

Sesión 2 de diciembre 2021

Tema 9: Detección, extracción de características, segmentación y reconocimiento de objetos (Parte 2)

El **reconocimiento de objetos** consiste en identificar los objetos en una imagen y su posición, a partir de algún conocimiento previo sobre los objetos.

Tiene dificultad debido a:

- Presencia de otros objetos no modelados
- cambio de iluminación
- cambio de punto de vista
- oclusión
- escala

Ej. de uso: vehículos sin conductor, visión o navegación robótica, identificación de enfermedades en imágenes médicas, reconocimiento facial...

Reconocimiento facial:

Se suele reconocer una posible posición en la imagen para una cara (mediante color de piel, identificación de ojos...) y luego se reconoce a la persona utilizando técnicas de aprendizaje

Ej. de uso: Desbloquear dispositivos móviles, identificación de personas en redes sociales, pagos...

Clasificación: asignar una etiqueta, clasificar en clases.

Localización: marcar con una caja un objeto en la imagen (posición dentro de la imagen).

Detección: combinación de las anteriores.

Técnicas:

1. *Conincidencia de plantillas:* se utiliza en problemas sencillos, se busca en la imagen una plantilla
2. *Segmentación de imágenes y análisis de blobs*
3. *reconocimiento de características*
4. *machine learning / Deep learning*

3.Reconocimiento de características:

Tenemos una imagen (modelo) de un objeto a reconocer.

Extraemos las características SIFT y las guardamos.

En una nueva imagen (escena) donde queremos buscar ese objeto, extraemos las características SIFT, y las comparamos con las características SIFT del modelo.

También se debe encontrar la transformación entre el modelo y escena (escala, rotación, traslación).

4. Machine learning / Deep learning:

Para machine learning se da un conjunto de imágenes iniciales (entrenamiento) , luego ante una nueva imagen lo que hacemos es extraer las características de esa nueva imagen y a partir de esas características encontradas durante el entrenamiento es capaz de decirnos a qué objeto pertenece sin embargo deep learning utiliza otro enfoque, emplean modelos como por ejemplo son las redes convolucionales lo hacen para aprender automáticamente las características inherentes en la imagen pueden aprender las diferencias entre objetos.

Reconocimiento de caras

Algoritmo viola&jones:

- Es muy rápido, en pocos segundos nos localiza caras en una imagen.
- Es muy fiable, muy pocos falsos positivos.

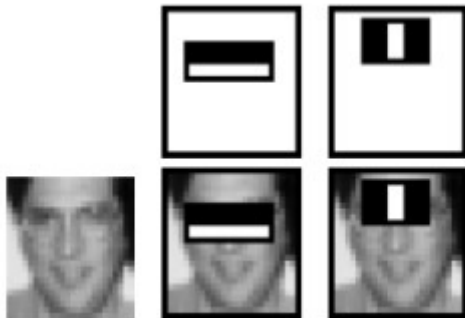
Puntos débiles:

- Solo sirve para caras frontales o con muy poco giro (necesita ver el puente de nariz y los ojos)
- Cambio de luminosidad.

Estructura del método:

- Extrae características (usando una imagen integral, extrae muchísimas características)
- Selecciona características (adaboost)
- Cascada de clasificadores (consigue más velocidad)
- Las caras comparten algunas características comunes:
La región del puente de nariz es más claro que los ojos
- La región de los pómulos es mas clara que los ojos

El algoritmo localiza nariz-ojos, y utiliza los valores claro-oscuro:



Calcula la diferencia de luminosidad de los píxeles del rectángulo.

Utilizando una imagen integral permite calcular la diferencia entre 2 zonas de la imagen utilizando solamente 2 accesos a memoria.

Reconocimiento de movimiento:

Se necesita capturar el esqueleto mediante algoritmos de entrenamiento para luego poder detectar su movimiento con tecnologías de video, sistemas de sensores/marcadores, o sistemas de mapas de profundidad (z-cam, Kinect)).

Enlace relacionado: <https://www.xataka.com/robotica-e-ia/algoritmos-reconocimiento-facial-estan-mejorando-su-precision-al-identificar-personas-mascarilla>

Las mascarillas han supuesto un problema para los sistemas de reconocimiento facial.

Estos algoritmos se valen de la posición de los ojos, la nariz, y la boca para identificar a los usuarios.

Según los estudios del NIST (National Institute of Standards and Technology), los algoritmos más precisos tienen una tasa de error del 0,3% al detectar a personas sin mascarilla. Con los algoritmos mas precisos y una mascarilla cubriendo el 70% de la cara, la tasa del error es del alrededor del 5%, los algoritmos menos precisos siguen fallando, con la tasa de error entre 10% y 40%.

Otras conclusiones se deben con el color y el tamaño de la mascarilla (las negras y rojas provocan mas fallos que las blancas y azules).