

Tarea 3 EL7008 – Primavera 2020

Detección de objetos usando RANSAC, Hough y descriptores SIFT

Profesor: Javier Ruiz del Solar

Auxiliar: Patricio Loncomilla

Ayudantes: Juan Pablo Cáceres, Hans Starke, Javier Smith, José Villagrán

Fecha enunciado: Jueves 15 de Octubre de 2020

Fecha entrega: Miércoles 4 de Noviembre de 2020

El objetivo de esta tarea es la detección de objetos usando puntos de interés y descriptores SIFT, usando tanto RANSAC como Hough. Para esto, se obtendrán los puntos de interés y descriptores (no es necesario programar esta parte), se generarán calces entre descriptores parecidos entre dos imágenes y se usará RANSAC o Hough para encontrar un grupo de calces que sean compatibles con una misma transformación de semejanza entre una imagen de prueba (pru) y otra de referencia (ref).

1.- Un par punto de interés – descriptor se define como:

$$(x, y, \sigma, \varphi, \vec{d})$$

donde (x, y, σ, φ) corresponde a la posición, escala y orientación asociada al punto de interés (se puede ver como la posición, largo y orientación de una flecha). El descriptor \vec{d} corresponde a un vector de características que describe la vecindad del punto de interés.

2.- Una transformación de semejanza se define como:

$$\begin{pmatrix} x_{PRU} \\ y_{PRU} \end{pmatrix} = e \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x_{REF} \\ y_{REF} \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$

donde e corresponde al factor de escala entre las imágenes, θ es una rotación, y (t_x, t_y) es la traslación.

3.- Una transformación afín se define como:

$$\begin{pmatrix} x_{PRU} \\ y_{PRU} \end{pmatrix} = \begin{pmatrix} m_{XX} & m_{XY} \\ m_{YX} & m_{YY} \end{pmatrix} \begin{pmatrix} x_{REF} \\ y_{REF} \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$

Alternativamente, se puede describir de este modo (dejando los parámetros de la transformación afín separados en un vector):

$$\begin{pmatrix} x_{REF} & y_{REF} & 0 & 0 & 1 & 0 \\ 0 & 0 & x_{REF} & y_{REF} & 0 & 1 \end{pmatrix} (m_{XX} \ m_{XY} \ m_{YX} \ m_{YY} \ t_x \ t_y)^T = \begin{pmatrix} x_{PRU} \\ y_{PRU} \end{pmatrix}$$

Lo cual se puede reescribir resumidamente como:

$$Ax = b$$

De este modo, si se conoce A y b , y se quiere resolver el siguiente problema:

$$\operatorname{argmin}_x \|Ax - b\|^2$$

la solución se puede obtener así:

$$x^* = (A^T A)^{-1} A^T b$$

4. La transformación de semejanza que proyecta un punto de interés $(x_{REF}, y_{REF}, \sigma_{REF}, \varphi_{REF})$ de la imagen de referencia en un punto de interés $(x_{PRU}, y_{PRU}, \sigma_{PRU}, \varphi_{PRU})$ en la imagen de prueba se define como:

$$e = \frac{\sigma_{PRU}}{\sigma_{REF}}$$

$$\theta = \varphi_{PRU} - \varphi_{REF}$$

$$t_x = x_{PRU} - e(x_{REF} \cos(\theta) - y_{REF} \sin(\theta))$$

$$t_y = y_{PRU} - e(x_{REF} \sin(\theta) + y_{REF} \cos(\theta))$$

5. El error de proyección, dada una transformación (e, θ, t_x, t_y) , se define como:

$$e_{PROY} = \left\| e \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x_{REF} \\ y_{REF} \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} - \begin{pmatrix} x_{PRU} \\ y_{PRU} \end{pmatrix} \right\|$$

6. Para RANSAC: Dado un conjunto de calces, el consenso asociado a una transformación corresponde a un subconjunto de los calces cuyo error de proyección es menor a un umbral, es decir, los que son compatibles con la transformación. RANSAC intenta encontrar una transformación con el mayor consenso posible.

7. Para Hough: Se discretiza el espacio de las transformaciones de semejanza en celdas. La celda por la cual debe votar una transformación en el espacio de Hough se determina del siguiente modo (con $dxBin$ en pixeles y considerando un offset de 500):

$$i = \text{floor} \left(\frac{t_x}{dxBin} + 0.5 \right) + 500$$

$$j = \text{floor} \left(\frac{t_y}{dxBin} + 0.5 \right) + 500$$

$$k = \text{floor} \left(\frac{\theta}{dangBin} + 0.5 \right) + 500$$

$$z = \text{floor} \left(\frac{\log(e)}{\log(2.0)} + 0.5 \right) + 500$$

Hough (en el caso particular de esta tarea) funciona haciendo que cada hipótesis de transformación de semejanza vote por una celda, y luego intenta encontrar celdas tales que la probabilidad de que representen una transformación correcta sea alta. En el caso de esta tarea, se debe encontrar la celda con el mayor número de votos.

Se pide realizar los siguientes pasos, cuyos resultados deben ser incluidos en el informe:

- 1) Implementar la función que permite generar una transformación de semejanza a partir de un calce. Se debe notar que en OpenCV las orientaciones de los descriptores son entregadas en grados.
- 2) Obtener la transformación de semejanza con mayor consenso, y los calces compatibles usando RANSAC:
 - a. Programar la selección de calces al azar para generar la hipótesis
 - b. Programar el cálculo del consenso (calces compatibles con la hipótesis)
 - c. Programar la selección de la hipótesis con el mayor consenso
 - d. Programar la selección de los calces compatibles con la mejor hipótesis
 - e. Probar con los 4 pares de imágenes entregados, graficando los calces
- 3) Obtener la transformación de semejanza con mayor votación en el espacio de Hough y los calces compatibles:
 - a. Programar el cálculo de los índices de celdas (i, j, k, z) a partir de un calce
 - b. Programar el almacenamiento de los índices en una lista
 - c. Programar la selección de los calces compatibles con la celda más votada

- d. Probar con los 4 pares de imágenes, graficando los calces
- 4) A partir de un conjunto de calces, calcular y graficar la transformación afín asociada usando mínimos cuadrados:
 - a. Implementar la función que encuentra una transformación afín usando mínimos cuadrados a partir de un conjunto de calces
 - b. Implementar la función que dibuja el romboide correspondiente a una transformación afín
 - c. A partir de los calces encontrados en el punto 2e, calcular una transformación afín usando mínimos cuadrados y graficar el romboide asociado. Hacer esto para los cuatro pares de imágenes.
 - d. A partir de los calces encontrados en el punto 3d, calcular una transformación afín usando mínimos cuadrados y graficar el romboide asociado. Hacer esto para los cuatro pares de imágenes.
- 5) Evaluar el efecto de variar los tres parámetros indicados de RANSAC (error máximo para inliers, consenso mínimo y número máximo de intentos) sobre las detecciones obtenidas, usando un par de imágenes en que haya una detección inicial correcta.
- 6) Evaluar el efecto de variar los tres parámetros indicados de Hough (tamaño del bin en posición, en orientación y número mínimo de votos) sobre las detecciones obtenidas, usando un par de imágenes en que haya una detección inicial correcta.
- 7) Analizar las ventajas y debilidades relativas de Hough y RANSAC, considerando los resultados obtenidos sobre los 4 pares de imágenes.

Se entrega un código base en Python que ya calcula los puntos de interés, los descriptores y los calces iniciales. Las funciones que calculan la transformación a partir de un calce, la que realiza RANSAC, la que realiza Hough, la que calcula una transformación afín por mínimos cuadrados y la que grafica el romboide deben ser implementadas para lograr que el sistema funcione adecuadamente.

El código además incluye indicaciones que pueden ayudar a implementar la tarea.

El informe debe contener como mínimo: introducción, partes relevantes del código dentro de cada sección del informe, resultados (mostrando las imágenes resultantes), análisis de los resultados y conclusiones generales.

Los informes y los códigos deben ser entregados a más tardar el día miércoles 4 de Noviembre a las 23:59, mediante u-cursos. Cada día de retraso (incluyendo fines de semana) será castigado con un punto de descuento en la nota.

Importante: La evaluación de esta tarea considerará el correcto funcionamiento del sistema, la inclusión de los resultados de los pasos pedidos en el informe, la calidad de los experimentos realizados y de su análisis, la inclusión de las partes importantes del código en las secciones del informe correspondientes, así como la prolijidad y calidad del mismo.

Se agrega una pauta que indica la estructura esperada del informe

Parámetros recomendados para RANSAC: 100 intentos, umbral de error de posición 100, umbral de consenso 30