

Enunciado de la segunda práctica

Fecha entrega: Domingo 3 de diciembre, 23:55 a través de AG

Tarea: Control de reproducción de video mediante gestos con hand-pose-detection

En esta práctica, usarás los landmarks (puntos clave) detectados por el modelo de hand-pose-detection para construir una aplicación que controle la reproducción de un video a través de gestos. Implementarás tres gestos del estado del arte que permiten:


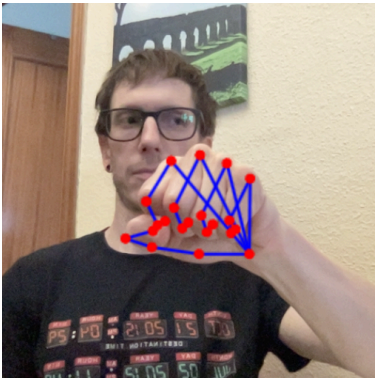
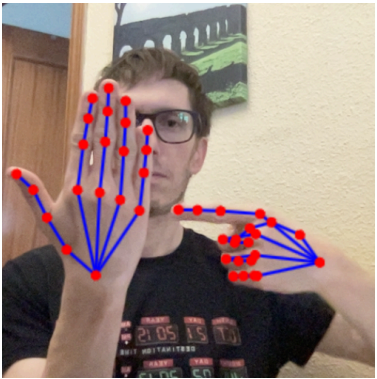
1. Mover el video adelante o atrás.
2. Cerrar el video.
3. Subir o bajar el volumen.

Los gestos de esta práctica están fundamentados en estudios de educación centrados en las personas [1,2]. Estos estudios se realizan para identificar gestos naturales e intuitivos que las personas utilizan de manera espontánea para interactuar con sistemas digitales.

Además se requiere la implementación de un 4 gesto a elección del estudiante.

Descripción de los Gestos

A continuación, se describen los gestos que implementarás junto con imágenes de ejemplo y sugerencias para su implementación. Atención, las imágenes son espejos y los gestos con una mano se realizan con la mano derecha.

1. Mover el video adelante o atrás	2. Cerrar el video	3. Subir o bajar el volumen
		

1. Mover el video adelante o atrás

La palma de la mano derecha está extendida, con los dedos abiertos y apuntando hacia arriba. Mueve la mano hacia la derecha en el eje X para mover adelante el video. Mueve la mano hacia la izquierda en el eje X para mover atrás

Implementación sugerida para detectar la palma hacia la izquierda: el reconocimiento de la orientación de la palma (hacia la izquierda) se puede lograr calculando el vector normal de la palma a partir de los landmarks3D proporcionados por el modelo Hand-Pose-Detection. Este enfoque utiliza las posiciones de los landmarks para definir el plano de la palma y calcular su orientación.

Atención:

Los landmarks 3D que forman el plano de la palma son:

- Wrist (keypoint 0): Base de la palma.
- Index MCP (keypoint 5): Base del dedo índice.
- Pinky MCP (keypoint 17): Base del dedo meñique.

Normalizar el vector asegurando que tiene una longitud de 1 (dividir cada componente por la magnitud del vector).

Si la componente x está cerca del valor 1, la palma está orientada hacia la izquierda ya que el vector normal apunta hacia la derecha desde la perspectiva del usuario).

Agrega umbrales para robustez.

Recuerda que los valores de landmarks3D son en metros.

Como puedo detectar que la palma está abierta con los dedos apuntando hacia arriba?

Implementación sugerida para detectar el movimiento: usa el landmark de la muñeca (wrist, keypoint 0) para detectar el movimiento horizontal. Decide que valores devueltos por el modelo son los mas adecuados en este caso: landmarks o landmarks3D

2. Cerrar el video

Forma un puño con la mano derecha (todos los dedos cerrados) para cerrar la reproducción (detener y ocultar el reproductor).

Implementación sugerida: Verifica que todos los landmarks de las puntas de los dedos (keypoints 4, 8, 12, 16, 20) estén muy cerca de sus respectivos landmarks en las bases de los dedos (keypoints 3, 6, 10, 14, 18). Si la distancia promedio entre las puntas de los dedos y sus bases ($\frac{\sum d_{ij}}{n}$) es menor que un umbral definido, identifica el gesto como un puño.

3. Subir o bajar el volumen

La mano derecha con el índice extendido hacia la izquierda, los demás dedos doblados. La mano izquierda abierta con la palma hacia la pantalla. Subir volumen: Mueve la mano derecha hacia arriba. Mueve la mano derecha hacia abajo

Pistas Técnicas

- Usa buffers de fotogramas (por ejemplo, los últimos 5) para suavizar los movimientos detectados y evitar detecciones erráticas.
- Usa la etiqueta `<video>` y el API [HTMLMediaElement](#) para la reproducción del video
 - Usa `video.currentTime += delta` para avanzar o retroceder.
 - Usa `video.pause()` para detener la reproducción
 - Ajusta `video.volume` dentro del rango `[0, 1]`.

Plantilla para la práctica: <https://aulaglobal.uc3m.es/mod/resource/view.php?id=5270282>

Normas de entrega:

Se entregará en AG un fichero comprimido (.zip) con:

- Código fuente: Incluye todos los archivos necesarios para ejecutar la aplicación (HTML, CSS, y JavaScript).
- Memoria (.pdf): El documento debe contener:
 - Cómo se implementaron las detecciones y dificultades.
 - Reflexión honesta sobre qué aprendieron del proceso.
 - Cómo esta práctica conecta con otras áreas de su formación o intereses personales.
- Breve video enseñando la interacción con el sistema

Puntuación

- Cada gesto con su funcionalidad: 0.4 (total 1.6)
- Experiencia de usuario: 0.2
- Memoria: 0.2

[1] Radu-Daniel Vatavu and Ionut-Alexandru Zaiti. 2014. Leap gestures for TV: insights from an elicitation study. In Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '14). Association for Computing Machinery, New York, NY, USA, 131–138. <https://doi.org/10.1145/2602299.2602316>

[2] Xiang Li, Yuzheng Chen, and Xiaohang Tang. 2022. GesPlayer: Using Augmented Gestures to Empower Video Players. In Companion Proceedings of the 2022 Conference on Interactive Surfaces and Spaces (ISS Companion '22). Association for Computing Machinery, New York, NY, USA, 4–8. <https://doi.org/10.1145/3532104.3571456>