Assessment of pain expression in infant cry signals using Empirical Mode Decomposition

[1]B. Mijović, [2]M. Silva, [3,4]B.R.H. Van den Bergh, [5]K. Allegaert , [2]J.M. Aerts, [2]D. Berckmans, [1]S. Van Huffel

1 Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven, Leuven, Belgium
2 Division M3-BIORES: Measure, Model & Manage Bioresponses, Katholieke Universiteit Leuven, Leuven, Belgium
3 Department of Psychology, Tilburg University, Tilburg, The Netherlands
4 Department of Psychology, Katholieke Universiteit Leuven, Leuven, Belgium
5 Department of women and child, Group Biomedical Sciences, Katholieke Universiteit Leuven, Leuven, Belgium

Summary

Background **:** The presence of decoupling, i.e. the absence of coupling between fundamental frequency variation and intensity contour during phonetic crying, and its extent, reflects the degree of maturation of the central nervous system.

Objectives **:** The aim of this work was to evaluate whether Empirical Mode Decomposition (EMD) is a suitable technique for analyzing infant cries. We hereby wanted to assess the existence and extent of decoupling in term neonates and whether an association between decoupling (derived from EMD) and clinical pain expression could be unveiled.

Methods **:** To assess decoupling in healthy term neonates during procedural pain, 24 newborns were videotaped and crying was recorded during venous blood sampling. Besides acoustic analysis, pain expression was quantified based on the Modified Behavioral Pain Scale (MBPS). Fundamental frequency and the intensity contour of the cry signals were extracted by applying the EMD to the data, and the correlation between the two was studied.

Results **:** Based on data collected in healthy term neonates, correlation coefficients varied between 0.39 and 0.83. The degree of decoupling displayed extended variability between the neonates and also in different cry bouts in a crying sequence within an individual neonate.

Conclusion **:** When considering the individual ratio between the mean correlation of cry bouts during a crying sequence and their standard deviation, there seems to be a positive trend with increasing MBPS value. This might indicate that higher stressed subjects have less consistency in the investigated acoustic cry features, concluding that EMD has potential in the assessment of infant cry analysis.

Keywords **:** Empirical Mode Decomposition, phonetic crying, pain expression, neonates.

Correspondence to:

Bogdan Mijovic, Student,

Department of Electrical Engineering,

Katholieke Universiteit Leuven

Kasteelpark Arenberg 10,

3000 Leuven,

Belgium,

Telephone : +32 16 32-18-57,

e-mail address: bogdan.mijovic@esat.kuelueven.be

I. INTRODUCTION

One of the most important tools for infants to communicate is vocalization. In neonates, vocalization is mainly crying, and research has been done on the relationship between stress states and acoustic cry features. Bellieni et al. were able to document that newborns with a higher DAN (Douleur Aigue du Nouveau-né) score during venous blood sampling (heel lancing) also had a higher fundamental frequency and a correlation was found between the pain intensity (DAN-score) and the normalized Root Mean Square value (RMS) [1]. Facchini et al. looked to the relation between the DAN score and the occurrence of noise patterns in the sound spectrograms [2]. These noise patterns were chaotic and showed discontinuous parts in the crying signal, caused by a highly turbulent flow in the larynx. Results indicated that newborns with a higher DAN score showed more noise patterns.

Baby cries are the result of the interaction between control of different areas in the brain, respiratory control and vocal fold vibrations. At early stage, it is believed that a cry is the result of respiratory action and the effect of air going through a pipe, causing the vocal folds to vibrate, resulting in a cry bout [3], this is the vocalization produced during one expiration. The more the neural system matures, the more laryngeal control can be exerted resulting in manipulation and modulation of the cry signal, but observations on maturational aspects of the cry signal are limited and conflicting [1, 4]. The development of the central neural system has been linked with the extent of vocalization control [4]. Based on the cry production model of Golub described in the work of Moller and Schonweiler [3] and Barr et al. [5], decoupling between variations in fundamental frequency and intensity contour during crying indicates cortical control. Decoupling is referred to as the absence of coupling between fundamental frequency variation and intensity contour during phonetic crying. In this cry model, cry production is controlled by a three level processor structure, in which the lower level can be separated into the subglottal, glottal and supraglottal cry production areas. The mid-processor is involved in physiological inputs like pain, blood levels and respiratory constraints, while the feed-back processes are part of the upper-processor level, especially auditory feedback.

In the current study, we want to assess the relation between the energy envelope of the cry bout and frequency variation within the same bout using the method recently introduced by Huang [6], called Empirical Mode Decomposition (EMD). This method can decompose the signal into different oscillatory modes, and we are able to follow amplitudes and frequencies of each of these modes separately.

Furthermore, the relation between the resulting correlation of these variables with clinical indicators of pain (MBPS) after venous blood sampling (or stress expression) is evaluated.

II. MATERIALS AND METHODS

1. Data Set

The vocalization was recorded in 24 term neonates in whom a venous blood sampling was performed on the 3rd day of life for routine metabolic screening. All babies were born in the University Hospital of Leuven (U.Z. Leuven). The study was approved by the Ethical Board of the University Hospitals, and neonates were only included after written consent of the parents. In order to participate in the project, babies had to have minimal gestational age at least 37 weeks, no complications during pregnancy, a 1- and 5- minutes APGAR score above 7, and a minimal birth weight of 2.5 *kg*.

The vocalization from the babies was recorded up to three minutes after the venous puncture, with a sampling rate of 44000 *Hz* at a distance of 0.5 meters from the babies head.

In infant cry analysis there are three types of vocalization as described in [7]: phonation (voiced cries), hyperphonation (high pitched cries) and disphonation (turbulence, voiceless cries). In order to evaluate the relation between the variation in the frequency and amplitude of the cry, only manually selected voiced cry bouts (phonations) are used.

2. Preprocessing

In the preprocessing step all the signals were first band-pass filtered using a 6th order Butterworth filter with lower and upper cut-off frequencies being 200 *Hz* and 2500 *Hz* respectively. After filtering, the signal was downsampled to 11025 *Hz*.

3. Ensemble Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is a recently developed technique, proposed by Huang *et al.* [6] for decomposing any complicated time series into a finite set of Intrinsic Mode Functions (IMF) during a so-called sifting process. Intrinsic mode functions are meant to be mono-component, orthogonal to each other and a set of IMF's should be complete. Here the term *mono-component* means that all the IMF's contain only one frequency at the time, which is called *Instantaneous Frequency* (IF). Orthogonal property states that different IMF's do not have similar frequency content. The amplitude of the IMF is the power of the frequency component at the particular time instant and is called *instantaneous amplitude* (IA).

The advantage of the EMD to other available methods for time-frequency analysis (such as Short Time Fourier Transform (STFT) or Wavelet transform) is that EMD is completely data driven and does not impose any predefined assumption about the signal. Therefore, this method allows for properties of the signal to be extracted in a natural way.

Since a voice signal is known to be composed of a fundamental (pitch) frequency and its harmonics, we are able by means of EMD to follow amplitudes and frequencies of only one of these modes at the time instead of looking to all of them, mixed up in the original signal itself.

One of the major drawbacks of the original EMD is that it is highly sensitive to noise. This leads to the frequent appearance of mode mixing, which is defined as a single IMF either consisting of widely disparate scales, or a signal of a similar scale residing in different IMF components. To overcome the problem we used in this paper the advanced, noise-assisted data analysis method recently proposed by Huang et al. [8], called *Ensemble Empirical Mode Decomposition* (EEMD). This method defines the IMF as an ensemble of trials, each consisting of the EMD of the signal plus white noise of finite amplitude. The algorithm is outlined below.

1. An amount of white noise (in our case with standard deviation being 20% of the standard deviation of the signal) is added to the original signal itself
2. EMD is applied and IMF set is derived
3. Steps 1 and 2 are repeated a few times (in our case 100), which leads to an ensemble of 100 different sets of IMF's
4. Average over the ensemble in order to obtain a set of averaged IMF's

Taking into account properties of white noise, noisy components are expected to cancel out and a set of noise-free IMF's is derived. Increasing the ensemble size would give more accurate estimates of IMF's, but the computational efficiency is decreased. The EEMD technique showed to perform well with voice signals [8], and therefore we decided to use it in this study.

D. Extracting Frequencies and Amplitudes

After the EEMD has been applied and the set of IMF's has been extracted, instantaneous frequencies (IF) and instantaneous amplitudes (IA) have been computed. The regular way to compute IA's and IF's from the IMF's is to apply the Hilbert Transform (HT), which well behaves when applied to IMF's since they are monotonic and locally symmetric to the zero level. However, in time instants where large and fast changes in amplitude occur, cubic spline fitting, which is

incorporated in the sifting process is not able to follow these changes and hence IMF's are not always ideal. Due to this phenomenon, the HT will give non-accurate instantaneous frequencies in those places.

To overcome these problems the time-frequency spectrum has to be averaged in some way. This can be achieved using Gaussian windowing as in [6] which will concentrate the averaged spectrum of the particular window in the center of that window.

Another way to average the spectrum is to reassign it, as proposed by Flandrin [9,10]. This method relies on the principle that frequency values around the observed point have no reason to be symmetrically distributed around that point, as Gaussian averaging supposes. Therefore their average should not be assigned to this point, but rather to the center of gravity of the observed domain. In this work we used the reassigned spectrogram.

4. Computing Correlations

After extracting the IMF's, and deriving the reassigned spectrogram as mentioned in the previous sections, we follow the changes in those which contain the information about the fundamental frequency (between 400 and 900 *Hz*) and studied how these changes correlate with the *Intensity Contour* (*IC*), which is the envelope of the signal. To compute the envelope of the signal we first computed the absolute value of the signal, and then filtered it with the 8th order low-pass Butterworth filter at cut-off frequency 80 Hz.

When envelopes and frequencies have been computed we were interested in the correlations between both. To see how they correlate we computed the averaged values of the frequencies and amplitudes over half-overlapping windowed epochs of 512 data points (i.e. 46 *ms*) using a Hamming window, and then computed the correlations between those averaged values.

III. RESULTS

In this section we present the results of signal processing as well as the computed correlations between the intensity contour of the voice signal and the instantaneous frequency of the modes containing the fundamental frequency. The 3rd and 4th intrinsic modes were observed to carry most of the signal power as well as the fundamental frequency information (400 to 900 *Hz*). This is illustrated in Figure 2.

In Figure 1 and Figure 2 we show the spectrogram of a randomly chosen cry bout from the database before and after reassignment respectively. It is obvious that in the reassigned spectrogram blurring is significantly suppressed, which allows

us to observe frequencies of the intrinsic modes more accurately and to easily compare it with the intensity contour of the voice signal.

Since we had 800 cry bouts for 24 babies, we don't show here for the sake of readability the correlations for every cry bout separately.

Instead in Table I we present the mean correlations between the fundamental frequency and the intensity contour over all cry bouts for each subject. Corresponding standard deviations are also given in the third column of the table. Additionally, we wanted to check whether there is a relation between the correlation coefficient of the babies and their MBPS score. Therefore, in the last column in Table I we add the corresponding MBPS values. To better visualize any relationship between the results in Table I,, we plot the ratio of the mean correlation value and standard deviation from Table I against the corresponding MBPS value (Figure 3). The correlation coefficient between the two mentioned measures is 0.55 with the p-value of 0.006. The results are discussed in the following section.

IV. Discussion and Conclusion

It is shown that EMD is a technique that can be successfully applied on newborn cry signals. More specifically, it was found that the 3rd and 4th IMF contain most of the signal power which is related to the fundamental frequency. The frequency content of these modes also confirms that. In the particular case illustrated in figure 2, the reassigned spectrogram shows that IMF 3 and 4 seem to partly overlap in the frequency domain, however in different time instants. Since the fundamental frequency of the cry is located in this particular frequency band, those two modes mutually switch the essential information on the fundamental, implying that the fundamental frequency can be found in both of these modes. This is not always the case, since sometimes only one of these modes is dominant.

When calculating the mean value of the correlation between the IC and frequency variation a great spread is observed, ranging between 0.39 (±0.22) and 0.83 (±0.07).

In addition, a considerably high standard deviation of the correlations is observed compared to the mean correlation per subject. This indicates a high intra-subject variability. Taking the ratio between the mean correlation value and the standard deviation gives information about the consistency of the investigated cry bouts during a crying sequence. By plotting this ratio against the MBPS values (Figure 3), a trend can be seen that infants with a higher MBPS score show a higher mean/standard deviation ratio. This indicates that babies with higher MBPS score have higher mean and lower standard

deviation of correlation between IC and frequency. In other words the IC and frequency are more coupled in more stressed infants. This is a careful statement as more data is required to validate this. Wolf suggested that prior to one month of age, infant cries are highly reflective and undifferentiated, but that shortly afterwards cries progressively reflect various psycho-physiological states like hunger and pain [4]. In contrast, Bellieni et al. were able to document a correlation between cry features and DAN score in term neonates [1].

The same problem, using a classical autoregressive spectral estimation method, has been introduced in [11], including a more detailed clinical interpretation of the problem. The use of EMD allows to better extract, through the selection of one IMF, the essential information on the fundamental frequency and the intensity contour. Compared to [11], EMD more clearly relates correlation between frequency and intensity contour with MBPS, thereby showing amore positive trend.

ACKNOWLEDGMENTS

Research supported by Research council KUL: GOA-AMBioRICS, GOA-MANET, CoE EF/05/006, by DWTC: IUAP P6/04 (DYSCO); by EU: Neuromath (COST-BM0601), by ESA: Cardiovascular Control (Prodex-8 C90242)

KA is supported by the Fund for Scientific Research, Flanders (Belgium) (FWO Vlaanderen) by a Fundamental Clinical Investigatorship (1800209N)

REFERENCES

1. Bellieni CV, Sisto R, Codelli BM, Buonocore G. Cry features reflect pain intensity in term newborns: an alarm threshold. Pediatric Research. Jan 2004;55(1):142-6
2. Facchini A, Bellieni CV, Marchettini N, Pulselli FM, Tiezzi EB. Relating pain intensity of newborns to onset of non-linear phenomena in cry recordings. Phys Lett. 2005;338:332-7
3. Moller S, Schonweiler R. Analysis of infant cries for the early detection of hearing impairment. Speech Com. 1999;28:175-93
4. Wolf P. The natural history of crying and other vocalization in early infancy. Determinants of infant behavior. 1969;81-109
5. Barr RG, Hopkins B, Green JA. Crying as a sign, a symptom and a signal.Mac Keith Press. 2000;143-6

6. Huang NE, Shen Z, Long HR, Wu MC, Shih HH, Zheng Q et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc London. 1998;454:903-995
7. LaGasse L, Neal A, Lester B. Assessment of infant cry: acoustic cry analysis and parental perception. Ment Retard Dev Disabil Res Rev. 2005;11(1):83-93
8. Zhaohua W, Huang NE. Ensemble empirical mode decomposition: A noise assisted data analysis method. Centre for Ocean-Land-Atmosphere Studies, Tech Rep 193.Available from: http://perso.ens-lyon.fr/patrick.flandrin/EEMD.pdf
9. Rilling G, Flandrin P, Goncalves P. On Empirical Mode Decomposition and its algorithms. IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP03 [internet]. 2003 Jun; Available from: http://perso.ens-lyon.fr/patrick.flandrin/NSIP03.pdf
10. Auger F, Flandrin P. The why and how of time-frequency reassignment. Proc on IEEE International Symposium on Time-Frequency and Time-Scale Anal. Oct 1994;197-200. doi: 10.1109/TFSA.1994.467259. ISSBN: 0-7803-2127-8
11. Silva M, Mijović B, Van Den Bergh B.R.H, Allegaert K, Aerts J.M, Van Huffel S, Berckmans D. Decoupling between fundamental frequency and energy envelope of neonate cries. Early Human Development. Accepted.