

Reconocimiento de voz mediante Modelos Ocultos de Markov

Adriana Becerra, Marcela Gómez, María Fernanda Ordóñez, Byron Macas, Ing. Janeth Chio

August 3, 2009

Abstract

El presente proyecto tiene como finalidad mostrar la aplicabilidad de los Modelos Ocultos de Markov en el reconocimiento de voz, se hace una breve descripción de los conceptos básicos necesarios para realizar, los algoritmos que se utilizan y una descripción de herramientas HTK para el diseño y manipulación de HMMs

HMMs, Reconocimiento de voz, HTK, Viterbi

1 Introduction

El reconocimiento de voz en la actualidad es uno de los temas de investigación más relevante al cuál se le pueden aplicar una serie de técnicas y modelos; en éste paper especificamos la aplicabilidad de modelos ocultos de Markov para su desarrollo. Estos modelos son una extensión de las Cadenas de Markov (Hidden Markov Model) asumen que el sistema estudiado sigue un proceso de Markov con parámetros desconocidos.

Los principales algoritmos que intervienen dentro de éste modelo son: El algoritmo de Viterbi, que es un algoritmo que permite encontrar la secuencia de estados más probable en un HMM a partir de la observación ; éste algoritmo parte de un estado inicial y, teniendo en cuenta la probabilidad de transición entre estados, la probabilidad de emisión de estos estados y las probabilidades que gobiernan la concatenación de modelos representativos de palabras, obtiene de manera recurrente la secuencia de estados más probable. Los modelos que subyacen a esta secuencia de estados más probable son los que determinan la transcripción de la secuencia que estamos reconociendo. A más de éste algoritmo citaremos de una manera general otros como los algoritmos forward y backward y el algoritmo de Baum Welch.

El objetivo principal de éste paper consiste en buscar soluciones originales y eficaces al problema de reconocimiento de habla en presencia de ruidos aditivos

*abecerra@utpl.edu.ec, mtgomez@utpl.edu.ec, mfordonez@utpl.edu.ec, bg-macas@utpl.edu.ec

y de las distorsiones propias de los sistemas de comunicaciones inalámbricos actuales, para lo cual realizamos una simulación de reconocimiento de voz utilizando ésta técnica mediante el software de simulación Matlab.

2 Modelos Ocultos de Markov

Los modelos ocultos de Markov (HMM), es un proceso estocástico que consta de un proceso de Markov no observado (oculto) y un proceso observado O cuyos estados son dependientes estocásticamente de los estados ocultos. [VIR] La tarea fundamental consiste en determinar los parámetros ocultos a partir de los parámetros observados. La diferencia fundamental respecto a un modelo de Markov habitual consiste en que los estados no son directamente visibles para el observador, pero sí lo son las variables influenciadas por el estado. Cada estado tiene una distribución de probabilidad asociada sobre el conjunto de posibles valores de salida. La secuencia de valores de salida generados a partir de un HMM nos dará cierta información sobre la secuencia de estados. [EM] Una de las aplicaciones más utilizadas de estos modelos, es el reconocimiento del habla, esta técnica ha permitido modelar adecuadamente la gran variabilidad en el tiempo de la señal de voz [RE].

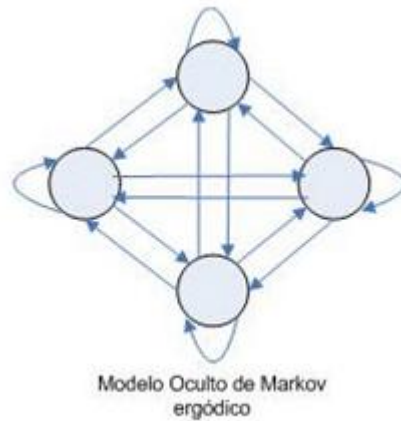
2.1 Arquitectura de los HMM

Esta arquitectura se da de acuerdo al número de estados que lo componen y las transiciones permitidas entre dichos estados. Los principales modelos existentes son : Modelos de izquierda a derecha y Modelos ergódicos.

- Modelos de izquierda a derecha



- Modelos Ergódicos



2.2 Tipos de Modelos Ocultos de Markov

- HMM Discretos
- HMM Continuos
- HMM Semicontinuos

2.3 Algoritmos HMM

- Algoritmo Viterbi

El algoritmo de Viterbi nos permite encontrar la secuencia de estados más probable en un Modelo Oculto de Markov (MOM) a partir de una observación, éste obtiene la secuencia óptima que mejor explica la secuencia de observaciones.

Los pasos a seguir son los siguientes:

1. Se elige un MOM en el cual no existe información apriori del problema.
2. Se crean todas las transiciones entre estados con igual probabilidad.
3. Se utiliza el algoritmo para alinear las secuencias de entrenamiento al modelo actual.
4. Luego se toma las estadísticas correspondientes a los valores que se quiere estimar.
5. Número de veces que el estado i -ésimo aparece en la secuencia.
6. Número de transiciones de estado i -ésimo a un estado j -ésimo durante el alineamiento de la secuencia.
7. Cálculo de las frecuencias relativas, que luego se utilizan en la estimación del modelo

8. Se repiten los pasos de alineamiento, el cálculo de estimadores y la re-estimación de probabilidades

- Algoritmo Baum Welch

Algoritmo Baum-Welch Toda la información necesitada para realizar la re-estimación de parámetros utilizando el algoritmo Baum-Welch está en marcha. Los pasos en éste algoritmo pueden ser resumidos así:

1. Para cada parámetro del vector/matriz que requiere re-estimación , se asigna almacenamiento para el numerador y denominador, de la forma :Para almacenar direcciones son referenciados como acumuladores
2. Calcula las probabilidades hacia atrás y hacia adelante para los estados j y tiempos t
3. Para cada estado j y tiempo t , usa la probabilidad $L_j(t)$ y el vector O_t de observaciones actuales para actualizar los acumuladores para ése estado
4. Utilice el acumulador final de los valores para calcular nuevos valores de los parámetros
5. Si el valor de $P = P(O|M)$ para ésta iteración no es superior a la del valor de la iteración anterior y luego se detiene, de lo contrario repita los pasos anteriores utilizando una nueva estimación de parámetros Toda ésta descripción descrita anteriormente asume que los parámetros para un HMM se vuelven a estimar a partir de una única secuencia de observación, que es un solo ejemplo de la palabra hablada. En practica, son muchos los ejemplos necesitados para hacer una buena estimación de parámetros, sin embargo, el uso de múltiples secuencias de observación añade no una complejidad adicional al algoritmo. Los pasos 2 y 3 son simplemente repetidos para cada secuencia de de entrenamiento.

3 Metodología

3.1 DESCRIPCIÓN GENERAL DEL SISTEMA DE RE-CONOCIMIENTO

Debido a la inercia inherente a los órganos articulatorios, es posible suponer que las características de la señal no varían apreciablemente en un intervalo suficientemente corto de tiempo (del orden de 20 ms) y, por tanto, es posible realizar un análisis espectral cuasi-estacionario sobre segmentos de señal de esta duración temporal. La evolución temporal de las características espectrales se obtiene repitiendo el análisis sobre segmentos consecutivos de la señal, que suelen tomarse con un cierto solapamiento temporal. De esta forma, a partir de una señal de voz se obtiene una secuencia de espectros, que suele denominarse espectrograma. En los sistemas de reconocimiento del habla mediante técnicas de comparación de patrones se aborda el proceso de reconocimiento sin realizar un

modelado de la evolución temporal de esta secuencia de espectros. Los patrones de referencia y de test consisten simplemente en secuencias de espectros y el proceso de comparación se limita a calcular la distancia acumulada entre dichos patrones a lo largo del camino óptimo dado por el algoritmo de programación dinámica . Salvo en el caso de aplicar técnicas de agrupamiento para obtener los patrones de referencia a partir de varias pronunciaciones, la variabilidad de la señal de voz sólo es tomada en cuenta en el alineamiento temporal no lineal de los patrones que realiza el citado algoritmo de programación dinámica.

En los sistemas de reconocimiento del habla basados en los modelos ocultos de Markov, se modela la evolución temporal de la secuencia de espectros obtenida de la señal de voz mediante un HMM con el fin de contemplar estocásticamente las diversas fuentes de variabilidad de la señal. Este modelado consiste en la asociación de los estados del HMM a los diferentes tramos de la señal, de forma que las probabilidades de generación de observaciones modelan la variabilidad estadística de las características espectrales de cada tramo, mientras que las probabilidades de transición modelan su secuenciamiento y duración.[PER]

En el siguiente gráfico se representa el proceso de reconocimiento mediante éstos modelos

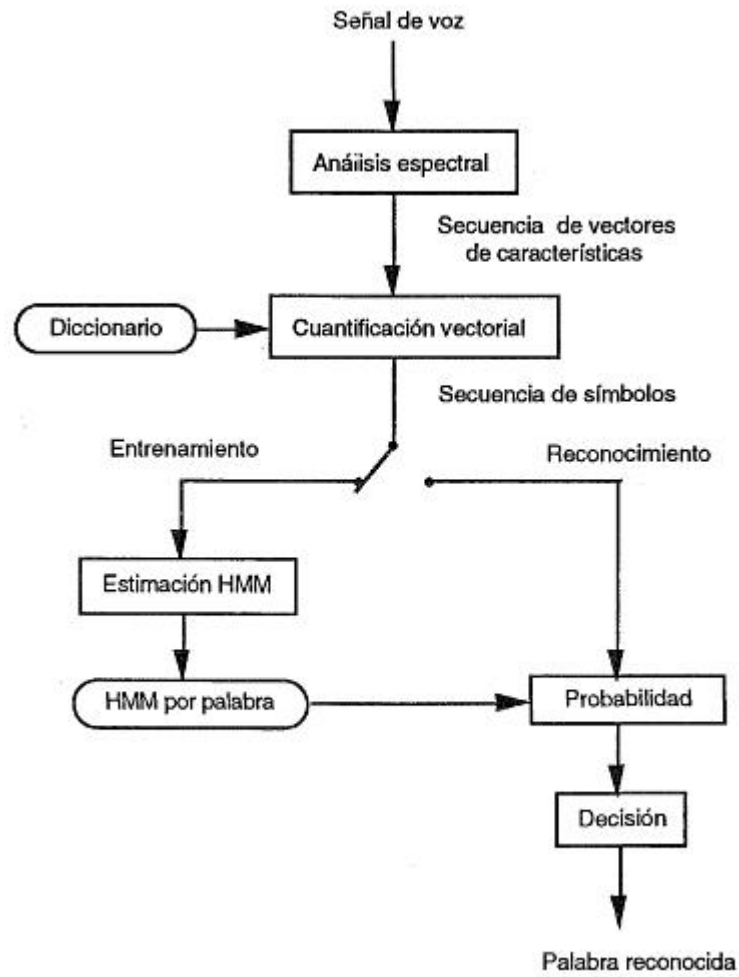


Figura1. Proceso de reconocimiento de palabras [PER]

4 Results

Para realizar la simulación de reconocimiento de voz utilizamos la herramienta Matlab ejecutando la función LPC que es una técnica de predicción lineal. Se basan en la posibilidad de modelar el tracto vocal como una serie de cilindros huecos de diámetro variable. Las propiedades de la onda sonora, después de haber pasado por todos ellos, pueden predecirse teniendo en cuenta que cada uno de los cilindros condiciona la forma de la onda sonora al entrar en el siguiente. Este cálculo complejo se simplifica mediante la utilización de la predicción lineal, y aprovechando la periodicidad de la onda sonora. [GONZ].

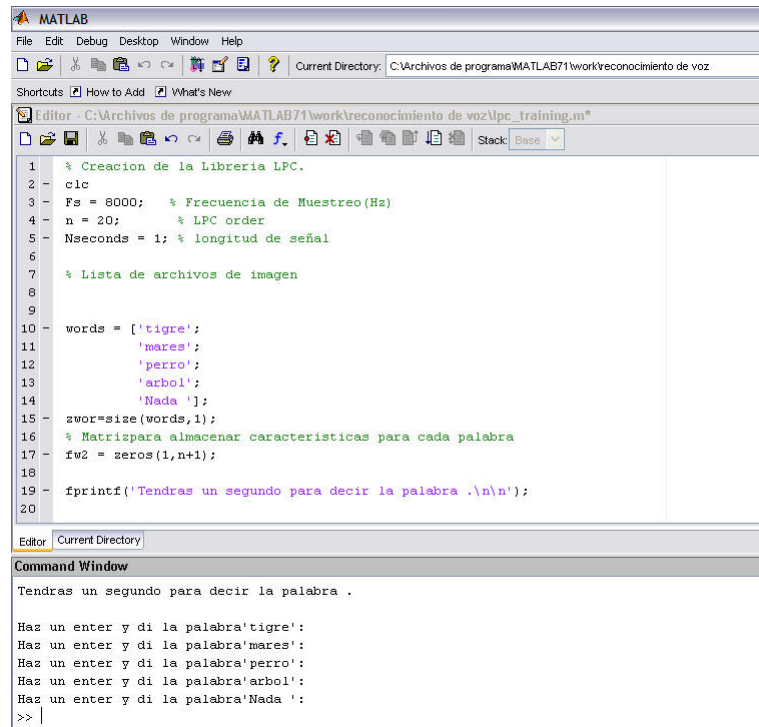


Figura2. Captura del conjunto de entrenamiento de la aplicación

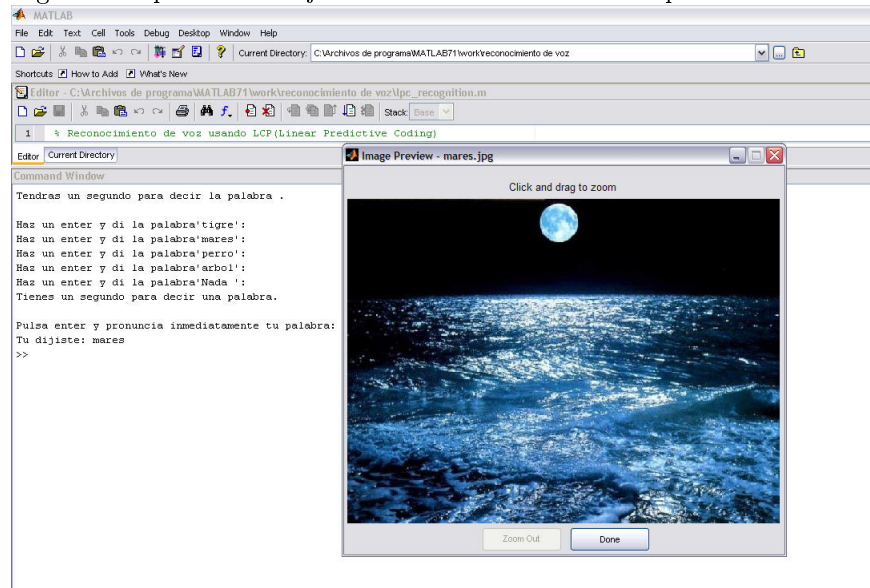


Figura 3. Captura del Funcionamiento de la aplicación

5 Conclusiones

- Mediante el presente paper comprendemos mejor el funcionamiento de los Modelos Ocultos de Markov y su aplicabilidad en la vida real.
- Existen varias arquitecturas de los Modelos Ocultos de Markov a aplicar dependiendo del ambiente en el que se vaya a trabajar y la aplicación que se quiera realizar.
- Existen varios algoritmos que se pueden aplicar , todo dependiendo de los resultados que se quiera obtener , por ejemplo el ahorro de tiempo, número de repeticiones , etc.
- Se deben analizar el conjunto de observaciones iniciales antes de aplicar o seleccionar un HMM.
- En lo referente a la aplicación ,debido a la sensibilidad de los dispositivos de entrada se debe evitar en lo posible el ruido que se produce en el entorno para evitar malas interpretaciones del simulador.
- El proceso de reconocimiento de voz es una tarea complicada puesto que requiere de varias consideraciones y procesos inteligentes para lograr su perfección.
- En los sistemas de reconocimiento del habla basados en los modelos ocultos de Markov, se modela la evolución temporal de la secuencia de espectros obtenida de la señal de voz mediante un HMM con el fin de contemplar estocásticamente las diversas fuentes de variabilidad de la seña

6 Referencias

[AM] Alvares Mauricio , ANÁLISIS DE PROXIMIDAD DE MODELOS OCULTOS DE Markov PARA LA IDENTIFICACIÓN DE FUENTES DE ESPIGAS

[AP] Aprendizaje del modelo. Disponible en: <http://www.virtual.unal.edu.co/cursos/ingenieria/2001832/le>

[APL] http://es.wikipedia.org/wiki/Algoritmo_de_Viterbi#Aplicaci.C3.B3n_del_algoritmo_de_Viterbi

[BA] Algoritmo de Baum-Welch. Disponible en: http://es.wikipedia.org/wiki/Algoritmo_de_Baum-Welch

[BE] Bergasa Luis Miguel. Introducción a los Modelos Ocultos de Markov.

Disponible en: http://www.depeca.uah.es/docencia/doctorado/cursos04_05/82854/docus/HMM.pdf

[CY] Chen Yang, Suling .- Recognition of Hand-drawn Sketches with Markov

Random Fields

[DE] DEIBY ALEXANDER FANDIÑO R.- RECONOCIMIENTO AUTOMÁTICO DE VOZ (ESTADO DEL ARTE)

[DI] Carrilo Aguilar Roberto. Diseño y manipulación de modelos ocultos de Markov, utilizando herramientas HTK. Una tutoría. Disponible en: <http://dialnet.unirioja.es/servlet/articulo?codigo=2317870>

[EM] Modelos ocultos de Markov. Disponible en: <http://supervisadaextraccionrecuperacioninformacion.iesp>

- [FO] ESCOLANO , O Colomina .- Aprendizaje: Modelos Ocultos de Markov (HMMs) [GAV] <http://www.gavab.etsii.urjc.es/recursos.html#hmm>
- [GJ] GUEVARA, Jorge Luis .- Modelos Ocultos de Markov [MIT] Reconocimiento automático del habla.- Disponible en: <http://mit.ocw.universia.net/6.345/NR/rdonlyres/Electrical-Engineering-and-Computer-Science/6-345Automatic-Speech-RecognitionSpring2003/4C063D4A-3B8B-4F05-B475-85AAD0D66586/0/lecture10.pdf>
- [JF] VELEZ J.F, MORENO A.B.- HMMSim: An Interactive Simulation Tool for Teaching Discrete Hidden Markov Models. Proc. Congreso Iberoamericano de Informática Educativa 2002 (IE - 2002), Vigo (Spain), 2002. ISBN 848158-227-1.
- [MO] Modelos ocultos de Markov. Disponible en: <http://advancedtech.wordpress.com/2008/08/06/modelos-ocultos-de-Markov/>
- [PR] Problemas básicos de los HMMs. Disponible en: <http://www.virtual.unal.edu.co/cursos/ingenieria/200183/>
- [RE] Modelos ocultos de Markov para el reconocimiento automático del habla. Una breve introducción.
- [RW] ROSELLI T, WERNER. Algoritmos de Markov.
- [VALDIVIEZO, 2008] VALDIVIEZO, Daniel. Modelos Ocultos de Markov – Arquitectura. 2008. Disponible en: <http://danielvn7.wordpress.com/2008/07/03/modelos-ocultos-de-markovarquitectura/>
- [VIR] Modelos ocultos de Markov. Disponible en: <http://www.virtual.unal.edu.co/cursos/ingenieria/200183/>
- [WI] Reconocimiento del habla. Disponible en: http://es.wikipedia.org/wiki/Reconocimiento_del_habla
- [GONZ] E. GONZÁLEZ y J. CALERO: Aplicaciones de la Tecnología del Habla. Comunicaciones de Telefónica I+D, vol. 5, núm. 2 (1994).
- [PER] Francisco Javier Hernando Pericas.- TECNICAS DE PROCESADO Y REPRESENTACION DE LA SEÑAL DE VOZ PARA EL RECONOCIMIENTO DEL HABLA EN AMBIENTES RUIDOSOS