Application of Pattern Recognition Techniques to the Classification of Full-Term and Preterm Infant Cry

*Silvia Orlandi, †Carlos Alberto Reyes Garcia, *,‡Andrea Bandini, §Gianpaolo Donzelli, and *Claudia Manfredi,
*§*Firenze*, ‡*Bologna, Italy*, †*Puebla, Mexico*

Summary: Objectives. Scientific and clinical advances in perinatology and neonatology have enhanced the chances of survival of preterm and very low weight neonates. Infant cry analysis is a suitable noninvasive complementary tool to assess the neurologic state of infants particularly important in the case of preterm neonates. This article aims at exploiting differences between full-term and preterm infant cry with robust automatic acoustical analysis and data mining techniques.

Study design. Twenty-two acoustical parameters are estimated in more than 3000 cry units from cry recordings of 28 full-term and 10 preterm newborns.

Methods. Feature extraction is performed through the *BioVoice* dedicated software tool, developed at the Biomedical Engineering Lab, University of Firenze, Italy. Classification and pattern recognition is based on genetic algorithms for the selection of the best attributes. Training is performed comparing four classifiers: Logistic Curve, Multilayer Perceptron, Support Vector Machine, and Random Forest and three different testing options: full training set, 10-fold cross-validation, and 66% split.

Results. Results show that the best feature set is made up by 10 parameters capable to assess differences between preterm and full-term newborns with about 87% of accuracy. Best results are obtained with the Random Forest method (receiver operating characteristic area, 0.94).

Conclusions. These 10 cry features might convey important additional information to assist the clinical specialist in the diagnosis and follow-up of possible delays or disorders in the neurologic development due to premature birth in this extremely vulnerable population of patients. The proposed approach is a first step toward an automatic infant cry recognition system for fast and proper identification of risk in preterm babies.

Key Words: Infant cry analysis–Preterm newborn–Automatic classification–Acoustical parameters–Feature selection.

INTRODUCTION

Scientific and clinical advances in perinatology and neonatology have enhanced the chances of survival of preterm and very low birth weight neonates. Clinical and ethical demands have emerged regarding the early assessment of these vulnerable children to detect markers of possible developmental deficits. The studies have shown that an early detection of the risks for vulnerable children would allow implementing prevention strategies and policies in childhood.[1]

The crying of newborns and infants is a functional expression of basic biological needs, and emotional or psychological conditions such as hunger, cold, pain, cramps, and even joy.[2] It requires a coordinated effort of several brain regions, mainly brainstem and limbic system and is linked to the breath and the lung mechanisms. Its characteristics reflect the development and possibly the integrity of the central nervous system. Thus, infant cry analysis is a suitable noninvasive complementary tool to assess the physical state of infants particularly important in the case of preterm neonates. Specifically, the distinction between a regular wailing and one with anomalies is of clinical interest.

Being cost-effective and contactless, the study of the newborn infant crying has had an outstanding growth in the last decades. Several studies concern both the subjective auditory analysis of voice and speech and the automatic acoustical analysis in adults. However, with respect to the newborn cry, few automated methods exist, some of them based on classical approaches such as Fourier transform and autocorrelation analysis[2–6] and other on parametric techniques.[7–9] Such methods allow estimating the main acoustical features such as the frequency of vibration of the vocal folds, the vocal tract resonance frequencies, the cry duration, and so forth. However, the high variability of the newborn cry signal has limited the development of methods as robust as those devoted to the analysis of adult voice for its automated analysis.

Preterm infants and infants with neurologic conditions may have different cry characteristics when compared to healthy full-term infants. Qualitative and quantitative research on possible neurophysiological differences between full-term and preterm infants has been carried on since the 1980s. Most of the studies investigated possible differences in infant gender, their neurophysiological maturity, and risks of brain damage for preterm infants caused by deoxygenation due to prolonged crying.[10–16]

In general, the automatic infant cry classification process is a pattern recognition problem. From the infant's cry (input pattern), the goal was to classify the kind of cry or pathology detected on the baby.

Accepted for publication August 7, 2015.
From the *Department of Information Engineering, Università degli Studi di Firenze, Firenze, Italy; †Computer Science Department, Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Puebla, Mexico; ‡Department of Electrical, Electronic and Information Engineering (DEI) "Guglielmo Marconi", University of Bologna, Bologna, Italy; and the §Department of Neuroscience, Psychology, Drug Research and Child Health, Università degli Studi di Firenze, Firenze, Italy.
Address correspondence and reprint requests to Claudia Manfredi, Department of Information Engineering, Università degli Studi di Firenze, Via di S. Marta 3, 50139 Firenze, Italy. E-mail: claudia.manfredi@unifi.it
Journal of Voice, Vol. ■, No. ■, pp. 1-8
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2015.08.007

Several authors have proposed classification methods for a wide range of pathologies. Reyes et al[17–19] have investigated normal, deaf, and asphyxiating newborns using classification methods such as neural networks, genetic selection, and fuzzy logic. Some of these powerful techniques are successfully applied also in the present work and will be shortly described in the next section. Poel and Ekkel[20] present results concerning the classification of newborn cry into normal and hypoxia-related disorder using Radial Basis Function Neural Networks with 85% overall classification performance. Lederman et al[21] propose the classification of infants with cleft palate on the basis of parallel Hidden Markov Models with an average of 91% correct classification rate in a subject- and age-dependent experiment. Mijovic et al[22] propose Empirical Mode Decomposition techniques to assess the existence and extent of decoupling in term neonates and its possible relation to clinical pain expression. Sahak et al[23] applied Combined Support Vector Machine (SVM) and Principal Component Analysis to recognize the infant cries with asphyxia with a classification accuracy of 95.86%. Zabidi et al[24] applied a new algorithm to optimize Mel frequency cepstrum coefficients to extract an optimal feature set for the diagnosis of hypothyroidism in infants using a Multilayer Perceptron (MLP) neural network. Nonaka et al[25] used a Hidden Markov Model architecture. The algorithm yields up to 95% classification precision (86% average) to identify expiratory and inspiratory phases from the baby cries. In the study by Hariharan et al,[26,27] a General Regression Neural Network is used as a classifier for discriminating normal cry signals and pathologic cry signals from deaf infants and babies with asphyxia. Etz et al[28] propose a decision tree to classify infant cries to find differences between infants with normal development, hearing impairment, and unilateral cleft lip and palate, whereas Alaie et al[29] apply Gaussian mixture models to distinguish between healthy full-term and premature infants and those with specific medical problems with a true positive (TP) rate of 80.77% and a true negative rate of 86.96%. Finally, in the study by Singh et al,[30] three different types of infant cries are considered: hunger, pain, and wet diaper. Gaussian mixture models are used to classify the previously mentioned cries.

This nonexhaustive list of studies shows that the newborn cry contains specific features that enable the classification of various diseases and conditions by automated techniques.

This article aims at highlighting and differentiating the features of newborn cry in the two groups of healthy term and preterm infants. To this aim, a robust tool (*BioVoice*) specifically developed for the acoustical analysis of newborn cry is applied[16] that provides 22 acoustical parameters both in time and frequency domain. The proposed classification method allows pointing out the relevant cry features capable to assess differences between preterm and full-term newborns with about 87% of accuracy. This result may be a valuable aid to the diagnosis of possible delays or disorders in the neurologic development due to premature birth in this extremely vulnerable population of patients.

MATERIALS AND METHODS

Recording protocol

Newborn cry signals were recorded in a quiet room of neonatology unit (S. Giovanni di Dio Hospital, Firenze, Italy) and neonatal intensive care unit (Children Hospital A. Meyer, Firenze, Italy), respectively, for term and preterm infants. All parents of the infants were native Italian speakers, and they signed informed consent to participate in this study.

A unidirectional microphone (Shure SM58; Shure Inc. Chicago, IL) was positioned at a fixed distance (25 cm) from the baby's mouth and equipped with Tascam US-144 (TEAC Corp. Montebello, CA) portable audio/musical instrument digital interface (96 kHz/24-bit recording). Recordings were stored on a multimedia laptop in a single channel audio track. The sampling rate was $Fs = 44$ kHz with 16-bit resolution. Each recording lasts at least 15 minutes to include several cry sequences. A cry sequence is defined here as a set of multiple cries, the so-called cry units (CUs). Cry sequences are spaced one from the other by a suitable amount of time, lasting more than 30 seconds.[31,32]

A CU is defined here as a high-energy voiced frame lasting at least 260 ms. This choice comes from literature where different time lengths are considered for CUs, ranging from 60 to 500 ms.[21,33–36] In fact, CUs of very short duration do not allow the assessment of some relevant features such as their melodic shape. Moreover, inspiratory sounds that have duration less than 200 ms must be disregarded.[33]

Database

We recorded 28 healthy term newborns (TN, 17 boys and 11 girls) and 10 preterm infants (PN, 5 boys and 5 girls). Gestational age of TN at birth was between 37 weeks and 2 days and 42 weeks; the weight was between 2400 and 4250 g. Gestational age of PN at birth was between 23 weeks and 5 days and 34 weeks. The weight at birth was between 590 g and 2700 g. At the recording time (20–30 days after birth), the PN gestational age was between 35 weeks and 1 day and 43 weeks and 1 day; the weight ranged between 1380 and 2430 g.

The TN infants were recorded within the first 2 days of life, whereas PN newborns could be recorded only about 20–30 days after birth, because of their long staying in the incubator. Specifically, the PN infants were recorded within the first 45 days after the normal end of pregnancy (37 weeks). PN infants were recorded in a quiet room in the Neonatal Intensive Care Unit at the Children Hospital A. Meyer, Firenze, Italy. The newborns were hospitalized because of prematurity and other diseases: respiratory distress, obstructive sleep apnea, hypoglycemia, bowel obstruction, bleeding, and anemia. TN infants were recorded in a quiet room of the neonatology clinic at S. Giovanni di Dio Hospital, Firenze, Italy. They did not suffer from any disease.

We collected an audio recording for each infant of at least 1 hour of duration consisting of at least 10% of crying. Recordings were performed before the afternoon feeding. Full-term infants usually cried without stimulation, whereas for preterm infants, sometimes a little solicitation was made on the foot