

Análisis Comparativo de Distintas Toolkits para el Reconocimiento Biométrico de Personas Mediante Voz

Silvia Ruiz, Ernesto Miranda, Mauro Herlein, Graciela Etchart, Carlos Alvez.

Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos

Av. Tavella 1424, Concordia, Entre Ríos - CP 3200

sruiz@fcad.uner.edu.ar, emiranda@fcad.uner.edu.ar, herlein.mauro@gmail.com, getchart@fcad.uner.edu.ar,
caralv@fcad.uner.edu.ar,

Resumen

El objetivo de este trabajo es realizar un análisis comparativo de distintas toolkits para el reconocimiento biométrico de personas mediante voz. Hoy en día los sistemas de identificación de personas se han convertido en una necesidad para la sociedad. A medida que avanza la tecnología y la aplicación de la misma en entornos tanto de ocio como de seguridad, la evolución en desarrollo biométrico es muy grande. Los sistemas de identificación o verificación tradicionales (tarjetas o claves) se han ido sustituyendo por sistemas automáticos de reconocimiento biométrico. La biometría de voz es muy utilizada en aplicaciones o sistemas relacionados con la seguridad, ya que cada individuo tiene características físicas diferentes. En este sentido, el presente trabajo propone realizar un estudio comparativo de distintas toolkits para el reconocimiento biométrico de personas mediante voz en base a criterios de evaluación previamente definidos que permitirán determinar la/s herramientas consideradas más eficaces.

Palabras clave: biometría, rasgos biométricos, voz, seguridad.

Contexto

Este trabajo surge en el marco del proyecto de investigación *PID 07/G044 "Gestión de Datos Biométricos en Bases de Datos Objeto-Relacionales"*, que da continuidad al Proyecto *PID 07/G035 "Identificación de personas mediante Sistemas Biométricos. Estudio de factibilidad y su implementación en organismos estatales"* [1][2][3][4][5].

El mismo constituye un futuro trabajo de tesis de la Maestría en Sistemas de Información dictada en la Facultad de Ciencias de la Administración perteneciente a la Universidad Nacional de Entre Ríos. El objetivo general es realizar un análisis comparativo de distintas toolkits para el reconocimiento biométrico de personas mediante voz, evaluando y seleccionando la/s herramienta/s consideradas más adecuadas para la implementación de un sistema de reconocimiento por voz.

Introducción

La importancia del uso de las tecnologías basadas en biometrías, se encuentra en su aplicación como mecanismo de control a través de una serie de medidas de características específicas que permite el reconocimiento de personas, para superar los problemas de vulnerabilidad e inconvenientes (gastos en su creación, control, administración, posibles extravíos, olvidos, etc.) del sistema de autenticación por contraseñas, tarjetas de acceso, entre otros.

Los sistemas de reconocimiento biométrico usan características fisiológicas o de comportamiento propias de cada individuo para identificarlo, es decir, se reconoce al usuario por lo que es en lugar de por lo que tiene o sabe [6]. Los rasgos fisiológicos presentan una reducida variabilidad a lo largo del tiempo, pese a que su adquisición es más invasiva y requiere de la cooperación de los sujetos. Por el contrario, los rasgos de comportamiento resultan menos invasivos aunque la exactitud de la identificación es menor debido a la variabilidad de los patrones de comportamiento.

De esta forma, el objetivo de todos ellos será obtener, a partir de la captura de un rasgo biométrico, una representación de cada individuo que resulte lo suficientemente discriminante respecto a las de los demás usuarios del sistema; de modo que, mediante técnicas de reconocimiento de patrones, el sistema sea capaz de determinar la identidad del usuario que está intentando acceder al sistema o simplemente si el usuario está registrado o no en él, dependiendo del tipo de aplicación para la que se utilice.

El reconocimiento de la voz es el proceso de reconocer automáticamente quién está hablando utilizando la información dada por las ondas de sonido emitidas [7][8]. En general, los sistemas de reconocimiento de voz tienen dos módulos principales: extracción de características y comparación de características. La extracción de características es el proceso por el cual extraemos una pequeña cantidad de datos de

la señal de voz que puede ser usada para representar a cada persona. La comparación de características involucra el proceso de identificar a la persona desconocida comparando las características extraídas de su voz, con las previamente obtenidas, que corresponden a las personas conocidas por sistema.

En biometrías de voz existen diferentes tipos:

- Dependiente de texto (modo limitado).
- Independiente del texto (modo ilimitado)

El principal tipo de aplicación de la interacción por voz son los conocidos como Sistemas de Diálogo Hablado (dependientes de texto). A la hora de implementarlos, principalmente se pueden optar por dos tipos de estrategias. La primera de ellas consiste en el uso de un lenguaje basado en comandos o palabras aisladas. Así, el usuario podría utilizar un comando como usar las palabras "Sí" o "No" en un diálogo dirigido, o utilizar como comandos una serie de números para seleccionar una opción en un sistema de asistencia telefónica en vez del sistema DTMF (Dual Tone Multiple Frequency) tradicional. La segunda opción es utilizar el lenguaje natural, o uno restringido a ciertas estructuras gramaticales.

Por otro lado, la Biometría independiente del texto resulta atractiva debido a que se puede aplicar a sistemas de seguridad, realmente hay ocasiones en las que se necesita mayor seguridad para ciertas actividades o simplemente es que se nos puede olvidar una clave o puede ser sustraída. Por eso se ha ido desarrollando proyectos biométricos, además de por la comodidad que le aporta al usuario, ya que no es necesario que recuerde una contraseña o que lleve una tarjeta o una identificación consigo.

La utilización de la voz como rasgo biométrico, además de cumplir propiedades como: universalidad, distintividad, estabilidad, evaluabilidad, no-intrusivo, etc.,

es un rasgo de fácil adquisición. Además, este tipo de biometría no necesita nada que la mayoría de los usuarios no tengan. Para poder utilizar este tipo de sistemas simplemente es necesario algo que permita realizar capturas de señales de voz, como un micrófono, ya sea en el móvil, en la tablet o en una computadora.

A través de los años se han desarrollado diferentes tipos de programas para el reconocimiento de voz que utilizan múltiples herramientas matemáticas como lo son: Dinamic Time Warping (DTW), modelos ocultos de Markov (HMM), redes neuronales, entre otros. Existen una serie de herramientas en la actualidad de bastante uso en el área de reconocimiento de voz tales como: HTK [9], CMU Sphinx [10], CSLU toolkit [11], entre otros.

En este sentido, en el presente trabajo se realizará un estudio comparativo de distintas toolkits para el reconocimiento de voz en base a los criterios de evaluación definidos, analizando sus fortalezas y debilidades.

Líneas de Investigación y Desarrollo

Una de las principales líneas de investigación de este trabajo son los sistemas de reconocimiento de personas mediante voz, utilizados ampliamente en la actualidad en aplicaciones o sistemas relacionados con la seguridad, ya que cada individuo tiene características físicas diferentes. No resulta sencillo realizar este tipo de aplicaciones ya que pueden existir inconvenientes tales como el ruido de fondo, enfermedades, edad, estados de ánimo en los cuales pueden cambiar la voz y también pueden existir diferencias entre sistemas de captación de la señal, ajenos al locutor.

El uso del procesamiento de la voz como un modo de interacción con los sistemas o de acceso a ellos, ofrece innumerables ventajas respecto a las interfaces clásicas de comunicación con los sistemas como por

ejemplo: bajo costo, facilidad de uso, aceptación por parte de los usuarios, captura y transmisión de manera simple, etc.

Resultados y Objetivos

El objetivo general de este trabajo es realizar un análisis comparativo de distintas toolkits para el reconocimiento biométrico de personas mediante voz. En principio, se realizará una investigación exploratoria de las distintas herramientas para el reconocimiento biométrico de personas mediante voz utilizadas en la actualidad. Para ello, se analizarán los sitios webs oficiales de las herramientas, manuales, foros, artículos, etc. que permitan conocer las principales características que presentan las distintas herramientas. Luego, se definirán los criterios de evaluación que permitirán realizar el análisis comparativo y finalmente se seleccionarán la/s herramienta/s consideradas más adecuadas para la implementación de un sistema de reconocimiento por voz.

Formación de Recursos Humanos

La estructura del equipo de investigación está formada por el Director, Co-Director y cuatro integrantes docentes. El Director del proyecto dirige la tesis de Maestría en Sistemas de Información (MSI) de la Facultad de Ciencias de la Administración de la UNER de tres integrantes y la Co-Directora, donde realizarán sus trabajos de investigación en el área del proyecto.

Además, el proyecto cuenta con un becario de Iniciación en la Investigación cuyas tareas están relacionadas con la captura, registración y almacenamiento de datos biométricos y un integrante interno alumno de la Licenciatura en Sistemas,

Referencias

1. Carlos E. Alvez, Marcelo G. Benedetto. "Los Sistemas Biométricos y su Factibilidad de Aplicación en los Organismos Estatales". XII Workshop de Investigadores en Ciencias de la Computación (WICC 2010), El Calafate, Santa Cruz Argentina, 5 y 6 de Mayo de 2010, Universidad Nacional de la Patagonia San Juan Bosco. Páginas 247-251.
2. Graciela Etchart, Lucas Luna, Carlos Leal, Marcelo Benedetto, Carlos Alvez. Sistemas de reconocimiento biométricos, importancia del uso de estándares en entes estatales. CGIV - XIII Workshop de Investigadores en Ciencias de la Computación (WICC 2011), 5 y 6 de Mayo de 2011. Universidad Nacional de Rosario. Rosario – Argentina. Páginas 339-343.
3. Graciela Etchart, Lucas Luna, Rafael Leal, Marcelo Benedetto, Carlos Alvez. "Sistema adecuado a estándares de reconocimiento de personas mediante el iris". CGIV - XIV Workshop de Investigadores en Ciencias de la Computación (WICC 2012), 25 y 26 de Abril de 2012. Universidad Nacional de Misiones. Posadas – Argentina. Páginas 321-325.
4. Graciela Etchart, Carlos Alvez, Marcelo Benedetto "Gestión de Datos Biométricos en Bases de Datos Objeto-Relacionales" XV Workshop de Investigadores en Ciencias de la Computación (WICC 2013). pp. 97-101. Paraná Entre Ríos.
5. Silvia Ruiz, Graciela Etchart, Carlos Alvez, Ernesto Miranda, Marcelo Benedetto, Juan José Aguirre " Representación e interoperabilidad de imágenes biométricas" XVII Workshop de Investigadores en Ciencias de la Computación (WICC 2015). Salta.
6. Anil K. Jain, Arun Ross, and Salil Prabhakar. An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology, 14(1), 2004.
7. Doddington, G. R. (1985). Speaker Recognition – Identifying people by their Voices. Proc. of the IEEE.
8. Furui, S. (1981). "Cepstral analysis technique for automatic speaker verification." IEEE Transactions on acoustic, speech and signal processing 29(2): 254- 277.
9. Cambridge University Engineering Department (CUED). Hidden Markov Model Toolkit (HTK), accedido 15 de marzo de 2017. <http://htk.eng.cam.ac.uk/>.
10. Carnegie Mellon University. CMUSphinx, accedido 15 de marzo de 2017. <http://cmusphinx.sourceforge.net/>.
11. Technology and Research Collaborations, Oregon Health & Science University CSLUtoolkit, accedido 15 de marzo de 2017. <http://www.cslu.ogi.edu/toolkit/>.