Acoustic Study of a Neonatal Intensive Care Unit: Preliminary Results

Ganna Raboshchuk[1], Climent Nadeu[1], Blanca Muñoz Mahamud[2], Ana Riverola de Veciana[2], and Santiago Navarro Hervas[2]

[1] TALP Research Center, Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, c/ Jordi Girona 1-3, 08034, Barcelona, Spain
{ganna.raboshchuk,climent.nadeu}@upc.edu
[2] Hospital Sant Joan de Déu Barcelona, Psg. Sant Joan de Déu, 2, 08950, Esplugues de Llobregat, Barcelona, Spain
{bmunozm,ariverola,snavarroh}@hsjdbcn.org

Abstract. The acoustic environment of a typical neonatal intensive care unit is very rich and may contain a large number of different sounds, which reflect the activities taking place in it. There exists a medical concern about the effect of that acoustical environment on a preterm, since loud or particular sounds may be harmful for the preterm's further neurological development. In the reported work, an audio database of ten acoustic scenarios was recorded using two microphones, one inside and the other outside the incubator. Preliminary results of the acoustical analysis of sounds are presented along with the experiments on automatic detection of the presence of vocalizations.

Keywords: Neonatal intensive care unit, acoustic analysis, acoustic event detection

1 Introduction

Newborns born between the 24 and 32 weeks of gestation (very low birth weight preterms) commonly have health problems and must be admitted to the Neonatal Intensive Care Unit (NICU), which, in most of the cases, is crucial for their survival. The increased survival and reduced neonatal morbidity of preterm infants in the past three decades has not always been accompanied by an improvement in their neurological development [1]. The negative or stressful environmental impact of Neonatal Intensive Care Units (NICUs) on the developing brain has been widely documented [2–5].

At present it is known that when a premature passes from an ideal intrauterine environment to an environment with multiple unexpected stimuli (light, noise, proprioceptive stimuli) this may have a negative effect on its neurodevelopment. Inadequate, unexpected noises replace natural hearing placental stimulation [6]. Usually, to respond to so many extraneous stimuli in an organized way is difficult for a preterm newborn. An important negative effect of noise in the NICU is the one it has on sleep, which is essential for the neurodevelopment of prematures [7].

The problem is aggravated because in NICUs the maximum noise limits recommended [8] are exceeded frequently [9]. Different ways that have been proposed to deal with this can be divided into two groups. The major group of methods is aimed at analyzing and changing the acoustical environment of the NICU (for example, by planning a more rational distribution of the wards [10] or controlling and reducing the activities taking place in it [11]). The other group of methods directly concerns the preterm baby and implies protecting a baby with special accessories (earmuffs) [12, 13]; but these methods have not proven their effectiveness. The line of work pursued in this paper, which consists of analyzing the acoustical environment of a preterm baby, refers to the first group of methods.

The acoustic environment of a preterm baby admitted into a NICU has been the object of a number of reported studies during the last two decades. These works analyze the environment placing a microphone both inside [14] and outside the incubator. Usually, sound is represented only by its intensity level and just a few works analyze sound spectra [15]. To our knowledge, very little studies considered the intensity levels of specific sounds [16] or analyzed specific conditions of the NICU acoustic environment [17].

The objective of this work is the analysis of the different scenarios from our NICU acoustic environment in order to describe the existing acoustic events and their spectro-temporal characteristics. That analysis is done both manually, from the labeled events, and by employing automatic recognition or detection systems. Unlike most previous works, the whole content of the audio signal is used, not only its intensity level. In this paper, we present some preliminary outcomes that include the creation of a (limited) annotated audio database, an initial taxonomy and sound description, and first results from automatic detection of vocalizations with Gaussian Mixture Models (GMM) using the produced database. The focus on vocalizations (speech, laughter, cries, cough, etc.) is due to the large extent they are present in the recordings, and the fact that an automatic monitoring system aiming to detect loud vocalizations may help to create a safer environment for a preterm in the NICU.

The paper is organized as follows. Section 2 contains the description of the considered acoustic scenarios and the produced database. In Section 3, the results of the acoustical analysis of the NICU are given. In Section 4, we present and discuss the experimental results of the vocalizations detection task.

2 Acoustic Scenarios and Database Acquisition

The most frequent procedures/activities that take place in the NICU can be described by a set of scenarios. For example, during the day, aproximately every 3 or 4 hours, every preterm baby in the NICU gets nursery care. There is a standard list of interventions that a nurse should do with any baby and other interventions that vary depending on the baby's needs. Each of these nursery care operations can be viewed as a different scenario. Other examples of scenarios are preterm's admission/relocation, kangaroo care, parents entering the NICU, pediatric observations or interventions and NICU cleaning.

Obviously, the number of scenarios needed to describe all the NICU's activities is large. So, at the first stage of the study, a subset of scenarios, which mostly correspond to the daily nursery care related activities, was selected. These scenarios are: Changing a diaper, Measuring blood pressure, Changing an oxygen sensor, Cleaning respiratory secretions, Measuring temperature, Changing temperature sensors, Weighting a newborn, Pediatric observation, Changing medications. Also, a Neutral scenario was defined for including the time periods when no other scenario takes place and the doors of the incubator are closed.

In total, ten recording sessions were carried out in the NICU of Hospital Sant Joan de Deu (HSJD) (see Figure 1), both in the morning and in the afternoon. The overall duration of the acquired audio data is 108.7 minutes. Two electret unidirectional microphones connected to the Olympus LS-5 Linear PCM Recorder were used to make recordings. One microphone was placed inside the incubator, close to the infant ear, and the other one outside the incubator, at aproximately 50cm distance above it.

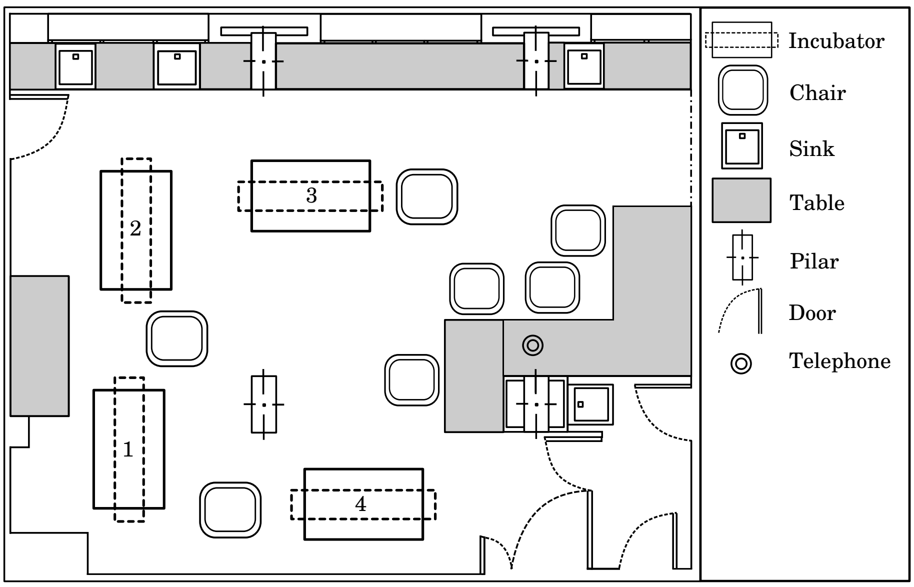

Fig. 1. The scheme of the NICU room with the positions of the incubators and other equipment.

Each recording session includes a subset of the defined scenarios (in average there are 8.6 samples per scenario). We have observed a strong inter-session variability caused by differences in: incubators used during recording sessions;

their position (one of the four positions depicted in Figure 1 were used); nursing procedures depending on the nurse personal style of work and the particular preterm needs; acoustic environment in general (number of people in the room, sounding alarms, etc.).

The audio data was manually annotated using the ELAN tool [18].

3 Acoustic Analysis

A typical NICU environment is characterized by a large diversity of sounds. During our study more than 65 different acoustic events happening in the NICU environment were found and this list is not exhaustive. This number was surprisingly high even for the medical staff working there.

In order to deal with such a variety in an organized way we present a general sound taxonomy in which sounds present in the NICU are structured into acoustically homogeneous classes or groups. The sound taxonomy is represented in Figure 2.

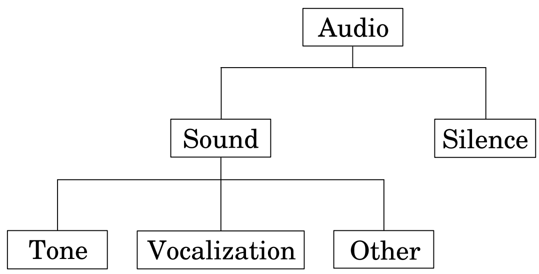

Fig. 2. A general sound taxonomy of a typical NICU.

In general, the set of all the sounds can be split into three main groups (tone, Vocalization, and other) which possess different acoustical properties.

Tone (in music, note) is a sound of distinct fundamental frequency and duration. From a semantical point of view these are informative sounds produced by devices. This group mostly comprises alarms; in our recordings we found at least 13 different types of alarms.

The vocalizations class includes all the sounds that are produced through the vocal tract, either by infant or adult. In this group, at least three semantical categories are distinguished: adult speech, infant vocalizations (mostly cries) and non-speech adult vocalizations (like cough or laughter).

The sounds which cannot be referred to the two previous groups (tone and vocalization) belong to the *other* group. A very specific sound found in this group is the ventilation noise produced by a continuous positive airway pressure (CPAP) device.

The sounds found in the NICU environment have diverse spectro-temporal caracteristics. Regarding the time dimension, we observe sounds which are continuous (like chair moving, or drawer) or impulsive (like knocks, steps, or door slam); sounds which are periodic (like alarms, or CPAP noise) or aperiodic (like spray, or plastic wrapping). With respect to the spectral domain, and referring to the three main groups of sounds defined in the taxonomy: the fundamental frequencies of alarms are concentrated in the range 1-3 kHz, and the content of vocalizations can be observed up to 8 kHz. And the group *other* is even more diverse, since it contains lower- ($\leq$3kHz) and higher-frequency sounds, as well as acoustic events whose content is dispersed across all the frequency range.

Most of the sounds (like steps, door slam, speech) are common to all the scenarios and, in principle, can happen at any period of time. But some of the sounds are specific to the scenario in which they occur, namely they are happening only in special conditions. Table 1 shows the specific sounds for the scenarios from the recorded database.

Table 1. Scenario-specific acoustic events.

Scenario	Specific events
1 Changing a diaper	Diaper wrapping
2 Measuring blood pressure	Putting on/off sphygmomanometer, rhytmic hissing, squeaks of the device
3 Changing an oxygen sensor	Taking off the sensor
4 Cleaning respiratory secretions	Unboxing the tube, clearing secretions (hissing)
5 Measuring temperature	Alarm of the thermometer
7 Weighting a newborn	Buttons of weights, diaper wrapping
9 Changing medications	Buttons and hits of the infusion pump

4 Detection experiments

4.1 Experimental setup and detection system

The experiments were carried out with the part of the recorded database that was labelled. In total 18 files from different recording sessions were used, which are: 14 samples of Neutral scenario, 3 samples of Changing a diaper, and 1 sample of Measuring blood pressure scenario. Only recordings made with the microphone placed outside the incubator were used to keep homogeneous recognition conditions. The original 44.1 kHz recordings were downsampled to 16kHz. The total amount of data used is slightly more than 20 minutes and 58.39% of this time is labelled as vocalizations.

A Gaussian Mixture Model (GMM) based classifier was used to perform vocalizations detection. There are two models: the vocalizations (VOC) model, and the non-vocalizations (NO-VOC) model. Each model consists of one Gaussian pdf with diagonal covariance matrix. The HTK toolkit [19] was employed for training and testing this GMM-based system.

The input signal is successively framed with a Hamming window, being the frame length 30ms and the frame shift 10ms. Concerning the features, 16 frequency-filtered log filter-bank energies (FF-LFBE) [20] along with their 16 first temporal derivatives are extracted from each frame. Therefore, the dimension of the feature vector is 32.

With the likelihoods obtained from the two models, each frame is classified as either VOC or NO-VOC. The resulting time string of hypothesized labels is then smoothed to force grouping of frame labels in order to obtain event labels. Within each of the consecutive non-overlapping smoothing windows of odd length the class to which the majority of the output frames belongs is determined. Then, all the frames within the window are assigned the label of that winner class.

As the dataset is small, in order to obtain more statistically relevant results, a 10-fold cross validation scheme was applied. Each time 9 sessions of data were used for training and 1 session for testing. All the parameters were empirically selected to maximize the recognition performance.

4.2 Results

The system detection performance was evaluated using an event-based or event level metric, which is calculated as the F-score between *Precision* and *Recall* by the equation:

$$Fscore = \frac{2 * Precision * Recall}{Precision + Recall}. \tag{1}$$

Since the accuracy in terms of beginning and end of the VOC time segments is not crucial for the task considered, when calculating *Precision* and *Recall* an event is regarded as correctly detected if at least one of its frames is correctly detected. Note that short events will be removed by the smoothing process after frame-level classification; otherwise, there would be a lot of false alarms in NO-VOC intervals.

The baseline recognition results resulting from the GMM frame classification, and the results obtained when applying a smoothing window are presented in Table 2.

Table 2. The performance of the vocalizations detection (VD) system.

Experiment	Event-based accuracy, %
Baseline	72.90
With output smoothing	96.82

It can be seen that the relative improvement after applying a smoothing window of 23 frames length is 32.92%. Figure 3 shows how the length of the smoothing window influences the detection accuracy of the system. As the window length increase the accuracy also increases, but in the limit all the outputs of the system would belong exclusively to one of the classes, which does not make sense.

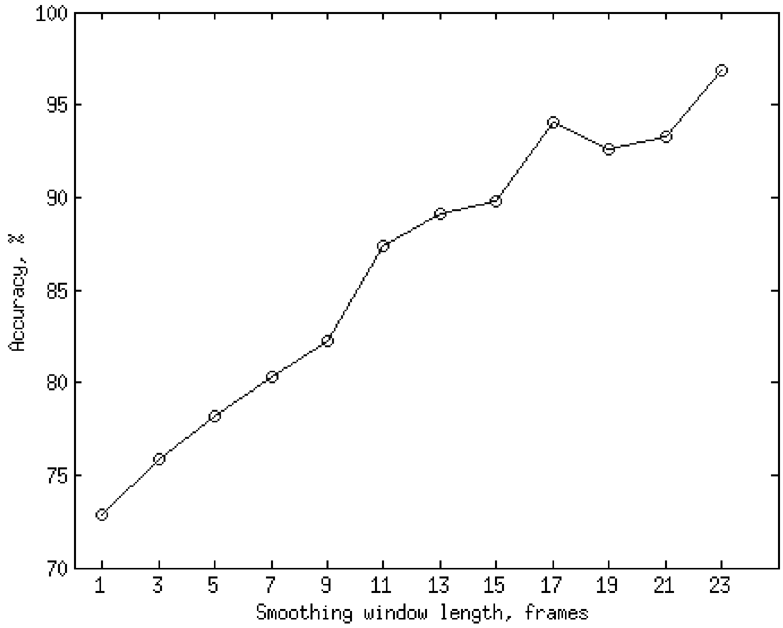

Fig. 3. Influence of the smoothing window length on detection rate.

5 Conclusions

This paper reports the preliminary results obtained by exploring the acoustic environment of a preterm in a NICU. The basic acoustic analysis carried out from a preliminary set of recordings shows the strong acoustic diversity of the NICU. Preliminary results on a vocalizations detection task are rather encouraging.

Apart from collecting a larger database, future work will be oriented on thoroughly analyzing the spectro-temporal characteristics of the typical acoustic events of the NICU, and enhancing the solution proposed for the vocalizations detection task. Also the problem of alarm detection will be investigated.

6 Acknowledgments

The authors are grateful to Vanessa Sancho Torrents and Cristina Borras Novell for their help with database production, and to Eva Bargallo for her contributions at the beginning of this work.

References

1. Wilson-Costello, D., Friedman, H., Minich, N., Fanaroff, A.A., Hack, M.: Improved Survival Rates With Increased Neurodevelopmental Disability for Extremely Low Birth Weight Infants in the 1990s. Pediatrics. 115, pp. 997–1003 (2005).
2. Gray, L., Philbin, M.K.: Effects of the neonatal intensive care unit on auditory attention and distraction. Clin Perinatol. 31, pp. 243–260, vi (2004).
3. Salavitabar, A., Haidet, K.K., Adkins, C.S., Susman, E.J., Palmer, C., Storm, H.: Preterm infants sympathetic arousal and associated behavioral responses to sound stimuli in the neonatal intensive care unit. Adv Neonatal Care. 10, pp. 158–166 (2010).
4. Wachman, E.M., Lahav, A.: The effects of noise on preterm infants in the NICU. Arch Dis Child Fetal Neonatal Ed. 96, pp. F305–F309 (2010).
5. Committee on Environmental Health: Noise: A Hazard for the Fetus and Newborn. Pediatrics. 100, pp. 724–727 (1997).
6. Thomas, K.A., Uran, A.: How the NICU environment sounds to a preterm infant: update. MCN Am J Matern Child Nurs. 32, pp. 250–253 (2007).
7. Graven, S.N., Browne, J.V.: Sleep and Brain Development. Newborn and Infant Nursing Reviews. 8, pp. 173–179 (2008).
8. White, R.D.: Recommended standards for the newborn ICU. Journal of Perinatology. 27, pp. S4–S19 (2012).
9. Krueger, C., Wall, S., Parker, L., Nealis, R.: Elevated sound levels within a busy NICU. Neonatal Netw. 24, pp. 33–37 (2005).
10. Philbin, M.K.: Planning the acoustic environment of a neonatal intensive care unit. Clinics in Perinatology. 31, 331352 (2004).
11. Jousselme, C., Vialet, R., Jouve, E., Lagier, P., Martin, C., Michel, F.: Efficacy and mode of action of a noise-sensor light alarm to decrease noise in the pediatric intensive care unit: a prospective, randomized study. Pediatr Crit Care Med. 12, pp. e69–72 (2011).
12. Aita, M., Johnston, C., Goulet, C., Oberlander, T.F., Snider, L.: Intervention minimizing preterm infants exposure to NICU light and noise. Clin Nurs Res. 22, pp. 337–358 (2013).
13. Duran, R., Ciftdemir, N.A., Ozbek, U.V., Berberolu, U., Duranku, F., St, N., Acuna, B.: The effects of noise reduction by earmuffs on the physiologic and behavioral responses in very low birth weight preterm infants. Int. J. Pediatr. Otorhinolaryngol. 76, pp. 1490–1493 (2012).
14. Benini, F., Magnavita, V., Lago, P., Arslan, E., Pisan, P.: Evaluation of noise in the neonatal intensive care unit. Am J Perinatol. 13, pp. 37–41 (1996).
15. Marik, P.E., Fuller, C., Levitov, A., Moll, E.: Neonatal incubators: a toxic sound environment for the preterm infant?. Pediatr Crit Care Med. 13, pp. 685–689 (2012).
16. Hassanein, S.M.A., El Raggal, N.M., Shalaby, A.A.: Neonatal nursery noise: practice-based learning and improvement. Journal of Maternal-Fetal and Neonatal Medicine. 26, pp. 392–395 (2013).

17. Trickey, A.W., Arnold, C.C., Parmar, A., Lasky, R.E.: Sound levels, staff perceptions, and patient outcomes during renovation near the neonatal intensive care unit. HERD. 5, pp. 76–87 (2012).
18. Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., Sloetjes, H.: Elan: a professional framework for multimodality research. In: Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC), pp. 1556–1559 (2006).
19. Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V., Woodland, P.: The HTK Book, version 3.4. Cambridge University Press (2006).
20. Nadeu, C., Macho, D., Hernando, J.: Time and frequency filtering of filter-bank energies for robust HMM speech recognition. Speech Communication. 34, p. 93–114 (2001).