

# DESARROLLO DE INTERFACES

TAREA I UD2



---

ALUMNO CESUR

25/26

Alejandro Muñoz de la Sierra

PROFESOR

Manuel Gómez Lora

# INTRODUCCION

Este caso práctico se centra en un tema que, aunque inicialmente pueda parecer cosa del futuro, está cada vez más presente en nuestra vida cotidiana: las interfaces naturales de usuario. Nos referimos a formas de interactuar con un sistema sin la necesidad del clásico teclado o ratón. La idea principal fue experimentar con métodos de interacción alternativos —como la voz o el movimiento corporal— aplicados a un escenario simple y muy reconocible: controlar el volumen de una televisión.

Para organizar el trabajo de forma clara, imaginamos el sistema como un pequeño “ecosistema” doméstico digital. En este ecosistema, un dispositivo, que sería la televisión, puede recibir órdenes a través de dos interfaces distintas: un control por voz y un control por gestos captados mediante una cámara. A partir de esta idea, decidimos dividir el proyecto en tres módulos principales:

MandoTelevision: Aquí reside la lógica del dispositivo.

MandoVoz: Se encarga del reconocimiento de los comandos hablados.

DeteccionMovimientoAR: Este gestiona los gestos mediante la imagen.

Este planteamiento modular nos ayudó a comprender cómo se comunican las distintas fuentes de entrada —audio y vídeo— con el procesado central y, finalmente, con las acciones que ejecuta el sistema. Personalmente, creo que este enfoque fue crucial para mantener la claridad y el orden en el proyecto.

# OBJETIVOS DEL CASO PRÁCTICO

Al iniciar el proyecto, nos planteamos una serie de objetivos que abarcaban tanto la parte conceptual como la técnica y la pedagógica.

## Objetivos conceptuales:

- Comprender qué son las interfaces alternativas de entrada.

- Conocer las bases del reconocimiento de voz.

- Entender cómo funciona la detección de movimiento usando visión por computador.

- Iniciarnos en modelos ligeros de machine learning como **Vosk**.

- Aprender a trabajar con librerías multimedia complejas, como **OpenCV** y **JavaCV**.

## Objetivos técnicos:

- Desarrollar un proyecto Java limpio y modular.

- Integrar dependencias externas a través de Maven.

- Capturar audio en tiempo real y procesarlo correctamente.

- Obtener vídeo desde la cámara del equipo y analizarlo.

- Implementar reglas de decisión basadas en gestos y comandos hablados.

- Gestionar adecuadamente los recursos (audio, vídeo y memoria) para evitar fallos.

## Objetivos pedagógicos:

- Mejorar la capacidad de documentar de forma clara y profesional.

- Afrontar problemas reales de configuración y aprender a resolverlos con criterio.

# ENTORNO DE DESARROLLO

Para el desarrollo nos apoyamos en un entorno bastante completo:

**JDK 25**, instalado manualmente en C:\Java\jdk-25. Optamos por esta instalación local en lugar de usar JavaSE-25 para evitar problemas y mantener estabilidad.

Eclipse IDE, configurado para trabajar con Maven.

Dependencias externas, gestionadas íntegramente desde el pom.xml:

**Vosk** (para reconocimiento de voz).

**JavaCV** (captura de vídeo).

**OpenCV** (procesamiento de imagen).

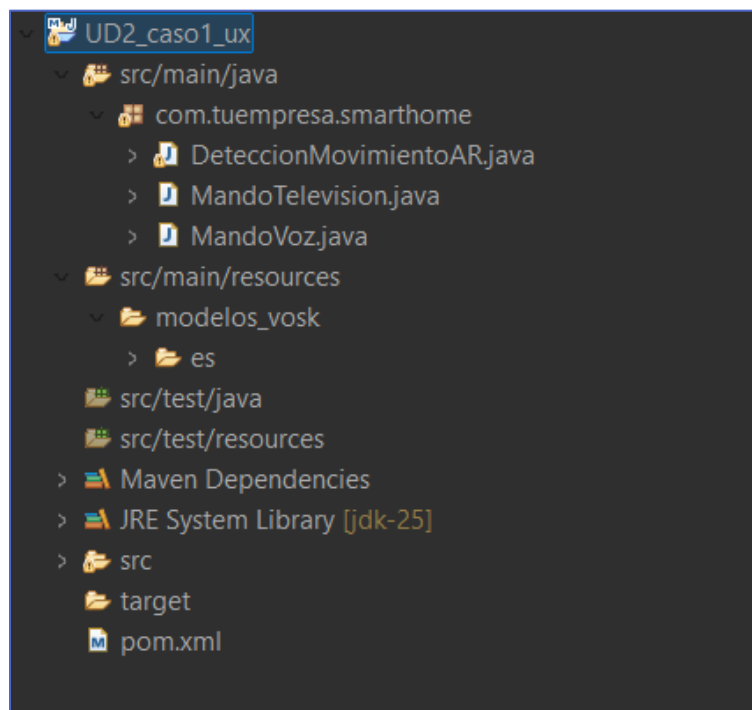
Modelos:

Modelo Vosk en español ubicado en: src/main/resources/modelos\_vosk/es

Hardware:

Micrófono integrado.

Webcam estándar (índice 0 en JavaCV).



# 03

## ARCHIVO POM.XML

```
MandoTelevisionj... MandoVoz.java DeteccionMovimie... UD2_caso1_ux/po... LibVosk.class UD2_caso1_ux/po... X
http://maven.apache.org/xsd/maven-4.0.0.xsd (xsi:schemaLocation with catalog)
1 <project xmlns="http://maven.apache.org/POM/4.0.0"
2   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3   xsi:schemaLocation="http://maven.apache.org/POM/4.0.0 http://maven.apache.org/xsd/maven-4.0.0.xsd">
4
5   <modelVersion>4.0.0</modelVersion>
6
7   <groupId>es.alejandro.dam</groupId>
8   <artifactId>UD2_caso1_ux</artifactId>
9   <version>0.0.1-SNAPSHOT</version>
10
11   <dependencies>
12     <dependency>
13       <groupId>com.alphacephi</groupId>
14       <artifactId>vosk</artifactId>
15       <version>0.3.45</version>
16     </dependency>
17
18     <dependency>
19       <groupId>org.bytedeco</groupId>
20       <artifactId>javacv-platform</artifactId>
21       <version>1.5.10</version>
22     </dependency>
23
24     <dependency>
25       <groupId>org.bytedeco</groupId>
26       <artifactId>opencv-platform</artifactId>
27       <version>4.9.0-1.5.10</version>
28     </dependency>
29   </dependencies>
30
31 </project>
32
```

```
http://maven.apache.org/xsd/maven-4.0.0.xsd (xsi:schemaLocation with catalog)
1 <project xmlns="http://maven.apache.org/POM/4.0.0"
2   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3   xsi:schemaLocation="http://maven.apache.org/POM/4.0.0 http://maven.apache.org/xsd/maven-4.0.0.xsd">
4
5   <modelVersion>4.0.0</modelVersion>
```

Declaramos, pues, que este archivo se adhiere al modelo de proyecto Maven 4.0.0.

Los atributos **xmlns** y **xsi:schemaLocation**, a mi parecer, facilitan que herramientas como Eclipse validen el XML.

Así, confirmamos que nuestro pom.xml, efectivamente, utiliza la **versión 4.0.0** del modelo de Maven.

```
7     <groupId>es.alejandro.dam</groupId>
8     <artifactId>UD2_caso1_ux</artifactId>
9     <version>0.0.1-SNAPSHOT</version>
```

**groupId:** este atributo identifica el grupo o paquete base de nuestro proyecto. Por lo general.

**artifactId:** el nombre único de nuestro proyecto, sin más.

**version:** indica, como bien sabemos, que se trata de un snapshot, o sea, una versión en desarrollo.

```
11     <dependencies>
12         <dependency>
13             <groupId>com.alphacepei</groupId>
14             <artifactId>vosk</artifactId>
15             <version>0.3.45</version>
16         </dependency>
17
```

Esta dependencia, crucial, añade **Vosk**, que nos permitirá reconocer comandos de voz en español, sin duda.

**Maven** descargará automáticamente esta librería desde los repositorios remotos.

```
18     <dependency>
19         <groupId>org.bytedeco</groupId>
20         <artifactId>javacv-platform</artifactId>
21         <version>1.5.10</version>
22     </dependency>
```

**JavaCV** es una librería para procesamiento de video y captura de cámara; una herramienta invaluable.

Incluyendo la versión platform, nos aseguramos de tener todos los binarios necesarios para Windows, Linux y macOS.

```
24     <dependency>
25         <groupId>org.bytedeco</groupId>
26         <artifactId>opencv-platform</artifactId>
27         <version>4.9.0-1.5.10</version>
28     </dependency>
29 </dependencies>
30
31 </project>
```

**OpenCV**, por su parte, proporciona funciones avanzadas de visión por computador. Importantísimo.

Lo usamos, concretamente, para detectar el movimiento de brazos y analizar fotogramas de la cámara, una tarea esencial.

Cierre del bloque de **proyecto Maven**. Fin.

Maven, ahora, conoce todas las dependencias, y Eclipse puede descargarlas y añadirlas al classpath automáticamente.

## CLASE MANDOTELEVISION

La base de todo es el objeto, **MandoTelevision**. como el director de orquesta, recibiendo órdenes de dos fuentes distintas.

Es nuestra simulación de un mando real. Guarda celosamente el volumen actual y, se asegura de que nunca caiga por debajo de cero. Además, nos dice qué módulos externos están en acción, mostrando mensajes informativos.

```
MandoTelevision.java x MandoVoz.java DetecccionMovimientoAR.java UD2_caso1_ux/pom.xml
1 package com.tuempresa.smarthome;
2
3 public class MandoTelevision {
4     private int volumen;
5
6     // Constructor
7     public MandoTelevision() {
8         this.volumen = 0;
9     }
10
11     // Método para subir el volumen
12     public void subirVolumen() {
13         this.volumen++;
14         System.out.println("Volumen actual: " + this.volumen);
15     }
16
17     // Método para bajar el volumen
18     public void bajarVolumen() {
19         if (this.volumen > 0) {
20             this.volumen--;
21             System.out.println("Volumen actual: " + this.volumen);
22         } else {
23             System.out.println("El volumen ya está en 0 y no se puede bajar más.");
24         }
25     }
26
27     // Método para obtener el volumen actual
28     public int obtenerVolumen() {
29         return this.volumen;
30     }
31 }
32
```

```
1 package com.tuempresa.smarthome;  
2
```

El **package**, como tal, agrupa las clases en lo que podríamos considerar una "carpeta virtual," denominada **com.tuempresa.smarthome**. Esta organización, a mi parecer, es crucial para mantener el proyecto ordenado y evitar posibles choques de nombres con otras librerías.

```
3 public class MandoTelevision {
```

Ahora bien, declaramos la clase pública **MandoTelevision**. Que sea pública, creo yo, es esencial, ya que otras clases, como **MandoVoz** o **DeteccionMovimientoAR**, deben tener la posibilidad de usarla.

```
3 public class MandoTelevision {  
4     private int volumen;  
5
```

Por otro lado, definimos un atributo volumen, que es privado. En mi opinión, hacerlo privado es una buena práctica, ya que así nos aseguramos de que solo se pueda modificar a través de los métodos de la clase, garantizando así la consistencia.

```
6 // Constructor  
7 public MandoTelevision() {  
8     this.volumen = 0;  
9 }
```

El constructor, por su parte, inicializa el volumen a 0 al momento de crear un objeto **MandoTelevision**. Me parece una buena manera de empezar con un valor conocido.

```

11 // Método para subir el volumen
12 public void subirVolumen() {
13     this.volumen++;
14     System.out.println("Volumen actual: " + this.volumen);
15 }

```

Tenemos, además, un método para subir el volumen. Este método incrementa el volumen en 1 y, además, muestra el valor por consola.

También hay un método para bajar el volumen. Este método, a diferencia del anterior, comprueba que el volumen no sea menor que 0 antes de decrementar, y proporciona feedback si el volumen ya está en 0.

```

17 // Método para bajar el volumen
18 public void bajarVolumen() {
19     if (this.volumen > 0) {
20         this.volumen--;
21         System.out.println("Volumen actual: " + this.volumen);
22     } else {
23         System.out.println("El volumen ya está en 0 y no se puede bajar más.");
24     }
25 }

```

Finalmente, encontramos un método para consultar el volumen actual. En mi opinión, este método es bastante útil para otras clases que necesiten conocer el volumen sin modificarlo.

```

27 // Método para obtener el volumen actual
28 public int obtenerVolumen() {
29     return this.volumen;
30 }
31 }

```

## CLASE MANDVOZ

Este objeto escucha a 16 kHz, digiere bloques de 4096 bytes y usa el modelo de **Vosk** para descifrar lo que le decimos. Luego, traduce esos comandos en acciones concretas sobre el televisor. Un intérprete de primera.

```

MandoTelevision.java  MandoVoz.java x  DeteccionMovimientoAR.java  UD2_caso1_ux/pom.xml  LibVosk.class
1  package com.tuempresa.smarthome;
2
3  import javax.sound.sampled.AudioFormat;
9
10 public class MandoVoz {
11
12     public static void main(String[] args) throws Exception {
13         System.out.println("Comenzando a procesar audio...");
14         MandoTelevision mando = new MandoTelevision();
15
16         // LibVosk.setLogLevel(0);
17         Model model = new Model("src/main/resources/modelos_vosk/es"); // modelo español
18         Recognizer recognizer = new Recognizer(model, 16000);
19
20         AudioFormat format = new AudioFormat(16000, 16, 1, true, false);
21         TargetDataLine microphone = AudioSystem.getTargetDataLine(format);
22         microphone.open(format);
23         microphone.start();
24
25         byte[] buffer = new byte[4096];
26         while (true) {
27             int bytesRead = microphone.read(buffer, 0, buffer.length);
28             if (bytesRead < 0)
29                 break;
30
31             if (recognizer.acceptWaveForm(buffer, bytesRead)) {
32                 String result = recognizer.getResult();
33                 if (result.contains("subir volumen")) {
34                     mando.subirVolumen();
35                 } else if (result.contains("bajar volumen")) {
36                     mando.bajarVolumen();
37                 }
38             }
39         }
40         microphone.close();
41         recognizer.close();
42         model.close();
43     }
44 }
45

```

Al pertenecer al mismo **paquete**, podemos acceder fácilmente a la clase **MandoTelevision** sin tener que importarla.

```
1 package com.tuempresa.smarthome;
2
3 import javax.sound.sampled.AudioFormat;
4 import javax.sound.sampled.AudioSystem;
5 import javax.sound.sampled.TargetDataLine;
6
7 import org.vosk.Model;
8 import org.vosk.Recognizer;
```

Para empezar, necesitamos importar ciertas clases. Estas clases nos permiten tanto capturar audio —con **AudioSystem** y **TargetDataLine**—, como entender lo que decimos, o sea, reconocer comandos de voz usando **Model** y **Recognizer**.

```
10 public class MandoVoz {
11
12     public static void main(String[] args) throws Exception {
13         System.out.println("Comenzando a procesar audio...");
14         MandoTelevision mando = new MandoTelevision();
```

Luego, creamos una instancia de **MandoTelevision** y, para que el usuario lo sepa, mostramos un mensajito de inicio.

```
16         // LibVosk.setLogLevel(0);
17         Model model = new Model("src/main/resources/modelos_vosk/es"); // modelo español
18         Recognizer recognizer = new Recognizer(model, 16000);
```

Ahora, algo crucial: cargamos el modelo **Vosk** en **español**. También, ajustamos el reconocedor a una frecuencia de muestreo de 16kHz, algo importante para la claridad.

```
20         AudioFormat format = new AudioFormat(16000, 16, 1, true, false);
21         TargetDataLine microphone = AudioSystem.getTargetDataLine(format);
22         microphone.open(format);
23         microphone.start();
```

Por supuesto, debemos configurar el micrófono para escuchar el audio en vivo, al momento.

```

25     byte[] buffer = new byte[4096];
26     while (true) {
27         int bytesRead = microphone.read(buffer, 0, buffer.length);
28         if (bytesRead < 0)
29             break;
30
31         if (recognizer.acceptWaveForm(buffer, bytesRead)) {
32             String result = recognizer.getResult();
33             if (result.contains("subir volumen")) {
34                 mando.subirVolumen();
35             } else if (result.contains("bajar volumen")) {
36                 mando.bajarVolumen();
37             }
38         }

```

En un ciclo continuo, como una escucha constante, leemos el audio y lo pasamos al reconocedor. Si Vosk "entiende" un comando, pues ejecutamos la acción que corresponda.

```

40         microphone.close();
41         recognizer.close();
42         model.close();
43     }

```

Finalmente, para evitar problemas y liberar memoria, cerramos con cuidado todos los recursos usados.

## CLASE

## DETECCION MOVIMIENTO AR

Aquí es donde la cámara entra en juego, conectándose con un simple **VideoCapture(0)**. Transformamos las imágenes a escala de grises y buscamos movimiento comparando fotogramas consecutivos. Analizamos los contornos más llamativos para reconocer gestos y aplicamos un filtro de repetición para evitar esos falsos positivos que tanto molestan.

La regla a implementar: mover la mano a la izquierda sube el volumen; a la derecha, lo baja.

```

1 package com.tuempresa.smarthome;
2
3 import org.bytedeco.javacv.CanvasFrame;
12
13 // @SuppressWarnings("resource")
14 public class DeteccionMovimientoAR {
15
16     public static void main(String[] args) throws FrameGrabber.Exception {
17         MandoTelevision mando = new MandoTelevision();
18         OpenCVFrameConverter.ToMat convertidor = new OpenCVFrameConverter.ToMat();
19         VideoCapture camara = new VideoCapture(0);
20
21         if (!camara.isOpened()) {
22             System.out.println("No se puede abrir la cámara");
23             return;
24         }
25
26         Mat frame = new Mat(), anterior = new Mat(), diferencia = new Mat();
27         camara.read(anterior);
28         opencv_imgproc.cvtColor(anterior, anterior, opencv_imgproc.COLOR_BGR2GRAY);
29
30         CanvasFrame lienzo = new CanvasFrame("Detección de Movimiento AR");
31         lienzo.setDefaultCloseOperation(javax.swing.JFrame.EXIT_ON_CLOSE);
32
33         // Contadores para filtrado temporal
34         int framesBrazoIzq = 0;
35         int framesBrazoDer = 0;
36         int minFramesParaAccion = 3;

```

```

37
38     try {
39         while (lienzo.isVisible() && camara.read(frame)) {
40             opencv_imgproc.cvtColor(frame, frame, opencv_imgproc.COLOR_BGR2GRAY);
41             opencv_core.absdiff(anterior, frame, diferencia);
42             opencv_imgproc.threshold(diferencia, diferencia, 25, 255, opencv_imgproc.THRESH_BINARY);
43
44             MatVector contornos = new MatVector();
45             opencv_imgproc.findContours(diferencia, contornos, new Mat(), opencv_imgproc.RETR_EXTERNAL,
46                 opencv_imgproc.CHAIN_APPROX_SIMPLE);
47
48             boolean brazoIzq = false, brazoDer = false;
49
50             int ancho = frame.cols();
51             int alto = frame.rows();
52             int limiteY = alto / 2; // Solo mitad superior (brazo levantado)
53
54             for (int i = 0; i < contornos.size(); i++) {
55                 Rect rect = opencv_imgproc.boundingRect(contornos.get(i));
56                 int centroX = rect.x() + rect.width() / 2;
57                 int centroY = rect.y() + rect.height() / 2;
58
59                 // Solo contornos grandes en la mitad superior
60                 if (centroY < limiteY && rect.height() > 50 && rect.width() > 20) {
61                     if (centroX < ancho / 2)
62                         brazoIzq = true; // brazo izquierdo levantado
63                     else
64                         brazoDer = true; // brazo derecho levantado
65                 }
66             }
67
68             // Filtrado temporal
69             if (brazoIzq)
70                 framesBrazoIzq++;
71             else
72                 framesBrazoIzq = 0;
73             if (brazoDer)
74                 framesBrazoDer++;
75             else
76                 framesBrazoDer = 0;
77
78             if (framesBrazoIzq >= minFramesParaAccion) {
79                 mando.subirVolumen();
80                 framesBrazoIzq = 0;
81             }
82
83             if (framesBrazoDer >= minFramesParaAccion) {
84                 mando.bajarVolumen();
85                 framesBrazoDer = 0;
86             }
87
88             lienzo.showImage(convertidor.convert(frame));
89             anterior = frame.clone();
90         }
91     } finally {
92         camara.release();
93         lienzo.dispose();
94     }
95 }
96 }

```

```

MandoTelevision.java  MandoVoz.java  DeteccionMovimientoAR.java
1 package com.tuempresa.smarthome;
2
3 import org.bytedeco.javacv.CanvasFrame;
4 import org.bytedeco.javacv.FrameGrabber;
5 import org.bytedeco.javacv.OpenCVFrameConverter;
6 import org.bytedeco.opencv.global.opencv_core;
7 import org.bytedeco.opencv.global.opencv_imgproc;
8 import org.bytedeco.opencv.opencv_core.Mat;
9 import org.bytedeco.opencv.opencv_core.MatVector;
10 import org.bytedeco.opencv.opencv_core.Rect;
11 import org.bytedeco.opencv.opencv_videoio.VideoCapture;
12

```

Vale, pues, arrancamos importando las librerías que nos permiten pillar el video (con **VideoCapture**), manipular cada fotograma (**Mat**, **Rect**) y pintar los resultados en un **CanvasFrame**.

```

13 // @SuppressWarnings("resource")
14 public class DeteccionMovimientoAR {
15
16     public static void main(String[] args) throws FrameGrabber.Exception {
17         MandoTelevision mando = new MandoTelevision();
18         OpenCVFrameConverter.ToMat convertidor = new OpenCVFrameConverter.ToMat();
19         VideoCapture camara = new VideoCapture(0);

```

Luego, preparamos el terreno: creamos el "mando", el conversor de **frames** y le decimos que use la cámara que viene por defecto (la 0).

```

21         if (!camara.isOpened()) {
22             System.out.println("No se puede abrir la cámara");
23             return;
24         }

```

Aquí: comprobamos que la cámara responda bien; si no, no podemos avanzar. Una vez seguros, leemos el primer **frame** y lo pasamos a **grises**. Para que el proceso de detectar movimiento sea más fácil, evitamos problemas con colores.

```

26         Mat frame = new Mat(), anterior = new Mat(), diferencia = new Mat();
27         camara.read(anterior);
28         opencv_imgproc.cvtColor(anterior, anterior, opencv_imgproc.COLOR_BGR2GRAY);
29

```

Abrimos una ventana para ir viendo cómo quedan los frames según los vamos trabajando.

```
30 CanvasFrame lienzo = new CanvasFrame("Detección de Movimiento AR");
31 lienzo.setDefaultCloseOperation(javax.swing.JFrame.EXIT_ON_CLOSE);
```

También, definimos unas variables para un "filtro temporal". Esto es importante, porque no queremos que un simple movimiento rápido active la acción, sino que el brazo tenga que estar arriba un rato.

```
33 // Contadores para filtrado temporal
34 int framesBrazoIzq = 0;
35 int framesBrazoDer = 0;
36 int minFramesParaAccion = 3;
```

Ahora sí, procesamos cada frame: lo convertimos a gris, calculamos la diferencia con el frame anterior y aplicamos un umbral para que las zonas donde hay movimiento se vean más claras.

```
38 try {
39     while (lienzo.isVisible() && camara.read(frame)) {
40         opencv_imgproc.cvtColor(frame, frame, opencv_imgproc.COLOR_BGR2GRAY);
41         opencv_core.absdiff(anterior, frame, diferencia);
42         opencv_imgproc.threshold(diferencia, diferencia, 255, opencv_imgproc.THRESH_BINARY);
43     }
```

Con esto, detectamos los bordes de los objetos que se mueven; estos contornos nos van a chivar si alguien ha levantado un brazo.

```
44 MatVector contornos = new MatVector();
45 opencv_imgproc.findContours(diferencia, contornos, new Mat(), opencv_imgproc.RETR_EXTERNAL,
46                             opencv_imgproc.CHAIN_APPROX_SIMPLE);
```

Necesitamos unas variables que nos digan si hemos visto el brazo izquierdo o el derecho.

```
48 boolean brazoIzq = false, brazoDer = false;
```

Analizamos cada contorno con lupa: si está en la parte de arriba de la imagen y es grande, lo consideramos un brazo levantado.

```

54         for (int i = 0; i < contornos.size(); i++) {
55             Rect rect = opencv_imgproc.boundingRect(contornos.get(i));
56             int centroX = rect.x() + rect.width() / 2;
57             int centroY = rect.y() + rect.height() / 2;
58
59             // Solo contornos grandes en la mitad superior
60             if (centroY < limiteY && rect.height() > 50 && rect.width() > 20) {
61                 if (centroX < ancho / 2)
62                     brazoIzq = true; // brazo izquierdo levantado
63                 else
64                     brazoDer = true; // brazo derecho levantado
65             }

```

Y, por último, usamos contadores para los frames consecutivos. Solo si el brazo está arriba durante suficientes frames, ejecutamos la acción, evitando que cualquier espasmo active algo por error.

```

68         // Filtrado temporal
69         if (brazoIzq)
70             framesBrazoIzq++;
71         else
72             framesBrazoIzq = 0;
73         if (brazoDer)
74             framesBrazoDer++;
75         else
76             framesBrazoDer = 0;
77
78         if (framesBrazoIzq >= minFramesParaAccion) {
79             mando.subirVolumen();
80             framesBrazoIzq = 0;
81         }
82
83         if (framesBrazoDer >= minFramesParaAccion) {
84             mando.bajarVolumen();
85             framesBrazoDer = 0;
86         }

```

Vamos mostrando el frame procesado y, para no perder información, lo clonamos para usarlo en la siguiente vuelta.

```

88         lienzo.showImage(convertidor.convert(frame));
89         anterior = frame.clone();

```

Para terminar, toca ser ordenados: liberamos todos los recursos, cerramos la cámara y la ventana. Y listo, personalmente, me parece una buena manera de detectar gestos simples.

```

91         } finally {
92             camara.release();
93             lienzo.dispose();
94         }

```

# PROBLEMAS ENCUENTRADOS

Nos topamos con unos cuantos problemas reales:

Problema 0: No descargaba desde el archivo pom.xml

Nos dimos cuenta de que el problema que Maven simplemente **no conseguía acceder al esquema a través de HTTPS**, probablemente por una de estas razones típicas:

- problemas de certificados SSL,
- restricciones de red,
- incompatibilidades momentáneas con el servidor,
- o incluso una caída temporal en el endpoint HTTPS de Maven.

Solución: Lo que hicimos fue modificar la línea del esquema para utilizar **HTTP** en lugar de HTTPS.

```
xsi:schemaLocation="http://maven.apache.org/POM/4.0.0  
http://maven.apache.org/xsd/maven-4.0.0.xsd"
```

Problema 1: Librerías

Intentamos meter las librerías a mano y... Errores por todos lados, culpa de los archivos nativos.

Solución: Crearlo como proyecto Maven.

## Problema 2: Vosk

El micrófono parecía funcionar, pero solo nos devolvía silencio.

Solución: Ajustamos el formato del TargetDataLine y aumentamos el tamaño de los buffers. Funcionó.

## Problema 3: OpenCV Muy Sensible

La cámara veía cualquier mínimo movimiento.

Solución: Escala de grises + absdiff + umbral de 25 + filtro temporal de 3 fotogramas consecutivos.

## Problema 4: Cámaras HD Lentas

A resoluciones altas, el programa se arrastraba.

Solución: Gris y regiones más pequeñas.

08

## PRUEBAS Y EJECUCION

### Prueba de Reconocimiento de Voz. Ejecución de la clase MandoVoz.java

```
Problems Javadoc Declaration Console Error Log
MandoVoz [Java Application] C:\Java\jdk-25\bin\javaw.exe (24 nov 2025, 10:59:27) [pid: 13564]
Comenzando a procesar audio...
WARNING: A restricted method in java.lang.System has been called
WARNING: java.lang.System::load has been called by com.sun.jna.Native in an unnamed module (file:/C:/Users/sibko/.m2/repository/net/java/dev/jna/jna/5.7.0)
WARNING: Use --enable-native-access=ALL-UNNAMED to avoid a warning for callers in this module
WARNING: Restricted methods will be blocked in a future release unless native access is enabled

LOG (VoskAPI:ReadDataFiles():model.cc:213) Decoding params beam=11 max-active=4000 lattice-beam=4
LOG (VoskAPI:ReadDataFiles():model.cc:216) Silence phones 1:2:3:4:5:6:7:8:9:10
LOG (VoskAPI:RemoveOrphanNodes():nnet-nnet.cc:948) Removed 0 orphan nodes.
LOG (VoskAPI:RemoveOrphanComponents():nnet-nnet.cc:847) Removing 0 orphan components.
LOG (VoskAPI:ReadDataFiles():model.cc:248) Loading i-vector extractor from src/main/resources/modelos_vosk/es/ivector/final.ie
LOG (VoskAPI:ComputeDerivedVars():ivector-extractor.cc:183) Computing derived variables for iVector extractor
LOG (VoskAPI:ComputeDerivedVars():ivector-extractor.cc:204) Done.
LOG (VoskAPI:ReadDataFiles():model.cc:282) Loading HCL and G from src/main/resources/modelos_vosk/es/graph/HCLr.fst src/main/resources/modelos_vosk/es/gr
LOG (VoskAPI:ReadDataFiles():model.cc:308) Loading winfo src/main/resources/modelos_vosk/es/graph/phones/word_boundary.int
Volumen actual: 1
Volumen actual: 0
Volumen actual: 1
Volumen actual: 2
Volumen actual: 1
Volumen actual: 0
El volumen ya está en 0 y no se puede bajar más.
```

Responde lento a las ordenes de voz, pero funciona.

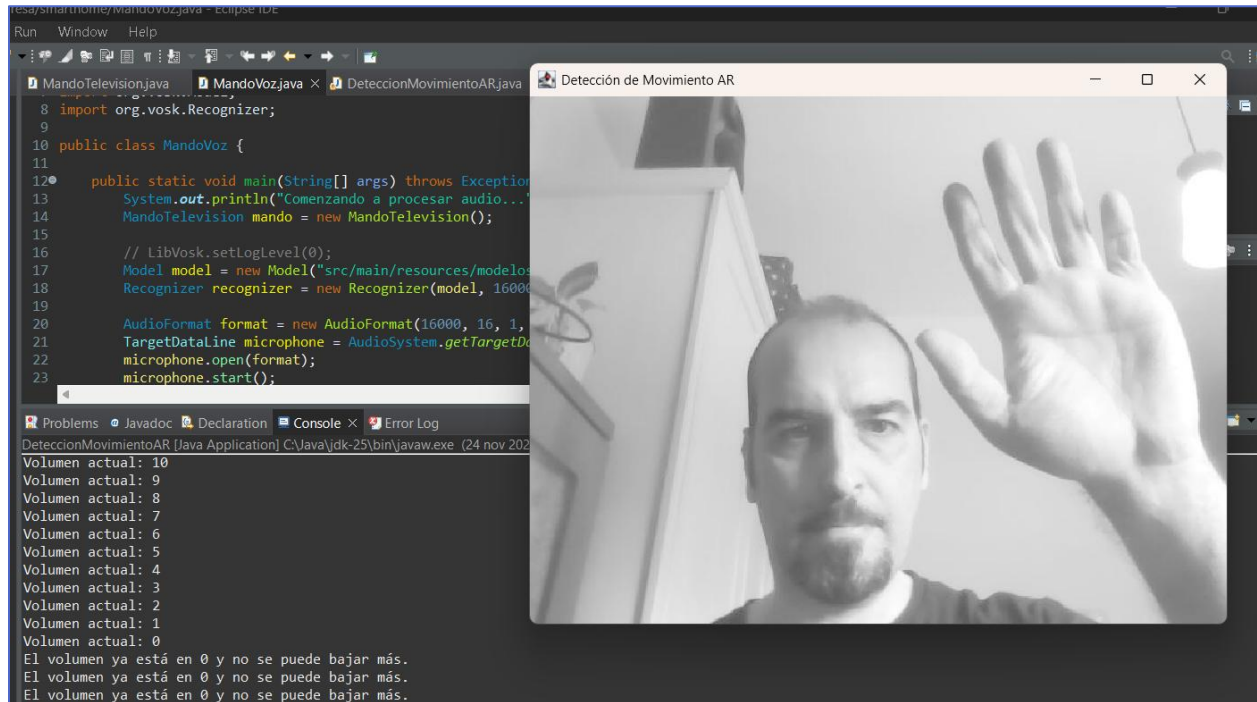
### Prueba de Detección de Gestos. Ejecución de la clase DeteccionMovimientoAR

Brazo derecho (elemento a la izquierda de la pantalla): subir volumen

```
MandoTelevision.java MandoVoz.java DeteccionMovimientoAR.java Detección de Movimiento AR
8 import org.vosk.Recognizer;
9
10 public class MandoVoz {
11
12 public static void main(String[] args) throws Exception
13     System.out.println("Comenzando a procesar audio...");
14     MandoTelevision mando = new MandoTelevision();
15
16     // LibVosk.setLogLevel(0);
17     Model model = new Model("src/main/resources/modelos_vosk/es/graph/HCLr.fst src/main/resources/modelos_vosk/es/gr
18     Recognizer recognizer = new Recognizer(model, 16000);
19
20     AudioFormat format = new AudioFormat(16000, 16, 1,
21     TargetDataLine microphone = AudioSystem.getTargetDataLine(format);
22     microphone.open(format);
23     microphone.start();
24
Problems Javadoc Declaration Console Error Log
DeteccionMovimientoAR [Java Application] C:\Java\jdk-25\bin\javaw.exe (24 nov 2025, 10:59:27) [pid: 13564]
WARNING: java.lang.System::load has been called by org.bytedeco.javacpp.Loader in an unnamed module (file:/C:/Users/sibko/.m2/repository/org/bytedeco/javacpp/1.5.10)
WARNING: Use --enable-native-access=ALL-UNNAMED to avoid a warning for callers in this module
WARNING: Restricted methods will be blocked in a future release unless native access is enabled

El volumen ya está en 0 y no se puede bajar más.
El volumen ya está en 0 y no se puede bajar más.
El volumen ya está en 0 y no se puede bajar más.
Volumen actual: 1
Volumen actual: 2
Volumen actual: 3
Volumen actual: 4
Volumen actual: 5
Volumen actual: 6
Volumen actual: 7
Volumen actual: 8
Volumen actual: 9
```

Brazo izquierdo (elemento a la derecha de la pantalla): bajar volumen



Ajustamos los tamaños de contorno mínimos para evitar que confundiera manos pequeñas, cabezas o sombras.

## CONCLUSIONES

Este proyecto nos ha permitido tocar una tecnología que solemos ver en dispositivos modernos, pero rara vez en proyectos educativos. Aprendimos a capturar audio y vídeo en tiempo real, a usar modelos de reconocimiento de voz locales, a detectar movimiento mediante diferencias de imagen y a estructurar un proyecto por módulos.

Además, Maven fue un salvavidas. Pero lo mejor de todo es la satisfacción de haber unido multimedia y machine learning en un proyecto Java funcional.

El sistema es ampliable: gestos nuevos, más comandos de voz, control de otros dispositivos... Pero lo importante es que ahora entendemos cada pieza y cómo encajan en este rompecabezas. Personalmente, creo que la curva de aprendizaje fue pronunciada, pero el resultado final bien valió el esfuerzo. La satisfacción de ver algo así funcionando es impagable.

## REFERENCIAS

<https://maven.apache.org/guides/getting-started/>

<https://maven.apache.org/pom.html>

<https://docs.oracle.com/en/java/javase/25/>

<https://help.eclipse.org/latest/index.jsp>

<https://alphacephei.com/vosk/install>

<https://alphacephei.com/vosk/>

<https://github.com/bytedeco/javacv/wiki>

<https://bytedeco.org/>