

Automatic Newborn Cry Analysis: a Non-Invasive Tool to Help Autism Early Diagnosis

S. Orlandi^{1,2}, C. Manfredi¹ *Member, IEEE*, L. Bocchi¹ *Member, IEEE* and M.L. Scattoni³

Abstract— Autism Spectrum Disorders (ASD) are complex developmental disabilities that cause problems with social interaction and communication. ASD are associated with motor development problems, more or less blurred, and with perceptual and sensory brain areas. Crying is the infant's earliest form of communication and recent studies connect this original "language" form with autism disturbs. Being completely non-invasive, cry analysis is an appealing approach for early ASD diagnosis to improve rehabilitation. To this aim, we have developed an automatic system to record newborn cry and movements, during the first six months of life with a specific recording protocol.

In this work we present first results of acoustic cry analysis in newborns classified as high-risk subjects being siblings of children already diagnosed as autistics. The work aims at finding possible early ASD signs in high-risk subjects as compared to a group of control subjects based on the fundamental frequency and the vocal tract resonance frequencies. Also, voiced and unvoiced parts of signal and cry-episodes duration are analyzed.

I. INTRODUCTION

Autism Spectrum Disorders (ASD) define a set of mental conditions characterized by abnormal development of some children's ability. These neuro-developmental disorders result in the alteration of communication skills, behavioural and social interaction. Autistic children are different from those considered "normal" for the activities carried out, fixed and repetitive interests, usually stereotyped. Autism was initially described by Kanner in 1943 as a triad of impairments and was defined as specific behavioural signs, and impairments in socialising and communicating. At present, autism comprises a spectrum of disorders and autistic subjects are classified based on different kind of capabilities that depend on their individual intellectual and social development. A lack of social and communication skills are amongst the earliest and most common signs of autism and children with autism generally don't share attention with other people and have difficulties to understand the mental state of others. There are some different theories about autism difficulties, one of these

called "theory of mind" explains some deficits in autism and considers that autistic sufferers have difficulties to interpret and interact with their environment and the people in it. These difficulties in comparison to a "normal" children produce a lack of understanding of their own and others motor intentions. In fact, some studies [1] explain that autistic individuals used advance information to plan their movements when the information was direct. However, their actions were less controlled when their strategies for movement were self-generated.

The true mechanisms involved in autism are still not completely understood but it is believed to be related to abnormal neurotransmission and specific brain regions, in particular limbic system, amygdal region, cerebellum and sub cortical areas that mediate motor control [2, 3].

ASDs involve an increasing number of children. The impairment of mental functions and difficulties in social interaction give reasons for finding non-invasive techniques that might aid in the early diagnosis of ASD. Kanner's triad and other studies don't focus on the early development of children, i.e. before their first year of life.

At present autism can be reliably diagnosed from the third year of age on [4] despite many claims that there are evidential signs long before this age.

However, since the first descriptions of autism, it was found that some interactive and communication difficulties are present, albeit mild, in the first months of life. Recent studies support the hypothesis of a strict relationship between autism and cry [5], [6], but at present no reliable method for a fully automated analysis is available. Besides, motor difficulties, hypotonia, motor apraxia and dyspraxia are main symptoms in autistic children [7].

Crying is really a form of language because is the first method of communication between parent and child: most parents can distinguish a painful cry from a hunger cry, also tested with acoustic analysis. Acoustic analysis to identify parameters such as fundamental frequency (F0), intensity, vocal tract resonance frequencies, length of each cry-episodes has been carried out.

In our project we search normative ranges for motor and acoustic parameters in a population of about 200 newborn/infants (male and females), as up to now ranges have been defined for older infants only. We aim at implementing the automatic analysis of general movements, i.e. particular movements in newborns that can be clearly detected in the first 6 months of life [8] and the study of abnormal cry features that occur in several neural disorders [9], according to a standard protocol defined in accordance with expert clinicians [10]. Indeed the correlation between cry and general movements has never been reported before though

*Research supported by the Italian Ministry of Health Grant (GR3), Young Researcher 2008, "Non-invasive tools for early detection of Autism Spectrum Disorders".

^{1, 2} S. Orlandi is with the Dept. of Electronics & Telecommunications, Università degli Studi di Firenze, Firenze, Italy and with Department. of Electronics, Computer Science and Systems, Università degli Studi di Bologna, Bologna, Italy (e-mail: silvia.orlandi@unifi.it)

¹ C. Manfredi is with the Dept. of Electronics & Telecommunications, Università degli Studi di Firenze, Firenze, Italy.

¹ L. Bocchi is with the Dept. of Electronics & Telecommunications, Università degli Studi di Firenze, Firenze, Italy.

³ M.L. Scattoni is with the Dept. of Cell Biology and Neurosciences, Istituto Superiore di Sanità, Roma, Italy.

very peculiar, due to the connection between crying and the functioning of brainstem and limbic system, both areas being compromised in children with ASD [3].

In this work we focus on the fundamental frequency development in newborns during the first six months of life and on its reliable estimation comparing control healthy subjects and “high-risk” infants, i.e. siblings of children already diagnosed as autistics. Also, the first and second resonance frequencies, the voiced and unvoiced parts of the signal and cry-episodes duration are analyzed.

II. MATERIAL AND METHODS

The detail of our project, acquisition protocol and acquisition system and some steps of audio signal processing have been already described in [6, 10]. For clarity we summarize the method and describe the relevant improvements in the following sections.

A. Acquisition Protocol

The project will recruit a set of about 200 control subjects and a set of high risk subjects to be followed prospectively (tentatively, 20 subjects). For accurate and detailed assessment of behavioural measures time points at 10 days, 6-12-18 and 24 weeks of child’s life were fixed. This would allow detecting early modifications of ASD, while avoiding biases. Presently, no automatic tool is available for early detection of autistic children. Recent advances have resulted from prospective studies carried out on high risk infants (later-born infant siblings of children diagnosed with ASD), based home videotapes, that suffer from different quality of recordings.

The acquisition procedure proposed here is totally not invasive, thus minimizing the ethical issues involved in the recruitment of control subjects and high risk infants.

Each subject is involved in a set of measures, scheduled every six weeks, starting a few days after birth up to the 24th week of life.

The protocol requires the recording of a set of personal and anamnesis data, together with the acquisition of a video signal and an audio track [10].

In this work, we present results from 12 control subjects and 2 high-risk infants, obtained during the 1st, 2nd, 3th, testing session, i.e. recorded at 10 days, 6th and 12th week of life.

Informed consent has been obtained from parents. The protocol was approved by the local ethical committee (Istituto Superiore di Sanità, IRCCS Fondazione Stella Maris, and IRCCS Pediatric Hospital Bambino Gesù).

B. Acquisition System and Infant Recorder

The acquisition system has been designed for being used in the patient home, thus minimizing the discomfort for the involved subjects and the impact of the external environment on children habits. Hence, the basic requirement is the ease of transport and assembly of the system.

It includes a laptop connected to a high-speed USB video camera (Logitech HD pro webcam C910), able to provide a 1280x1024 pixel video stream, an external audio acquisition

device (Tascam US-144-MK2), and a professional microphone (Shure SM58).

The system reported in Fig.1 shows the acquisition window of our software tool, named *Infant Recorder*.

Infant Recorder allows for the recording of personal and anamnesis data, of a video signal (general movements) and an audio track (newborn cry). It allows to manage several recording sessions for each subject and the generation of a database to avoid data loss or mislabelling.

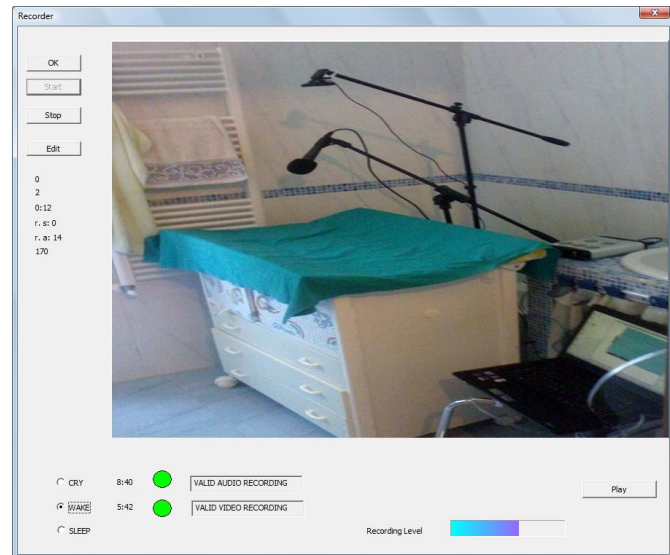


Fig. 1. Acquisition system and acquisition window.

Infant Recorder is written in C++ language using the MFC architecture for the user interface, and the OpenCv image processing library for video acquisition and recording.

A simpler interface with respect to [10] has been built, allowing to record transitions of the subject from calm state to crying state and vice versa. The user-friendly interface includes a main window for selecting the patient from the database, some dialog windows for editing text data, and an acquisition window (Fig. 2), that allows previewing video and audio quality before starting the recording, thus improving the quality of the recordings.

Other new features allow selecting the newborn’s state (crying or calm awake) for the acoustic and the general movement analysis.

A set of markers are stored together with the multimedia tracks, indicating the state transitions of the subject, easing the automatic analysis of the signals.

C. Data management

As the acquisition procedure is performed at patient’s home, the *Infant Recorder* uses a local database for storing the personal and anamnesis data. Multimedia files are stored in a dedicated folder, while the database contains only a reference to the file.

The constraints related to data security and privacy required using an encrypted file system (TrueCrypt) for storing all sensitive data. The central database is located on a shared server, which can be accessed through internet connection. Data inside the central database are anonymized and only an identifier is used to address subjects.



Fig. 2. Infant Recorder and acquisition window.

D. Data Analysis

The analysis of cry and features extraction has been performed using BioVoice, a custom software developed in Firenze under Matlab language.

In this work we have considered the following parameters: length and average amplitude of the cry episode, fundamental frequency F0 and the first two resonance frequencies (F1 and F2, respectively).

Detection of crying episodes (voiced frames lasting at least 150ms) is performed using a robust newly developed Voiced/Unvoiced detection procedure [6]. The procedure is summarized here: first of all the signal is band-pass filtered in the range 200 Hz–1500 Hz, to reduce noise and artefact. Voiced frames are selected according to a modified version of the method implemented in BioVoice tool for adult's voices. Modifications mainly concern selecting the most appropriate frame length and thresholds to deal with high-pitched, quasi-stationary signals as newborn infant cries are, disregarding noise and silence while retaining those irregularities that may be a diagnostic index of possible pathologies or malfunctioning.

On selected voiced frames the fundamental frequency F0 is estimated with a two-step procedure that was already found successful when analysing newborn cries [6, 11]. For F1 and F2 estimation a robust parametric technique is used, obtained by peak picking in the power spectral density (PSD) plot. To increase spectral resolution, PSD estimation was based on autoregressive (AR) models whose order q varies according to the signal characteristics.

III. RESULTS

At present we have collected data from 24 children, namely 17 control cases and 7 high-risk. Audio signals (maximum length: 5 minutes) consist of 90 recordings. From each recording, 90 seconds of cry have been analyzed. We have found two kinds of cry: hunger and boring cry.

Here we present the first results of the control group and of two high-risk subjects that resulted positive to the general

movements analysis carried out by clinicians. Results are limited to the first 12 weeks of life, as recordings started few months ago and are still going on.

The performance of the V/UV detector has been assessed by qualitative inspection of the audio signal superimposed on the output of the voice detector and correct extraction of cry episodes was already presented in [6].

BioVoice was capable to accurately identify cry events, while preserving possible irregularities that are characteristic of the signal itself, namely pitch jumps and abrupt transitions from different phonation types. These are also found with Praat [12]. However, Praat tends to disregard as unvoiced much more frames than BioVoice, which were instead considered significant for diagnosis by clinical experts.

First results show some differences between normal and high risk cases. As concerns fundamental frequency, results confirmed a lower frequency range in the high risk newborns with respect to control subjects [6]. Results at 10 days, 6 and 12 weeks of life are reported in Table I.

TABLE I.

FUNDAMENTAL FREQUENCY RESULTS

Fundamental Frequency						
	10 days		6 weeks		12 weeks	
	mean	std	mean	std	mean	std
17 Control subjects	428.2	145.9	447.9	166.8	477.2	188.7
2 H-R subjects	401.0	108.2	417.7	89.9	398.0	93.3

Results show that in most cases F0 is lower in high-risk infants than in normal subjects, while F1 and F2 reach higher values. An example is reported in Figs.3 and 4 that show the spectrogram of 30s of cry, respectively for one high-risk and a normal subject, both at 10 days of age, with the tracking of F0, F1 and F2 superimposed. Also, less cry episodes are found in the high-risk subject.

An important result concerns the number and length of cry episodes and their “melody” in hunger and boring cry: healthy newborns show large differences between the two cases, while high-risk subjects show a similar pattern in both cases. Moreover they present less cry episodes with long pauses between them.

Table II summarizes the results for the percentage of voiced intervals and the number of the cry episodes in 90 seconds of recording at 10 days, 6 weeks and 12 weeks of life.

The results of these first experiments indicate that appreciable differences can be found between normal and high-risk newborns. However, at present, the sample size, especially as concerns high risk subjects, is too small to assess the statistical significance of these differences. Work is in progress in order to increase the sample size and to investigate other cry parameters to identify precursors of autism, such as the melody and the time before the newborn

starts to cry, along with other parameters concerning movement.

TABLE II
CRY EPISODES RESULTS:
PERCENTAGE OF VOICED TRACTS AND NUMBER OF CRY- EPISODES

	Cry-episodes					
	10 days		6 weeks		12 weeks	
	%Voiced	episodes	%Voiced	episodes	%Voiced	episodes
17 C. Infants	48.5	46	39.8	46	49.0	48
2 H-R Infants	24.0	28	26.9	29	27.9	32

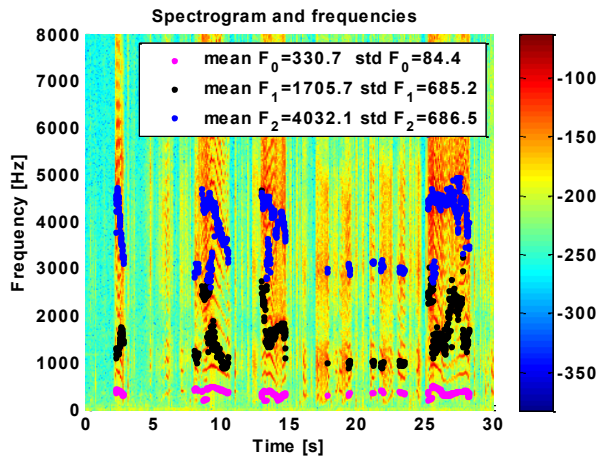


Fig. 3. Spectrogram of boring cry of high-risk infant at 10 days of life.

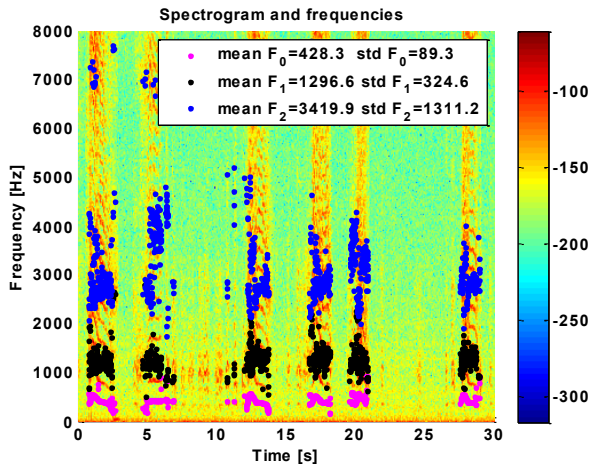


Fig. 4. Spectrogram of boring cry of control case at 10 days of life.

Due to the small number of data available, at present it is not possible to compare these results with those coming from infants suffering from, or in risk of, having health problems other than the autism disorders. This point will be exploited in future work.

If non-invasive parameters suited for an early diagnosis of autism spectrum disorders could be found, a more effective intervention could be applied for better recovering .

ACKNOWLEDGMENT

Supported by the Italian Ministry of Health Grant (GR3), Young Researcher 2008, “Non-invasive tools for early detection of Autism Spectrum Disorders”.

REFERENCES

- [1] C. Glazebrook, D. Gonzalez, S. Hansen, D. Elliott, “Current Perspectives on Motor Functioning in Infants, Children, and Adults With Autism Spectrum Disorders”, *Journal of the American Physical Therapy Association*, vol. 91, July 2011, pp. 1116-1129.
- [2] M. L. Bauman, T.L. Kemper, “Neuroanatomic observations of the brain in autism: a review and future directions”, *International Journal of Developmental Neuroscience*, vol. 23, no. 2-3, April-May 2005, pp. 183-187.
- [3] J. Schulkin, “Autism and the amygdala: An endocrine hypothesis”, *Brain and Cognition*, vol. 65, 2008, pp.87-99.
- [4] E. Werner, G. Dawson, J. Osterling, N. Dinno, “Brief report: recognition of autism spectrum disorder before one year of age: retrospective study based on home videotapes”, *Journal of autism developmental disorders*, vol.30, 2002, pp. 157-162.
- [5] P. Venuti, G. Esposito, Z. Giusti, “A qualitative analysis of crying and vocal distress in children with autism”, *Journal of Intellectual Disability Research*, vol.48, 2004, pp. 4-5.
- [6] S. Orlandi, L. Bocchi, C. Manfredi, M. Puopolo, A. Guzzetta, S. Vicari and M.L. Scattoni, “Study of cry patterns in infants at high risk for autism”, Proc. 7th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications, August 25-27, 2011, Firenze, Italy, pp. 7-10.
- [7] X. Ming, M. Brimacombe, G. C. Wagner, “Prevalence of motor impairment in autism spectrum disorders”, *Brain and Development*, vol. 29, no 9, October 2007, pp. 565-570.
- [8] P.F. Bolton, J. Golding, A. Emond, C. D. Steer, “Autism Spectrum Disorder and Autistic Traits in the Avon Longitudinal Study of Parents and Children: Precursors and Early Signs”, *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 51, no 3, March 2012, pp. 249-260.
- [9] K. Michelsson, O. Michelsson, “Phonation in the newborn, infant cry”, *Int J Pediatr Otorhinolaryngol*, vol. 49, suppl. 1, October 5, 1999, pp. 297-301.
- [10] L. Bocchi, S. Orlandi, C. Manfredi, M. Puopolo, A. Guzzetta, S. Vicari, M.L. Scattoni, “Early diagnosis of autism spectrum disorders – design of the data acquisition and management system”, 5th Europ. Conf. of IFMBE, 14-18 September 2011, Budapest, Hungary.
- [11] C. Manfredi, L. Bocchi, S. Orlandi, L. Spaccaterra, G.P. Donzelli, “High-resolution cry analysis in preterm newborn infants”, *Medical Engineering & Physics*, vol. 31, no.5, 2009, pp. 528-532.
- [12] O. Michael, “GSU Praat Tools: Scripts for modifying and analyzing sounds using Praat acoustics software”, *Behavior Research Methods*, Springer New York, vol. 40, no. 3, August 2008, pp. 822-829.