Infant Cry Recognition System:

A Comparison of System Performance based on CDHMM and ANN

Yosra Abdulaziz Mohammed, University of Fallujah, Baghdad, Iraq

ABSTRACT

Cries of infants can be seen as an indicator of pain. It has been proven that crying caused by pain, hunger, fear, stress, etc., show different cry patterns. The work presented here introduces a comparative study between the performance of two different classification techniques implemented in an automatic classification system for identifying two types of infants' cries, pain, and non-pain. The techniques are namely, Continuous Hidden Markov Models (CHMM) and Artificial Neural Networks (ANN). Two different sets of acoustic features were extracted from the cry samples, those are MFCC and LPCC, the feature vectors generated by each were eventually fed into the classification module for the purpose of training and testing. The results of this work showed that the system based on CDHMM have better performance than that based on ANN. CDHMM gives the best identification rate at 96.1%, which is much higher than 79% of ANN whereby in general the system based on MFCC features performed better than the one that utilizes LPCC features.

KEYWORDS

INTRODUCTION

For infants, crying is a communication tool, a very limited one, but similar to the way an adult communicates. They use cries to express their physical, emotional and psychological states and needs (Drummond & McBride, 1993). An infant may cry for a variety of reasons, and many scientists believe that there are different types of cries which reflects different states and needs of infants, thus it is possible to analyze and classify infant cries for clinical diagnosis purposes.

Based on the information carried by the crying wave, the infant's physical state can be determined, and thus it can be detected if the infant is suffering a physical pain or just hunger or anger. Given that the processing of the information in the infant cry is basically a kind of pattern recognition, the task was approached by using the same techniques used for automatic speech recognition.

Hidden Markov Model is based on double stochastic processes, whereby the first process produces a set of observations which in turns can be used indirectly to reveal another hidden process that describes the states evolution (Rabiner,1989). This technique has been used extensively to analyze audio signals such as for biomedical signal processing (Lederman, Cohen, & Zmora, 2002) and speech recognition (Al-Alaoui, Al-Kanj, Azar, & Yaacoub, 2008). Neural Networks are defined as systems which have the capability to model highly complex nonlinear problems and composed of many simple processing elements, that operate in parallel and whose function is determined by the

network's structure, the strength of its connections, and the processing carried out by the processing elements or nodes.

In this work, a series of an observable feature vector is used to reveal the cry model hence assists in its classification. First, the paper describes the overall architecture of an automatic recognition system which main task is to differentiate between an infant 'pain' cries from 'non-pain' cries. The performance of both systems is compared in terms of recognition accuracy, classification error rate and F-measure under the use of two different acoustic features, namely Mel Frequency Cepstral Coefficient (MFCC) and Linear Prediction Cepstral Coefficients (LPCC). Separate phases of system training and system testing are carried out on two different sample sets of infant cries recorded from a group of babies which ranges from newborns up to 12 months old.

The prime objective of this paper is to compare the performance of an automatic infant's cry classification system applying two different classification techniques, Artificial Neural Networks and continuous Hidden Markov Model.

BACKGROUND

A number of research work related to this line have been reported, whereby many of which are based on Artificial Neural Network (ANN) classification techniques. (Petroni, Malowany, Johnston, & Stevens,1995) for example, have used three different varieties of supervised ANN technique which include a simple feed-forward, a recurrent neural network (RNN) and a time-delay neural network (TDNN) in their infant cry classification system. In their study, they have attempted to recognize and classify three categories of cry, namely 'pain', 'fear' and 'hunger' and the results demonstrated that the highest classification rate was achieved by using feed-forward neural network. Another research work carried out by (Cano & Escobedo, 1999) used the Kohonen's self-organizing maps (SOM) which is basically a variety of unsupervised ANN technique to classify different infant cries. (Rosales-Pérez, Reyes-Garcia, Gonzalez, & Arch-Tirado, 2012) used Genetic Selection of a Fuzzy Model (GSFM) for classification of infant cry where GSFM selects a combination of feature selection methods, type of fuzzy processing, learning algorithm, and its associated parameters that best fit to the data and have obtained up to 99.42% in recognition accuracy. (Al-Azzawi, 2014) designed an automatic infant cry recognition system based on the fuzzy transform (F-transform) that classifies two different kinds of cries, which come from physiological status and medical disease, a supervised MLP scaled conjugate ANN was used and the classification accuracy obtained was 96%.

Apart from the traditional ANN approach, another infant cry classification technique studied is Support Vector Machine (SVM) which has been reported by (Barajas & Reyes, 2005). Here, a set of Mel Frequency Cepstral Coefficients (MFCC) was extracted from the audio samples as the input features. On the other hand, (Orozco & Reyes, 2003a), use the linear prediction technique to extract the acoustic features from the cry samples of which are then fed into a feed-forward neural network recognition module. Recently (CY. Chang, CW. Chang, Kathiravan, & Chen, 2017) have proposed an infant cry classification system to categorize the types of infant crying into hunger; pain; and feeling sleepy. Fifteen features were extracted from each crying frame and the sequential forward floating selection was adopted to pick out high discriminative features. The directed acyclic graph support vector machine was used to classify infant crying. The proposed system showed a classification accuracy of 92.17%.

(Orlandi, Reyes-Garcia, Bandini, Donzelli, & Manfredi, 2016) studied the differences between full-term and preterm infant cry comparing four classifiers: Logistic Curve, Multilayer Perceptron, Support Vector Machine, and Random Forest. They managed to assess differences between preterm and full-term newborns with about 87% of accuracy. Best results were obtained with the Random Forest method (receiver operating characteristic area, 0.94).

Related Content

Knowledge Sharing and Pervasive Computing: The Need for Trust and a Sense of History
Phillip W.J. Brook (2009). *Risk Assessment and Management in Pervasive Computing: Operational, Legal, Ethical, and Financial Perspectives (pp. 285-297).*
www.irma-international.org/chapter/knowledge-sharing-pervasive-computing/28461

An Agent-Based Approach to Adapt Multimedia Web Content in Ubiquitous Environment
Domenico Rosaciand Giuseppe M.L. Sarnè (2012). *Ubiquitous Multimedia and Mobile Agents: Models and Implementations (pp. 60-84).*
www.irma-international.org/chapter/agent-based-approach-adapt-multimedia/56420

Predicting Search Intent Based on In-Search Context for Exploratory Search
Vikram Singh (2019). *International Journal of Advanced Pervasive and Ubiquitous Computing (pp. 53-75).*
www.irma-international.org/article/predicting-search-intent-based-on-in-search-context-for-exploratory-search/233559

A Self-Learning Based Antenna System for Indoor Wireless Network
Wei Ni (2017). *International Journal of Advanced Pervasive and Ubiquitous Computing (pp. 78-87).*
www.irma-international.org/article/a-self-learning-based-antenna-system-for-indoor-wireless-network/189228

Robust Object Detection in Military Infrared Image
Tao Gao (2013). *International Journal of Advanced Pervasive and Ubiquitous Computing (pp. 30-36).*
www.irma-international.org/article/robust-object-detection-in-military-infrared-image/93000