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

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

1 Department of Electrical Engineering (ESAT), Kasteelpark Arenberg 10 bus 2446, B-3001 Leuven, Belgium
2 Department of Biosystems, Katholieke Universiteit Leuven, Kasteelpark Arenberg 30, B-3001 Leuven, Belgium
3 Department of Psychology, Tilburg University, P.O.Box 90153, 5000 LE Tilburg, The Netherlands and Department of Psychology, Katholieke Universiteit Leuven, Tiensestraat 102, B-3000 Leuven., Belgium
4 Department of women and child, Group Biomedical Sciences, Katholieke Universiteit Leuven, Herestraat 49, B-3000 Leuven, Belgium

Abstract— **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. 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). The aim of this work was to evaluate if 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. 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.**

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

A. 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 44100 *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 [6]: 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.

B. Preprocessing

In the preprocessing step all the signals were first band-pass filtered using 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*.

C. Ensemble Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is a recently developed technique, proposed by Huang *et al.* [7] 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).

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. [6], called *Ensemble Empirical Mode Decomposition* (EEMD). This method defines the IMF as an ensemble of trials, each consisting of the EMD of the signal with the added 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 [7], 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 have 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 sudden changes and hence IMF's are not always ideal. Due to this phenomenon, the HT will give non-accurate instantaneous frequencies and amplitudes 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 [8,9]. 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.

E. Computing Correlations

After extracting the IMF's, and deriving the reassigned spectrogram as mentioned in the previous sections, we realized that the 3rd and the 4th IMF carry most of the signal power, and therefore we decided to follow changes in these modes and to see how these changes correlate with the *Intensity Contour* (*IC*), which is the envelope of the signal. These modes also carry information about the fundamental frequency, according to their value (between 400 and 900 *Hz*) 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 the two. To see how they correlate we computed the averaged values of the frequencies and amplitudes over half-overlapping windowed epochs of 512 data points 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 3rd or 4th intrinsic mode.

In Fig. 1 and 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.

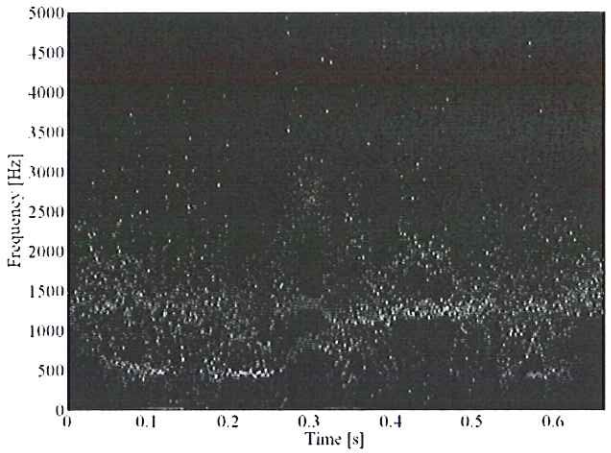


Fig. 1 Spectrogram of the EMD treated cry signal before reassignment.

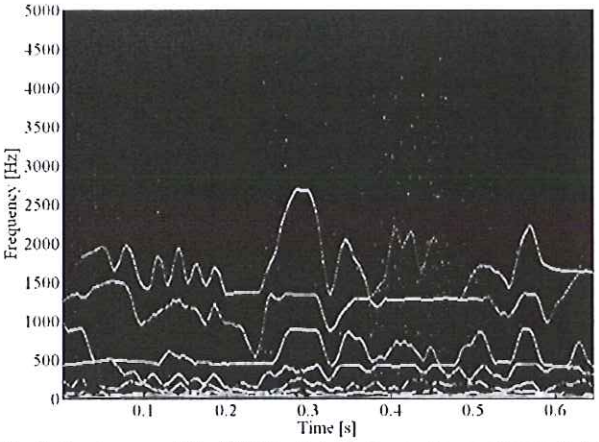


Fig. 2 Spectrogram of the EMD treated cry signal after reassignment. The different lines represent different IMF's

Instead in Table I we present the mean correlations between the fundamental frequency and the intensity contour over all cry bouts for each subject and the corresponding standard deviations. We also 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. Additionally, we plot the ratio of the mean value and standard deviation of the correlation coefficients of each cry bout per subject against the corresponding MBPS value in Fig 3. The results are discussed in the following section.

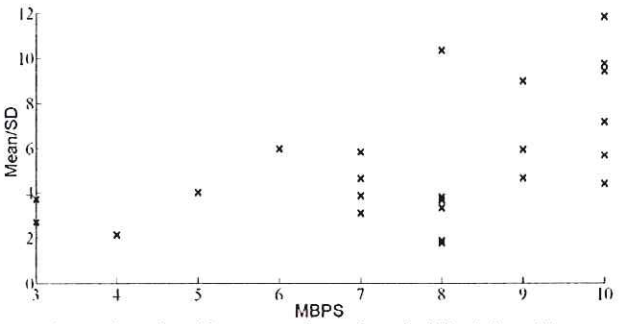


Fig. 3 The ratio of the mean value and standard deviation of the correlations over all cry bouts for a particular subject with respect to the MBPS value

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 frequency, implying that the fundamental frequency can be found in both of these modes. However, in some cases, only one mode 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).

TABLE I
MEAN VALUES AND STANDARD DEVIATIONS OF THE CORRELATIONS BETWEEN THE FUNDAMENTAL FREQUENCY AND THE INTENSITY CONTOUR AND MBPS VALUE FOR EACH SUBJECT

Subject No.	Mean value	Standard Deviation	MBPS
1	0.6116	0.2266	3
2	0.6584	0.1774	3
3	0.5629	0.2626	4
4	0.6831	0.1698	5
5	0.7754	0.1297	6
6	0.8138	0.1754	7
7	0.6806	0.1758	7
8	0.6213	0.2002	7
9	0.6832	0.1173	7
10	0.5641	0.3019	8
11	0.4684	0.1405	8
12	0.6359	0.1725	8
13	0.7286	0.1916	8
14	0.3852	0.2159	8
15	0.8226	0.0795	8
16	0.7754	0.1303	9
17	0.7514	0.0837	9
18	0.7765	0.1668	9
19	0.7946	0.1109	10
20	0.6969	0.1584	10
21	0.7097	0.0754	10
22	0.8391	0.0709	10
23	0.8052	0.0826	10
24	0.6906	0.1216	10

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, one might assume that more stressed infants have less consistent cry bouts. 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 (9). In contrast, Bellieni et al. were able to document a correlation between cry features and DAN score in term neonates (6). In our study, there seems to be a tend that EMD is able to relate the correlation between frequency and intensity contour with MBPS. In order to validate these findings, vocalizations should be monitored over several months to one year during early development.

ACKNOWLEDGMENTS

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

REFERENCES

[1] Bellieni CV, Sisto R, Codelli BM, Buonocore G 2004 Cry features reflect pain intensity in term newborns: an alarm threshold. Pediatr Res 55: 142-146

[2] Fecchini A, Bellieni CV, Marchettini N, Pulselli FM, Tiezzi EB 2005 Relating pain intensity of newborns to onset of nonlinear phenomena in cry recordings. Phys Lett A 338: 332-337

[3] Moller S, Schonweiler R 1999 Analysis of infant cries for the early detection of hearing impairment. Speech Commun 28:175-193

[4] Wolf P 1969 The natural history of crying and other vocalization in early infancy In: Foss B Determinants of infant behaviour, Methuen Press, London, pp 81-109

[5] LaGasse L, Neal A, Lester B 2005 Assessment of infant cry: acoustic cry analysis and parental perception. Ment Retard Dev Disabil Res Rev 11: 83-93

[6] Zhaohua W, Huang N E (2005) Ensemble empirical mode decomposition: A noise assisted data analysis method. Centre for Ocean-Land-Atmosphere Studies, Tech Rep 193.

[7] Huang N E, Shen Z, Long H R, Wu M C, Shih H H, Zheng Q, Yen N C, Thung C C, Liu H H (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc R Soc London 454:903-995

[8] Rilling G., Flandrin P. and Gon¸alv`s P., "On Empirical Mode Decompositionand its algorithms", in IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, Grado(I), 2003.

[9] Auger F. and Flandrin P., "The why and how of time-frequency reassignment", in IEEE International Symposium on Time-Frequency and Time-Scale Analysis, 1994, pp. 197-200.

Adress of the corresponding author

Bogdan Mijovic, Student,
Department of Electronic Engineering
Katholieke Universiteit Leuven
Kasteelpark Arenberg 10, 3000 Leuven, Belgium
Tel: +32-(0)16-321-067
Email: bogdan.mijovic@esat.kuelueven.be