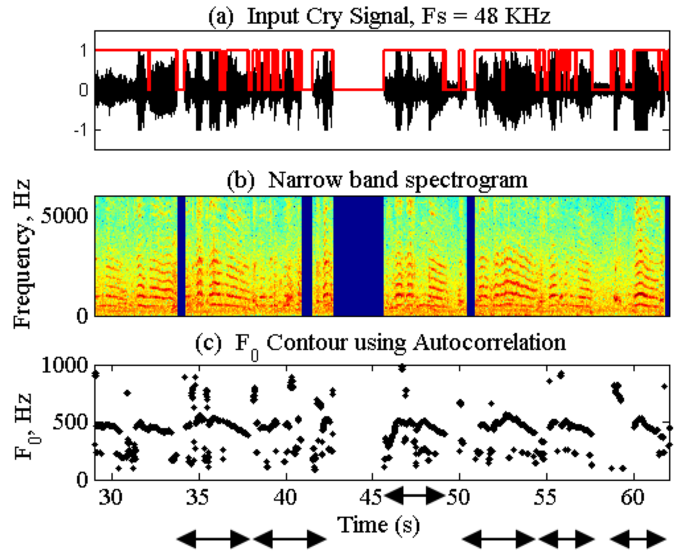

Fig. 2 Acoustic features for infant cry due to *Stranger's Anxiety* [Spkr 62, *Female*, 15 months]. (a) signal, (b) Narrow-band Spectrogram, (c) F_0 contour (using autocorrelation) [Notice inverted cup-shaped contours in the arrow-marked regions].

TABLE II *Analysis Parameters and their Notations*

Parameter Description	*Notation*
(a) Average F_0 for Cry-cause categories	$\mu_{F0_1}(Hz)$
(b) Std. Dev. of F_0 for Cry-cause categories	$\sigma_{F0_1}(Hz)$
(c) Average cry duration for Cry-cause categories	$\mu_{dur_1}(sec)$
(d) Average F_0 for each case (file)	$\mu_{F0_2}(Hz)$
(e) Std. Dev. of F_0 for each case (file)	$\sigma_{F0_2}(Hz)$
(f) Normalized Std. Dev. of F_0, $\sigma N_{F0_2} = \sigma_{F0_2}/\mu_{F0_2}$	$\sigma N_{F0_2}(Hz)$
(g) Average cry duration for each case (file)	$\mu_{dur_2}(sec)$
(h) Average F_0 at the cry signal segment level	$\mu_{F0_3}(Hz)$
(i) Std. Dev. of F_0 at the cry signal segment level	$\sigma_{F0_3}(Hz)$

TABLE III *Category-wise Average and Standard-deviation of the F_0 contour for infant cries, processed at cry segment level.*

Category	*(a)* μ_{F0_1} (Hz)	*(b)* σ_{F0_1} (Hz)	*(c)* μ_{dur_1} (sec)
Discomfort	516.30	274.53	1.23
Pain	658.08	363.09	1.02
Environmental Changes	465.50	252.25	1.09
Stranger's Anxiety	466.63	228.61	1.50

V. Observations from Acoustic Analysis

A. *Distinguishing Environmental Change cries from Stranger's Anxiety*

Our previous observations in [10] and [11] are corroborated in terms of the F_0 statistics. Further comparative analysis for the categories *Environmental Change* and *Stranger's Anxiety* showed very subtle differences with respect to the same features. For *Stranger's Anxiety*, one significant difference based upon the patterns of the cry melody contour is the presence of prominent, but relatively few arch shaped cry signal segments, similar to the ones found within the Pain cry cases, but in abundance. Hence, logically the $\sigma_{F0_1}(Hz)$ for this category should be more than that for *Environmental Change*, but the empirical verification suggests otherwise (see Table III). The reason could very well be the combined effect of different types of *hyper-phonated* (or Non-normal) cry sounds that make up most of the cry cases for *Environmental Change* along-with the scarcity of the arch-shaped cry patterns in the case of *Stranger's Anxiety*. A comparison between these characteristics is shown for brief portions for both the categories in Fig. 1 and 2, leading to statistical analysis at the cry signal segment level, shown in Fig. 3 and 4.

Presence of rising and falling contour shapes can be clearly seen from the spectrogram for the *Stranger's Anxiety* case considered here, as shown in Fig. 2, which is nearly absent for the *Environmental Change* in Fig. 1. This effect can be validated from the Mean-Error Line plots in Fig. 3 and 4, depicting the Average and Std. Dev. of F_0, denoted by $\mu_{F0_3}(Hz)$ and $\sigma_{F0_3}(Hz)$ respectively, along y-axis (in Hz), for the individual cry signal segments depicted along on the x-axis. As can be seen from the Fig. 3, the Standard Deviation for most of the cry signal segments, for the *Environmental Change*, is 100 - 125 Hz, which is less in comparison to that of $\approx>$ 200 Hz for the *cry segments* in the case of *Stranger's Anxiety* (Fig. 4). The quantified melody behaviour observed in the above cases reflects the general characteristics of all the cases considered for this study. With varying degree, these

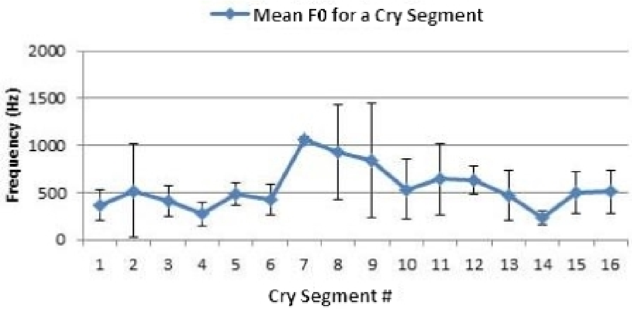

Fig. 3 Illustration of changes in F_0 for different segments of infant cry due to *Environmental Changes* [Spkr 44, male, 9 months]. Please note smaller deviations by vertical bar size.

TABLE IV *Quantitative analysis using parameters to measure melody contour characteristics - (c) Average, (d) Std. Dev., (e) Normalised Std. Dev., of F_0 and (f) Average cry segment-duration for different cases of Environmental Change.*

(a) Infant (M/F)	(b)Age (Months)	(c) μ_{F0_2} (Hz)	(d) σ_{F0_2} (Hz)	(e) σN_{F0_2} (Hz)	(f) μ_{dur_2} (sec)
S44 (M)	9	553.8	249.8	0.45	1.1
S30 (M)	11	348.8	193.7	0.56	1.2
S30 (M)	11	303.4	149.8	0.49	0.8
S70 (M)	12	539.1	199.5	0.37	0.7
S12 (M)	18	399.8	210.3	0.53	0.9
S60 (M)	19	573.6	329.5	0.57	1.3
S83 (F)	12	503.7	269.1	0.53	1.7
S78 (F)	20	428.1	224.8	0.53	0.9
S05 (F)	36	488.2	354.0	0.73	1.4
S59 (F)	44	516.8	342.2	0.66	1.0
Average		465.5	252.3	0.54	1.1

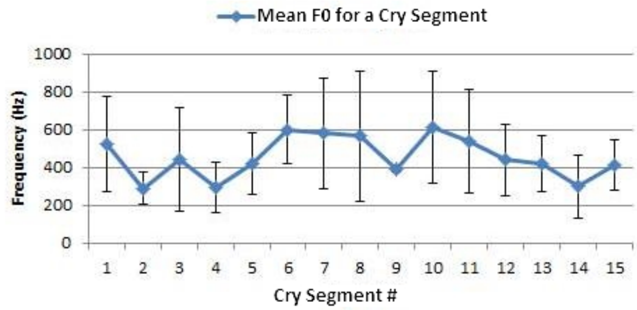

Fig. 4 Illustration of changes in F_0 for different segments of infant cry due to *Stranger's Anxiety* [Spkr 67, *Male*, 12 months]. Please note larger deviations by vertical bar size.

TABLE V *Quantitative analysis using parameters to measure melody contour characteristics - (c) Average, (d) Std. Dev., (e) Normalised Std. Dev., of F_0 and (f) Average cry segment-duration for different cases of Stranger's Anxiety.*

(a) Infant (M/F)	(b)Age (Months)	(c) μ_{F0_2} (Hz)	(d) σ_{F0_2} (Hz)	(e) σN_{F0_2} (Hz)	(f) μ_{dur_2} (sec)
S26 (M)	6	322.7	187.5	0.58	1.1
S67 (M)	12	455.3	197.2	0.43	1.4
S70 (M)	12	519.2	297.4	0.57	1.8
S60 (M)	19	824.8	227.4	0.28	1.7
S79 (M)	24	417.8	160.1	0.38	1.6
S28 (F)	2	302.0	182.4	0.60	2.4
S87 (F)	6	567.7	402.6	0.71	1.0
S85 (F)	13	501.5	270.3	0.54	1.2
S24 (F)	13	305.9	157.7	0.52	1.9
S69 (F)	15	449.6	203.6	0.45	1.0
Average		466.7	228.6	0.51	1.5

arch shaped cry signal segments are present within the cry signals for *Stranger's Anxiety*.

The fact that average Std. Dev. turns out to be lesser for *Stranger's Anxiety* than for *Environmental Change*, not only for the Std. Dev. of F_0 for the entire category, but also for individual cases $\sigma_{F0_2}(Hz)$ shown in Tables IV and V, appeals for highly localised melody contour pattern analysis when characterising for these two categories.

B. Pitch range comparison

(1) *Environmental Change and Stranger's Anxiety:* The values obtained from the statistical[1] analysis for various cases, suggest some insights. The minimum average F_0 for each case/file $\mu_{F0_2}(Hz)$, for the categories *Environmental Change* and *Stranger's Anxiety*, as can be observed from Tables IV and V, are close enough and are $\approx$ 303 Hz for both. But for the latter case, the maximum of this parameter can go even above 800 Hz. Overall, this could be possible due to the subtle stimulation at the beginning of the crying due to change in the external environment for both the cases, resulting in similar average F_0 at both category and case levels (see Table I, IV and V), but the gradual intensifying of the unpleasantaries due to a stranger's presence in case of *Stranger's Anxiety*, tends to induce significant changes in the characteristics like cry melody, pitch etc., as explained previously with the help of Fig. 1 and 2, ultimately leading off to a high F_0 contour.

(2) *Discomfort and Pain:* The case-wise dynamic range of F_0 for the *Discomfort* turns out to be $\approx$ 40 Hz below the corresponding values for the *Pain* category, with minimum/maximum deviations for both the categories as 154.69/586.11 Hz and 196.12/534.13 Hz respectively (Table VI and VII). Such drastic extension of the maximum pitch level for the *Discomfort* category could be attributed to the different types of the Hyper-phonated cry sounds present within, whereas for the *Pain* cry cases, the increment appears to be significantly through

[1] *Statistics*: The meanings of the Notations used can be referred from Section 3, Table II

TABLE VI *Quantitative analysis using parameters to measure melody contour characteristics - (c) Average, (d) Std. Dev., (e) Normalised Std. Dev., of F_0 and (f) Average cry segment-duration for Discomfort cases.*

(a) Infant (M/F)	*(b)Age (Months)*	*(c)* μ_{F0_2} (Hz)	*(d)* σ_{F0_2} (Hz)	*(e)* σN_{F0_2} (Hz)	*(f)* μ_{dur_2} (sec)
S55 (M)	9 (d)	391.7	169.5	0.43	0.6
S09 (M)	11	351.5	163.3	0.46	2.9
S57 (M)	15	541.3	216.1	0.40	1.3
S41 (M)	30	591.7	237.4	0.40	0.7
S34 (F)	7	288.5	154.7	0.54	1.4
S102 (F)	9	725.6	382.3	0.53	1.0
S16 (F)	9	329.4	175.8	0.53	1.3
S09 (F)	11	805.5	586.1	0.73	1.0
S29 (F)	12	762.1	465.1	0.61	1.0
S101 (F)	24	375.7	195.0	0.52	1.3
Average		516.3	274.5	0.53	1.2

TABLE VII *Quantitative analysis using parameters to measure melody contour characteristics - (c) Average, (d) Std. Dev., (e) Normalised Std. Dev., of F_0 and (f) Average cry segment-duration for different cases of Pain.*

(a) Infant (M/F)	*(b)Age (Months)*	*(c)* μ_{F0_2} (Hz)	*(d)* σ_{F0_2} (Hz)	*(e)* σN_{F0_2} (Hz)	*(f)* μ_{dur_2} (sec)
S40 (M)	1	844.9	398.7	0.47	1.0
S25 (M)	4	324.5	534.1	1.65	1.4
S70 (M)	12	720.3	328.4	1.46	1.00
S39 (M)	13	673.3	268.4	0.40	1.00
S32 (M)	16	708.9	416.9	0.59	1.2
S73 (F)	1	627.0	248.7	0.40	0.8
S15 (F)	5	718.5	403.9	0.56	1.0
S14 (F)	7	480.0	196.1	0.41	1.3
S07 (F)	18	812.7	456.8	0.56	1.0
S38 (F)	18	670.7	384.8	0.41	1.0
Average		658.1	363.1	0.61	1.0

normally Phonated cry sounds. This is validated from the $\sigma_{F0_1}(Hz)$, which when observed with the help of the Table III for *Pain* category, is well ahead than for the Discomfort by more than 75 Hz. Also, the *Normalized Std. Dev.* tabulated as column (e) in the Tables above, clearly has *Pain* with the highest value.

C. A perspective on Cry Duration

The comparison of the durations of the cry signal segments, identified by the VAD does not reflect any special significance. Since the Average cry duration for cry-cause categories for *Discomfort*, *Pain*, *Environmental Changes* and *Stranger's Anxiety* are obtained as 1.23 sec, 1.02 sec, 1.09 sec and 1.50 sec respectively, as can be observed from Table III. The sharp surge in the values for the first and last categories is due to exceptional cases having the Average cry duration for each case (file) $\mu_{dur_2}(sec)$ longer than the general trend, which is either due to sustained cry signal segment or the clubbing of two consecutive cry signal segments that are too close for the voice activity detection algorithm to distinguish.

D. Explanation of Possible Reasons

The relative cry intensity levels for the categories under focus can also be interpreted from the perspective of the level of uneasiness the infant might be experiencing when under their influence. For the category *Pain*, the infant would want to get rid of the unpleasantness developing within right away and hence the cry sound production will be through the system, the articulatory characteristics of which would already be in shape so as to convey the intense associated emotions in various ways, that manifest as different cry patterns within a cry signal, whereas for Stranger's Anxiety, the zenith of the cry intensity, lying in the cry segments, is not reached as early as for *Pain* because of the time taken by the infant to get over the external stimulations due to curious change of events accompanied along-with this Category. Categories like *Discomfort* and *Environmental Change* causes an infant to stay perplexed for a significant time, thereby not generating significant acoustic variations within the cry sound.

VI. Key Results and analysis

(1) Cry segment deviation is more for *Pain* than for *Discomfort* by 75 Hz.
(2) Cry segment deviation for *Environmental Change* is more than that for the *Stranger's Anxiety* by 23.65 Hz.
(3) *Core Cry* segment deviation observed for *Environmental Change* is found to be 100 Hz lesser than that of *Stranger's Anxiety*.
(4) The minimum average F_0 for both *Environmental Change* and *Stranger's Anxiety* is ≈ 303 Hz, whereas it's maximum reaches 800 Hz for the latter, suggesting more intensity in its cry patterns.
(5) Average F_0 dynamic range for *Discomfort* is lesser than that for *Pain* by ≈ 40 Hz. The key distinction here lies in the cry melody contours for the cry segments, which is reflected as greater Std. Dev. of F_0 for category average.
(6) Average Cry duration for all the categories is obtained to be 1.18 sec.

The acoustic characteristics of the cry sound produced by an infant vary dynamically with the age. Changes like highly fluctuating cry melody contour tend to get induced with time. Although, the size of infant cry production system is smaller than that of an adult that results in brief cry outbursts, the overall patterns do show some peculiarity in different situations. It is highly likely for a child of age for instance above 18-20 months to have acoustic characteristics significantly evolved from that of the age group for neonates. Therefore, majority of the cases considered for our quantitative analysis, i.e., 88 % are grouped between 8-20 months. Also, the trends observed in the current acoustic analysis are consistent for both Male and Female cases alike. It is during the non-severe cases like the ones due to *Discomfort* or when *Environment Changes*, that the melody and harmonic patterns during the core cries is relatively more stable (i.e. with less changes), whereas the same is observed to be of frequent and larger changes in the case of more severe categories like *Pain and Stranger's Anxiety*, occurring in the form of brief cry outbursts due to psychological reasons involved. The analysis done and the patterns observed strongly reflect the characteristics of 73 cases analysed, thus providing insightful inputs towards further exploratory work.

VII. Summary and Conclusion

Cry melody patterns can primarily be quantified using the Std. Dev. at the cry segment level. Average deviation for *Discomfort* is found to be 75 Hz below that for the *Pain* category, corroborating our previous results, whereas although limited, but despite the presence of fewer arch shaped cry patterns in the *Stranger's Anxiety*, and significant heterogeneity observed in the sound type within the cries belonging to the *Environmental change* category, the resulting average standard deviation is lesser for *Stranger's Anxiety* by 23.65 Hz. Hence, further approaches, as also stated earlier, need to be localised for melody contour pattern analysis for *Stranger's Anxiety* identification.

Besides major similarities, there are subtle differences that are evident because of the arch-shaped cry patterns for this category. Apart from the category specific behaviour, consistent observations concerning the F_0 deviations for the cry signal segments, are found to be clearly conclusive in categorising a cry signal as either *severe* or *non-severe*, the distinction that can prove to be helpful in a lot of cases.

The implications for the utility of such technology are of immense value. Early detection of the medical conditions of infants in cases where it's not that simple to know the cause can be crucial for a better healthcare. Thus making communication between an infant and his/her surrounding more meaningful. The Infant cry analysis needs to happen for the data which is sufficient and as real time as possible, specifically for the categories like Ailment where the dataset is hard to procure. A solution could be crowd-sourcing of real-time cry data collection by having parents/guardians record cries of their babies with the help of a mobile app and comprehensively annotate it. Cry signal characterization can extend towards analysis from system's perspective like using MFCC and Formant frequencies, giving rise to further prospects.

Acknowledgment

The authors are thankful to Dr. Nizam, M.B.B.S (DCH), Dr. Bhavya, M.D. (Ped.) and Dr. Venkat, M.D. (Ped.), the nursing staff of Pranaam Hospital, Madinaguda, Hyderabad, and parents of the infants for supporting in collection of the cry samples needed for the study.

References

[1] Linda L LaGasse, A Rebecca Neal, and Barry M Lester, "Assessment of infant cry: acoustic cry analysis and parental perception," in *Mental retardation and developmental disabilities research reviews*, vol. 11, no. 1, pp. 8393, 2005.

[2] Amy Neustein, Ed., "Advances in Speech Recognition: Mobile Environments, Call Centers and Clinics," Springer, New York, 2010.

[3] Y Skogsdal, M Eriksson, and J Schollin, "Analgesia in newborns given oral glucose," in *Acta Paediatrica*, vol. 86, no. 2, pp. 217-220, 1997.

[4] Raina P. Daga and Anagha M. Panditrao, Article: Acoustical analysis of pain cries in neonates: Fundamental frequency, IJCA Special Issue on Electronics, Information and Communication Engineering, vol. ICEICE, no. 3, pp. 1821, December 2011.

[5] S. Chandralingam, T. Anjaneyulu, and K. Satyanarayana, "Estimation of fundamental and formant frequencies of infants cries; a study of infants with congenital heart disorder," in *Indian Journal of Computer Science and Engineering*, 2012, vol. 3(4), pp. 574582.

[6] Rami Cohen and Yizhar Lavner, "Infant cry analysis and detection," in *IEEE 27th Convention of. Electrical & Electronics Engineers in Israel (IEEEI), 2012*, pp. 15.

[7] Y. Lavner, R. Cohen, D. Ruinskiy, and H. Ijzerman, "Baby cry detection in domestic environment using deep learning," in *2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE)*, Nov 2016, pp. 15.

[8] M. Petroni, M.E. Malowany, C.C. Johnston, and B.J. Stevens, "A new, robust vocal fundamental frequency (F_0) determination method for the analysis of infant cries," in *Proceedings IEEE Seventh Symposium on Computer-Based Medical Systems*, June 1994, pp. 223-228.

[9] Y. Mima and K. Arakawa, "Cause estimation of younger babies' cries from the frequency analyses of the voice - classification of hunger, sleepiness, and discomfort," in *International Symposium on Intelligent Signal Processing and Communications, 2006. ISPACS '06*, Dec 2006, pp. 29-32.

[10] Shubham Asthana, N Varma, and Vinay Kumar Mittal, "An investigation into classification of infant cries using modified signal processing methods," in *2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, Feb 2015, pp. 679-684.

[11] Vinay Kumar Mittal, "Discriminating the infant cry sounds due to pain vs. discomfort towards assisted clinical diagnosis," in *7th Workshop on Speech and Language Processing for Assistive Technologies, SLPAT 2016*, San Francisco, USA, 13 September 2016, pp. 37-42.

[12] Vinay Kumar Mittal, "Discriminating features of infant cry acoustic signal towards automated diagnosis of cause of crying," in *10th International Symposium on Chinese Spoken Language Processing (ISCSLP 2016)*, Tianjin, October 17-20 2016.

[13] Vinay Kumar Mittal and B. Yegnanarayana, "Study of characteristics of aperiodicity in Noh voices," *The Journal of the Acoustical Society of America*, vol. 137, no. 6, 2015.

[14] Vinay Kumar Mittal and B Yegnanarayana, "Effect of glottal dynamics in the production of shouted speech," *The Journal of the Acoustical Society of America*, vol. 133, no. 5, pp. 3050-3061, 2013.

[15] Vinay Kumar Mittal and Bayya Yegnanarayana, "Analysis of production characteristics of laughter," *Computer Speech & Language*, vol. 30, no. 1, pp. 99-115, 2015.

[16] V. P. Singh, J. M. S. Rohith and V. K. Mittal, "Preliminary analysis of cough sounds," in *2015 Annual IEEE India Conference (INDICON)*, New Delhi, 2015, pp. 1-6. doi: 10.1109/INDICON.2015.7443512

[17] Shivam Sharma, Shubham Asthana, and VK Mittal, "A database of infant cry sounds to study the likely cause of cry," in *12th International Conference on Natural Language Processing (ICON 2015)*, IIITM-K, Trivandrum, India.

[18] Simon Haykin, An introduction to analog and digital communications, Wiley New York, 1989, ch. xvi, pp. 652.

[19] Alan V Oppenheim, Ronald W Schafer, John R Buck, and et al, Discretetime signal processing, Prentice-hall Englewood Cliffs, 1989, vol. 2, ch. 3, pp. 71.

[20] Alain de Cheveigne, "Yin, a fundamental frequency estimator for speech and music," *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 19171930, April 2002.

[21] S. Sharma, P. R. Myakala, R. Nalumachu, SV Gangashetty and V. K. Mittal, "A Study on Acoustic Features of Infant Cry Signal for Different Causes of Crying", 3^{rd} *international workshop on Affective Social Multimedia Computing (ASMMC 2017)*, Sweden, Aug. 2017.

[22] S. Sharma and V. K. Mittal, "A Qualitative Assessment of Different Sound Types of an Infant Cry", *4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON 2017)*, India, Oct. 2017.