

Universidad de las Américas

Ingeniería de Sistemas

Tratamiento y reconocimiento de señales de voz. Implementación informática y aplicación.

Trabajo de Titulación presentado en conformidad a los requisitos
Para obtener el título de Ingeniero de Sistemas

Profesor Guía: Dr. Danilo Gortaire J.

Autores:

María Isabel Molineros

Angel Quimbiulco

2004

1. INTRODUCCIÓN

El habla, medio de comunicación por excelencia entre los seres humanos, está siendo velozmente implantado en la interacción cotidiana entre el hombre y la máquina. Las encuestas de usuarios realizadas por la revista electrónica TidBits a finales del año 2000 situaban el reconocimiento del habla como la "manzana más deseada" de la tecnología informática del nuevo siglo. Además, la domótica ha iniciado la implementación de las 'casas inteligentes' en las que habitaremos y con las que hablaremos en el futuro. Se acerca el momento de realizar lo que numerosas páginas de ciencia-ficción ya habían plasmado, el gran sueño del hombre de comunicarse con sus máquinas de la forma más natural: conversando. [Tec]

La posibilidad de "hablar con el ordenador", imitando las habilidades naturales del ser humano se ha convertido en uno de los retos más complejos con los que se enfrentan científicos e ingenieros, los cuales deben resolver problemas como la dispersión de conocimientos involucrados en cada área, de manera que están obligados a descubrir, adquirir y asimilar diversos contenidos como redes neuronales, fonética acústica y matemática, además para solucionar el problema de la Comprensión de la Lengua Hablada es necesaria la combinación de distintas fuentes de conocimiento lingüístico (fonológico¹, morfológico², léxico³, sintáctico⁴, semántico⁵) y de comportamiento humano.

Las ventajas que se esperan de la comunicación hombre-máquina por medio del habla son múltiples, debido a que este modo de relación libera completamente a la persona del uso de la vista y de las manos, es decir de la pantalla y el teclado, dejando a la persona en libertad de movimientos.

¹ Relativo a la Fonología. Rama que estudia los sonidos dentro de una lengua estableciendo las normas para su ordenamiento

² Relativo a la morfología. Parte de la gramática que trata de la forma de las palabras.

³ Conjunto de palabras o locuciones de una lengua.

⁴ Relativo a la sintaxis. Parte de la gramática que estudia las oraciones y sus clases.

⁵ Referente al significado de las palabras.

La voz⁶ puede ser transportada por medios simples y muy extendidos como la red telefónica. Estas y otras ventajas se traducen en una gran variedad de aplicaciones actuales y de futuro. Por ejemplo en el campo industrial, es muy útil poder controlar una máquina-herramienta, un robot, o proporcionar datos a un sistema de diseño asistido por computadora e, incluso, poder iluminar las distintas dependencias de una casa conservando la independencia de movimientos de las personas que la habitan, mediante la interacción hablada a través de la red telefónica. Esta también permite simplificar el acceso a sistemas automáticos de reservaciones, de aceptación de pedidos o acceso a bancos de información.

Los científicos actualmente encaminan sus esfuerzos hacia la obtención de un mayor conocimiento sobre diversos paradigmas de Reconocimiento Automático del Habla (RAH) como son los Modelos Ocultos de Markov (HMM), las Redes Neuronales (NN), o la combinación de Modelos Ocultos de Markov y Redes Neuronales (HMM + NN), que permitan reconocer la voz con un mayor nivel acústico, es decir, decodificar la señal acústica de entrada producida por el hablante.

⁶ En el transcurso del trabajo emplearemos indistintamente las palabras voz y habla.