

# LA EVOLUCIÓN DEL RETRATO HABLADO: DEL LÁPIZ Y EL PAPEL A LOS ALGORITMOS GENÉTICOS\*

○ Luis Fernando Cuevas Remigio\*\*, Katya Rodríguez Vázquez\*\*\*, Arodi Farrera\*\*\*\*, Sergio Padilla Renaud\*\*\*\*\* y Germán Palafox Palafox\*\*\*\*\*

\* Los autores agradecen el apoyo otorgado por el proyecto PAPI-IT IN-101620, UNAM.

\*\* Universidad Nacional Autónoma de México

\*\*\* Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México.

\*\*\*\* Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México.

\*\*\*\*\* Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México.

\*\*\*\*\* Facultad de Psicología, Universidad Nacional Autónoma de México.

## PALABRAS CLAVE

## KEYWORDS

○ **Retrato hablado**

*Facial composite*

○ **Rostro**

*Face*

○ **Programación evolutiva**

*Evolutionary programming*

○ **Algoritmos genéticos**

*Genetic algorithms*

○ **Identificación**

*Identification*

**Resumen.** El retrato hablado es una técnica frecuentemente utilizada en la investigación policial. Sin embargo, la investigación de laboratorio indica que el nivel de identificación de un rostro, a partir de un retrato hablado, es bajo. Una de las razones de esto es que el retrato hablado tradicional no se aproxima a los procesos de reconocimiento facial humano. En las últimas dos décadas se han creado sistemas de cuarta generación, basados en programación evolutiva con algoritmos genéticos, cuya principal característica es la combinación y evolución de rostros completos que gradualmente pueden converger en un rostro objetivo. El sistema Caramex II, basado en las características antropométricas del rostro de la población mexicana, pretende utilizar una aproximación evolutiva para construir retratos hablados.

**Abstract.** Facial composite is a technique frequently used in police investigation. However, laboratory research indicates that the level of identification of a face, from a facial composite, is low. One reason for this is that the traditional sketch does not approximate human facial recognition processes. In the last two decades there have been fourth generation systems, based on evolutionary programming with genetic algorithms, whose main characteristic is the combination and evolution of complete faces that gradually can converge on an objective face. The Caramex II system, based on the anthropometric characteristics of the face of the Mexican population, intends to use an evolutionary approach to construct facial composites.

Fecha de recepción: 30 de octubre de 2020

Fecha de aceptación: 11 de noviembre de 2020

## SUMARIO:

**I. Introducción. II. Los problemas del retrato hablado. III. Retrato hablado de cuarta generación. IV. Análisis de componentes principales. V. Evolución de rostros. VI. Softwares de cuarta generación. VII. Caramex. VIII. Caramex-II. IX. Conclusiones. X. Fuentes de consulta**

---

### I. INTRODUCCIÓN

El retrato compuesto forense, comúnmente conocido como retrato hablado, es la técnica por medio de la cual una víctima o testigo de un evento delictivo proporciona la descripción verbal de los rasgos faciales de un sospechoso a un experto en arte forense, quien intenta recrear el rostro de dicho sospechoso a través de un dibujo. Desde que se introdujo esta técnica como parte de la investigación policial, ha existido una serie de cambios o generaciones en la forma de realizar el retrato hablado. La primera generación comenzó a finales de siglo XIX, cuando diferentes departamentos de policía en Europa y en Estados Unidos solicitaron a retratistas o caricaturistas de periódicos su colaboración para dibujar los rostros de sospechosos de algún delito (Taylor, 2001). Esta primera etapa se extendió hasta la década de los cincuenta del siglo XX, en la cual se va profesionalizando y especializando la labor de los artistas forenses.

La segunda generación comenzó con la introducción de técnicas mecánicas para la construcción del retrato hablado, como los sistemas Identikit y Photofit (Davies y Valentine, 2007). Estos sistemas contaban con un catálogo de imágenes de diferentes rasgos faciales (ojos, narices, bocas, etc.), con el objetivo de que las víctimas o testigos de un delito seleccionaran estos rasgos, tomando como criterio su grado de semejanza con el de un sospechoso, y construir su rostro (Davies y Valentine, 2007). En el caso de Identikit, estos rasgos eran dibujos impresos en acetatos transparentes que se podían sobreponer para formar el rostro. Este sistema contaba con un lápiz especial de cera, con el cual se podían añadir elementos a la composición final, como marcas, lunares, arrugas o cicatrices para mejorar el parecido. En el caso de Photofit, eran imágenes de fotografías de rasgos faciales impresas en cartón con un recorte especial para que encajaran, a manera de un rompecabezas, y poder formar el rostro del sospechoso (Taylor,

2001). Esta segunda generación se extendió hasta la década de 1980, cuando aparecieron los primeros sistemas informáticos.

En la tercera generación destacan programas como Mac-Mug-Pro y E-Fit (Davies y Valentine, 2007), muy similares a los sistemas mecánicos, debido a que contaban con un catálogo de rasgos faciales que debían ser seleccionados por las víctimas o testigos, pero que contaban con la ventaja de poder utilizar algunas herramientas de edición de imágenes que permitían aumentar o disminuir el tamaño de los rasgos, aclarar u oscurecer la piel o agregar elementos, como arrugas o bigotes (Davies y Valentine, 2007). Esta tercera generación se extendió hasta el final del siglo XX.

Con el cambio de siglo y con el avance en la capacidad de las computadoras para procesar imágenes y el desarrollo de la inteligencia artificial, se diseñaron nuevas técnicas de construcción de retratos hablados. La característica principal de esta nueva generación se fundamenta en una aproximación *holística* o global de construir los retratos hablados, es decir, se deja de lado la descripción verbal de las víctimas o testigos y se prioriza una manera más natural o cercana a la percepción facial humana. Esto se realiza a través de la selección y combinación de rostros completos. Las víctimas o testigos que utilizan estos nuevos sistemas deben seleccionar rostros que, *grosso modo*, se asemejen al rostro de un sospechoso. Estos sistemas utilizan una programación con algoritmos genéticos (Holland, 1975; Goldberg, 1989), que les permiten combinar y evolucionar estos rostros hasta converger en un rostro objetivo.

---

## II. LOS PROBLEMAS DEL RETRATO HABLADO

A pesar de ser una de las técnicas periciales más empleadas en la investigación policial, el retrato hablado presenta una serie de problemas que poca atención ha recibido por parte de la autoridad (Davies y Valentine, 2007). Por ejemplo, la organización no gubernamental *Innocence Project* informó en un estudio que, de 367 casos de identificación incorrecta en Estados Unidos, el 69% implicó la declaración de un solo testigo, y el 27% de estos casos involucró el uso de un retrato hablado (Innocence Project, 2016). Además, la investigación psicológica de laboratorio sobre reconocimiento de retratos hablados muestra un desempeño pobre por parte de los participantes en diversos experimentos, para identificar un rostro conocido a partir de un retrato hablado. Por ejemplo, en su investigación, Davies (1986)

confrontó el desempeño de dos grupos de participantes para construir, con ayuda de expertos en arte forense o con Photofit, un grupo de rostros desconocidos para ellos. Los rostros fueron una serie de fotografías mostradas durante un minuto a los participantes. Los retratos hablados elaborados por los artistas forenses y con Photofit, fueron luego mostrados a otro grupo de participantes que sí conocían los rostros, a partir de los cuales se elaboraron los retratos hablados para que intentaran reconocerlos. Sus resultados mostraron que los retratos hablados hechos con arte forense fueron mejor identificados que aquellos elaborados con Photofit. Sin embargo, los retratos hablados hechos por artistas forenses, que mayor calificación de identificación obtuvieron, fueron aquellos elaborados en presencia del rostro, es decir, observándolo en todo momento, que aquellos hechos a partir de la memoria de los participantes.

En otra investigación parecida, pero utilizando esta vez Identikit, Laughery y Fowler (1980) hallaron nuevamente que los retratos hablados hechos con arte forense fueron mejor identificados que aquellos hechos con Identikit. Sin embargo, otra vez los retratos hablados que mayores puntajes de identificación obtuvieron fueron los hechos en presencia del rostro, en lugar de a partir de la memoria de los participantes. Por otro lado, Davies, van der Willick y Morrison (2000) compararon el sistema mecánico Photofit con el sistema informático E-fit para construir retratos hablados de rostros conocidos y desconocidos para sus participantes. Los retratos hablados fueron también realizados en presencia del rostro y a partir de la memoria de los participantes. Sus resultados indicaron que los retratos hablados elaborados en presencia del rostro y elaborados con E-fit fueron mejor identificados que aquellos hechos con Photofit. Por su parte, aquellos retratos elaborados de memoria obtuvieron bajos puntajes de identificación, independientemente del sistema utilizado, además de que no hubo diferencias significativas para los rostros conocidos o desconocidos. Otras investigaciones han reportado resultados similares (Brace, Pike, Allen y Kemp, 2006; Brace, Pike y Kemp, 2000; Davies y Oldman, 1999). De manera general, la tendencia que se encuentra en esta clase de investigaciones muestra que los niveles de identificación de un rostro a partir de un retrato hablado en condiciones ideales, es decir, construido en presencia del rostro, oscilan en alrededor del 20%, mientras que aquellos elaborados simulando situaciones forenses reales alcanzan menos del 5% (Zahradnikova, Duchovicova y Schreiber, 2016).

Existen dos principales factores que explicarían el bajo nivel de identificación de esta técnica. El primero de ellos es la forma en la que se construyen los retratos hablados a partir de rasgos individuales. La mayoría de los sistemas de construcción de retratos hablados requiere que las víctimas o testigos de un delito proporcionen una descripción verbal de un rostro o que seleccione, de entre un catálogo, aquellos rasgos faciales que más se asemejen al de un sospechoso. Sin embargo, existe abundante investigación psicológica que indica que la percepción de un rostro es *holística* o global; es decir, como una sola unidad perceptual, más que como un conjunto de rasgos individuales agrupados (Behrmann, Richler, Avidan y Kimchi, 2015; Tanaka y Farah, 1993; Tanaka y Simonyi, 2016). Algunas investigaciones que apoyan esta idea son las que han acuñado el término *efecto de inversión del rostro* (Boutsen y Humphreys, 2003; Carbon y Leder, 2005; Thompson, 1988; Valentine, 1988), el cual se refiere a la tendencia de las personas a no percibir grandes alteraciones en la imagen de un rostro (como los ojos y la boca volteados), cuando este se presenta de manera invertida. Solo hasta que se coloca en posición vertical, con la cabeza hacia arriba, es cuando las personas pueden percibir lo grotesco de la apariencia de este rostro.

La explicación que proporcionan los investigadores para este efecto indica que, cuando un rostro se muestra de forma invertida, se interrumpe la capacidad de percepción holística y se utiliza un procesamiento individual de cada rasgo facial. Sin embargo, este procesamiento por rasgos no es del todo eficiente, por lo cual no permite detectar los rasgos alterados de un rostro en forma invertida. Otra evidencia del procesamiento holístico proviene del denominado “efecto de composición del rostro” (Young, Hellawell y Hay, 2013), que se refiere a la tendencia de las personas a percibir como un solo rostro a la imagen compuesta por dos mitades, una superior y otra inferior, de dos rostros distintos. En algunos experimentos de tiempo de reacción se ha encontrado que, cuando se utilizan rostros de celebridades que se combinan entre sí para formar rostros compuestos, los participantes tienen una mayor latencia para identificar a las celebridades a partir de la mitad superior, cuando estas mitades están completamente alineadas. Sin embargo, cuando las mitades se muestran desalineadas, el tiempo de respuesta disminuye (Hole, 1994; Rossion, 2008; Young, Hellawell y Hay, 2013). Además, este efecto desaparece cuando los rostros compuestos se muestran de manera invertida (Hole, 1994, Rossion, 2008).

Otro de los efectos que apoya el procesamiento holístico es el denominado “efecto del todo sobre sus partes” (Tanaka y Farah, 1993), que se refiere

a la tendencia a reconocer, de manera más precisa, un rasgo facial, como una nariz, cuando esta se encuentra en el contexto de un rostro completo que cuando se muestra de manera aislada, como en los catálogos de rasgos faciales de los sistemas mecánicos de retrato hablado. Por ejemplo, Tanaka y Farah (1993) le solicitaron a un grupo de participantes que intentaran memorizar una serie de rostros. Luego, evaluaron la memoria de sus participantes para determinar si podían reconocer tres rasgos faciales de los rostros aprendidos (nariz, boca y ojos) en dos condiciones distintas: rostro completo y rasgo aislado. En la condición de rostro completo se presentaba un par de rostros iguales, a excepción de un rasgo distinto, por ejemplo, la nariz, mientras que, en la condición de rasgo aislado, solo se presentaban dos rasgos, por ejemplo, dos narices. La tarea de los participantes era identificar, en ambas condiciones, el rasgo de los rostros aprendidos. Sus resultados mostraron que los participantes fueron más precisos para reconocer un rasgo cuando este se presentaba en el contexto de un rostro completo que de manera aislada.

El segundo factor relacionado con el bajo nivel de identificación de un retrato hablado es el componente verbal de aquel. Muchas investigaciones (Dodson, Johnson y Schooler, 1997; Fallshore y Schooler, 1995; Ryan y Schooler, 1998; Schooler y Engstler-Schooler, 1990; Sporer, 2007) sugieren que describir un rostro desconocido que se ha observado brevemente puede afectar su posterior reconocimiento. Por ejemplo, Schooler y Engstler-Schooler (1990), en uno de sus experimentos, mostraron a un grupo de participantes un video de 30 segundos, en el cual aparecía una persona asaltando un banco. Luego de observar este video, la mitad de sus participantes llevó a cabo una tarea de lectura durante cinco minutos. A la otra mitad se le pidió que proporcionara una descripción por escrito, lo más detallada posible, del rostro del asaltante durante cinco minutos. Al término de ambas tareas, los dos grupos fueron sometidos a una tarea de reconocimiento, en la cual tenían que tratar de identificar el rostro del asaltante del video de entre otros ocho rostros distractores. Sus resultados mostraron que los participantes que proporcionaron la descripción detallada del rostro del asaltante pudieron identificarlo correctamente en un 38%, mientras que los participantes de la actividad de lectura pudieron reconocerlo en un 64%. Los investigadores denominaron a este efecto “ensombrecimiento verbal”, y consideraron que se presenta debido a que la memoria verbal de un rostro interfiere con la memoria visual de este; como no existen las palabras lo suficientemente precisas para describir la forma y el tamaño de

los rasgos faciales, esta descripción puede alterar la representación visual original del rostro por otra sesgada y menos precisa. Por su parte, Wells, Charman y Olson (2005) demostraron un efecto similar durante la construcción de retratos hablados.

---

### III. RETRATO HABLADO DE CUARTA GENERACIÓN

En las últimas dos décadas se ha desarrollado una serie de *softwares* de cuarta generación que intentan abatir las limitaciones de los métodos tradicionales de construcción de retratos hablados (Frowd *et al.*, 2015a; Frowd, Erickson, Lampinen *et al.*, 2015b; Davies y Valentine, 2007). Estos sistemas, principalmente Evo-Fit y EFIT-V (Frowd *et al.*, 2015; Valentine, 2010; Zahradnikova, Duchovicova y Schreiber, 2018), utilizan rostros completos para la construcción de los retratos y una programación con algoritmos genéticos (Holland, 1975; Goldberg, 1989). Estas características permiten, en primer lugar, que los testigos seleccionen rostros completos, tomando en cuenta si presentan rasgos similares a los del rostro de un sospechoso. Por otro lado, y a través de su programación con algoritmos genéticos, estos *softwares* combinan los rostros seleccionados por los testigos para gradualmente converger en un retrato, en teoría, mucho más parecido al rostro del sospechoso. De acuerdo con los investigadores (Frowd *et al.*, 2015a; Frowd, Erickson, Lampinen *et al.*, 2015b; Davies y Valentine, 2007; Valentine *et al.*, 2010; Zahradnikova, Duchovicova y Schreiber, 2016), la metodología que emplean estos *softwares* hace que la descripción verbal de un rostro sea innecesaria durante la construcción del retrato y la selección y combinación de rostros completos sea más cercana al reconocimiento facial humano.

La teoría psicológica de la que parten estos *softwares* se basa en la percepción holística o global del rostro, y que la representación mental de cada uno de estos, que se guarda en la memoria, está organizada en un *espacio facial multidimensional* (Valentine, 1991, 2002; Valentine, *et al.*, 2010). De acuerdo con esta teoría, cada rostro es almacenado en términos de sus valores (o vectores) en un amplio rango de dimensiones faciales. Por ejemplo, una persona ha observado una cantidad considerable de rostros a lo largo de su vida y ha podido percatarse que existen narices que son muy cortas o largas, además de otras que estarían entre esos extremos. Según la teoría del espacio facial, esta característica de las narices podría representarse en la memoria como un continuo o una dimensión. Si además se considera

otro rasgo facial, como la separación entre los ojos, se podría establecer otra dimensión; es decir, hay rostros cuyos ojos parecen estar muy juntos o separados, además de rostros con valores entre estos extremos. De esta manera, se podrían agregar más valores o dimensiones de rasgos faciales, como el tamaño de las orejas, la anchura de la boca, lo largo de la cara, la edad aparente, etc. Según esta teoría, todos estos valores faciales estarían almacenados en la memoria en un espacio multidimensional. De esta manera, cuando una persona considera que dos rostros son muy parecidos entre sí, sería porque sus vectores dimensionales son muy similares.

#### IV. ANÁLISIS DE COMPONENTES PRINCIPALES

La teoría del espacio facial puede ser modelada matemáticamente a través de un método denominado “análisis de componentes principales” (PCA, por sus siglas en inglés), que es la base para los *softwares* de cuarta generación. Un rostro está compuesto por una gran cantidad de rasgos faciales (ojos, nariz, boca, lunares, etc.). La variabilidad que existe en cada uno de estos rasgos agrega una dimensión al espacio facial multidimensional. Entre más rasgos consideremos, mayor será el número de combinaciones posibles que podríamos tener para formar un rostro y, por tanto, más se complica el análisis. Cuando se trabaja con imágenes, esta dimensionalidad es aún mayor, pues, en lugar de tener una serie de rasgos, estamos trabajando con miles de píxeles, y cada uno de ellos cuenta como una dimensión; por lo que un conjunto cualquiera de imágenes (no es la cantidad de imágenes, sino la cantidad de variables lo que incrementa la dimensionalidad) representa un espacio multidimensional de información que requiere grandes demandas de memoria.

El PCA se utiliza para reducir la variación de dicho espacio de información a unas cuantas variables nuevas, conocidas como componentes principales. Cada una de estas nuevas variables es una combinación lineal de las variables originales; es decir, cada componente principal toma en cuenta todas las variables originales a la vez. Por ello, se le podría considerar como una representación holística del rostro (es decir, en lugar de centrarnos en las variables por separado, las concentramos todas en un componente principal). Este análisis se utiliza entonces para reducir a unos cuantos componentes o conjunto de dimensiones, conocidos como *eigenvalues* (o *eigenfaces* en este contexto), una serie de imágenes de rostros (Sirovich

y Kirby, 1987; Turk y Pentland, 1991). Cada *eigenface* es holístico, es decir, contiene una representación “resumida” de la imagen facial a partir de la cual se generó. Cuando se toma una muestra amplia de rostros y se les aplica un PCA, algunos de estos componentes principales pueden interpretarse como si codificaran alguna característica particular, como el sexo, la textura de la piel, el color del cabello, la edad, entre otros. Además, cuando se realiza un PCA, los *eigenfaces* resultantes se codifican como una serie de pesos o valores (*weights*), los cuales pueden ser recombinados para generar rostros nuevos (Craw y Cameron, 1991, 1992). Sin embargo, cuando se recombinan tales pesos para formar rostros nuevos, estos pueden parecer desproporcionados, amorfos o difuminados (Troje y Vetter, 1996). Para evitar esta situación, Craw y Cameron (1991) propusieron el análisis separado para la forma y la textura de los rostros.

Primero, la forma se obtiene a través de una serie de *landmarks* o marcadores de referencia que se colocan sobre el rostro, para indicarle al programa la posición de los rasgos faciales. Estos *landmarks* están codificados en valores de coordenadas  $x$  e  $y$ . Este método genera que el análisis de los rostros que lleva a cabo el PCA se estandarice a una forma facial única (denominada *shape-free*), donde los *landmarks* se localizan en la misma posición en cada una de las imágenes de la muestra de rostros. Generalmente, esta forma facial única es el promedio de la forma del conjunto de los diferentes rostros que constituyen la muestra (es decir, *morph*). Por su parte, la textura está representada por la escala de grises o la información de color de la imagen de los rostros. Cada *landmark* proporciona información de la textura dependiendo de su posición. De esta manera, diferentes rostros presentarán una variación distinta en la información de textura. Posteriormente, textura y forma pueden ser combinados con la aplicación de otro PCA como parte de un algoritmo denominado *active appearance model*, que proporciona un solo conjunto de parámetros compactados de modo óptimo para esta información (Cootes, Edwards y Taylor, 1998; Cootes y Taylor, 2001). Uno de los problemas con esta metodología es que no es posible reconstruir adecuadamente el tipo de cabello. Para ese fin, algunos sistemas recurren a un catálogo de imágenes con diferentes estilos y cortes de cabello, que pueden ser colocados al término de la construcción del retrato.

## V. EVOLUCIÓN DE ROSTROS

El PCA se puede combinar con algoritmos evolutivos para converger en la imagen de un rostro objetivo (*target face*). Los algoritmos evolutivos son llamados así porque emplean dos conceptos fundamentales de la evolución: variación aleatoria (mutación) y selección. La construcción de retratos compuestos comienza con la generación de un conjunto aleatorio de imágenes faciales (artificiales) dentro del espacio facial del PCA. Posteriormente, el usuario selecciona los rostros que considera más similares a los de un rostro objetivo. En el primer conjunto de rostros que muestra el espacio facial habrá un amplio rango de rostros, y ninguno necesariamente se parecerá al rostro objetivo. Sin embargo, las selecciones de rostros que haga el usuario servirán para “engendrar” un nuevo conjunto de rostros, introduciendo mutaciones del rostro o rostros “padre”. El proceso se repite hasta que las siguientes “generaciones” de rostros comiencen a parecerse entre sí, y termina cuando el usuario considera que ya no puede elegir entre las nuevas generaciones de imágenes, porque todos reproducen igual de bien el rostro objetivo, o bien, porque el proceso fracasó en reproducirlo.

Gibson, Pallares-Bejarano y Solomon (2003) identifican tres tipos de algoritmos evolutivos indispensables para la construcción de retratos hablados:

1. *Scale rating*. En este algoritmo, todas las imágenes, en cada generación, son calificadas en una escala de similitud respecto del rostro objetivo. Los rostros mejor calificados son seleccionados para engendrar la siguiente selección, habilitando tanto el cruce (*crossover*) como la mutación (Hancock, *et al.*, 2000).
2. *Select Multiple Mutate*. En este algoritmo, el testigo selecciona el rostro más parecido, el cual es reproducido con mutaciones aleatorias en todos los rostros de la siguiente generación, excepto en uno (Tredoux *et al.*, 1999).
3. *Follow the leader*. Este algoritmo muestra una serie de rostros nuevos con el mejor parecido posible, alcanzado por el sistema respecto al rostro objetivo. El testigo debe, entonces, simplemente elegir aquellos rostros que, a su juicio, reproducen de mejor manera el rostro objetivo. Este proceso se repite para alimentar a las siguientes generaciones y obtener un registro histórico de evolución, que sirve de guía para el proceso.

La ventaja de este procedimiento para construir retratos forenses es que permite hacer modificaciones globales u holísticas de un rostro, el cual se aproxima mucho más al proceso de reconocimiento facial humano, además de prescindir por completo de la descripción verbal. Sin embargo, algunos testigos señalan que ciertos rasgos del rostro, en su etapa de evolución final, requieren la modificación de un rasgo específico. Esto no es posible a través de esta metodología. Para ello, se utilizan programas de edición de imágenes, como Photoshop, pero trabajando con la configuración global del rostro, y no con partes aisladas.

---

## VI. SOFTWARES DE CUARTA GENERACIÓN

Existen hasta el momento cuatro sistemas de construcción de retratos hablados de cuarta generación. Estos *softwares* utilizan, de manera general, el procedimiento descrito en la sección anterior, pero difieren en el empleo de algunos algoritmos para optimizar su tarea. El primero de ellos fue desarrollado en la Universidad de Ciudad del Cabo (Tredoux *et al.*, 2006) y se denomina ID. Este *software* utiliza un algoritmo denominado PBIL (*population based incremental learning*), el cual guía la búsqueda de un rostro objetivo dentro de su espacio facial de la siguiente manera:

1. El sistema genera una pequeña muestra aleatoria de rostros, los cuales son mostrados al usuario.
2. El usuario selecciona el rostro que más se asemeje a un rostro objetivo.
3. El sistema usa los coeficientes del rostro seleccionado como una nueva *semilla* o punto de partida para generar nuevos rostros que mostrar al usuario.

Sin embargo, el sistema requiere de 100 a 150 generaciones para producir un rostro aproximado al rostro objetivo. Por esta razón, y considerando que en un contexto real este número de generaciones es proclive a la fatiga y a terminar la búsqueda del rostro objetivo antes de converger en una aproximación adecuada, los investigadores introdujeron un nuevo algoritmo denominado *M-choice*, que optimiza y reduce el tiempo de búsqueda. Este algoritmo opera de la siguiente forma:

1. El sistema genera una pequeña muestra aleatoria de rostros, los cuales son mostrados al usuario.
2. El usuario selecciona una serie de rostros que son similares al rostro objetivo en un orden de similitud; es decir, la primera selección es el rostro más similar, la segunda es la segunda mejor similitud, etc. El número de rostros seleccionados puede variar de generación en generación.
3. El sistema calcula la media y la varianza de estas selecciones y las utiliza como base para crear nuevas generaciones de rostros para mostrar al usuario.

Tredoux *et al.* (2006) evaluaron este sistema en dos experimentos para determinar su eficiencia. En el primero se comparó el desempeño de los algoritmos PBIL y M-choice, utilizando la información de solo la textura o textura y forma para construir los retratos. En este experimento, un grupo de 30 participantes observó dos rostros durante 30 segundos. Posteriormente, se les explicó el funcionamiento de ID y cómo se construían los retratos con este sistema. Una vez comprendida esta parte, los participantes debieron construir los rostros observados de memoria. La mitad de los participantes utilizó el algoritmo PBIL, y la otra mitad M-choice. Una vez construidos los retratos, otro grupo de participantes tuvo que evaluar su parecido con el rostro a partir del cual se elaboraron. Para ello, se diseñó una tarea en la computadora, en la que se mostraban a la derecha los rostros originales y, a la izquierda, una serie de cuatro retratos (distractores) y un retrato construido con el sistema ID. Debajo de cada retrato se mostraba una escala de similitud de 7 puntos, donde 1 representó *nada parecido* y 7, *muy parecido*. La tarea de los participantes fue observar las imágenes y determinar el grado de similitud utilizando la escala mostrada. Sus resultados arrojaron que, en todas las condiciones, excepto aquellas que utilizaban el algoritmo M-choice, con solo la información de textura, se obtuvieron puntuaciones de reconocimiento superior a las esperadas por azar. Además, se encontró que el algoritmo de M-Choice fue superior a PBIL cuando este podía utilizar la información de textura y forma para construir los retratos. En su segundo experimento, compararon el sistema ID con otro *software* de construcción de retratos hablados de tercera generación, basado en la selección de rasgos llamado FACES (IQ Biometrix). Sus resultados mostraron un desempeño superior de ID sobre FACES; sin embargo, los retratos construidos con ID no obtuvieron calificaciones muy altas de similitud por los participantes del experimento.

El otro sistema fue desarrollado por la Universidad Técnica de Estambul (Kurt *et al.*, 2006) y se llama INIH (*Interactive Nature-Inspired Heuristics*). Este *software* utiliza cinco diferentes algoritmos evolutivos para construir los retratos compuestos. El primero de ellos, *Interactive Generational Genetic Algorithm* (IGGA), emplea la siguiente estrategia:

1. Las nuevas generaciones de rostros son creadas a través de la selección que hace el usuario de los rostros mostrados por el sistema. Luego, estos se combinan y mutan.
2. El valor de ajuste (similitud) es obtenido directamente del usuario; es decir, cada rostro mostrado por el sistema es calificado por aquel con base en el grado de similitud de este con respecto a un rostro objetivo.
3. A través de una “competición” binaria basada en las calificaciones de similitud, se establecen los rostros “padres”.
4. Luego se aplica una *mutación gaussiana*; es decir, la asignación de un valor aleatorio basado en una distribución normal o de Gauss, para crear la nueva generación de rostros (*offsprings*).

*Interactive Steady-State Genetic Algorithm (ISSGA):*

1. La población ahora consiste en 3 rostros (dos padres y un descendiente).
2. El usuario debe entonces elegir el peor individuo, es decir, el rostro que menos se asemeje al rostro objetivo.
3. Entonces se genera un “hijo” como resultado de la recombinación de los genes de los rostros “padre”, y aplicando una *mutación gaussiana*.

*Interactive Evolutionary Strategy (IES):*

1. Las nuevas generaciones de rostros se crean a partir de las operaciones de cruce (*crossover*) y *mutación*.
2. Se eligen pares de padres para recombinación de manera aleatoria.
3. Se aplica una *mutación gaussiana*.
4. En cada etapa se selecciona un número de rostros de la población mostrada.

*Interactive Differential Strategy (IDS):*

1. En este caso, las *mutaciones* preceden a las operaciones de cruce.

2. La diferencia de los valores de pesos (*weights*) entre dos rostros candidatos se agrega a un tercero, para obtener otro rostro padre para una nueva recombinación genética.
3. En cada etapa se selecciona un número de rostros de la población mostrada.

#### *Interactive Particle Swarm Optimization (IPSO):*

1. La nueva población de soluciones potenciales se genera en cada iteración.
2. El usuario debe seleccionar, en cada etapa, la mejor opción de esta población.
3. En la etapa final, se muestran al usuario las mejores opciones de rostros seleccionadas, para elegir la mejor solución global.

Para evaluar este sistema, los investigadores desarrollaron un experimento, en el cual, a un grupo de participantes le mostraron una serie de rostros para luego construirlos de memoria, utilizando uno de los cinco algoritmos mencionados anteriormente. Una vez construidos estos retratos, se mostraron a otro grupo de participantes que conocían los rostros, a partir de los cuales se elaboraron, para que intentaran reconocerlos a través de proporcionar su nombre. Los resultados mostraron que los algoritmos ISSGA e IES fueron los que mayor porcentaje de reconocimiento obtuvieron. Sin embargo, el algoritmo ISSGA requirió de un mayor número de generaciones para llegar a una buena aproximación. Los investigadores reconocen que el sistema debe mejorarse para que sea competitivo con otros (Kurt *et al.*, 2006; Zahradnikova, Duchovicova y Schreiber, 2016).

El otro sistema, denominado inicialmente EigenFIT y luego E-FIT V, fue desarrollado en la Universidad de Kent (Gibson *et al.*, 2009; Solomon *et al.*, 2005), y se trató del primer *software* comercial de cuarta generación para construir retratos hablados. Este sistema estuvo originalmente basado en el algoritmo *Full Scale Rating Algorithm* (FSR), el cual operaba de la siguiente manera:

1. Cada individuo o rostro del espacio facial se calificaba con una escala del 1 al 10, con base en el grado de semejanza con respecto a un rostro objetivo.

2. Se utilizaban operaciones de cruce y mutación, basadas en las calificaciones proporcionadas por el usuario para engendrar a la siguiente generación.

Debido a la lenta velocidad de convergencia para alcanzar un retrato que se asemejara al rostro objetivo, se decidió implementar el algoritmo *Follow the leader* (FTL), el cual opera de la siguiente manera:

1. Se muestran al usuario dos rostros, con la indicación de que seleccione el mejor.
2. Se genera un descendiente (*offspring*) en cada iteración, como resultado de la selección de mejor semejanza por parte del usuario, además de un nuevo rostro de las iteraciones previas.
3. En esta implementación no existe una función de ajuste ni se preservan los mejores individuos para la siguiente generación.

Por último, los investigadores propusieron el algoritmo *Select Multiply and Mutate* como una estrategia que combinaba el porcentaje satisfactorio de convergencia y su simplicidad cognitiva. Este algoritmo opera de la siguiente manera:

1. Se muestran al usuario nueve rostros para que elija al mejor individuo.
2. Para evitar perder el mejor rostro o individuo una vez que se aplican los procesos evolutivos, este se copia a la siguiente generación sin cambios.
3. Luego, la nueva generación se compone de ocho individuos adicionales, clones del rostro mejor calificado.

En modificaciones posteriores (Solomon *et al.*, 2013), los investigadores lograron producir retratos de apariencia aceptable al rostro objetivo en 42 generaciones. Sin embargo, como el sistema cuenta con diferentes herramientas para modificar directamente los rostros, se puede alcanzar un retrato de apariencia aceptable en 25 iteraciones, en los cuales se muestran y se evalúan 225 rostros. Sin embargo, y a pesar de que muchos departamentos de policía en Europa y Estados Unidos utilizan este *software*, no existe mucha información pública de su desempeño. Solamente un estudio (Valentine *et al.*, 2010) reportó un 20% de identificación correcta en una tarea que implicaba proporcionar los nombres de los actores de una serie

conocidos por sus participantes, y cuyos retratos habían sido elaborados con este *software*. Sin embargo, cuando a los participantes no se les proporcionaba ninguna pista de a quiénes pertenecían los retratos, el porcentaje de identificación disminuyó hasta 0.8 %, a pesar de que los participantes conocían a los actores de la serie. Por el contrario, evaluaciones del desempeño de este *software* en contextos de testigos reales, llevados a cabo por la policía del oeste de Yorkshire en Reino Unido, entre 2010 y 2011, mostraron un 55% de identificación correcta de sospechosos (Solomon *et al.*, 2013).

Finalmente, está el sistema llamado Evo-Fit (Frowd *et al.*, 2015), el cual fue desarrollado en la Universidad de Stirling y la Universidad Central de Lancashire. Este *software* requiere que se establezca el rango de edad aproximado de un rostro objetivo. Posteriormente, el sistema le mostrará al usuario una serie de rostros al azar dentro del rango de edad seleccionado; enseguida, el usuario debe seleccionar el tipo de cabello del rostro objetivo entre un catálogo con diferentes estilos. La forma en que este sistema construye los retratos es a través de la selección de la forma del rostro y, luego, la textura de la piel, para finalmente combinar ambas. Este sistema utiliza un algoritmo llamado *Interactive Evolutionary Algorithm*, el cual opera de la siguiente manera:

1. Al inicio, se le muestra al usuario una serie de 10 a 32 rostros generados al azar. Luego, el usuario debe seleccionar aquellos rostros que por su forma se asemejen lo más posible al rostro objetivo, y luego por la textura. Finalmente, el usuario debe seleccionar aquellos rostros que mejor combinen forma y textura para obtener un solo individuo.
2. Para evitar que el mejor individuo o rostro se pierdan durante el proceso de evolución, este se copia y se transmite a la siguiente generación.
3. Los rostros de las diferentes generaciones son creados como resultado de los procesos evolutivos de cruce y mutación. Entonces, los nuevos rostros se obtienen de la combinación de los genes (coeficientes) de los rostros seleccionados por el usuario.

La investigación llevada a cabo por Frowd *et al.* (2001) indica que se requieren 10 generaciones para obtener un retrato aproximado de un rostro objetivo, si la población se fija en 160 rostros a evaluar por el usuario. Este *software* se emplea en diferentes departamentos de policía de Reino Unido y, a diferencia de los otros sistemas, existe una gran cantidad de investigación

sobre su desempeño (Frowd *et al.*, 2015). Para evaluar este sistema y otros similares, Frowd *et al.* (2005b) propusieron un protocolo de referencia (*gold standard*), el cual implica dos grupos de participantes en dos tareas distintas. El primer grupo participa en la construcción de los retratos a través de observar un rostro desconocido durante 1 minuto. Luego de un intervalo de 48 horas, los participantes son entrevistados sobre los rostros que observaron con una *entrevista cognitiva* (Fisher y Geiselman, 1992; Frowd *et al.*, 2008), la cual ha mostrado ser efectiva para ayudar a víctimas y testigos a recordar más detalles de un evento delictivo que hayan presenciado. Posteriormente, cada participante construirá de memoria el rostro que observó con Evo-Fit (o con el sistema disponible). Los rostros construidos con este sistema serán luego evaluados para su reconocimiento por el segundo grupo de participantes.

Este segundo grupo debe conocer los rostros a partir de los cuales se elaboraron los retratos, ya sea porque son profesores, celebridades, compañeros de escuela o trabajo, que solo este grupo conoce. Para determinar que un retrato ha sido correctamente identificado, los participantes de este segundo grupo deberán proporcionar el nombre de la persona que está representada en esos retratos. En las primeras versiones de Evo-Fit (Frowd *et al.*, 2004), este sistema alcanzó una tasa de identificación correcta de apenas 9.46%. En una versión posterior (Frowd *et al.*, 2007a), y siguiendo su protocolo de referencia propuesto, se incrementó la identificación a 11%. Este incremento se debió, de acuerdo con los investigadores, a una serie de agregados o manipulaciones de las “dimensiones holísticas” de los retratos (Frowd *et al.*, 2006). Estas dimensiones holísticas implicaban la posibilidad de modificar el retrato final, de tal manera que se pudiera aumentar o disminuir la edad aparente del rostro, el peso, y aclarar u oscurecer la piel, además de otras dimensiones psicológicas, como atractivo, introversión, amenaza o masculinidad del rostro. Esto último se conseguía agregando ciertas sobras o cambios sutiles en la expresión facial, como el ceño fruncido o una ligera sonrisa. En las primeras versiones de Evo-Fit, se podía comenzar la construcción de un retrato eligiendo una serie de opciones, como el tipo de cabello, cuello u orejas; es decir, comenzando por lo que se consideran los rasgos externos del rostro. Sin embargo, estos rasgos externos son, en general, bien recordados por las personas que han observado un rostro desconocido brevemente. En cambio, los rasgos internos (cejas, ojos, nariz y boca) presentan mayor dificultad para ser recordados (Ellis, Shepherd, Davies, 1979). Para fomentar la atención en los rasgos internos

durante los procesos evolutivos de construcción de un retrato, Frowd *et al.* (2008) modificaron este procedimiento, de tal manera que, en su nueva versión, la construcción del retrato comenzaba eligiendo un tipo de cabello particular, para luego mostrarse difuminado durante todo el proceso de evolución del rostro, destacando solo los rasgos internos. Hasta el final de la construcción del retrato se mostraba el rostro de forma completa. Con estas modificaciones al inicio, además de los agregados de las dimensiones holísticas al final de la construcción del retrato, se incrementó la identificación correcta hasta un 24.5% (Frowd *et al.*, 2011b).

Investigaciones posteriores mostraron que, si se omitían completamente los rasgos externos hasta el final de la construcción del retrato, se podía incrementar la identificación hasta un 42% (Frowd *et al.*, 2012d). Un porcentaje similar de incremento se obtuvo cuando se introdujeron cambios en la forma de presentar al público el retrato, y en la manera de entrevistar a los participantes. Por ejemplo, para la presentación pública del retrato se utilizaba una versión en la que se exageraban ciertos rasgos (caricaturización), o se presentaba con movimiento. Además, se introdujo una modificación a la entrevista cognitiva, llamada entrevista cognitiva holística, que enfatizaba los aspectos globales para recordar de un rostro (Frowd *et al.*, 2007b; Frowd *et al.*, 2012a). En los últimos años, este sistema se ha ido modificando y mejorando hasta alcanzar tasas de identificación correcta del 74% (Fondarella *et al.*, 2015; Frowd *et al.*, 2013; Frowd *et al.*, 2014a).

En una revisión de su efectividad con testigos reales, el departamento de policía de Humberside, en el norte de Inglaterra, reportó un 60% de arrestos de sospechosos utilizando este sistema, en un periodo de 12 meses en 2010 (Frowd *et al.*, 2012b).

---

## VII. CARAMEX

En México se ha desarrollado un acervo de imágenes de rasgos faciales para la generación de retratos hablados de sospechosos basado en las características antropométricas del rostro de esta población (Serrano, Villanueva, Luy y Link, 1997; Serrano, 2013; Villanueva, 2010). Este proyecto se denominó “La cara del mexicano” (Villanueva, 2010) y es, hasta el momento, una aproximación única a la construcción de retratos hablados en el mundo. Dada la gran variabilidad genética que representa el mestizaje

de la población mexicana, era importante, de acuerdo con los investigadores, contar con un acervo de imágenes que tomara en cuenta esta situación.

Este proyecto involucró la colaboración del Instituto de Investigaciones Antropológicas (IIA) de la Universidad Nacional Autónoma de México (UNAM) y la entonces Procuraduría General de Justicia del Distrito Federal (PGJDF). La investigación fue llevada a cabo entre 1993 y 1997, e implicó la obtención de un registro fotográfico del rostro de 2,890 individuos de diferentes regiones de México. Cada una de estas fotos fue tomada de frente y de perfil izquierdo con iguales condiciones de iluminación y distancia lente-rostro. En total, se recopilaron 5,780 imágenes. A partir de estas, se obtuvieron 21 mediciones antropométricas del rostro, de las que derivaron diferentes índices morfológicos y valoraciones morfoscópicas, como la forma de la inserción del cabello, el tipo de calvicie, la distribución del vello facial, la forma del dorso de la nariz, entre otras. Posteriormente, todos estos índices fueron sometidos a un análisis estadístico multivariado con el objetivo de establecer los factores tipológicos más representativos del rostro, y obtener imágenes prototípicas de cada uno de los rasgos faciales para hombres y mujeres. El resultado fue un acervo fotográfico, denominado Caramex, el cual contiene 586 archivos, de los cuales 405 están distribuidos en 26 directorios con rasgos faciales, que abarcan desde la forma general de la cara hasta arrugas o lunares (Villanueva, 2010).

La manera de construir los retratos hablados con este acervo implica la utilización del programa PhotoShop para la manipulación y edición de cada imagen. El proceso requiere que los participantes-testigos observen, primero, las diferentes formas del rostro y seleccionen la que consideran más parecida a la de la persona que observaron, para, posteriormente, ir seleccionando el resto de los rasgos faciales y colocarlos sobre esta forma facial. Este acervo, listo para usarse, fue entregado a la PGJ en 1996, y en los años subsecuentes a otros departamentos de policía de México. Según Santiago (2004, en Serrano, 2013: 9), en 2004, solo en la Ciudad de México, “cada mes, medio centenar de peritos recurrían, en casi mil casos, a Caramex para tratar de ubicar a una persona”.

---

## VIII. CARAMEX-II

Actualmente, en el Instituto de Investigación en Matemáticas Aplicadas y en Sistemas de la Universidad Nacional Autónoma de México se desarrolla

un *software* para la construcción de retratos. Al igual que otros *softwares* de cuarta generación, este se apoya en algoritmos de reconocimiento facial y creación de redes neuronales, a través de bibliotecas como Dlib, OpenCV y TensorFlow. Este procedimiento sigue los siguientes pasos:

1. Gracias al conjunto de imágenes de Caramex, es posible contar con una colección de rasgos “típicos” de la cara del mexicano y sus *landmarks* o marcadores de referencia asociados y, a su vez, por medio del *toolkit* de detección de Dlib, se reconocen los rasgos faciales principales (ojos, nariz, boca, contorno facial) dentro del catálogo de imágenes de rostros de Caramex. Los rasgos detectados se recortan y ajustan para formar combinaciones en diferentes formas de cara, para obtener una vasta colección de rostros posibles.
2. Una vez obtenida esta galería de imágenes, se introduce a una red neuronal convolucional generada en la biblioteca de código abierto para aprendizaje automático TensorFlow, donde las imágenes son transformadas en matrices de píxeles con valores que van del 0 al 1, tras una normalización de los valores originales de 0 a 255. Por medio de varias convoluciones, se toman conjuntos de píxeles relativamente cercanos de la imagen, y se van multiplicando escalarmente por una pequeña matriz llamada *kernel*, la cual, aunada a otros *kernels*, funciona como filtro. En las primeras convoluciones es posible detectar características primitivas, como líneas o curvas. Al paso de más capas en las convoluciones, será capaz de reconocer formas más complejas. Al final, este conjunto de imágenes sirve para entrenar y probar el funcionamiento de la red neuronal artificial.
3. Un grupo de estos retratos es presentado para que el usuario seleccione los más parecidos al rostro que se busca identificar. Esta selección es analizada por la red neuronal.
4. Cuando la red convolucional ya está entrenada, es capaz de distinguir y clasificar rostros generados automáticamente a través de una técnica de “morphing” implementada con algoritmos en la biblioteca OpenCV, convirtiendo las imágenes en arreglos de números flotantes y tomando *landmarks* de referencia para encontrar puntos dentro de triángulos de Delaunay (Borut 2005; Mohammadzade *et al.*, 2018), para encontrar transformaciones afines que puedan generar y ajustar a una nueva imagen del promedio de las 2 o más imágenes seleccionadas con anterioridad, logrando así mezclar los rostros.

5. Ya que el sistema clasifica los rostros escogidos, es posible saber con qué clase de rasgos fue formado dicho rostro, y sugerir una nueva combinación de rasgos para obtener rostros semejantes a los escogidos previamente. Este paso se repite hasta que el *morphing* de imágenes sea lo más parecido al retrato buscado.

---

## IX. CONCLUSIONES

El retrato hablado es una técnica ampliamente utilizada en la investigación policial. Sin embargo, se ha evaluado muy poco su eficacia en escenarios reales de la comisión de un delito. En México, se siguen utilizando el arte forense y el catálogo de imágenes Caramex, que desde una perspectiva cognitiva siguen siendo métodos tradicionales que muy poco tienen que ver con los procesos de percepción facial humana. Es en este sentido que las nuevas metodologías de los sistemas de cuarta generación ofrecen una alternativa a los métodos tradicionales. Su principal ventaja radica en la utilización de rostros completos que permiten que la carga de trabajo recaiga sobre la memoria visual de una víctima o testigo, y no sobre su capacidad de descripción verbal susceptible de error, o de contaminar su memoria visual del rostro de un sospechoso.

La investigación que lleva a cabo el grupo del IIMAS presenta, en este mismo sentido, una ventaja extra. Para construir sus retratos hablados, la mayoría de los sistemas de cuarta generación utilizan como base de datos una serie de rostros relativamente pequeños, sin un criterio claro de inclusión. Por su parte, Caramex II cuenta con una sólida investigación estadística de los rasgos faciales más representativos de la población mexicana. La utilización de técnicas de inteligencia artificial, programación evolutiva, algoritmos genéticos, procesamiento de imágenes y el conocimiento sobre percepción facial humana, puede proporcionar un salto enorme en el desarrollo de nuevas estrategias de construcción de retratos hablados. Pero aún se requiere mayor investigación.

---

## X. FUENTES DE CONSULTA

Behrmann, M., Richler, J.J., Avidan, G., y Kimchi, R. (2015). "Holistic face perception". *Oxford handbook of perceptual organization*, 758-774.

- Boutsen, L., y Humphreys, G.W. (2003). "The effect of inversion on the encoding of normal and 'Thatcherized' faces". *The Quarterly Journal of Experimental Psychology*, 56(6), 955-975.
- Brace, N.A., Pike, G.E., Allen, P., y Kemp, R.I. (2006). "Identifying composites of famous faces: Investigating memory, language, and system issues". *Psychology, Crime & Law*, 12(4), 351-366.
- Brace, N., Pike, G., y Kemp, R. (2000). "Investigating E-FIT using famous faces". *Forensic psychology and law*, 272-276.
- Carbon, C.C., y Leder, H. (2005). "When feature information comes first! Early processing of inverted faces". *Perception*, 34(9), 1117-1134.
- Cootes, T.F., y Taylor, C.J. (2001, July). "Statistical models of appearance for medical image analysis and computer vision". *Medical Imaging 2001: Image Processing* (Vol. 4322, pp. 236-248). International Society for Optics and Photonics.
- Cootes, T.F., Edwards, G.J., y Taylor, C.J. (1998, June). "Active appearance models". *European conference on computer vision* (pp. 484-498). Springer, Berlin, Heidelberg.
- Craw, I., y Cameron, P. (1991). "Parameterising images for recognition and reconstruction". In *BMVC91* (pp. 367-370). Springer, London.
- Craw, I., y Cameron, P. (1992). "Face recognition by computer". In *BMVC92* (pp. 498-507). Springer, London.
- Davies, G. (1986). "Capturing likeness in eyewitness composites: The police artist and his rivals". *Medicine, Science, and the Law*, 26(4), 283-290.
- Davies, G., y Oldman, H. (1999). "The impact of character attribution on composite production: A real world effect?" *Current Psychology*, 18(1), 128-139.
- Davies, G., y Valentine, T. (2007). "Facial composites: Forensic utility and psychological research". *The Handbook of Eyewitness Psychology: Volume II* (pp. 73-98). Psychology Press.
- Davies, G., Van der Willik, P., y Morrison, L.J. (2000). "Facial composite production: A comparison of mechanical and computer-driven systems". *Journal of Applied Psychology*, 85(1), 119.
- Dodson, C.S., Johnson, M.K., y Schooler, J.W. (1997). "The verbal overshadowing effect: Why descriptions impair face recognition". *Memory & Cognition*, 25(2), 129-139.
- Ellis, H.D. (1986). "Introduction to aspects of face processing: Ten questions in need of answers". *Aspects of face processing* (pp. 3-13). Springer, Dordrecht.

- Ellis, H.D., Shepherd, J.W., y Davies, G.M. (1979). "Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition". *Perception*, 8(4), 431-439.
- Fallshore, M. y Schooler, J.W. (1995). "Verbal vulnerability of perceptual expertise". *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(6), 1608.
- Fisher, R.P., y Geiselman, R.E. (1992). *Memory enhancing techniques for investigative interviewing: The cognitive interview*. Charles C Thomas Publisher.
- Fodarella, C., Kuivaniemi-Smith, H., Gawrylowicz, J., y Frowd, C. (2015). "Forensic procedures for facial-composite construction". *Journal of Forensic Practice*, 17(4), 259-270.
- Frowd C.D. (2001). "EvoFIT: a holistic, evolutionary facial imaging system". PhD thesis, University of Stirling.
- Frowd, C.D., Valentine, T. y Davis, J. (2015a). "Facial composites and techniques to improve image recognizability". *Forensic facial identification: Theory and practice of identification from eyewitnesses, composites, and CCTV*, 43-70.
- Frowd, C.D., Erickson, W.B., Lampinen, J.M., Skelton, F.C., McIntyre, A.H. y Hancock, P.J. (2015b). "A decade of evolving composites: regression and meta-analysis". *Journal of Forensic Practice*.
- Frowd, C.D., Skelton, F., Hepton, G., Holden, L., Minahil, S., Pitchford, M. y Hancock, P.J. (2013). "Whole-face procedures for recovering facial images from memory". *Science & Justice*, 53(2), 89-97.
- Frowd, C.D., Carson, D., Ness, H., Richardson, J., Morrison, L., McLanaghan, S. y Hancock, P. (2005b). "A forensically valid comparison of facial composite systems". *Psychology, Crime & Law*, 11(1), 33-52
- Frowd, C.D., Hancock, P.J. y Carson, D. (2004). "EvoFIT: A holistic, evolutionary facial imaging technique for creating composites". *ACM Transactions on applied perception (TAP)*, 1(1), 19-39.
- Frowd, C.D., Bruce, V., Ness, H., Bowie, L., Paterson, J., Thomson-Bogner, C. & Hancock, P.J. (2007a). "Parallel approaches to composite production: interfaces that behave contrary to expectation". *Ergonomics*, 50(4), 562-585.
- Frowd, C., Bruce, V., Ross, D., McIntyre, A., y Hancock, P.J. (2007b). "An application of caricature: how to improve the recognition of facial composites". *Visual Cognition*, 15(8), 954-984.
- Frowd, C., Park, J., McIntyre, A., Bruce, V., Pitchford, M., Fields, S. y Hancock, P.J. (2008, August). "Effecting an improvement to the fitness

- function. How to evolve a more identifiable face”. *2008 Bio-inspired, Learning and Intelligent Systems for Security* (pp. 3-10). IEEE.
- Frowd, C.D., Pitchford, M., Bruce, V., Jackson, S., Hepton, G., Greenall, M. y Hancock, P.J. (2011b). “The psychology of face construction: giving evolution a helping hand”. *Applied Cognitive Psychology*, 25(2), 195-203.
- Frowd, C., Nelson, L., Skelton, F., Noyce, R., Atkins, R., Heard, P. y Hancock, P.J. (2012a). “Interviewing techniques for Darwinian facial-composite systems”. *Applied Cognitive Psychology*, 26(4), 576-584.
- Frowd, C.D., Pitchford, M., Skelton, F., Petkovic, A., Prosser, C. y Coates, B. (2012b, September). “Catching even more offenders with EvoFIT facial composites”. *2012 Third International Conference on Emerging Security Technologies* (pp. 20-26). IEEE.
- Frowd, C.D., Skelton, F., Atherton, C., Pitchford, M., Hepton, G., Holden, L. y Hancock, P.J. (2012d). “Recovering faces from memory: The distracting influence of external facial features”. *Journal of Experimental Psychology: Applied*, 18(2), 224.
- Frowd, C.D., Jones, S., Fodarella, C., Skelton, F., Fields, S., Williams, A. y Date, L. (2014). “Configural and featural information in facial-composite images”. *Science & Justice*, 54(3), 215-227.
- Frowd, C.D., Bruce, V., McIntyre, A.H., Ross, D., Fields, S., Plenderleith, Y. y Hancock, P.J. (2006). “Implementing holistic dimensions for a facial composite system”. *Journal of Multimedia*, 1(3), 42-51.
- Gibson, S.J., Solomon, C.J., Maylin, M.I. y Clark, C. (2009). “New methodology in facial composite construction: From theory to practice”. *International Journal of Electronic Security and Digital Forensics*, 2(2), 156-168.
- Hancock, P.J., Bruce, V. y Burton, A.M. (2000). “Recognition of unfamiliar faces”. *Trends in cognitive sciences*, 4(9), 330-337.
- Hole, G.J. (1994). “Configurational factors in the perception of unfamiliar faces”. *Perception*, 23(1), 65-74.
- Innocence Project (2016): *Reevaluating lineups: why witness make mistakes and how to reduce the chance of a misidentification*. Recuperado de: <https://www.innocenceproject.org/reevaluating-lineups-why-witnesses-make-mistakes-and-how-to-reduce-the-chance-of-a-misidentification/>
- Kurt, B., Etaner-Uyar, A.S., Akbal, T., Demir, N., Kanlikilicer, A.E., Kus, M.C. y Ulu, F.H. (2006, September). “Active appearance model-based facial composite generation with interactive nature-inspired heuristics”. *International Workshop on Multimedia Content Representation, Classification and Security* (pp. 183-190). Springer, Berlin, Heidelberg.

- Laughery, K.R. y Fowler, R.H. (1980). "Sketch artist and Identi-kit procedures for recalling faces". *Journal of Applied Psychology*, 65(3), 307.
- Mohammadzade, H., Sayyafan, A. y Ghogh, B. (2018). "Pixel-Level Alignment of Facial Images for High Accuracy Recognition Using Ensemble of Patches". *Journal of the Optical Society of America A*. 35. 10.1364/JOSAA.35.001149.
- Rossion, B. (2008). "Picture-plane inversion leads to qualitative changes of face perception". *Acta psychologica*, 128(2), 274-289.
- Ryan, R.S., y Schooler, J.W. (1998). "Whom do words hurt? Individual differences in susceptibility to verbal overshadowing". *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 12(7), S105-S125.
- Schooler, J.W., y Engstler-Schooler, T.Y. (1990). "Verbal overshadowing of visual memories: Some things are better left unsaid". *Cognitive psychology*, 22(1), 36-71.
- Serrano, S.C. (2013). "Un sistema automatizado de identificación de rasgos faciales (retrato hablado) para la población mexicana". *La bibliotecología y la documentación en el contexto de la internacionalización y el acceso abierto*. (Coord.) Jaime Ríos Ortega y César Augusto Ramírez Velázquez. México, UNAM, Instituto de Investigaciones Bibliotecológicas y de la Información.
- Serrano, C., Villanueva, M., Luy, J. y Link, K.F. (1997). "El proyecto 'La cara del mexicano'. Un sistema de retrato hablado asistido por computadora para la población mexicana." *Boletín de Enlaces*: 26-28, Coordinación de Humanidades, UNAM, México.
- Sirovich, L. y Kirby, M. (1987). "Low-dimensional procedure for the characterization of human faces". *Josa a*, 4(3), 519-524.
- Solomon, C.J., Gibson, S.J. y Pallares-Bejarano, A. (2005). "Eigenfit—the generation of photographic-quality facial composites". *The Journal of Forensic Science*.
- Solomon, C.J., Gibson, S.J. y Mist, J.J. (2013). "Interactive evolutionary generation of facial composites for locating suspects in criminal investigations". *Applied Soft Computing*, 13(7), 3298-3306.
- Sporer, S.L. (2007). "Person descriptions as retrieval cues: Do they really help?" *Psychology, Crime & Law*, 13(6), 591-609.
- Tanaka, J.W. y Farah, M.J. (1993). "Parts and wholes in face recognition". *The Quarterly journal of experimental psychology*, 46(2), 225-245.

- Tanaka, J.W. y Simonyi, D. (2016). “The ‘parts and wholes’ of face recognition: A review of the literature”. *Quarterly Journal of Experimental Psychology*, 69(10), 1876-1889.
- Taylor, K.T. (2000). *Forensic art and illustration*. CRC Press.
- Thompson, P. (1980). “Margaret Thatcher: a new illusion”. *Perception*.
- Tredoux, C. (1999). “Statistical considerations when determining measures of lineup size and lineup bias”. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 13(S1), S9-S26.
- Tredoux, C., Nunez, D., Oxtoby, O. y Prag, B. (2006). “An evaluation of ID: an eigenface based construction system: reviewed article.” *South African Computer Journal*, 2006(37), 90-97.
- Troje, N.F. y Vetter, T. (1996). *Representations of human faces*. Max-Planck-Institut Tubingen
- Turk, M. y Pentland, A. (1991). “Eigenfaces for recognition”. *Journal of cognitive neuroscience*, 3(1), 71-86.
- Valentine, T. (1988). “Upside-down faces: A review of the effect of inversion upon face recognition”. *British journal of psychology*, 79(4), 471-491.
- Valentine, T. (1991). “A unified account of the effects of distinctiveness, inversion, and race in face recognition”. *The Quarterly Journal of Experimental Psychology Section A*, 43(2), 161-204.
- Valentine, T. (Ed.). (2002). *Cognitive and computational aspects of face recognition: Explorations in face space*. Routledge.
- Valentine, T., Davis, J.P., Thorner, K., Solomon, C. y Gibson, S. (2010). “Evolving and combining facial composites: Between-witness and within-witness morphs compared”. *Journal of Experimental Psychology: Applied*, 16(1), 72.
- Villanueva, S.G. (2010). *Morfología facial. Estudios en población mexicana a través de fotografías digitales*. México, UNAM, Instituto de Investigaciones Antropológicas.
- Wells, G.L., Charman, S.D., y Olson, E.A. (2005). “Building face composites can harm lineup identification performance”. *Journal of experimental psychology: Applied*, 11(3), 147.
- Zahradnikova, B., Duchovicova, S. y Schreiber, P. (2018). “Facial composite systems”. *Artificial Intelligence Review*, 49(1), 131-152.
- Žalik, Borut (2005). “An efficient sweep-line Delaunay triangulation algorithm”. *Computer-Aided Design*, 37 (10), pp. 1027-1038.

