Evaluación de los sistemas biométricos en aplicaciones reales

PID_00215066

Francesc Serratosa

Universitat Oberta de Catalunya

Índice

Introducción

El objetivo final de los sistemas biométricos es aumentar la seguridad en otros sistemas.

Aumentar la seguridad de otros sistemas

Si se aplica un sistema de identificación de personas a través del iris en los aeropuertos es para aumentar la seguridad en un país. En este caso, la seguridad en el país es el sistema en sí al que se desea aumentar la seguridad.

Por este motivo, es imprescindible garantizar que el sistema biométrico en sí sea lo más seguro posible y tener herramientas para evaluar esta seguridad. En este módulo, vamos a explicar las herramientas que se han desarrollado para mostrar la calidad del sistema biométrico.

En el presente módulo, hemos puesto conjuntamente la evaluación de los sistemas biométricos y la exposición de varios sistemas biométricos reales. Esto ha sido así para mostrar la importancia que tiene esta evaluación con los grandes despliegues. Estas aplicaciones biométricas de gran trascendencia social, económica o política con millones de personas involucradas no se podrían llevar a cabo si no fuera por unos bajísimos errores biométricos.

Está claro que se acepta basar la seguridad de un sistema (por ejemplo unas elecciones de un presidente) con un método biométrico si los errores del propio sistema biométrico son muy inferiores a otros métodos clásicos no automatizados o semiautomatizados.

Objetivos

Los objetivos de este módulo son explicar los errores que pueden aparecer en un sistema biométrico así como explicar los mecanismos que se han planteado para evaluar la bondad de los sistemas biométricos. También deseamos mostrar unos cuantos ejemplos de aplicaciones biométricas reales y a gran escala para que el estudiante sea capaz de:

1. Clasificar los errores que pueden aparecer en un sistema biométrico y determinar en qué condiciones aparecen estos errores.

2. Evaluar una aplicación biométrica para saber su bondad. Conocer las métricas para evaluar y comparar la bondad de los sistemas biométricos.

3. Conocer unos cuantos ejemplos de grandes sistemas biométricos.

1. Errores de los sistemas biométricos

La promesa con la que se fundamenta el reconocimiento biométrico es que, dada una nueva muestra, el sistema biométrico ofrece siempre la decisión correcta, tanto sea en verificación (es o no es la persona) como en identificación (devuelve la identificación de la persona). En la práctica, un sistema biométrico es un sistema de reconocimiento de patrones que inevitablemente toma decisiones incorrectas. Por ello, es fundamental entender por qué un sistema biométrico comete errores y modelar esos errores para poder conocer su magnitud.

Referencia bibliográfica

El tratamiento comprensivo del sistemas biométricos está detallado en la ISO/IEC 19795-2, 2007.

1.1. Razones de los errores de los sistemas biométricos

Las razones de los errores de los sistemas biométricos se deben a una serie de limitaciones, que se detallan a continuación:

- **Limitación de la información**: La información invariante y distintiva contenida en una muestra biométrica está inherentemente limitada debido a la capacidad intrínseca de la señal del identificador o sensor biométrico. Por ejemplo, la información distintiva en la geometría de la mano es menor que en los dedos. En consecuencia, las muestras de la geometría de la mano pueden distinguir menos identificaciones que los dedos aunque sea en condiciones ideales. La limitación de la información también puede provenir de una presentación pobre del rasgo biométrico al sensor por parte de los usuarios o de la adquisición de la señal inconsistente. Muestras adquiridas de forma diferente de un rasgo biométrico limitan la invariancia a lo largo de diferentes muestras del mismo usuario.

- **Limitación de la representación**: La representación ideal se tendría que diseñar de tal modo que retuviera toda la invariancia así como la información discriminatoria de las muestras tomadas. Los módulos actuales de extracción de características, típicamente basadas en modelos simplistas de la señal biométrica, fallan al capturar toda la riqueza de la información en una señal biométrica real. Así, se incluyen características erróneas y se excluyen características verdaderas. En consecuencia, una fracción del espacio legítimo de las muestras no puede ser representado por el sistema biométrico, y así aparecen errores de representación.

- **Limitación en la invariancia**: Finalmente, dado un esquema representativo, el diseño de un comparador ideal tendría que modelar a la perfección la relación de invariancia a lo largo de diferentes muestras del mismo usuario (misma identificación), aunque las muestras hayan sido adquiridas en diferentes condiciones. Otra vez, en la práctica (debido a la incapacidad de adquirir un número suficientemente grande de muestras o la varianza

en las condiciones de la captura de las muestras), el comparador puede ser que no modele las relaciones de invariancia y así aparecen errores en el comparador.

El reto es poder llegar a una representación realista e invariante del rasgo biométrico proveniente de muestras nuevas bajo condiciones no controladas (o casi no controladas). Entonces, hay que estimar formalmente la información discriminatoria en la señal de las muestras. Esta tarea es realmente difícil en un sistema de identificación a gran escala donde el número de usuarios matriculados puede ser enorme (más de 50 millones).

1.2. Tipos de errores en los sistemas biométricos

En este apartado, se detallan los errores que aparecen en las diferentes etapas de un sistema biométrico.

Ved también

Ved la figura 10 del módulo "La biometría para la identificación de las personas" de esta asignatura.

1) Errores en el módulo de captura: En los sistemas completamente automáticos, los datos se capturan sin supervisión de ningún experto. Estos sistemas biométricos suelen usar un dispositivo del tipo *live-scan* que detecta automáticamente la presencia de un rasgo biométrico cuando aparece en el punto de mira del dispositivo. Estos tipos de captura pueden producir dos tipos de errores, que son el **fallo en la detección (FD)** y el **fallo en la captura (FC)**. El fallo en la detección aparece cuando el rasgo biométrico se acerca al sensor pero el sensor no es capaz de detectar su presencia. El fallo en la captura aparece cuando el sistema se ha dado cuenta de que hay un rasgo biométrico pero no puede capturar la muestra. Normalmente, la razón entre estos dos fallos es inversamente proporcional.

2) Errores en el extractor de características: Después de capturar la muestra, el sistema la envía al extractor de características. Si la imagen tiene muy poca calidad (atención, en el caso de la voz sería la señal acústica), el extractor no es capaz de extraer ninguna característica coherente. Este error se conoce como **fallo de proceso (FP)**. Debido a que el módulo de captura y el extractor de características se usan en los tres procesos básicos (matrícula, verificación e identificación), normalmente se juntan en una sola medida llamada **fallo de adquisición (FA)**. Un porcentaje de fallos de adquisición alto respecto al número de veces que se han adquirido muestras provoca una bajada importante en el rendimiento del sistema y también la frustración de los usuarios, que a su vez genera rechazo al sistema biométrico (reducción de aceptabilidad). Una forma de disminuir este porcentaje de fallos de adquisición es permitir que el sistema genere un conjunto de características aunque la imagen sea mala, es decir, que la calidad de estas características sea baja. El problema está en que, entonces, el módulo de comparación sufre una carga adicional y las comparaciones pueden devolver salidas erróneas.

3) Errores en el módulo de creación de la plantilla: El módulo de creación de la plantilla también puede fallar, dado un conjunto de características de diferentes muestras. Estos fallos aparecen cuando las características se han extraído en una situación muy ruidosa y por lo tanto hay poca coherencia entre las muestras. Este fallo se denomina **fallo de matriculación (FM)**, puesto que la plantilla solo se genera en el proceso de matriculación. Parecido al fallo de adquisición, si el fallo de matriculación se desactiva o se ponen unos límites de calidad muy bajos, entonces aparecen muchos más errores en la comparación.

4) Errores en el módulo de comparación: El módulo de comparación genera un resultado dada una muestra y una plantilla. Este resultado suele tener un valor dentro del rango de 0 a 1 y representa una probabilidad o una distancia. Los posibles fallos del módulo de comparación dependen de si nos encontramos en un proceso de verificación o identificación. Pasemos a describirlos.

Verificación

En un **proceso de verificación**, tras calcular la distancia o probabilidad, se aplica un umbral, modificable externamente, para tomar una decisión final. Si la distancia (o probabilidad) es inferior (o superior) al umbral, entonces se considera que la plantilla y la muestra provienen del mismo individuo. De lo contrario, se considera que son de diferentes individuos. En este proceso, nos encontramos ante cuatro combinaciones, dos de las cuales generan errores:

a) El usuario se identifica correctamente y presenta al sistema sus rasgos biométricos:

- El sistema devuelve correctamente que hay etiquetado, es decir, que los rasgos biométricos pertenecen a la identificación presentada. No hay error y se denomina **aceptación correcta**[1] **(AC)**.

(1)En inglés, *correct acceptance.*

- El sistema devuelve erróneamente que los rasgos biométricos no son de la persona con la identificación presentada al sistema. Este es un **falso rechazo** o **error de no etiquetado**[2] **(FR)**.

(2)En inglés, *false non-match* o *false rejection.*

b) El usuario se identifica de forma fraudulenta (por ejemplo, introduce la identificación de otra persona que sabe que tiene unos permisos especiales) y presenta al sistema sus rasgos biométricos:

- El sistema devuelve correctamente que los rasgos biométricos pertenecen a otra identificación. No hay error y se denomina **rechazo correcto**[3] **(RC)**.

(3)En inglés, *correct rejection*

- El sistema devuelve erróneamente que sí, que los rasgos biométricos son de la persona con la identificación presentada al sistema. Este es un error de **falsa aceptación** o **falso etiquetado**[4] **(FA)**.

(4)En inglés, *false match* o *false acceptance.*

Identificación

En un **proceso de identificación**, a la hora de definir los errores, nos encontramos ante seis combinaciones, tres de las cuales generan errores y una no es posible.

a) La persona cuyos rasgos biométricos se buscan se había matriculado en el sistema (su plantilla está en la base de datos):

- El sistema devuelve la identificación de la persona cuya busca se hace. No hay error: **aceptación correcta (AC)**.

- El sistema devuelve otra identificación: **error de identificación positivo**[5] **(FPI)**. Lo que ha pasado es que hay una plantilla de otra persona que por error ha devuelto una distancia menor que la plantilla correcta.

(5)En inglés, *false positive identification*.

- El sistema devuelve que no hay ninguna plantilla con esos rasgos biométricos: **error de rechazo**[6] **(FR)**. Este caso solo puede aparecer cuando el sistema de identificación dispone de un umbral, como en la verificación. Podría ser que este error desapareciera si el umbral de la distancia fuera menos restrictivo, es decir, si aumentamos su valor.

(6)En inglés, *false rejection*.

b) La persona cuyos rasgos biométricos se buscan no se había matriculado en el sistema (su plantilla no está en la base de datos):

- El sistema devuelve la identificación de la persona cuya busca se hace. Esta combinación no es posible. Si la identificación no ha sido nunca entrada porque el usuario no se ha matriculado, entonces no puede devolver nunca su identificación.

- El sistema devuelve otra identificación: **error de identificación negativo**[7] **(FNI)**.

(7)En inglés, *false negative identification*.

- El sistema devuelve que no hay ninguna plantilla con estos rasgos biométricos: **rechazo correcto (CR)**. Parecido al error de rechazo, solo puede aparecer esta situación si hay un umbral de aceptación. Ninguna plantilla ha devuelto una distancia menor al umbral de aceptación. Si se aumentara el umbral de aceptación para intentar que desaparezcan los errores de rechazo, entonces nos podríamos encontrar con que algunos rechazos correctos desaparecerían.

1.3. Modelización de los errores

En el apartado anterior, hemos detallado qué tipo de errores pueden aparecer en un sistema biométrico así como las causas que los generan. A continuación, vamos a llevar a cabo un estudio más científico de estos errores en el supuesto de que nos encontremos en un sistema de verificación y en un sistema de identificación.

1.3.1. Sistema de verificación

Supongamos que la plantilla almacenada en la base de datos de una persona es T y la muestra que deseamos verificar es I. Además, supongamos que tenemos una función de similitud (se define como la inversa de una distancia) entre una muestra y una plantilla $S(I,T)$. S toma valores dentro del rango de 0 a 1. Cuanto mayor es S, más se parece la muestra a la plantilla, es decir, más probabilidad hay de que pertenezcan a la misma persona. Entonces tenemos dos posibles hipótesis:

- $\boldsymbol{H_0}$: $I \neq T$: La muestra que queremos verificar no pertenece a la misma persona con la que se ha generado la plantilla.

- $\boldsymbol{H_1}$: $I = T$: La muestra que queremos verificar es de la misma persona con la que se ha generado la plantilla.

Las posibles respuestas del sistema biométrico son las siguientes:

- $\boldsymbol{D_0}$: No hay etiquetado. El sistema considera que pertenecen a personas diferentes.

- $\boldsymbol{D_1}$: Hay etiquetado. El sistema considera que provienen de la misma persona.

Considerando las hipótesis y las salidas del sistema, nos encontramos con los errores que ya hemos mencionado:

- **Falsa aceptación (FA)**: También llamado *error de tipo I*. El sistema devuelve D_1 cuando la hipótesis era H_0.

- **Falso rechazo (FR)**: También llamado *error de tipo II*. El sistema devuelve D_0 cuando la hipótesis era H_1.

La **razón de error de etiquetado (*FMR* o FAR[8])** es la probabilidad de un error de tipo I. Matemáticamente:

$$FMR = \text{probabilidad } (D_1|H_0).$$

La **razón de error de no etiquetado (*FNMR* o FRR[9])** es la probabilidad de un error de tipo II. Matemáticamente:

$$FNMR = \text{probabilidad } (D_0|H_1).$$

[8]Acrónimo del inglés *false match rate* o *false acceptance rate.*

[9]Acrónimo del inglés *false non-match rate* o *false rejection rate.*

Para poder evaluar la precisión de un sistema de verificación biométrico es necesario recoger un número muy elevado de comparaciones entre muestras y plantillas de la misma persona y también un número muy elevado de comparaciones entre muestras y plantillas de diferentes personas. El conjunto de las primeras muestras se denomina **distribución genuina** y matemáticamente se representa con la distribución $p(s|H_1)$. El conjunto de las segundas muestras se denomina **distribución impostora** y matemáticamente se representa con la distribución $p(s|H_0)$. De este modo, podemos definir las razones de los errores con las siguientes funciones:

$$FNMR = \int_0^t p(s|H_1)ds \qquad 2.1$$

y

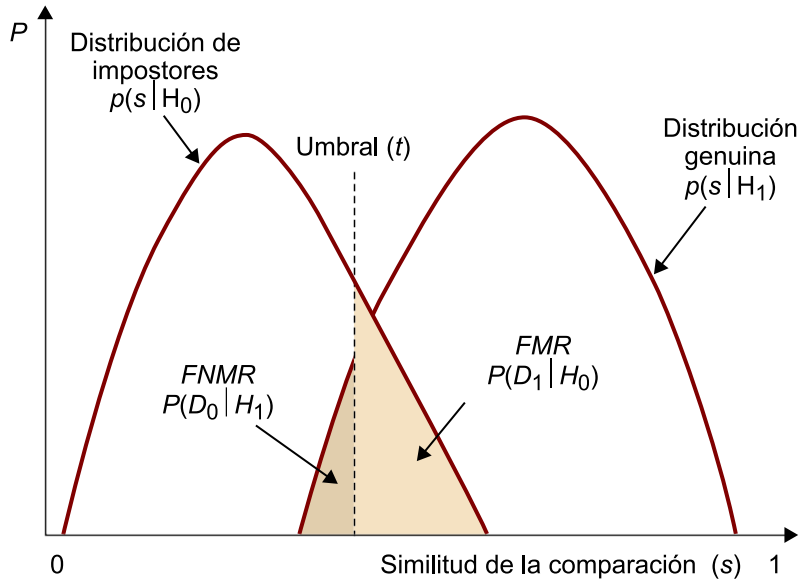
$$FMR = \int_t^1 p(s|H_0)ds \qquad 2.2$$

donde t es el umbral de aceptación impuesto por el administrador del sistema.

El anexo muestra un método práctico para calcular estas probabilidades.

La figura 1 muestra las distribuciones impostoras y genuinas respecto al valor de la similitud de la comparación (*matching score*). En un sistema real, las muestras que pertenecen a la distribución genuina suelen tener una similitud mayor (o una distancia menor, más a la derecha en la figura) que las muestras que pertenecen a la distribución impostora (con una distancia mayor, más a la izquierda de la figura).

Figura 1. Distribución de las poblaciones impostoras y las genuinas respecto a la similitud



Hay una disyuntiva muy importante entre la *FMR* (o también llamada *false acceptance rate*, *FAR*) y la *FNMR* (o también llamada *false rejection rate*, *FRR*) en cada sistema biométrico. De hecho, tal y como se ve en las fórmulas, las dos dependen del umbral de aceptación *t*, que es una entrada en el módulo de comparación (ver la figura 10 del módulo "La biometría para la identificación de las personas"). Por eso, en realidad, tendríamos que escribir *FMR*(*t*) y *FNMR*(*t*). Como se ve en la figura, la *FNMR*(*t*) (o *FRR*(*t*)) es el área marcada por la distribución genuina y el umbral *t*. Y la *FMR*(*t*) (o *FAR*(*t*)) es el área marcada por la distribución impostora y también por el umbral *t*. Si se reduce *t* para volver al sistema más tolerante a las variaciones de entrada y al ruido, entonces la *FMR*(*t*) aumenta. Por otro lado, si aumentamos *t* para hacer al sistema más seguro, entonces aumentamos la *FNMR*(*t*). El administrador del sistema no puede saber anticipadamente dónde se va a desplegar el sistema ni qué respuesta tendrán sus usuarios. Por este motivo, es difícil imponer inicialmente el umbral *t*. Para poder mostrar la bonanza de un sistema de verificación independientemente del umbral se han definido las dos funciones siguientes:

- ***Receiver operating characteristic* (ROC)**: La ROC es una curva en un plano bidimensional marcada por los puntos *FMR*(*t*) y 1 – *FNMR*(*t*) por varios valores de *t*. El valor 1 – *FNMR*(*t*) se denomina la bondad del test. Esta gráfica muestra la *FMR* respecto a la bondad del test.

- ***Detection-error trade-off* (DEC)**: La DEC es una curva parecida a la ROC pero marcada por los puntos *FMR*(*t*) y *FNMR*(*t*). La DEC es interesante para mostrar la relación entre los dos tipos de errores puesto que el objetivo es rebajar al máximo ambos errores. La figura 3c muestra un ejemplo de DEC.

Además de estas gráficas, cuando queremos analizar un sistema biométrico, se suelen dar cuatro índices globales. Cuando recibimos información de un sistema biométrico por parte de una persona de la empresa, es importante usar

o considerar estos índices con cautela puesto que por lo general se han llevado a cabo científicamente, pero con bases de datos controladas por los propios desarrolladores del sistema. Los índices son los siguientes:

- **Razón de error equivalente (EER[10])**: Indica la razón de error para todos los valores del umbral donde la FMR es igual a la *FNMR*:

  $$EER = FMR(t) \text{ tal que } FMR(t) = FNMR(t) \text{ para todo } t.$$

(10)Acrónimo del inglés *equal-error rate*.

- **Cero FNMR**: Se define como el valor menor de la *FMR* en el que no hay errores de etiquetado, FNMR = ∅ (o también FRR = 0).

- **Cero FMR**: Se define como el valor menor de la *FNMR* en el que no hay errores de no etiquetado, FAR = ∅ (o también FAR = 0).

- **Separabilidad**: Si asumimos que las poblaciones genuina e impostora generan distribuciones normales (distribuciones gaussianas), entonces podemos analizar lo muy separadas que están o, dicho de otro modo, el poco solapamiento que encontramos entre ambas poblaciones. Cuanto más solapamiento, más errores se generarán en el proceso de reconocimiento.

$$S = \frac{\| \hat{x}_{impostores} - \hat{x}_{genuinos} \|}{\sqrt{\frac{\sigma^2_{impostores} + \sigma^2_{genuinos}}{2}}} \qquad 2.3$$

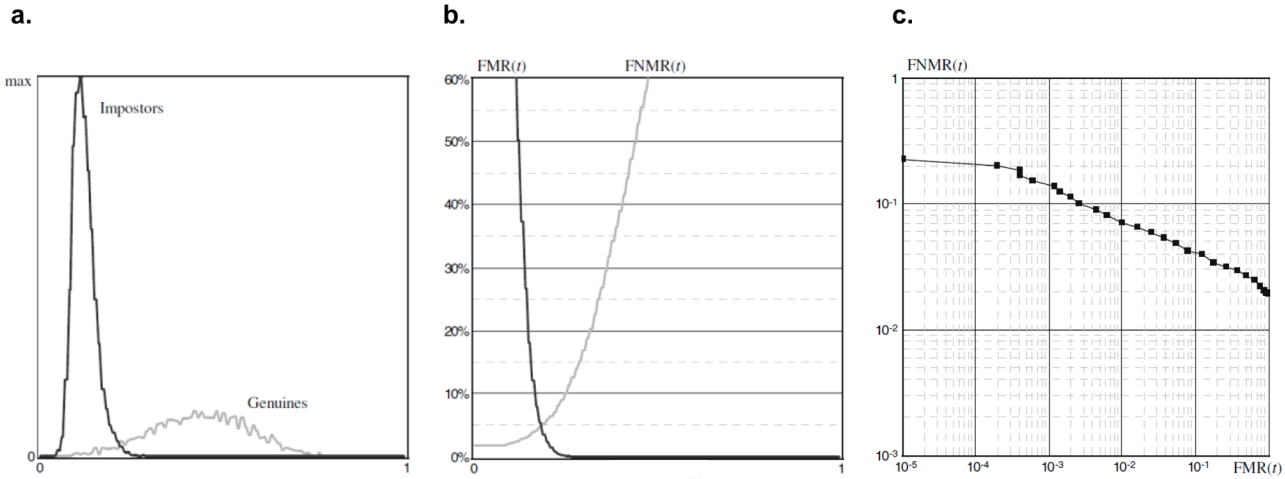
donde $\hat{x}$ representa la media y σ representa la desviación estándar de la población impostora o genuina.

La figura 2 muestra los resultados de un algoritmo de comparación de huellas dactilares presentado en la Fingerprint Verification Competition (FVC) del año 2002. Los datos que se muestran se calcularon con 2.800 parejas de huellas dactilares genuinas (pertenecían al mismo dedo) y 4.950 parejas impostoras (pertenecían a diferentes individuos).

FVC

La **FVC** es una competición a la que empresas y centros de investigación pueden enviar sus algoritmos ya compilados (no se envía el código fuente) para comprobar su funcionamiento y su bondad.

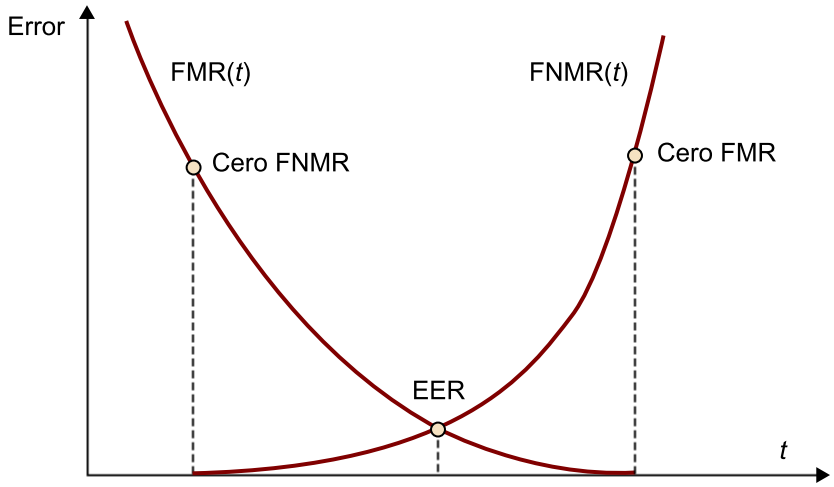
Figura 2. Algunos resultados del algoritmo de comparación FVC



a. Distribuciones genuinas e impostoras respecto al umbral de similitud; b. evolución del porcentaje de error de los errores FMR y FNMR respecto al umbral; c. la curva DEC obtenida con los mismos datos con los que se ha computado la figura b

La figura 3 muestra el porcentaje de los errores FMR y FNMR respecto al umbral. También se muestra el punto donde se define la EER así como los valores cero FNMR y cero FMR. Fijaos en que el valor cero FNMR (o cero FMR) se ubica exactamente en el punto donde la gráfica del FNMR (o FMR) pasa a valer 0.

Figura 3. Ejemplo de obtención de los valores globales cero FNMR, cero FMR y EER con las curvas de la razón de los errores respecto al umbral de similitud



Los requerimientos de precisión de un sistema de verificación biométrico dependen mucho de la aplicación.

En aplicaciones forenses tales como la identificación de criminales, lo que deseamos es no dejar de identificar a un criminal aunque exista el riesgo de tener que examinar manualmente gran cantidad de potenciales falsas aceptaciones identificadas por el sistema. Esto implica que lo que nos preocupa es que la *FNMR* sea alta y, por lo tanto, pondremos un umbral de similitud bajo. Otro extremo sería un control de acceso de alta seguridad. El principal objetivo es que no entren impostores. En tal caso, nos preocupa que la *FMR* sea alta. Claramente, si imponemos que el umbral de similitud sea muy alto, entonces

haremos bajar la *FMR* pero esto implicará que algunas veces habrá personas autorizadas a quienes no se permitirá el acceso. Visualizando la figura 3 se ve muy claro este concepto.

1.3.2. Sistema de identificación

Supongamos que una muestra que se quiere identificar se compara con *N* plantillas de la base de datos y supongamos también que estas comparaciones son independientes entre ellas. En los sistemas de identificación, como ya hemos visto, se definen tres tipos de errores que están relacionados con los errores de etiquetado y de no etiquetado:

- **Error de identificación positivo (FPI)**: Se relaciona directamente con el error de no etiquetado. Tenemos la muestra correcta y la plantilla existe, pero el sistema no es capaz de encontrar el etiquetado correcto y retorna una identificación incorrecta.

- **Error de rechazo (FR)**: Es parecido al error de identificación positivo, por lo tanto, también se relaciona directamente con el error de no etiquetado, ya que tenemos la muestra correcta y la plantilla existe pero el comparador devuelve una similitud inferior al umbral (o una distancia mayor al umbral). Por lo tanto, decide que no existe aquella persona.

- **Error de identificación negativo (FNI)**: Se relaciona con el error de etiquetado. En este caso, no tenemos la plantilla correcta y el sistema devuelve un etiquetado incorrecto y la similitud es mayor que el umbral de similitud.

Dados estos errores, podemos definir las dos probabilidades siguientes:

La **razón de error de identificación positivo (*FPIR*[11])** engloba la razón de error de los dos primeros errores mencionados, el **error de identificación positivo** y el **errorde rechazo**. Este error depende del número de plantillas *N* con los que se ha comparado y se aproxima a la probabilidad de error de etiquetado (FMR) calculado para esta base de datos (ecuación 2.2). Se define del modo siguiente:

(11)Acrónimo del inglés *false positive identification-error rate*.

$$FPIR_{\mathrm{N}} = 1 - (1 - FMR)^{\mathrm{N}}$$

Estos errores aparecen cuando la muestra se etiqueta erróneamente con una o más plantillas de la base de datos. Por eso, la $FPIR_N$ se calcula como uno menos la probabilidad de que no se haga ningún etiquetado falso con ninguna de las plantillas. La expresión (1 – *FMR*) es la probabilidad de que la muestra no se etiquete falsamente con una de las plantillas de la base de datos. Si la *FMR* es muy pequeña, entonces esta expresión general se puede aproximar por $FPIR_N \approx N \cdot FMR$. Y de este modo podemos establecer que la probabilidad de

los errores de etiquetado positivos se incrementa linealmente con el tamaño de la base de datos. Esta aproximación se basa en solo considerar el primer término del binomio de Newton. Si deseáramos una aproximación más acertada, podríamos usar los dos primeros términos del binomio de Newton y la expresión quedaría así:

$$FPIR_N \approx N \cdot FMR - \frac{N \cdot (N-1)}{2} \cdot FMR^2 \qquad 2.4$$

Con esta segunda aproximación, el valor que se obtiene es un poco menor puesto que tiene un término más restando.

La **razón de error de identificación negativo (*FNIR*[12])** es más sencilla de calcular puesto que se considera exactamente igual a la probabilidad de error de no etiquetado (*FNMR*) calculado en esta base de datos; *FNIR* = *FNMR*. Esto se debe a que la probabilidad de un error de identificación negativo cuando se está buscando la plantilla en las *N* plantillas de la base de datos es el mismo que la *FNMR* en el modo de verificación.

(12)Acrónimo del inglés *false negative identification-error rate*.

2. Evaluación de un sistema biométrico

La bondad de un sistema biométrico depende drásticamente de muchas variables: la composición de la población (ocupación, edad, sexo, demografía, raza, entre otros), el entorno, el modo de hacer las pruebas así como otras restricciones específicas de la aplicación. En una situación ideal, se querría caracterizar el rendimiento en un modelo independiente a la aplicación. Así, se podría predecir el rendimiento en una aplicación real.

Se han aplicado técnicas de modelado rigurosas para caracterizar la adquisición de los datos y el proceso de comparación. Con estas técnicas, se ha logrado extrapolar los resultados obtenidos en el laboratorio como si fueran aplicaciones reales obteniendo bastantes buenos resultados. Hoy en día, se están realizando evaluaciones comparativas en bases de datos reducidas. Los ejemplos más claros son la Fingerprint Verification Competition (FVC) (ya tratada) y la Iris Verification Competition (IVC). De los resultados obtenidos en estas competiciones puede depender que un sistema se pueda llevar al mercado comercial o no. Debido a la importancia de poder evaluar la precisión de los sistemas biométricos, se pueden definir tres tipos de evaluaciones:

1) Evaluación de la tecnología: El objetivo es evaluar la calidad de los algoritmos dada una tecnología específica. No se evalúa todo el sistema sino algoritmo a algoritmo. Todos los algoritmos se comparan dados los mismos sensores, base de datos y cualquier aspecto que pueda afectar a los resultados. La base de datos se divide en dos partes. Normalmente, todos los datos se generan a la vez y la partición se lleva a cabo de una manera aleatoria. La primera parte compone la **base de datos de aprendizaje** (***learning database***). Forma la parte de los datos que los usuarios pueden usar para poder hacer la puesta a punto del algoritmo y extraer el máximo rendimiento. La segunda parte compone la **base de datos de test** (***test database***). Forma la parte de los datos que los evaluadores usan para hacer las pruebas finales. Los participantes no han podido usarla ni visualizarla antes de las pruebas. Debido a que los datos quedan disponibles para toda la comunidad científica, tras los experimentos se pueden repetir. Algunos libros de biometría incorporan DVD con estos datos. Es una **evaluación repetible**.

Web recomendada

Acceded a la web de la Fingerprint Verification Competition y consultad las diferentes competiciones. https://biolab.csr.unibo.it/FVCOnGoing/UI/Form/Home.aspx

2) Evaluación del escenario: El objetivo de este tipo de evaluación es determinar el rendimiento completo de todo el sistema en un prototipo de laboratorio o en un simulador de aplicaciones. El test se lleva a cabo en un sistema completo, pero en unas condiciones controladas aunque intenta simular una situación del mundo real. La comparación siempre se lleva a cabo con los mismos sensores biométricos y la misma población. Es una **evaluación repetible**.

3) Evaluación del funcionamiento: El objetivo de esta evaluación es determinar el rendimiento del sistema completo en una situación real de entorno específico y una población específica. Es una **evaluación no repetible** debido a que puede haber parámetros no documentados o desconocidos. No hay una base de datos inicial.

3. Primeras grandes aplicaciones reales

La lista siguiente muestra unos cuantos ejemplos de aplicaciones desarrolladas a gran escala. No pretende ser una lista exhaustiva sino unos cuantos ejemplos para mostrar que la biometría se está aplicando y ya hace un tiempo que se aplica en problemas reales y en todo el mundo. Además, se han seleccionado las aplicaciones que no están relacionadas con el control de acceso ni seguridad, que son a las que estamos más habituados.

1) Sudáfrica. Verificación de huellas. La primera aplicación a gran escala de la biometría utilizando huellas digitales fue la distribución de pensiones en áreas rurales de Sudáfrica. En 1990, cada pensionista tenía sus huellas digitales registradas. Se almacenaron en una tarjeta personal y se verificaron antes de entregar la pensión para garantizar que quien llevaba la tarjeta era el propietario de ella. Este sistema redujo el fraude considerablemente.

2) México. Verificación de huellas. El Instituto Electoral Federal instaló 2.000 aparatos de control biométrico para la verificación de las tarjetas de identidad de los votantes. El propósito no era identificar al votante por el nombre, sino comprobar que la persona, a pesar del nombre, tuviera derecho a votar. En apariencia, la operación fue un éxito.

3) Uganda. Verificación de caras. Para luchar contra el fraude electoral, el presidente de Uganda decidió tener un sistema de reconocimiento de caras instalado en los centros electorales para las elecciones generales de junio del 2001. En dos meses, once millones de votantes fueron fotografiados para crear una base de datos solo para las votaciones. Este sistema se encuentra todavía en vigor.

4) Malasia. Verificación de huellas. Desde el 2001, a cada habitante de Malasia de más de 12 años le ha sido expedida una tarjeta de identificación biométrica, la MyKad, que contiene datos como la fecha y el lugar de nacimiento, el sexo, el nombre de los padres, la etnia de origen, la religión, una fotografía y las huellas. Es una tarjeta con múltiples fines, puesto que sirve como carné de conducir, pasaporte, tarjeta de pago electrónico y también contiene información médica de emergencia.

Web recomendada

Ved Mykad en la Wikipedia.

5) Afganistán. Verificación del iris. El Alto Comisionado para los Refugiados (HCR) utilizó la biometría en el 2003 para ayudar a las familias afganas a volver a su país después de una larga estancia en Pakistán. El personal del HCR fotografió el iris de las posibles personas que volverían. Cuando estuvieron preparadas para volver, su identificación se hizo antes de que les dieran el di-

nero para el transporte, los cupones de la comida y las necesidades básicas. La agencia sostiene que de este modo se ahorró millones de dólares al prevenir fraudes de identidad.

6) Australia. Verificación del iris. En el año 2003, se probó en Australia un sistema biométrico para dispensar metadona. Se fotografió el iris de los drogadictos que eligieron participar en el programa para su identificación posterior en las farmacias que participaban en el estudio, donde los pacientes recibirían la dosis exacta prescrita por el médico. La utilización de este sistema de control es particularmente útil en las farmacias grandes, donde los farmacéuticos no conocen a todos los pacientes.

7) Europa. Verificación de huellas. En un intento de reducir las múltiples solicitudes de asilo político en los diferentes países de Europa, la Comunidad Europea ha establecido una base de datos centralizada, la Eurodac, que contiene las huellas de toda persona que pide asilo. Antes de enero del 2003, cuando el sistema se volvió operativo, se estimaba que un 80% de las 500.000 solicitudes anuales se pedían a varios países, mientras que ahora se ha reducido a un 11% de las 280.000 solicitudes. La base de datos Eurodac no se puede emparejar con otras bases de datos.

Web recomendada

Ved Eurodac en la Wikipedia.

8) La India. Verificación de huellas. Algunas ceremonias religiosas en el templo de Tirumala en la India pueden atraer a 150.000 peregrinos al día. Las autoridades han adoptado un sistema biométrico para facilitar la gestión de la multitud. Los peregrinos registran sus huellas dactilares con antelación. El día de las ceremonias, se identifican por las huellas y se les autoriza a entrar al templo.

9) Japón. Verificación de la cara, de las huellas. En Kioto, se llevó a cabo un experimento para permitir a las personas mayores beneficiarse de los servicios sociales sin salir de sus hogares. Las personas inscritas conectan con trabajadores del Ayuntamiento y se identifican mostrando la cara en una webcam y poniendo el dedo en un sensor de huellas.

10) Indonesia, Tailandia. Verificación del ADN, huellas. Para identificar los cuerpos después del tsunami de diciembre del 2004, los expertos recopilaron muestras de ADN que después se compararon con el ADN de las familias que buscaban a un familiar desaparecido y rastros de huellas digitales con las huellas de personas desaparecidas que las tenían registradas en los documentos de identificación.

11) Texas, EE. UU. Verificación de huellas. Medicaid es un programa estadounidense que da prestaciones sanitarias a las personas con ingresos bajos. Para ofrecer una mejor protección de la información médica de los pacientes y reducir el fraude, el estado de Texas ha estado probando un proyecto piloto

Web recomendada

Ved Medicaid en la Wikipedia.

desde el año 2004. Las huellas digitales de las personas aptas para Medicaid se almacenan en su tarjeta Medicaid para comprobarlas antes de recibir las prestaciones.

Resumen

En este módulo, hemos descrito cómo evaluar los sistemas biométricos. Antes de nada, hemos estudiado qué errores pueden aparecen en los sistemas biométricos. Después, hemos visto que hay métricas globales y que hay métricas que son gráficas como la ROC o la DEC. La evaluación de los sistemas biométricos es fundamental puesto que siempre tenemos que ver la biometría como una ciencia muy aplicada.

También hemos descrito ejemplos de grandes despliegues. Estos ejemplos se tienen que ver simplemente como una muestra de las posibilidades que tiene la biometría para el reconocimiento de las personas. Otros sistemas se han puesto en funcionamiento y surgen algunos nuevos.

Actividades

1. Las razones de los errores en los sistemas biométricos se clasifican en tres tipos diferentes de limitaciones. Describidlas y resumidlas.

2. Dadas las limitaciones anteriores, aparecen cuatro tipos de errores en los módulos (o procesos) de un sistema biométrico. Describid los módulos así como los posibles errores.

3. Los sistemas biométricos disponen de un umbral (normalmente no es visible para el usuario ni siquiera para el administrador del sistema) a partir del cual se considera que, dada una comparación, las dos muestras provienen del mismo individuo o no. Relacionad este umbral con los errores generales de los sistemas biométricos de verificación, el error de etiquetado y el error de no etiquetado.

4. Haced el mismo ejercicio que en el punto 3 pero para los sistemas biométricos de identificación con los errores: error de etiquetado positivo, error de rechazo y error de identificación negativo.

5. Explicad qué es la distribución genuina y qué es la distribución impostora y cómo se modelan matemáticamente.

6. Dados los valores de similitud genuinos {3, 3, 5, 5, 6, 6, 6, 7, 9} y los valores de similitud impostores {1, 2, 2, 3, 3, 3, 4, 4, 5}, dibujad la función de distribución de la *FNMR* y de la *FMR* (figura 1). ¿Cuál es el valor del umbral que hace que el error general (*FNMR* + *FMR*) sea mínimo? ¿Cuál es el valor del umbral mínimo para que la *FNMR* = 0? En este umbral, ¿qué valor toma la *FMR*? ¿Cuál es el valor del umbral máximo para que la *FMR* = 0? En este umbral, ¿qué valor toma la *FNMR*?

7. Dibujad la ROC y la DEC con las distribuciones de población del ejercicio 6. Suponed los diez umbrales siguientes {0,5; 1,5; 2,5; 3,5; 4,5; 5,5; 6,5; 7,5; 8,5; 9,5}.

8. Dadas las poblaciones de los dos ejercicios anteriores, decid cuál es la EER, el cero FNMR, el cero FMR y la separabilidad de las poblaciones.

9. Dibujad la figura de la DEC y de la ROC dada la tabla 1 del anexo.

10. Describid los tres tipos de evaluaciones: tecnología, escenario y funcionamiento.

11. Buscad aplicaciones reales donde se apliquen técnicas biométricas y agrupadlas por rasgos biométricos específicos.

Abreviaturas

AC aceptación correcta

DEC *detection-error trade-off*

EER *equal-error rate* (razón de error equivalente)

FA falsa aceptación

FC fallo de captura

FD fallo de detección

FM fallo de matriculación

FMR *false match rate* (razón de error de etiquetado)

FNI error de identificación negativo

FNIR *false negative identification-error rate* (razón de error de identificación negativo)

FNMR *false non-match rate* (razón de error de no etiquetado)

FP fallo de proceso

FPI error de identificación positivo

FPIR *false positive identification-error rate* (razón de error de identificación positivo)

FR falso rechazo

RC rechazo correcto

ROC receiver operaring characteristic

Bibliografía

Jain, Anil; Bolle, Ruud; Pankanti, Sharath (1999). *Biometrics. Personal identification in networked society*. Kluweer Academic Publishers.

Jain, Anil; Flynn, Patrick; Ros, Arun (ed.) (2008). *Handbook of biometrics*. Springer.

Nanavati, Samir; Thieme, Michael; Nanavati, Raj (2002). *Biometrics. Identity verification in a networked world*. Wiley Computer Publishing.

Ross, Arun; Nandakumar, Karthik; Jain, Anil (2006). *Handbook of multibiometrics*. Springer.

Wayman, James; Jain, Anil; Maltoni, Davide; Maio, Dario (ed.) (2005). *Biometric systems. Technology, design and performance evaluation*. Springer.

Zhang, David (2000). *Automated biometrics. Technologies and systems*. Kluweer Academic Publishers.

Evaluación de un sistema biométrico: obtención del FMR, el FNMR y la DEC

En este apartado, describimos cómo obtener la DEC a partir de la matriz de similitud. Es decir, la matriz en la que las filas y las columnas representan registros concretos de rasgos biométricos y las celdas son la similitud entre los rasgos biométricos. La tabla 1 muestra un ejemplo de la matriz de similitud.

Supongamos que tenemos una base de datos con tres personas y que cada persona ha realizado tres registros. Es usual que se pidan tres registros a la hora de matricularse para garantizar tener más información de los rasgos biométricos. Estos nueve registros quedan representados por las columnas de la tabla 1. Por otro lado, supongamos que cinco personas han mostrado sus rasgos biométricos para intentar acceder a la base de datos. Las tres primeras personas son las mismas que se han matriculado pero las dos últimas no se han matriculado. Estos cinco intentos de verificación se representan en las cinco filas de la tabla 1.

La tabla 1 se llama *matriz de similitud*, ya que en cada celda está la similitud (obtenida en el modelo de comparación) entre los rasgos biométricos representada en la respectiva fila y columna. En este caso mostramos el valor de similitud × 10.

Tabla 1. Matriz de similitud: tres personas grabadas tres veces; cinco intentos de verificación

	Persona 1			**Persona 2**			**Persona 3**		
Persona 1	**3**	**3**	**4**	3	2	3	3	0	1
Persona 2	3	2	4	**1**	**3**	**4**	2	3	6
Persona 3	1	4	1	2	1	2	**3**	**8**	**5**
Persona 4	3	1	9	2	6	3	5	0	7
Persona 5	2	1	3	3	0	0	1	2	2

Los intentos de identificación genuinos son los valores marcados en negrita. Corresponde a la persona 1 o la persona 2 o la persona 3 cuando dicen que son quienes realmente son y se lleva a cabo la comparación con sus rasgos biométricos. Fijaos en que se tiene que considerar que solo hay tres intentos genuinos y no nueve a pesar de se hayan hecho nueve comparaciones. Los intentos de identificación fraudulentos son el resto de los valores. La persona presenta sus rasgos biométricos al sistema pero dice que es otra persona. Hay 3 × 5 identificaciones menos las 3 genuinas = 12.

El resultado de una verificación es binario: "Se acepta que la persona es quien dice que es" o "No se acepta que la persona es quien dice que es". Esta decisión se toma aplicando el umbral del módulo de comparación (impuesto por el administrador del sistema) y comprobando los tres registros de la persona que el usuario dice que es. Con solo que una de las tres comparaciones esté por encima del umbral, entonces ya consideramos que tenemos una petición correcta: "Se acepta que la persona es quien dice que es". Por otro lado, es necesario que las tres comparaciones estén por debajo del umbral para que consideremos que la petición no es correcta: "No se acepta que la persona es quien dice que es".

Supongamos que el umbral es 4,5 (ha sido multiplicado por 10). Entonces la tabla 2 muestra el resultado de la verificación.

Tabla 2. Resultado de la verificación dada la matriz de similitud de la tabla 1 y el umbral 4,5

Umbral 4,5	Persona 1	Persona 2	Persona 3
Persona 1	~~**diferente persona**~~	diferente persona	diferente persona
Persona 2	diferente persona	~~**diferente persona**~~	**misma persona**
Persona 3	diferente persona	diferente persona	misma persona
Persona 4	**misma persona**	**misma persona**	**misma persona**
Persona 5	diferente persona	diferente persona	diferente persona

Los errores están marcados en negrita. Las falsas aceptaciones son las celdas con valores en negrita y subrayados. Es decir, impostores en los que uno de los tres resultados de la comparación estaba por encima del umbral de similitud. Los falsos rechazos son las celdas con valores en negrita y tachados. Es decir, verificaciones genuinas de que el umbral está por encima de los tres valores (ver la tabla 3).

Tabla 3. Tabla de errores

4.5	Persona 1	Persona 2	Persona 3
Persona 1	FR	RC	RC
Persona 2	RC	FR	FA
Persona 3	RC	RC	AC
Persona 4	FA	FA	FA
Persona 5	RC	RC	RC

FR: falso rechazo
FA: falsa aceptación
RC: rechazo correcto
AC: aceptación correcta

La razón de error de etiquetado (*FMR*) se calcula como el número de falsas aceptaciones (celdas en negrita y subrayadas) dividido por la población impostora. Con el umbral 4,5 toma el valor:

$$FMR = 4/12 = 1/3.$$

La razón de error de no etiquetado (FNMR) se calcula como el número de falsos rechazos (celdas en negrita y tasadas) dividido por la población genuina. Con el umbral 4,5 toma el valor:

$$FNMR = 2/3.$$

Para dibujar la DEC, tendríamos que calcular el valor de la *FMR* y de la *FNMR* para varios valores del umbral, por ejemplo en este caso: 0,5; 1,5; 2,5; 3,5; 4,5; 5,5; 6,5; 7,5; 8,5; 9,5. Con los diez pares de valores de la *FMR* y de la *FNMR* obtenidos, dibujaríamos la gráfica de la DEC.