

Detección y seguimiento del rostro para interacción

Modesto Castrillón Santana



SIANI
INSTITUTO UNIVERSITARIO
INGENIERIA COMPUTACIONAL



UNIVERSIDAD DE LAS PALMAS
DE GRAN CANARIA

Contenidos

- Proyectos previos
- Seguimiento
- Detección de personas
- Aplicaciones



Proyectos previos: DESEO

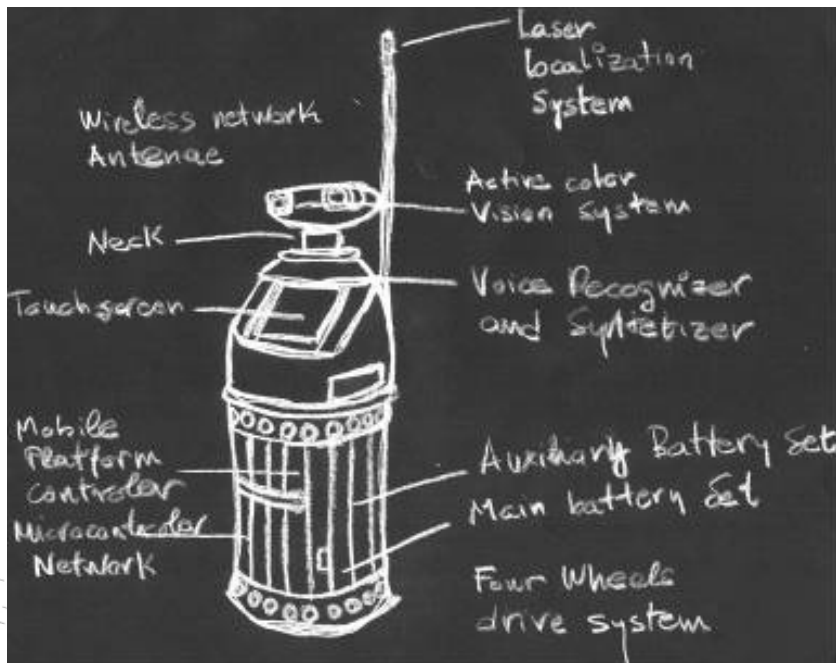
- DEtección y SEguimiento de Objetos

DESEO capaz de seguir personas en tiempo real haciendo uso de DSPs



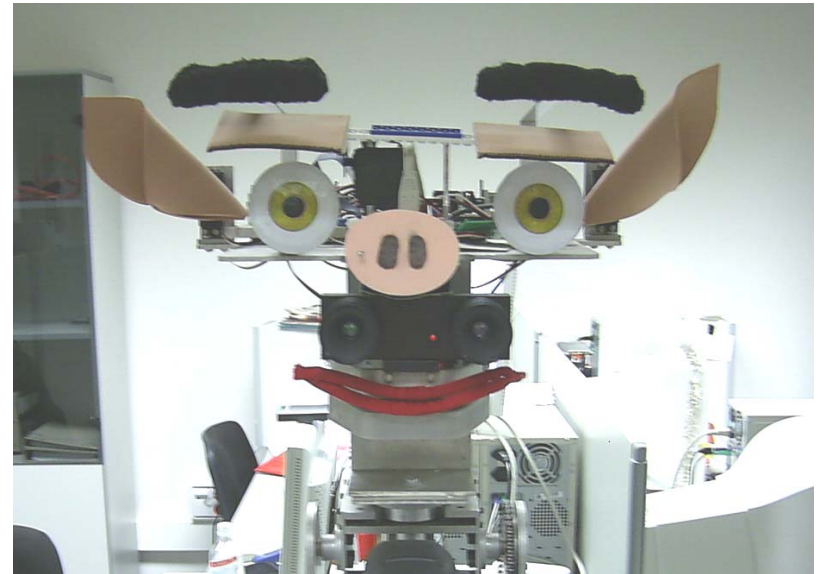
Proyectos previos: eldi

Mascota del Museo Elder de Ciencia y Tecnología de Las Palmas de Gran Canaria. Operando diariamente desde Dic. de 1999 hasta mediados 2001.



Proyectos previos: CASIMIRO

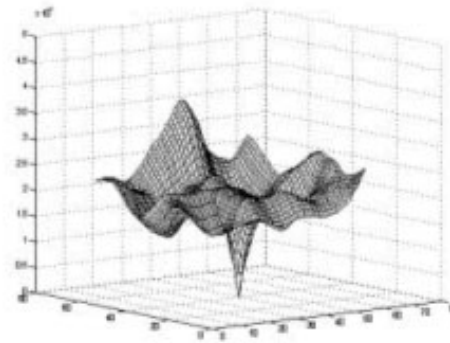
- Detección de personas y localización
 - Visión Omnidireccional
 - Localización de sonido
- Fijación
 - Atención Audiovisual
 - Habitación
- Identificación
 - Detección del rostro
 - Identificación
 - Propietario
- Reconocimiento de acciones
 - Afirmaciones y negaciones
 - Acciones básicas
- Habilidades de comunicación
 - Emociones
 - Habla



Seguimiento

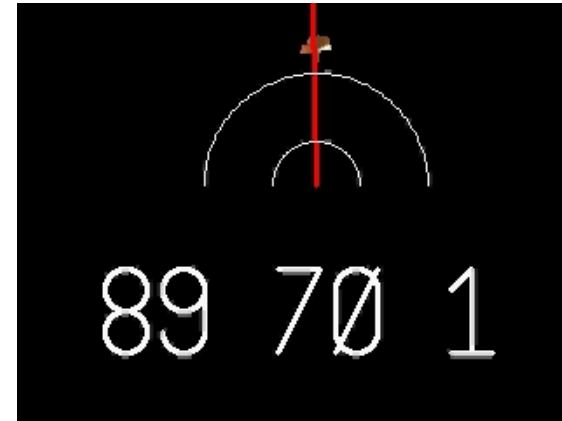
- Patrones

$$D(u, v) = \sum_A |I(u + i, v + j) - P(i, j)|$$



Seguimiento

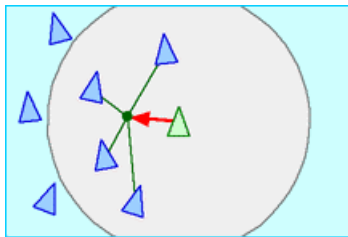
- Omnidireccional



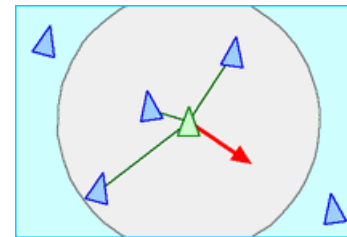
Seguimiento

- Boids

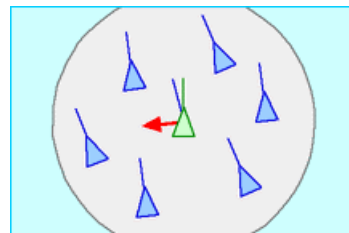
Considerado como una propiedad emergente de un conjunto de partículas, donde el comportamiento de cada una viene afectado también por el del grupo.



Cohesión: tendencia a moverse hacia la posición promedio del grupo



Separación: tendencia a evitar acercarse en exceso



Alineamiento: tendencia a dirigirse en similar dirección que el grupo



Seguimiento

- Boids

Dada la imagen CIE L*a*b Image, S_c , S_v , S_{pc} , para cada partícula:

Regla 1: Color y Topografía.

Movimiento hacia elementos atractivos dado el encaje de color o su alto valor de gradiente. La vecindad define la velocidad.

Regla 2: Agrupamiento

Movimiento hacia el centroide de la manada (S_c)

Regla 3: Alineamiento

Adaptación a la velocidad media de la manada (S_v)

Regla 4: Predicción

Movimiento hacia el centroide previsto (S_{pc})



Detección de personas

Técnicas:

- Implícitas o basadas en patrones, p.e. Viola-Jones [1]
- Explícitas o basadas en conocimiento: movimiento, color, etc.

Esquema de Viola-Jones aplicado a:

- Rostros frontales [1]
- Cabeza y hombros [2], así como piernas y brazos [3].
- Elementos faciales: ojos, nariz y boca [4].

[1] P. Viola and M. J. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. In Computer Vision and Pattern Recognition, 2001.

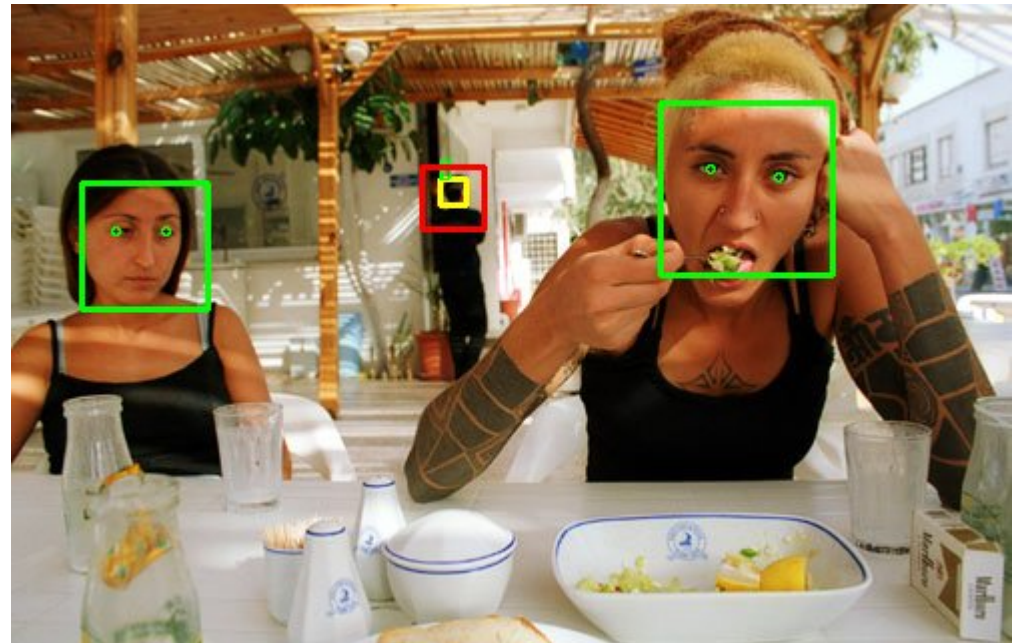
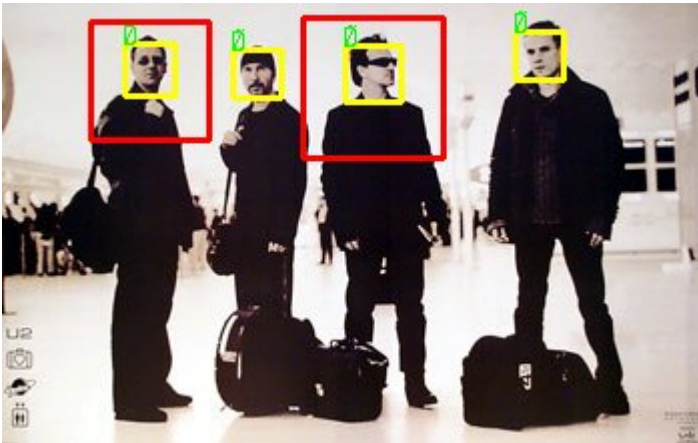
[2] [H. Kruppa, M. Castrillón, B. Schiele. "Fast and Robust Face Finding via Local Context" VS-PETS Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, Nice \(France\), 2003.](#)

[3] [M. Castrillón Santana and Q. Vuong. Combining Human Perception and Geometric Restrictions for Automatic Pedestrian Detection, CAEPIA05](#)

[4] Clasificadores para OpenCV [disponibles](#)

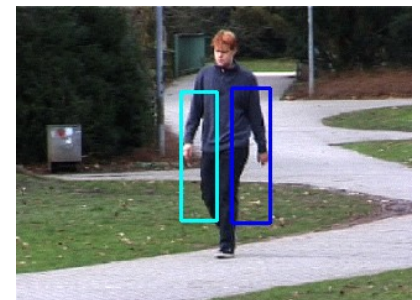
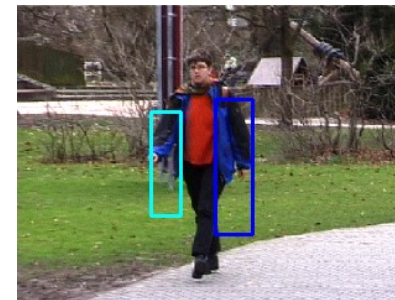
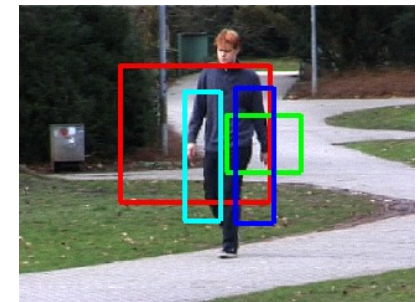
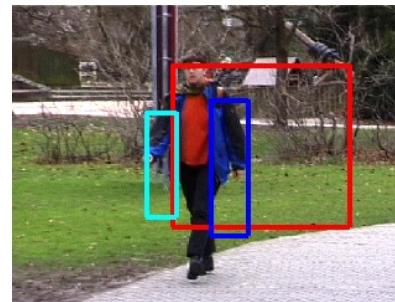
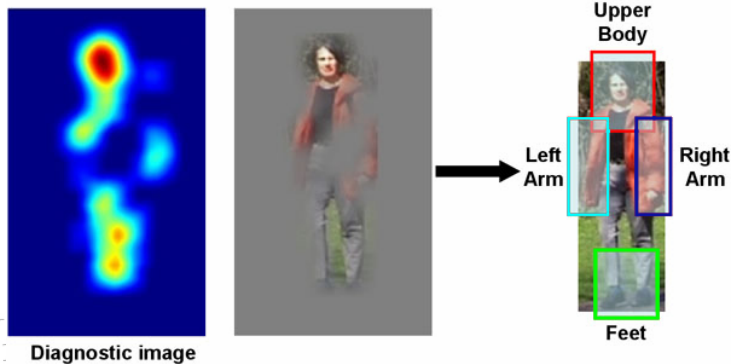
DetECCIÓN DE PERSONAS

El detector basado en el contexto local permite detectar caras/cabezas más allá de la limitación de caras frontales.



DetECCIÓN DE PERSONAS

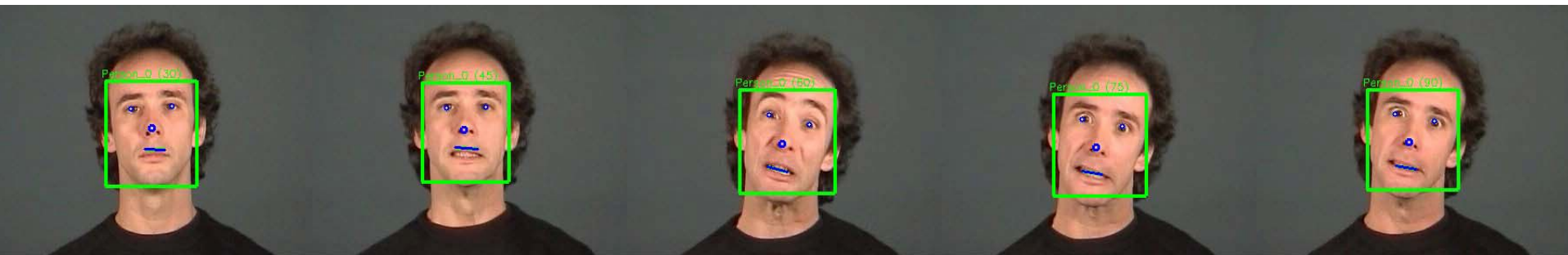
- Uso de burbujas [1] para determinar las zonas que permiten a un sujeto detectar la presencia de personas.
- Entrenar y aplicar múltiples detectores sobre una imagen.
- Tras aplicar restricciones geométricas se consiguen eliminar falsas detecciones [2].



DetECCIÓN DE PERSONAS: ELTOS. FACIALES

Clasificadores individuales de libre disposición obtenidos a partir de 7000 rostros anotados manualmente.

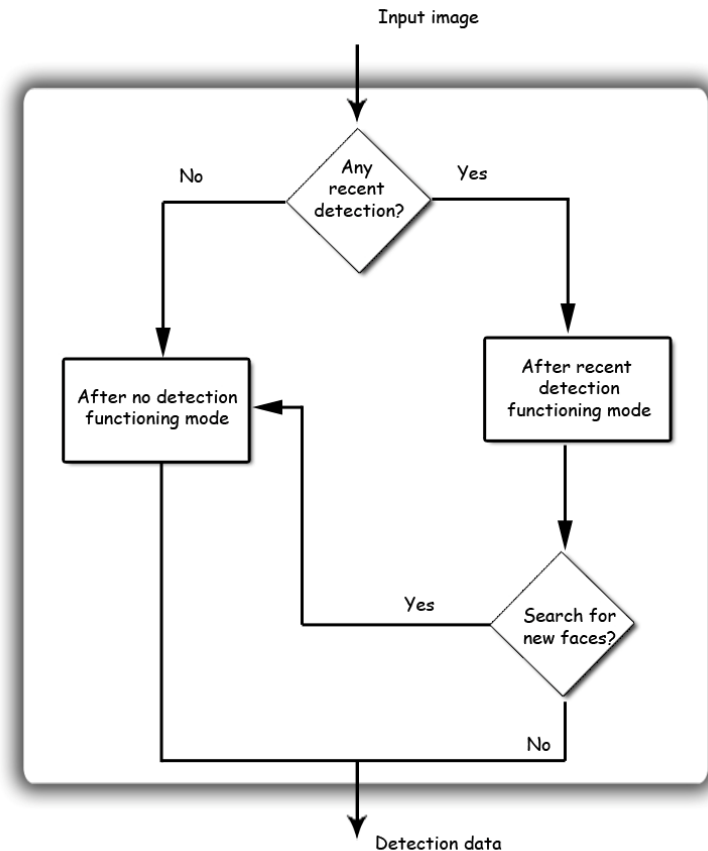
Múltiples candidatos de elementos faciales se escogen en base a la estructura promedio.



DetECCIÓN DE PERSONAS: ENCARA2

ENCARA2 combina distintas técnicas con el objetivo de aprovechar sus ventajas individuales:

- Detectores basados en el marco de Viola-Jones [2,3], sirven de inicializadores.
- Distintas técnicas (Color de piel, seguimiento multinivel, ojos, rostro, etc.) se combinan de forma robusta haciendo uso de la coherencia temporal en vídeo.



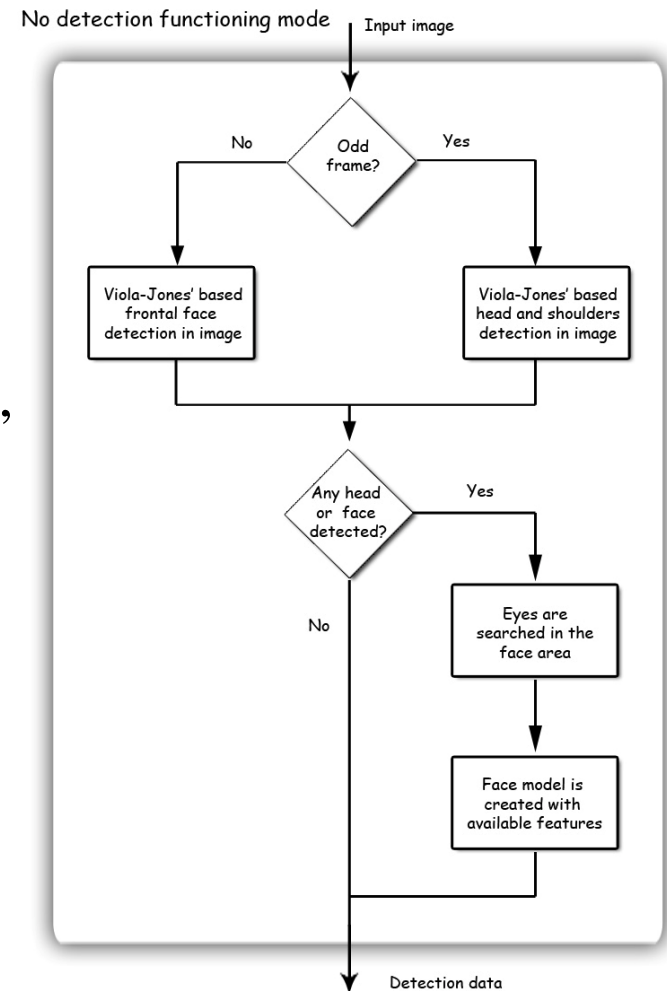
- [1] [M. Castrillón, O. Déniz, C. Guerra, M. Hernández "ENCARA2: Real-time Detection of Multiple Faces at Different Resolutions in Video Streams" Journal of Visual Communication and Image Representation, ISSN 1047-3203, vol 18, issue 2, pp. 130-140, April 2007.](#)
- [2] P. Viola and M. J. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. In Computer Vision and Pattern Recognition, 2001.
- [3] [H. Kruppa, M. Castrillón, B. Schiele. "Fast and Robust Face Finding via Local Context" VS-PETS Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, Nice \(France\), 2003.](#)

DetECCIÓN DE PERSONAS: ENCARA2

Una primera detección permite la creación del modelo particular para la cara detectada:

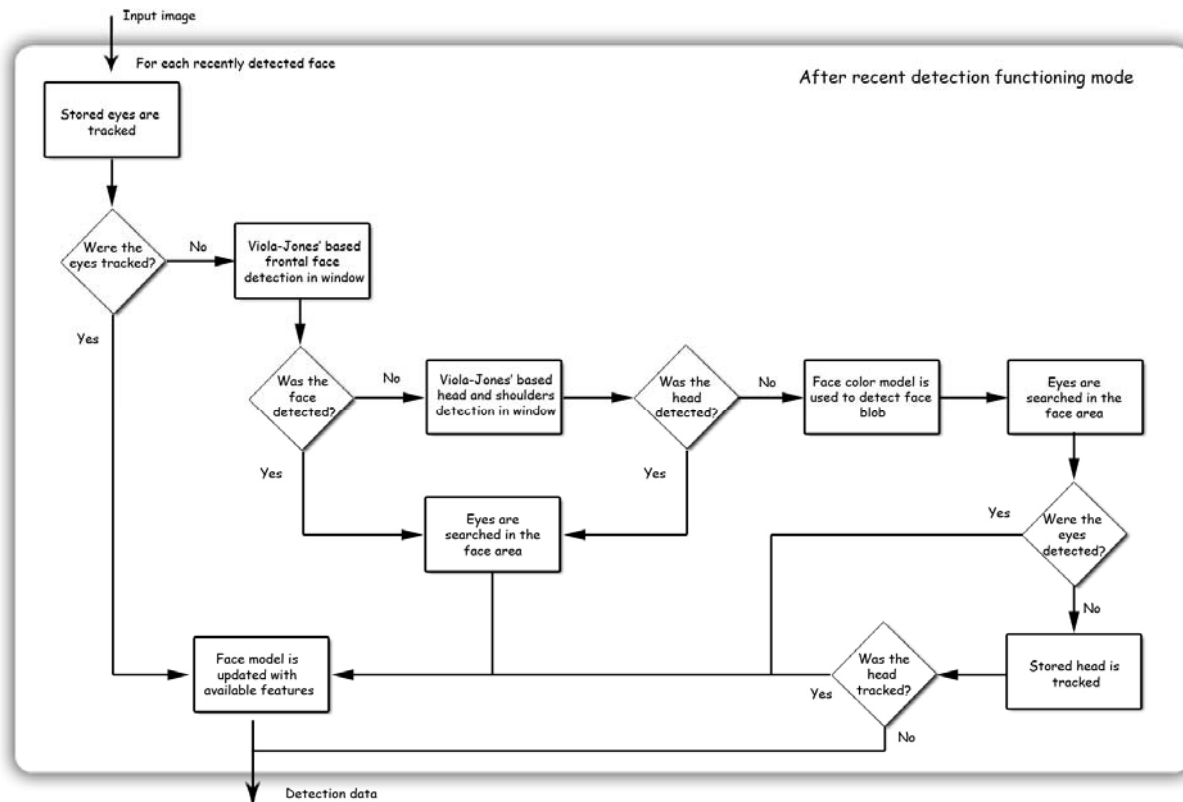
$$f = \left\langle pos, size, red, green, tshirt_{color}, leye_{pos}, leye_{pat}, \right. \\ \left. reye_{pos}, reye_{pat}, nose_{pos}, nose_{pat}, \right. \\ \left. mouth_{pos}, mouth_{pat}, face_{pat} \right\rangle$$

La coherencia temporal se emplea en las imágenes posteriores.

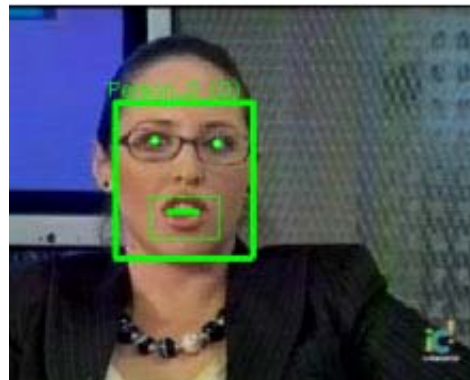


DetECCIÓN DE PERSONAS: ENCARA2

Una detección provoca una búsqueda localizada en las siguientes imágenes, aplicando en primer lugar las técnicas más rápidas.



Detección de personas: ENCARA2



Aplicaciones

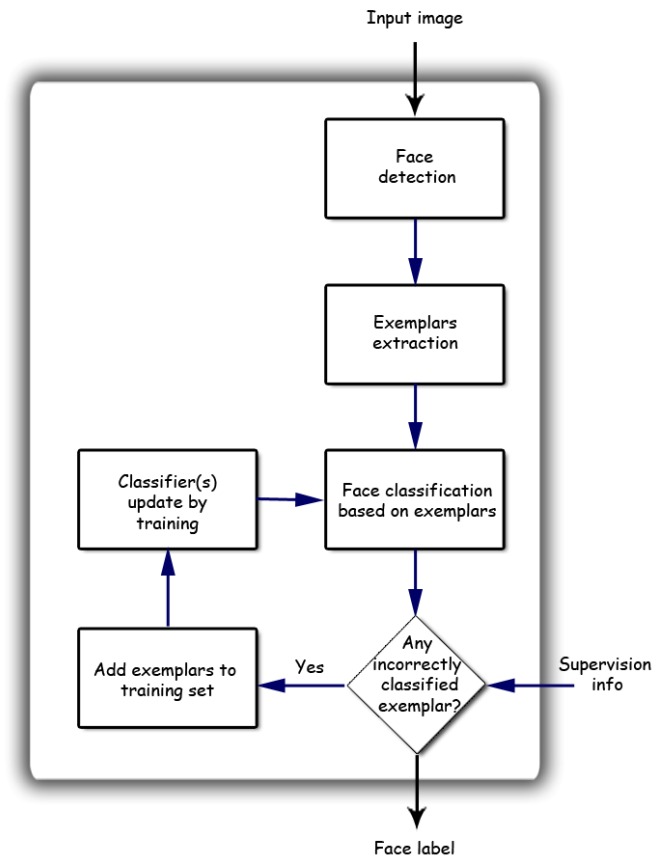
- Reconocimiento

Aprendizaje desde cero.

Extracción automática de los ejemplares.

Supervisión para aumentar la experiencia del sistema.

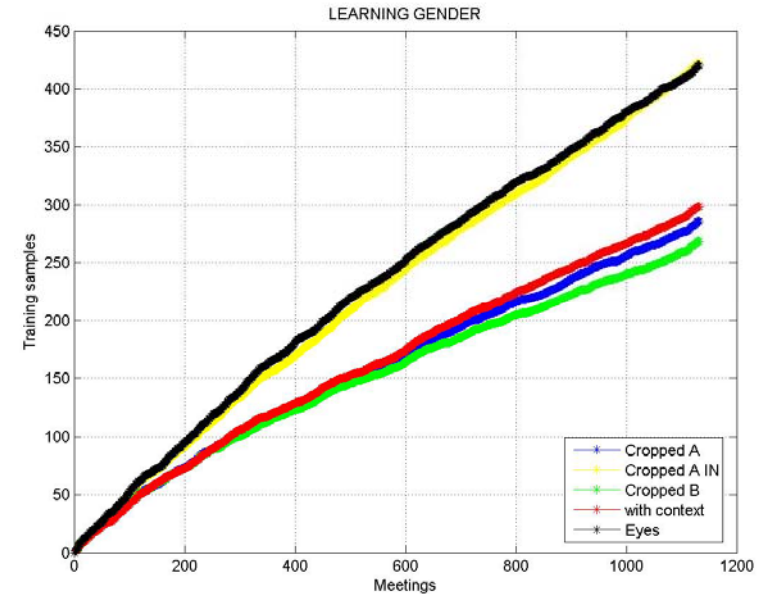
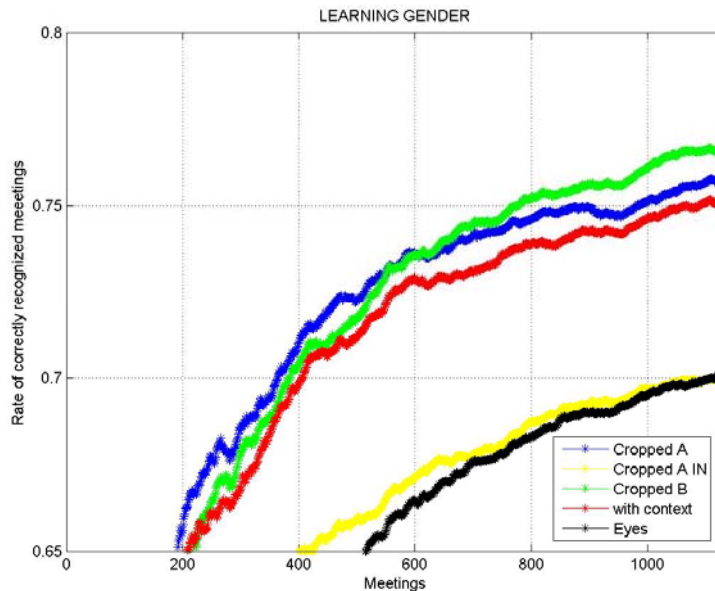
Aproximación basada en PCA incremental y SVM.



Aplicaciones

- Reconocimiento: Género

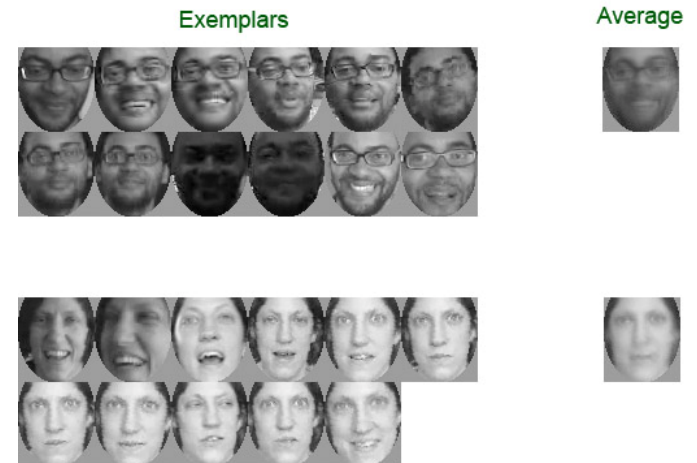
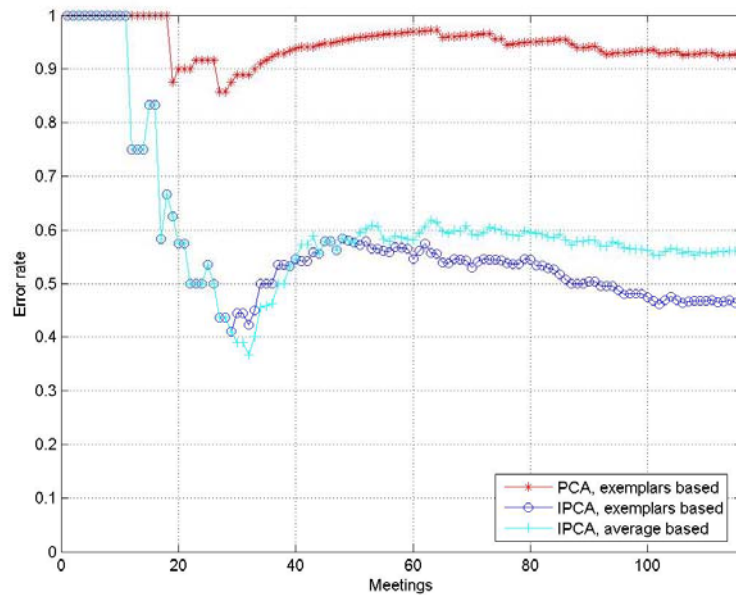
- 1130 encuentros.
- PCA incremental (IPCA)
- Tendencia de reducción del error.
- Entrenamiento (15 msecs), actualización PCA 50-80 msecs.



Aplicaciones

- Identidad

- Cinco encuentros con el mismo individuo (24 identidades).
- Pruebas con ejemplares e imagen promedio
- PCA e IPCA



Aplicaciones

- Integración

- El seguimiento de actores humanos en 3D [1] suele requerir inicialización manual.
- El rostro como punto de partida para localizar el torso y los brazos.



[1] J. Schmidt, B. Kwolok, and J. Fritsch. Kernel Particle Filter for Real-Time 3D Body Tracking in Monocular Color Images. FGR , pp 567-572, 2006.

Aplicaciones

