

## EJEMPLO GLOBAL DE SONIDO

### **FUNCIONALIDAD:**

Esta librería permite al robot Nao ser capaz de interaccionar con un humano, a través del reconocimiento de voz y de la propia habla del robot, para luego ser capaz de seguir con la cabeza o el cuerpo un sonido, una pelota roja o una cara.

Por tanto, el funcionamiento general de la librería se puede dividir en 2 etapas:

- Interacción humano-robot: reconocimiento y sintetización de voz por parte del robot.
- Seguimiento de un sonido, pelota o cara.

### **INTERACCIÓN HUMANO-ROBOT:**

Esta etapa se realiza con la ayuda del módulo “ALSpeechRecognition” y del módulo “ALTextToSpeech” presentes en la API NAOqi Audio del robot.

Por una parte “ALSpeechRecognition” se basa en sofisticadas tecnologías de reconocimiento de voz proporcionadas por: GRUPO ACAPELA para Nao Versión 3.x y NUANCE para Nao Versión 4.

Su principio de operación:

- Antes de comenzar, “ALSpeechRecognition” necesita ser alimentado por la lista de frases que deben ser reconocidas. De esta forma se creará el diccionario particular de Nao y sólo reconocerá las palabras o frases contenidas en él.
- Una vez iniciado, “ALSpeechRecognition” coloca en la clave ‘SpeechDetected’, un booleano que especifica si se escucha o no por un altavoz. Cuando se haya escuchado algo, el robot emitirá un pitido para dejar constancia de ello.
- Si se oye por un altavoz, el elemento de la lista que mejor se adapte a lo que se escucha por el robot se coloca en la clave ‘WordRecognized’, que se organiza de la siguiente manera:

***[Frase\_1, confianza\_1, frase\_2, confianza\_2, ..., frase\_n, confianza\_n]***

dónde:

***Frase\_i*** es una de las frases predefinidas.

***Confianza<sub>i</sub>*** es una estimación de la probabilidad de que esta frase sea realmente lo que ha sido pronunciado por el hablante humano.

Hai que aclarar que las diferentes hipótesis contenidas en 'WordRecognized' están ordenadas para que las frases más probables aparezcan primero; por tanto bastará con acceder a la primera posición de este vector para conocer la palabra o frase que ha escuchado Nao.

Por otra parte "ALTextToSpeech" el módulo "ALTextToSpeech" permite al robot hablar. Funciona enviando comandos a un motor de conversión de texto a voz proporcionado por ACAPELA, y autoriza también la personalización de voz. El resultado de la síntesis se envía a los altavoces de Nao haciendo así que este hable.

### **SEGUIMIENTO DE UN SONIDO, PELOTA O CARA:**

Para esta etapa final se cuenta con la ayuda del módulo "ALTracker" presente en la API NAOqi Trackers. El módulo "ALTracker" es un seguidor genérico que permite al robot seguir diferentes objetivos (pelota roja, cara, punto de referencia, sonido...) usando diferentes medios (cabeza, cuerpo entero, movimiento, etc). Así, el objetivo principal de este módulo es establecer un puente entre la detección de objetivos y el movimiento con el fin de hacer que el robot mantenga a la vista el objetivo en el centro de la cámara, de la forma en se muestra en la figura 1.

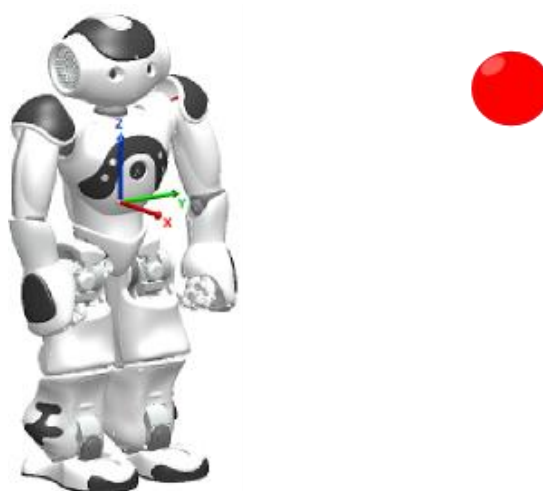


Figura 1. Posición del robot al seguir una pelota de color.

"ALTracker" puede rastrear el objetivo con cualquiera de los modos mostrados en la tabla 1 y puede a su vez seguir cualquiera de los objetivos mostrados en la tabla 2.

Modo	El robot seguirá....	Comentario
<b>Cabeza</b>	Con la cabeza solamente (modo por defecto)	Las dos juntas de la cabeza se controlan para seguir el objetivo.
<b>Cuerpo Entero</b>	Con todo el cuerpo y piernas fijas	El robot mantiene el equilibrio de forma autónoma y adapta su postura para rastrear el objetivo.
<b>Movimiento</b>	Con movimiento	El robot se mueve para mantener una distancia fija con el objeto.

Tabla 1. Modos de seguimiento

Objetivo	Parámetros	Comentario
<b>Pelota Roja</b>	Diámetro de la pelota (metros)	Se utiliza para calcular la distancia entre el robot y la pelota.
<b>Cara</b>	Anchura de la cara (metros)	Se utiliza para calcular la distancia entre el robot y la cara.
<b>Sonido</b>	[Distancia, confianza]	Distancia se usa para estimar la posición del sonido y confianza para filtrar su localización.

Tabla 2. Posibles objetivos de seguimiento

### **FUNCIONAMIENTO:**

A partir de los módulos explicados antes y si se juntan las dos etapas comentadas, se obtiene una librería que permite realizar un ejemplo final cuyo funcionamiento se resume en el diagrama de la figura 2 y que es el siguiente:

1. Se realiza la configuración de los módulos seleccionando los lenguajes para el reconocedor y el sintetizador de voz. Se añade además el vocabulario a reconocer deseado.

2. El robot interactúa con el humano para preguntarle acerca del objetivo a seguir (pelota, sonido o cara) y acerca del modo de seguimiento (sólo cabeza o movimiento con cuerpo hacia el objetivo). Nao preguntará hasta asegurarse de que ha entendido correctamente lo que tiene que hacer.
3. Con los datos obtenidos del usuario en el paso anterior, se configura el seguidor en función del objetivo a seguir y del modo de movimiento deseado.
4. El robot realizará el seguimiento según la configuración deseada hasta que se pare el programa (Ctrl + C).

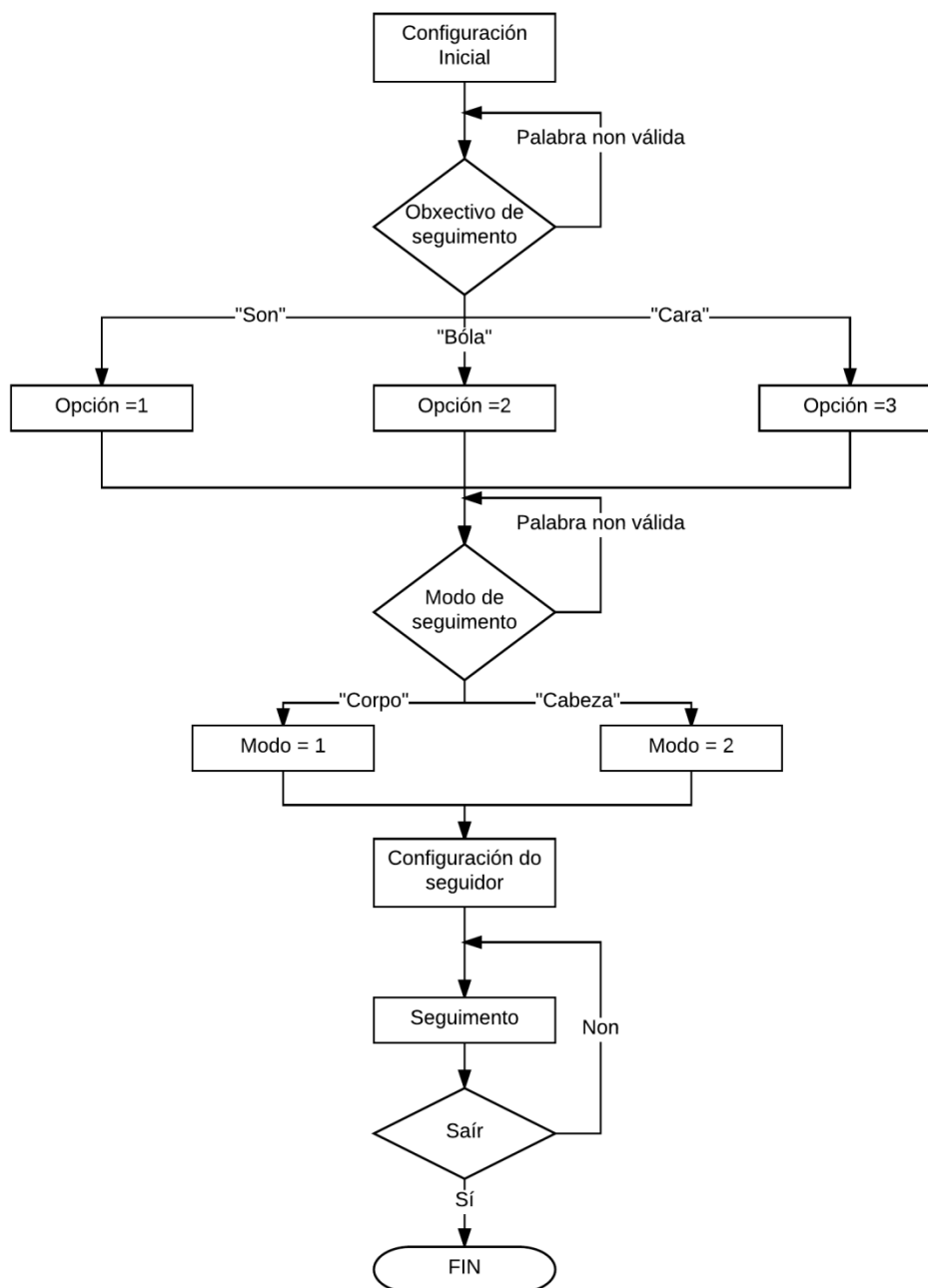


Figura 2. Diagrama con el funcionamiento de la librería.