NEWBORN'S PATHOLOGICAL CRY IDENTIFICATION SYSTEM

Yasmina Kheddache, Chakib Tadj

École de Technologie Supérieure - Electrical Engineering Department
1100, Notre-Dame Ouest, Montreal, Qc, H3C 1K3 Canada

ABSTRACT

In this paper we compare the performance of an identification system of the pathological and normal cries of the newborn, using various methods of characterisation of cries. This system is similar to a speaker identification system. It contains two main parts namely a cry signal characterisation and modeling. We used Mel-Frequency Cestrum Coefficients and Mel Frequency Discret Wavelet Coefficients to characterize the newborn cry signals. We also applied Best Structure Abstract Tree algorithm and the Principal Component Analysis to reduce the number of Wavelet packet transform WPT coefficients. In this study a Probabilistic Neural Network classifier is used. The best result obtained is 96.99 % of correct identification using Best Structure Abstract Tree algorithm.

Keywords Classification, pathologic cry, WPT, PCA, Best abstract Tree.

1. INTRODUCTION

The cry of babies is the first sign of life to the birth. It is an early behavior which plays a fundamental part in survival, health and the development of child. It is also less explicit mean of communication to indicate the distress and the pain. The majority of the babies reached of diseases seem in good health at birth [1]. Our interest in this research project is to analyze the sounds of the new-born cries and to develop a diagnostic system which allows paediatricians and neonatologists detect pathologies affecting newborns in an early stage. For this purpose, we present in this paper the performance of an identification system of a newborn's pathological and normal cry using various techniques of characterisation of cry signals. Pathologic cries concern different pathologies. In this context, we study four types of analysis to characterize acoustic features of these signals, namely, a) MFCC (Mel Frequency Cepstral Coefficients) is based on Fourier analysis and filter bank on a Mel scale. b) MFDWC (Mel Frequency Discrete Wavelet Coefficients) based on wavelet analysis on a Mel scale. c) MFDWC and the application of BSAT (*Best Structure Abstract Tree)* algorithm presented in [2] to modify the structure of the admissible tree and selected nodes. Finally d) application of PCA *(Principal Component Analysis)* on the MFDWC.

This paper is organised as follows: section 2 presents a review of previous works. Section 3 presents a brief definition of the basic techniques used to design a newborn's pathological cry identification system (PCIS). The next sections present all the experiments and results obtained. We conclude this paper with a discussion and a conclusion.

2. PREVIOUS WORKS

Analysis of newborn's cry signals is an important help to clinical diagnosis, because the characteristics of these signals reflect the central nervous system integrity [3]. The researches in this area are divided into three main categories: research on the reasons of cry, research on the development of the cry and research on the relationship between the disease and the characteristics of the cry.

The classification of the cries falls within the automation of a natural perception of cries and its modelling using advanced signal processing techniques. Cohen et al [4] suggested an algorithm based on a simple minimum distance classifier. This algorithm has shown to be successful in classification of hunger and pain cries of healthy full-term infants. Fausto et al. [5] developed an automatic recognition system for three types of newborn cries (normal, deaf, neonatal asphyxiate). They used the linear predictive coding (LPC) coefficients and MFCC and generic characteristics selection system combined with "feed forward input delay neural network". They obtained around 96% of cries recognition. Another approach was also presented by Lederman et al. [6]; they used LPC and MFCC coefficients and Gaussian mixture models (GMM). It was applied on cry signals of infants with cleft lip and cleft palate. They obtained 91% of recognition in subject- and age-dependent. Hariharan et al. [7] presented another system based on wavelet packet transform and a probabilistic neural network (PNN) to identify neonatal asphyxiate / healthy babies cry. They obtained 99% of recognition.

978-1-4673-0382-8/12/$31.00 ©2012 IEEE

3. TECHNIQUES OF PARAMETERIZATION USED

The objective of the acoustic analysis is to extract the most representative coefficients of the cry signal. Several methods are used for this purpose.

3.1. MFCC Coefficients

Represent the most used features in the most recent speaker recognition systems [8]. They have the advantage to be slightly correlated to each other. They can also represent properly various models of cries [6], because they allow decoupling between the characteristics of the vocal tract and the characteristics generated by the source of excitation. The cepstrum of the signal is defined as the inverse Fourier transform of the logarithm of the spectral power density.

3.2. Wavelet Transform

The wavelet transform allows the characterisation of a signal in the temporal and frequency domain. We find several types namely: the continuous wavelet transform (CWT), the discrete wavelet transform (DWT), and the wavelet packet transform (WPT). Each transformation has its advantages and its inconveniences. The WPT has the advantage of providing level by level complete decomposition of the studied signal.

3.3. Principal Component Analysis

The PCA is a method used generally for dimensionality reduction of data. It is based on linear transformation and decomposition of correlated observations into new uncorrelated variables of maximum variance and which are linear combinations of the variables of origins [9]. These new variables are called principal component. In our work, the PCA is used to find a better wavelet basic structure.

3.4. The Algorithm of the Best Structure Abstract Tree (BSAT)

It is an algorithm which allows building the best structure of abstract tree according to a criterion from entropy or energy [1]. We adapted this algorithm and applied for a PCIS as follow:

- Pre-processing of the signal.
- Decomposition of the signal in wavelet packets.
- Calculation for all the nodes of the tree the values of the parameters using the selected criterion (entropy or energies).
- Concatenation of all values to obtain a corpus and a tree for each class of cries (healthy babies cries and sick babies cries) [1].
- Application of the PCA for a given rate of variance, to seek the nodes (wavelet coefficients) which allow a better structure of tree. This structure is specific to each class of cries.

3.5. PCA-Based Algorithm to Reduce WPT Coefficients Dimension

This algorithm allows reduction of the WPT coefficients size for two classes of cries using PCA. The mean steps of this algorithm are:

- Pre-processing of the signal.
- Decomposition of the signal in wavelet packets.
- Calculation for all the nodes of the tree the values of the parameters using the entropy.
- Convert each matrix of WPT (entropy) coefficients specific to a cry sample into only one vector.
- The vectors of WPT (entropy) coefficients of all cries samples are disposed in one matrix N×M (N: number of WPT coefficients (63×Nb of segments of 20 ms in sample cry), M: number of vectors agree to the number of cries sample).
- Application of the PCA for matrix of WPT (entropy) coefficient to reduce row dimension of this matrix.

4. EXPERIMENTAL IMPLEMENTATION

The adopted methodology for the design of a PCIS is represented on the simplified diagram blocks illustrated in Figure 1. Matlab is used for the development of all blocks of the system. We used also “Wavelet Toolbox” and “Neural Network Toolbox”. Praat is used for the segmentation of the cry signals. This is a freeware program for the analysis and reconstruction of acoustic speech signals.

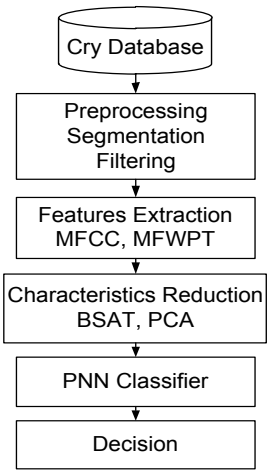


Figure 1: Synoptic of the system

5. DATABASE

Database used contains cry signals of 29 new-born babies. These cries were collected in neonatology department at Saint-Justine hospital in Montreal. They concern full-term and preterm newborn, healthy newborn and newborn who present some diseases as shown in

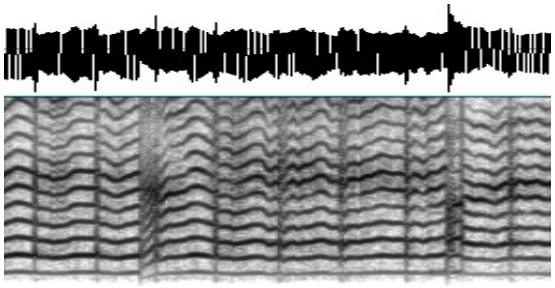
Table 1. The spectrograms of different pathologic cries are show in Figure 2.

Table 1: Summary of available pathologies.

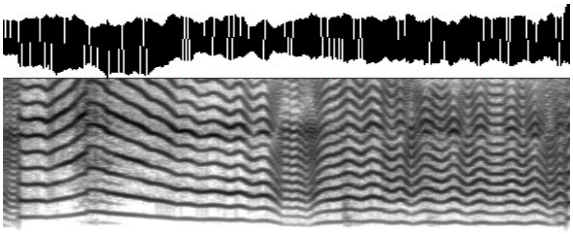
Categories	Pathology
Full-Term	Healthy
	Tetralogy of Fallot
	Thrombosis in the vena cava
	Bovine proteins allergy
Premature	Healthy
	Tetralogy of Fallot
	Anomaly chromosomes X

The conditions in which the cries are registered are: hunger, sampling blood, change of diapers. The recording of cries is done using a small recorder, at a distance of 10 to 30 cm of babies' mouth with a sampling rate of 44.1 kHz. For each baby, three recordings of duration 2 to 3 minutes are made with at least one hour interval after each recording session (over a period of ten days at most). The time, date and gender, date of birth, diagnosis, and reason of cries are noted for each episode of cry.

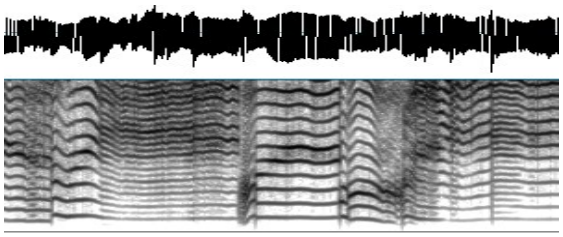
The samples of cries are distributed in two classes, healthy cries and pathological cries. To test the influence of the length's cry samples on the performance of the PCIS, two subsets are used: In the first one, the cry samples were divided into one second segments to create a database of 664 samples of cries (332 healthy and 332 pathological). In the second one, the cry samples were divided into segment of 3 seconds to create 188 samples (94 healthy and 94 pathological). For the test of the identification system ¼ of each subset are used for the tests and ¾ for training. The choice of 1s and 3s samples is to test the influence of the length's cry samples on the performance of the PCIS. It is important to note that the subset with 1s (664) sample is larger then 3 s subset (188*3).



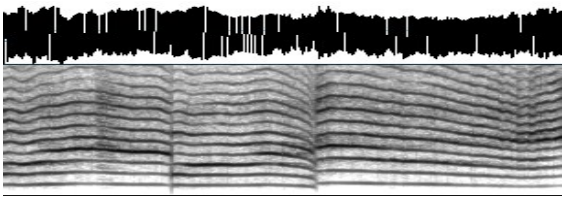
a- Premature healthy baby



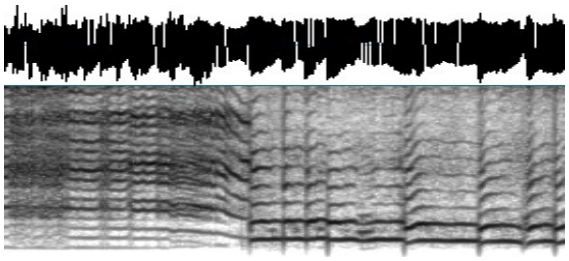
b- Premature baby with Tetralogy of Fallot



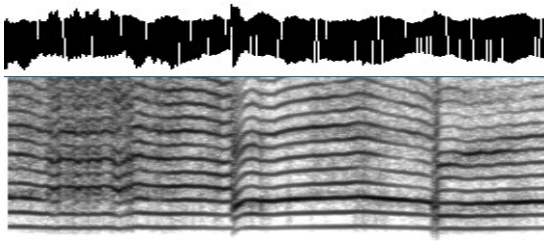
c- Premature baby with Anomaly chromosomes X



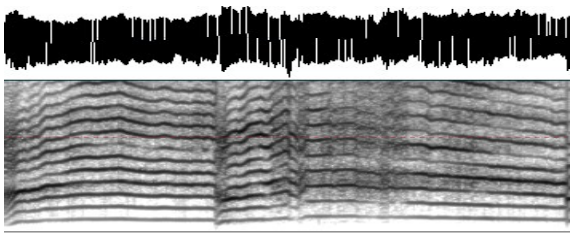
d- Full-Term healthy baby



e-Full-Term baby with Thrombosis in the vena cava



f- Full-Term baby with Tetralogy of Fallot



g- Full-Term baby with Bovine proteins allergy

Figure 2: Spectrograms of different pathologic cries

6. PREPROCESSING AND FEATURES CHARACTERISTIC EXTRACTION

We used a high-pass filter with a cut-off frequency of 150 Hz to reduce an important part of the ambient noise. We also used a pre-emphasis filter to eliminate the spectral combination from the lips and larynx.

MFCC and WTP coefficients are extracted using 20ms interlaced frame with 10ms recovering, providing 99 windows for one second cry samples and 299

windows for 3 seconds cry samples. Thereafter each frame is multiplied by a Hamming window to keep the continuity of the frame.

6.1. MFCC Extraction

We extracted 12 MFCC parameters for each 20ms of a cry signal. We obtained an MFCC matrix of 12 lines ×99 columns for each sample of 1second and a MFCC matrix of 12 lines ×299 columns for each sample of 3secondes. The two sets of characteristics once obtained are used as an input of the classifier.

6.2. MFDWC Extraction

The features characteristics are obtained from the wavelet packet decomposition of cry samples. We used Daubechies wavelet of order 8. The overall decomposition in wavelet packet is made for 5 levels. The general structure tree obtained is represented on Figure 3.

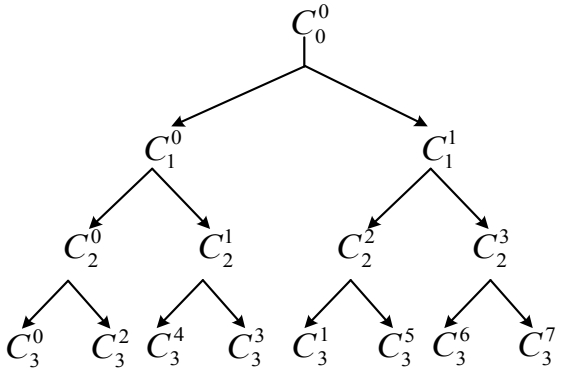


Figure 3: General structure of a wavelet packet decomposition on 3 levels.

Hence we obtain 63 WPT coefficients for each 20ms segment for all interlaced segments of a cry sample. We obtained a WPT matrix of 99 lines×63 columns for each sample of one second and a WPT matrix of 299 lines× 63 columns for each sample of 3secondes. Thereafter we calculate for each node the value of the entropy and energy. The matrix of entropy and energy obtained for all the cry samples are structured in ways to build two various admissible trees: a) Mel scale admissible tree by cry's samples and b) Mel scale admissible tree by cry's classes.

6.2.1 *Mel Scale Admissible Tree by Cry's Samples (TbCS)*

The first structure consists of converting each matrix of WPT (entropy) coefficients specific to a cry sample into only one vector of 6237=63×99 dimension for the one second samples and 18837=63×299 for the 3 second samples. This disposition of data is used as input of the classifier [6]. We also apply the PCA on this disposition of data before using them as an input of the classifier.

6.2.2 *Mel Scale Admissible Tree by Cry's Classes (TbCC)*

The second structure consists of concatenating all energy matrices obtained for the samples of the same class to obtain a tree by class of cries [2]. These data are used to find a better structure of tree adapted to each class according to algorithm BSAT [2].

7. APPLICATION OF PCA

To obtain the best base of decomposition, and to search out the best adapted approximation coefficients to represent the signal, we apply the PCA for the two structures of admissible trees found at the preceding step. In the TbCS case, we use the PCA to reduce the dimension of the training and test data separately by using the singular value decomposition (SVD) of the matrix of covariance.

In the TbCC case, we apply the PCA for the (training and test) data. For a given rate of variance, we obtain a specific structure of abstract tree to each class. Once the admissible tree is obtained, we apply the discrete cosine transform to keep 12 coefficients DCT (Discrete Cosine Transform).

8. PROBABILISTIC NEURAL NETWORK (PNN)

PNN is generally used for classification problems in the medical domain [10, 11]. The use of the PNN classifier is motivated by its speed and simplicity of the training process. There is no need for weight adaptation.
PNN is able to learn new incoming training data without having to repeat the whole training process. These characteristics are ideal for real time applications [10].

The PNN has three layers: the Input layer, Radial Basis Layer and the Competitive Layer. The second layer which is a radial network gives information on the resemblance between the input data and the data file used at the training phase. The third layer produced output vector of probability. The assigned class is the one with the highest probability [12]. Classifier PNN is obtained using function newpnn () in Matlab. The inputs of classifier PNN in our study are the vectors of the characteristics obtained at the preceding step, according to various cases presented previously.

9. EXPERIMENTAL RESULTS

To evaluate the efficiency of the studied system, four experiments were performed. The first experiment consists in using MFCC coefficients as vectors of features characteristic for the training and the test of the two classes of cries (pathologic and normal).

The second experiment consists in using WPT (entropy) coefficients as vectors of features characteristic for the training and the test of the two classes of cries.

The third experiment consists in applying the PCA to reduce the dimension of the WPT (entropy) coefficients of training and test data. The WPT (entropy) coefficients obtained are used as vectors of features characteristic for the training and the test of the two classes of cries.

The fourth experiment consists in applying BSAT algorithm to reduce the dimension of WPT (energy) coefficients of training and test for the two classes of cries.

The test of studied system was performed with cross-validation method. The results of the correct identification rate (Overall accuracy) for all experiments are shown in Table 3 for the 1second samples and Table 4 for the 3second samples. To evaluate the performance of this system, other measures such as specificity, sensitivity are also calculated from true positive (TP), true negative (TN), false positive (FP), and false negative (FN) as presented in Table 2 [7].

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP}+\text{FN})}$$

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN}+\text{FP})}$$

$$\text{Overall accuracy} = \frac{\text{TP}+\text{TN}}{(\text{TP}+\text{TN}+\text{FP}+\text{FN})}$$

Table 2: Confusion matrix.

Actual classification	Predicted classification	
	Pathological	Normal
Pathological	TP	FN
Normal	FP	TN

10. DISCUSSION

The results obtained concern mainly the use of PCA and BSAT algorithm to reduce the dimension of WPT coefficients. MFCC parameters are also used to compare the performance of the PCIS. The best performance with MFCCs is obtained for one second samples. The use of WPT (entropy and energy) coefficients as features characteristics is tested in several ways. By using WPT (entropy) coefficients directly without any reduction, the best performance is 86.36% for the 3second samples. The second test corresponds to application of PCA to reduce WPT (entropy) coefficients. This test gives a better identification rate for the two classes of cries for 3second samples. The BSAT Algorithm is also tested. The test of BSAT algorithm corresponds to its application for the training and test data for each class of cries. The correct identification rate found is 96.99% for the 1second samples.

Table 3: Results for various experiments (1second samples).

	Sensitivity	Specificity	Overall accuracy
MFCC	80.72	86.75	83.73
MFDWC	56.10	89.02	72.56
MFDWC + PCA	68.57	84.29	76.43
MFDWC + BSAT	93.97	100	96.98

Table 4: Results for various experiments (3second samples).

	Sensitivity	Specificity	Overall accuracy
MFCC	78.31	69.88	74.10
MFDWC	81.82	90.91	86.36
MFDWC + PCA	81.82	81.82	81.82
MFDWC + BSAT	100	83.91	91.95

11. CONCLUSION

In this experimental work, we study the performance of a PCIS using a PNN classifier and several methods of characterisation of newborn cries. The various methods tested are MFCC, MFDWC as well as the application of the PCA and BSAT algorithm to reduce the dimension of WPT (entropy and energy) coefficients.

The results from this study showed that the performance of PCIS depends on length of cry sample and on the technique of parameterization used. We also deduce from the experimental results that the use of PCA does not necessarily improve the performance of a PCIS. In fact, it sometimes performs better and sometimes worse according to the length of cry sample. Application of PCA-based BSAT algorithm generated better results.

Result obtained using MFDWC and PNN classifier is different compared with the result presented in the reference [7]. It is important to note that in this work, we are dealing with full-term and premature babies, both

with and without pathologies. This increases significantly the complexity of a PCIS. However in [7], cry signals studied (neonatal asphyxiate / healthy babies cry) contain significant differences in the acoustic signals [14] and that make the identification process easier.

The use of the wavelet coefficients and the PCA are very effective for the development of a pathological cry detection system. This work is still in progress. The results obtained by using WPT coefficients as features characteristics of newborn cries can be improved using other tree structures. We also expect to improve the results by using other classifier and other acoustic characteristics in addition to WPT coefficients and PCA.

In order to achieve a reliable evaluation of the classification performances we are still recording a larger database with a greater variety of pathologies and more subjects for each pathology.

ACKNOWLEDGEMENTS

This research work has been made possible through the funding provided by Bill & Melinda Gates Foundation under the Grand Challenges Explorations program.

REFERENCES

[1] BM lester, L. LaGasse , "crying", Elsevier 2008 pp 80-90.

[2] C.Tadj, A.Benlahouar, "Speaker Characterization Using Principal Component Analysis and Wavelet Transform for Speaker Verification". Euro speech, 2003.

[3] Linda L. LaGasse, A. Rebecca Neal, Barry M. Lester, "Assessment of infant cry: Acoustic cry analysis and Parental Perception", MRDD research reviews 2005, pp.83-93.

[4] Cohen A, Zmora E, "Automatic classification of infants' hunger and pain cry", In: Cappellini V, Constantinides AG (eds) Proceedings of the international conference on digital signal processing, Elsevier, Amsterdam, 1984.

[5] Orion Fausto, Reyes Galaviz, Carlos Alberto reyes Garcia, "Infant cry classification to identify hypo acoustics and asphyxia comparing an evolutionary-neural system with a network system", Springer-Verlag Berlin Heidelberg, pp. 949-958, 2005.

[6] Dror Lederman, Ehud Zmora, Stephane Hauschildt, Angelika Stellzig-Eisenhauer, Kathleen Wermke, "Classification of cries of infants with cleft-palate using parallel hidden markov models" Med Bio Eng Comput 2008, pp.965-975.

[7] M.Hariharan, Sazali Yaacob, Saidatul Ardeenaawatie Awang "Pathological infant cry analysis using wavelet packet transform and probabilistic neural network", Expert Systems with Applications 38 (2011) 15377–15382.

[8] D.E. Strurim, W.M.Campbell, D.A.Reynolds, "classification methods for speaker recognition", Springer-Verlag Berlin 2007, pp.278-297.

[9] J.Edward Jackson, "A user's guide to principal components", Wiley series in probability and statistics 2003.

[10] Walter P. Sweeney Jr., Mohamad T. Musavi, John N. "Guidi Classification of Chromosomes Using a Probabilistic Neural Network", 1994 Wiley-Liss, Inc.

[11] Mohd Fauzi Othman, Mohd Ariffanan Mohd Basri "Probabilistic Neural Network For Brain Tumor Classification", 2011 Second International Conference on Intelligent Systems, Modelling and Simulation.

[12] Neural Network Toolbox™ 6 User's Guide.

[13] Michelsson, K. "Cry analysis of symptomless low birth weight neonates and of asphyxiated newborn infants". Acta paediatrica Scandinavica, Vol. 216, pp. 1- 45, 1971.