Identification of Diseases in Newborns Using Advanced Acoustic Features of Cry Signals

Yasmina Kheddache*, Chakib Tadj**

* *Faculty of Science and Technology , Ziane Achour University, 3117 Djelfa, Algeria*

**Department of Electrical Engineering , École de technologie supérieure, H3C 1K3 Montréal (Qc), Canada*

Corresponding Author:

Yasmina Kheddache,

e-mail: yasmina.kheddache.1@etsmtl.net

fax: + (001) 514 396-8684

Abstract

Our challenge in the current study is to extend research on the cries of newborns for the early diagnosis of different pathologies. This paper proposes a recognition system for healthy and pathological cries using a probabilistic neural network classifier. Two different kinds of features have been used to characterize newborn cry signals: 1) acoustic features such as fundamental frequency glide (F_{0glide}) and resonance frequencies dysregulation (RFs_{dys}); 2) conventional features such as mel-frequency cestrum coefficients.

This paper describes the automatic estimation of the proposed characteristics and the performance evaluation of these features in identifying pathological cries. The adopted methods for $F_{0glides}$ and RFs_{dys} estimation are based on the derived function of the F_0 contour and the jump "J" of the RFs between two subsequent tunings, respectively. The database used contains 3250 cry samples of full-term and preterm newborns, and includes healthy and pathologic cries.

The obtained results indicate the important association between the quantified features and some studied pathologies, and also an improvement in the identification of pathologic cries. The best result obtained is 88.71% for the correct identification of health status of preterm newborns, and 82% for the correct identification of full-term infants with a specific disease. We conclude that using the proposed characteristics improves the diagnosis of pathologies in newborns. Moreover, the method applied in the estimation of these characteristics allows us to extend this study to other uninvestigated pathologies.

1. Introduction

The majority of sick babies appear healthy at birth. Thus, early diagnosis of hidden pathologies for a quick and effective treatment during their first week of life is crucial, as it could save the lives of these babies. However, the realization of a systematic neonatal diagnosis procedure for all newborns demands high costs because it involves the participation of numerous health professionals and specialized equipment. Thus, our aim is to develop a low-cost diagnostic system that allows pediatricians to detect pathologies affecting newborns using spontaneous cry signals.

Cry signal analysis is a valuable tool for predicting neonatal diseases. It allows for the recognition of a sick infant when signs of illnesses are absent. In the previous studies of infant cries, two different approaches have been adopted: 1) an automatic recognition that consist of cry classification using advanced signal processing techniques [1, 2, 3, 4, and 5], and 2) a spectral cry analysis based on observing the spectrograms of cry signals and software tools [6,7,8,9, and 10].

However, the cries of newborns provide important acoustic parameters that are not considered while monitoring the first days of infant life and in the standard measurements of the Apgar score (appearance, pulse, grimace, activity, and respiration), which is used to verify a baby's health immediately after birth. Moreover, studies regarding the development of reliable estimation procedures for the most important acoustics characteristics as well as the identification of their pathological markers are scarce.

Unlike previous works, the aim of our study is to propose an automated tool that uses important characteristics of cry signals to support the diagnosis of diseases in newborn infants. These characteristics are described in Section 4.2.1. They include acoustic parameters that qualify the vocal tract, such as mel-frequency cestrum coefficients (MFCCs) and resonance frequencies dysregulation (RFs_{dys}), and the vocal fold, such as fundamental frequency glide (F_{0glide}).

F_{0glide} has been associated with central nervous system diseases, asphyxia, and malformations of the orolaryngeal tract; RFs_{dys} indicate poor neural control of the vocal tract and breathing, and has been associated with hyperbilirubinemia, and prenatal tobacco and cocaine exposure [11]. MFCCs can represent properly various models of cries [2], they allow decoupling between the features of the vocal tract and the features generated by the source of excitation.

To differentiate pathologic cries from healthy ones, we strengthened the previous studies with data based on cry analysis. Hence, the primary purpose of our work is to complete a detailed analysis of acoustic phenomena produced in newborn cries, focusing specifically on the prevalence of F_{0glide} and RFs_{dys}. Moreover, we study the effectiveness of their use as an input of the proposed newborn's pathological cry identification system (PCIS).

We performed specific procedures to estimate the average percentage of F_{0glide}, the rising and falling times of F_{0glide} (T_{glide}), and the percentage of RFs_{dys} in the cry signals of healthy and sick infants. By implementing these procedures as described in Section 4.2.2 and by using the median and interquartile range, we managed to 1) determine the quantitative relationship among these acoustic parameters and the various studied pathologies, 2) facilitate their use in the proposed diagnostic system.

This paper is organized as follows. Section 2 provides an overview of previous works. In Section 3, the details regarding the database used are presented . In Section 4, we explain our methodology for cry signal feature extraction. Section 5 describes the proposed PCIS. In Section 6, we provide a statistical analysis of the estimated characteristics and also an evaluation of the obtained classification results. Finally, the conclusion is presented in Section 7, which also summarizes the suggestions and future directions for further research.

2. Previous studies

Studies regarding neonatal cry had focused primarily on the classification of cries and also on the relationship between the characteristics of cry signals and diseases. These two research directions progressed separately. Hence, the characteristics of cries that are closely related to the pathologies of newborns have not been considered in the classification of cry studies.

In cry recognition studies, different types of classifiers have been used, such as artificial neural networks [12], Gaussian mixture model [13], hidden Markov model [2], probabilistic neural network (PNN) [3], support vector machine [14], and radial basis function networks[15]. The developed automatic recognition systems were applied on cry signals of normal infants and infants suffering from deafness, asphyxia, cleft lip, or cleft palate. The most used features were MFCCs, wavelet packet transform coefficients, and linear predictive coding coefficients.

According to our knowledge, none have considered the acoustic characteristics of cry signals in the classification of pathologic cries.

In this study, we are interested in the early diagnosis of newborn pathologies using the acoustic characteristics of cry signals. Most direct methods that have been addressed in different studies to assess the acoustic properties of cry signals were based on spectrographic analysis using visual techniques [8, 9]. The studied parameters were manually measurable in the spectrogram of the cry signal. These studies were primarily focused on babies at risk with low neurological factors such as prematurity, hyperbilirubinemia, prenatal exposure to drugs, and also neurological damages such as Krabbe's disease, Down syndrome, asphyxia, and meningitis [16]. Several characteristics have been associated with these diseases, such as hyperphonic cries, F_0 irregularity, extremely high-pitched cries, dysphonic cries, change in phonation mode, variability of F_0, glide of F_0, and dysregulation of RFs [7,8,10]. In our previous studies [10,17,18], we adopted an automatic approach for the evaluation of the prevalence of these features. The results obtained indicated that the studied characteristics depend on the pathology itself.

We have also provided the experimental results corresponding to a proposed PCIS to separate pathologic cries from healthy ones, with and without using the studied characteristics such as the average percentage of hyperphonic cries and F_0 irregularity [17]. The results of cry classification were better when these features were used in the PCIS.

3. Database and cry recording

The acoustic analysis presented herein was performed on a real database. The recordings of cry signals were created specifically to study the possibilities of the early diagnosis of various pathologies using spontaneous cries during the first days of infant life. They were performed in the Pediatrics Department at the Sainte-Justine Hospital in Montreal, using a small recorder, at a distance of 10 cm from a baby's mouth with a sampling rate of 44.1 kHz and a resolution of 16 bits. This database is similar to that used in [17], [18], and [19]. A description of the recording sessions was presented in a previous work [13]. The recordings were performed with background noise; they have been performed for healthy and pathological newborns cries for different kinds of cries, such as hunger, sampling blood, and change of diapers. The categories of health conditions considered in our cry database are as follows: healthy, heart problems, neurological disorders, respiratory diseases, and blood abnormalities. The constructed dataset contains 3250

cry samples of 1-s duration from 66 babies. Among them were preterm and full-term newborns, from 1-day to 1-month old. It is distributed by pathologies and gestational age, as illustrated in Table 1. It is noteworthy that the age of the babies does not exceed one month because the infants acquire voluntary control of their vocal tracts beyond this age [10].

All the samples in the database have been used to investigate and analyze the proposed acoustic features in cry signals. However, because more healthy samples are available than pathology samples, and the number of cry samples associated with each of the diseases are different, only part of this database has been used in the study of the newborn's PCIS, as shown in Table 4 in Section 6.2.

Table 1. Studied pathologies for different gestational ages

Gestational age	Pathology	Sample size
Full-term newborn (t)	Healthy	1010
	Hyperbilirubinemia	250
	Vena cava thrombosis	77
	Meningitis	115
	Peritonitis	20
	Asphyxia	190
	Lingual frenum	141
Preterm newborn (P)	Healthy	764
	IUGR- microcephaly (in utero growth retardation)	78
	Tetralogy of Fallot	53
	Gastroschisis	134
	IUGR-asphyxia (intra-uterine growth retardation)	148
	RDS (respiratory distress syndrome)	270

4. Cry signal features extraction

Feature extraction is a crucial step in the classification task; it affects the accuracy and reliability of pattern recognition. In our work, new features were extracted in addition to MFCCs. They were used to characterize different models of pathological cries and also cries from healthy newborns. These new high level features are the F_{0glide}, T_{glide}, and RFs_{dys}.

4.1 MFCC

MFCCs are the most used feature coefficients in automatic infant cry classification [1, 2, 4]. Considering the statistical stationary of cry signals in short periods of time, MFCCs were extracted using a frame duration of 23.2 ms (1024 samples) with 50% recovery and a sampling rate of 44.1 kHz.

For every cry signal, we extracted 13 MFCC parameters for each frame of 23.2 ms. We obtained an MFCC matrix of 13 lines × N columns. N corresponds to the total number of frames in the whole cry signal.

The calculation of MFCCs was performed as follows:

- Perform a glottal inverse filtering to attenuate the influence of the vocal tract.

- Multiply each frame by the Hamming window.

- Estimate the power spectrum sequence using fast Fourier transform (FFT).

- Convert the frequency outputs by the FFT onto the mel scale and obtain the logarithm of all filterbank energies.

- Apply the discrete cosine transform on the log filterbank energies.

Hence, MFCCs are obtained from the following relation:

$$c_n = \sum_{k=1}^{k} \log(S_k) \cos\left[i(k - \frac{1}{2})\frac{\pi}{k} \right], i = 1, 2...., M$$

where K is the number of filterbanks, which is 20 in this study. M is the length of the cepstrum that is chosen to be 13, and S_k represents the energy output of the kth triangular band pass filter.

4.2 Acoustic features

The set of acoustic parameters used in our study is presented in Table 2. The first five characteristics were investigated previously in [19]. The automatic estimation of the remaining features are described below.

Table 2. Studied cry characteristics

Characteristics	Definitions.
Fundamental frequency (F_0)	Vibration frequency (in Hz) of the vocal folds.
F_0 Harmonics	Multiples of fundamental frequency.
RF_1, RF_2 (Hz)	The first and second vocal tract resonance frequencies.
Tuning (TUP)	Blocks of 1024 samples during which an RF remains close to the harmonics of F_0 (distance < 100 Hz).
Transition (TRP)	Blocks of 1024 samples between two subsequent TUPs.
F_{0glide}	Rapid increase or decrease in F_0 of 600 Hz or more.
Glide duration (T_{glide})	Defined by the start time and finish time of F_0 glide.
Dysregulation of RFs ($RFs_{dys)}$	High frequency variability of RFs (RF_1, RF_2).

4.2.1 Features description

Research in acoustic phonetics was focused on the mechanisms of speech production rather than cry production in newborns. Moreover, the acoustic properties of vocal folds and many aspects of source-filter interaction have not been clearly defined in newborn cries. Therefore, we have investigated the acoustic properties of speech and singing to understand the F_{0glide} and RFs_{dys} phenomena in cry signals.

In vowel and voiced consonant production theory, the glottis operates independently of the vocal tract [20]. This theory concerns male speech more than female and child speeches. However, glide production is influenced by the interaction between the vocal tract and vocal folds [21]. When the oscillation frequency of the vocal fold F_0 approaches the vocal tract resonance RFs (RF_1, RF_2), a nonlinear source-filter system coupling occurs. This interaction is more important for female and child speeches and even more for singing [22]. It has been mentioned [23] that the

rising and falling F_0 occur with the upward and downward movement of the larynx respectively, and that the control mechanisms of F_0 are influenced by the movements of the laryngeal frame and cervical spine.

Based on the average pressure decrease across a narrow constriction in normal voice, Stevens [24] defined glides as a class of consonants with a constriction that is not sufficiently narrow. Thus, the glides were produced with a greater degree of constriction in the vocal tract, and their categories were differentiated by the aerodynamics of different degrees of vocal tract constriction.

In cry signals, F_{0glide} was defined as a rapid change in F_0 (> 600 Hz) in 0.1 s. It can be either rising or falling glides [8]. This characteristic is shown in Figure 1 (a) that represents a spectrogram of a hyperphonic and dysphonic cry for a newborn suffering from asphyxia. We can observe the upward glide of 1102-2390 Hz at 0.053-0.210 s from the F_0 contour of this cry, and thus the duration of glide (0.157 s) is higher than 0.1 s.

Some studies have investigated the relationship between the glottal properties and vocal tract acoustics of an opera singer [25, 26, 27, and 28]. According to [25], the increase or decrease in resonance frequencies is related to the possibility of changes in the configuration of organs involved in voice production as follows: glottal-opening duration, glottal vibratory amplitude, glottal area, laryngeal height, mouth opening, and tongue shape. Because the tuning (TUP) phenomenon is defined as an adjustment of the RF such that it is close to F_0 or one of its harmonics, it has been shown that professional singers use many resonance tuning strategies for vowels (RF_1: $2F_0$, RF_1: $3F_0$, RF_2: $4F_0$, RF_2: $5F_0$, RF_2: $6F_0$, and RF_2: $8F_0$ tuning) [25] .

In our previous work [19], an automated approach has been used to estimate and evaluate the duration and percentage of TUP between RFs and the harmonics of F_0 for healthy and pathologic newborn cries. The obtained results encouraged us to further investigate the variability of RFs, termed RFs_{dys} here. Furthermore, this characteristic has been associated with neonatal diseases [10]. However, to our best knowledge, none have evaluated the variation pattern of RFs_{dys} whether in cry signals or in speech signals.

To quantify RFs_{dys} and analyze their variation patterns according to the studied pathologies, we define this characteristic by the RF jumps between two successive TUPs. Therefore, during the transition period TRP as shown in Figure 1 (b), we observed that RF_1 suddenly jumped downward from 3178 Hz to attain 2196 Hz at t = 0.031 s. Because F_0 = 450 Hz, this jump

occurred between two successive tunings with the 7th and 5th harmonics of F_0, ($RF_1:7F_0$ and $RF_1:5F_0$). Thus, in this case, the jump "J" is equal to two.
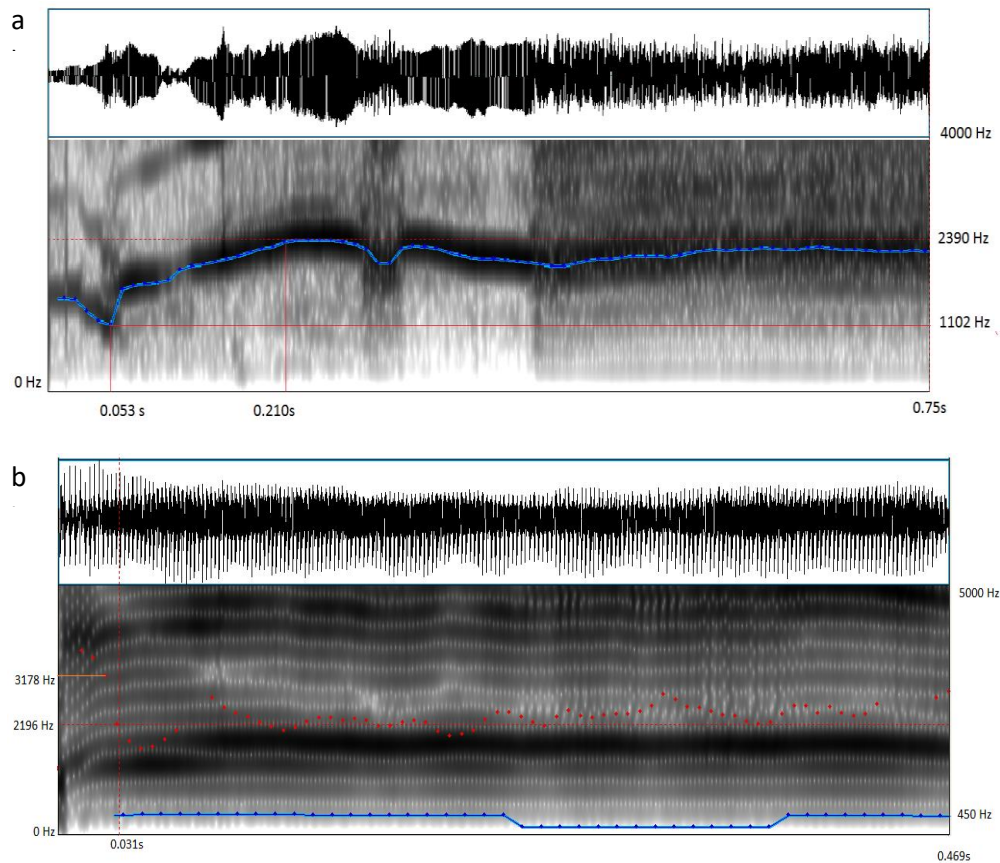

Figure 1. Spectrograms of cry signals using PRAAT [29].

a) Rising F_{0glide} in the cry of full-term infant suffering from asphyxia.

b) RF_{1dys} between two successive TUPs in the cry of preterm infant with gastroschisis.

4.2.2 Feature estimation

The adopted approach for the estimation of F_{0glide}, T_{glide}, and RFs_{dys} is illustrated in Figure 2, and the principal steps are detailed below.
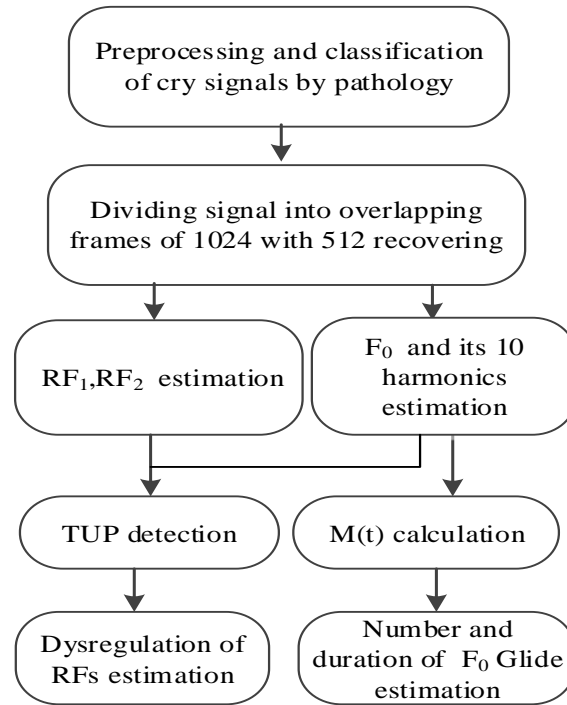
Figure 2. Flowchart of adopted approach.

In the first step, cry signals were processed manually using PRAAT (a freeware program for the analysis and reconstruction of acoustic speech signals) [29]. The automated preprocessing method is currently being investigated by other researchers studying the PICS [30]. Because the recorded sounds include the background noises, speech, sound of medical equipment, and silence, some of these parts may distort the results of the analysis. Therefore, this step consists of noise filtering and segmentation of recordings into useful and non-useful segments. Subsequently, the database was organized by pathology and gestational age.

In the second step, we applied the following procedure for acoustic characteristic measurement using self-developed functions in Matlab. This has been performed for healthy and pathologic cries by pathology and gestational age. Each cry segment of 1-s duration was divided into 50% overlapping frames of 1024 samples. Each frame was multiplied by the Hamming window. Thereafter, for each data frame, we estimated F_0 and its 10 harmonics using the simple inverse filtering tracking algorithm [31, 32]; further, we estimated the RFs (RF_1, RF_2) using the modified covariance method based on the autoregressive power spectral density (AR-PSD) [33, 34]. The primary steps of the estimation algorithm of F_0 and RFs are detailed in [18]. The performances of

these algorithms were tested on a real newborn cry database [16, 32]. The derived function M(t) was estimated using F_0 contour measurements and subsequently used for F_{0glide} detection.

The TUPs and the jumps of RFs were estimated using the harmonics of F_0 and RFs measurements. In contrast to a previous study [35], where the TUPs were investigated using a special visualization concept, the TUPs in our work were detected automatically according to the given definitions in Table 2 and subsequently used for RFs_{dys} estimation.

For each cry sample, $F_{0glides}$, T_{glide}, and RFs_{dys} (RF_{1dys}, RF_{2dys}) were calculated using the adopted algorithms detailed below.

4.2.2.1 F_{0glide} and T_{glide} estimation

The following approach was adopted in this work to detect F_{0glide} and its corresponding T_{glide} in 1-s cry segments:

(1) Estimate F_0 and obtain the derived function M(t) of F_0.

$$M(t) = \frac{dF_0(t)}{dt}$$

(2) Obtain the time indexes "x" when M (t) is equal to zero.

$$M(x) = 0$$

(3) Calculate the corresponding $F_0(x)$.

(4) Identify the glides of F_0 by the following condition:

$$\left| F_0(x(i)) - F_0(x(i+1)) \right| \geq 600 Hz, \; i=1\ldots\ldots N\text{-}1$$

N: number of zero values at the function M(t)

(5) Obtain the duration T_{glide} of each F_{0glide}, by the following formula:

$$T_{glide} = T(x(i+1)) - T(x(i))$$

T(*x*): the times at index « *x*».

(6) Calculate the average percentage of F_{0glide} ($P_{F0glide}$) and average of T_{glide} (A_{Tglide}) by pathology and gestational age, using these formulas:

$$P_{F0glide} = \frac{N_{glide}}{N_{Total}},$$

$$A_{Tglide} = \frac{\sum T_{glide}}{N_{glide}}$$

N_{Total} : Total number of segments in cry samples of 1-s duration.

N_{glide}: Number of F_{0glide} that occurred.

(7) Calculate the median and interquartile range of the evaluated characteristics $P_{F0glide}$ and A_{Tglide}.

Herein, we did not consider the time constraint of 0.1 s for the durations of F_{0glide}. Thus, T_{glide}, which could be of any length, was estimated and analyzed according to the pathologies.

4.2.2.2 RFs$_{dys}$ estimation

The estimation of RFs$_{dys}$ was performed by computing the jump "J" of the RFs between two subsequent TUPs or during the transition duration (TRP). "Jump" represents the number of harmonics between two subsequent TUPs. This study examines the Js in the range of 1 to 9. In our work, the TUPs and TRPs were detected automatically and used for the evaluation of RFs$_{dys}$. RFs$_{dys}$ were separately investigated for RF$_{1dys}$ and RF$_{2dys}$ with the same procedure for each data frame in 1-s cry segments as follow:

(1) Estimate F_0 and its 10 harmonics.

(2) Estimate RF$_1$ and RF$_2$.

(3) Identify TUPs separately for RF$_1$ and RF$_2$ by the following conditions.

$$|n.F_0(p) - RF_1(p)| \leq 100, n \in [1,...9], p \in [1.....N_{Total}]$$

$$|n.F_0(p) - RF_2(p)| \leq 100, n \in [1,...9], p \in [1.....N_{Total}]$$

n: order of harmonic (1 to 9).

p: index of frame in cry segment

(4) Find the time indexes of the start and the end of TRPs, i.e., (s_1, e_1) and (s_2, e_2), for RF$_1$ and RF$_2$, respectively.

(5) Calculate the corresponding RF_1 (s_1), RF_1 (e_1), RF_2 (s_2), and RF_2 (e_2).

(6) Obtain the jumps J_1 and J_2 for RF_1 and RF_2 ,respectively, using the following formulas, respectively:

$$C_{RF1}(z_1) = RF_1(s_1(z_1)) - RF_1(e_1(z_1)) = J_1(z_1) \times F_0 \ z_1 = [1.......N_{1TRP}]$$
$$C_{RF2}(z_2) = RF_2(s_2(z_2)) - RF_2(e_2(z_2)) = J_2(z_2) \times F_0 \ z_2 = [1.......N_{2TRP}]$$

where C_{RF1}, C_{RF2} are the differences in RFs, i.e., RF_1, RF_2, respectively, during the TRPs. N_{1TRP}, N_{2TRP} represent the number of TRPs for RF_1, RF_2, respectively.

Thereafter, for each jump J_1 and J_2 in the range of 1 to 9, the average percentages of RF_{1dys} (P_{RF1dys}) and of RF_{2dys} (P_{RF2dys}) were estimated for each pathology and gestational age, using these formulas, respectively:

$$P_{RF_{2dys}}(J_1) = \frac{N_{J_1}}{N_{Total}}, J_1 \in [1.....9], \ P_{RF_{2dys}}(J_2) = \frac{N_{J_2}}{N_{Total}}, J_2 \in [1.....9]$$

N_{J1}, N_{J2} are the number of jumps J_1 and J_2, respectively, in cry samples of 1-s duration.

N_{Total} : Total number of segments in cry samples of 1-s duration.

(7) Calculate the median and interquartile range of the evaluated characteristics P_{RF1dys} and P_{RF2dys}.

5. PCIS description

The proposed diagnostic system was built around an efficient PNN classifier and significant infant cry features. In this study, we investigate the effectiveness of the extracted characteristics ($P_{F0glide}+A_{Tglide}+P_{RF1dys}+P_{RF2dys}$) concatenated with MFCC features for the classification of healthy and sick newborn infants according to their cries. Thus, the sets of characteristics once obtained were used as inputs to the PNN classifier.

The PNN classifier is used widely for classification problems in the medical domain [36, 37]. It is ideal for real-time applications and is computationally inexpensive. It can learn using the conjugate gradient method, new incoming training data without having to repeat the whole training process [38]. The PNN classifier implements the Bayesian decision rule. Its network architecture is based on three layers: (1) the input layer that computes the distances between the

input vector and the training input; (2) radial basis layer that produces a vector of probabilities; (3) competitive layer that selects the maximum of these probabilities. The network classifies the input vector, the assigned class as those with the highest probability [36].

Matlab was used for the development of the test system. The primary steps are as follows:

- wav files recovery from each folder of the studied pathologies.

- division of 1-s cry signals into overlapping frames of 1024 samples, with 50% recovery.

For each 1-s cry segment, the following were performed:

- estimation of $P_{F0glide}$, AT_{glide}, P_{RF1dys}, and P_{RF2dys}.

- extraction of 13 MFCCs for each interlaced frame.

- conversion of each matrix of the MFCCs specific to a cry sample into only one vector.

- concatenation of each vector with its corresponding $P_{F0glide}$, A_{Tglide}, P_{RF1dys}, and P_{RF2dys}. The set of obtained vectors were selected as the final inputs to the PNN classifier.

- application of five-fold cross validation; four folds for training set and one fold for testing set.

- data classification using the PNN classifier; the strategy proposed by [36] paper and the function newpnn () in Matlab [39] were used.

- calculation of the correct identification rate namely, accuracy.

6. Results and discussion

6.1 Acoustic and statistical analysis

The cry signals were analyzed using the studied characteristics ($P_{F0glide}$, A_{Tglide}, P_{RF1dys}, and P_{RF2dys}). Moreover, they have been investigated according to the gestational age (full-term and preterm) and diseases afflicting newborns. The cry features given in Table 2 were estimated for the complete database presented in Table 1.

An example of F_{0glide} is shown in Figure 3; it represents an estimate of F_0 for a 1-s cry of a newborn suffering from asphyxia. The irregular contour of F_0 presents the rising glide between 0.21 s and 0.24 s and the falling glide between 0.58 s and 0.63 s.
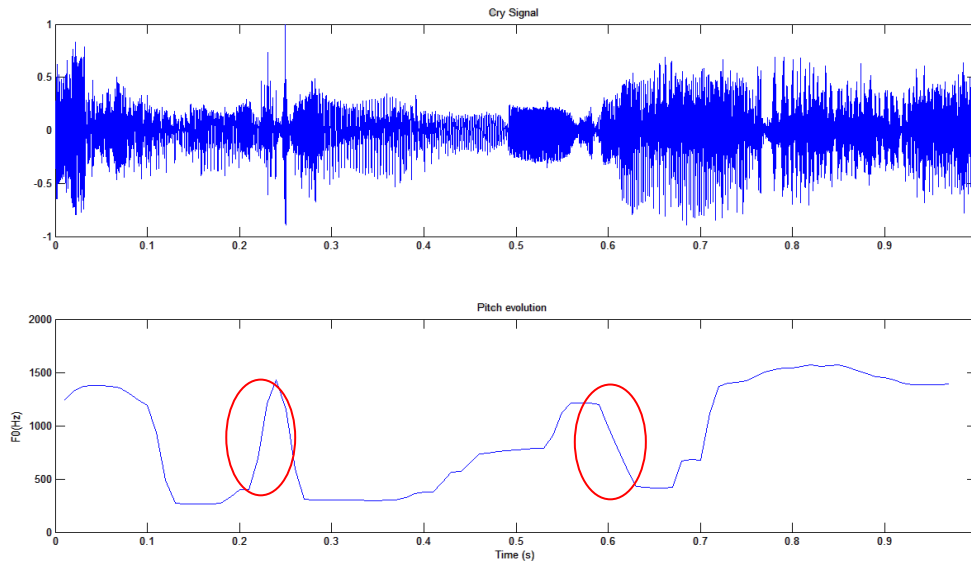
Figure 3. Estimated F_0 for a cry signal of a newborn suffering from asphyxia; rising and falling F_{0glide} (red ellipses).

The average percentage of F_{0glide} ($P_{F0glide}$) and the average durations of the glides (A_{Tglide}) are given in Table 3 and presented by pathology and gestational age in Figures 4(a) and 4(b), respectively.

The obtained results indicate that the cry signals do not present the glides of F0 in both healthy premature and full-term newborns, and also in the studied cases of lingual frenum and meningitis. They also indicate that, in a group of full-term newborns, the highest median and interquartile range of PF0glide are found in peritonitis with average glide durations compared to other diseases. The lowest percentage is found in the case of vena cava thrombosis with longer durations of F0glide. In the same group of full-term newborns, a small difference occurred in the occurrence of F0glide and also ATglide between the cries of hyperbilirubinemia and asphyxia diseases.

Table3. $P_{F0glide}$, and A_{Tglide} estimation results

Pathology	$P_{F0glide}$ (%)		A_{Tglide} (s)	
	Median	*Interquartile Range*	*Median*	*Interquartile Range*
Healthy (t)	0	0	0	0
Hyperbiluribinemia(t)	8.33	3.9	0.16	0.10
Vena cava thrombosis (t)	7.14	4.87	0.25	0.11
Meningitis (t)	0	0	0	0.00
Peritonitis (t)	12.5	6.52	0.22	0.12
Asphyxia (t)	8	4.17	0.15	0.09
Lingual frenum (t)	0	0	0	0.00
Healthy (P)	0	0	0	0.00
IUGR-microcephaly (P)	18.46	12.5	0.28	0.10
Tetralogy of Fallot (P)	10	4.91	0.31	0.17
Gastroschisis (P)	16.66	8	0.27	0.11
IUGR-asphyxia (P)	11.11	4.95	0.23	0.12
RDS (P)	10	5.71	0.18	0.11

The highest median and interquartile ranges of PF_{0glide} among the two groups of newborns were observed in the cries of preterm newborns affected by IUGR-microcephaly disease with rather long periods of glides. However, Figure 4 and Table 3 show that the highest median and interquartile range of A_{Tglide} characterize the tetralogy of Fallot disease, which presents the same $P_{F0glide}$ as RDS disease compared to other pathologies.
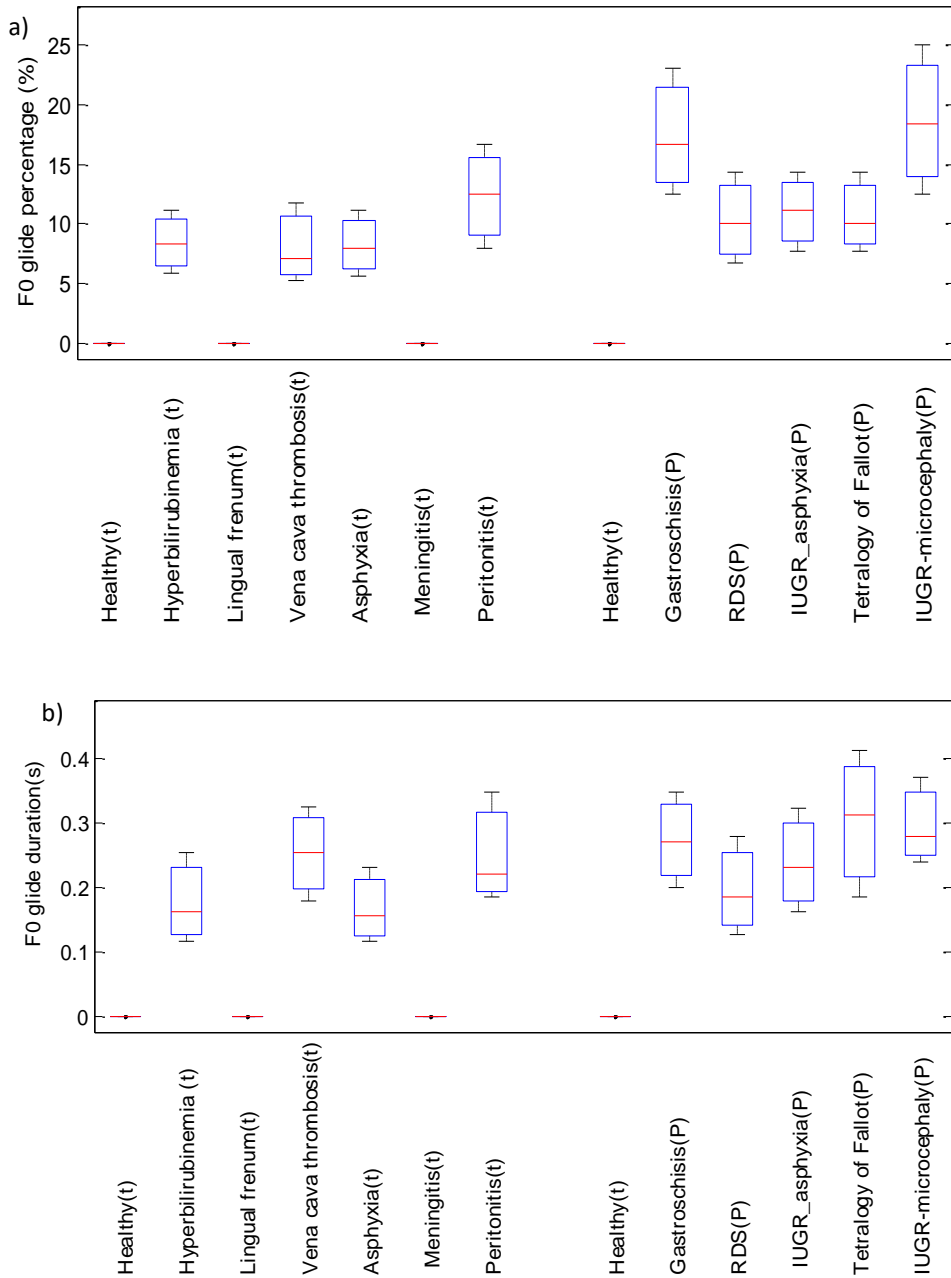
Figure 4. Box-and-whiskers plots for a) $P_{F0glide}$, b) T_{glide}, by pathology.

Concerning the RFs_{dys}, examples of this feature are shown in Figure 5, which represents the estimates of F_0, their harmonics, RF_1, and RF_2 for the 1-s cry of a newborn suffering from hyperbilirubinemia. An RF_{1dys} was observed at approximately 0.38 s between the third and sixth harmonics, "J = 3" and also an RF_{2dys} at approximately 0.26 s between the seventh and ninth harmonics "J = 2".
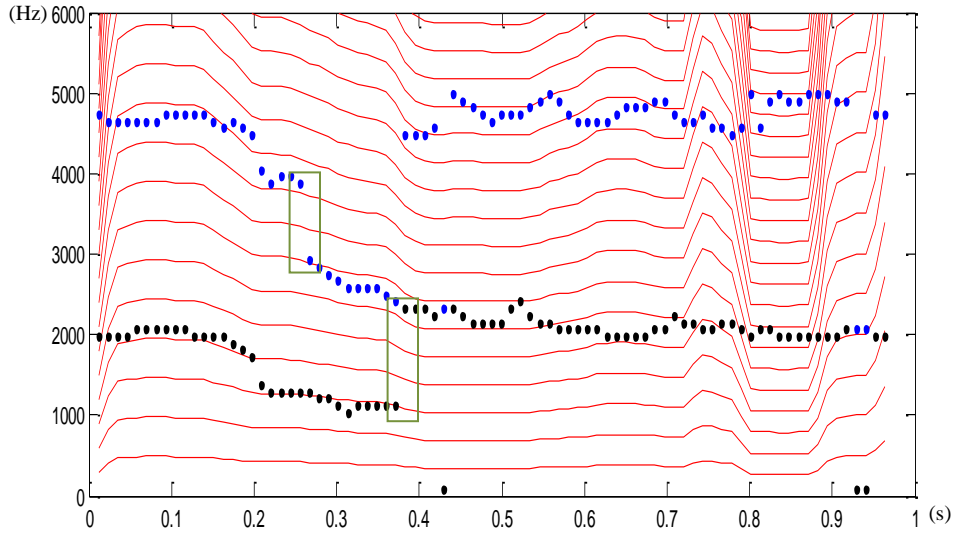
Figure 5. Full-term newborn suffering from hyperbilirubinemia,

F_0 and their harmonics (red lines), RFs: RF_1 (black dots) and RF_2 (blue dots), RFs_{dys} between

TUPs (green rectangles).

Comparative results are presented in Figure 6. They exhibit the variation patterns of RFs_{dys} by pathology. This dysregulation is expressed by the percentage of jump "J" previously defined in Section 3.4. Figures 6(a) and (b) show that the majority of cries of preterm newborns are characterized by more RF_{1dys} when J = 1 than cries of full-term newborns. The percentage of RF_1 jumps is more important for all jumps J = 1 to 9 in the cries of healthy preterm newborns than in the cries of healthy full-term newborns.

According to Figure 6(a), the cry signals of vena cava thrombosis and asphyxia diseases present the most RF_{1dys}, with the highest percentage of jump for "J" from 1 to 9, except for "J" of 3 and 4 jumps, where the largest percentage is found in cry signals of lingual frenum pathology. In the case of hyperbilirubinemia and for "J" equal to 1, 2, and 3 jumps, the cry signals were qualified by more RF_{1dys} than those of healthy newborns. Further, in peritonitis disease, the percentage of jumps for "J" of 1, 2, 5, and 6 is higher than those of healthy newborns. Unlike these latest pathologies, the cry signals in case of meningitis present less RF_{1dys} for all jumps "J" (1 to 9) than in healthy newborn. From Figure 6 (b), we can observe the highest RF_{1dys} for jumps J = 1 to 6 in the case of tetralogy of Fallot disease compared to the other studied cases. We also noticed a

19

higher percentage of RF_1 jumps for J = 1 in both cases of RDS and gastroschisis, and less RF_{1dys} was observed in the cry signals of IUGR–microcephaly and IUGR–asphyxia compared to the cries of healthy preterm newborns. The variation patterns of RF_{2dys} for full-term and preterm newborn cry signals are shown in Figure 6 (c and d, respectively). We observed that the percentage of RF_2 jumps depends on the studied pathologies. According to the jumps "J", it can be less or more important than in those in healthy newborn cries. In the cases of full and preterm newborns, all studied diseases are characterized by low difference in the percentage of RF_2 jumps compared to healthy preterm newborn cries, except for the cry signals of tetralogy of Fallot disease that presents a greater percentage of RF_{2dys} than the healthy ones for J = 1 to 5 jumps.
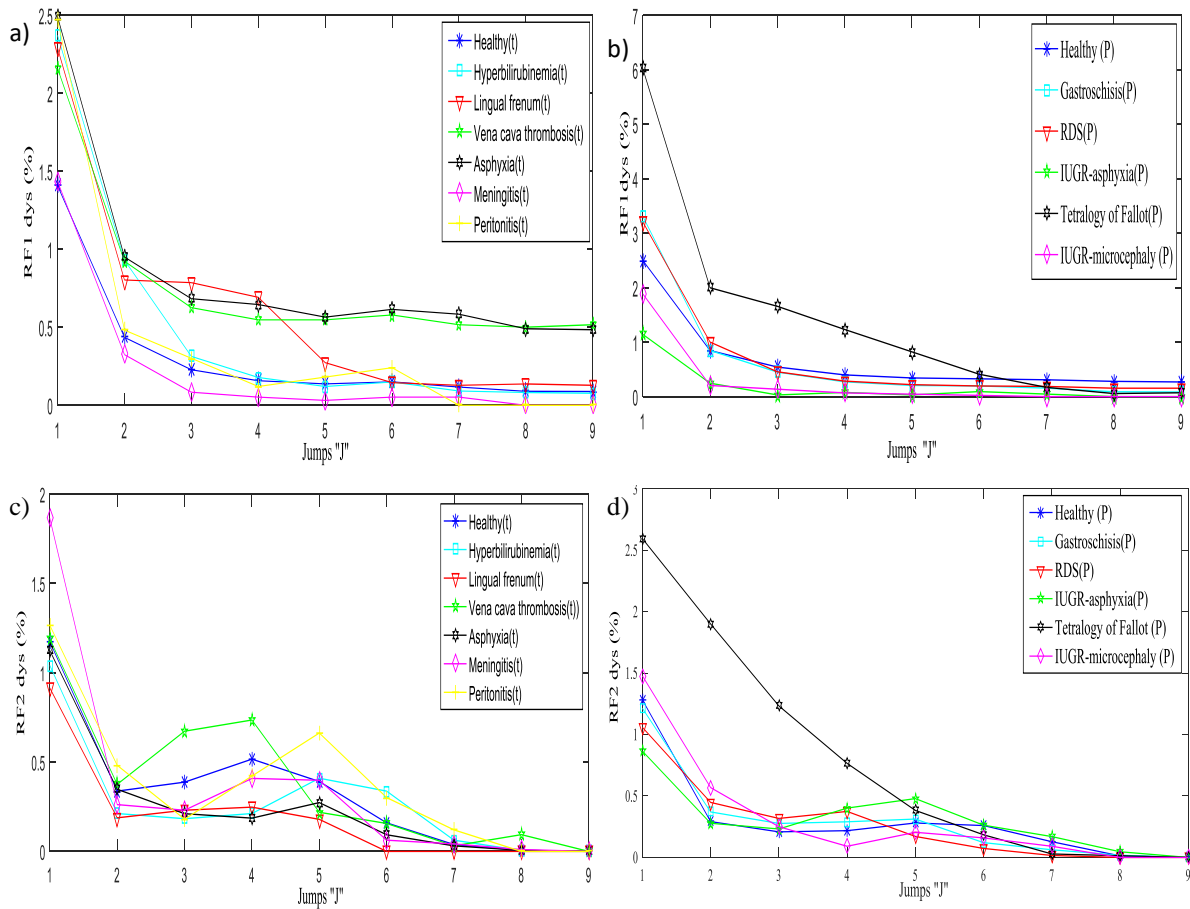

Figure 6. Average percentage of RFs_{dys} by jumps "J".

a) RF_{1dys} for full-term newborns, b) RF_{1dys} for preterm newborns,

c) RF_{2dys} for full-term newborns, d) RF_{2dys} preterm newborns.

6.2 Classification results

In this work, we present the results of healthy and sick (with specific diseases, see Table 1) infant identification systems. To test and validate the use of the studied characteristics, two experiments were performed in both cases of full-term and preterm infants: 1) separate cries by health status of infant (pathologic cries from healthy ones); 2) separate cries into those of healthy infants and those sick with specific diseases. The effectiveness of the studied characteristics ($P_{F0glide}$, A_{Tglide}, P_{RF1dys}, and P_{RF2dys}) in the recognition of pathological cries was compared to the use of MFCC parameters as the input to the PCIS. The performance of the PCIS was evaluated based on the rate of correct identification (overall accuracy) [4].

The distribution of cry samples according to classes and pathologies are shown in Table 4. In the first experiment, the cry samples are distributed into two classes: healthy cries and pathological cries. In the second experiment, the cry samples are distributed in five classes according to the used pathologies. There are many healthy samples in our database and limited numbers of cry samples from sick infants. However, the five-fold cross validation has been used, in which each fold is of equal size and contains the same percentage of samples as that of each target class.

Table 4. Distribution of cry samples per class

Experiments	Number of Classes	Gestational age	Samples number Per class	Classes
First	2 classes	Preterm	50×5/250	Pathologic (Tetralogy of Fallot, Gastroschisis, IUGR-microcephaly, RDS, IUGR-asphyxia) /healthy
		Full term	50×5/250	Pathologic (hyperbilirubinemia,lingual frenum, vena cava thrombosis, asphyxia, meningitis) /healthy
Second	5 classes	Preterm	50×5	Tetralogy of Fallo /gastroschisis /RDS /IUGR-asphyxia/healthy
		Full-term	50×5	Hyperbiluribinemia/vena cava thrombosis/asphyxia/meningitis/healthy

The best results of the correct identification rate obtained for all implemented experiments are shown in Table 5. The use of MFCCs as feature characteristics is tested for the two experiments. The best performance obtained (72%) is found in the identification of full-term infants pathologies. The second test corresponds to the use of the studied features in addition to MFCCs. This test provides a better identification rate (88.71%) of health status of preterm infants.

According to the comparative results presented in Table 5, the use of the studied characteristics ($P_{F0glide}$, A_{Tglide}, P_{RF1dys}, and P_{RF2dys}) in addition to MFCCs improves the performance of the PCIS compared to the use of MFCCs without the studied characteristics compared to the MFCC features only; in the first experiment, the overall accuracy had improved by 18.5% points for the identification of health status of preterm infants and by 7% points for the identification of the health status of full-term infants. In the second experiment, the overall accuracy improved by 12% points for the identification of preterm infants with a specific disease, and 10% points for the identification of full-term infants with a specific disease.

Table 5. Overall accuracy results

Experiments	Classes number	Gestational age	Features	
			MFCC	$PF_{0glide} + A_{Tglide} + P_{RF1dys} + P_{RF2dys} + MFCC$
First	2 classes	Preterm	70.16 %	88.71 %
		Full term	60.00 %	67.00 %
Second	5 classes	Preterm	60.00 %	72.00 %
		Full term	72.00 %	82.00 %

7. Conclusion

In this work, we proposed an automated classification method for different cry patterns. We used new features that improved the accuracy of the recognition of pathological cries. This study focused on the following: 1) improvement of PCIS performances using studied features, 2)

description of our method for estimating the proposed features, and 3) description of associations between infant medical conditions and the considered characteristics.

The proposed methods to estimate these features permitted the analysis of a considerable number of pathologic cry signals in a short time, and the extension of the study for other pathologies not considered. Hence, we are recording a larger database with a greater variety of pathologies and more subjects for each pathology.

We successfully identified the occurrence of the proposed characteristics in cry signals according to the studied pathologies. The glide of F_0 was found absent in healthy newborn cry signals and also in some of the pathologies. The obtained results suggested that the prevalence of relatively slow glides increased in sick preterm infants, especially in babies with diseases affecting the CNS such as IUGR-microcephaly. Our results indicated a complex pattern variation of RF_{2dys} in the pathologic cry signals compared to that of RF_{1dys} .

From the presented work and results obtained in this study, we conclude that, occurrences of F_{glides} and RF_{1dys} in addition to other characteristics investigated in our previous works were found to be useful in the diagnosis of the studied pathologies.

The encouraging classification results indicated the highly discriminatory nature of the proposed features. These characteristics can be explored as inputs to a diagnostic system using other modeling and classification methods to ultimately provide a basis for alerting healthcare workers to intervene. We also expect to improve the results by studying other classifiers.

The results obtained using the PNN classifier and the proposed acoustic features were not compared with other results of previous studies. It is noteworthy that in this study, cases of full-term and premature babies, both with and without pathologies were investigated, thus increased the complexity of a PCIS significantly. For example, the cry signals of neonatal asphyxiate and healthy babies contain significant differences in the acoustic signals, thus easing the identification process.

In conclusion, our research underlined the importance of using acoustic features in the task of cry recognition and provided support and additional results to investigations regarding neonatal cries.

REFERENCES

[1] O.Fausto, R.Galaviz, C.A.Reyes Garcia, Infant cry classification to identify hypo acoustics and asphyxia comparing an evolutionary-neural system with a network system, Springer-Verlag Berlin Heidelberg. 2005, 949-958.

[2] D. Lederman, E.Zmora, S. Hauschildt, A. Stellzig-Eisenhauer, K. Wermke, Classification of cries of infants with cleft-palate using parallel hidden markov models, Med Bio Eng Comput. 2008, 965-975.

[3] M.Hariharan, S.Yaacob, S.Ardeena awatie Awang, Pathological infant cry analysis using wavelet packet transform and probabilistic neural network, Expert Systems with Applications. 2011, 38, 15377–15382.

[4] Kheddache Y, Tadj C , "Newborn's pathological cry identification system", The 11 th international conference on information sciences, signal processing and their applications. 2012 pp 1051-1056

[5] Manfredi.C, Andrea.B, Melino.D, Viellevoyec.R, Masendu. K, Silvia. O, Automated detection and classification of basic shapes of newborn cry melody, Biomedical Signal Processing and Control, 2018, pp 174–181,

[6] LaGasse LL, Neal AR, Lester BM, Assessment of the infant cry: acoustic cry analysis and parental perception, Ment Retard Dev Disabil Res Rev. 2005, 11, 83–93.

[7] S.Orlandia, A.Bandiniab , F.F.Fiaschia, C.Manfredia,Testing software tools for newborn cry analysis using synthetic signals, Biomedical Signal Processing and Control. 2017, pp 16-22.

[8] Wasz-Hockert O, Michelsson K, Lind J. Twenty-five years of scandinavian cry research, In: Lester BM, Boukydis CFZ, eds. Infant crying. New York, NY: Plenum; 1985, 83–104

[9] Michelson K, Todd de Barra H, Michelson O, Sound Spectrographic cry analysis and mothers perception of their infant's crying. In: Lewis FR, ed. Focus on nonverbal communication research. New York, NY: Nova Science Publishers. 2007, pp 31–64.

[10] Lester BM, LaGasse LL. Crying. In: Haith MM, Benson JB, eds. Encyclopedia of infant and early childhood development. San Diego, CA: Academic Press; 2008, pp 332–343.

[11] Corwin, Michael J, Herbert Kayne, Barry M Lester, Carol Sepkoski, Sarah McLaughlin et Howard L Golub.. Effects of in utero cocaine exposure on newborn acoustical cry characteristics ». 1992. Pediatrics, vol. 89, no 6, pp 1199-1203.

[12] Orozco,J.,Garcia,C.A.R.,. Detecting pathologies from infant cry applying scale conjugate gradient neural networks. In: presented at the Europe an Symposium on Artificial Neural Networks. Bruges-Belgium. 2003

[13] Alaie, Hesam Farsaie, et Chakib Tadj. « Cry-based classification of healthy and sick infants using adapted boosting mixture learning method for gaussian mixture models ». Modelling and Simulation in Engineering, vol. 2012, p. 55.

[14] Amaro- Camargo,E ., Reyes García,C.. Applying statistical vectors of acoustic characteristics for the automatic classification of infant cry. In:Huang, D.-S., Heutte, L. ,Loog,M.(Eds.), Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues, 4681. Springer, Berlin/Heidelberg, 2007 pp.1078–1085.

[15] Cano Ortiz, SergioD, DanielI Escobedo Beceiro et Taco Ekkel.. « A Radial Basis Function Network Oriented for Infant Cry Classification ». In Progress in Pattern Recognition, Image Analysis and Applications,. 2004 Vol. 3287, pp 374- 380.

[16]. Fisichelli, Vincent R., S.Karelitz. Frequency spectra of the cries of normal infants and those with down's syndrome. Psychonomic Science.1966, 6, pp 195-196.

[17] Kheddache Y, Tadj C. Frequential characterization of healthy and pathologic newborn cries. Am J Biomed Eng. 2013, 3, pp 182–193.

[18] Kheddache Y, Tadj C, Acoustic measures of the cry characteristics of healthy newborns and newborns with pathologies, J Biomed Sci Eng. 2013,6 , pp 796–804.

[19] Kheddache Y, Tadj C, Resonance frequencies behaviour in pathologic cries of newborns, Journal of Voice, 2015, 29, pp 1-12

[20] A. Aalto, D. Aalto, J. Malinen, M Vainio, Interaction of vocal fold and vocal tract oscillations, Proceedings of the 24th Nordic Seminar on Computational Mechanics, J. Freund and R. Kouhia (Eds.) Aalto University, 2011, pp.1-4

[21] A. Aalto, D. Aalto, J. Malinen, M Vainio, Modal locking between vocal fold and vocal tract ascillations, arXiv preprint arXiv:1506:01395

[22] Ingo R. Titze, Nonlinear source–filter coupling in phonation: Theory, J. Acoust. Soc. Am, 2008, pp 1902-1915.

[23] Kiyoshi honda, Hiroyuki hirai, Shinobu masaki, Yasuhiro shimada, Role of vertical larynx movement and cervical lordosis in F_0 control, language and speech, 1999, pp 401– 411

[24] stevens , K.N. Acoustic phonetics, Cambridge, Massachusetts: MIT Press, (1998).

[25] Nathalie Henrich Bernardoni, John Smith, Joe Wolfe, Vocal tract resonances in singing: variation with laryngeal mechanism for male operatic singers in chest and falsetto registers, J. Acoust. Soc. Am, 2014, pp 491-501.

[26] Echternach, M., Sundberg, J., Arndt, S., Markl, M., Schumacher, M., and Richter B, Vocal tract in female registers–A dynamic real-time MRI study, J. Voice, 2010, pp 133–139.

[27] Sundberg, J., La, F. M. B., and Gill, B. P, Professional male singers' formant tuning strategies for the vowel /a/," Logoped. Phoniatr. Vocol. 2011, pp 156–167.

[28] Johan Sundberg, Filipa M. B. Brian P. Gill, Formant Tuning Strategies in Professional Male Opera Singers Journal of Voice, 2013. pp 278-288.

[29] P. Boersma, D. Weenink, Praat: doing phonetics by computer. [Online]. Available: http://www.praat.org/

[30] Abou-Abbas.L,.Tadj.C, Gargour.C, Montazeri.L, Expiratory and Inspiratory Cries Detection Using Different Signals' Decomposition Techniques, Journal of Voice, 2017 259.pp 13–28

[31] Markel, J.D, The SIFT algorithm for fundamental frequency estimation, IEEE Transactions on Audio Electroacoustic, 1972 , pp 367-377.

[32] Lederman D, Estimation of infants' cry fundamental frequency using a modified SIFT algorithm, 2010, arXiv: 1009.2796.

[33] Manfredi C, Bocchi L, Orlandi S, Spaccaterra L, Donzelli GP, High resolution cry analysis in preterm newborn infants, Med Eng Phys, 2009, pp 528–532.

[34] A. Fort, A. Ismaellit, C. Manfredi, P. Bruscaglionit, Parametric and non-parametric estimation of speech formants: application to infant cry, Med. Eng. phys. 1996, pp 677-691

[35] Wermk K, Mende W, Kempf A, Manfredi C, Bruscaglioni P, Stellzig- Eisenhauer A, Interaction patterns between melodies and resonance frequencies in infants' pre-speech utterances, In Proceedings of the 4th International Workshop.Models and Analysis of Vocal Emissions for Biomedical Applications, 2005. pp187–190,.

[36] Walter P. Sweeney Jr., Mohamad T. Musavi, John N. Guidi . Classification of chromosomes using a probabilistic neural network, Wiley-Liss, Inc. 1994 pp. 17-24,.

[37] Mohd Fauzi Othman, Mohd Ariffanan Mohd Basri . Probabilistic neural network for brain tumor classification. Second International Conference on Intelligent Systems, Modelling and Simulation, 2011 pp.136-138,.

[38] Maciej KusyEmail author, Roman Zajdel. " Probabilistic neural network training procedure based on Q(0)-learning algorithm in medical data classification", Applied Intelligence, 2014 pp. 837–854.

[39] Neural Network Toolbox™ 6 User's Guide.