

Tema 5: Ejercicio Filtro de Partículas y de Kalman

Motivación

El sistema de seguimiento propuesto se basa en la detección de las manos en la secuencia de vídeo. Para ello, se ha aprovechado un modelo ya implementado disponible en la librería de “mediapipe”. Con este tipo de detector, se ha hecho un sencillo juego de “Atrapar la galleta”, cuyas instrucciones de juego son las siguientes:

- Se detectan un máximo de dos manos en la pantalla, y se dibuja su esqueleto. Es preferible que sea una mano izquierda y una derecha, para que el indicador de puntuación funcione correctamente.
- Al empezar el juego, la galleta se inicializa de forma aleatoria en la pantalla.
- Los jugadores tienen que coger la galleta, y evitar que el adversario se la arrebate.
- A la hora de iniciar el juego, se puede estipular el tipo de filtro a usar: de Kalman o de partículas.

La idea del juego se sacó de clase, donde la profesora la planteó como posibilidad. La motivación, es realizar un juego donde se ponga en práctica el tipo de filtro estudiado, y donde este sirve para mejorar la experiencia de realidad aumentada.

Implementación

Como ya se ha comentado, se usa la librería “mediapipe”, que incluye funciones para detectar manos de forma sencilla. Al tratarse de un juego competitivo, no se ha inicializado la galleta en ninguna mano, sino de forma aleatoria en la pantalla. Aun así, es sencillo adecuar el código para que empiece en posesión de algún jugador.

Cuando se detectan las manos, se consigue el esqueleto entero de las mismas, que aprovechamos para conseguir las coordenadas de la falange baja del dedo corazón, que se usará como posición de la mano.

Filtro de Kalman

Se calcula, en cada iteración, la distancia euclidiana de las dos manos a la galleta. Se declara ganadora aquella mano que se encuentre más cerca de la galleta, y por tanto esa mano se queda con la galleta. Mientras tenga la galleta esa mano, será muy difícil que se la quiten, ya que la galleta realiza un seguimiento de esa mano. La única forma, será por tanto que la mano adversaria se ponga delante de esta, momento en que la primera mano deja de ser detectada, y la galleta pasa por tanto a la nueva.

Filtro de partículas

En este caso, se calcula la distancia euclidiana de cada partícula a la galleta, la cual cosa ya representa una mejora muy grande en cuanto a rendimiento respecto a la implementación inicial (para cada partícula, se llamaba 121 veces al método `esError()`, que comprobaba si un píxel era del color de la bola a detectar).

El número de partículas usado es de 300, con media 0 y varianza 10. El número de partículas tiene un impacto directo sobre los fps del juego, y se ha considerado que con 300 funciona lo suficientemente fluido. De estos parámetros, se puede decir que:

- Aumentando el número de partículas se puede mejorar un poco más la precisión, aunque también se puede reducir el número a menos de cien partículas, y seguirá funcionando relativamente bien.
- La varianza puede aumentarse más, con lo que aumentará el área donde se inicializarán partículas, aunque tiene un impacto directo sobre la precisión de la detección, que “vibrará” más cuanto mayor sea la varianza.

Resultados

De los dos filtros, los mejores resultados los proporciona el filtro de Kalman, que consigue trayectorias mucho más suavizadas para los objetos que el filtro de partículas. Además, consigue predecir bastante bien la posición de la galleta en momentos de oclusión, y durante movimientos bruscos de la mano.

El mayor problema del filtro de partículas es en momentos de oclusión, donde la galleta se vuelve loca por la pantalla, hasta que se vuelve a detectar la mano. También presenta problemas cuando los movimientos son muy rápidos o bruscos, que puede perder por un momento el objeto detectado.

Finalmente, acabamos obteniendo un juego bastante decente, donde el filtro de Kalman proporciona un suavizado de los movimientos y predicciones en momentos de oclusión. Además, el impacto en los fps de este filtro no representa un problema. Se pueden encontrar adjuntados dos vídeos de ejemplo, capturando el seguimiento de la mano por la pantalla y viendo cómo responde cada filtro ante oclusiones, así como el desarrollo normal de una partida de robar la galleta.