

# Newborn infant's cry analysis

Anshu Chittora<sup>1</sup> · Hemant A. Patil<sup>1</sup>

Received: 16 May 2016 / Accepted: 4 October 2016 / Published online: 19 October 2016  
© Springer Science+Business Media New York 2016

**Abstract** The very first cry or the birth cry of an infant carries significant information about the health of an infant and hence, it is considered as the vital parameter in deciding the Apgar count. As an infant grows, the cry acoustics changes with the integration of vocal tract system. Infants are found to produce many sounds apart from crying, which reflect the learning mechanism of the infants of the language spoken in his or her surroundings or the environment. Along with this, infants who have distinct cry sounds or who require large amount of stimulation to produce a cry, are found to be at risk of sudden infant death syndrome (SIDS) or possible neurological disorders. In this paper, newborn infant cries are analyzed using features derived from fundamental frequency ( $F_0$ ) contour or pitch contour, energy of the cry signal in different frequency subbands and unvoicing present in the infant's cry. For the extraction of fundamental frequency, modified autocorrelation method is used and shown to perform better than traditional autocorrelation-based method. To identify the significance of these features in identifying the reason of crying, ANOVA analysis is applied on these features. It is observed that the  $F_0$  features are not of significance in the newborn cry analysis and presence of unvoicing in the infant's cry varies with the maturity of central nervous system (CNS) and is a discriminative feature of prime importance in newborn's cry analysis. In birth cries, the

mean percentage of unvoicing is 84.4 % which drops to 67.7 % in normal infants (20 days–3 months). Birth cry analysis shows that there is very less voicing and hence, less vibration of the vocal folds.

**Keywords** Infant's cry · Newborn · Fundamental frequency ( $F_0$ ) · ANOVA

## 1 Introduction

Infant cry analysis is a multidisciplinary area, where contributions have been made by the paediatrics, linguists, psychologists, neuroscientists and the engineers. Though, most of the contributions in this field are from the domain of paediatrics, now-a-days interest is taken by researchers from other domains for infant cry research. Infant cry analysis is necessary to understand the needs of the infants and identifying the pathological infants in the initial stages so that they can be treated in the initial stages of disease and can be protected from possible temporary or permanent disorders. In newborns, the cry characteristics, such as, kitten-like cry is an indicator of the possibility of infant suffering from genetic disorders. Similarly, a hoarse cry is an indicator of cramped muscles. Another important aspect in infant cry analysis is the identification of infants who at risk of sudden infant death syndrome (SIDS) (Corwin et al. 1995). SIDS is the condition in which an infant die all of a sudden and the reason of death remains unidentified even after autopsy. Hence, research in infant cry analysis may prove helpful in developing some applications or devices which can monitor the activities of the infants and help the parents.

The work done in infant cry analysis is mostly towards the analysis and classification of cry types. Cry has been

---

✉ Anshu Chittora  
anshu\_chittora@daiict.ac.in  
Hemant A. Patil  
hemant\_patil@daiict.ac.in

<sup>1</sup> Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), Infocity, Gandhinagar, Gujarat, India

divided in different cry types such as hunger, pain, pleasure and birth cries. Estimation of fundamental frequency ( $F_0$ ) for infant cry signal is proposed in (Petroni et al. 1994). Another method, which was used for  $F_0$  estimation, is average magnitude difference function (AMDF) along with simple inverse filter tracking (Manfredi Claudia et al. 2006). Research has been done in the area of classifying normal infant's cries from the pathological infant's cries (Chittora and Patil 2015a). Most of the work is done towards classifying normal infant's cries from the cries of deaf infant/infant with hearing disorder. The spectral feature set, namely, Mel frequency cepstral coefficients (MFCC) has been used as a state-of-the-art feature set for the classification task with various classifiers (Garcia and Garcia 2003; Reyes Galaviz Orion Fausto et al. 2008). Another feature set used for the normal and deaf babies classification is short-time Fourier transform (STFT) features with time delay neural networks (TDNN), general regression neural networks (GRNN) and multi layer perceptron (MLP) neural networks (Hariharan et al. 2012). Three class classifications is performed for classification of normal, deaf and asphyxiated infants using features such as MFCC and wavelet-based features (Saraswathy et al. 2012). Classification of normal and asphyxia is also attempted using MFCC features in Ali et al. (2012). Classification of asthma and HIE infant cries is reported in Chittora and Patil (2014, 2015b). In (Lederman Dror 2002), the work is done to classify normal and infants with cleft palate, preterm infants and sick infants (cri-du-chat and down's syndrome) using MFCC, linear prediction coefficients (LPC), linear prediction cepstral coefficients (LPCC) and fundamental frequency ( $F_0$ ) features using hidden Markov models (HMM). MFCC feature set has also been used to analyze cries of infants suffering from hypothyroidism (Zabidi et al. 2009). In last few decades, attempts have been made to classify and analyze different infant cry types. The cry types defined by several researchers for infants are hunger, pain, pleasure, discomfort, fear, anger and birth cry. Classification of fear, anger and pain cries using MFCC features is reported in Petroni et al. (1995, 2009). Hunger vs. no hunger and pain vs. no pain cries are classified using MFCC feature set with support vector machines (SVM) classifier and NN ensembles (Barajas-Montiel and Reyes-Garcia 2005; Singh et al. 2013). Analysis of pain and manipulation cries (cry during cloth changing) is performed using pitch ( $F_0$ ), formant frequencies  $F_1$ ,  $F_2$  and  $F_3$  (Baeck and Souza 2001).

Some researchers have worked in the analysis of first cry of the infants. Most of the work in this direction is done by the medical practitioners and researchers. In (Nicollas et al. 2012) authors have used larynx of two dead newborns to generate sounds by applying air pressure. Their finding shows that the role of larynx is same as excised organ, free

of neurologic control. Their role in first cry is not to vibrate by themselves, however, to generate aerodynamic perturbations generating supraglottic vibrations. Complex interactions are responsible for the nonlinear phenomenon found in first cry signal. Neurological control and regulation is absent in first cry. In another study, researchers have used the newborns' cries to find out the effect of prenatal exposure to cocaine (Bauer et al. 1994). In this paper, distinction between first cry and other cry types is reported using different features and effectiveness of these features in infant development is presented. In our earlier paper, we have shown the importance of the feature unvoicing percentage in the infant cry for the study of infant cry pathologies and development (Chittora and Patil 2015c). However, along with this feature, other features are used in this paper and found useful in the newborn infant cry analysis.

The rest of the paper is organized as follows: estimation of fundamental frequency ( $F_0$ ) using modified autocorrelation method is presented in Sect. 2. Feature extraction and experimental results are explained in Sects. 3 and 4, respectively. Finally, the paper is summarized in Sect. 6 along with future directions.

## 2 Fundamental frequency ( $F_0$ ) estimation using modified autocorrelation method

The autocorrelation method of the pitch estimation is widely used for pitch estimation in speech-related applications (Rabiner 1977). In autocorrelation method of pitch estimation, the speech signal is divided in small frames because the speech is a non-stationary signal. For a small frame of speech such as 20–30 ms [comprising of 2–3 pitch periods ( $T_0$ )], after pre-processing which includes passing the signal through a lowpass filter, autocorrelation is found. Periodicity which is observed in the periodic signal is also observed in its autocorrelation function. The autocorrelation function is symmetric, distance between two highest peaks is calculated which is equal to the pitch period ( $T_0$ ) of the signal. Autocorrelation function method of  $F_0$  estimation does not work well for infant cry analysis because for noisy infant cry signals, sometimes false spurious peaks are present, which gives misleading false peaks and thereby, high frequency values. In this paper, fundamental frequency ( $F_0$ ) contour is estimated using modified autocorrelation method. In the pre-processing stage, the infant cry signal is passed through a 4th order Butterworth low-pass filter with a cutoff frequency of 1 kHz, in order to remove high frequency harmonics present in the signal. The filtered signal is then segmented in small frames of duration 30 ms with an overlap of 50 %. On each of the cry signal frame, modified autocorrelation method is applied

and peaks corresponding to the pitch values are identified and pitch is estimated. In modified autocorrelation method of pitch extraction, the signal is clipped by a reference level  $C_L$ . The clipping level  $C_L$  is chosen as the 25 % of the maximum peak sample values. The resulting signal is given by:

$$y(n) = clc[x(n)] = \begin{cases} (x(n) - C_L), & x(n) \geq C_L \\ 0, & |x(n)| < C_L \\ (x(n) + C_L), & x(n) \leq -C_L. \end{cases}$$

For the clipped signal  $y(n)$ , the autocorrelation function is found using the formula:

$$R'(m) = \sum_{n=0}^{N-1-m} y(n) \cdot y(n + m), \quad 0 \leq m \leq M_0,$$

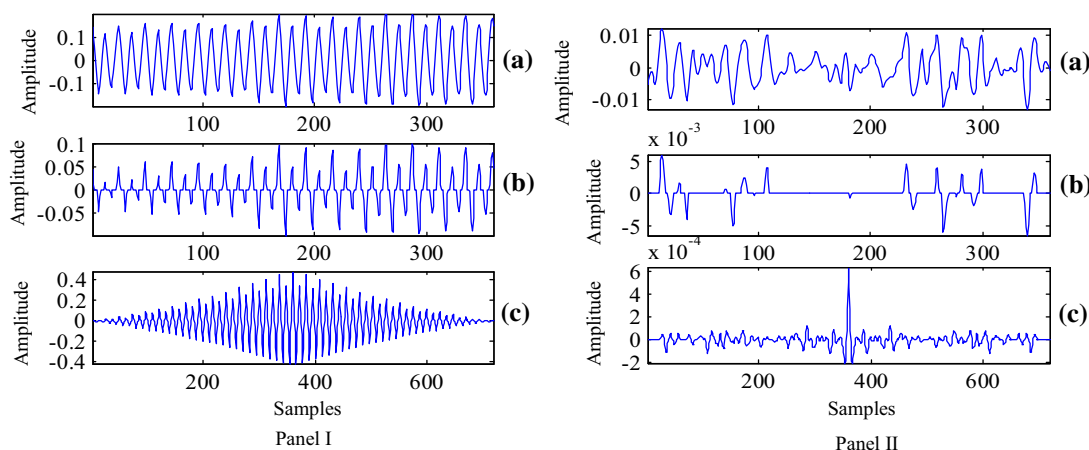
where  $N$  is the length of the sequence,  $M_0$  is the number of autocorrelation points to be computed,  $m$  is lag or delay. Clipping of the signal removes the added noise effects and hence, it performs better than autocorrelation method of pitch estimation. From the autocorrelation function applied on clipped signal, the peaks are identified. The difference of the peak locations gives the estimate of the pitch or  $F_0$  of the signal. The examples of the modified autocorrelation method for pitch extraction applied to voiced and unvoiced segments of the cry signal are shown in Fig. 1. We can observe that for unvoiced segments, the autocorrelation function have very less number of peaks and thus, the segments which have less than 6 number of peaks are taken as unvoiced segments and pitch is taken as zero for them.

In Fig. 1, the modified autocorrelation method is illustrated for the voiced and unvoiced segments. In the proposed method, clipping level was suggested as 64 % of the maximum peak amplitude. In case of infant cry signals, it was observed by intensive computer simulation that

keeping such a high threshold for clipping is removing most of the peaks of the signal, thereby does not work for pitch ( $F_0$ ) estimation. By iterative method, we decided the threshold for clipping as 25 % and this is found to give best results for  $F_0$  estimation. To compare the performance of the  $F_0$  extraction with the standard autocorrelation method, spectrogram is used. In infants, reference glottal flow waveform for comparing the performance of the  $F_0$  methods is not available. The glottal flow waveform cannot be acquired from the infants by non-invasive methods and hence, it limits the availability of the glottal flow waveform for infants. Thus, to compare the performances of the two  $F_0$  estimation algorithms, we used spectrogram. If the estimated harmonics match with the harmonics present in the spectrogram, we can say that the algorithm is better. This decision is made after observing the matching of harmonics with spectrogram for many infant cry samples in order to have decision which is statistically significant. From Fig. 2, it can be observed that the modified autocorrelation-based method of  $F_0$  extraction works better than state-of-the-art method, i.e., autocorrelation method of  $F_0$  extraction.

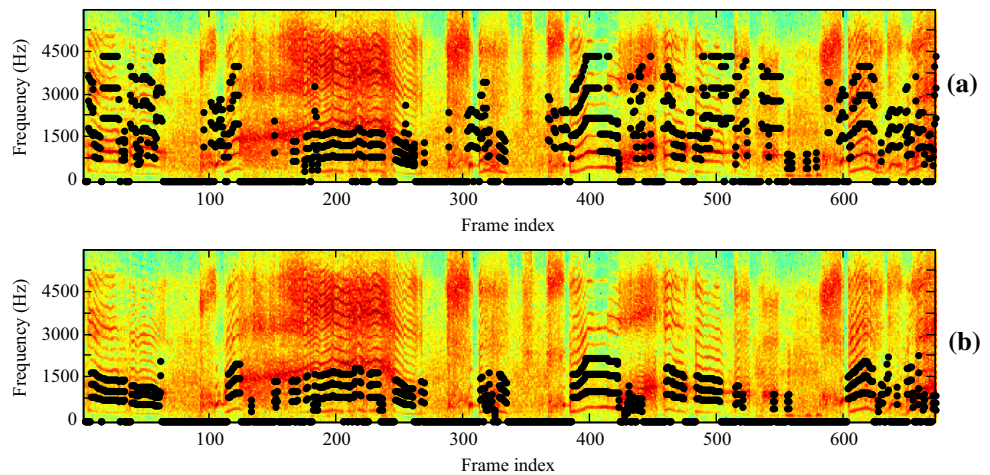
### 3 Feature extraction

Database: In this study, infant cry data was collected from three hospitals of Visakhapatnam, India. Data was recorded with a handheld Cenix recorder (Model: VR-P2340) with external microphone with sampling frequency of 12 kHz and 12-bit PCM quantization (Buddha and Patil 2007). The pain cries of normal infants were recorded during vaccination, birth cries were recorded the nursing home, hunger cries are recorded when the infant cries because of hunger (duration of last feed is used as an indicator for the identification of hunger cry), cries while passing the urine



**Fig. 1** Modified autocorrelation algorithm for  $F_0$  extraction for Panel I: voiced segments and Panel II: unvoiced frame. In all the subfigures **a** time-domain signal, **b** center clipped signal of **a** and **c** autocorrelation function of **b**

**Fig. 2** Comparison of pitch ( $F_0$ ) extraction methods  
**a** autocorrelation method and  
**b** modified autocorrelation method



**Table 1** Corpus statistics for infant cry analysis

Class	Number of infants
Newborn full term normal birth	45
Newborn pre-term	36

was recorded when infant passed the urine in routine course or while bathing. From one infant, sometimes more than one cry is also recorded. The duration of the cries varies from 30 to 50 s. The corpus statistics are given in Table 1. From this corpus, cry types are separated as shown in Table 2 for different reasons of crying and age. Most of infants considered in this study are below 1 month of age.

It is known that our ears are sensitive to two parameters, namely, loudness and pitch ( $F_0$ ). Loudness is associated with the amplitude of the signal, it is a perceptual feature which is recently found to be associated with the strength of excitation (SoE) (Seshadri and Yegnanarayana 2009). However, pitch ( $F_0$ ) is also a perceptual feature and is associated with the  $F_0$  of the signal. Hence, to extract information of these two parameters, energy and  $F_0$ -related parameters are estimated and using them different cry signals are analyzed. For each of the cry sample,  $F_0$

contour is calculated using the modified autocorrelation method and following features are estimated:

1. Minimum of  $F_0$  contour
2. Maximum of  $F_0$  contour
3. Mean of  $F_0$  contour
4. Median of  $F_0$  contour
5. Normalized energy of the signal ( $E$ )
6. Normalized energy in 0–2 kHz ( $E_1$ )
7. Normalized energy in 2–4 kHz ( $E_2$ )
8. Normalized energy in 4–6 kHz ( $E_3$ )
9. Unvoicing percentage in the total cry (UV ratio)

The normalized energy of the signal is defined as the energy of signal divided by the length of the signal, i.e.,

$$E = \frac{1}{n} |X(\omega)|^2,$$

where  $E$  is the normalized energy,  $n$  is the number of cry segments and  $X(\omega)$  is the short-time Fourier transform (STFT) of the signal. The normalized energy of the signal is calculated for the three sub-bands, namely, (1) 0–2 kHz, (2) 2–4 kHz and (3) 4–6 kHz (because the data is recorded at 12 kHz sampling frequency and hence, the maximum available bandwidth is 6 kHz). The unvoicing regions are

**Table 2** Distribution of cry samples of newborn infant's cries

S. No.	Type of cry	Age after birth	No. of subjects	No. of samples
1.	Full term infant's birth cry (BC)	After birth	16	49
2.	Premature infant's cry (PC)	<20 days (mostly <10 days of age)	7	50
3.	Newborn hunger cry (H)	<10 days age	16	30
4.	Newborn pain cry (P)	<3 months (mostly <1 month)	17 (4 infants (6 cries) are older than 1 month)	19
5.	Newborn cry due to urination (U)	<20 days	12	24
6.	Newborn cry due to wet diaper (W)	<20 days	4	5
7.	Normal after birth (N)	S. No. 3–6 combined		

identified as the segments where the number of peaks in the autocorrelation function is less than 6, thereby giving zero pitch frequency. The sum of frames with zero pitch values divided by the total number of frames present in the cry is considered as the unvoicing ratio of the cry signal.

Different cry types defined in Table 2 are analyzed using these features and analysis of variance (ANOVA) analysis is used to find the significance of these features in various infant cry types. The analysis and the results are given in the next Section.

### 4 Experimental results

Different cry features are analyzed for the reasons of crying of an infant for following cases:

1. Full term birth cry vs. premature newborn’s cry
2. Full term birth cry vs. newborn’s pain cry
3. Full term birth cry vs. newborn’s hunger cry
4. Newborn’s pain cry vs. newborn’s hunger cry
5. Newborn’s pain cry vs. newborn’s cry due to wet diaper
6. Newborn’s pain cry vs. newborn’s cry during passing the urine
7. Newborn’s cry due to wet diaper vs. newborn’s cry during passing the urine
8. Newborn’s hunger cry vs. newborn’s cry due to wet diaper
9. Newborn’s hunger cry vs. newborn’s cry during passing the urine
10. Newborn’s birth cry vs. newborn’s other reasons of crying (hunger\wet diaper\passing urine\pain).

The mean values of the above features along with the standard deviation are given in Table 3.

For the simplification purpose, the analysis is taken separately for the  $F_0$ -based features and remaining features.

#### 4.1 Analysis using fundamental frequency ( $F_0$ )-based features

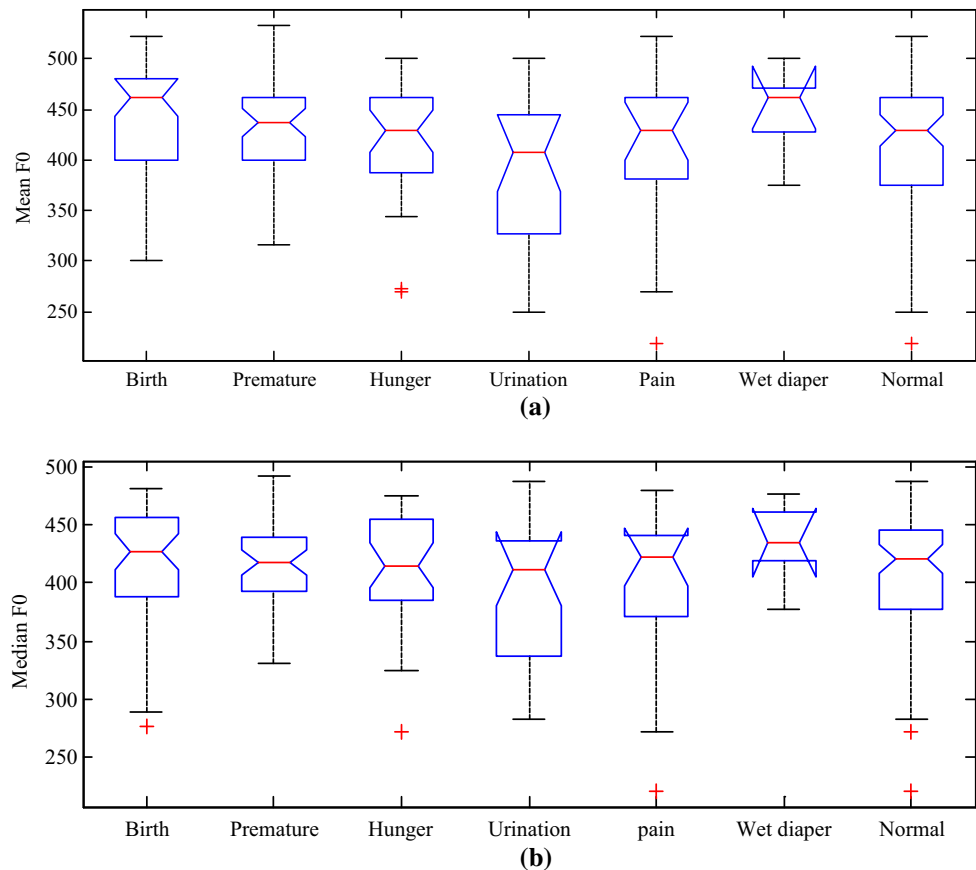
From Table 3 and Fig. 3, it can be observed that the minimum  $F_0$ , maximum  $F_0$  and median of  $F_0$  are almost similar in all the cases. Thus, these features cannot be used to characterize or discriminate a particular cry type. However, mean  $F_0$  feature is showing differences in some cry types such as newborn’s birth cry has mean  $F_0$  of 436.22 Hz while this parameter is 411.15 Hz for the normal newborn’s cry. Differences in the hunger cry and pain cries of the newborns are also observed. In hunger cries, the mean value of the  $F_0$  is  $425.19 \pm 55$  Hz, mean  $F_0$  is  $387.48 \pm 72$  Hz for urination cries. Significant differences are not found in the different features of  $F_0$ , based on the

**Table 3** Mean values of the features for different infant cry types

S. No.	Type of cry signal	Min ( $F_0$ )	Max ( $F_0$ )	Mean ( $F_0$ )	Med. ( $F_0$ )	E	E1	E2	E3	UV ratio
1.	BC	167.86 ± 18	561.16 ± 38	436.22 ± 57	417.70 ± 46	0.487 ± 0.2	0.167 ± 0.05	0.245 ± 0.12	0.078 ± 0.04	0.844 ± 0.07
2.	PC	169.80 ± 20	564.95 ± 16	429.31 ± 46	416.42 ± 35	0.344 ± 0.17	0.169 ± 0.07	0.119 ± 0.07	0.057 ± 0.04	0.687 ± 0.14
3.	H	175.18 ± 36	550.36 ± 40	425.19 ± 55	413.30 ± 48	0.416 ± 0.17	0.199 ± 0.04	0.1543 ± 0.09	0.0657 ± 0.05	0.693 ± 0.14
4.	P	160.00 ± 11	556.28 ± 55	409.06 ± 76	397.38 ± 66	0.636 ± 0.23	0.238 ± 0.06	0.279 ± 0.13	0.123 ± 0.07	0.688 ± 0.14
5.	U	173.91 ± 34	552.62 ± 48	387.48 ± 74	388.45 ± 63	0.435 ± 0.15	0.198 ± 0.05	0.168 ± 0.08	0.070 ± 0.04	0.674 ± 0.21
6.	W	166.14 ± 15	571.43 ± 0	448.50 ± 41	435.30 ± 33	0.663 ± 0.09	0.261 ± 0.03	0.306 ± 0.07	0.100 ± 0.02	0.548 ± 0.11
7.	N	170.51 ± 30	553.85 ± 45	411.15 ± 67	403.19 ± 57	0.491 ± 0.20	0.212 ± 0.05	0.198 ± 0.11	0.083 ± 0.05	0.677 ± 0.17

BC Birth cry; PC Pre-mature infant’s cry; H Hunger cry; P Pain cry; U Urination cry; W Wet diaper cry; N Normal Infant’s cry

**Fig. 3** Boxplot for the  $F_0$  features **a** mean  $F_0$  and **b** median  $F_0$



**Table 4** ANOVA analysis of the newborn infant's cry

S. No.	Class	Min $F_0$	Max $F_0$	Mean $F_0$	Med. $F_0$	$E$	$E1$	$E2$	$E3$	UV ratio
1.	Full term birth vs. premature	0.62	0.52	0.516	0.88	<b>0.0004</b>	0.85	<b>4.08e-8</b>	<b>0.0247</b>	<b>6.21e-10</b>
2.	Full term birth vs. pain	0.0926	0.68	0.117	0.16	<b>0.0129</b>	<b>6.84e-05</b>	0.34	<b>0.0016</b>	<b>2.29e-7</b>
3.	Full term birth vs. hunger	0.24	0.24	0.41	0.69	0.1191	<b>0.017</b>	<b>0.0015</b>	0.256	<b>5.09e-8</b>
4.	Pain vs. hunger	0.0881	0.668	0.402	0.33	<b>0.0005</b>	<b>0.021</b>	<b>0.0004</b>	<b>0.0027</b>	0.91
5.	Pain vs. wet diaper	0.3463	0.55	0.28	0.23	0.81	0.46	0.66	0.4962	0.06
6.	Pain vs. urination	0.107	0.82	0.35	0.65	<b>0.0015</b>	<b>0.03</b>	<b>0.0016</b>	<b>0.0042</b>	0.809
7.	Wet diaper vs. urination	0.6331	0.404	0.0899	0.127	<b>0.0036</b>	<b>0.0195</b>	<b>0.0023</b>	0.1372	0.22
8.	Hunger vs. wet diaper	0.59	0.26	0.3922	0.34	<b>0.0047</b>	<b>0.0103</b>	<b>0.0029</b>	0.18	<b>0.04</b>
9.	Hunger vs. urination	0.8976	0.85	<b>0.039</b>	0.11	0.6792	0.97	0.58	0.70	0.70
10.	Newborn birth vs. normal	0.589	0.35	<b>0.035</b>	0.14	0.919	<b>4.29e-5</b>	<b>0.037</b>	0.57	<b>2.18e-9</b>

Bold values indicate that the feature is of prime importance in distinguishing the two cry types considered

reason of crying, except in the two cases mentioned above. In the birth cries as well, these features do not change with the gestation age (GA). These parameters are almost similar for normal full term as well as for premature babies. In newborn cries, mean  $F_0$  lies in the range of 400–600 Hz (Michelson and Michelson 1999). Thus, the results obtained here are in agreement with the previous studies.

The ANOVA analysis of the parameters derived from the  $F_0$  contour also suggests the similar results. The results of ANOVA analysis are given in Table 4 for all the

features. Here, we have considered 95 % confidence interval in ANOVA analysis which means features which give  $p$  value less than 0.05 are of significance in the analysis of those particular cry types.

#### 4.2 Analysis using normalized energy-based features

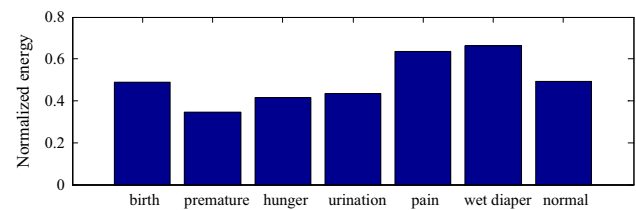
Analysis is done for various cry types based on normalized energy-based features. The mean values and standard

deviations of the features are also mentioned in the Table 3. From the Table 3, bar plots are drawn for the energy features to illustrate their importance in the cry of an infant.

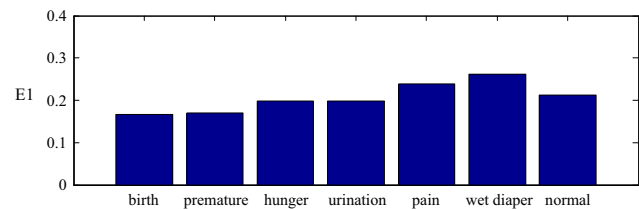
From Fig. 4, it can be observed that normalized energy of the pain and wet diaper cries are higher than other cry types. The energy is lowest in the premature infant's cries. The energy of full-term birth cries is higher than the premature infant's cries. Comparing the distribution of the energy of the cry signals in the three frequency bands as shown in Figs. 5, 6 and 7, we can observe that the pain cries and wet diaper cries have highest energy in all the sub-bands. Moreover, most of the energy lies in the 2–4 kHz sub-band in all cries. In premature infants, distribution of energy is higher in lower frequency bands compared to normal full term infant's birth cries (as shown in Fig. 7), where the distribution of energy is higher in the mid frequency band (2–4 kHz) (as shown in Fig. 8). In hunger and urination cries, distribution of energy is more in lower frequency bands (0–2 kHz) compared to pain and wet diaper cries where energy in 2–4 kHz band is higher. In the high frequency bands (4–6 kHz), the distribution of energy is very low for infant's cries except for pain and wet diaper cries as shown in Fig. 7.

Results of ANOVA analysis are shown in Table 4. It can be observed that the normal infant's birth cries are distinct from the premature infant's cries. Because of higher energy of normal full term infants, we can distinguish their cries from premature infants, who have low energy in the cries. The reason of cry can also be identified from the energy feature. Hunger cries are found distinct from the pain cries and wet diaper cries are found different from the crying during the passing of urine. In case of birth cry and premature infants' cries, it is observed that the energy difference is very high and this result in identification of the cries by auditory analysis as well. The differences in the two cry patterns are there in the mid- frequency bands. In the band 2–4 kHz, the energy of the birth cry is higher than the pre-mature infant's cry and in other bands, the distribution of energy is same for both the cries. Birth cries of normal full term infants and pain cries are characterized by high energy of the signal as shown in Fig. 8a. ANOVA analysis in the three frequency bands show that the two cry can be characterized by the distribution of energy in the low and high frequency bands. The energy is high in low and high frequency bands in pain cries compared to birth cries as shown in Fig. 8b, d.

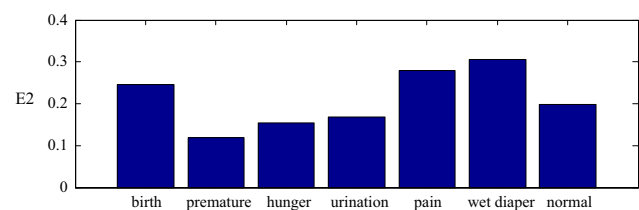
Analysis of hunger, pain, wet diaper and urination cries shows that distribution of energy is similar in hunger and urination cries as well as in pain and wet diaper cries. These two groups of the cries are distinct from each other on the basis of total normalized energy as well as energy in their respective bands. However, it is difficult to



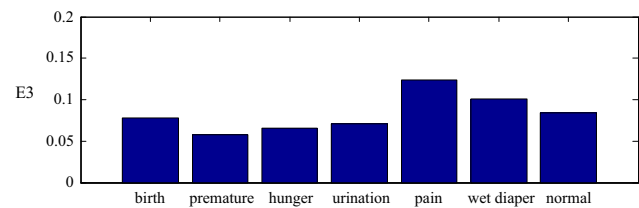
**Fig. 4** Bar plot of mean values of normalized energy values for different cry types. Y-axis represents the normalized energy of the signal



**Fig. 5** Bar plot of mean values of E1 for different types of cries. Y-axis represents the normalized value of feature E1

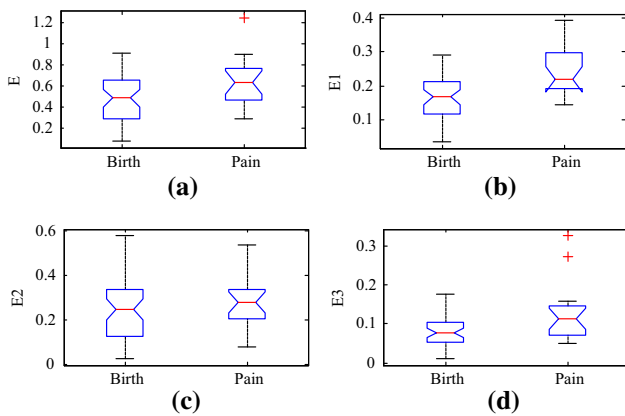


**Fig. 6** Bar plot of mean values of E2 for different types of cries. Y-axis represents the normalized value of feature E2

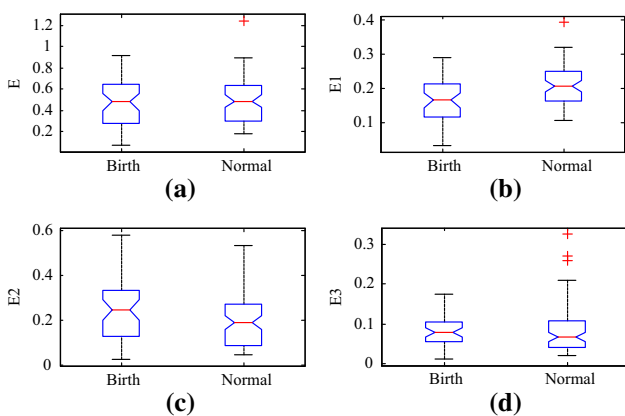


**Fig. 7** Bar plot of mean values of E3 for different types of cries. Y-axis represents the normalized value of feature E3

characterize differences in hunger and urination cries using energy-based features. Similar is the case for the classification of pain and wet diaper cries, where the energy in all the bands is almost similar irrespective of the reason of crying. Normal full-term birth cries are different from the other reasons of crying such as hunger, pain, wet diaper and urination named here as normal cry, on the basis of  $E1$  and  $E2$ . In birth cries,  $E2$  is higher than the other reasons of crying. However, in normal crying (due to other reasons of crying) energy  $E1$  is higher than birth cries of full term healthy infants as shown in Fig. 9.



**Fig. 8** Boxplots of normalized **a**  $E$ , **b**  $E1$ , **c**  $E2$  and **d**  $E3$  for birth and pain cries

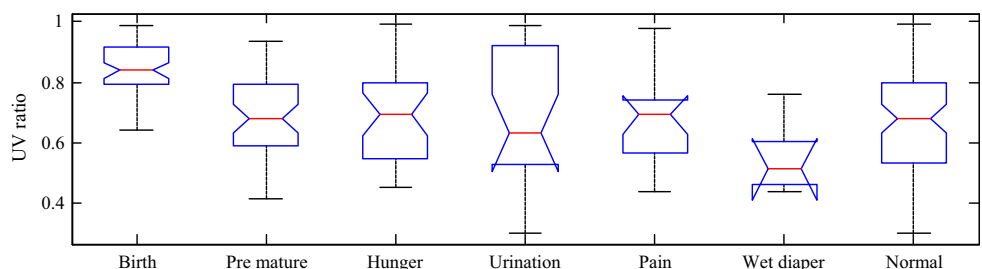


**Fig. 9** Boxplots of normalized **a**  $E$ , **b**  $E1$ , **c**  $E2$  and **d**  $E3$  for birth and normal cries

### 4.3 Analysis using unvoicing ratio of the cry

From Fig. 10 and Table 3, we can observe that the birth cries are characterized by very high unvoicing ratio. Compared to cries due to hunger, pain, wet diaper and urination, this higher unvoicing ratio makes birth cries distinct from other cry types. This feature is found to be useful in classifying the reason of crying also where energy-based features are not working. Similar energy-level of cries can be classified according to the ratio of crying present in the cry. Pain and wet diaper cries which

**Fig. 10** Boxplot for the UV ratio in the cries



have similar energy in all the frequency bands can be distinguished by using UV ratio analysis. In pain cries, UV ratio is higher than the wet diaper cries. Similarly, between wet diaper and hunger cries, hunger cries are found to have more unvoicing and can be distinguished from cries due to wet diaper.

## 5 Summary and discussions

In this study, newborn infants cries are analyzed for the various reasons of crying such as hunger, pain, wet diaper and while passing the urine. These are the various reasons of crying in a newborn. For the analysis of the cries, features used are the  $F_0$ -based features, energy-based features and the unvoicing ratio of the cry segments. Some important results from the above analysis are as follows:

1. Birth cry can be characterized by high energy and high unvoicing ratio. The reason for this is, as soon as the newborn come to the external world from the mother’s warm womb; it is his or her response to the external stimulation. At birth, there is poor regulation of central nervous system (CNS) over vocal folds working. At birth cry, lungs open up for the first time and breaths air instead of sack fluid (Lester Barry 1985).
2. Most of the energy in birth cry is located in the frequency band 2–4 kHz. However, normal infant’s cry is having its maximum distribution of energy in 0–2 kHz (i.e., normal, hunger, urinating). Pain cry is also having the same characteristics of having higher  $E2$  than  $E1$ .
3. Compared to other infant cry types, pain cries and wet diaper cries have higher energy distribution in 4–6 kHz frequency range. Higher energy in higher frequency ranges asks for the attention of the care taker and informs that a quick action is required. In the other words, higher frequency content in the cry reflects urgency of the attention and discomfort to the infants.
4. Characteristics of hunger cry and cry during passing the urine found to be similar on all the parameters. Similarly, pain cries and wet diaper cries have similar characteristics.



5. Hunger cry and cry during passing the urine can be distinguished from each other using mean  $F_0$  parameter. Remaining parameters are same for them.
6. Unvoicing ratio in infants is an indicator of maturity of infant's vocal production system. In birth cry, high unvoicing indicate that, in birth cry vocal folds movement is very irregular which results in poor voiced quality of the cry. With the production of the birth cry, infant's neural system integrates and within few days cries become rhythmic.
7. Wet diaper cries can be distinguished from the pain cries based on the feature of unvoicing ratio. In pain cries, it is found to be higher than wet diaper cries.
8. Mean  $F_0$  in newborn birth cries is higher than the normal infant's cries. There are no significant differences in the birth cries of newborns and those of premature infants cries. This indicates that until infant achieves a minimum gestation age (GA), vocal folds do not vibrate to produce voiced cry sounds.
9.  $F_0$ - related features are not useful in identifying the reason of crying in newborns, though it is a useful parameter in infant's (more than 1 month of age) cry analysis for understanding the reason of cry.

In future, authors would like to come up with classification of infant cries using these features. In addition, we would like to direct our efforts towards finding differences in male and female infant cries.

**Acknowledgments** Authors would like to thank DA-IICT, Gandhinagar, India, for providing necessary resources for this study. We also like to thank Department of Electronics and Information Technology (DeitY) and Department of Science and Technology (DST), Government of India, New Delhi, India for partial support in providing resources for carrying out this research work. We acknowledge the help given by the members of Speech Research Lab, DA-IICT, Gandhinagar.

## References

- Ali, M. Z. M., Mansor, W., Lee, Y. K., & Zabidi, A. (2012). Asphyxiated infant cry classification using simulink model. In *IEEE 8th International Colloquium on Signal Processing and its Applications*, Malacca, Malaysia, 491–494.
- Baek, H. E., & Souza, M. N. (2001). Study of acoustic features of newborn cries that correlate with the context. In *Proceedings of 23rd IEEE Annual International Conference of EMBS*, Istanbul, 2174–2177.
- Barajas-Montiel, S. E., & Reyes-Garcia, C. A. (2005). Identifying pain and hunger in infant cry with classifiers ensembles. *International Conference on Intelligent Agents, Web Technologies and Internet Commerce, Vienna, 2*, 770–775.
- Bauer, H. R., & Zimmerman, L. (1994). Newborn human cries: Prenatal cocaine exposed and nonexposed. *The Journal of the Acoustical Society of America*, 95(5), 3013.
- Buddha, N., & Patil, H. A. (2007). Corpora for analysis of infant cry. In *International Conference on Speech Databases and Assessments, Oriental COCODA*, Hanoi, Vietnam, 43–48.
- Chittora, A., & Patil, H. A. (2014). Use of glottal inverse filtering for asthma and HIE infant cries classification. In *International Conference on Asian Language Processing (IALP)* (pp. 158–161). Kuching, Sarawak.
- Chittora, A., & Patil, H. A. (2015a). Significance of unvoiced segments and fundamental frequency for infant cry analysis. In P. Kral & V. Matousek (Eds.), *Text, Speech and Dialogue (TSD)*, New York: Springer. LNAI 9302, pp. 273–281.
- Chittora, A., & Patil, H. A. (2015b). Classification of normal and pathological infant cries using bispectrum features. In *23rd European Signal Processing Conference (EUSIPCO)*(pp. 639–643). Nice, France.
- Chittora, A., & Patil, H. A. (2015c). Modified group delay-based features for Asthma and HIE infant cries classification. In P. Kral & V. Matousek (Eds.), *18th International Conference on Text, Speech and Dialogue (TSD)* (pp. 595–602), Lecture Notes in Artificial Intelligence (LNAI) New York: Springer.
- Corwin, M. J., et al. (1995). Newborn acoustic cry characteristics of infants subsequently dying of sudden infant death syndrome. *Pediatrics*, 96(1), 73–77.
- Garcia, J. O., & Garcia, C. A. R. (2003). Mel-frequency cepstrum coefficients extraction from infant cry for classification of normal and pathological cry with feedforward neural networks. In *Proceedings of the International Joint Conference on Neural Networks* (pp. 3140–3145). Portland.
- Hariharan, M., Sindhu, R., & Yaacob, S. (2012). Normal and hypoacoustic infant cry signal classification using time-frequency analysis and general regression neural networks. *Journal of Computer Methods and Programs in Biomedicine, Amsterdam: Elsevier*, 108(2), 559–569.
- Lederman, D. (2002). Automatic classification of infants' cry. Masters Thesis, Ben-Gurion University of the Negev.
- Lester, B. M. (1985). Introduction- There's more to crying than meets the ear. In B. M. Lester & Z. C. F. Boukydis (Eds.), *Infant crying- Theoretical and Research Perspective*. New York: Plenum Press.
- Manfredi, C., Tocchioni, V., & Bocchi, L. (2006). A robust tool for newborn infant cry analysis. In *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)* (pp. 509–512). New York.
- Michelson, K., & Michelson, O. (1999). Phonation in newborn. *International Journal of Pediatric Otorhinolaryngology*, 49(1), S297–S301.
- Nicollas, R., Giordano, J., & Ouaknine, M. (2012). The very first cry: A multidisciplinary approach toward a model. In *Annals of Otolaryngology, Rhinology and Laryngology*, Annals Pub. Co., 121(12), 821–826.
- Petroni, M., Malowany, M. E., Johnston, C. C., & Stevens, B. J. (1994). A Crosscorrelation based method for improved visualization of infant cry vocalizations. In *Canadian Conference on Electrical and Computer Engineering* (pp. 25–28).
- Petroni, M., Malowany, M. E., Johnston, C. C., & Stevens, B. J. (1995). A comparison of neural network architectures for the classification of three types of infant cry vocalizations. In *IEEE 17th Annual Conference Engineering in Medicine and Biology Society* (pp. 821–822). Canada.
- Petroni, M., Malowany, A. S., Johnston, C. C., & Stevens, B. J. (2009). Classification of infant cry vocalizations using artificial neural networks, In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Georgia, USA, Vol. 5, 3475–3478.
- Rabiner, L. R. (1977). On the use of autocorrelation analysis for pitch detection. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 25(1), 24–33.
- Reyes Galaviz O. F., Cano-Ortiz S. D., & Reyes-Garcia C. A. (2008). Evolutionary-neural system to classify infant cry units for

- pathologies identification in recently born babies. In *7th Mexican International Conference on Artificial Intelligence* (pp. 330–335).
- Saraswathy, J., Hariharan, M., Vijejan, V., Yaacob, S., & Khairunizam, W. (2012). Performance comparison of Daubechies wavelet family in infant cry classification. In *8th International Colloquium on Signal Processing and its Applications* (pp. 451–455).
- Seshadri, G., & Yegnanarayana, B. (2009). Perceived loudness of speech based on the characteristics of glottal excitation source. *The Journal of the Acoustical Society of America*, 126(4), 2061–2071.
- Singh, A. K., Mukhopadhyay, J., Kumar, S. S., & Rao, K. S. (2013). Infant cry recognition using excitation source features. In *IEEE India conference (INDICON)* (pp. 1–5). Mumbai, India.
- Zabidi, A., Mansor, W., Khuan, L. Y., Sahak, R., & Rahman, F. Y. (2009). Mel-frequency cepstrum coefficient analysis of infant cry with hypothyroidism. In *5th International Colloquium on Signal Processing & Its Applications (CSPA)* (pp. 204–208).