

ARQUITECTURA GENÉRICA PARA EL ALMACENAMIENTO DE DATOS BIOMÉTRICOS

Silvia Ruiz, Ernesto Miranda, Graciela Etchart, Carlos Alvez, Marcelo Benedetto, Juan José Aguirre, Santiago Delfin, Mauro Herlein

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

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

silruiz@fcad.uner.edu.ar, emiranda@fcad.uner.edu.ar, getchart@fcad.uner.edu.ar,
caralv@fcad.uner.edu.ar, marben@fcad.uner.edu.ar, juaagu@fcad.uner.edu.ar,
toti.delfin@gmail.com, herlein.mauro@gmail.com

Resumen

El objetivo de este trabajo es el desarrollo de una arquitectura genérica utilizando la tecnología de base de datos objeto relacional (BDOR), conforme a normas internacionales, para la identificación de las personas por medio de reconocimiento biométrico. La necesidad de contar con modelos basados en estándares, es permitir la interoperabilidad entre sistemas, para de esa manera, facilitar la búsqueda y el intercambio de datos. En el presente proyecto, se ha propuesto una arquitectura genérica para Iris, basada en el registro Tipo 17 del estándar ANSI/NIST-ITL 1-2011. La idea es extender esta arquitectura para el rasgo de voz con el objeto de incrementar la fiabilidad del sistema utilizando el reconocimiento multimodal.

Palabras clave: modelo objeto-relacional base de datos biométricas, interoperabilidad.

Contexto

Este trabajo se da en el marco del Proyecto PID 07/G044 “*Gestión de datos biométricos en base 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 objetivo general del presente proyecto es desarrollar sistemas para la identificación de personas mediante el reconocimiento de iris y voz que permitan la interoperabilidad entre organismos acorde a normas internacionales. Este trabajo, se enfoca principalmente en el desarrollo de arquitecturas de bases de datos para la representación e intercambio de datos biométricos. Como caso de aplicación, se tratarán imágenes de iris para el reconocimiento biométrico [6][7].

Introducción

La relevancia del uso de tecnologías basadas en biometrías se encuentra en su aplicación como mecanismo de control que a través de una serie de medidas de características específicas 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.

En el campo de la biometría, los estándares definen, entre otros aspectos,

los formatos de intercambio para el almacenamiento y la transmisión de los datos biométricos. Dichos formatos son importantes para aplicaciones en las que el ancho de banda es limitado y/o cuando el registro del rasgo biométrico debe ser almacenado en dispositivos con limitada capacidad.

En los últimos años, la Argentina ha demostrado ser un país con potencial en la región para implementar y desarrollar soluciones de biometría tanto para el sector público como privado [8][9].

Los organismos públicos que emplean esta tecnología necesitan consultar las bases de datos biométricas de manera eficiente e intercambiar datos con otros organismos asociados. De allí la necesidad de contar con modelos de datos y con estándares que permitan la interoperabilidad entre sistemas y así facilitar la búsqueda e intercambio de datos en bases biométricas. El ANSI/NIST-ITL 1-2011 [10], contempla este tipo de aspectos (entre otros). Esta norma biométrica publicada en noviembre de 2011, es el estándar más utilizado por entes estatales, y sobre el mismo se viene trabajando en el grupo de investigación del proyecto [3]. En diciembre de 2013, una actualización del estándar incorporó la especificación para el registro Tipo 11 referido al tratamiento de la voz como rasgo biométrico, lo cual es importante porque es uno de los rasgos sobre el que se trabajará en el presente proyecto.

La elección de la voz, como otro rasgo biométrico a fusionar con el iris, se debe que, además de cumplir propiedades como: universalidad, distintividad, estabilidad, evaluabilidad, no-intrusivo, etc., es un rasgo de fácil adquisición. Puede ser adquirida de una manera muy sencilla sin métodos invasivos, ni dispositivos especializados. Esto hace de la voz, un rasgo biométrico ideal para la arquitectura propuesta.

Actualmente, en el grupo de investigación se ha propuesto una arquitectura genérica para la gestión de imágenes de iris [11].

En la misma (Figura 1), las tareas de generación, transmisión y recepción de registros de Tipo 17 del estándar se encuentran contempladas dentro del sistema de intercambio de imagen del iris. Este sistema tiene un módulo que se encarga de la generación y transmisión de los registros necesarios para una transacción ANSI / NIST ITL 1-2011 y otro módulo para la recepción de estos registros a partir de organismos asociados. Tanto la estructura de registro de Tipo 17 como los metadatos para el reconocimiento del iris, son estructuras complejas. Esto puede causar algunos problemas cuando se trabaja con el modelo de datos relacional, debido a las limitaciones impuestas por el mismo. Por lo tanto, aquí se ha optado por la tecnología objeto-relacional (OR) [12].

Por otra parte, el intercambio de imágenes requiere en muchos casos transacciones web, con lo cual el tamaño de las mismas es importante. Un aspecto a tener especialmente en cuenta es la calidad de representación de las imágenes. En este sentido, los estándares de intercambio de datos también exigen que los formatos de compresión utilicen métodos que cumplan determinados requisitos. El estándar ANSI/NIST-ITL 1-2011, para imágenes de iris, sólo admite JPEG 2000 (Estándar ISO/IEC 15444:2004) [13] y PNG (Estándar ISO/IEC 15948:2004) [14]. Este aspecto también es contemplado en la arquitectura propuesta.

El estándar internacional (JPEG 2000) ISO/IEC 15444:2004 es un sistema de codificación de imágenes, que permite una compresión, transmisión y almacenamiento eficiente de imágenes fijas y de secuencias de imágenes. Con este estándar es posible adecuar la secuencia de bits que representa la información de la imagen comprimida a las características del canal de transmisión, del dispositivo de almacenamiento, o del dispositivo de visualización, independientemente del

tamaño, el número de componentes, y la precisión de la imagen original.

El estándar internacional ISO/IEC 15948: 2004 PNG (Gráficos de Red Portables) especifica un flujo de datos y un formato de archivo asociado para una imagen sin pérdidas, comprimida, portátil y que puede ser transmitida a través de Internet. Entre las principales características del formato PNG se encuentra la portabilidad. El formato es flexible a aceptar agregados y extensiones futuras sin comprometer el intercambio de los flujos de datos PNG.

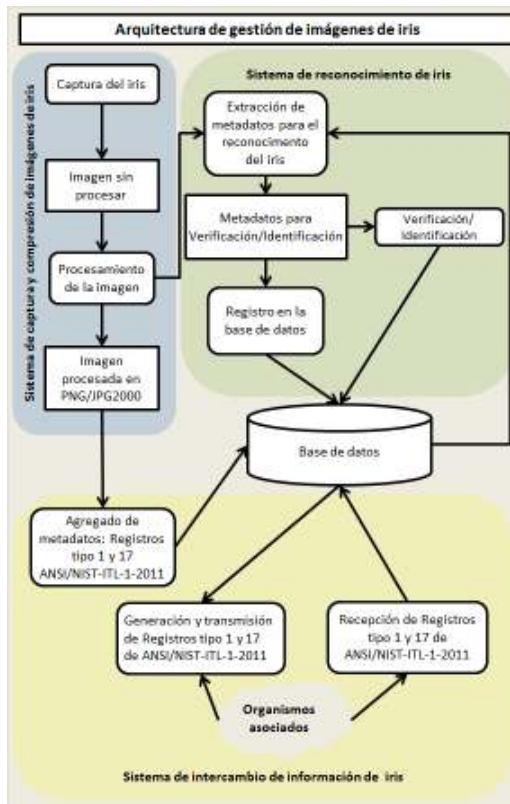


Figura 1: Arquitectura para gestión de imágenes de iris

Líneas de Investigación y Desarrollo

Una de las líneas más importantes, consideradas en el proyecto, consiste en la construcción de modelos para la representación de datos relacionados con los datos biométricos, de manera que los mismos puedan compartirse, consultarse, recuperarse y compararse de manera

simple y eficiente. En particular, el reconocimiento del iris es ampliamente utilizado para fines de seguridad. Asimismo, como se mencionó con anterioridad, especialmente en organismos gubernamentales, es muy importante considerar el nivel de estandarización alcanzado, sobre todo, en lo referente al formato de intercambio de datos y a una arquitectura de base de datos que facilite el almacenamiento y la recuperación de los datos biométricos. En nuestro caso, por su capacidad de adaptación a diferentes dominios, se utilizará el modelo de base de datos objeto relacional.

Otra línea de investigación vigente está orientada a sistemas que combinan diferentes rasgos biométricos o bases de datos biométricas multimodales. Estos sistemas son más precisos y seguros dado que superan algunas limitaciones de los sistemas que utilizan un único rasgo biométrico. Las limitaciones de un rasgo pueden ser: no universalidad, ruidos en los datos, suplantación de identidad, entre otros [15]. A pesar de que el tiempo de procesamiento, el hardware requerido y los requisitos de almacenamiento en bases de datos multimodales pueden ser mayores a los empleados por las bases de datos unimodales, las ventajas de aquellas hacen que se prefiera su utilización. [16] Por lo antes expuesto, se extenderá la arquitectura propuesta para el tratamiento del rasgo de voz (registro Tipo 11 del estándar ANSI/NIST-ITL 1-2011).

Resultados y Objetivos

El objetivo general del Proyecto 07/G044 es desarrollar sistemas para la identificación de personas mediante el reconocimiento de iris y voz que permitan la interoperabilidad entre organismos. El objetivo específico en que se está trabajando y que se presenta en este artículo, es el desarrollo de una arquitectura de base de datos que facilite

el almacenamiento y la recuperación de los datos biométricos.

En esta etapa del proyecto se encuentra en desarrollo una arquitectura genérica utilizando la tecnología de las BDOR, de acuerdo con las normas internacionales, para la identificación de las personas por medio de reconocimiento de iris. Para dicha arquitectura se propuso un modelo expresado en un diagrama de clase UML en el que se definen los tipos de datos de dominio a utilizar en la arquitectura.

La arquitectura presentada se adaptará para el soporte de datos biométricos de voz. Esto deberá contemplar, al igual que con el iris, lo especificado en el estándar ANSI/NIST-ITL 1-2011, los métodos para extracción de características y el almacenamiento de las secuencias de los vectores a utilizar para el reconocimiento [17].

En etapas posteriores se trabajará en la implementación de los métodos que gestionen los tipos específicos para los datos biométricos crudos, los tipos de datos para la codificación de los mismos, y todos los metadatos necesarios para generar los registros de transacciones ANSI/NIST-ITL1-2011. Además, se estudiarán alternativas para la creación de índices de dominio que permitan mejorar los tiempos de respuestas en los procesos de reconocimiento.

Formación de Recursos Humanos

La estructura del equipo de investigación está formada por el director y co-director, 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 un integrante. Además, dos integrantes y la Co-Directora se encuentran cursando la MSI, 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 alumno, de la Licenciatura en Sistemas, que colabora en la implementación de la arquitectura propuesta.

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. Daugman, J.: High condence visual recognition of persons by a test of statistical independence. IEEE Transactions on Pattern Analysis and Machine Intelligence 15(11), 1148-1161 (November 1993).
7. J. Daugman and C. Downing, "Effect of severe image compression on iris recognition performance," IEEE Transactions on Information Forensics and Security, vol. 3, no. 1, p. 52-61, 2008.
8. Casal Gabriel, Revolva Mercedes. Biometrías. Herramientas para la Identidad y la Seguridad Pública. Jefatura de Gabinete de Ministros. Presidencia de la Nación. Noviembre de 2010.
9. Julio Fuoco, Tendencias Biométricas, desafíos y oportunidades. En Biometrías 2. Jefatura de Gabinete de Ministros. Presidencia de la Nación. Octubre de 2011
10. Wing B. ANSI/NIST-ITL 1-2011. Update: 2013. Information Technology: American National Standard for Information Systems Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information. December, 2013.
11. Carlos Alvez, Graciela Etchart, Silvia Ruiz, Ernesto Miranda, Juan Aguirre, Marcelo Benedetto, Santiago Delfin "Iris Information Management in Object-Relational Databases". XXI Congreso Argentino de Ciencias de la Computación (Junín, 2015), pp. 741-750. Junin, Argentina. Octubre 2015.
12. Melton Jim, "(ISO-ANSI Working Draft) Foundation (SQL/ Foundation)", ISO/IEC 9075-2:2003 (E), United States of America (ANSI), (2003).
13. ISO/IEC 15444-1:2004 Information Technology-JPEG 2000 Image Coding System: Core Coding System, Second Edition, 2004-09-15.
14. ISO/IEC 15948:2004 Information Technology-Computer Graphics and Image Processing-Portable Netware Graphics (PNG): Functional Specification, First Edition, 2004-03-01.
15. Alvez C., Benedetto M., Berón G., Etchart G., Luna L. y Leal C. Desarrollo de un sistema multi-biométrico mediante reconocimiento de iris y voz, adecuado a estándares, para su aplicación en organismos públicos. SIE 2011 - Simposio de Informática en el Estado. Córdoba, 31 de Agosto, 01 y 02 de Septiembre de 2011. 40° JAIIO. Páginas: 206 - 220.
16. M. Ghayoumi, "A review of multimodal biometric systems: Fusion methods and their applications," Computer and Information Science (ICIS), 2015 IEEE/ACIS 14th International Conference on, Las Vegas, NV, 2015, pp. 131-136
17. Reynolds, D. A., Campbell, W., Gleason, T. T., Quillen, C., Sturim, D., Torres-Carrasquillo, P., Adami, A., 2005. The 2004 MIT Lincoln Laboratory speaker recognition system. In: Proc. of IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing, ICASSP. Vol. 1. pp. 177-180.