

# Sesión 11-Tema 9: Detección, extracción de características y segmentación

## Tema 9: Detección, extracción de características y segmentación – Resumen

### Detección– Transformada de Hough (T.H)

Uno de los problemas en visión automática es la detección de puntos de interés (p.d.i) en las imágenes procesadas, véase bordes, rectas, cuadrados, círculos, etc. Para la detección de estos existen varios métodos, entre ellos la Transformada de Hough.

Basado en lo que devuelve Canny al aplicarlo a imágenes, trata de corregir los errores presentados por el mismo, ruido, líneas discontinuas, etc. Hough detecta las figuras marcadas siempre y cuando se puedan expresar de manera sencilla matemáticamente. Además, tiene un sistema de votaciones para decidir cómo es realmente la figura entre otras posibilidades.

T.H. trabaja con coordenadas polares en un espacio parametrizado. Una recta en el espacio de coordenadas (X, Y), se representa cómo:

$$r = x \cdot \cos \theta + y \cdot \sin \theta$$

Dónde 'r' es la distancia que hay entre el origen de coordenadas y la recta, y 'θ' es el ángulo que tiene con respecto al origen.

Para detectar que tipo de figura es, recorre la imagen de Canny, y cada vez que encuentra un punto perteneciente a un borde de la imagen, sobre ese punto, calcula diferentes rectas con diferentes orientaciones, y vota cada una de las rectas con 1. Cuando encuentre un segundo punto, hace lo mismo, y al final, la recta en la que haya encontrado más puntos tendrá más votos, por lo que es la ganadora. Al final, donde interseccionen las diferentes rectas con mayor puntuación, dará la figura correspondiente.



Figura 1: Transformada de Hough

### Extracción de características – Algoritmo Sift

A la hora de extraer esos p.d.i., debemos aplicar algoritmos capaces de detectar la misma característica en diferentes imágenes, de modo que hagan 'matching' los p.d.i. de ambas imágenes que sean similares. Para ello estudiaremos el algoritmo Sift.

Presentado en 2004, Sift (Scale Invariant Feature Transform), define los p.d.i. (bordes, texturas, esquinas, etc. Contengan mucha información de la imagen). A través de la vecindad del punto a evaluar. El punto se presenta en un Vector descriptor, el cual almacena la información del punto y su vecindad (128 píxeles).

Es capaz de localizar las características de la imagen en múltiples escalas de la misma imagen. Primero analiza una ventana pequeña de la imagen, después una ventana mediana y otra grande, y se queda con la ventana que maximiza la información de la imagen, de modo que, una imagen dada en múltiples escalas es capaz de encontrar el mismo p.d.i. sin importar la escala dada.

Una vez localizada, aplica filtros blur con diferentes magnitudes a la imagen, calcula las diferencias entre imágenes par a par, superpone los resultados dados (restas de imágenes) y comprueba el punto P. Si es su plano, el plano + 1 y el plano - 1 ese punto es un Máx. o Mín. con respecto a sus vecinos, es un p.d.i. También se suele aplicar un umbral para reducir el número de estos.

Lo siguiente es obtener la orientación del punto, con el p.d.i. y los vecinos, calcula el histograma de las posibles rotaciones del punto, aquel con la barra más elevado es la orientación actual. Una vez que tenemos la localización (Transformada de Hough), escala y orientación (Sift), debemos obtener el vector descriptor del punto de interés.

Calcula el gradiente de los vecinos, obtiene el histograma con la frecuencia (magnitud) de los gradientes, y a partir de ello construye el vector, el cual tendrá un tamaño de  $4 \times 4 \times 8 = 128$  datos (píxeles). Funciona muy bien a la hora de hacer matching en imágenes con una alta oclusión.



*Figura 2: SIFT permite de manera directa poder encontrar un objeto específico en una escena. Se puede apreciar su robustez ante rotación, oclusión y escala.*

**Segmentación de imágenes**

Ahora debemos extraer partes de imágenes con p.d.i. similares, de modo que nos sea más fácil trabajar sobre esas imágenes (reconocimiento de tumores en medicina). Para ello hay 2 algoritmo, k-medias y segmentación basada en regiones.

El algoritmo k-medias trata de lo siguiente, se establece un nº k, el cual será el nº de clasificadores que contendrá el algoritmo. Estos puntos se deben establecer con mucho cuidado en el espacio de coordenadas de la imagen, si no, puede que no encuentre una solución óptima. Para ello, se pueden establecer de manera aleatoria o pseudoaleatoria (dividiendo la imagen en k sectores, y en cada sector un k).

Una vez establecidos, etiquetamos los datos según la cercanía con el cluster (k) más cercano, normalmente por distancia euclídea. Una vez etiquetados, establecemos una nueva posición para el cluster con la media entre las posiciones de los datos etiquetados con ese cluster. Repetimos el algoritmo hasta lograr un conjunto estable de resultados, de este modo, clasificamos los datos según su cluster y podemos extraer los datos.

La segmentación por regiones es lo más utilizado en este caso, existen dos tipos, el crecimiento de regiones y la partición de regiones. El crecimiento de regiones es la más utilizada, esta planta una o varias semillas en la imagen (en varios puntos/píxeles), si los vecinos cumplen parte de las características del punto, aumenta la región de tamaño y comprueba de nuevo los vecinos. Si no cumple, puede comenzar una nueva región con las características de ese punto.



*Figura 3: Segmentación de imagen por regiones.*

La partición por regiones comprueba si todos los píxeles de la imagen comparten características, sino cumplen, reduce la sección, así hasta que todos los píxeles cumplen con las características. Se suelen combinar ambos algoritmos para obtener resultados más precisos, cómo detección de tumores en imágenes y viceversa.

Segmentación