

ANÁLISIS DE ESCENARIOS PARA RECONOCIMIENTO BIOMÉTRICO A DISTANCIA ADECUADOS PARA LA ADQUISICIÓN DE IMÁGENES MMW

M. Moreno-Moreno, J. Fierrez, P. Tome, J. Ortega-García.

{miriam.moreno, julian.fierrez, pedro.tome, javier.ortega}@uam.es

Avda. Francisco Tomás y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain

Abstract— Con objeto de analizar en qué medida puede ayudar el uso de imágenes adquiridas en el rango de ondas milimétricas (MMW) al reconocimiento biométrico de personas a distancia, se presenta un análisis experimental de tres escenarios con diferentes distancias de captura adecuadas para la adquisición de imágenes MMW. En concreto se realizan experimentos de reconocimiento facial a corta, media y larga distancia entre la cámara y el sujeto a identificar, siendo los escenarios de media y larga distancia los más adecuados para las capturas en MMW. Los tres escenarios consideran plantillas registradas en condiciones controladas, y para su estudio se usan datos del NIST Multiple Biometric Grand Challenge. Este enfoque permite: 1) aproximarnos al problema del reconocimiento biométrico de imágenes en MMW como complemento a biometrías tradicionales como es el reconocimiento facial, y 2) entender los factores de mayor variabilidad que afectan al reconocimiento de personas a distancia, que serán los factores a compensar con las tecnologías complementarias al reconocimiento, como MMW. El análisis de escenarios se realiza cuantitativamente de dos maneras. Primero se analiza el contenido de información en caras segmentadas en los diferentes escenarios. A continuación se evalúa el rendimiento en cada escenario de tres comparadores, uno comercial, y dos aproximaciones estándar usando características (PCA y DCT) y comparadores (SVM y GMM) tradicionales. Los resultados muestran hasta qué punto las características de la adquisición influyen en el rendimiento de la verificación del reconocimiento facial a distancia, y esto sirve de punto de partida para comprender en qué medida pueden ayudar al reconocimiento tecnologías complementarias como MMW.

I. INTRODUCCIÓN

Las imágenes adquiridas a frecuencias de milimétricas (MMW) presentan propiedades únicas, que permiten su aplicación en diferentes ámbitos (seguridad, inspección no intrusiva, medicina y biometría entre otras) [1], [2]. A pesar del actual desarrollo de los sistemas de adquisición de este tipo de imágenes [3], [4], no existe ninguna base de datos de imágenes de personas en este rango espectral. Ello es debido, en parte, a los problemas de privacidad que presentan dichas imágenes. Este hecho limita considerablemente el desarrollo de sistemas biométricos basados en imágenes adquiridas en este rango espectral, los cuales necesitan un número suficiente de imágenes de diferentes usuarios. Dado que la adquisición de estas imágenes se realiza a diferentes distancias y puesto que

no existe ninguna base de datos de imágenes en MMW, el trabajo presentado se aproxima al problema del reconocimiento biométrico de personas con imágenes en MMW, pero usando imágenes adquiridas en el visible en un escenario similar al de la adquisición de imágenes en MMW. En concreto el artículo realiza un análisis de tres escenarios para reconocimiento facial a distancia.

La cara y el iris son dos de los rasgos biométricos usados actualmente en muchas aplicaciones de reconocimiento de usuario a distancia [5], [6]. Una nueva línea de investigación que está ganando popularidad se centra en el uso de estos rasgos biométricos en escenarios menos controlados de una manera no-invasiva incluyendo adquisición en movimiento ("On the Move") y a distancia ("At a Distance") [7]. Este tipo de escenarios está todavía en su infancia y se necesita mucha investigación y desarrollo para alcanzar los niveles de precisión y de rendimiento que requieren las aplicaciones prácticas.

El nuevo campo de la biometría a distancia se ha hecho posible primordialmente gracias a: 1) los recientes avances en los sensores, y 2) los nuevos algoritmos y métodos para tratar con factores variables (e.g. iluminación, movimiento, pose, distancia a la cámara), que en este caso están menos controlados que en las situaciones consideradas comúnmente en biometría.

Como resultado del interés en estas aplicaciones biométricas a distancia, hay un número creciente de trabajos de investigación que estudian cómo compensar las principales degradaciones encontradas en escenarios no controlados [8]. Sin embargo, apenas existe conocimiento experimental sobre los factores de variabilidad en escenarios específicos, que puedan ayudar en la creación de métodos robustos para la biometría a distancia (diseñada para aplicaciones específicas de importancia práctica). La contribución de este artículo es hacia este fin, analizando cuantitativamente tres escenarios del reconocimiento facial a distancia, a saber: corta, media y larga distancia entre el usuario y la cámara; siendo los escenarios de media y larga distancia los más adecuados para incorporar imágenes del cuerpo en MMW. Este análisis se realiza a dos niveles: 1) los principales datos estadísticos de

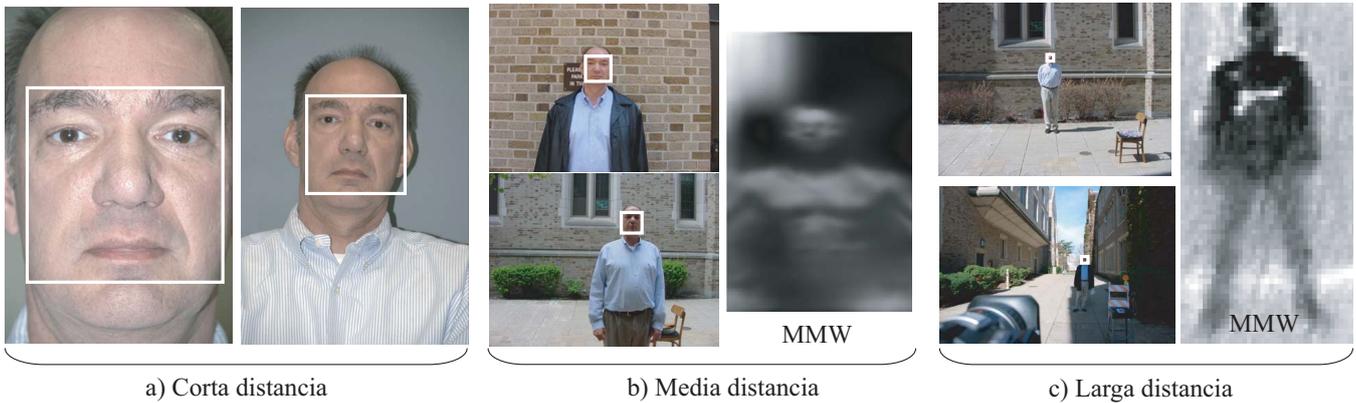


Fig. 1. Ejemplo de imágenes de los tres escenarios: a) corta distancia, b) media distancia, y c) larga distancia.

la información de contenido, y 2) el rendimiento de varios sistemas de reconocimiento: uno comercial, y dos basados en características (PCA, DCT) y comparadores tradicionales (SVM, GMM).

Los escenarios bajo estudio han sido extraídos de NIST Multiple Biometric Grand Challenge [9], el cual se centra en el reconocimiento biométrico a distancia usando el iris y la cara. En particular, usamos un subconjunto de estos datos compuesto de imágenes de un total de 112 usuarios adquiridos a diferentes distancias y condiciones variables de iluminación, pose/ángulo de la cara, y expresión facial.

Este análisis de escenarios para una biometría tradicional, como es el reconocimiento facial, permite comprender las limitaciones de la tecnología actual y sirve de punto de partida para ver en qué medida la tecnología MMW puede solventar dichas limitaciones.

El artículo se estructura como sigue. La Sec. II describe los conjuntos de datos y escenarios bajo estudio. La Sec. III analiza los principales datos estadísticos de los escenarios. La Sec. IV estudia el rendimiento de los tres sistemas de reconocimiento considerados en los diferentes escenarios. Finalmente la Sec. V muestra los resultados experimentales y esboza el trabajo futuro.

II. DEFINICIÓN DE LOS ESCENARIOS

Los tres escenarios considerados son: 1) "corta" distancia, en la que los hombros pueden estar presentes; 2) "media" distancia, incluyendo la parte superior del cuerpo; y 3) "larga" distancia, incluyendo el cuerpo entero (ver Fig. 1). En esta figura, además de un par de imágenes en cada escenario en el rango visible, se muestran dos imágenes adquiridas en el rango de MMW mediante un radiómetro trabajando a 94 GHz, cortesía de AlfaImaging¹. Una de ellas (a media distancia) se adquirió en exteriores y la otra (a larga distancia) en interiores. Se comprueba el parecido de ambos tipos de imágenes (en visible y en MMW) en los escenarios de media y larga distancia, de ahí el interés en el análisis del reconocimiento biométrico usando imágenes en el visible en estos escenarios,

como base sobre la que comparar en qué medida puede ayudar al reconocimiento clásico la incorporación de imágenes adquiridas en el rango de los GHz.

Usando estos tres escenarios generales, hemos marcado manualmente las 3482 imágenes de los 147 usuarios presentes en la base de datos NISR MBGC v2.0 Face Stills [9]. Algunas se muestran en la Fig. 1. Una porción del conjunto de imágenes fue desechada (360 imágenes de 89 usuarios), porque la cara estaba ocluida o la iluminación degradaba completamente la imagen de la cara. Además, aunque esta información no se usa en el presente artículo, todas las imágenes fueron marcadas como adquiridas en interiores o exteriores.

Finalmente, para posibilitar experimentos de verificación considerando el registro con imágenes adquiridas a corta distancia y realizando experimentos de evaluación en escenarios a corta, media y larga distancia, mantenemos sólo los usuarios con al menos dos imágenes a corta y al menos una imagen en los otros dos escenarios. El proceso de selección de datos está resumido en la Tabla I, donde se puede observar que los tres escenarios considerados resultan en 112 usuarios y 2964 imágenes.

III. ANÁLISIS DE ESCENARIOS: DATOS ESTADÍSTICOS

Primero se segmentan y se localizan las caras (áreas rectangulares) en los tres escenarios usando el software VeriLook SDK (ver Sec. IV-A). Los resultados de la segmentación se muestran en la Tabla II, donde los errores de segmentación crecen significativamente según la distancia cara-cámara aumenta, desde sólo un 1.43% en corta distancia a un 82.57% en larga distancia. Los errores de segmentación en este caso significan que el software VeriLook no puede encontrar la cara en la imagen. Para todas las caras detectadas por VeriLook, llevamos a cabo una comprobación visual, donde se observaron 3 y 10 errores de segmentación para media y larga distancia respectivamente.

Estos errores de segmentación fueron corregidos manualmente marcando los ojos. El área de la cara se estimó basándose en la distancia marcada entre los ojos.

¹<http://www.alfaimaging.com/>

Núm. usuarios	Corta distancia	Media distancia	Larga distancia	Imágenes Rechazadas	Total
147	1539	870	713	360	3482
	<i>Al menos 2 imágenes por usuario</i>	<i>Al menos 1 imagen por usuario</i>			
112	1468	836	660		2964

TABLE I

NÚMERO DE IMÁGENES DE CADA ESCENARIO DEFINIDAS DEL NIST MBGC V2.0 FACE VISIBLE STILLS.

	Corta distancia	Media distancia	Larga distancia	Imágenes Rechazadas	Total
Núm. Imágenes	1468	836	660	360	3324
Errores	21	151	545		848
Errores(%)	1.43%	18.06%	82.57%		

TABLE II

RESULTADOS DE SEGMENTACIÓN BASADOS EN ERRORES PRODUCIDOS POR EL EXTRACTOR DE CARAS DE VERILOOK SDK.

Se comprobó que el tamaño de las caras segmentadas decrece con la distancia de adquisición. En concreto el tamaño promedio de la cara (en píxeles) para cada escenario es: 988×988 para corta, 261×261 para media, y 78×78 para larga distancia.

Otro dato estadístico que calculamos para los tres escenarios es el promedio del índice de calidad proporcionado por VeriLook (0 = mínima, 100 = máxima): 73.93 para corta, 68.77 para media, y 66.50 para larga distancia (ver Fig. 2, calculados sólo para las caras correctamente segmentadas por VeriLook). Como afirman los proveedores de VeriLook, este índice de calidad considera factores como la iluminación, pose y expresión.

IV. ANÁLISIS DE ESCENARIOS: EVALUACIÓN DE RENDIMIENTO DE VERIFICACIÓN

A. Sistemas de Reconocimiento Facial

- **VeriLook SDK.** Sistema de reconocimiento facial comercial desarrollado por Neurotechnology².
- **PCA-SVM system.** Este sistema de verificación usa PCA (Principal Component Analysis). El sistema usa imágenes normalizadas y recortadas, para entrenar un espacio de vectores PCA donde el 96% de la varianza se mantiene. Esto conduce a un sistema donde el espacio original de 5120 dimensiones se reduce a 249 dimensiones. Las comparaciones entre caras se calculan en este espacio de vectores PCA usando un clasificador SVM con núcleo lineal.
- **DCT-GMM system.** Este sistema de verificación usa imágenes de caras divididas en bloques, resultando en 285 bloques por cara segmentada. Se obtiene un vector de características de cada bloque aplicando la Transformada Discreta del Coseno (DCT); de la que sólo los 15 primeros coeficientes se mantienen. Los bloques se usan para derivar un modelo global GMM Ω_w y un modelo adaptado al usuario GMM Ω_c [10]. De experimentos previos se obtuvo

²<http://www.neurotechnology.com/>

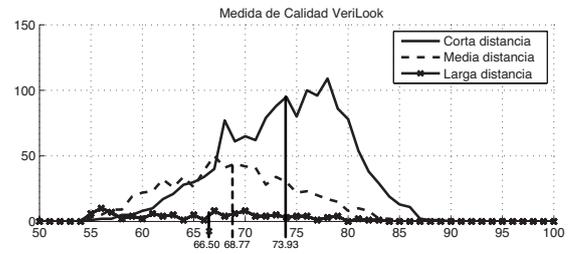


Fig. 2. Histograma de las medidas de calidad producidas por VeriLook SDK.

que usando ($M = 1024$) componentes de mezcla por GMM se obtenían los mejores resultados. El vector de características obtenido con la DCT de cada bloque se compara con ambos Ω_w y Ω_c para producir una puntuación de similitud logarítmica.

B. Protocolo Experimental

- *Close2close.* Este protocolo nos dará una idea sobre el rendimiento de los sistemas en condiciones ideales (ambas: registro y evaluación usando imágenes a corta distancia). Alrededor de la mitad de los datos (754 imágenes) se usa para el ajuste del sistema (entrenamiento del subespacio PCA, SVM, etc.), y el resto (714 imágenes) se usa para evaluar el rendimiento.
- *Close2medium,* y *close2far.* Estos dos protocolos usan como conjunto de datos de ajuste el conjunto completo de datos a corta distancia (1468 imágenes de la cara). Para evaluar el rendimiento de los sistemas, usamos los otros dos conjuntos de datos: 836 imágenes a media distancia para *close2medium,* y 680 imágenes a larga distancia para *close2far.*

C. Resultados

En la Fig. 3 se muestra el rendimiento de verificación para los tres escenarios considerados: *close2close,* *close2medium,* y *close2far.* Primero observamos que VeriLook es el mejor de los tres sistemas en *close2close* con una Equal Error Rate (EER) alrededor de 7%. Al mismo tiempo este sistema comercial es el que más se degrada en condiciones no controladas, con una EER cercana a 40% en *close2far,* mucho peor que los otros dos sistemas más sencillos. Este resultado corrobora la importancia de analizar y tratar adecuadamente los factores de variabilidad que surgen en la biometría a distancia.

Es evidente el empeoramiento que sufre el rendimiento de verificación de los sistemas al pasar del escenario de corta distancia al de media distancia y muy especialmente al pasar al escenario de larga distancia (ver Fig. 3). Una posible mejora del rendimiento en estos dos últimos escenarios (a media y larga distancia) es la fusión de rasgos biométricos adquiridos en el espectro visible (en este caso la cara) con rasgos adquiridos en el rango de MMW (como pueda ser la imagen del contorno del cuerpo, dada la transparencia de la vestimenta a esas frecuencias) [11].

También se observa en la Fig. 3 que el sistema basado en GMM funciona mejor a larga distancia que los otros sistemas,

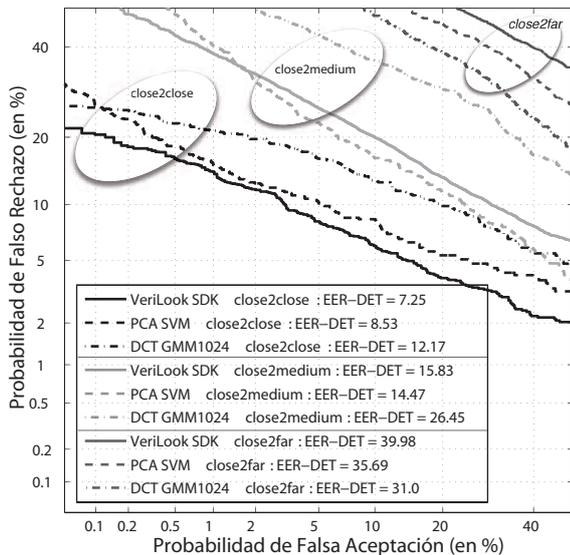


Fig. 3. Resultados del rendimiento de verificación para los tres escenarios y los tres sistemas considerados.

aunque es menos preciso en *close2close* y *close2medium*. Este resultado demuestra el mayor poder de generalización de esta aproximación al reconocimiento, y su robustez frente a condiciones de adquisición no controladas.

V. DISCUSIÓN Y TRABAJO FUTURO

Se ha estudiado el rendimiento de diferentes sistemas de reconocimiento facial que usan imágenes adquiridas en el espectro visible en tres escenarios distintos. Esto constituye la base sobre la que comparar el comportamiento de sistemas que aprovechen la información de imágenes adquiridas en el rango de los GHz. En concreto, en el escenario *close2far* la biometría clásica, en este caso el reconocimiento facial, no proporciona buenos rendimientos, que pueden ser mejorados mediante la fusión con imágenes en MMW. De la misma manera, este trabajo, conforma una aproximación experimental hacia el entendimiento de los factores de variabilidad en reconocimiento facial a distancia. En particular, se ha realizado un análisis de datos de tres escenarios realistas de adquisición a diferentes distancias (corta, media y larga, las dos últimas adecuadas para la adquisición con MMW), como primer paso hacia la creación de un método de reconocimiento adecuado capaz de trabajar en escenarios menos controlados. Este análisis se ha realizado para un subconjunto de datos del NIST MBGC v2.0 Face Stills.

Nuestro análisis se ha centrado en: 1) datos estadísticos (tamaños de las caras segmentadas y medidas de calidad), y 2) evaluación del rendimiento de verificación de los tres sistemas. Los resultados muestran que los sistemas considerados se degradan significativamente en el escenario a larga distancia, siendo la aproximación más simple, la más robusta a condiciones no controladas.

Vale la pena resaltar que los tres escenarios considerados en el presente artículo difieren no sólo en el factor distancia,

sino también en la iluminación y la pose (siendo la variación en la iluminación mucho mayor a larga distancia que a corta distancia). Estos tres factores (distancia, iluminación y pose) son claves en las imágenes adquiridas en la banda de MMW, especialmente en la imágenes adquiridas en exteriores [2], [12]. Basándonos en los datos estadísticos obtenidos y los resultados de la evaluación del rendimiento, el estudio individual de los efectos de los factores mencionados se propone como trabajo futuro para el reconocimiento biométrico en el espectro visible a distancia. En el caso de reconocimiento biométrico mediante el uso de imágenes en MMW, se propone la adquisición de una base de datos de imágenes de usuarios de cuerpo entero, cuya silueta se pueda usar para la extracción de información que pueda ser utilizada en un sistema biométrico o en combinación con otro tipo de información proveniente de otros rasgos biométricos (*Soft Biometrics*) [13], [14], [11], como la facial considerada en este trabajo, de forma que se compensen sus limitaciones, que como hemos observado son muy severas para adquisiciones a cierta distancia.

AGRADECIMIENTOS

Este trabajo ha sido financiado parcialmente por los proyectos Bio-Challenge (TEC2009-11186), Contexts (S2009/TIC-1485), TeraSense (CSD2008-00068) y "Cátedra UAM-Telefónica". M. M.-M. y P. T. están financiados por un contrato CPI de la CAM y una beca FPU, respectivamente.

REFERENCES

- [1] R. Appleby and R.N. Anderton, "Millimeter-wave and submillimeter-wave imaging for security and surveillance," *Proc. of the IEEE*, vol. 95, no. 8, pp. 1683–1690, 2007.
- [2] L. Yujiri et al., "Passive millimeter wave imaging," *Microwave Magazine, IEEE*, vol. 4, no. 3, pp. 39–50, 2003.
- [3] D.M. Sheen et al., "Active millimeter-wave standoff and portal imaging techniques for personnel screening," in *Technologies for Homeland Security, 2009. HST '09. IEEE Conference on*, 2009, pp. 440–447.
- [4] J.N. Mait et al., "94 ghz imager with extended depth of field," *Antennas and Prop., IEEE Trans. on*, vol. 57, no. 6, pp. 1713–19, 2009.
- [5] W. Zhao et al., "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, 2003.
- [6] J.R. Matey et al., "Iris on the move: Acquisition of images for iris recognition in less constrained environments," *Proc. of the IEEE*, vol. 94, no. 11, pp. 1936–1947, 2006.
- [7] Z. Li Stan et al., *Handbook of Remote Biometrics for Surveillance and Security*, chapter Biometrics at a Distance: Issues, Challenges, and Prospects, pp. 3–21, Springer, 2009.
- [8] Robust2008, "Robust biometrics: Understanding science & technology," <http://biometrics.cylab.cmu.edu/ROBUST2008>.
- [9] MBGC, "Multiple biometric grand challenge," *NIST - National Institute of Standard and Technology*, <http://face.nist.gov/mbgc/>.
- [10] J. Galbally et al., "On the vulnerability of face verification systems to hill-climbing attacks," *Pattern Recognition*, vol. 43, no. 3, pp. 1027–1038, 2010.
- [11] J. Fierrez-Aguilar et al., "Discriminative multimodal biometric authentication based on quality measures," *Pattern Recognition*, vol. 38, no. 5, pp. 777–779, 2005.
- [12] D. T. Petkie et al., "Active and passive imaging in the thz spectral region: Phenomenology, dynamic range, modes, and illumination," *J. Opt. Soc. Am. B*, vol. 25, no. 9, pp. 1523–1531, 2008.
- [13] Bo Ye and Yumei Wen, "A new gait recognition method based on body contour," in *Control, Automation, Robotics and Vision*, 2006, pp. 1–6.
- [14] Anil K. Jain et al., "Can soft biometric traits assist user recognition?," 2004, vol. 5404, pp. 561–572, SPIE.