

UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO  
FACULTAD DE INGENIERIA ELÉCTRICA ELECTRÓNICA INFORMÁTICA Y  
MECÁNICA  
ESCUELA PROFESIONAL DE INGENIERIA INFORMÁTICA Y DE SISTEMAS



TESIS

---

**“IDENTIFICACIÓN DE LOCUTOR USANDO CODEBOOKS  
DE COEFICIENTES CEPSTRALES EN LAS FRECUENCIAS  
DE MEL Y MODELOS OCULTOS DE MARKOV”**

---

Presentado por:

**Br. JHON DENNIS AUCCAPUMA GAMARRA**

**Br. ERROL WILDERD MAMANI CONDORI**

Para Optar el título profesional de:

**INGENIERO INFORMÁTICO Y DE SISTEMAS**

Asesor:

**Mgt. RONY VILLAFUERTE SERNA**

Co-Asesor:

**Mcs. YESHICA ISELA ORMEÑO AYALA**

CUSCO-PERÚ

2016

# Resumen

El habla es un tipo de señal complicada producto del resultado de una serie de transformaciones ocurridas en diferentes niveles: semántica, lingüística y acústica. Estas transformaciones conducen a diferencias en las características de un individuo ampliamente estudiadas por la Biometría. La identificación de locutor (identificar quién es la persona que emitió la voz) en síntesis, es un análisis detallado de las características del habla de cada individuo basado en puntuaciones. En este contexto este proyecto se centró en la identificación de locutor mediante Coeficientes Cepstrales en las Frecuencias Mel o Mel Frequency Cepstral Coefficients (MFCC) y Modelos Ocultos de Markov o Hidden Markov Model (HMM). Iniciando el procesamiento de voz, para obtener las características más importantes de un individuo se utilizó Coeficientes Cepstrales en las Frecuencias Mel, debido a que en la actualidad otorgan los mejores resultados en el análisis Cepstral según el estado de arte, posteriormente; se hizo uso de la cuantificación vectorial o Vector quantization (VQ) que por medio del algoritmo de clasificación K-means, divide el conjunto de vectores característicos en un número determinado de vectores representativos, los cuales mejoran sustancialmente el tiempo de procesamiento. Para el modelamiento de los vectores representativos se hizo uso de los Modelos Ocultos de Markov; los HMMs son entrenados para generar el modelo oculto del locutor el cual estará formado por la secuencia de observaciones (símbolos de observación) y la secuencia de estados, para luego encontrar la secuencia de estados con mayor probabilidad, la identidad de un locutor se determina mediante el modelo que obtenga la máxima probabilidad (puntuación) determinado por el algoritmo de Viterbi. Adicionalmente se estimaron los parámetros de los módulos de pre procesamiento, extracción de características, pos procesamiento y el cuantificador vectorial basado en codebooks; para sugerir el tamaño de codebook más adecuado y los parámetros con los que se obtenga buenos resultados en la identificación de un locutor; para cada módulo se describió la teoría y la implementación del código fuente en Java. Finalmente, nuestros resultados experimentales muestran los parámetros con los que se obtienen buenos resultados teniendo un 90% de aceptación para un grupo reducido y cerrado de 5 personas en condiciones reales (con ruido de fondo), con una tendencia de decrecimiento a medida que aumenta el número de población y una mayor efectividad en condiciones ideales (ambiente cerrado y sin ruido de fondo).

**Palabras Clave:** Biometría, *MFCC*, *HMM*, *VQ*, *K-means*.