

Hybrid System of Optimal Self Organizing Maps and Hidden Markov Model for Arabic Digits Recognition

¹ZAKARIAE EN-NAIMANI, ²MOHAMED LAZAAR, ³MOHAMED ETTAOUIL

^{1,3}Modelling and Scientific Computing Laboratory,
Faculty of Science and Technology,

University Sidi Mohammed ben Abdellah, Fez, MOROCCO

²National School of Applied Sciences (ENSA)

Abdelmalek Essaadi University, Tetouan, MOROCCO

¹z.ennaimani@gmail.com

²lazaarmd@gmail.com,

³mohamedettaouil@yahoo.fr

Abstract: - Thanks to Automatic Speech Recognition (ASR), a lot of machines can nowadays emulate human being ability to understand and speak natural language. However, ASR problematic could be as interesting as it is difficult. Its difficulty is precisely due to the complexity of speech processing, which takes into consideration many aspects: acoustic, phonetic, syntactic, etc. Thus, the most commonly used technology, in the context of speech recognition, is based on statistical models. Especially, the Hidden Markov Models which are capable of simultaneously modeling frequency and temporal characteristics of the speech signal. There is also the alternative of using Neuronal Networks. But another interesting framework applied in ASR is indeed the hybrid Artificial Neural Network (ANN) and Hidden Markov Model (HMM) speech recognizer that improves the accuracy of the two models. In the present work, we propose an Arabic digits recognition system based on hybrid Optimal Artificial Neural Network and Hidden Markov Model (OANN/HMM). The main innovation in this work is the use of an optimal neural network to determine the optimal groups, unlike in classical Kohonen approach. The numerical results are powerful and show the practical interest of our approach.

Key-Words: - Automatic Speech Recognition, Hidden Markov Models, Self Organizing Maps, Vector Quantization.

1 Introduction

Speech recognition is very useful in many applications, in our daily life, including command and control, dictation, transcription of recorded speech, searching audio documents and interactive spoken dialogues [3]. It has interested many scientists since the early 1950's. In fact because speech is the most essential medium of human being communication, he always tries to replicate this ability on machines. However, the processing of speech signal has to consider its different aspects, which complicates the task. Indeed, the signal information could be analyzed from different sides: acoustic, phonetic, phonologic, morphologic, syntactic, semantic and pragmatic [8] [29].

Artificial Neural Network (ANN) often called as Neural Network (NN), is a computational model or mathematical model based on biological neural networks.

The Artificial Neural Networks (ANN) are a very powerful tool to deal with many applications, and they have proved their effectiveness in several

research areas such as analysis and image compression, recognition of writing, speech recognition, speech compression, video compression, signal analysis, process control, robotics, and research on the Web [25][20][28].

The Hidden Markov Model has become one of the most powerful statistical methods for modeling speech signals [15] [26] [27]. Its principles have been successfully used in automatic speech recognition, formant and pitch tracking, speech enhancement, speech synthesis, statistical language modeling, part-of-speech tagging, spoken language understanding, and machine translation. The Hidden Markov Model allows the temporal variation, with local spectral variability modeled using flexible distributions such as mixtures of Gaussian densities [12]. New techniques have emerged to increase the system performance speech recognition. These are based on the hybridization of Artificial Neural Network and Hidden Markov Model (ANN/HMM). The approaches ANN/HMM represent a competitive alternative to standard systems [4] [11] [16].

