

Procesamiento Digital de Imágenes

Detección de Rostro

Detección de Rostro

- Durante las décadas de los 70 y 80 del siglo pasado se utilizaban plantillas y mediciones de características geométricas de partes del rostro para detectar y reconocer caras.
- En 1994 Yang y Huang [1] propusieron un método jerárquico basado en conocimiento que va refinando la detección del rostro en cada etapa.
- Li y colaboradores [2] presentaron un método para detectar y seguir rostros en videos en color utilizando detección de la piel a través del color junto con un modelo facial para detectar rostros dentro de regiones de la piel.
- Viola y Jones presentaron un método de detección de objetos [3], que también serviría para detección de rostros que, aún hoy, es muy utilizado.

Algoritmo de Viola-Jones

Idea Principal

Se desplaza una ventana a lo largo de toda la imagen y se evalúa un modelo del rostro en cada ubicación.

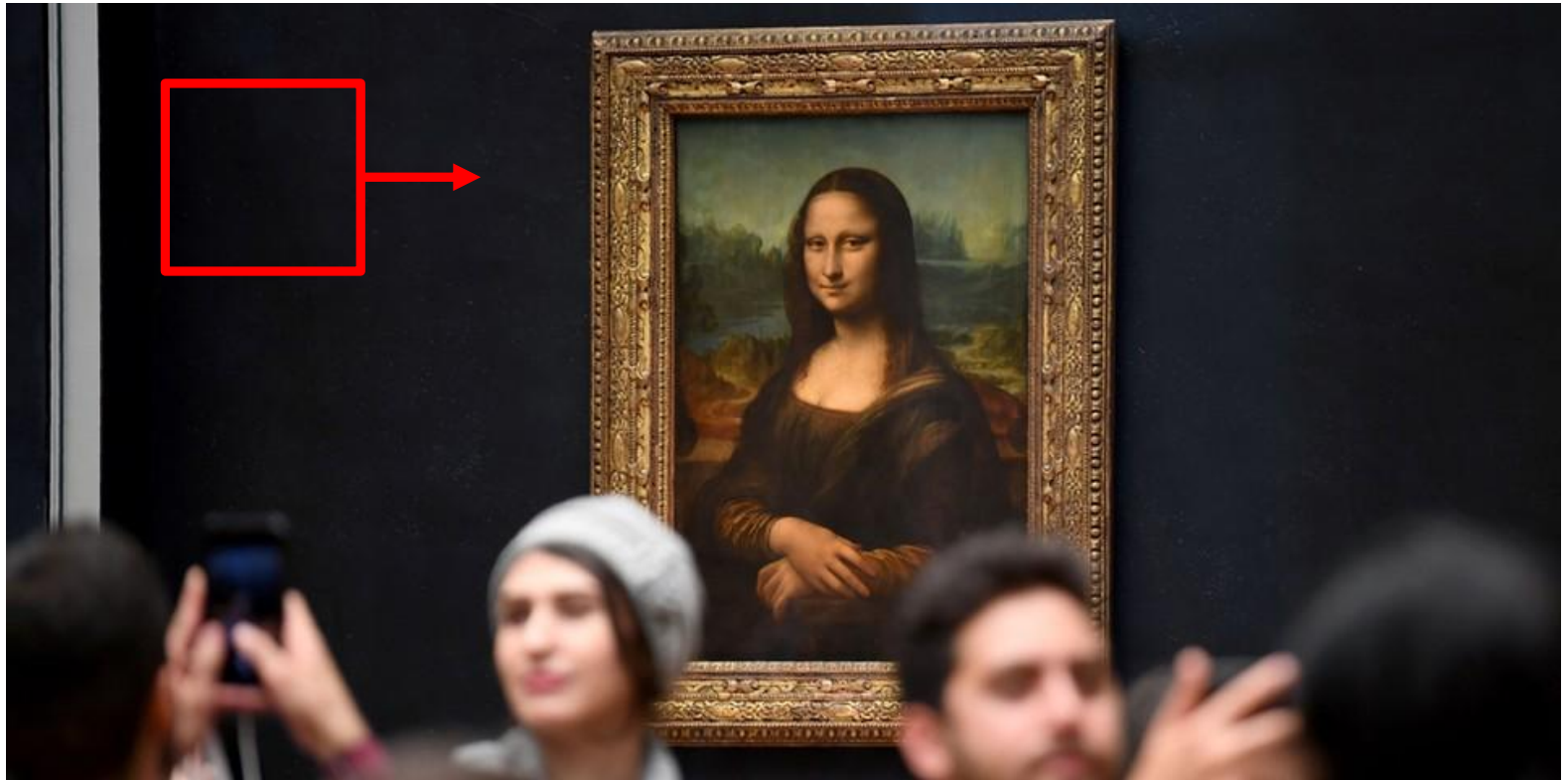


Figura 1: Ventana deslizante para el cálculo de característica

Algoritmo de Viola-Jones

- En el desarrollo de su método, Viola y Jones plantean el uso de características simples que se basan en los niveles de intensidad de los píxeles de la imagen en escala de grises.
- Se utilizan las llamadas **características de Haar** (en inglés Haar-like features), las cuales permiten obtener información de una zona concreta mediante una operación aritmética simple.
- Las características de Haar están definidas por estructuras simples divididas en dos o más regiones rectangulares, como se muestra en la Figura 2.
- El valor de la característica es la diferencia entre la suma de las intensidades de los píxeles contenidos en la región negativa (región negra) y la suma de las intensidades de los píxeles en la región positiva (región blanca).

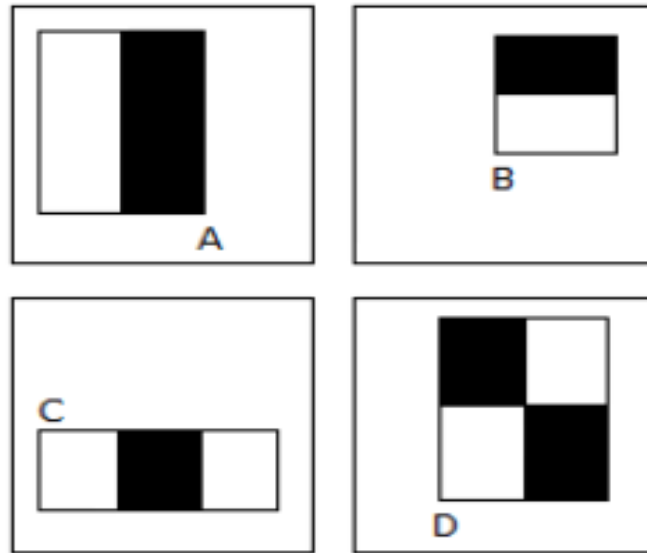


Figura 2: Ejemplos de regiones rectangulares empleadas en la extracción de las características Haar

- Las características Haar capturan muchos rasgos presentes en una cara. En la Figura 3, se muestra como ejemplo que la zona de los ojos suele ser más oscura que la zona superior de las mejillas y el puente nasal suele ser más claro que las mejillas.



Figura 3: Ejemplos de características Haar aplicadas al rostro de una persona

- El cálculo de estas características se puede realizar rápidamente gracias a la utilización de la **imagen integral**. Dada una imagen f de tamaño $n \times m$, su integral F en el punto (x, y) se calcula sumando todos los píxeles a la izquierda y por encima de (x, y) .

1	0	2
3	2	1
1	4	0

1	1	3
4	6	9
5	11	20

Figura 4: Imagen Integral

- La Imagen Integral puede calcularse en una sola pasada de la imagen
- Se construye un clasificador “fuerte” como una combinación lineal de clasificadores “débiles “ (Algoritmo AdaBoost).

Clasificadores en cascada

Viola-Jones presenta un método para combinar clasificadores en una estructura en forma de cascada lo que permite descartar rápidamente regiones que no son de interés en la imagen, denominadas *background*, es decir, todo aquello que no es un rostro, y centrarse en el procesamiento en aquellas zonas que posiblemente contengan una cara. En la cascada, los clasificadores son progresivamente más complejos y tienen una tasa menor de falsos positivos.

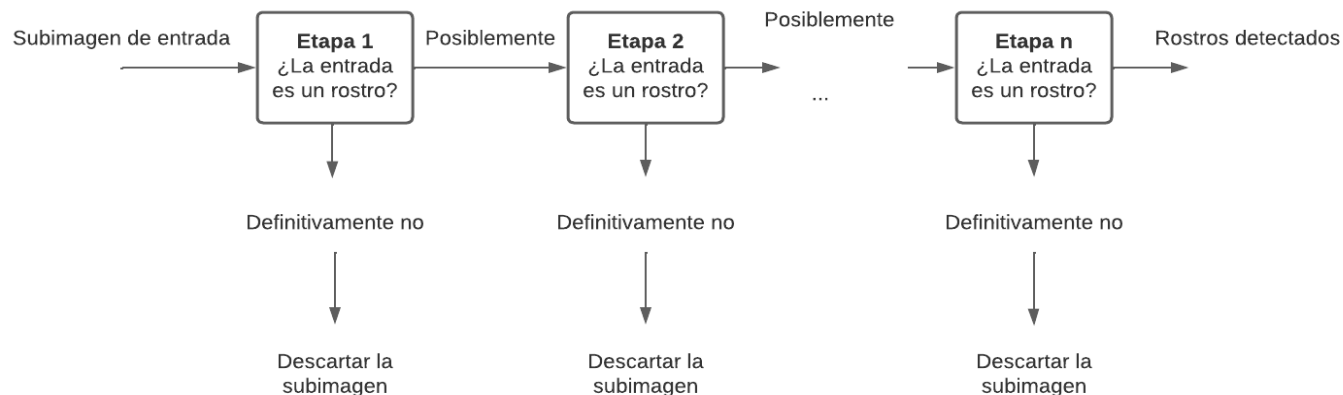


Figura 5: Clasificadores en cascada para detección de rostro

Para la implementación del método se realiza una exploración dentro de la imagen, desplazando una ventana que es donde se evalúan las características para encontrar los rostros dentro de la misma. Los meta-parámetros del algoritmo son el tamaño inicial de la ventana de exploración de la imagen, y la forma de desplazamiento de la misma.

Ejemplo de clasificación con características Haar

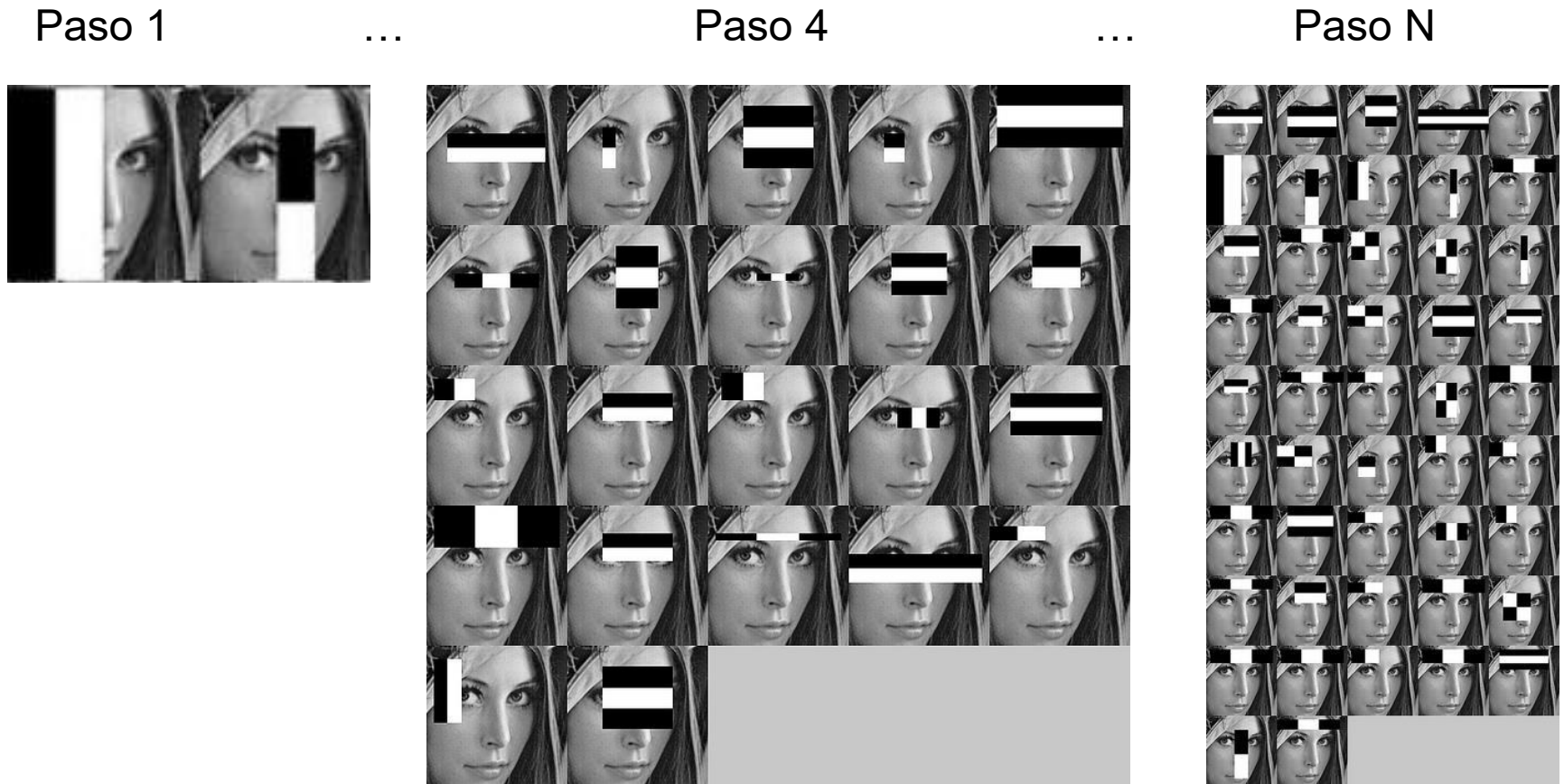


Figura 6: Etapa de extracción de características Haar.

Visualización:

<https://vimeo.com/12774628>

Resumen

Pros

- Computación extremadamente rápida
- Selección eficiente de características
- Detector invariante a escalado y localización
- Se escalan las características pero no la imagen (e.g. imagen piramidal)
- Es un esquema de detección genérico, por lo que puede entrenarse para detectar otros objetos

Cons

- El detector funciona solo para vistas frontales del rostro ($<45^\circ$)
- Es sensible a las condiciones de iluminación
- Pueden aparecer múltiples detecciones del mismo rostro debido a ventanas que se solapan

Referencias

- [1] G. Yuan & H. T. S. Yang, «Human face detection in a complex background,» *Pattern Recognition*, 27(1), 53-63, 1994.
- [2] A. G. & O. G. Y. Li, «Detecting and tracking human faces in videos,» In *Proceedings of the 15th International Conference on Pattern Recognition ICPR-2000*, 2000, pp. 807-810, Vol.1,
doi: 10.1109/ICPR.2000.905528.
- [3] P. Viola and M. Jones, «Rapid object detection using a boosted cascade of simple features,» In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2001*, 2001, Vol. I, pp. 512-518,
doi: 10.1109/CVPR.2001.990517.