

PERCEPCIÓN

Curso 2020/2021

Práctica Final - 3 en raya

AUTOR1: Allae Ghayat Zayoun

AUTOR2: Álvaro Torné Zambrano

Puerto Real, 2021

Índice

1. Introducción	2
1.1. Descripción y objetivo	2
2. Metodología	2
2.1. GUI	2
2.2. Reconocedor de imágenes	2
2.3. Reconocedor de voz	3
3. Resultados	3

1. Introducción

1.1. Descripción y objetivo

Nuestro proyecto consiste en una interfaz con la que se podrá jugar al clásico juego del 3 en raya. En él, será necesario dos jugadores ya que no está permitido el poder jugar contra el ordenador.

Mediante diferentes técnicas de reconocimiento de voz, y de una forma muy sencilla, cada uno de los jugadores podrá realizar su jugada simplemente diciéndolo. También, se ha implementado la opción de poder terminar una partida ya comenzada en un tablero real. Con el objetivo de poder reutilizar dicho tablero cada jugador dispondrá de monedas, en principio de 10 y 50 céntimos, para completar el tablero.

Como en el juego original, gana el primero que complete una línea, ya sea horizontal, vertical o diagonal, a menos que se termine en un empate. Finalmente, añadir que el proyecto está diseñado e implementado de tal forma que su escalabilidad sea lo más sencilla posible en caso de querer emular el 4 en raya o superior.

2. Metodología

2.1. GUI

Para la interacción con el usuario hemos implementado una simple interfaz que consta de 1 imagen, en la que se irán mostrando los resultados y dos botones, uno para iniciar una partida desde cero y otro botón para continuar con una partida cargando un tablero previamente.

El funcionamiento de la GUI es bastante sencillo e intuitivo ya que disponemos de un constructor de bloques donde podemos ir arrastrando los botones, etiquetas de textos, imágenes, etc...

Una vez que hemos decidido que elementos tendrá nuestro panel (funciona como el frontend de nuestra interfaz), la GUI de matlab nos aporta sus respectivas funciones (backend) donde podremos implementar diferentes funcionalidades. Como el objetivo de la práctica es el reconocedor de voz y de imagen, y no la implementación de la interfaz, consideramos que no hace falta profundizar más en la explicación de esta.

2.2. Reconocedor de imágenes

Para poder reconocer el tablero y las fichas hemos aplicado varias técnicas. Primero partimos de una foto del tablero y de su estado. Luego, como casi todas

las veces, la pasamos a escala de grises y obtenemos los bordes de la imagen aplicando una Laplaciana suavizada con un filtro gaussiano, es decir, el operador LoG. Solo queda obtener los círculos de las monedas, que representan las fichas, y sus características.

Para ello, aplicamos la función "imfindcircles", obteniendo los centros de los círculos y sus radios. Gracias a los radios sabremos de que tipo de monedas se trata, y por tanto, de que jugador es. Es por ello que las monedas usadas son de 10 y 50 céntimos, ya que son fácilmente distinguibles.

Finalmente, antes de llegar a este punto, consideramos otras opciones. La primera fue la de no usar monedas, sino pintar mediante círculos de diferente color (azul y rojo) las fichas. Sin embargo, nos encontramos con el inconveniente de que se tendría que pintar un tablero por partida, lo que podría provocar que se cambiarán sus dimensiones y no se pudiera saber en que posición está cada ficha.

Además, al ser formas irregulares difícilmente se podría asemejar a un círculo. El último problema que encontramos al probar esta forma fue reconocer los colores. Si bien es verdad que al pasarlo a escala de grises hay cierta diferencia, esta depende en gran medida de la claridad de la imagen. Es por ello, que optamos por usar monedas y diferenciar mediante tamaño y no por el color.

2.3. Reconocedor de voz

Para el reconocedor de voz hemos utilizado las técnicas vistas en las prácticas. Estas consisten en obtener las características delta cepstrum de las palabras a reconocer, que son las palabras "uno", "dos" y "tres", con las que se podrá designar la fila y columna a elegir.

Una vez tenemos dichas características almacenadas y empieza la partida, obtenemos las mismas características para la palabra dicha por el jugador, comparando ambas mediante el método DTW, Dynamic Time Warping, con lo que obtenemos la palabra dicha.

En todo el proceso de obtener la muestra aplicamos diferentes técnicas. Estas son la preénfasis, segmentación, inventanado y cálculo del inicio y fin de la palabra.

3. Resultados

Una vez finalizado el proyecto hemos podido ver que el aspecto más determinante a la hora de una buena experiencia por parte de los usuarios es el

reconocedor de voz. Esto se debe a que es lo único que puede fallar en el programa. Si la fotografía del tablero ha sido tomada con una adecuada iluminación e inclinación no habrá ningún problema. En cambio, para el reconocedor de voz, es necesario una correcta articulación y que se diga toda la palabra en el tiempo en el que se toma la muestra, para evitar que salga cortada y no pueda ser identificada correctamente.

Para comprobar la efectividad del reconocedor de voz realizamos una prueba, en la que se grabó 60 veces los números, 20 cada uno, pudimos comprobar que acertó 54 de las 60 veces, siendo las palabras "dos" y "tres" las que más confunde. Es por ello por lo que optamos en usar solamente tres palabras, en vez de nueve (una para cada casilla), ya que el error y confusión aumenta considerablemente.

Por otra parte, si el usuario no indica el número en el tiempo establecido, 1 segundo, la señal tendrá un valor máximo muy pequeño. Por tanto, para evitar errores de ejecución y de confusión de palabra, si dicho valor no supera un umbral establecido a priori se repite el proceso hasta que el jugador diga la palabra que crea necesaria.