

Proyecto 1: detección y clasificación de formas

5/8/21



Alejandro Silvestri

Visión artificial

Contenido

[Introducción](#)

[Proyecto](#)

[Proceso](#)

[Output](#)

Introducción

La detección y la clasificación son problemas generales de la visión artificial. Detección significa localizar objetos en la imagen y obtener sus coordenadas en píxeles, y usualmente más datos relativos a su tamaño y forma. Clasificación significa reconocimiento del objeto dentro de una lista corta de objetos conocidos, y usualmente implica también señalar los casos en que el objeto no se pudo reconocer.

En visión artificial se llama **ambiente controlado** a la escena especialmente preparada para el proceso que se requiere, de manera que el software no tenga que compensar defectos de la imagen. Implica, entre otros:

1. iluminar adecuadamente (usualmente con varias fuentes) para evitar sombras que puedan confundir el proceso, evitar el degradado de la iluminación, y lograr el contraste adecuado
2. elegir criteriosamente la posición de una cámara fija, usualmente frontal cuando la escena es plana (es decir, con la cámara perpendicular al plano de la escena)
3. elegir color y textura del fondo para no confundir al proceso; típicamente se usa un fondo de color homogéneo (liso) de un color de contraste (por ejemplo, una mesa negra para reconocer objetos claros, un dibujo con fibra negra sobre una hoja blanca)
4. región de interés definida: o bien la escena comprende la imagen completa, o bien se recorta programáticamente el rectángulo de la **región de interés** (ROI), descartando el resto, de modo que al final la escena comprende la totalidad de la zona procesada

Este proyecto plantea la detección y clasificación de **contornos** de objetos en un ambiente controlado. Está diseñado para ser un proyecto inicial, entregable en dos semanas.

Proyecto

Los alumnos elegirán al menos tres objetos reconocibles por su contorno, usarán una imagen de cada uno como referencia de cada tipo de objeto, y les asignarán un nombre a cada uno.

Prepararán el ambiente controlado para evitar correcciones por software que de otro modo pueden resultar muy demandantes, y así poder concentrarse en el objetivo del proyecto.

El sistema detectará y clasificará los objetos en la imagen de la webcam en tiempo real.

Proceso

Se sugieren los siguientes pasos para el proceso de la imagen de la webcam:

1. Convertir la imagen a monocromática
2. Aplicar un *threshold* con umbral ajustable con una barra de desplazamiento
 - a. se pueden incluir opciones de ajuste automático
3. Aplicar operaciones morfológicas para eliminar ruido de la imagen
 - a. es buena idea incluir una barra de desplazamiento para ajustar el tamaño del elemento estructural
4. El sistema puede obtener varios contornos en una misma imagen, y los debe procesar a todos individualmente
5. Filtrar contornos que se pueden descartar de antemano
 - a. para quitar contornos espúreos indeseables
 - b. por ejemplo, por tener un área muy pequeña
6. Compara cada contorno con todos los objetos de referencia, usando `matchShapes()`
 - a. establecer un umbral de distancia máxima de validez
 - i. es conveniente una barra deslizante para ajustar este valor

-
- ii. uno global o uno diferente para cada objeto de referencia
 - b. clasificar la forma según la menor distancia entre los candidatos válidos
 - i. si no hay ninguno, la forma es desconocida
7. Generar la imagen anotada

Output

La salida del sistema es una ventana con la imagen original anotada de la siguiente manera:

- **localización de objetos relevantes**
 - en verde para objetos reconocidos
 - alternativamente se puede asignar un color particular a cada clase de objeto
 - en rojo para objetos desconocidos
 - evitando mostrar contornos de ruidos y elementos espúreos
 - se puede anotar el contorno del objeto o un rectángulo que lo contenga
- etiqueta con el nombre del objeto

Se puede decorar la imagen con anotaciones adicionales. Las anotaciones se pueden hacer sobre la imagen completa o sobre la región de interés, y se puede usar la versión color o en escala de grises. Es válido y deseable mostrar otras ventanas con pasos intermedios.

ves un borde, una descripción de ese borde y con machine learning haces una clasificación