Available online at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/bbe

Original Research Article

Time–frequency analysis in infant cry classification using quadratic time frequency distributions

Q1 J. Saraswathy [a,*], M. Hariharan [b], Wan Khairunizam [a], J. Sarojini [c], N. Thiyagar [d], Y. Sazali [e], Shafriza Nisha [a]

[a] School of Mechatronic Engineering, University Malaysia Perlis (UniMAP), Campus Pauh Putra, 02600 Perlis, Malaysia
[b] Department of Biomedical Engineering, Faculty of Engineering & Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamil Nadu, India
[c] School of Bioprocess Engineering, University Malaysia Perlis (UniMAP), Campus Jejawi 3, 02600 Arau, Perlis, Malaysia
Q2 [d] Department of Pediatrics, Hospital Sultanah Bahiyah, 05460 Alor Setar, Kedah, Malaysia
[e] Universiti Kuala Lumpur Malaysian Spanish Institute, Kulim Hi-Tech Park, 09000 Kulim, Kedah, Malaysia

ARTICLE INFO

Article history:
Received 16 December 2017
Received in revised form
3 April 2018
Accepted 9 May 2018
Available online xxx

Keywords:
Infant cry
Time–frequency analysis
Quadratic time–frequency distributions
t–f based feature extraction
Classification

ABSTRACT

This paper presents a new investigation of time–frequency (*t–f*) based signal processing approach using quadratic time–frequency distributions (QTFDs) namely spectrogram (SPEC), Wigner–Ville distribution (WVD), Smoothed–Wigner Ville distribution (SWVD), Choi–William distribution (CWD) and modified B-distribution (MBD) for classification of infant cry signals. *t–f* approaches have proved as an efficient approach for applications involving the non stationary signals. In feature extraction, a cluster of *t–f* based features were extracted by extending the time-domain and frequency-domain features to the joint *t–f* domain from the generated *t–f* representation. Conventional features such as mel-frequency cepstral coefficients (MFCCs) and linear prediction coefficients (LPCs) were also extracted in order to compare the effectiveness of the *t–f* methods. The efficacy of the extracted feature vectors was validated using probabilistic neural network (PNN) and general regression neural network (GRNN). The proposed methodology was implemented to classify different sets of binary classification problems of infant cry signals from different native. The best empirical result of above 90% was reported and revealed the good potential of *t–f* methods in the context of infant cry classification.

* *Corresponding author at*: School of Mechatronic Engineering, University Malaysia Perlis (UniMAP), Campus Pauh Putra, 02600 Perlis, Malaysia.
E-mail addresses: wathy_87@ymail.com (J. Saraswathy), hariharan.m@ktr.srmuniv.ac.in (M. Hariharan).
https://doi.org/10.1016/j.bbe.2018.05.002

Please cite this article in press as: Saraswathy J, et al. Time–frequency analysis in infant cry classification using quadratic time frequency distributions. Biocybern Biomed Eng (2018), https://doi.org/10.1016/j.bbe.2018.05.002

1. Introduction

Research on infant cry has been started in early of the 60s and it still carried out intensively in order to invent an automated tool for significantly discriminating different conditions of physical and physiological of infants which are mostly due to the organic disturbances, feed management, sleep management, maternal health and sensorimotor integration [1]. Toward this intent, there are some tools proposed and commercialized in the market, which developed mainly for discriminating the different physical behaviors of infants namely hunger, pain, discomfort, bore and anger [2,3]. It is declared that, these existing tools role very meaningful among caregivers, especially new parents for distinguishing their newborn cries and reacting quickly toward them without any uncertainties. Furthermore, by concerning on user accessibility and flexibility currently the handset models for sale slot in with the user freely infant cry translator applications.

Although there is some infant cry classification based systems available, the development of an automatic tool which can significantly aid the medical persons in diagnosis the pathological conditions of newborns at their early life time is still a challenging task. Hence, research works on infant cry constantly carried out by proposing different artificial and digital signal processing techniques in order to draw north worthy affects [4–8]. Particularly, different variable aspects of cry features such as pitch information [9–11], noise concentration [12], spectral energy features [13], harmonic analysis based attributes [14], linear prediction cepstral coefficients (LPCs) [15,16] and mel-frequency cepstral coefficients (MFCCs) [17–22] have been suggested to characterize the cry signals. However, most of the studies concentrated mainly on the application of MFCCs and LPCs in analyzing the non-stationary characteristics of newborn cry signals even though these methods are limited to non-stationary based analysis. Besides, it is important to highlight that, eventhough promising classification results are reported using the conventional features (MFCCs and LPCs) researchers have encountered the complexity issues, due to the high dimensionality of the extracted feature vectors. Hence in the literature, generally the MFCCs and LPCs feature extraction is enclosed with the dimensionality reduction process [17–20].

Generally, in applications involving the multi component signals like infant cry signals, the time–frequency (*t–f*) based techniques are demonstrated as knowledgeable rather than the methods based on time or frequency domains [23–25]. Nevertheless, the application of *t–f* based approaches in the development of infant cry based classification tools are not greatly highlighted and is identifies as a gap in this line of research. Although the existence of *t–f* based analyses is not much, a number of past studies have attempted the *t–f* based analyses in the classification of different cry utterances, and reported incentive results toward the hypothesis that *t–f* methods are superior for the non-stationary analysis. Briefly it can be said that, in the literature, different interpretations and representations of infant cry signals are endeavored using some *t–f* based techniques namely wavelet packet transform (WPT), short time Fourier transform (STFT) and empirical mode decomposition (EMD), and the outcomes from the previous works persuaded to explore and investigate more on the effectiveness of *t–f* methods in infant cry classification [26–33].

Furthermore, in the development of a reliable cry based classification system, the significant aspects as shown in Fig. 1 namely gender, age, native, size, severity and *N*-way classification should be concerned as these aspects may facade the originality of the cry signals. In literatures, although great concern not given on the fore mentioned aspects in the process of automatic classification, their effects on discrimination of the infant cry signals is experimentally discussed and emphasized [34–39].

Hence, by taking the aforementioned issues into concern, the aim of the current study is set into twofold:

1. Application of *t–f* based signal processing techniques and *t–f* based feature extraction for infant cry classification – a new study of time–frequency (*t–f*) based signal processing approach which is not attempted yet in literature using different members of Quadratic time–frequency distributions (QTFDs) including spectrogram (SPEC), Wigner–Ville distribution (WVD), Smoothed–Wigner Ville distribution (SWVD), Choi–William distribution (CWD) and Modified B-distribution (MBD) is performed. In feature extraction, a cluster of *t–f* based features are extracted from the *t–f* representation matrix by extending the time-domain and frequency-domain features to the joint (*t–f*) domain. The extracted features are subjected for classification using different supervised classifiers namely probabilistic neural network (PNN) and general regression neural network (GRNN).
2. Cross classification and study among different natives of cry signals – preliminary, classification among different natives of cry datasets which are available for research purposes (Mexico, Hungary and Malaysia) is performed. Initially, using the accessed cry datasets different binary classes are attempted in order to study the association among the different nativity of cry signals in binary cases.

2. Database

The investigational cry signals are accessed from three different origins of databases namely Mexico, Hungary and Malaysia. The detailed description of these databases and isolation of data sets for analysis is given in the ensuing sections.

2.1. *Mexican database*

A subset of samples including normal and deaf cries from the *Baby Chillanto* database which is a property of the Instituto Nacional de Astrofisica Optica y Electronica (INAOE)–CONACYT, Mexico is considered for investigation [6,15,17]. It comprises of 507 and 879 of normal and deaf cry samples respectively with the length of 1 s. The normal cry signals are recorded from 38 babies whereas the deaf cry signals recorded from 6 deaf babies. The cry samples are recorded directly by trained pediatricians and tagged in the instant of their recording.

Please cite this article in press as: Saraswathy J, et al. Time–frequency analysis in infant cry classification using quadratic time frequency distributions. Biocybern Biomed Eng (2018), https://doi.org/10.1016/j.bbe.2018.05.002