

Sesión 25 de noviembre 2021

## **Tema 9: Detección, extracción de características, segmentación y reconocimiento de objetos**

### **Transformada de Hough:**

Permiten identificar primitivas geométricas sencillas en la imagen.

Se utiliza un proceso de “votación”. Cada punto de la imagen votará por aquella ecuación que cumpla. Daremos como salida las celdas con suficiente soporte o evidencia, esto es, aquellas con un “número suficiente de votos”.

Este algoritmo tiene varios **puntos críticos**:

**Ineficiencia:** La complejidad espacial de la construcción del espacio de parámetros aumenta de forma factorial con el número de estos.

En consecuencia, el método exhaustivo solamente es aplicable a **primitivas sencillas** e incluso en estos casos se hace necesaria una discretización muy somera para satisfacer los **requerimientos de memoria**.

**Ajuste de sensibilidad:** aunque es un factor de menor importancia, el ajuste del tamaño de celda y del número mínimo de votos es clave para evitar un elevado número de falsos positivos.

### **Algoritmo SIFT**

Permite la extracción de **características de una imagen**, localizar en una imagen puntos importantes o relevantes, o si son **estables a algunas transformaciones de imagen**.

Para que sea capaz de detectar características de la imagen tiene un **descriptor** que contendrá algún tipo de información de la imagen alrededor del punto.

### **Localización multiescala**

Busca diferencias entre las escalas, si un píxel es u min o Max entre sus vecinos, es un punto de interés de la imagen, y para que sea multiescala, escala las imágenes a diferentes tamaños, para escalar los puntos de interés a diferentes tamaños. Así el algoritmo será capaz de encontrar los mismos puntos clave sin importar el tamaño de imagen.

Un problema puede ser que se encuentren muchos puntos, pero muchos son inestables, se eliminan los que tienen un contraste bajo o los que se encuentran sobre una arista mediante un filtrado.

Es necesario definir la **orientación** de la característica de la imagen, utiliza un histograma de orientaciones ponderado en la escala calculada y se calcula la orientación dominante, con esto se podrá detectar el punto de interés, aunque este en diferentes rotaciones.

La **segmentación de imágenes** es el proceso de extraer zonas de la imagen con el mismo color/nivel de gris/textura para identificarlas automáticamente.

#### **Algoritmo K-Medias:**

Primero se buscan k puntos

luego para cada punto mira cual es su grupo mas cercano (el k más cercano)

y luego esas K se mueven hasta el punto medio

esto se repite hasta que no haya ningún cambio del conjunto de datos o hasta un numero de pasos que se le indique.

la pertenencia de un dato a un conjunto dependerá del problema, se puede calcular mediante la distancia o la probabilidad de pertenencia a un clúster, usando su media y varianza.

#### **Puntos críticos:**

-Debemos indicar las K.

-Una mala inicialización puede llevar más tiempo.

-Puede no encontrar la solución mas optima dependiendo de la inicialización de las K.

Se puede utilizar distancia o probabilidad para determinar si un dato pertenece a un conjunto.

#### **Segmentación basada en regiones**

el objetivo es encontrar regiones de la imagen homogéneas según algún criterio, hay dos maneras de encontrarlo:

1. **Según el crecimiento de regiones:** empezamos con regiones pequeñas y las hacemos crecer o las mezclamos, usando un criterio de similaridad, para el cual deberemos definir un test para saber si un nuevo pixel y, adyacente a algún pixel de R se puede añadir a la región. Este método se puede hacer empezando por cualquier píxel de la imagen, pero es mejor lanzar varios puntos de partida.

2. **La partición de regiones:** empezamos con regiones grandes y las vamos dividiendo, usando un criterio de homogeneidad, este proceso es el inverso del anterior, partimos de toda la imagen, y utilizamos un criterio para partir la región.

-si se cumple el criterio, no se sigue dividiendo,

-si no se cumple, se divide.

Tiene una complejidad mayor al tener que manejar una estructura de datos adicional.

---

**Enlace relacionado:** <https://programmerclick.com/article/1147927769/>

Este artículo entra en profundidad en el algoritmo Sift, nos da una explicación general, SIFT (Scale Invariant Feature Transform) significa Scale Invariant Feature Transform. El operador SIFT utiliza un vector de características de 128 dimensiones para describir los puntos de características detectados en una imagen, por lo que una imagen Después del algoritmo SIFT, se expresa como un conjunto de vectores de características de 128 dimensiones.

Además nos introduce a la definición matemática de la escala de una imagen:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

Siendo  $I(x, y)$  la imagen original, y  $G(x, y, \sigma)$  la función gaussiana con espacio de escala variable.

$\sigma$  representa el tamaño del espacio de escala,  $*$  representa la operación de convolución vista en el tema anterior. La función gaussiana se puede definir como:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$