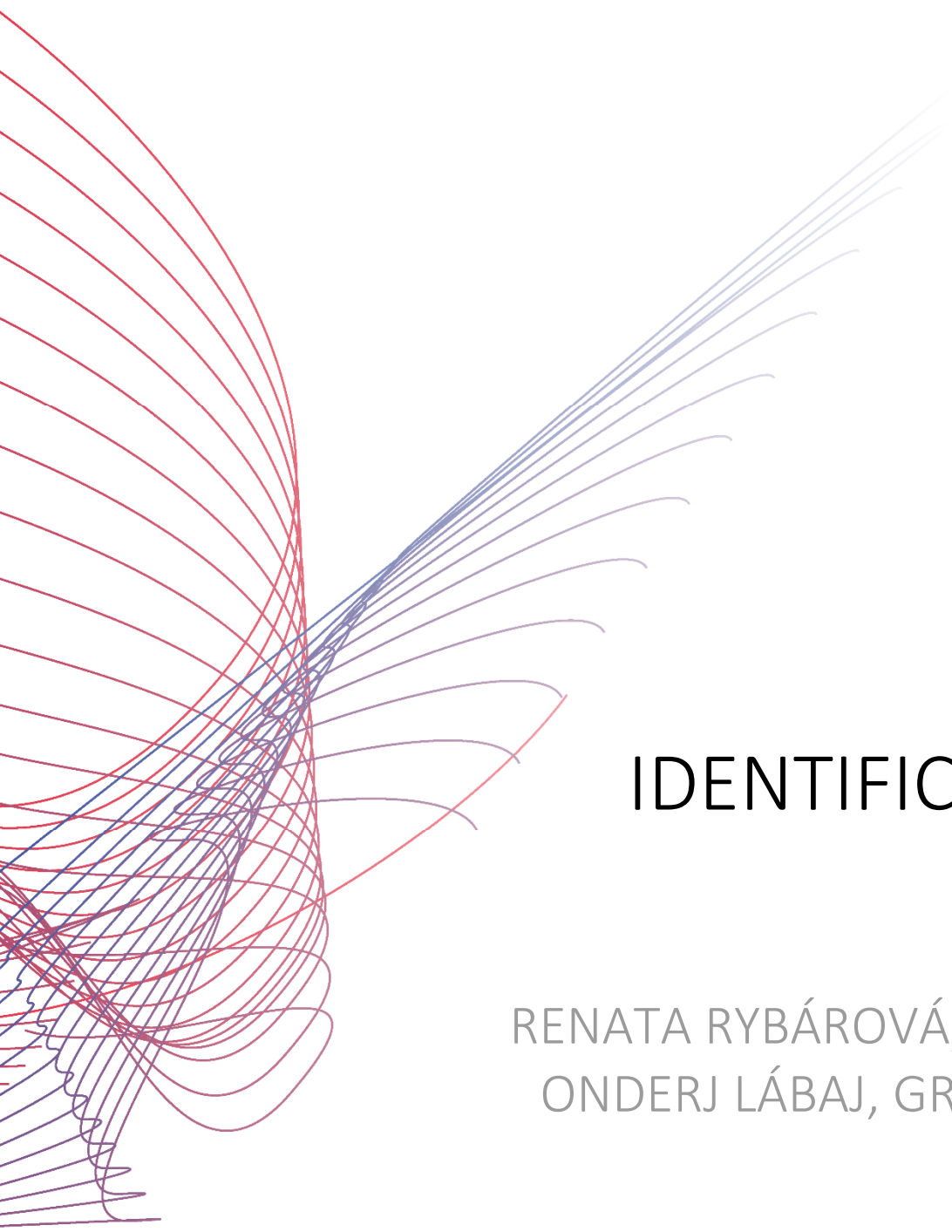
en cs de sk es

TECH pedia



IDENTIFICACIÓN DE USUARIO

RENATA RYBÁROVÁ, JURAJ KAČUR,
ONDERJ LÁBAJ, GREGOR ROZINAJ

Título: Identificación de usuario

Autor: Renata Rybárová, Juraj Kačur, Onderj Lábaj, Gregor Rozinaj

Traducido por: Miquel Soriano

Publicado por: České vysoké učení technické v Praze

Fakulta elektrotechnická

Dirección de contacto: Technická 2, Praha 6, Czech Republic

Número de teléfono: +420 224352084

Print: (only electronic form)

Número de páginas: 43

Edición: Primera edición, 2017

ISBN 978-80-01-06239-5

TechPedia
European Virtual Learning Platform for Electrical and Information Engineering

http://www.techpedia.eu

Erasmus+

El presente proyecto ha sido financiado con el apoyo de la Comisión Europea.
Esta publicación (comunicación) es responsabilidad exclusiva de su autor. La Comisión no es responsable del uso que pueda hacerse de la información aquí difundida.

NOTAS EXPLICATIVAS

	Definición
	Interesante
	Nota
	Ejemplo
	Resumen
	Ventajas
	Desventajas

ANOTACIÓN

La identificación del usuario, la autorización y la autenticación aseguran que el sistema sólo es utilizado por ciertos usuarios y sólo los comandos que están debidamente autorizados se llevan a cabo. La identificación del hablante tiende a proporcionar identificación básica de los posibles usuarios situados en el área de instalación del sistema. Esto sería adecuado para tareas de identificación, tales como la carga del perfil personal. El enfoque de detección de rostros tiene como objetivo proporcionar una identificación del usuario más fiable basada en los rostros de los usuarios que contienen muchas más características que se pueden parametrizar en comparación con el enfoque de identificación de voz. Además, el reconocimiento facial 3D amplía aún más las posibilidades de extracción de características con el fin de identificar con mayor precisión las personas particulares y puede ser utilizado tanto para la autenticación de nivel más alto (y autorización) como para las aplicaciones más exigentes (por ejemplo, el acceso a la cuenta bancaria, etc.).

OBJETIVOS

El objetivo principal del módulo es introducir al estudiante a los fundamentos de los procesos de identificación de usuario, autenticación y autorización. El estudiante estará claramente familiarizado con los principios básicos de identificación del hablante, la identificación del usuario en base a técnicas de 2D y 3D, métodos de autenticación y autorización de usuarios.

LITERATURA

[1] Abate, Andrea F.; Nappi, Michele; Riccio, Daniel; Sabatino, Gabriele. 2D and 3D face recognition: A survey In: Pattern Recognition Letters, Volume 28, Issue 14, 15 October 2007, Pages 1885–1906. available at www.sciencedirect.com.

[2] T. Kinnunen, H. Li, An overview of text-independent speaker recognition: from features to supervectors, Speech communication, Vol. 52, pp. 12-40, Elsevier, 2010

[3] Probst, Michael; Schumann, Sebastian; Rozinaj, Gregor; Minarik, Ivan; Rybárová, Renata; Oravec, Miloš. EVALUATION: Final Multimodal Interface for User/Group-Aware Personalisation, Deliverable 5.5.1, available at http://www.hbb-next.eu/index.php/documents, Decmber 2013.

[4] Bán, Jozef; Féder, Matej; Oravec, Miloš; Pavlovičová, Jarmila. Face Recognition of Images Corrupted by Transmission Errors. In: Redžúr 2012: proceedings; 6th International Workshop on Multimedia and Signal Processing. April 11, 2012, Vienna, Austria. Bratislava: Nakladateľstvo STU, 2012. pp. 15-18, ISBN 978-80-227-3686-2

[5] Rozinaj, Gregor; Minarik, Ivan; Rybárová, Renata; Pavlovičová, Jarmila; Mármol, Félix Gómez; Tormo, Ginés Dólera, Gülbahar, Mark; Schumann, Sebastian. DESIGN AND

PROTOCOL: Final User ID, Profile, Application Reputation Framework, Deliverable 3.4.1, available at http://www.hbb-next.eu/index.php/documents, Decmber 2013.

[6] Schneier, Bruce. Sensible Authentication, ACM Queue 1, Volume 1 Issue 10, February 2004. Pages 74.

[7] McCue, A. Is Your Cat a Target for Password-Stealing Hackers?, Silicon.com, 11 August 2004.

[8] Haskett, J.A., Pass-Algorithms: A User Validation Scheme Based on Knowledge of Secret Algorithms In Communications of the ACM 27, 1984.

[9] Madigan, A. Picture Memory - Memory and Cognition: Essays in Honour of Allan Paivio Erlbaum, 1983.

[10] Cranor, L.F.; Garfinkel, S. Security and Usability, O'Reilly, August 2005. ISBN 0-596-00827-9.

[11] Vacca, J.R. Computer and Information Security Handbook, Morgan Kaufmann, 2009. ISBN 978-0-12-374354-1.

[12] Gattiker, U. E. The Information Security Dictionary, KLUWER ACADEMICPUBLISHERS, 2004. ISBN 1-4020-7927-3.

Indice

1 Identificación de usuario

La identificación del usuario es una de las características clave que aseguran que el sistema o aplicación sólo realiza los comandos que están debidamente autorizados. El tipo de autenticación más utilizado es la contraseña, pero con el desarrollo de las tecnologías de la información y los algoritmos de protección de seguridad, los sistemas y aplicaciones comienzan a utilizar la autenticación basada en los factores biométricos.

El uso de datos biométricos elimina efectivamente los posibles riesgos asociados con tecnologías menos avanzadas que se basan en lo que una persona tiene o sabe más que en lo que una persona realmente es [1]. Es una tecnología muy atractiva y popular, ya que puede ser integrada en cualquier aplicación o sistema que requiera control de seguridad o de acceso.

La identificación del hablante tiende a proporcionar identificación básica de los posibles usuarios situados en el área de instalación del sistema. El enfoque de detección de rostros tiene como objetivo proporcionar una identificación del usuario más fiable basada en los rostros de los usuarios que contienen muchas más características que se pueden parametrizar en comparación con el enfoque de identificación de voz. Además, el reconocimiento facial 3D amplía aún más las posibilidades de extracción de características con el fin de identificar con mayor precisión a las personas particulares y puede ser utilizado tanto para la autenticación de nivel más alto (y autorización) como para las aplicaciones más exigentes (por ejemplo, el acceso a la cuenta bancaria, etc.). Por razones de seguridad, la autenticación de reconocimiento facial 3D se puede mejorar con, por ejemplo el seguimiento de movimiento ocular o el reconocimiento del iris. Este enfoque puede simular autenticación de múltiples factores (login y además token) necesario para una autenticación de nivel más alto.

El reconocimiento del iris es extremadamente preciso, pero costoso de implementar y no muy aceptado por la gente. Las huellas dactilares son fiables y no intrusivas, pero no es adecuado para las personas no predispuestas a la colaboración. Por el contrario, el reconocimiento de la cara presenta ser un buen compromiso entre la fiabilidad y la aceptación social [1].

2 Identificación del hablante

2.1 Visión general de la identificación del hablante

La identificación del hablante es una parte de un concepto más amplio conocido como el reconocimiento de voz. Comprende dos tareas importantes y de alguna forma similares, pero diferentes, a saber: la identificación del hablante y la verificación del hablante. La primera de ellas básicamente apunta a una asignación de decidir automáticamente que la muestra de voz probada pertenece a un individuo a partir de un conjunto de usuarios almacenados en una base de datos durante una fase de inscripción (formación).

Opcionalmente, si la confianza de una decisión final es demasiado baja no se reconoce a nadie. Esta tarea se conoce a menudo como un problema de grupo cerrado, ya que existe un conjunto fijo de usuarios que pueden reconocerse.

Por otra parte, el proceso de verificación de la mano evalúa si el individuo probado es el que él o ella dice ser.

Como hay muchos usuarios (posiblemente 7.000 millones de personas que viven en el mundo), es imposible que se hayan medido las características o modelos para todo el mundo. Por lo tanto, se denomina a esta tarea como un problema en grupo abierto. En esta situación el modelo de un hablante en general es vital para determinar los umbrales adecuados de aceptación / rechazo.

El reconocimiento del habla es un problema bastante difícil debido a muchas razones mencionadas más adelante en el texto y ha sido objeto de investigación científica seria a lo largo de más de 40 años por muchos equipos de investigación. Como están surgiendo tecnologías nuevas y accesibles, esta técnica está encontrando rápidamente aplicación en muchas áreas, sólo por mencionar algunas de ellas:

- ciencia forense
- método natural y no invasivo para el acceso seguro y protección de datos y servicios
- indexación automática de voz y grabaciones de audio almacenados en bases de datos
- rango de aplicaciones en industrias de juegos
- ayuda a discapacitados

: Debido a la amplia gama de problemas que deben ser abordados hay muchas soluciones y técnicas comunes relacionadas con el problema de la identificación del hablante. Estas pueden ser clasificadas en 3 grupos principales:

- **Características de la voz** - adecuadas para el reconocimiento del hablante o tareas de identificación del hablante

- **La normalización de características / de compensación de modelo-** que pretende suprimir la variabilidad del período de sesiones

- **Clasificación y algoritmos de toma de decisiones**- las decisiones basadas en características y modelos representa el mejor aproximación a una muestra desconocida

Por último, la tarea de identificación del hablante se divide en dos grandes grupos, es decir, dependiente del texto y el problema de texto independiente. En el primer grupo el proceso de identificación no asume ningún texto específico, mientras que en la clase siguiente, los sistemas requieren de texto preciso para ser pronunciado. Obviamente el sistema dependiente del texto alcanza una mejor precisión.

2.2 Propiedades de las señales de voz

Las señales de voz genuinas son creadas y producidas por los seres humanos; más precisamente por su aparato vocal y sus cerebros que son únicos para cada individuo. Ambas fases naturalmente dejan sus marcas en la señal audible, y por lo tanto el habla puede considerarse como una señal biométrica.

A pesar de que el objetivo principal de las señales de voz es transmitir la información léxica que contienen. Excepto la parte léxica, que está más o menos dada por la secuencia de diferentes posiciones de los órganos vocales, la voz contiene información biométrica sobre cualquier hablante representada principalmente por diferentes formas, tamaños, pesos y las características de los órganos vocales, el estado de ánimo real de una persona (la entonación, el ritmo del habla, estrés, etc.), y el fondo social de una persona (el dialecto, vocabulario, etc.).

Sin embargo, estas diferentes piezas de información se codifican en la señal de voz mediante una transformación difícil que se cree que es irreversible y no se conoce. Por lo tanto, extraer sólo la información que es necesaria para una tarea en particular (léxico, identificación, estado de ánimo, estado de salud, ...) es un problema difícil. Además, el habla presenta gran variabilidad en función del hablante, que viene dada por el estado de ánimo, la salud y el estado físico u otras condiciones. Finalmente la forma acústica de una señal de voz puede verse seriamente alterada por las diferencias en los dispositivos de grabación, sala donde se registró y si existe ruido de fondo.

Las modificaciones del habla que no están relacionados con el hablante (dispositivos, habitación, etc.) se denominan variabilidad de la sesión. Este aspecto crea problemas y debe ser tratado en consecuencia repitiendo las medidas en una situación en la que las condiciones de no coincidan con las de sesiones anteriores.

2.3 Extracción de características

Debido a las características variables del habla y a muchas condiciones adversas mencionadas en el texto anterior, ha habido muchas técnicas de extracción inventadas a través del tiempo. Básicamente, una buena caracterización del habla debe ser:

- discriminativa
- robusta frente a distintos tipos de ruido de fondo
- insensible a los cambios causados por los dispositivos y lugares de grabación
- suprimir la variabilidad del hablante
- fácil de procesar

Como hay muchas características diferentes de hablantes que tienen diferentes significados físicos, distinguimos 3 tipos de características (desde el punto de vista de reconocimiento de voz):

- Acústicas
- Prosódicas
- De alto nivel

A nivel acústico, las características asociadas a la voz en espacios de tiempo de corta duración están relacionadas con las características físicas del aparato vocal. Estos métodos representan principalmente formas espectrales modificadas (envolvente) extraídas de intervalos que van de 10 ms a 30 ms. Además, aplican diferentes principios psicoacústicos para aumentar su robustez. En la actualidad, los más comunes son *Mel frequency cepstral coefficients* (**MFCC**), *Perceptual Linear Prediction* **(PLP)**, o *Cepstral Linear Prediction coefficients* **(CLPC)**. MFCC y PLP intentan capturar envolventes espectrales modificadas siguiendo algunos principios psicoacústicos como bandas críticas, la percepción humana de las frecuencias, la curva de igual sonoridad, la conversión de las intensidades de volumen, etc.

Como son capaces de extraer envolventes espectrales, preservan y hacen hincapié en la ubicación, anchuras y formas de las frecuencias que son vitales para la percepción de las diferencias entre los sonidos. Así que son muy importantes para los sistemas de reconocimiento de voz. Incluso juegan un papel importante en el problema de reconocimiento del hablante. Se puede explicar el modo en que son capaces de capturar ligeras diferencias en localizaciones y formas de las frecuencias que varían de una persona otra como las diferencias observadas entre teléfonos particulares. La caracterización CLPC se basa en la modelización del mecanismo de producción del habla en lugar del proceso de la audiencia y la percepción. Finalmente, para abarcar la dinámica de las características acústicas en el tiempo, se pueden evaluar coeficientes diferenciales y de aceleración. A medida que se cubren intervalos de tiempo más largos, se pueden detectar diferencias en la co-articulación que son específicas para un hablante particular. El nivel prosódico se centra más en el estilo de hablar, el estado de ánimo de un hablante, hábitos de habla

específicos, las condiciones físicas y de salud, etc. Obviamente esta información sólo se detecta y se puede extraer utilizando intervalos de tiempo más largos que se extienden a varios segundos de discurso. Las características favoritas para este nivel son: la dinámica del habla, el ritmo del habla, la modulación de la frecuencia fundamental, tipo de pausas que se hizo al hablar, etc. Sin embargo, estas características son más difíciles de medir y calificar que las de nivel acústico. Así, hay varios métodos para extraer y evaluarlos durante intervalos de tiempo apropiados. Los enfoques más comunes son la función de autocorrelación, *Average Magnitude Difference Function* **(AMDF)**, el filtrado inverso para la detección de la frecuencia fundamental, energía para la dinámica del habla y así sucesivamente. Sin embargo, hay muchas modificaciones tanto de autocorrelación como de AMDF.

2.4 Algoritmos de clasificación y toma de decisiones

Después de la extracción de características la fase de normalización / compensación (que se presentará en la próxima sección) debe ser utilizada para decidir qué usuario (características o modelos) son los encierra una incógnita. Aún es posible rechazar cualquier si el partido / la confianza grabado es demasiado bajo. Hay varias técnicas de clasificación de éxito que difieren en su complejidad, así como en su funcionalidad y en lo que suponen los datos procesados. Estos métodos se clasifican en varias categorías principales, cada una con sus pros y contras de la siguiente manera:

- **Métodos no paramétricos**- no imponen ninguna restricción a los datos, por lo que no utilizan ningún modelo para describir el espacio. Su principal representante es el *K-nearest neighbour method* **(KNN)** KNN encuentra los k vectores más cercanos al desconocido y en base a ellos utiliza algunos criterios para decidir sobre el resultado.

- **Métodos basados en parámetros**- asumir alguna estructura del espacio de características y modelarlo por algunos parámetros. La más exitosa y de uso frecuente es la mezcla de distribuciones gaussianas. Este modelo se denomina *Gaussian mixture model* **(GMM)**.. En ausencia de datos y utilizando modelos adecuados que coinciden con el espacio, estos métodos son superiores a los no paramétricos.

- **Métodos discriminativos** – intentar modelar / dividir el espacio de características de modo que el error de clasificación sea tan pequeño como sea posible. Es bastante fácil de hacer esto en la parte de la formación de la base de datos, pero es más complicado hacerlo (sólo estimar) para los datos que no se ven. En tal situación se requiere una buena capacidad de generalización (bajo error para las muestras que no se ven). Los más representativos de este gruo son las redes neuronales: *neural networks* (**NN**) y las máquinas de vectores: *support vector machines* (**SVM**). Ambos NN y SVM pueden dar excelentes resultados en condiciones específicas.

- **Métodos generativos** - Si los modelos se adaptan perfectamente a los datos (en la realidad no lo hacen), es posible construir un clasificador óptimo que alcanza el mínimo coste posible que definió Bayes, por lo que se conoce como la clasificación de Bayes. Una vez más, el modelo más exitoso es el mencionado anteriormente GMM.

Se ha descubierto que es beneficioso utilizar modelos generales de hablantes universales que han sido creados por las muestras de formación de muchos oradores.

2.5 Compensación de entorno

Para disminuir la variabilidad de sesión causada por diferentes condiciones de ensayo de estudios (ruidos de fondo, diferentes parámetros acústicos de dispositivos de grabación y salas) se han definido y empleado varios conceptos. Los métodos más básicos normalizan de manera uniforme la dinámica de la señal mediante la manipulación de la potencia total o tratan de igualar la potencia de cada banda de frecuencia de un espectro de voz promedio utilizando normalmente la resta del valor medio de los picos. Además, es posible utilizar técnicas de filtrado fijas que hacen hincapié en una señal de voz general, como la amplificación de la modulación de espectro del habla o el filtrado *relative spectral analysis* (**RASTA)**. Otros métodos más sofisticados intentan encontrar características de inscripción o mapeo de transformaciones óptimas a las características observadas en el entorno de trabajo (los llamados métodos de función de mapeo) o para transformar los modelos integrales de hablantes para que coincidan con el modelo del entorno de trabajo (se llama síntesis del modelo de hablante). Sin embargo, estos métodos se basan en matemáticas avanzadas, y adaptan su comportamiento a los datos entrantes. Así, si el entorno de trabajo está cambiando a la vez que su asignación óptima.

Otra solución menos sofisticada, pero útil en algunos momentos, es tener muestras pregrabados de voz (características o modelos) en diferentes condiciones y previamente al reconocimiento detectar el apropiado. A continuación, se utiliza el mejor ambiente para una grabación en particular. Es obvio que los mejores resultados se observan cuando hay una coincidencia entre los entornos de trabajo y las pruebas.

Para obtener una visión más detallada sobre el tema de reconocimiento del hablante, por favor, estudiar la referencia [2].

3 Reconocimiento facial

La cara se ha convertido en uno de los elementos biométricos más populares y los sistemas de reconocimiento biométrico facial para la identificación personal se utilizan cada vez más en una amplia gama de aplicaciones. El desarrollo de algoritmos y métodos de reconocimiento hace posible el uso de los sistemas de identificación y verificación en el campo comercial. Sin embargo, estos sistemas no logran tasas de reconocimiento comparables en condiciones no controladas y sin restricciones. El reconocimiento facial en estas condiciones sigue siendo un problema difícil, a pesar de los recientes avances.

Los sistemas biométricos de identificación personal, que son desarrollados por varios proveedores, logran alta precisión en el reconocimiento facial. La mayoría de estas aplicaciones requieren [3]:

- sistemas de reconocimiento que pueden reconocer varias caras de un fotograma de vídeo o una imagen
- alta tasa de reconocimiento
- invariancia de iluminación
- estabilidad ante expresiones faciales cambiantes
- reconocimiento en tiempo real, etc.

Así que podemos ver que hay varios factores que pueden afectar el rendimiento y la precisión de los sistemas de reconocimiento facial [1]:

- Variaciones de **iluminación** debido a las propiedades de reflectancia de la piel y debido al control interno de la cámara. Varios métodos 2D hacen bien tareas de reconocimiento sólo bajo una variación de iluminación moderada.
- Los **cambios de postura** afectan el proceso de autenticación, porque introducen deformaciones de objetos. Los métodos de detección deben resolver el problema teniendo en cuenta diversos ángulos de visión cuando el objeto está colocado (es decir, visto desde las cámaras de seguridad). Por otro lado, los algoritmos son relativamente robustos a la expresión facial (excepto a algunas expresiones extremas como el grito).
- El **tiempo de retardo** es también un factor importante, teniendo en cuenta que la cara cambia con el tiempo, de una manera no lineal durante largos períodos (variaciones de edad). En general, este problema es difícil de resolver con respecto a los otros.

3.1 Métodos de reconocimiento facial

Los sistemas de reconocimiento facial se dividen en dos categorías: Verificación e identificación.

La verificación de la cara es 1: 1. En este proceso la imagen facial cuya identidad está siendo reclamada se compara con una plantilla de imagen facial.

Por el contrario, en la identificación de la cara 1: N, la imagen facial se compara con todas las plantillas de imagen en una base de datos faciales de cara a determinar la identidad.

En el caso de que no sabemos si la cara está en la base de datos del sistema, el proceso es el siguiente. La imagen de la cara también se compara con todas las imágenes de la cara en la base de datos, y se realiza la evaluación de la probabilidad para cada caso. Todas estas probabilidades se clasifican numéricamente: el valor más alto es el primero. En caso de que la probabilidad es mayor que un umbral determinado, el sistema nos avisa sobre el resultado [1].

Los métodos de reconocimiento facial 2D básicos seleccionados:

- Métodos de proyección Lineal/no lineal
    - *Análisis de los componentes principales :Principal Component Analysis* (**PCA**) - el método basado en PCA se denomina Eigenface. La idea principal del PCA es descomponer un espacio de datos en una combinación lineal de una pequeña colección de bases, que son pares ortogonales y que captan las direcciones de máxima varianza en el conjunto de entrenamiento [4].
    - *Análisis de los components del Kernel principal : Kernel Principal Component Analysis* (**KPCA**) - Es un método de extracción de características no lineal. El IPCA puede extraer el conjunto de características que es más adecuado en la categorización que el PCA convencional. IPCA ha sido ampliamente utilizado en el caso de reconocimiento facial con expresión y bajo iluminación variante [4].
    - *Análisis discriminante lineal : Linear Discriminant Analysis* (**LDA**) -se ha propuesto como una alternativa mejor a la PCA. Proporciona la discriminación entre las clases, mientras que el PCA se ocupa de los datos de entrada en su totalidad, sin prestar atención a la estructura subyacente. De hecho, el principal objetivo del LDA consiste en encontrar una base de vectores que proporcionan la mejor discriminación entre las clases, tratando de maximizar las diferencias entre la clase, minimizando los de dentro de clase [1].
    - *Vectores comunes discriminantes ; Discriminant Common Vectors* (**DCV**) - la idea principal de DCV consiste en la recogida de las similitudes entre los elementos de la misma clase eliminando sus diferencias [1].
- Las redes neuronales - es una solución no lineal, se utiliza también en otros problemas de reconocimiento de patrones. La ventaja de los clasificadores

neuronales sobre los lineales es que pueden reducir errores de clasificación entre clases de vecinos. La idea básica es considerar una red con una neurona para cada pixel en la imagen. Sin embargo, debido a las dimensiones del patrón, las redes neuronales no están directamente entrenadas con las imágenes de entrada, pero la aplicación de una técnica de reducción de dimensionalidad se utiliza antes del entrenamiento [1].

- Fractales y Sistemas de función iterativa : *iterated function systems (IFS)* –La teoría IFS se ha desarrollado principalmente en el área de la codificación de imagen y últimamente se ha extendido a la indexación de la imagen. El código Fractal de una imagen es invariante con respecto a un amplio conjunto de transformaciones globales, como rotaciones, escalamiento contraste, etc. El IFS fractal de una imagen de la cara se utiliza para entrenar a las redes neuronales, donde se utiliza como clasificador [1].

3.2 Extracción de características

Algunos algoritmos de reconocimiento facial se basan en características extraídas de una imagen de la cara del sujeto - en las características faciales. Por ejemplo, un algoritmo puede analizar la posición relativa, el tamaño y / o la forma de los ojos, nariz, boca, pómulos y de la mandíbula. Estas características se utilizan entonces durante la búsqueda en el grupo de imágenes para el juego de características. Otros algoritmos normalizan una galería de imágenes de la cara y luego se comprimen los datos faciales, salvando sólo los datos de la imagen que es útil para el reconocimiento de rostros. Una imagen probada se compara entonces con los datos faciales.

Antes de la extracción de características de todas las imágenes deben ser pre-procesadas y normalizadas.

Una parte del pre-procesamiento es la reducción de la dimensión de todas las imágenes de entrada a un tamaño definido. También puede aplicarse una ecualización de histograma adaptativo con contraste limitado: *contrast limited adaptive histogram equalization* (**CLAHE**). Las imágenes normalizadas se pueden enmascarar, omitir el fondo y dejar sólo la región de la cara.

El objetivo principal del proceso de normalización es minimizar las variaciones incontroladas que se producen durante el proceso de adquisición y mantener las variaciones observadas en las diferencias de características faciales entre los individuos.

El cambio de actitud o postura puede llevar a diferencias en imágenes.

La extracción de características consiste en reducir la cantidad de recursos necesarios para describir un conjunto grande de datos. Durante el reconocimiento de rostros, se realiza el análisis de la gran cantidad de datos. El análisis con un gran número de variables generalmente requiere una gran cantidad de memoria y potencia de cálculo. La extracción de características está relacionada con la reducción de variables y datos.

Para la extracción de rasgo facial se usan los métodos de detección de bordes. Muy buenos resultados se consiguen también por patrones binarios locales: *local binary patterns* (**LBP**).

La detección de bordes es el nombre de un conjunto de métodos matemáticos cuyo objetivo principal es detectar puntos en una imagen digital, donde el brillo cambia bruscamente. Estos puntos de imagen con el cambio de brillo se organizan normalmente en una serie de segmentos de línea curvados denominados bordes.

Las funciones que se utilizan con mayor frecuencia para las detecciones de borde son operadores Sobel (también llamado filtro de Sobel), el operador de Prewitt o los filtros de Gabor.

La extracción desde caras pre-procesadas se puede hacer a través de histogramas LBP como características. Los histogramas LBP se consideran una de las mejores características para el reconocimiento de caras, incluso cuando sólo un número limitado de muestras está disponible y además puede ser fácilmente calculada en tiempo real [5] (Fig. 2.1).

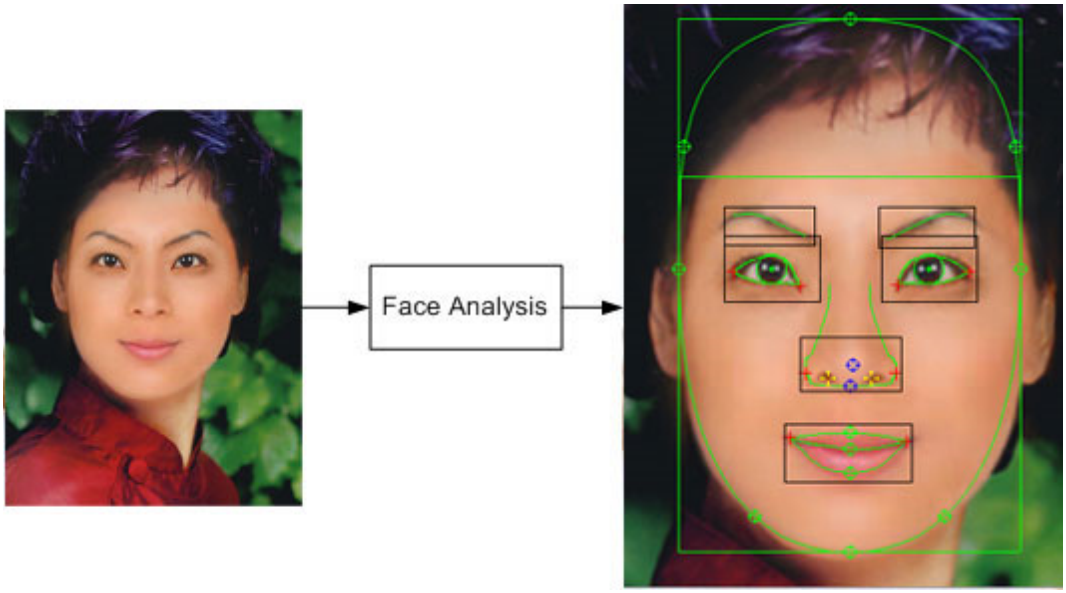


Fig. 2.1 – Ejemplo de resultado de extracción de características

3.3 Clasificación de caras

Un sistema de reconocimiento facial funciona normalmente en dos fases principales. La primera fase es un proceso de formación y el segundo es la clasificación de los usuarios. Los métodos de reconocimiento facial modernos funcionan correctamente cuando hasta 10 imágenes de una persona están disponibles en la etapa de entrenamiento. Incluso se han desarrollado numerosas técnicas para el reconocimiento facial a partir de una sola imagen por persona. El proceso de formación debe ser completamente automatizado y los usuarios tienen que ser capaces de controlarlo. El proceso de capacitación utiliza algoritmos de clustering.

El objetivo principal de todos los algoritmos de clustering es identificar grupos o clases en el conjunto de datos de entrada. Hay muchos algoritmos de agrupamiento o clustering. Estos algoritmos se pueden dividir en dos grupos: De partición y algoritmos jerárquicos [5].

Como ejemplo de algoritmo de agrupamiento podemos mencionar K-means. Otro algoritmo utilizado para la agrupación es el *mapa auto organizado* (SOM), perteneciente a las técnicas de r*edes neuronales o agrupación espacial basadas en la densidad con ruido* (DBSCAN).

Para la clasificación de las características extraídas de caras tenemos dos métodos, dependiendo del número de imágenes de capacitación y número de identidades que se va a utilizar en el sistema :

- Máquinas de vectores soporte - se utiliza cuando sólo se considera un número de identidades en el sistema relativamente pequeño. Su principal problema es que el entrenamiento consume mucho tiempo del modelo cuando se utiliza gran número de muestras.
- Vecino-k más cercano (con el uso de la distancia Chi-cuadrado) - Este algoritmo puede ser paralelizado y se utiliza en sistemas distribuidos fácilmente. La formación se hace simplemente mediante la inserción de características en la base de datos [5].

3.4 Localización y reconocimiento facial

Los sistemas de reconocimiento facial biométrico son ampliamente utilizados en muchos tipos diferentes de aplicaciones. En la actualidad, un televisor inteligente con sistema de reconocimiento facial es un ejemplo típico de dicha aplicación. El reconocimiento facial en la televisión inteligente se utiliza para la autenticación espectador y basado en esto, servicios personalizados o diferentes recomendaciones pueden ser proporcionados. Los sistemas de reconocimiento facial se deben dar en tiempo real y deben ser capaces de reconocer una o más identidades. La mayor parte de estos sistemas incluyen también la interfaz gráfica de usuario para el proceso de formación automática (Fig. 2.1).

Por lo general, la tarea de reconocimiento facial 2D requiere un procesamiento de la entrada de una cámara. El principal proceso de reconocimiento facial consiste en los siguientes subprocesos:

- Adquisición de imagen - lee una imagen de la cámara, la convierte al formato del sistema y pasa al proceso del sistema
- **localización facial** - localiza las caras en la imagen y coordenadas asociadas con la imagen. Dependiendo de la cámara que se utiliza se implementa un algoritmo de localización.
- **Proceso de entrenamiento** – se usan algoritmos de clustering o clasificación, ej. K-means
- **pre-procesado** de caras localizadas, incluye ecualización de histograma
- **normalización** – ej. redimensionado
- **extracción de características** - extrae características de caras, pre-procesado, puede utilizar LPB.
- **clasificación facial** - utilizar métodos como Máquinas de Vectores Soporte o vecino-K más cercano.
- **Seguimiento facial** – sólo se realiza un seguimiento de caras frontales en la imagen, porque la gran mayoría de los métodos de reconocimiento son fiables solamente con el uso de imágenes de caras frontales. Una vez que el rostro ha sido reconocido, se hace un seguimiento, lo que ahorra significativamente recursos computacionales y pueden seguir el tema, incluso después de cambios en la pose [3]. Así la información sobre el usuario reconocido se envía como salida del sistema.

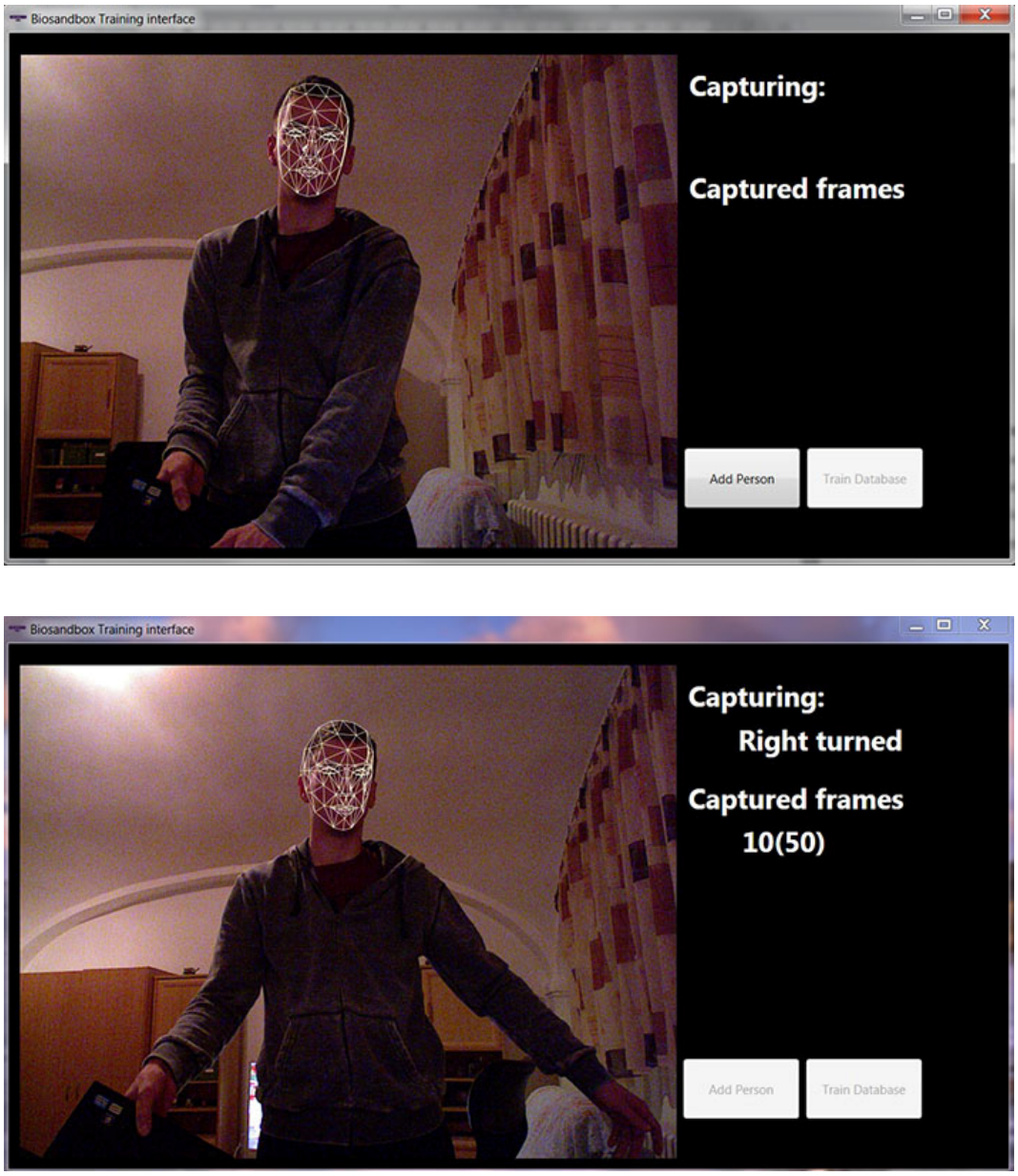


Fig. 2.2 – Ejemplo de entrenamiento GUI de Sistema de reconocimiento facial

3.5 Reconocimiento de Iris

El iris es uno de los rasgos biométricos más populares. La combinación de digitalización sin contacto, estabilidad en el tiempo y la alta precisión de reconocimiento permite su uso en la vigilancia, así como en aplicaciones de seguridad.

Se ha demostrado que la precisión de reconocimiento del iris depende de la calidad de la imagen del iris y del pre-procesamiento de imagen capturada. Para reducir la influencia negativa de la iluminación, se recomienda una cámara de detección de luz **NIR** (Infrarrojo cercano) (Fig. 2.3). El uso de la luz NIR permite añadir la fuente de luz, además, sin influir en la comodidad de detección.

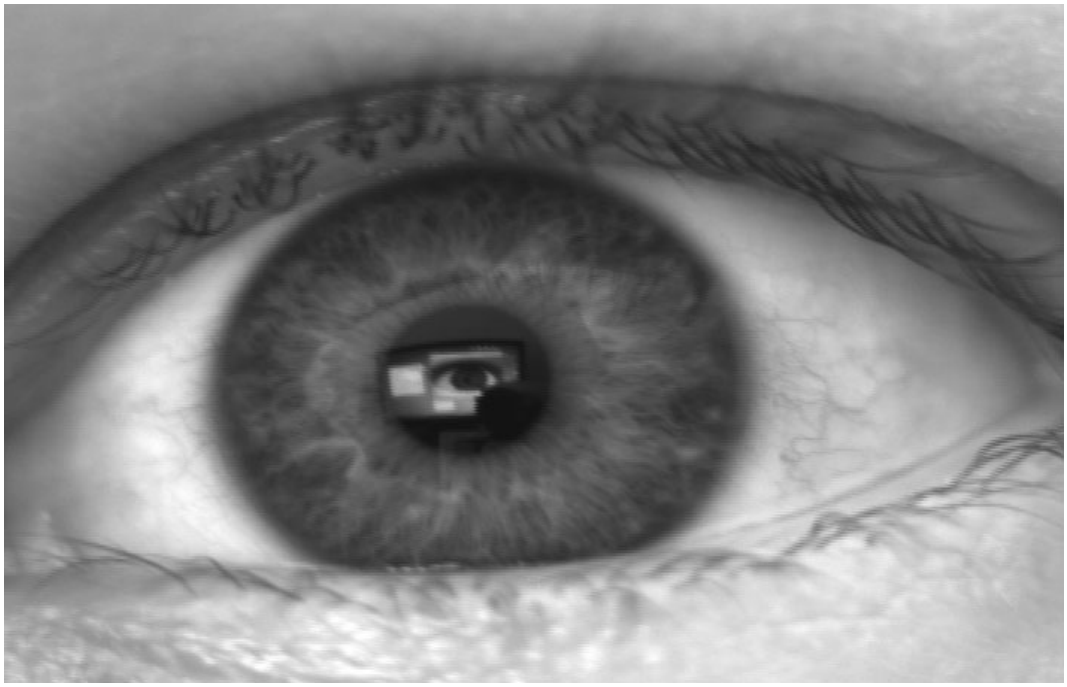


Fig. 2.3 – Ejemplo de foto tomada con una cámara Guppy F-038B NIR

La identificación basada en el iris consiste en la localización del iris, extracción de características y clasificación. Uno de los sistemas más exitosos presenta un 100 por ciento de precisión en ambientes controlados. Sin embargo, es necesario mejorar la localización y la normalización para la aplicación en la vida real. Este sistema utiliza filtros de Gabor para extracción de características donde las señales filtradas se cuantifican en dos niveles. Por este procedimiento, se obtienen cadenas de dígitos binarios (características). Para cotejar las muestras más cercanas se utilizan el método KNN y el reconocimiento de distancia de hamming.

4 Reconocimiento facial 3D

El reconocimiento facial basado en el reconocimiento 2D de la cara es un enfoque común y natural. Los resultados de reconocimiento facial 3D dan en general mayor seguridad que el enfoque de reconocimiento facial 2D.

Las técnicas basadas ende reconocimiento facial 3D deben poseer varias propiedades tales como robustez con respecto a las variaciones de iluminación así como la posición, la rotación y el escalado del modelo original dentro de un marco de referencia absoluto [1].

4.1 Métodos 3D de reconocimiento facial

El reconocimiento facial 3D, en comparación el reconocimiento facial 2D, utiliza más información sobre las características faciales. Del mismo modo, ambos enfoques tienen pre-procesamiento básico como la normalización del tamaño de la cara, la rotación a una posición neutral etc. Esta mayor cantidad de información añadida, no sólo la que usa el reconocimiento 2D sino también sobre el análisis en profundidad, mejora el análisis. Las principales ventajas son :

- No se ve afectado por las variaciones de iluminación o el uso de los cosméticos
- Menos sensible a las variaciones de apariencia
- Se manejan mejor los cambios de postura
- La naturaleza proyectiva de imágenes 2D
- Simplifica la cara y la detección de características faciales, plantean la estimación y suponen una compensación

Los métodos de reconocimiento facial 2D básicos seleccionados:

- Basados en la superficie de reconocimiento facial 3D -este enfoque se basa en las técnicas de reconocimiento de objetos 3D clásicos (Fig. 3.1). Hay varios tipos de métodos de reconocimiento basados en:
    - Uso de características de curvatura locales, que son invariantes a la rotación (por ejemplo, la curva del perfil de la cara)
    - Uso de la coincidencia punto a punto (polígono de varios puntos de la cara significativos)

Fig. 3.1 –Reconocimiento 3D basado en superficie

- Reconocimiento de caras 3D basado en la Apariencia - (Fig. 3.2). este método se ocupa de Eigenfaces y Fisherfaces. Se requiere alineación exacta de la sonda y las imágenes en la base de datos. Los rasgos faciales como los ojos, la boca, etc. son localizados y utilizados para el reconocimiento. El método es fácil de implementar y no requiere mucho tiempo (Fig. 3.2).

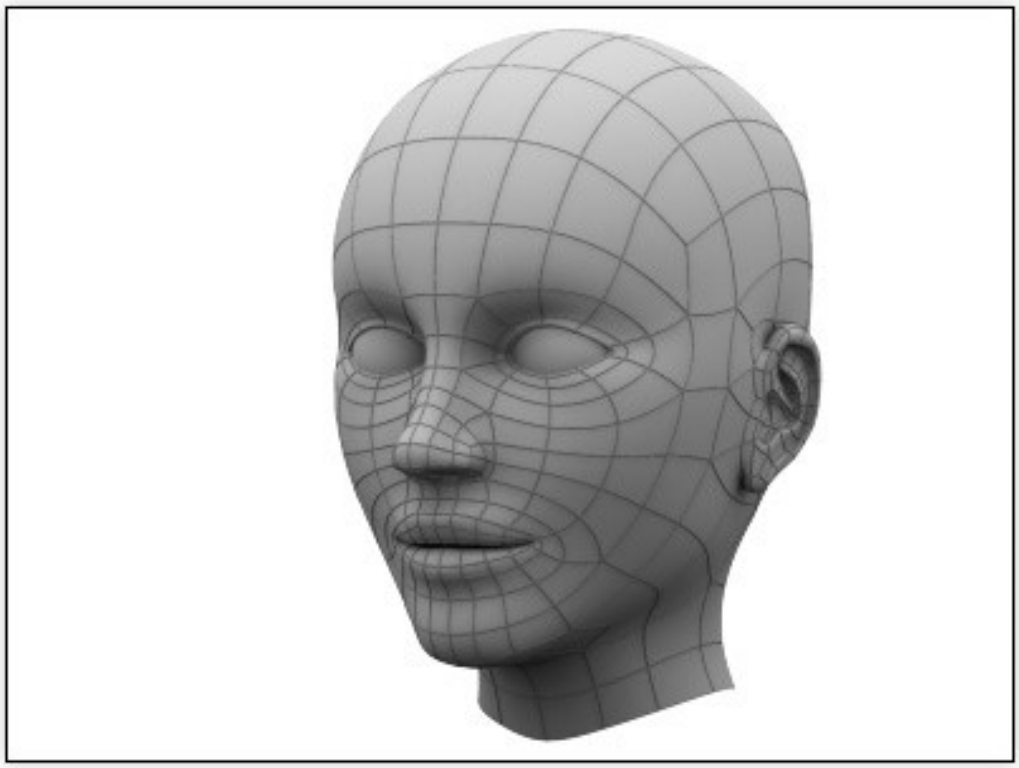


Fig. 3.2 –Reconocimiento de caras 3D basado en la Apariencia

- Reconocimiento de caras 3D basado en el modelo - Este método se basa en el análisis por método de síntesis, donde se produce el modelo de la cara anotado en 3D, que luego se compara con los modelos en la base de datos. No es adecuado para aplicación en tiempo real (Fig. 3.3).

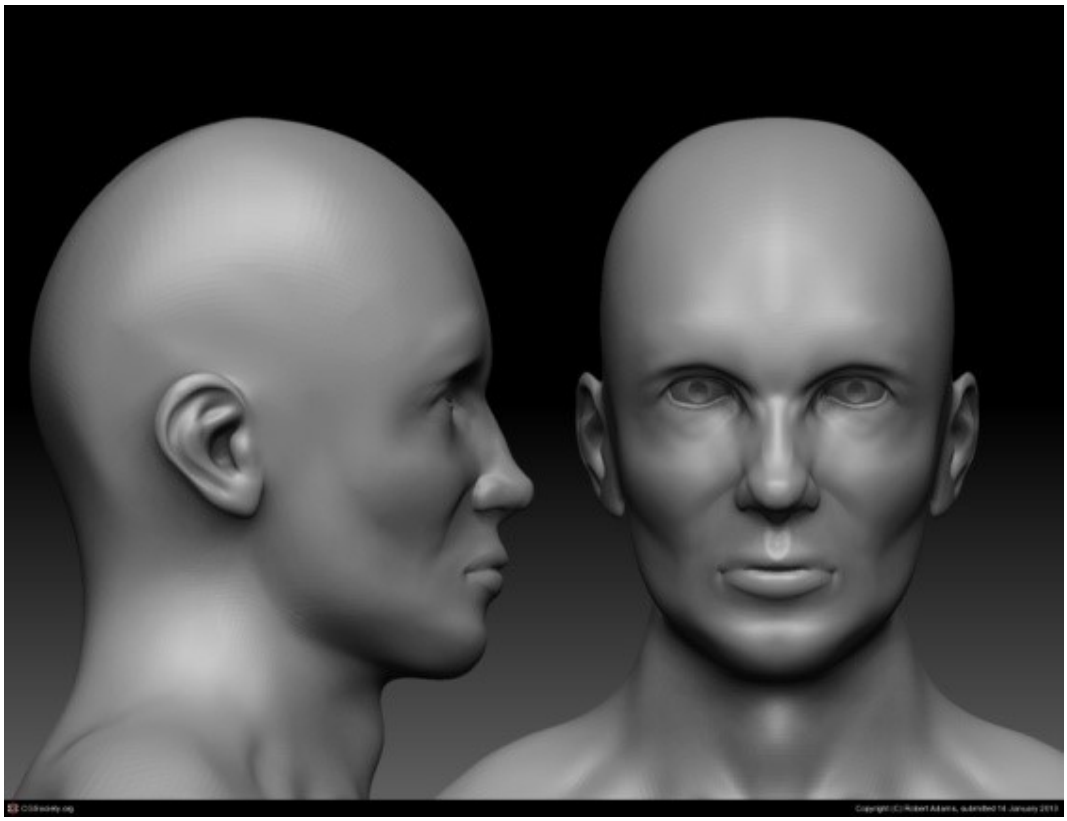


Fig. 3.3 –Reconocimiento de caras 3D basado en el modelo

4.2 Pre-procesamiento y registro de datos

Al principio de todo el proceso se captura la superficie facial 3D (un ejemplo de creación de la cara 3D son las fotos Fig 3.4 -. Fig. 3.6). Hay varias maneras diferentes de lograr esta tarea, por ejemplo cámaras estéreo, cámara de profundidad, láser, escáner óptico o láser, etc.

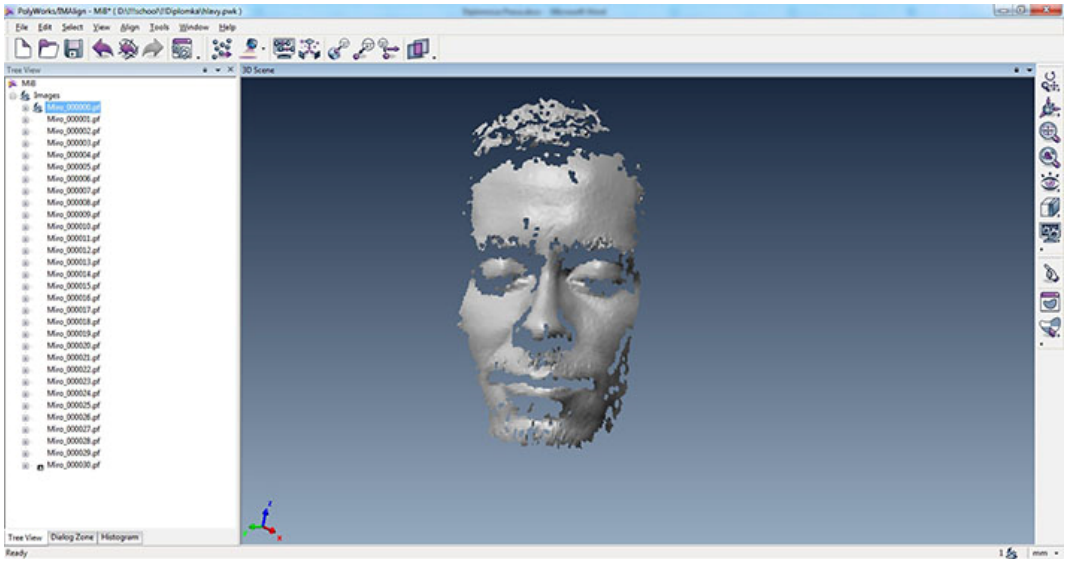


Fig. 3.4 – Escáner

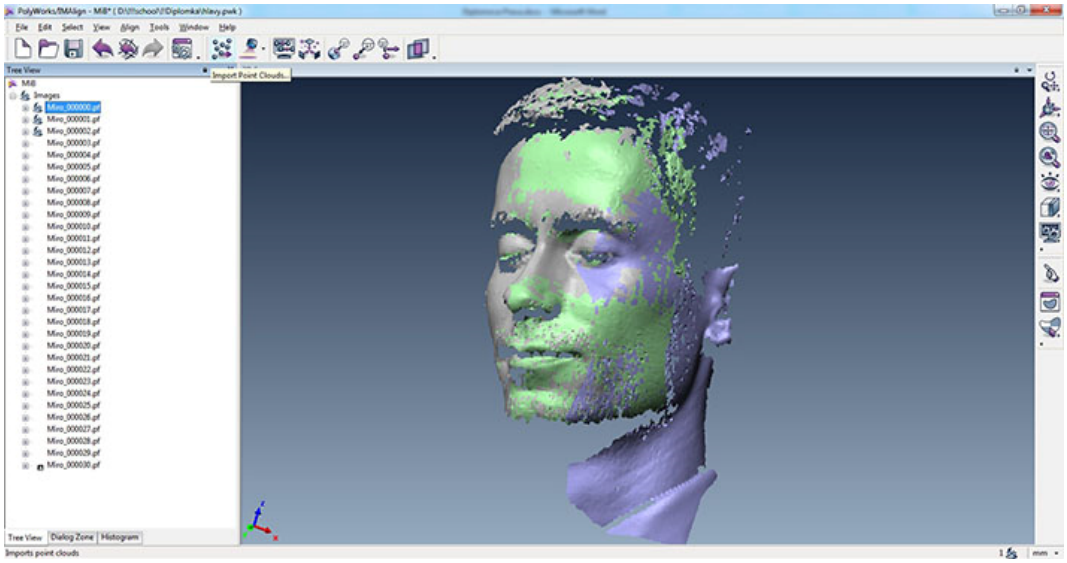


Fig. 3.5 – Varias fotos escaneadas crean una cara

Fig. 3.6 – Modelo de cara final 3D

Sólo se necesita la cara de la imagen capturada. El cultivo de la cara es necesario. Cada cara se encuentra en el rectángulo, que consta de 4 puntos en la cabeza. Los bordes laterales constan de puntos cuya posición está más a la izquierda y a la derecha. El punto más alto hace que el borde superior y el borde inferior estén en el punto más bajo. A continuación, el cultivo se basa en este rectángulo hecho por estos 4 puntos.

Los datos capturados son pre-procesados posteriormente el uso de algoritmos de extracción de características.

El objetivo de la extracción de características es extraer la información compacta de las imágenes que son relevantes para distinguir entre caras de diferentes personas y que es estable en términos fotométricos y ante variaciones geométricas en las imágenes.

Como características se pueden utilizar puntos faciales (la cabeza, la frente, los ojos, la barbilla, la nariz, la barbilla, la boca, etc.) y las distancias entre estos puntos seleccionados en el espacio Euclídeo 3D (Fig. 3.7).

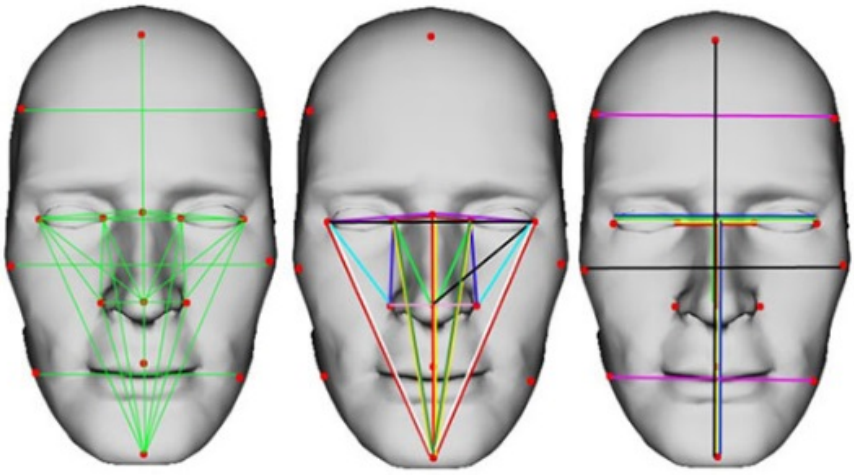


Fig. 3.7 – Ejemplo de características faciales

4.3 Aplicaciones de reconocimiento facial 3D

El reconocimiento facial 3D puede ser también utilizado en muchas aplicaciones como el acceso seguro a los sistemas o la identificación por la televisión inteligente y permite ir de compras en línea (es decir, se puede permitir solamente a los padres y no a los niños, etc.).

La tarea de reconocimiento facial 3D requiere como reconocimiento facial 2D una entrada de una cámara. Para el reconocimiento 3D, la cara y la superficie facial 3D deben ser capturadas. El principal proceso de reconocimiento facial consiste en los siguientes subprocesos:

- **Captura de superficie facial 3D** – Hay varias maneras de lograr esta tarea, por ejemplo las cámaras estéreo, láser o cámaras de profundidad (es decir, del sensor Kinect), etc.
- **Pre-procesado** - los datos capturados son pre-procesados posteriormente
- **Extracción de características** - el propósito de la extracción de características es extraer la información compacta de las imágenes que son relevantes para distinguir entre las caras de diferentes personas y es estable en términos fotométricos y ante variaciones geométricas en las imágenes
- **Medida de distancia -**el último paso del reconocimiento de la cara 3D es la medición de la distancia entre la cara 3D de la prueba y caras 3D almacenadas dentro de la base de datos. Existen varias técnicas para medir la distancia. El método más simple es la medición de distancias a nivel local y mundial de dos caras donde se necesita usar correctamente elementos muy precisos para determinar los puntos faciales (como ojos, nariz, boca, mentón, orejas, etc.) y medir sus distancias dadas por las métricas establecidas. Los métodos más sofisticados son clasificadores del vecino más cercano, técnicas que incluyen máquinas de vectores soporte, etc.

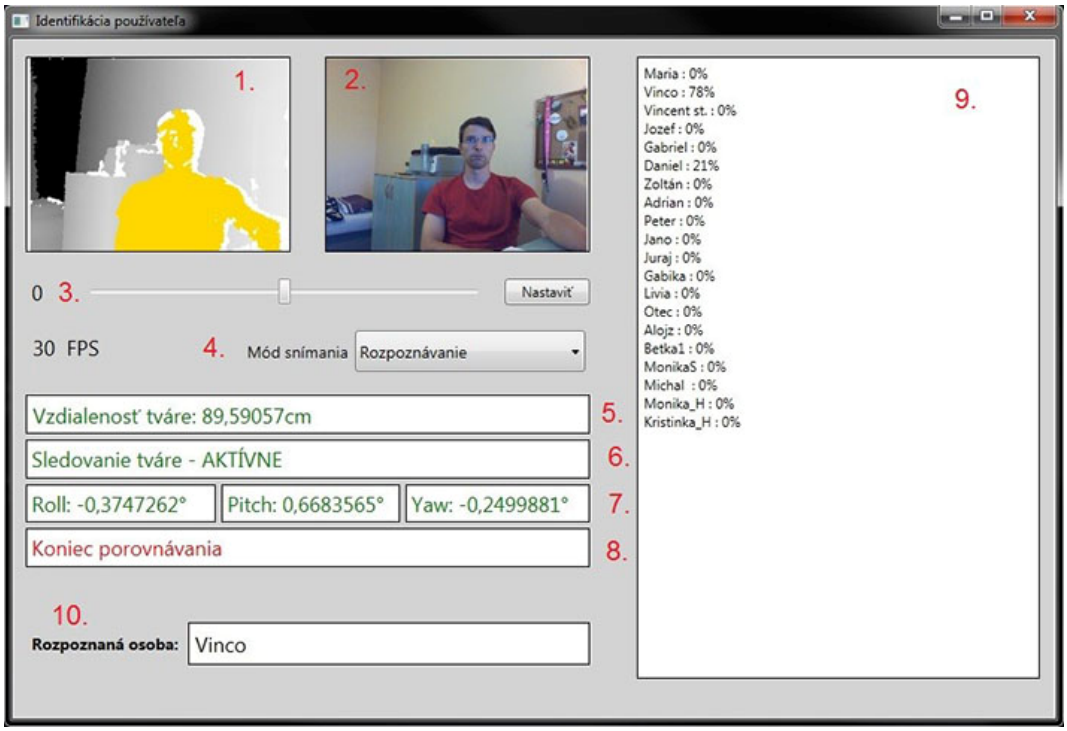


Fig. 3.8 – Ejemplo de GUI para reconocimiento facial 3D

5 Autenticación

Se propone el sistema de seguridad de acceso con la obligación de permitir el acceso sólo a usuarios autorizados cuya identidad pueda verificarse antes. Hay esencialmente tres pasos distintos, a saber, la identificación, autenticación y autorización [6].

Identificación – el usuario es identificado por la cadena de fichas o de identificación (número de teléfono o dirección de correo electrónico)

Autenticación –Una vez aceptada cadena de identificación o ficha, el usuario tiene que probar su identidad.

Autorización –Permitir o no permitir el acceso del usuario al contenido solicitado o para un conjunto de acciones en el marco basado en sus derechos de acceso.

El sistema puede autenticar a los usuarios con base al supuesto de que los usuarios saben algo (memometrics), reconocer algo (cognometrics), si tiene algo que es característico para cada persona (biometría). En las tres formas se comparten secretos del sistema y del usuario (es decir, clave de autenticación).

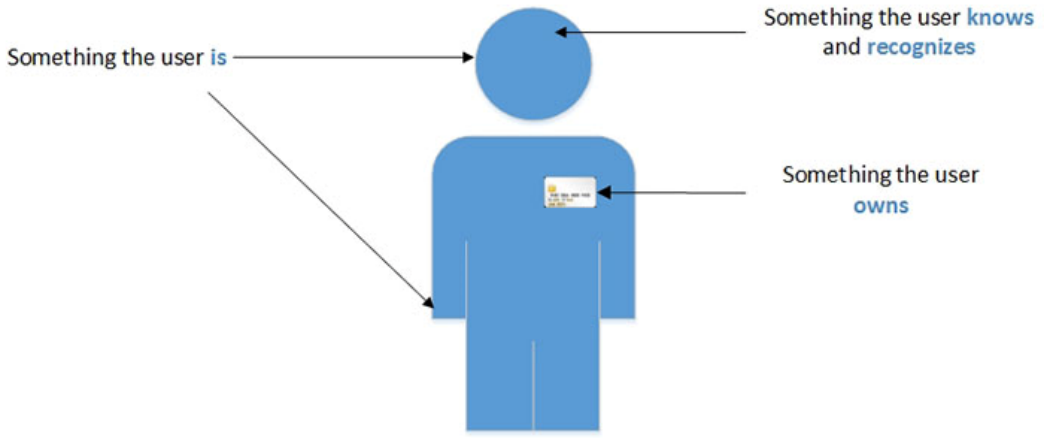


Fig. 4.1 – Opciones de autenticación de usuario

5.1 Tipos de mecanismos de autenticación

Sobre la base de los tipos de autenticación, se pueden enumerar los siguientes grupos.

Biométricos

La biometría es la comparación de las características anatómicas, fisiológicas y de comportamiento de una persona. Los mecanismos de autenticación biométrica se dividen en dos categorías básicas:

- **Biometría conductual** - basada en los movimientos, es decir, el usuario maneja el ratón del ordenador, la latencia, o la dinámica de pulsaciones de teclas o la dinámica de la firma.
- **Características fisiológicas** - basado en huellas digitales, voz, características de la cara, la mano o la geometría del dedo o incluso la forma de la oreja del usuario.

Es difícil comparar las tecnologías biométricas. Cada una tiene un rango diferente de precisión, fiabilidad y facilidad de uso. Un caso de uso de la técnica biométrica es la detección de rostros. Otros métodos requieren una posición específica del cuerpo para el sensor (detección del iris), y son por lo tanto menos cómodos de usar, aunque se pueden obtener resultados más precisos.

Memométrica

Este tipo de mecanismo de autenticación se basa en la generación de secuencias aleatorias de letras o números, llamados contraseña. Un caso es la palabra o el PIN si es una expresión numérica o frase de contraseña si contiene más de una palabra. Las contraseñas pueden ser también una forma semántica.

Tipos de Password o contraseñas:

- password aleatorio– el tipo más popular de la autenticación con un alto nivel de seguridad [7].
- password semántico – se basan en proceso deductivo, se hacen preguntas al usuario con el objetivo de obtener una respuesta precisa (Fig. 4.2) [8].

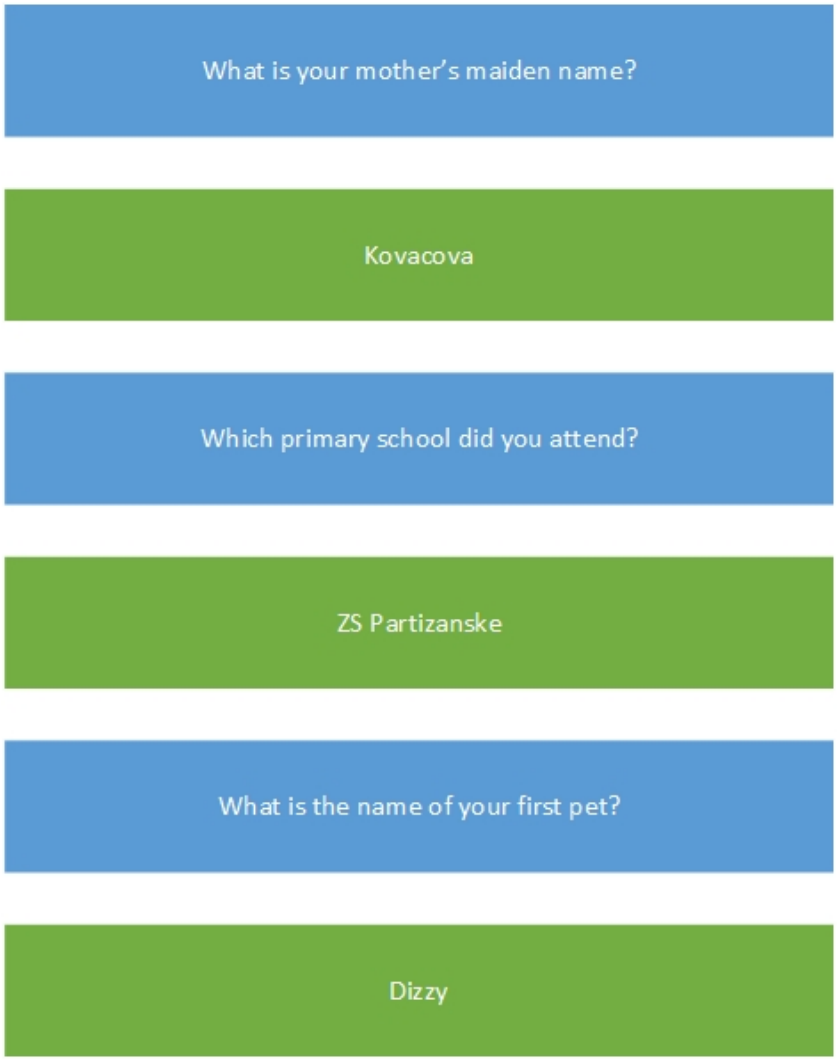


Fig. 4.2 – Principio básico de la contraseña semántica

Cognometría

La idea gráfica de autenticación se basa en la memoria visual del usuario. Los estudios científicos apuntan al hecho de que el ser humano tiene una enorme y prácticamente ilimitada posibilidad para recordar las imágenes [9].

Los códigos gráficos están ganando popularidad, especialmente en el caso de las tecnologías móviles o tabletas, es decir, para desbloquear el teléfono móvil. Hay dos principios fundamentales:

- **códigos gráficos basados en el reconocimiento** - el usuario selecciona la imagen de destino entre la cantidad de elementos perturbadores de la escena. Este enfoque se basa puramente en la memoria visual. El objetivo es reconocer un objeto visto previamente entre otros.
- **códigos gráficos basados en la posición** – el usuario, con este principio, debe dibujar un patrón, por lo general en la red, lo que requiere memoria visual-espacial y movimiento exacto.

Propiedad

La autenticación puede basarse en algo que un usuario posee. Este objeto es simbólico. Un buen ejemplo es el SecureID token de seguridad RSA en la Fig. 4.3. [15]

Fig. 4.3 – Ejemplo de Token example: SecureID – seguridad RSA

Símbolo o token a través de una función criptográfica que combina la cerradura y una clave secreta, crea un código numérico que aparece en la pantalla LCD. Para autenticar el número de tipo de usuario de SecurID. El servidor de autenticación también conoce la clave secreta almacenada en la señal del usuario, así como la hora y la fecha. Con base a este conocimiento el servidor de autenticación realiza las mismas funciones criptográficas. Para que la autenticación tenga éxito, el valor generado debe coincidir con el valor que se introduce por el usuario.

Otro tipo de token de autenticación es el que tiene la interfaz **(Universal Serial Bus) USB**.

Los Tokens se proporcionan como *software* (**SW**) o *hardware* (**HW**).

La principal desventaja del token HW es que el usuario tiene que llevar siempre la misma.

Las fichas de SW se almacenan en los usuarios de PC o portátil. En este caso, el usuario puede acceder al sistema sólo desde el PC donde se almacena el token.

5.2 Factores humanos en el proceso de autenticación

Varios escenarios de autenticación utilizan métodos de encriptación de clave pública (criptografía de clave pública). Por ejemplo, un usuario tiene una tarjeta inteligente, que lleva la correspondiente clave pública y una clave privada. Durante el proceso de autenticación de usuario, el sistema envía un desafío al azar. El usuario firma el reto con su clave privada y envía el resultado. El sistema verifica la firma con una clave pública. De esta manera, el sistema puede verificar que el usuario posee la clave privada correcta sin la necesidad de aceptar su llave. En lugar de almacenar la clave pública en un archivo en el sistema remoto, la tarjeta inteligente puede presentar un desafío firmado en el certificado de clave pública, firmado por un tercero. Esto se llama infraestructura de clave pública (PKI) estándar y se basa en las especificaciones del UIT-T.

La Fig. 4.4 muestra las entidades involucradas en el proceso de autenticación. En cada paso de este proceso, un potencial atacante puede obtener acceso a la clave de autenticación.

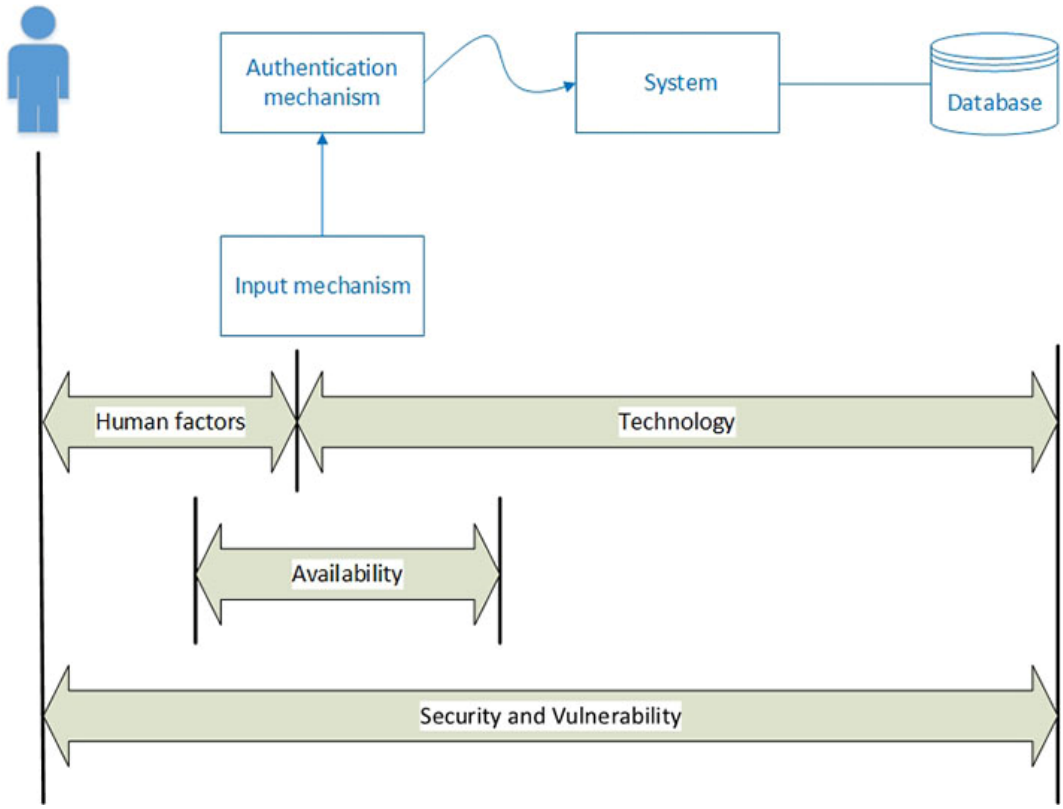


Fig. 4.4 – Entidades involucradas en el proceso de autenticación

La zona más frágil es el dispositivo de entrada y el usuario. Si la autenticación se basa en el conocimiento (contraseñas, PIN, etc.), el usuario tiene que recordar la clave secreta. Recordar contraseñas es difícil para muchas personas, a menudo comparten su contraseña con alguien o la escriben en un papel en la oficina.

La seguridad no se puede resolver solamente con el hardware, ya que los usuarios son una parte del proceso de autenticación [10].

6 Autorización

Autorización significa la verificación de la persona en la entrada (a la red o servicio), basada en derechos de acceso. Además, se define a qué información se puede acceder y qué acciones se pueden realizar por el usuario identificado y autenticado.

6.1 Modelo de autorización

Se utilizan modelos de autorización (Fig. 5.1) para el control de las normas de acceso al sistema (u objeto) y sus servicios, definidos por el sistema de seguridad. Los modelos de autorización básicos son [11]:

- *Discretionary Access Control* (**DAC**) – permite que el sistema (u objeto) propietario define quién puede o no puede acceder al sistema
- *Mandatory Access Control* (**MAC**) – el acceso de los usuarios se define a través de las clasificaciones.
- *Role-Based Access Control* (**RBAC**) – son los más utilizados. Los usuarios se dividen en grupos con un papel definido. El usuario puede acceder al sistema en base a su papel.
- **Task Based Access Control** (**TBAC**) – en este modelo se usa el número de acceso del usuario al sistema. Si se alcanza el valor definido, se rechaza el siguiente acceso.
- **Attribute based Access Control** (**ABAC**) – para controlar los usuarios de acceso se utilizan atributos.

Todos los modelos mencionados se pueden combinar.

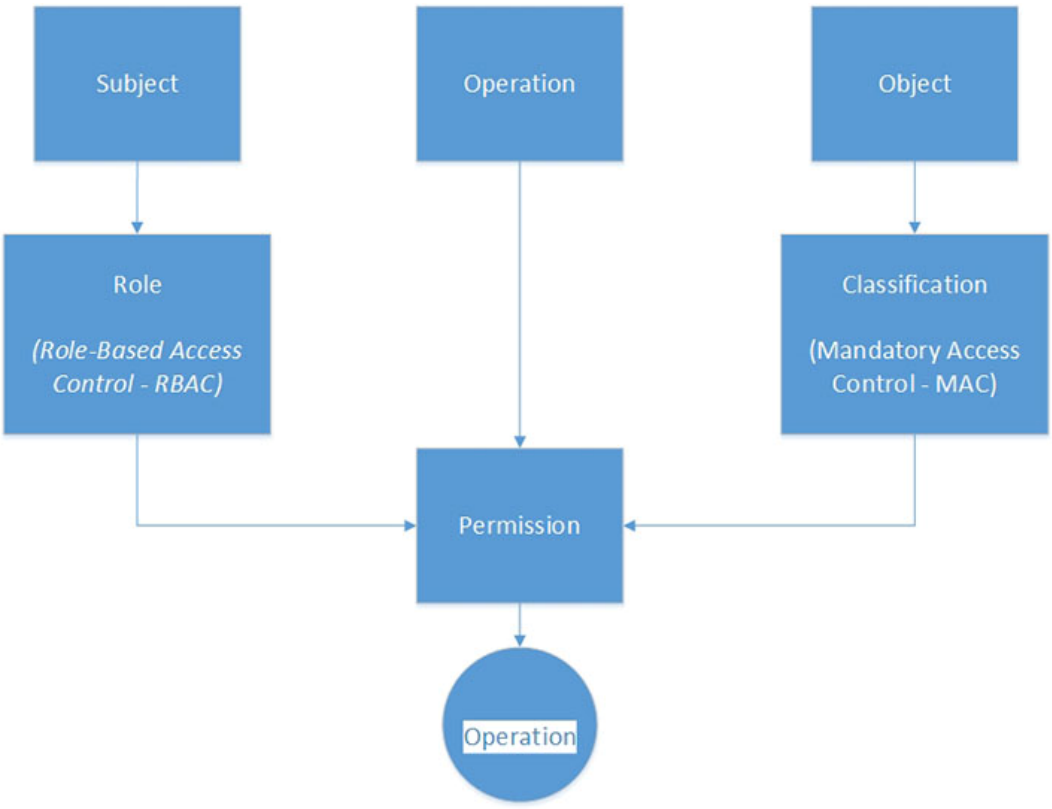


Fig. 5.1 – Modelo de autorización

6.2 Reglas de gestión de acceso

Una de las técnicas más comunes de control de acceso (Fig. 5.2) es la matriz de acceso. Las filas de la matriz representan opciones de usuario y las columnas representan los objetos de usuario. Esta técnica se denomina a menudo como lista de control de acceso (ACL) [12].

El control de acceso que depende del contenido es una técnica nueva en la que un usuario puede acceder a información más detallada o a objetos de datos como un usuario diferente. Esta decisión puede depender de factores tales como la edad, el terminal utilizado, acceder a un punto, la dirección IP de acceso del usuario y el tiempo.

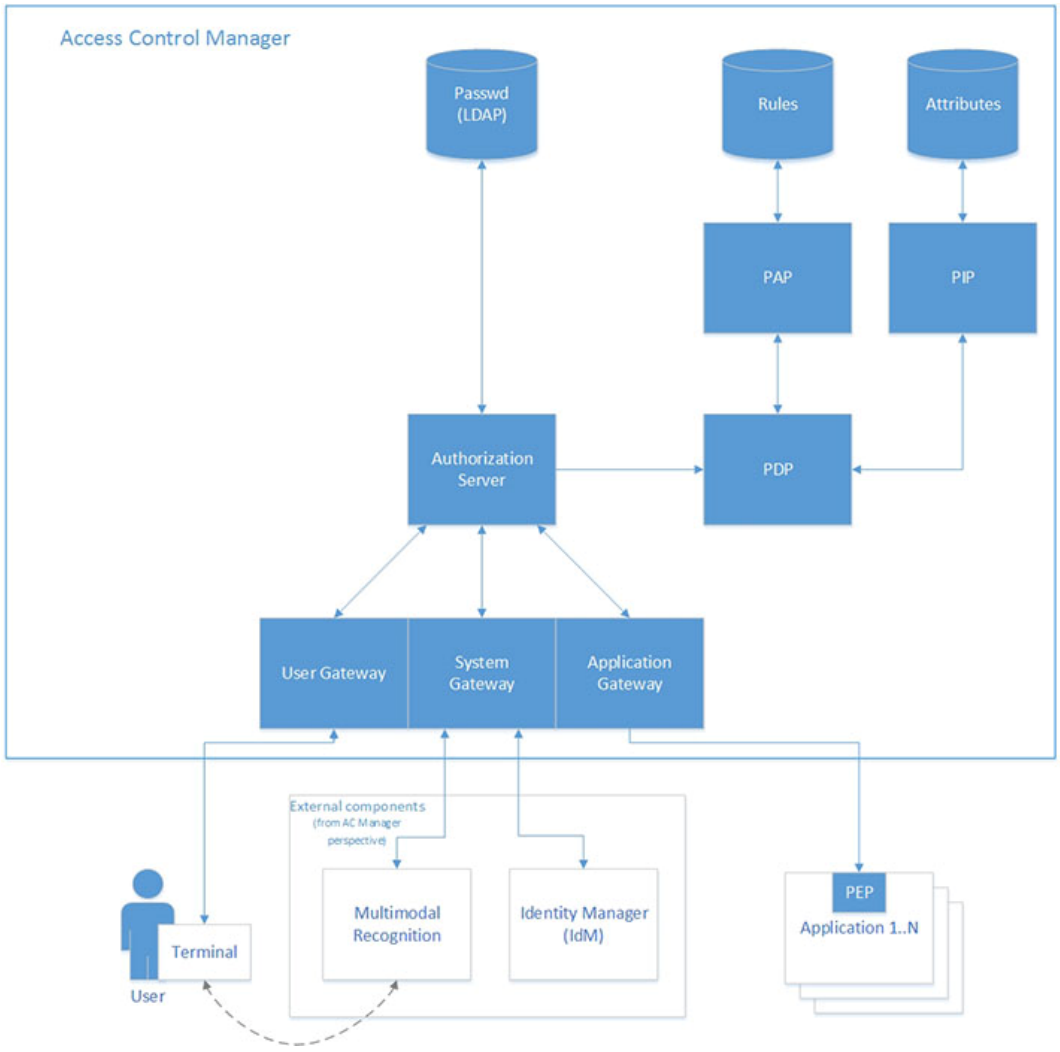


Fig. 5.2 – Modelo de autorización

6.3 Derechos o privilegios de acceso

El proceso de decisión cuando se recibe una solicitud de acceso a un sistema en particular, la aplicación o el contenido de información de la aplicación, en ciertas etapas, dependerá de los derechos de acceso, dispuestos en un archivo de autorización. Las reglas de asignación se basan en los modelos descritos en la sección 5.1 Autorización de Modelo.

En el sistema se utiliza el modelo RBAC y se definen tres funciones:

- administrador
- propietario del grupo
- usuario del grupo

El administrador asigna un propietario o derechos de acceso a los usuarios a las aplicaciones en el sistema. El propietario del grupo también puede asignar derechos de acceso para cada usuario para aplicaciones específicas en el sistema. Si el administrador ha asignado previamente al propietario del grupo los derechos de agregar, modificar y borrar el contenido de una aplicación particular, el propietario del grupo puede ceder aún más esos derechos a un usuario. En ese caso el usuario también puede convertirse en un colaborador de contenido, es decir, también puede actuar como titular de los datos. Un ejemplo de este tipo de aplicaciones es un servicio para contenido multimedia compartido.