

Sesión 10-Tema 8: Visión Artificial

Tema 8: Sistemas de Percepción: Visión Artificial – Resumen

Los sistemas de percepción nos permiten detectar el entorno que nos rodea de manera artificial, entre ellos se pueden encontrar cámaras de visión, micrófonos (reconocimiento de voz), sensores sonar (obtener distancias de un objeto con el tiempo que tarda en volver la onda emitida), infrarrojos, etc..

El símil biológico a la visión artificial es el ojo humano. Diferentes sistemas capaces de reproducir el comportamiento del ojo humano pueden ser:

- CCD o CMOS: Incluidos en cámaras digitales actuales, capturan la imagen con mejor o menor calidad.
- Pin-hole: Procesar la imagen a través de una retícula pequeña, la cual discretiza la entrada. Empleado en los objetivos de las cámaras digitales.

Una imagen representada en datos es una matriz con valores entre 0 y 255 si hablamos en escala de grises. Cuanto mayor sea el valor, más blanco es el color que representa. Si la imagen viene con 3 canales de colores, en cada posición de la matriz tendremos 3 valores, uno para cada tipo de color RGB.

Según el tipo de visión artificial, existen diferentes operaciones básicas, de nivel bajo nos encontramos los histogramas, ruido y suavizado de imágenes. De nivel medio y alto, detección de aristas, puntos característicos, reconocimiento de escenas, etc.

Cómo hemos comentado antes, existen varios modelos de colores:

- Binarias: Blanco o negro (0, 1 -> binarizadas)
- Escala de grises: Valores de 0 a 255 -> Tiene Relación de Orden Total ($x_1 < x_2 \rightarrow g(x_1) < g(x_2)$)
- 3 canales de colores RGB: Valor de 0 a 255 por cada canal.
- HSV

Podemos aplicar diferentes filtros a las imágenes mediante una convolución de una función sobre la imagen. Si $f(x)$ obtiene los datos de (valores) de un píxel en pantalla, y $g(x)$ es un filtro a aplicar en ese píxel, el proceso de aplicar el filtro a la entrada y obtener una salida, eso se llama convolución. La función $g(x)$ tiene varios sinónimos, tales como: Función de convulsión, filtro, elemento estimulante, máscara de convulsión, etc.



Figura 1: Convulsión

La imagen representa cómo se aplica el filtro a la entrada. Selecciona el punto donde va a aplicar la función $g(x)$, mira el entorno de vecindad, aplica el filtro en cuestión y obtenemos el nuevo punto con el filtro aplicado. Para ser más exactos, la función de convulsión es la siguiente:



Figura 2: Ecuación de convulsión

Existen distintos tipos de filtros:

- Promedio: Sustituye el punto con el promedio de los 9 puntos.

- Gaussiano: Función gaussiana de media 0 y varianza σ
- Media: La media de los valores de los 9 puntos.
- Mediana: Mediana de los 9 puntos. Elimina muy bien efectos de ruido, y 'sal y pimienta', pero pierde nitidez.

El histograma de una imagen es una función definida por K niveles, la cual indica el nº de píxeles que tienen cierta intensidad, es decir, la frecuencia con la que aparecen ciertos valores en la escala de grises. Si los valores se concentran en el primer cuartil, la imagen es muy oscura, mientras que si se concentran en el último, la imagen es muy clara.

Cuando tenemos alguno de estos problemas, la imagen recibe un proceso de ecualización, tratamos de abrir la representación de los grises, distribuyéndolos en todo el espectro y volvemos a mapear la imagen con la distribución hecha. Sirve tanto para aclarar imágenes oscuras como para lo contrario, la función es la siguiente:



Figura 3: Función de mapeo después de ecualizar

La detección de contornos en las imágenes se basa en detectar cambios radicales de intensidad en los valores de los píxeles, de este modo detectamos las aristas presentes en esta. Para la detección de los contornos existen diferentes métodos:

- Gradiente: Buscamos diferencias en el eje horizontal mediante el gradiente en X (mirar el píxel y el píxel de la derecha) y diferencias en el eje vertical (píxel y píxel de arriba), la unión de los gradientes nos muestra una imagen con los contornos resaltados.
- Operadores de Sobel: Calcula el contorno de la imagen con el promedio de esta.
- Detector Canny: Identifica bien el alto contraste de manera local, además suaviza la imagen resultante aplicando el filtro de gauss (convulsión).
- Puntos esquina: Basado en los niveles de grises de la vecindad del punto para detectar mejor las esquinas.
- Detector Nitzberg-Harris: Gradiente + filtrado lineal (gauss)

Harris Corner Detector

Índice Wiki individual