

Cómo estimar la edad de las personas utilizando imágenes de su rostro

Enrique Alegre, Rocío Aláiz-Rodríguez, Laura Fernández-Robles, Oscar García-Olalla, Diego García-Ordas, Maite García-Ordas y Víctor González-Castro

Universidad de León

1 Introducción

AL avanzar la edad la apariencia del rostro humano va desarrollando considerables cambios. Cambios lentos que no son reversibles por medios naturales. Sin embargo, los signos de envejecimiento del rostro son incontables y personales por lo que la estimación correcta de la edad se convierte en una tarea de alta dificultad. Sumado a ello, la genética y el estilo de vida de la persona afecta enormemente a la apariencia del rostro. Un mayor grado de estrés, el consumo de cigarrillos o los efectos de la radiación ultravioleta son factores que aceleran los signos de envejecimiento, pudiendo llevar a pensar que una persona es más de 10 años mayor de su edad real. Además, la postura, la iluminación, el maquillaje o las operaciones quirúrgicas, entre otros factores, dificultarán aún más la estimación correcta de la edad a partir de una cara.



Figura 1: Efectos de la cirugía estética (izquierda) (Troell, 2014) y del maquillaje (derecha) (Begoña, 2013) en la edad percibida.

Los principales signos de envejecimiento de un rostro son el crecimiento craneofacial, el envejecimiento de la piel debido a la aparición de arrugas, cambio de color de la piel y manchas y la reducción de la elasticidad del músculo (Dib, 2011). La distribución de la cara sufre ligeras variaciones apreciables en el contorno de la cara y en la forma de las características faciales como son los ojos, nariz o boca.

1. VISIÓN ARTIFICIAL

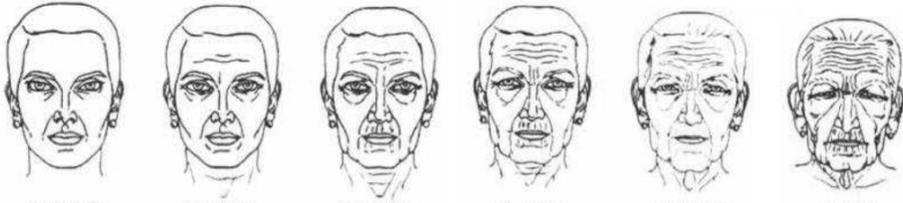


Figura 2: Esquemas de la edad de un rostro desde los 30 a los 80 años (Fu, Guo & Huang, 2010; Gonzalez-Ulloa & Flores, 1965).

La estimación de la edad tiene como fin el etiquetar una cara bien con la edad exacta (año) o bien dentro un grupo de edades (rango de años). Si bien la primera opción es más precisa también sufrirá de mayor número de errores. Para entender la dificultad a la hora de estimar la edad de un rostro hay que diferenciar entre los siguientes cuatro conceptos sobre la edad humana (Fu, Guo & Huang, 2010):

- Edad real: La edad de un individuo, años acumulados desde su nacimiento.
- Edad aparente: La edad mostrada por la apariencia visual de un rostro.
- Edad percibida: La edad de un individuo calculada por personas a partir de la apariencia visual.
- Edad estimada: La edad de un individuo reconocida por una máquina a partir de la apariencia visual.

Analizando todos estos factores, podemos afirmar que la estimación de la edad a partir de un rostro es un reto no sólo cuando se pretende realizar de manera automática mediante un ordenador y software sino incluso también cuando la estima una persona.

2 Dificultades

Una de las principales dificultades con las que nos podemos encontrar a la hora de diseñar un sistema de estimación de edad es la falta de conjuntos de datos etiquetados. Una buena colección de datos debería constar de sujetos con un amplio rango de edades y además una imagen de cada edad de cada individuo.

Otro de los problemas es la descomposición de las características faciales que contienen información relevante sobre la edad de la persona. A menudo, las personas utilizan otros indicadores además del rostro para estimar la edad de un

1. VISIÓN ARTIFICIAL

individuo, como el color del pelo, el tono de la voz. El sistema de estimación de edad mediante visión por computador podría no ser tan robusto debido a la falta de estos otros indicadores.

Las diferentes poses, gestos y condiciones de iluminación bajo las que se encuentran los rostros suponen también un problema a la hora de entrenar los clasificadores para la estimación de edad, sin embargo el principal problema es que no todo el mundo envejece de la misma manera ni a la misma velocidad, por lo tanto hay un margen de error entre la edad real y la edad aparente. En ocasiones una clasificación binaria es suficiente (Por ejemplo, para saber si un individuo es mayor de 18 años o no) sin embargo en otras ocasiones se requiere una clasificación difusa (Por ejemplo, hay un 85% de posibilidades de que el individuo tenga 18 años) (El Dib, 2011).



Figura 3: Diferentes edades de un mismo individuo.

3 Aplicaciones de la estimación automática de la edad

La detección de edad a partir de imágenes faciales ha abierto un camino a multitud de aplicaciones prácticas destinadas a mejorar la calidad de vida de las personas. A continuación se comenta alguna de ellas.

ECRM

La administración electrónica de relaciones con los clientes (ECRM) es un sistema que utiliza tecnología puntera de visión artificial para establecer relaciones con clientes y proporcionarles un servicio individualizado según sus necesidades específicas. Los clientes pueden entrar dentro de diferentes categorías según su edad: bebés, adolescentes, adultos y personas de la tercera edad. Mediante estos sistemas, los productos y servicios se personalizan de acuerdo con cada grupo de edad de forma que las empresas tras estos servicios puedan optimizar sus ingresos.

Gracias a estos sistemas, las campañas de marketing pueden orientarse hacia audiencias específicas en función de la edad de sus clientes, ya que les permiten conocer qué grupos de edad están más interesados en unos productos o en otros (El Dib, 2011).

1. VISIÓN ARTIFICIAL

Ayuda a personas mayores con necesidades especiales de asistencia



Figura 4: Robot enfermera RIBA.

Una de las principales aplicaciones de la estimación de edad es el cuidado de personas mayores con necesidades especiales, mediante robots enfermeros en los hospitales, unidades de cuidados intensivos inteligentes, o incluso para proporcionar una atención más humanizada con avatares virtuales que adoptan una u otra forma dependiendo de la edad del paciente (El Dib, 2011).

Sistemas de vigilancia y control de la seguridad

La importancia de los sistemas de vigilancia y seguridad está cobrando mucha importancia últimamente, debido al incremento de actividades delictivas y amenazas



Figura 5: Diferentes tipos de cámaras de vigilancia.

que podrán poner en riesgo la vida de las personas. Estos sistemas se podrían utilizar también para vigilar otros aspectos más allá de la seguridad, como por ejemplo la venta automatizada de alcohol y tabaco a menores en las máquinas mediante la estimación de edad de sus rostros. Otra de las aplicaciones de la estimación de edad en vigilancia podría ser el control de acceso de menores a contenidos de adultos en internet. Como se puede observar, la inclusión de estimación de edad en un sistema de vigilancia puede tener multitud de

aplicaciones (Geng, Fu & Smith-Miles, 2010; El Dib, 2011).

Recuperación de información

Internet es la mayor base de datos de rostros existentes, con billones de rostros y creciendo día a día. Estas imágenes de rostros suelen contener metadatos y etiquetas que pueden contener información útil. Haciendo uso de la estimación

1. VISIÓN ARTIFICIAL

de edad a la hora de recuperar datos se pueden realizar búsquedas de rostros por rangos de edades (El Dib, 2011).

Detección de pornografía infantil

Otra de las aplicaciones más importantes de la estimación de edad es la detección automática entre niños y adultos en contenidos sexuales explícitos. Mediante estos sistemas se ponen a disposición de las autoridades nuevas herramientas que les permiten realizar su trabajo de una manera más cómoda y rápida que la categorización manual de cientos de miles de imágenes (Tanner, 2011).



Figura 6: ASASEC European Project: Advisory System Against Sexual Exploitation of Children

4 Representación de la edad a partir de imágenes

Modelos antropométricos

Los primeros trabajos conocidos en el área de clasificación de la edad se basan en la teoría de desarrollo craneo-facial (Alley, 1988), que utiliza un modelo matemático para describir el crecimiento de la cabeza de una persona desde su infancia hasta que es adulto. Según dicho modelo el perfil de una cabeza está determinado por un par (r, θ) , donde r la distancia del origen a un punto P del perfil, y θ el ángulo que forma el segmento OP con el eje vertical. Si llamamos k a un parámetro que se incrementa con el tiempo, los crecimientos sucesivos del perfil, (r', θ') , se modelan como:

$$\begin{cases} \theta' &= \theta \\ r' &= r (1 + k (1 - \cos \theta)) \end{cases}$$

Dicho modelo ha dejado de utilizarse para estimar edades por dos motivos. En primer lugar, el modelo no es capaz de describir el perfil de la cabeza de manera natural, en especial en edades adultas. Además, en ocasiones el perfil de la cabeza es difícil de extraer de imágenes de caras en 2D.

1. VISIÓN ARTIFICIAL

Sucesivos trabajos de estimación de la edad han utilizado la llamada “antropometría de la cara”, que se basa en la extracción y estudio de tamaños y proporciones en la misma. Farkas definió la antropometría en función de medidas de 57 puntos de interés en caras de personas de diferentes edades (Farkas, 1994). De este modo, para caracterizar la edad de una persona se miden distancias entre algunos de estos puntos y se consideran las proporciones entre ellas.

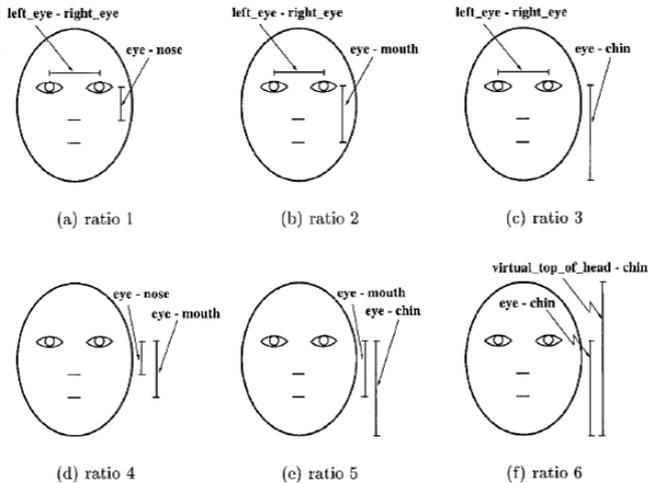


Figura 7: Caracterización de la antropometría mediante relaciones de medidas.

Esta idea es utilizada por Kwon y Lobo para distinguir bebés de adultos automáticamente. En concreto, en sus trabajos (Kwon & Lobo, 1994, 1999) extraen 6 proporciones diferentes de imágenes de caras (ver figura 7). Ramanathan y Chellappa (Ramanathan & Chellappa, 2006) usaron la misma idea (en este caso, utilizando 8 proporciones (ver figura 8) para modelar la progresión de edades comprendidas entre 0 y 18 años.

Sin embargo, los modelos antropométricos presentan varios inconvenientes:

1. Son solamente útiles en la detección de la edad de personas jóvenes ya que en la edad adulta la forma de la cabeza no cambia demasiado. De hecho, Kwon y Lobo tuvieron que desarrollar un sistema para detectar arrugas en las caras y así distinguir los adultos jóvenes de personas mayores en los casos que las medidas antropométricas no se correspondían a las de un menor.
2. La detección de los puntos de interés en las imágenes (previo al cálculo de las medidas y sus proporciones) es muy sensible a la pose de la cabeza, lo

1. VISIÓN ARTIFICIAL

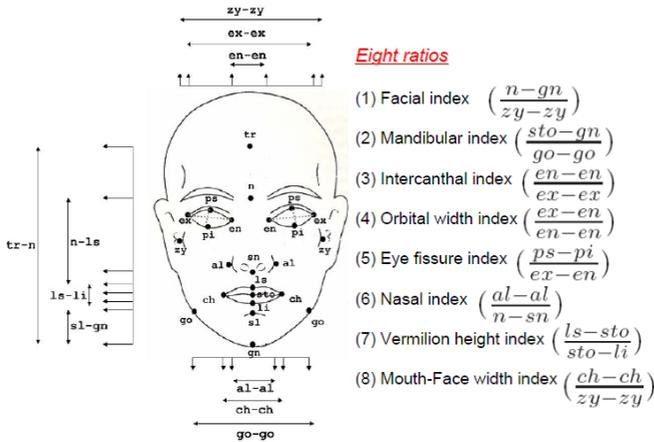


Figura 8: Proporciones usadas en el trabajo de Ramanathan y Chellappa.

que dificulta su aplicación en situaciones “reales”.

3. Este método no es universal, pues las medidas dadas por Farkas son válidas para personas de origen caucásico (europeos o descendientes). Sería precisa una adaptación a personas de otros orígenes.

Modelos de apariencia activos



Figura 9: Imagen etiquetada con 122 puntos de referencia.

Los modelos de apariencia activa están basados en modelos estadísticos del rostro y fueron propuestos inicialmente para codificar imágenes de rostros (Cootes, Edwards & Taylor, 2001). Se basa en puntos de referencia en la imagen que son marcados de manera manual en la etapa de entrenamiento. Teniendo en cuenta el análisis de componentes estadísticos aplicado sobre un conjunto de imágenes de entrenamiento formado por imágenes de rostros, se extrae un modelo estadístico de la forma y un modelo de intensidad, independientes el uno del otro.

1. VISIÓN ARTIFICIAL

Los MAAs (Modelos de Apariencia Activos) han sido utilizados en múltiples ocasiones para codificar rostros de manera satisfactoria. Entrando en temas relacionados con la edad, Lanitis et al (Lanitis, Taylor & Cootes, 2002) extendieron los modelos de apariencia activos añadiendo una función de edad para explicar la variación de los rostros respecto a la edad. Esta función define la relación entre la edad de los individuos y la descripción de las imágenes. En comparación con los modelos antropométricos, los MAAs pueden tratar con cualquier edad, en vez de tan solo con los jóvenes. Además, estos métodos consideran tanto la forma como la textura en vez de las características geométricas de la cara y son robustos frente a posturas cosa que los métodos antropométricos no. Estas técnicas son aplicables a la estimación precisa de la edad, en cuanto que cada imagen de test es etiquetada con un específico valor escogido de un rango continuo. Las mejoras más importantes de estos métodos de representación de la edad aportan son las siguientes:

- Prueba que la relación entre la edad y el rostro se puede representar por una función cuadrática.
- Hace frente a los valores atípicos en el rango de edades.
- Trata con parámetros de alta dimensionalidad.

Subespacio del patrón de envejecimiento

Geng y su grupo de investigación (Geng, Fu & Smith-Miles, 2010; Geng, Zhou & Smith-Miles, 2007), introdujeron el término AGES (AGing pattErn Subspace) que trata con una secuencia de imágenes de individuos con diferente edad usadas de manera conjunta para modelar el proceso del envejecimiento en vez de utilizar cada imagen por separado.

Un patrón de envejecimiento se define como una secuencia de imágenes de la misma persona ordenadas temporalmente. Se llama patrón de envejecimiento completo cuando se dispone de una imagen por año del individuo y patrón de envejecimiento incompleto en el caso contrario. El método AGES es capaz de simular las imágenes de los rostros no almacenados usando algoritmos de aprendizaje iterativos basados en EM (Esperanza-Maximización).

Este método consta de dos etapas: La etapa de aprendizaje y la etapa de estimación de la edad.

En la etapa de aprendizaje, la técnica de PCA (Análisis de Componentes Principales) se utiliza para obtener una representación del subespacio. Lo primero

1. VISIÓN ARTIFICIAL

que se hace es utilizar la técnica basada en EM para reconstruir las caras que faltan en el patrón de envejecimiento. Una vez que se ha obtenido una reconstrucción cuyo error es aceptable se pasa a la etapa siguiente.



Figura 10: Ejemplo de patrones de envejecimiento.

Para realizar la estimación de la edad, una imagen de test es evaluada y asignada a un determinado patrón de envejecimiento en una posición concreta. Para eso se compara con cada una de las imágenes que forman el patrón y se selecciona la que supone un error menor en la reconstrucción. El error medio que se obtiene con este método es de 6.77 años. El principal problema que presenta es que es necesario disponer de un conjunto de imágenes de diferentes edades para cada persona. Además, otra desventaja que tiene es que no utiliza información de textura por lo que características como las arrugas no se tienen en cuenta.

Manifold de edad

Este método se basa en la idea de que en vez de “aprender” un patrón de envejecimiento para cada individuo, se podría modelar un único patrón para varios individuos a lo largo de diferentes edades. Así, cada edad se representaría utilizando “diferentes caras”. De esta manera el reconocimiento de la edad es muy flexible en comparación con el método del subespacio del patrón de envejecimiento (Geng, Zhou & Smith-Miles, 2007). El manifold de edad (Fu, Xu & Huang, 2007) es una forma de modelar dicho patrón de edad común. Dicho método no necesita un número mínimo de imágenes de un individuo determinado. De hecho, permitiría

1. VISIÓN ARTIFICIAL

que, en un caso extremo, una persona solo “aportase” una imagen de una edad, lo que es muy práctico a la hora de crear bases de datos de imágenes.

Si visualizamos cada imagen (cuya edad real es conocida) como un punto en un espacio de D dimensiones (esto es, cada imagen se representa mediante un vector D -dimensional), el objetivo del manifold de edad es encontrar una representación de los datos en una dimensión d más pequeña ($d \ll D$) capturando a la vez la distribución intrínseca de los datos y su estructura geométrica subyacente. En la literatura se han propuesto diversas técnicas, como OLPP (Orthogonal Locality Preserving Projections) (Cai, He, Han & Zhang, 2006), CEA (Conformal Embedding Analysis) (Fu & Huang, 2008), LARR (Locally Adjusted Robust Regression) (Guo, Fu, Dyer & Huang, 2008) o SSE (Synchronized Submanifold Embedding) (Yan et al., 2009), que han estimado edades en diversas bases de datos de imágenes de personas con errores medios cercanos a 5 años.

La visualización de los datos obtenida con OLPP y CEA muestra que representan el conjunto de datos “mejor” que técnicas clásicas de reducción de dimensionalidad como Principal Component Analysis (PCA) (figura 11).

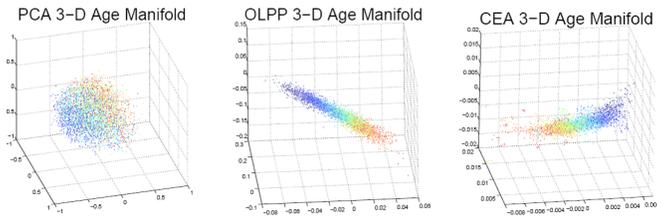


Figura 11: Manifold de edad obtenido mediante PCA (izquierda), OLPP (centro) y CEA (derecha).

La principal ventaja del manifold de edad es que permite reducir la dimensionalidad de los datos de imágenes para estimar la edad preservando la estructura geométrica de dichos datos. Además, no necesitan modelar patrones de edad específicos de cada persona ni incluir imágenes de la misma persona en diferentes edades. El único requisito para asegurar la suficiente robustez estadística del modelo es que el conjunto de imágenes utilizado para ello sea grande.

Modelos de apariencia

Los modelos de apariencia combinan descripción de textura con descripción de forma. Aunque existen muchos métodos, como se ha visto en secciones anteriores, la extracción de características faciales relacionadas con el envejecimiento, está

1. VISIÓN ARTIFICIAL

más centrada en el modelo de apariencia.

En sistemas existentes de estimación de la edad ya se han usado tanto descriptores globales como locales. El LBP (Local Binary Pattern) (Ahonen, Hadid & Pietikainen, 2006) que es un descriptor de textura muy robusto, también ha sido usado para la extracción de características de apariencia en sistemas de estimación de la edad y ha alcanzado una tasa de acierto del 80% usando las imágenes de la base de datos FERET y clasificando con la técnica del vecino más cercano y un 80%-90% usando FERET y PIE y clasificando con AdaBoost (Yang & Ai, 2007).

En las siguientes imágenes se puede ver un ejemplo de los resultados obtenidos usando la base de datos FERET (figura 12) y la base de datos PIE (figura 13).

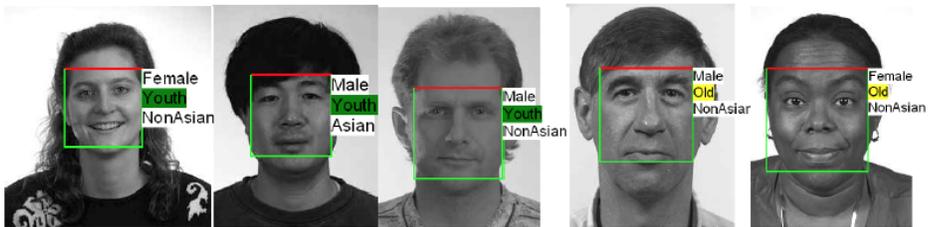


Figura 12: Resultados para las imágenes de FERET usando LBP y AdaBoost.

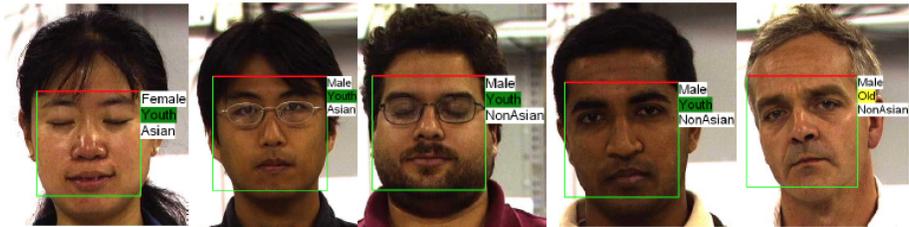


Figura 13: Resultados para las imágenes de PIE usando LBP y AdaBoost.

Con el paso del tiempo, se hacen evidentes variaciones en la apariencia y por esta razón, Yan y su grupo de investigación (Yan, Zhou, Liu, Hasegawa-Johnson & Huang, 2008; Yan, Liu & Huang, 2008b), propusieron usar Spatially Flexible Path (SFP) como descriptor de características. Este método usa información de regiones locales de la imagen y su posición. Esto es útil en caso de imágenes con oclusiones o con diferentes posiciones de la cara, ya que al ser un método robusto por trabajar a nivel local, se pueden describir este tipo de imágenes de manera eficiente. Además, también es capaz de dar más importancia a la característica discriminante del conjunto de características cuando se tiene un

1. VISIÓN ARTIFICIAL

conjunto de muestras insuficiente. Evaluado con un modelo Gaussiano mixto (GMM), este sistema obtiene un error absoluto medio de 4.95 años en la base de datos FG-NET. Usando la base de datos YGA, se obtiene un error absoluto medio de 4.94 años en el caso de rostros de hombres y 4.38 en el caso de las mujeres.

Por otra parte, Suo y su grupo (Suo et al., 2008), proponen diseñar características faciales gráficas (topología, geometría, fotometría, y configuración) basadas en el modelo de rostros jerárquico multiresolución que desarrollaron (Suo, Min, Zhu, Shan & Chen, 2007). Básicamente lo que hace este modelo, es aplicar diferentes filtros a diferentes parámetros y en diferentes niveles de extracción de características. Este sistema obtiene un error absoluto medio de 5.974 años usando la base de datos FG-NET y clasificando con redes neuronales (un perceptrón multicapa en este caso). Usando su propia base de datos obtuvieron un error de absoluto medio de 4.68 años.

Guo y su grupo (Guo, Mu, Fu & Huang, 2009), por otra parte, desarrollaron un método para estimar la edad a partir del rostro llamado BIF (Biologically Inspired Features) (Riesenhuber & Poggio, 1999; Serre, Wolf, Bileschi, Riesenhuber & Poggio, 2007). La idea básica es que el modelo contiene capas alternadas, unas simples (S) y otras complejas (C). La complejidad de las capas va aumentando según nos movemos desde el córtex visual primario al córtex inferior temporal. La primera capa, S1, se crea con filtros Gabor sobre la imagen de entrada. La segunda capa, C1, se obtiene aplicando una operación MAX sobre la capa S1. Este tipo de características tienen la ventaja de que son robustas a pequeñas traslaciones, rotaciones, escalado, capturando con eficacia patrones de edad. Si combinamos este modelo BIF con un clasificador SVM (Support Vector Machines), se obtiene un error absoluto medio de 4.77 años para la base de datos FG-NET. Si se prueba sobre las imágenes de la base de datos YGA, se obtiene un error absoluto medio de 3.91 en el caso de rostros femeninos y 3.47 en el caso de rostros masculinos.

5 Algoritmos para estimar la edad

Desde un punto de vista formal, la solución de este problema implica la obtención de una función matemática que establece la relación entre el espacio de observaciones de entrada y la salida correspondiente a la edad (o de grupo de edad).

Sin embargo, en una aplicación de estas características no es posible disponer de modelos matemáticos completos que caractericen con precisión la distribución estadística de los datos. Por esta razón, se puede recurrir a procedimientos basados en el aprendizaje a partir de un conjunto de ejemplos y así, plantear la estimación de la edad a partir del rostro como un problema particular de reconocimiento de patrones.

De esta forma, si el objetivo es la asignación de un caso a una de entre un conjunto de posibles categorías de edad predefinidas se puede abordar con técnicas de clasificación o bien, optar por modelos de regresión si lo que se pretende es la asignación de un valor numérico correspondiente a la edad. Igualmente, la combinación de ambas técnicas parece ser una aproximación prometedora (Tanner, 2011).



Figura 14: Categorización de los algoritmos de estimación de edad.

Cualquiera de las técnicas empleadas requiere disponer de un conjunto de imágenes etiquetado representativo pues esto es fundamental para poder ajustar los algoritmos de estimación de edad (Fu, Guo & Huang, 2010; Geng, Fu & Smith-Miles, 2010; Tanner, 2011) y obtener modelos con un rendimiento aceptable. Una base de imágenes ideal debería cubrir un amplio rango de edad, contener imágenes de diferentes personas y disponer al menos de una imagen de cada individuo para todas las edades consideradas. En la práctica, resulta extremadamente costoso y laborioso conseguir un conjunto de imágenes de entrenamiento suficientemente grande, especialmente cuando se precisan series cronológicas de un mismo individuo. Por ello, es frecuente disponer de un número limitado de imágenes para cada persona y que las imágenes correspondientes a edad avanzada sean escasas.

Dadas las peculiaridades de este contexto, la aplicación directa de las técnicas clásicas de clasificación o regresión resulta en la mayoría de las ocasiones difícil o ineficiente. Es necesario, pues, adaptarlas a las características de este escenario.

Clasificación

Lanitis et al. (Lanitis, Taylor & Cootes, 2002) evalúan diferentes clasificadores para la estimación automática de la edad a partir de imágenes del rostro. Se explora tanto la clasificación en una sola etapa como el empleo de métodos de categorización jerárquicos donde un clasificador global proporciona una estimación aproximada de un rango de edad que posteriormente se refina con un clasificador específico para dicho rango. En sus pruebas experimentales se emplean

1. VISIÓN ARTIFICIAL

Clasificadores cuadráticos, Clasificadores de mínima distancia, Redes Neuronales supervisadas con arquitectura de perceptron multicapa (MLP, MultiLayer Perceptron) (Kanno, Akiba, Teramachi, Nagahashi & Takeshi, 2001; Lanitis, Draganova & Christodoulou, 2004) y una aproximación no supervisada con los Mapas autoorganizativos de Kohonen (SOM, Self Organizing Map).



Figura 15: Estimación de edad en una etapa.

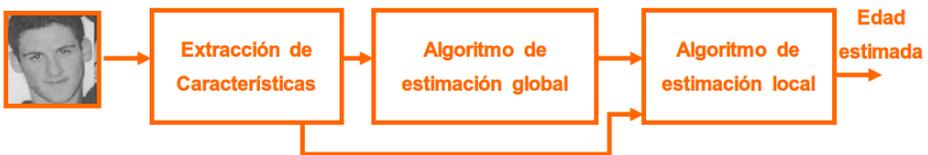


Figura 16: Estimación de edad jerárquica.

Otros clasificadores también empleados para la estimación de edad son las Máquinas de vectores soporte (SVM, Support Vector Machine) (Guo, Fu, Dyer & Huang, 2008).

Modelos de regresión

Para un escenario de estimación de edad planteado como un problema de regresión, Lanitis et al. (Lanitis, Taylor & Cootes, 2002) proponen una Función de Edad con tres formulaciones diferentes (lineal, cuadrática y cúbica) donde los parámetros del modelo se ajustan con un algoritmo genético. Basándose en este trabajo se propone un Modelo de Regresión Múltiple (MLR, Multiple Linear Regression) (Fu, Xu & Huang, 2007; Fu & Huang, 2008) con mejoras de rendimiento significativas respecto a otras técnicas. Las máquinas de vectores soporte para regresión (SVR, Suporta Vector Regression) se han empleado con éxito en los problemas de estimación de edad obteniendo errores absolutos que oscilan entre 5 y 7 años dependiendo de la base de datos de imagen empleada. El SVR constituye un método de regresión muy robusto para problemas cuya salida es unidimensional, pero su extensión a un caso general de múltiples salidas no es trivial. En estos escenarios de múltiple salida resulta útil la técnica de Regresión

1. VISIÓN ARTIFICIAL

imágenes del rostro, podemos concluir diciendo que el problema es interesante y complejo a la vez. Es complejo porque incluso un ser humano se equivoca con frecuencia al realizar dicha estimación pudiendo, en algunos casos, llegar a cometer errores de hasta 10 años. Estos errores se deben a que los signos del envejecimiento son muy personales y dependen de factores como la genética, el estrés o el estilo de vida. Además, hay un problema añadido para estimar la edad utilizando únicamente imágenes de la cara y es que los seres humanos utilizamos otro tipo de información como puede ser el pelo o el tono de voz por lo que prescindir de esos indicadores hace que el sistema de visión se encuentre en desventaja.

Hemos visto que hace ya bastantes años que se empezó a estudiar cómo clasificar la edad en función de factores antropométricos. Los primeros trabajos (Alley, 1988) se basaron en el desarrollo craneo-facial aunque es un modelo que dejó de utilizarse porque no permite describir el perfil de la cabeza de una forma natural y también porque dicho perfil es difícil de extraer a partir de imágenes de caras en 2D. Después se pasó a estimar la edad utilizando la antropometría de la cara, extrayendo y estudiando tamaños y proporciones de la misma (Farkas, 1994), pero también se ha observado que estos modelos antropométricos sólo son útiles para detectar la edad de personas jóvenes y son muy sensibles a la pose, por lo que su uso es limitado.

Un método que ha sido muy popular y que se sigue utilizando para estimar la edad son los modelos de apariencia activos (MAA). Se basan en crear un modelo estadístico de la forma del rostro y otro modelo de la intensidad de los píxeles obteniendo características de textura, analizando para ello un conjunto de puntos de referencia del rostro. Empleados inicialmente para el reconocimiento de caras fueron extendidos (Lanitis, Taylor & Cootes, 2002) añadiéndoles una función que permite explicar la variación de la edad. En relación con los modelos antropométricos presentan la ventaja de que se pueden utilizar para cualquier edad, no sólo para jóvenes, y que son robustos a las posturas del rostro.

Una propuesta diferente (Geng, Zhou, Zhang, Li & Dai, 2006), llamada subespacio del patrón de envejecimiento (AGES) se basa en utilizar secuencias de imágenes de individuos ordenadas temporalmente para modelar el proceso de envejecimiento. Los errores medios obtenidos con este método son de 6.77 años. Sus inconvenientes son que se necesita un conjunto de imágenes de diferentes edades para cada persona y que, al no utilizar textura, no tiene en cuenta la información de las arrugas.

Una extensión de la idea anterior es la del manifold (espacio) de edad propuesta por Fu y colegas (Fu, Xu & Huang, 2007). En esta ocasión en vez de aprender un patrón de envejecimiento para cada individuo se modela un único patrón para varios individuos, a lo largo de diferentes edades, representando cada

1. VISIÓN ARTIFICIAL

edad mediante “diferentes caras”. Aunque requiere disponer de un conjunto elevado de caras para modelar el espacio, un individuo podría aportar una única cara, lo que facilita mucho crear estas colecciones. Basándose en esta idea se han propuesto diversos métodos, alguno de los cuales obtiene errores medios de entre 4.4 (hombres) y 5 (mujeres) años.

La propuesta más reciente, y que es la que mejores resultados obtiene hasta la fecha, se basa en los modelos de apariencia que combinan la descripción de características de textura con características de forma de la cara. Dentro de los numerosos métodos de este tipo destaca el presentado por Guo y su grupo (Guo, Mu, Fu & Huang, 2009) llamado BIF (Biologically Inspired Features). Extraen numerosas características básicas utilizando filtros de Gabor en un proceso por capas que simula la forma de funcionar el córtex visual humano. Dependiendo del conjunto de datos utilizado para validar el método, los errores medios obtenidos, en años, van desde los 3.47 (masculinos en YGA) hasta 4.77 en FG-NET.

Finalmente comentar que cualquiera de los métodos propuestos requiere disponer de un conjunto representativo de rostros, etiquetado, que permita aplicar técnicas de clasificación o de regresión. Las técnicas de clasificación permiten, dado un rostro, indicar a qué clase (normalmente rango) de edad pertenece, y se utilizan sobre todo redes neuronales, mapas SOM y máquinas de vector de soporte (SVM). Las técnicas de regresión devuelven un valor de número de años para un rostro dado y se utilizan modelos de regresión múltiple, regresión basada en imagen y muy frecuentemente máquinas de vector de soporte para regresión (SVR). También aparecen otras soluciones que combinan los métodos anteriores, de diferentes maneras, a las que se considera métodos híbridos.

Como conclusión final, podemos decir que la estimación de la edad a partir de imágenes del rostro es un proceso complejo, que requiere disponer de colecciones etiquetadas de rostros de los que se conoce la edad. Es necesario obtener características del rostro, a ser posible basándose tanto en la forma como en la textura y posteriormente hay que clasificar o estimar la edad utilizando la información extraída de la imagen de la que se quiere conocer la edad y comparándola con la información conocida de las colecciones etiquetadas. Los estudios realizados hasta ahora presentan errores en la estimación o la clasificación ligeramente superiores a los que comete un ser humano, y que van desde aproximadamente los 3 años en el mejor de los casos hasta sistemas que se equivocan en 4 o 5 años al estimar la edad. ■

Referencias

- Ahonen, T., Hadid, A. & Pietikainen, M. (2006). Face Description with Local Binary Patterns: Application to Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12), 2037–2041.
- Alley, T. R. (1988). *Social and Applied Aspects of Perceiving Faces*. Lawrence Erlbaum Assoc.
- Begoña, M. d. (2013). *Maquillaje de ojos con waterline para pieles más maduras: El Antes y Después de Begoña*. Recuperado el 27 de febrero de 2014, de <http://www.aparichimakeup.com/2013/09/maquillaje-de-ojos-con-waterline-para.html>
- Cai, D., He, X., Han, J. & Zhang, H. J. (2006). Orthogonal Laplacian faces for Face Recognition. *IEEE Transactions on Image Processing*. 15(11), 3608–3614.
- Cootes, T., Edwards, G. & Taylor, C. (2001). Active Appearance Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6), 681–688.
- Dib, M. Y. (2011). *Automatic facial age estimation*. Faculty of Computers and Information, Universidad del Cairo. Cairo, Egipto.
- Farkas, L. (1994). *Anthropometry of the Head and Face*. Raven Press.
- Fu, Y., Guo, G. & Huang, T. S. (2010). Age synthesis and estimation via faces: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 32(11).
- Fu, Y., Xu, Y. & Huang, T.S. (2007). Estimating Human Ages by Manifold Analysis of Face Pictures and Regression on Aging Features. En *Proceedings of the IEEE International Conference on Multimedia and Expo*. pp. 1383–1386.
- Fu, Y. & Huang, T. S. (2008). Human Age Estimation with Regression on Discriminative Aging Manifold. *IEEE Transactions in Multimedia*. 10(4), 578–584.
- Geng, X. Fu, Y. & Smith-Miles, K. (2010). *Automatic Facial Age Estimation*. Paper presentado en 11th Pacific Rim International Conference on Artificial Intelligence.
- Geng, X., Zhou, Z. H., Zhang, Y., Li, G. & Dai, H. (2006). Learning from Facial Aging Patterns for Automatic Age Estimation. En *Proceedings of the 14th annual ACM international conference on Multimedia*. pp. 307–316.

1. VISIÓN ARTIFICIAL

- Geng, X. Zhou, Z. H. & Smith-Miles, K. (2007). Automatic Age Estimation Based on Facial Aging Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12), 2234–2240.
- Gonzalez-Ulloa, M. & Flores, E. (1965). Senility of the face basic study to understand its causes and effects. *Plastic and reconstructive surgery*. 36, 239–246.
- Guo, G., Fu, Y., Dyer, C. & Huang, T. S. (2008). Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression. *IEEE Transactions in Imaging Processing*. 17(7), 1178–1188.
- Guo, G., Mu, Y., Fu, Y. & Huang, T. S. (2009). Human Age Estimation Using Bio-Inspired Features. En *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 112–119.
- Hu, H., Otto, C. & Jain, A. K. (2013). Age estimation from face images: Human vs. machine performance. En *Proceedings of the IEEE International Conference on Biometrics*.
- Kanno, T., Akiba, M., Teramachi, Y., Nagahashi, H. & Takeshi, A. (2001). Classification of age group based on facial images of young males by using neural networks. *IEICE Transactions on Information and Systems*. 84(8), 1094–1101.
- Kwon, Y. & Lobo, N. (1994). Age Classification from Facial Images. En *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 762–767.
- Kwon, Y. & Lobo, N. (1999). Age Classification from Facial Images. *Computer Vision and Image Understanding*, 74(1), 1–21.
- Lanitis, A., Taylor, C. & Cootes, T. (2002). Toward Automatic Simulation of Aging Effects on Face Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 24(4), 442–455.
- Lanitis, A., Draganova, C. & Christodoulou, C. (2004). Comparing different classifiers for automatic age estimation. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 34(1), 621–628.
- Ramanathan, N. & Chellappa, R. (2006). Modeling Age Progression in Young Faces. En *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 387–394.
- Riesenhuber, M. & Poggio, T. (1999). Hierarchical Models of Object Recognition in Cortex. *Nature Neuroscience*, 2(11), 1019–1025.

1. VISIÓN ARTIFICIAL

- Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M. & Poggio, T. (2007). Robust object recognition with cortex-like mechanisms. En *Proceedings of the IEEE International Conference on Multimedia and Expo*, 29(3), 411–426.
- Suo, J., Wu, T., Zhu, S., Shan, S., Chen, X. & Gao, W. (2008). Design Sparse Features for Age Estimation Using Hierarchical Face Model. En *Proceedings of the IEEE Conference on Automatic Face and Gesture Recognition*. pp. 1–6.
- Suo, J., Min, F., Zhu, S., Shan, S. & Chen, X. (2007). A Multi-Resolution Dynamic Model for Face Aging Simulation. En *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1–8.
- Tanner, K. (2011). Modeling Automated Detection of Children in Images. (Tesis de máster, Universidad de Rhode Island, 2011). *Dissertations and Master's Theses*, AAI1498256.
- Troell, R. (2014). “Borrar” 10 años de envejecimiento y las arrugas con un “estiramiento facial”. Recuperado el 27 de febrero de 2014, de <http://www.drтроell.com/es/procedimientos/facial-cosmetico/estiramiento-facial-mediofacial/>
- Yan, S., Wang, H., Fu, Y., Yan, J., Tang, X. & Huang, T. S. (2009). Synchronized Submanifold Embedding for Person Independent Pose Estimation and Beyond. *IEEE Transactions in Image Processing*, 18(1), 202–210.
- Yang, Z. & Ai, H. (2007). Demographic Classification with Local Binary Patterns. En *Proceedings of International Conference on Biometrics*. Vol. 4642, 464–473.
- Yan, S., Zhou, X., Liu, M., Hasegawa-Johnson, M. & Huang, T. S. (2008). Regression from Patch-Kernel. En *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1–8.
- Yan, S., Liu, M. & Huang, T. S. (2008). Extracting Age Information from Local Spatially Flexible Patches. En *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing*. pp. 737–740.
- El Dib, M. Y. (2011). *Automatic Facial Age Estimation*. Tesis de máster no publicada. Universidad del Cairo, Cairo, Egipto.
- Zhou, S. K., Georgescu, B., Zhou, X. S. & Comaniciu, D. (2005). Image based regression using boosting method. *IEEE Internation Conference on Computer Vision*. Vol. 1, 451–548.