Chapter XII

Advanced Techniques in Speech Recognition

Jose Luis Oropeza-Rodriguez, National Polytechnic Institute, Mexico

Sergio Suárez-Gerra, National Polytechnic Institute, Mexico

Abstract

During the last 30 years, people have tried to communicate in an oral form with the computers, developing for this end an important amount of automatic speech recognition algorithms. Because of this, software such as the Dragon Dictate and the IBM Via Voice are already available to interact easily with the computer in oral form. However, during the last several years ASR has not reported important advances, not only due to the advances obtained until now, but also because the scientific community working in this area does not have founded another tool so powerful as HMM, despite a great number of alternatives that have been proposed since HMM appeared. This chapter presents the main elements required to create a practical ASR using HMM. The basic principles of the continuous density hidden Markov models (CDHMM) are also given.
Introduction

From readings in speech recognition such as those provided in the previous chapter or those provided by Waibel and Lee (1990), it follows that automatic speech recognition (ASR) is based mainly by the following approaches: template-based, knowledge-based, stochastic, and connectionist, principally. During a long time, 20 years ago, the interest was based on the data time warping (DTW) algorithm; it was used for a long time because it represented a good alignment between the speech signal and the time.

The template-based approach was mentioned in the previous chapter, and its most representative algorithm is the linear coding predictive (LPC). The LPC was widely used to solve problems of the verbal commands with a set of limited instructions. After that, in the 1980s, the stochastic approach using hidden Markov models was used to efficiently solve the alignment mentioned previously and the relationship with the significance of the word or message that corpus of voices needed. Actually, research centers are interested in realizing studies about the interaction of neural networks (connectionist approach) with ASR.

This chapter is focused on hidden Markov models because their applications actually are vast; HTK is the most important result of that. The discussion presented in this chapter pretends to give to the reader an idea of automatic speech recognition using HMMs, in both continuous and discontinuous speech recognition.

Finally, we present a new approach used in the training stage; the inclusion of the knowledge-based system has the goal of adding a priori knowledge to this stage, so that not only the segmentation of the corpus is important, but also the knowledge of the grammatical structure. For that, we use the 10 rules identified into the Spanish language that permit splitting a word into the syllables that it constitutes.

The Units and Models into Automatic Speech Recognition

Automatic speech recognition systems (ASRs) are implemented using the phoneme-like fundamental speech unit, being the more applications based on the hidden Markov models

Figure 1. Concatenation in hidden Markov models

Copyright © 2007, Idea Group Inc. Copying or distributing in print or electronic forms without written permission of Idea Group Inc. is prohibited.
Related Content

Interactive Question Answering
www.igi-global.com/chapter/interactive-question-answering/70067?camid=4v1a

Cocktail Party Problem: Source Separation Issues and Computational Methods
www.igi-global.com/chapter/cocktail-party-problem/45481?camid=4v1a

Data Hiding for Text and Binary Files
www.igi-global.com/chapter/data-hiding-for-text-and-binary-files/108790?camid=4v1a

Information Extraction from Text and Beyond
www.igi-global.com/chapter/information-extraction-text-beyond/64578?camid=4v1a