

Entorno experimental para el prototipado de aplicaciones biométricas multimodales en dispositivos móviles

Álvaro Hernández-Trapote¹, Rubén Fernández¹, Beatriz López-Mencia¹,
Álvaro Sigüenza¹, Luís Hernández¹, Javier Caminero², Doroteo Torre-Toledano³

¹ GAPS, Señales Sistemas y Radiocomunicaciones. Universidad Politécnica de Madrid,
Ciudad Universitaria s/n, 28040 Madrid, Spain

{alvaro, ruben, beatriz, alvaro.sigüenza, luis }@gaps.ssr.upm.es

² Telefónica I+D, Emilio Vargas 6. 28043 Madrid, Spain. fjcg@tid.es

³ ATVS - Biometric Recognition Group, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Calle Francisco Tomas y Valiente 11, 28049 Madrid, Spain. doroteo.torre@uam.es

Abstract. En este artículo hemos presentado un entorno que soporta la creación de prototipos de aplicaciones biométricas en escenarios móviles. Para ello se facilitan diversos componentes que por un lado, acceden a los recursos audiovisuales de una PDA, y por otro conectan con un servidor para proceder al tratamiento de la información biométrica. Como ejemplo de aplicación práctico del entorno de experimentación se ha desarrollado una utilidad con la que se ha recopilado una base de datos biométrica multimodal de voz y caras en la que han participado 72 personas.

Keywords: entorno experimental, biometría, dispositivos móviles, bases de datos biométricas.

1 Introducción

En la actualidad existe una marcada tendencia hacia la miniaturización y la creación de entornos y ambientes ubicuos. Esta tendencia hace que, cada vez más, los dispositivos móviles y el desarrollo de servicios para éstos empiecen a atraer mayor atención dentro del panorama industrial. Debido a este interés tan extendido, las garantías de fiabilidad y seguridad de la información intercambiada entre dispositivos portátiles se han convertido en un aspecto clave para los usuarios [1], especialmente en determinados servicios que requieren cierto grado de confidencialidad (operaciones bancarias, acceso seguro a servicios personales...). Así, actualmente en aplicaciones desarrolladas para teléfonos móviles, PDAs o Pocket PCs se está empezando a incorporar tecnologías biométricas para satisfacer esta necesidad de seguridad [2], [3]; además, las características propias de estos dispositivos (cámara/s incorporada/s, micrófonos, etc.) hacen que la inclusión de algunos modos biométricos no requiera de ningún hardware adicional, lo que constituye una ventaja para el desarrollo de estas tecnologías en dispositivos móviles.

Sin embargo, a pesar de la madurez de las tecnologías biométricas, el desarrollo de este mercado ha sido muy escaso en comparación con las expectativas previamente creadas. Existiendo una clara carencia de entornos experimentales que den soporte tanto a la evaluación de las diferentes tecnologías biométricas, como al análisis de usabilidad y aceptabilidad de interfaces que incorporen dichas tecnologías en dispositivos móviles. Así, informes de organismos oficiales como el National Science and Technology Council (NSTC) fijan como retos para el desarrollo de las tecnologías biométricas esta necesidad de creación de nuevos entornos y arquitecturas abiertas para la experimentación, así como el análisis de interfaces de usuario intuitivos seguido de la especificación de guías de diseño para su futura adopción [4]. Actualmente podemos encontrar interesantes desarrollos en este sentido, que hacen uso de iniciativas como el BioAPI [5], interfaz creado para la estandarización de la comunicación con los sensores biométricos, y que permiten implementar de una manera fácil y flexible aplicaciones biométricas en diversos entornos. Por ejemplo, en [6] se describe una arquitectura Java que hace uso del BioAPI para el acceso a los sensores y con la que se pueden diseñar aplicaciones biométricas que se ejecuten sobre ordenadores de escritorio. No obstante aún son necesarios esfuerzos para la creación de plataformas biométricas aplicables a entornos móviles.

En este artículo presentamos un entorno experimental que proporciona recursos para trabajar con las tecnologías de reconocimiento de voz y cara sobre dispositivos móviles. En la Sección 2 describimos este entorno experimental junto a la arquitectura que lo soporta. Destacándose sus posibilidades para permitir un fácil y rápido prototipado de aplicaciones que pueden utilizarse para diversas tareas como la evaluación de diferentes tecnologías biométricas, estudios de usabilidad de interfaces biométricos, o la captura de bases de datos biométricas siempre sobre escenarios móviles, con posibilidad de combinación de modalidades (multimodalidad), y entornos de trabajo realistas, etc. Como ejemplo de aplicación de nuestro entorno en la Sección 3 describimos el desarrollo de una aplicación para la captura de una base de datos de voz y caras sobre una PDA; adicionalmente, se explica cómo la aplicación se utilizó para la recolección de una base de datos de 72 usuarios. Finalmente en la Sección 4 presentamos las conclusiones de nuestro trabajo y cómo planteamos las líneas futuras de investigación.

2 Entorno experimental

El entorno experimental que hemos desarrollado está implementado sobre una arquitectura distribuida cliente-servidor (ver Figura 1). La decisión de adoptar esta configuración surgió con el objetivo de liberar de carga de procesamiento al dispositivo móvil. En este contexto de un entorno distribuido, se desarrollaron dos controles ActiveX que conforman la base principal de la arquitectura en el lado cliente, y que implementan las funciones de obtención de los rasgos biométricos, y de envío al servidor para su tratamiento. Una vez que la información biométrica ha sido capturada en el lado cliente, se envía al servidor donde se procesa según la utilidad de la aplicación que se haya diseñado. En la Figura 1 se puede observar que la comunicación entre cliente y servidor se lleva a cabo a través de una conexión WIFI.

No se han tomado precauciones especiales de encriptación en la transmisión más allá del uso del estándar WPA, y es que a pesar de ser éste un aspecto crítico en los sistemas de seguridad biométrica, su estudio está lejos del objetivo de este entorno que, como ya hemos indicado anteriormente, se orienta al apoyo de la investigación en tecnologías biométricas y a estudios de usabilidad para interfaces seguros sobre terminales móviles.

Antes de describir en más detalle los elementos que forman la arquitectura, creemos interesante presentar cuáles fueron los criterios de diseño que fijamos para la implementación del entorno de experimentación:

- *Experimentación en dispositivos móviles*, como ya se ha comentado anteriormente, los dispositivos móviles forman parte de un sector que reclama mayor seguridad, y la biometría puede ser una opción muy interesante para cubrir esta necesidad. La oferta de sistemas operativos para dispositivos móviles es variada, pero la presencia dominante en éstos del SO Windows Mobile y las facilidades en cuanto a portabilidad de los componentes ActiveX nos llevaron a decidimos por implementar objetos COM que ejecutaran las funciones en el cliente. Al mismo tiempo, estos dispositivos nos permitirán llevar a cabo experimentos con usuarios bajo diferentes condiciones experimentales.
- *Posibilidad de incluir elementos interactivos*, en la Sección 1 se comentaba la falta de estudios de usabilidad sobre interfaces de usuario para tecnologías biométricas. En este tipo de estudios es realmente interesante incluir elementos interactivos como avatares (p.ej.: software Flipz [7] para generar vídeos de bustos parlantes en formato Flash) o componentes multimedia (p.ej.: objetos Windows Media Player, ver Sección 3). Una de las mejores plataformas para este cometido son los navegadores, y más aún cuando se prevén utilizar módulos ActiveX.
- *Facilitar la creación de interfaces*, teniendo en cuenta que se utilizarán componentes ActiveX embebidos en un navegador, los interfaces se podrán generar utilizando recursos web, como el HTML para la presentación gráfica, y el lenguaje Javascript para la interacción y acceso a los componentes activos.

A continuación presentaremos en más detalle los elementos que componen la arquitectura.

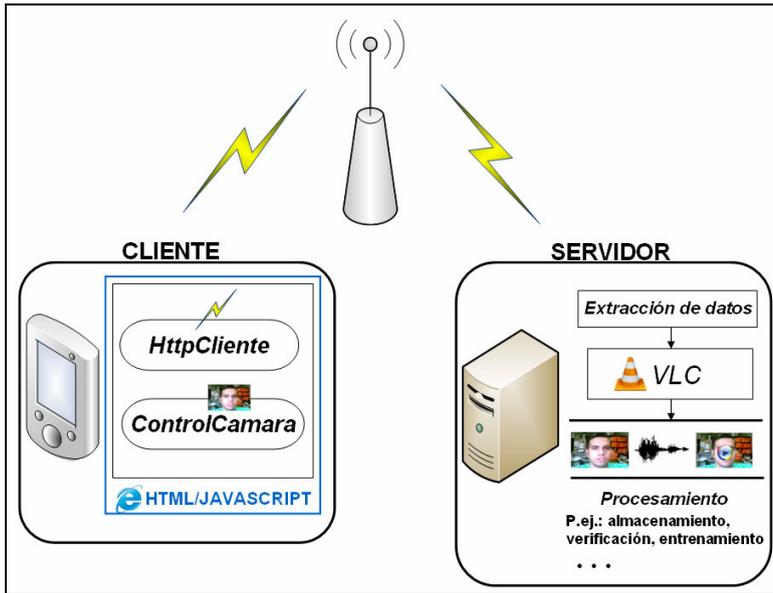


Figura 1. Representación de la arquitectura para el estudio de sistemas de seguridad biométrica sobre dispositivos móviles

2.1 Descripción de la arquitectura en el lado del cliente

Como hemos comentado anteriormente, la base principal del entorno en el lado cliente son dos componentes ActiveX que permitirán acceder a los recursos de la cámara del dispositivo móvil, y enviar las muestras recogidas al servidor. Ambos archivos han sido comprimidos en un archivo *cabinet* (.CAB) para su distribución e instalación. El primero de los componentes ActiveX es 'ClienteCamara', que implementa los métodos listados en la Tabla 1. 'ClienteCamara' está pensado para obtener fotos y vídeos (con audio incorporado) de los usuarios y está basado en OpenCV, un conjunto de librerías de libre distribución para el tratamiento de imágenes [8]. El ActiveX acepta un parámetro (*VideoRenderFile*) mediante el que podemos definir la ruta del archivo temporal en el que se irá guardando la captura de la foto o la grabación del vídeo. Además, para facilitar la gestión de la interacción 'ClienteCamara' lanza un evento cuando se termina un proceso de captura.

Una vez obtenidos los rasgos biométricos del usuario, éstos pueden ser enviados al servidor para su procesamiento. El encargado de realizar esta operación es el componente ActiveX de nombre 'ClienteHTTP'. Como se puede intuir de la definición de los métodos de la Tabla 2, su funcionalidad consiste en realizar el envío al servidor, mediante protocolo HTTP y una petición POST, de cadenas de caracteres o archivos (fotos, vídeos, etc.) previamente capturados en el lado cliente.

Tabla 1. Definición y descripción de los métodos accesibles en el componente ActiveX 'ClienteCamara'

| Método | Descripción del método |
|---|--|
| InitGraph () | Construye los objetos gráficos internos del componente. Al crear el control ActiveX este método se llama automáticamente, por lo que sólo sería necesario volver a invocarlo en el caso de que detenga el control mediante <i>ShutdownGraph ()</i> |
| SetMessage (message) | Presenta el mensaje <i>message</i> en la ventana del objeto ActiveX. Este mensaje no se mostrará por pantalla hasta que no llamemos a <i>ShowMessage (true)</i> |
| TakePicture (file) | Toma una foto y la guarda en el archivo <i>file</i> |
| StartVideoCapture (file) | Inicia la captura de un vídeo e indica en qué fichero, <i>file</i> , se almacenará |
| StopVideoCapture () | Para la captura del vídeo |
| ShutDownGraph () | Detiene los objetos de control de la cámara |
| Showimage (file) | Muestra la foto del archivo <i>file</i> en la pantalla del ActiveX |
| ShowMessage (bShow) | Habilita la presentación por pantalla del mensaje definido mediante el método <i>SetMessage (message)</i> |
| ShowPreviewWindow (bShow) | Habilita la ventana de previsualización de la cámara |
| SetPreviewPosition (Left, Top, Width, Height) | Establece el tamaño y la posición de la ventana de previsualización dentro del componente ActiveX |

Tabla 2. Definición y descripción de los métodos accesibles en el componente ActiveX 'ClienteHTTP'

| Método | Descripción del método |
|----------------------------|---|
| Connect (address , port) | Conecta con el servidor alojado en la dirección <i>address</i> por el puerto <i>port</i> |
| Close () | Cierra la conexión con el servidor |
| InitializePostArguments () | Inicializa los argumentos de la llamada POST. Es necesario llamar a este método antes de hacer alguna llamada <i>AddPostArguments()</i> |
| AddPostArguments () | Añade cadenas de caracteres y archivos a las llamadas POST al servidor |

2.2 Descripción de la arquitectura en el lado del servidor

En el lado del servidor se ha utilizado la tecnología Java Servlet. Quizás es necesario recordar que la tecnología Servlet de Java consiste en pequeños programas en lenguaje Java que se activan con una petición desde el cliente. Para manejar e invocar dichos servlets hemos utilizado un contenedor de servlets como Apache Tomcat [9]. En nuestra arquitectura, el servidor alojará un servlet 'genérico' que tiene asignadas básicamente dos tareas:

- En primer lugar nuestro servlet recogerá los mensajes y archivos enviados desde el terminal móvil. Para ello se ha utilizado el paquete *FileUpload* del proyecto

Apache Commons [10]. Este paquete pone a disposición del desarrollador una serie de clases que permiten implementar fácilmente un procedimiento para extraer los objetos que viajen en la petición POST.

- Por otro lado, una vez recogidos las fotos y los vídeos, el servlet podría implementar una serie de métodos para el procesamiento de la información capturada en función del prototipo que se quiera implementar (p.ej.: almacenamiento para la creación de una base de datos, procesos de entrenamiento o verificación para un sistema de autenticación biométrica). Como ejemplo nuestro entorno proporciona un método que haciendo uso del programa de código abierto VLC [11] extrae la información del audio a partir del vídeo capturado en el cliente.

3 Caso práctico: base de datos biométrica multimodal en un entorno móvil

Para demostrar la versatilidad y funcionalidad del entorno experimental implementado describimos en este apartado la adquisición de una base de datos en un entorno real para un sistema de verificación que emplea las tecnologías de reconocimiento de locutor y cara sobre una PDA. Esta base de datos ha sido capturada intentando simular unas condiciones de adquisición lo más realistas posibles, incluyendo:

- Escenarios de captura remotos sin supervisión de un experto.
- Diferentes condiciones de iluminación (artificial y natural) y presencia de ruido en las locuciones de voz, para valorar la robustez de los algoritmos de verificación biométrica frente a dichas fuentes de variabilidad.
- Usuarios sin experiencia previa en el uso de dispositivos biométricos.

3.1 Bases de datos biométricas multimodales en entornos móviles

Uno de los problemas para el desarrollo de las tecnologías biométricas en entornos móviles es la escasa cantidad de recursos y bases de datos capturadas en estas condiciones, importantes tanto para contrastar experimentos, como para utilizarlas en la evaluación de los distintos modos biométricos en función de las condiciones de entorno. Entre estas iniciativas podemos resaltar algunas interesantes como [12], donde intentan simular las condiciones de un entorno móvil capturando la información biométrica mediante una webcam conectada a un portátil. En el caso del proyecto *SecurePhone*, nos encontramos con una base de datos que sí fue capturada mediante una PDA [13] y que contiene 60 usuarios. Sin embargo todavía existe una carencia de este tipo de recursos que se hace aún más evidente si queremos disponer de bases de datos en castellano.

3.2 Recogida de los datos

Nuestra base de datos está formada por 72 participantes (46 hombres y 26 mujeres). Para crear la base de datos, y con el fin de capturar la variabilidad temporal intra-usuario, tal como se propone en [13], hemos dividido el proceso de captura en dos sesiones, dejando un espacio de tiempo entre sesiones de una semana. Además, para cada una de las sesiones anteriores, se capturaban los rasgos biométricos del participante (voz y vídeo) en dos condiciones (o escenarios) diferentes para intentar simular las condiciones reales de funcionamiento del sistema en la PDA.

El *Escenario A* se lleva a cabo bajo condiciones ambientales controladas, esto es, el participante es instado a situarse en lugares bien iluminados donde el foco de luz le viene de frente y de forma homogénea (preferentemente luz exterior natural) y además se solicitaba al usuario que intentara controlar el entorno para protegerlo frente al ruido. En el *Escenario B*, pretendemos que la captura se realice bajo condiciones de iluminación y ruido ambiental realistas, de forma que la calidad de las imágenes y grabaciones va a ser inferior. Ya que en una situación real, el usuario no va a poder siempre acceder al sistema en las condiciones ideales del *Escenario A*, los usuarios también grababan su información biométrica en un entorno donde las fuentes de luz no están controladas, de modo que pueden aparecer sombras en las caras del usuario que dificulten el reconocimiento. Para las grabaciones de voz tampoco se prestará especial atención al ruido de fondo presente en el momento de la grabación.

Para cada escenario (A y B), se capturaron imágenes estáticas de la cara (en vista frontal) y secuencias audiovisuales (vídeos) como explicamos a continuación:

Secuencias audiovisuales:

- Se capturaron 10 vídeos de cada participante.
- Los usuarios eran instados a repetir una secuencia de 5 dígitos por cada vídeo.
- Todos los participantes repetían las mismas 10 secuencias numéricas, de tal forma que, tal como estaban distribuidos los dígitos, asegurábamos que se realizasen al menos 5 repeticiones de cada número. Además, en el diseño de las secuencias se tuvo en cuenta las posiciones de los dígitos dentro de ellas, con el fin de incluir las posibles diferencias de pronunciación y de transiciones que deberían producirse, según sean dichos los dígitos al principio, en una posición intermedia de la grabación o al final de la misma.
- El sistema de captura comprime el vídeo en el perfil simple del codificador de *Windows Media Video 9 Series*, cuyas características son un *bit rate* de 96Kbps y una resolución de 176x144 @ 15Hz (QCIF), degradando la calidad del vídeo y audio capturado, lo que va a afectar a la robustez del proceso de verificación.

Imágenes estáticas:

- 10 capturas.
- 1 imagen facial estática frontal por cada captura.
- La imagen se recoge con formato 24-bit RGB y una resolución de 176x144.

El dispositivo empleado para la captura de la base de datos ha sido una PDA modelo Fujitsu Siemens Pocket LOOX T830.

3.3 Aplicación de captura de la base de datos

Utilizando como punto de partida el entorno distribuido experimental descrito en la Sección 2, se implementó una aplicación Web para la recogida de la base de datos de datos biométricos. Esta herramienta consiste en un interfaz HTML (Figura 2), que mediante Javascript accede a los métodos y eventos implementados en los controles *ActiveX* detallados anteriormente. En concreto se utilizaron:

- El control '*ClienteCamara*' para acceder a los recursos de la cámara de la PDA y permitir la grabación de los videos y la captura de las imágenes de la cara.
- El control '*ClienteHTTP*', que permite la comunicación entre la PDA y el servidor. La aplicación proporciona la posibilidad de obtener las fotografías y los vídeos de los usuarios de una forma ordenada y fácil. Así, el interfaz está diseñado de tal manera que el usuario puede elegir la sesión y el escenario en los que está realizando la captura en ese momento.
- Por último, y como ejemplo de componente multimedia para facilitar la interacción del usuario con el interfaz, se incluyó el control *ActiveX* del reproductor Windows Media Player '*wmp.dll*' para la previsualización de los videos durante la captura de la base de datos. De esta forma, se permite la posibilidad de regrabado si existe algún elemento (ruido, deficientes condiciones de iluminación, presencia de algún agente externo, etc.) que interfiera en el proceso de captura.



Figura 2. Interfaz para la captura de la base de datos

4 Conclusiones

Durante los últimos años se han producido importantes avances en el desarrollo de aplicaciones biométricas para dispositivos móviles; incluso existen algunos productos comerciales que están empezando a estar disponibles. Sin embargo todavía hay algunos aspectos que necesitan ser estudiados para garantizar la viabilidad y el buen funcionamiento de estas aplicaciones en escenarios reales [14]. Para intentar contribuir en esta línea de investigación, hemos desarrollado un entorno experimental que pueda ser un soporte para el desarrollo rápido de prototipos de aplicaciones biométricas sobre entornos móviles. Este entorno pretende ser sencillo y flexible para que pueda ser empleado para plantear estudios o contribuciones en diferentes áreas, como por ejemplo, análisis de usabilidad, pruebas en entornos reales, creación de bases de datos biométricas multimodales, etc.

Como caso práctico del entorno experimental propuesto, hemos elaborado una aplicación de captura de bases de datos biométricas multimodales. Con esta herramienta hemos elaborado una base de datos que consiste en imágenes de caras y vídeos de 72 usuarios en condiciones reales. La información biométrica fue recogida bajo dos escenarios distintos, uno de ellos con un entorno controlado y el otro bajo condiciones más realistas. Esperamos que estos recursos puedan servir para promover el estudio de las tecnologías biométricas en entornos móviles.

Con el entorno experimental que hemos presentado pretendemos implementar un escenario simulado donde un usuario, a través de una PDA, acceda a un área personal utilizando sus rasgos biométricos (p.ej.: caras y voz). Posteriormente, prevemos llevar a cabo análisis de usabilidad que creemos que nos permitirán obtener conocimientos sobre aquellos aspectos de la interacción que puedan estar relacionados con la calidad de servicio percibida por el usuario.

Agradecimientos. Las actividades descritas en este artículo han sido financiadas por el Ministerio Español de Ciencia y Tecnología bajo el proyecto TEC2006-13170-C02-01.

Referencias

1. Koreman, J., Morris, A.C., Wu, D., Jassim, S., Sellahewa, H., Ehlers, J., Chollet, G., Aversano, G., Bredin, H., Garcia-Salicetti, S., Allano, L. Ly Van, B. y Dorizzi, B., "Multi-modal biometric authentication on the SecurePhone PDA", Proc. Of Multi-Modal User Authentication Workshop (MMUA), Toulouse, France. (2006)
2. Oki Press Releases 2006. Vodafone K.K. Selects Oki Electric's Face Sensing Engine for Face Recognition on Mobile Phone. Información disponible en: <http://www.oki.com/en/press/2006/z05134e.html>
3. OMRON R&D. "OKAO Vision" Face Sensing Technology. Información disponible en: http://www.omron.com/r_d/vision/01.html
4. National Science and Technology Council Subcommittee on Biometrics. "The National Biometrics Challenge". Agosto 2006. Información disponible en <http://www.biometrics.gov/Documents/biochallengedoc.pdf>
5. BioAPI Consortium, <http://www.bioapi.org/>
6. Otero-Muras, E, Gonzalez-Agulla, E, Alba-Castro, J.L., Garcia-Mateo, C., Marquez-Florez, O.W, "An Open Framework For Distributed Biometric Authentication In A Web Environment". Annals of Telecommunications. Vol. 62, No. 1-2. Special issue on multimodal biometrics, (2007)
7. Flipz Software Technology, Disponible en: <http://www.flipz.tv/>
8. Open Computer Vision Library. Disponible en: <http://sourceforge.net/projects/opencvlibrary/>
9. The Apache Software Foundation. Apache Tomcat. <http://tomcat.apache.org/>
10. The Commons FileUpload package. Disponible en: <http://commons.apache.org/fileupload/>
11. VideoLAN - VLC media player. <http://www.videolan.org/>
12. Hazen T. J, y Schultz D., "Multi-Modal User Authentication from Video for Mobile or Variable-Environment Applications", Interspeech 2007, pages 1246-1249, (2007)
13. Morris, A.C., Koreman, J., Sellahewa, H. Ehlers, J., Jassim, S., Allano, L. y Garcia-Salicetti, S., "The SecurePhone PDA database, experimental protocol and automatic test procedure for multi-modal user authentication", Tech. Report, Saarland University , Institute of Phonetics, (2006)
14. International Biometric Group, "Survey: Biometrics enter mobile world", Biometric Technology Today – September (2005).