



FACULTAD DE INFORMÁTICA
DEPARTAMENTO DE COMPUTACIÓN

TRABAJO DE FIN DE GRADO
DE INGENIERÍA INFORMÁTICA

*Desarrollo de subsistema de interacción
humano-máquina para el ROBOBO 2.0*

Autor: Llamas Luaces, Luis Felipe
Tutor: Bellas Bouza, Francisco Javier
Directores: Varela Fernández, Gervasio
Prieto García, Abraham

A Coruña, a 31 de agosto de 2016.

Información general

Título del proyecto: “Desarrollo de subsistema de interacción humano-máquina para el ROBOBO 2.0 ”

Clase de proyecto: Proyecto de desarrollo en investigación

Nombre del alumno: Llamas Luaces, Luis Felipe

Nombre del tutor: Bellas Bouza, Francisco Javier

Nombre de los directores: Varela Fernández, Gervasio
Prieto García, Abraham

Miembros del tribunal:

Miembros suplentes:

Fecha de lectura:

Calificación:

Dr. Bellas Bouza, Francisco Javier

CERTIFICA

Que la memoria titulada “**Desarrollo de subsistema de interacción humano-máquina para el ROBOBO 2.0**” ha sido realizada por Luis Felipe Llamas Luaces con D.N.I. 48113017-F bajo la dirección del Dr. Francisco Javier Bellas Bouza. La presente constituye la documentación que, con mi autorización, entrega el mencionado alumno para optar a la titulación de Ingeniería en Informática.

A Coruña, a 31 de agosto de 2016.

Firmado:

Francisco Javier Bellas Bouza

Sigue rascando, hay millones de premios.

Agradecimientos

Resumen

ABSTRACT.tex

Palabras clave:

- ✓ Robótica autónoma
- ✓ Robótica educativa
- ✓ Interacción humano robot
- ✓ Android

Índice general

	Página
1. Introducción	1
1.1. Motivación	1
1.1.1. Robobo 1.0	2
1.1.2. Robobo 2.0	3
1.2. Objetivos	6
1.3. Estructura de la memoria	6
1.4. Herramientas utilizadas	7
2. Fundamentos teóricos	9
2.1. ROBOBO! Framework	9
2.1.1. Android	13
2.1.2. ROS	14
3. Antecedentes	17
3.1. Interacción humano robot	17
3.1.1. HRI en la industria	18
3.1.2. HRI en robots asistenciales	19
3.1.3. HRI en robots de entretenimiento	21
3.1.4. HRI en robots educativos	24
4. Fundamentos tecnológicos	29
4.0.1. Sphinx	29
4.0.2. TarsosDSP	29
4.0.3. OpenCV	29
4.0.4. Gmail Background	29

5. Desarrollo	31
5.1. Modelo conceptual para el subsistema de interacción	31
5.2. Arquitectura Global	33
5.2.1. Estructura de un módulo	34
5.3. Metodología	35
5.3.1. Primera iteración	35
5.3.2. Segunda iteración	35
5.3.3. Tercera iteración	36
5.3.4. Cuarta iteración	37
5.4. Librerías de interacción	38
5.4.1. Paquete Speech	38
5.4.1.1. Módulo recognition	38
5.4.1.2. Módulo Production	41
5.4.2. Paquete Touch	44
5.4.2.1. Módulo Touch	44
5.4.3. Paquete Sound	46
5.4.3.1. Módulo Sound Dispatcher	46
5.4.3.2. Módulo Pitch Detection	48
5.4.3.3. Módulo Note Detection	50
5.4.3.4. Módulo Clap Detection	50
5.4.3.5. Módulo NoteGenerator	51
5.4.3.6. Módulo EmotionSound	53
5.4.4. Paquete Vision	55
5.4.4.1. Módulo Basic Camera	55
5.4.4.2. Módulo Face Detection	57
5.4.4.3. Módulo ColorDetector	58
5.4.5. Paquete Messaging	59
5.4.5.1. Módulo email	59
6. Resultados y pruebas	61
6.1. Ejemplos de uso	61
6.1.1. Simon dice musical	61
6.1.2. Robobo Vigilante	63
6.1.3. Robobo Mascota	63
6.2. Problemas conocidos	63

7. Conclusiones	65
7.1. Trabajo Futuro	65
A. Manual de uso	67
Bibliografía	69

Índice de figuras

Figura	Página
1.1. Robobo 1.0	3
1.2. Robobo 2.0	4
1.3. Disposición de los sensores en el ROBOBO 2.0	5
2.1. Estructura de ROBOBO! framework dentro del entorno ROBOBO	10
2.2. Clase IRob	12
2.3. Clase IRobMovementModule	12
2.4. Interfaz IModule	13
3.1. Esquema de una jaula de seguridad para un robot industrial	18
3.2. Robot terapeutico Paro	20
3.3. Esquema de características del robot NAO	21
3.4. Sistema quirúrgico daVinci	22
3.5. Robot mascota Aibo de Sony	23
3.6. Robot mascota Smartpet de Bandai	23
3.7. Beebot	24
3.8. Escornabot	24
3.9. Kit básico Lego Mindstorms EV3	25
3.10. Robot Zowi de BQ	26
3.11. Robot Thymio	27
3.12. Robots Dash y Dot	28
5.1. Estructura de paquetes dentro del paquete HRI del Robobo	33
5.2. Paquete Speech	38
5.3. Módulo SpeechRecognition	39
5.4. Módulo SpeechProduction	42

5.5.	Módulo Touch	44
5.6.	Paquete Sound	46
5.7.	Módulo SoundDispatcher	47
5.8.	Módulo PitchDetection	48
5.9.	Módulo NoteDetection	49
5.10.	Módulo ClapDetection	51
5.11.	Módulo NoteGenerator	52
5.12.	Módulo EmotionSound	53
5.13.	Paquete Vision	55
5.14.	Módulo BasicCamera	56
5.15.	Módulo FaceDetection	57
5.16.	Módulo ColorDetection	58
5.17.	Módulo Email	60
6.1.	Escala pentatonica menor de La	63

Capítulo 1

Introducción

1.1. Motivación

El Grupo Integrado de Ingeniería (GII) de la Universidade da Coruña (UDC) viene desarrollando una línea de investigación en Robótica y Cognición desde hace más de 15 años, que tiene como objetivo general desarrollar sistemas reales que puedan responder con el mayor grado de autonomía posible a las condiciones cambiantes del medio, sin tener que precisar de la ayuda de un operador o programador para establecer las nuevas estrategias necesarias para la consecución de la tarea encomendada. Esta línea de investigación tiene otra línea derivada que se denomina Mecanismos Cognitivos, en la cual se centra en el estudio e implementación de funcionalidades cognitivas de alto nivel en robots autónomos, utilizando los procesos cognitivos humanos como inspiración para desarrollar un mecanismo cognitivo completo para un robot. Uno de los proyectos asociados a la línea de Mecanismos Cognitivos es DREAM [1] (Deferred Restructuring of Experience in Autonomous Machines), financiado por la Unión Europea, y que se centra en la integración de procesos cognitivos basados en el sueño dentro de una arquitectura cognitiva para robots autónomos. El objetivo de utilizar estos procesos reside en la capacidad de consolidar el aprendizaje adquirido durante la operación en "tiempo de vida" que se ha visto que existe en los procesos asociados al sueño en el cerebro humano. De esta forma, la experiencia adquirida por el robot se analiza y procesa en otra escala temporal que proporciona nuevas representaciones de más alto nivel. Uno de los principales problemas a solucionar en el proyecto DREAM se centra en el aprendizaje a partir de la interacción con humanos. En este sentido, se requieren numerosos datos de interacción en

tiempo real en diversas situaciones para poder mejorar la arquitectura, y para lograrlos se ha diseñado la iniciativa «adopt a robot». Esta iniciativa se centra en proporcionar a los centros educativos un robot de bajo coste que contenga la arquitectura cognitiva desarrollada en el DREAM, de modo que los alumnos puedan interactuar con el robot libremente y se consiga gran variedad de datos para mejorar dicha arquitectura. Este robot está en desarrollo actualmente en el GII bajo el nombre de ROBOBO, y es en el que se centrará el presente TFG. El ROBOBO se basa en el uso de una plataforma motorizada y sensorizada, llamada ROB en el sistema, que se conecta a un Smartphone, llamado OBO, que se encarga del procesado de los datos recogidos desde la plataforma y de controlar las acciones de la misma. Este tipo de arquitectura presenta numerosas ventajas que resultaron interesantes a la hora de escogerlo para su uso en DREAM:

- Bajo coste de la plataforma
- Gran capacidad de sensorización, cualquier smartphone de gama media cuenta con acelerómetros, giroscopios, gps, cámara y micrófono, y de comunicaciones, wifi, gsm y bluetooth.
- Robot fácilmente actualizable, mediante el cambio de smartphone, sin necesidad de adaptar el software antiguo.
- Gran disponibilidad de smartphones, todos los usuarios potenciales poseen uno, además Android, sistema operativo escogido para el sistema, cuenta con un 82.8 % de la cuota de mercado.

1.1.1. Robobo 1.0

La primera versión de la plataforma desarrollada por el GII fue el ROBOBO 1.0 (Figura 1.1), que sirvió como versión conceptual para comprobar la viabilidad del sistema plataforma + smartphone, y presentaba las siguientes características:

- 9 LEDs (diodos emisores de luz) RGB para interactuar con el usuario, que cambiaban de color con la proximidad de objetos.
- 9 sensores IR de proximidad para proporcionar capacidad de movimiento autónomo: 2 en la parte frontal, 4 en los laterales y los últimos tres en la parte trasera; de manera que proporcionaba una visión general del entorno.

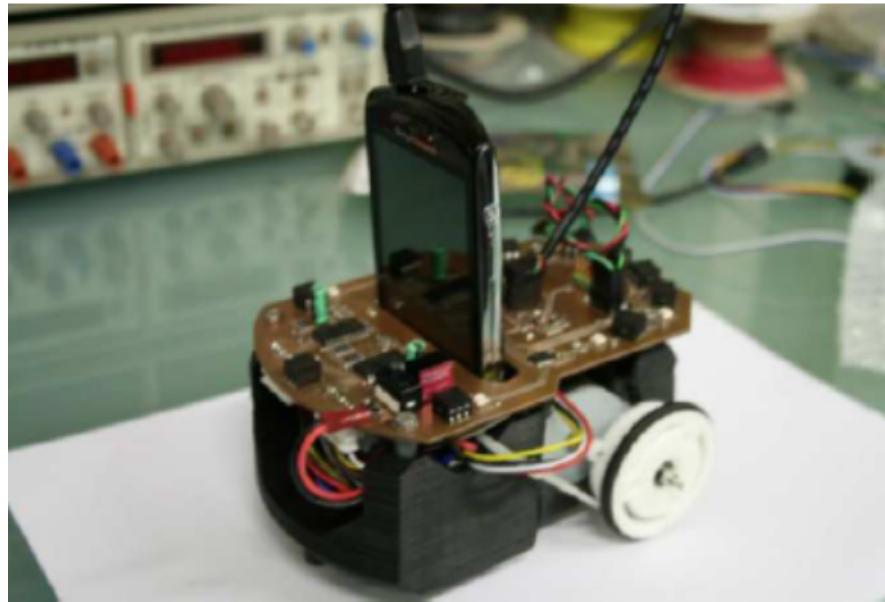


Figura 1.1: Robobo 1.0

- 2 motores paso a paso (convierten impulsos eléctricos en desplazamientos angulares discretos, es decir, pueden avanzar un ángulo concreto en función de la señal recibida) que aplicaban movimiento a las ruedas, con un paso de 1/8, con gran precisión de giro.
- Compatibilidad con dispositivos Android.
- Capacidad de interacción con otros dispositivos ROS.
- Conexión USB para la comunicación con el teléfono inteligente.

1.1.2. Robobo 2.0

La segunda versión, el ROBOBO 2.0 (figura 1.2), trata de solventar las carencias de la primera versión y añadir ciertas mejoras. Tras este rediseño, las características del robot pasaron a ser:

- LEDs: pasan a ser 21, 9 de ellos formando una matriz de comunicación para dar una información avanzada mediante colores y formas.
 - Motores de las ruedas: Pasan a ser motores de corriente continua, menos voluminosos y más eficientes energéticamente. Cada motor cuenta con un en-
-



Figura 1.2: Robobo 2.0

coder magnético, que transforma el movimiento del motor en pulsos eléctricos, que pueden ser interpretados por la parte software del robot.

- Comunicación: Se elimina la comunicación USB a favor de una conexión Bluetooth, lo cual permite utilizar diferentes smartphones con distintas posiciones de los puertos.
- Plataforma universal para smartphones: Se incluye un adaptador universal para teléfonos móviles, montado sobre una plataforma motorizada que permite el movimiento del smartphone sobre el chasis del robot. Esta plataforma tiene dos grados de libertad, rotación e inclinación.
- Sensores: Modificadas tanto el tipo, la cantidad y la posición. Se colocaron en la siguiente disposición (figura 1.3):
 - 3 sensores para la detección de colisiones en la parte delantera; uno frontal y los otros dos cercanos a las ruedas, de tal forma que el ROBOBO no pudiera chocar, frontalmente con ninguna superficie.
 - 2 sensores inclinados hacia abajo para la detección de caídas, en la parte delantera, colocados cerca de las ruedas, es decir, la zona con

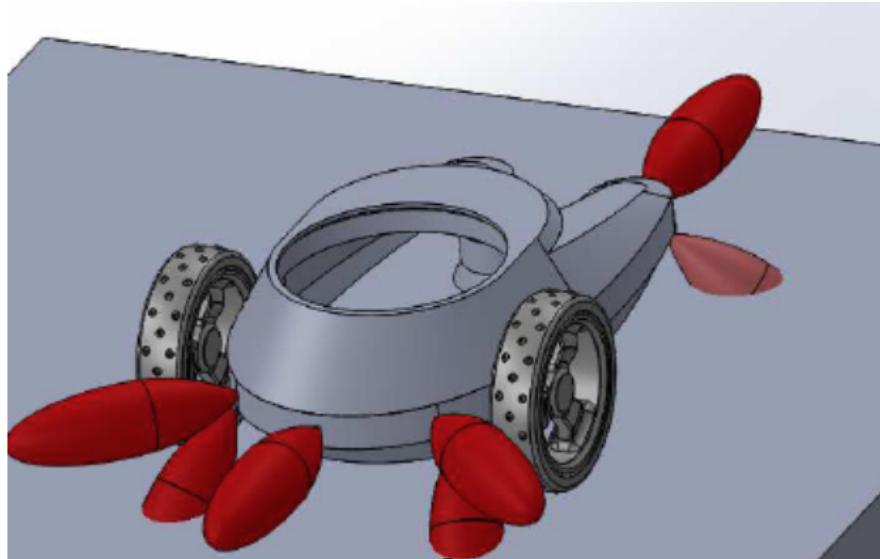


Figura 1.3: Disposición de los sensores en el ROBOBO 2.0

mayor peligro de caídas.

- 2 sensores de caídas y 2 de colisiones, de forma similar a los anteriores, en la zona anterior del robot.
- Unión sensores-microcontrolador: se han cambiado las resistencias pull-up anteriores por un multiplexor. Se consigue, así, un único elemento de muchas entradas y una única salida capaz de permitir la transmisión de una, y solo una, de las entradas hacia la salida.
- Microcontrolador: PIC32MX534F064H, de la familia de microcontroladores de 32 bits, económico, con alta capacidad de procesamiento, diversidad de interfaces de comunicación y multitud de puertos de entrada y salida; además de que su entorno de desarrollo y su compilador son gratuitos. Este microcontrolador se programa mediante el protocolo de comunicación I2C y logra el contacto entre sensores y microcontrolador mediante buses línea de reloj “SCL” y buses línea de datos “SDA”.

Esta versión del ROBOBO, aún en pleno desarrollo, es la que se utilizará en este trabajo de final de grado que se centrará en el desarrollo de librerías de interacción básica entre el robot y humanos, enfocándose específicamente en interacción por voz, táctil, sonido e imagen, dando lugar a módulos simples que

más tarde podrán ser combinados en comportamientos interactivos más complejos. El desarrollo fundamental de este trabajo se realizará en Android, ya que inicialmente el ROBOBO solo soportará teléfonos con este sistema operativo.

1.2. Objetivos

El objetivo de este proyecto, como se comentó en la sección anterior, es el desarrollo de librerías que permitan una interacción básica entre el ROBOBO y humanos. Estas librerías serán desarrolladas en Android, sistema operativo que soporta el ROBOBO a día de hoy. Se desarrollarán 5 librerías Android, cada una enfocada a un tipo de interacción diferente con el robot:

- Librería de interacción por voz
- Librería de interacción por sonido
- Librería de interacción táctil
- Librería de interacción por imagen
- Librería de interacción mediante mensajes

Finalmente se implementarán una serie de ejemplos para demostrar el funcionamiento de las diferentes librerías de forma conjunta.

1.3. Estructura de la memoria

La memoria del trabajo está estructurada en 6 capítulos, siendo el primero de ellos esta introducción.

- Segundo capítulo: **Fundamentos teóricos**, se habla sobre el ROBOBO! Framework, marco de desarrollo empleado para el desarrollo de los módulos que componen este trabajo, cómo interactúa con las diferentes partes del sistema ROBOBO y las tecnologías en las que se basa.
- Tercer capítulo: **Antecedentes**, se hablará sobre el campo de la interacción humano robot y se expondrán ejemplos de los diferentes modos de interacción que surgen en los diferentes campos de la robótica.

- Cuarto capítulo: ***Fundamentos Tecnológicos*** se dará un repaso a las diferentes tecnologías empleadas a la hora del desarrollo de los módulos de interacción.
- Quinto capítulo: ***Desarrollo***, se elabora el desarrollo técnico del trabajo, el modelo conceptual, la arquitectura escogida, la metodología seguida y los detalles de implementación.
- Sexto capítulo: ***Resultados y pruebas***, se comentan los resultados del desarrollo y se exponen las aplicaciones de ejemplo desarrolladas utilizando los módulos de interacción. También se exponen los problemas conocidos aún sin solución.
- Séptimo capítulo: ***Conclusiones***, en este último capítulo se comenta el resultado general del trabajo así como el trabajo futuro a desarrollar.

1.4. Herramientas utilizadas

Para el desarrollo de las librerías y aplicaciones Android que conforman el presente trabajo, se ha empleado el entorno de desarrollo **Android Studio**, basado en el IntelliJ IDEA de JetBrains. Actualmente Android Studio es el entorno de desarrollo oficial para Android.

Para la creación de los diagramas se ha empleado el software **StarUML**.

Para la redacción de la memoria se ha empleado el software **Texpad**.

Capítulo 2

Fundamentos teóricos

EN este capítulo se introducen los fundamentos teóricos sobre los que se realizará el trabajo, en concreto acerca del ROBOBO! Framework, un framework de desarrollo creado específicamente para el ROBOBO.

2.1. ROBOBO! Framework

El ROBOBO! Framework es el marco de trabajo empleado para desarrollar el software para el ROBOBO y cuyas funcionalidades se expandirán mediante los módulos de interacción que se desarrollarán en este trabajo. El framework está enfocado al desarrollo de aplicaciones Android, sistema operativo del que se hablará en la sección 2.1.1, mediante el lenguaje de programación Java.

En la figura 2.1 se puede ver como se estructura el sistema ROBOBO y cual es el rol del framework dentro del mismo.

La *ROBOBO! Custom App*, de la figura, representa la aplicación final que el usuario instalará en su teléfono móvil y está construida sobre el framework, utilizando los diferentes módulos que este ofrece.

Dentro del framework se pueden encontrar diferentes módulos que proporcionan funcionalidades para el desarrollo de aplicaciones. Dos de ellos resultan especialmente importantes:

ros-module:

El *ros-module*, proporciona una interfaz para pasar de acciones nativas del ROBOBO a conceptos propios de ROS, servicios o temas. Este modulo da acceso

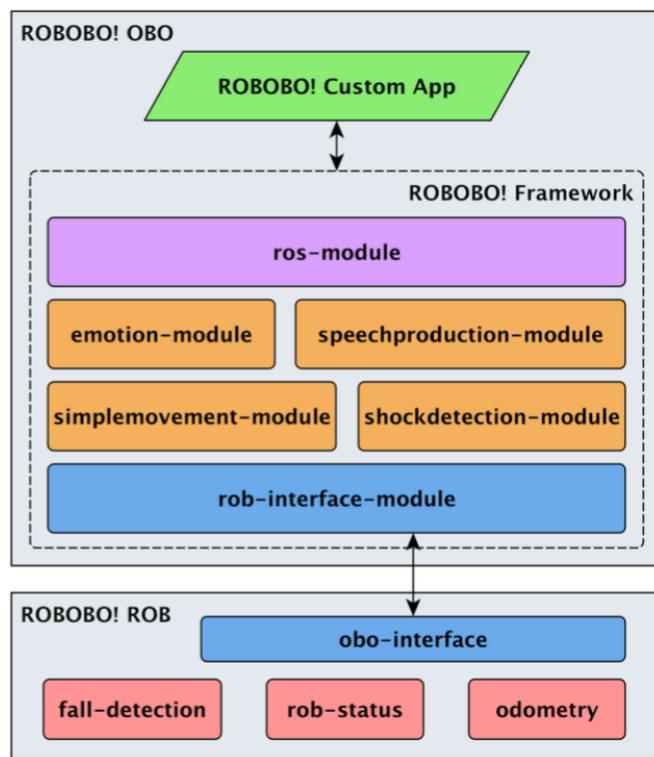


Figura 2.1: Estructura de ROBOBO! framework dentro del entorno ROBOBO

desde ROS a:

- Las funcionalidades del ROB
- Los módulos del ROBOBO
- La cámara del dispositivo
- Los datos de los sensores del Smartphone

Permite desarrollar tanto aplicaciones remotas, que no están alojadas en el smartphone pero hacen uso de los servicios o temas ejecutándose en el ROBOBO, como aplicaciones ROS nativas.

Más adelante, en la sección 2.1.2, se hablará con mas detalle de ROS .

rob-interface-module:

Este modulo nos permite obtener la interfaz de control de bajo nivel del robot, la clase *IRob*(figura 2.2) , mediante el uso de esta interfaz se puede:

Recibir mensajes periódicos del estado del ROB:

- Datos de los sensores de infrarrojos
- Información de detección de caídas
- Datos del movimiento de los motores
- Estado de la batería

Enviar comandos al ROB:

- Movimiento de los motores
- Cambio de colores de los LEDs
- Modificación de la periodicidad de los mensajes de estado
- Cambiar el modo del ROB

Existen módulos para facilitar el uso del robot y evitar tener que usar la clase *IRob*, un ejemplo es el *rob-movement-module*(figura 2.3) que proporciona una interfaz de alto nivel para manejar los movimientos del robot.

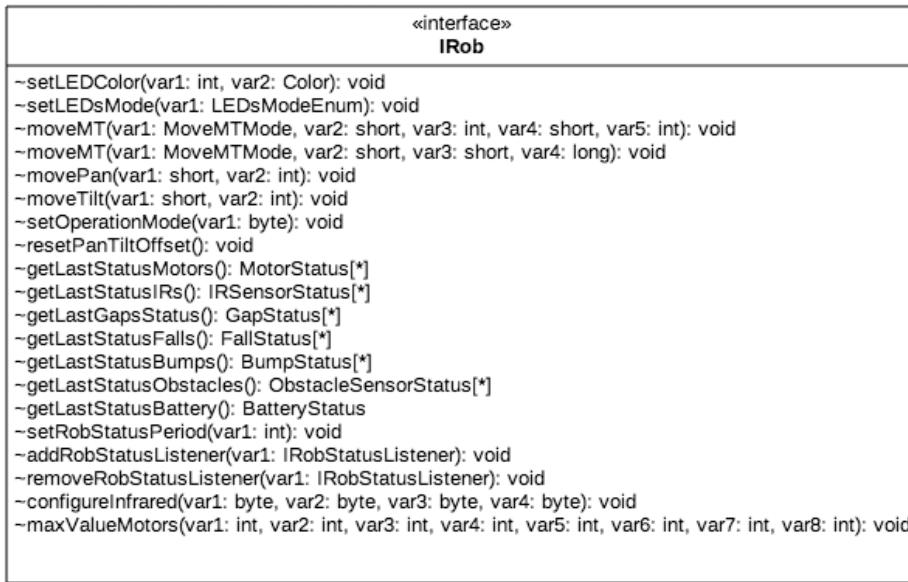


Figura 2.2: Clase IRob

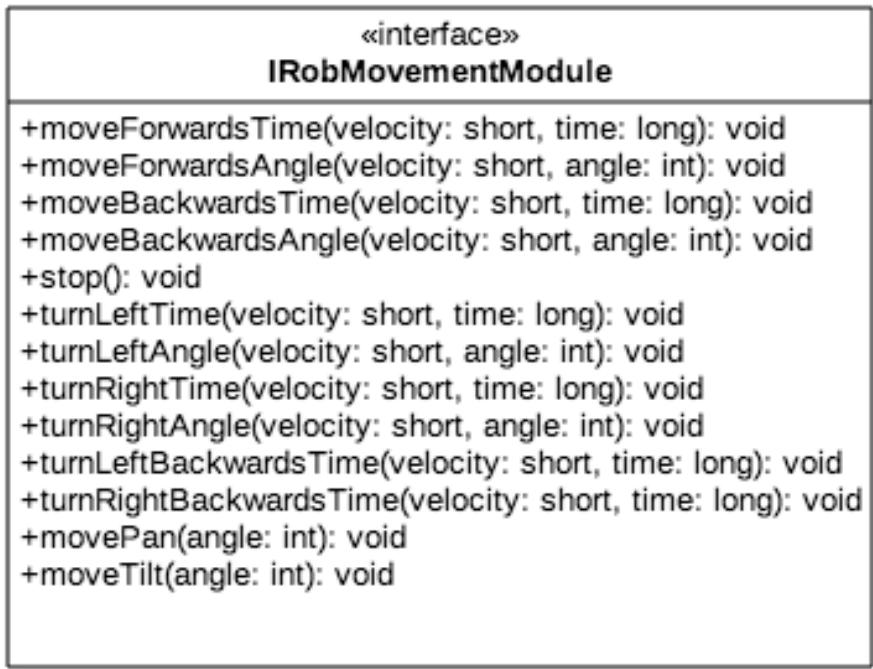


Figura 2.3: Clase IRobMovementModule

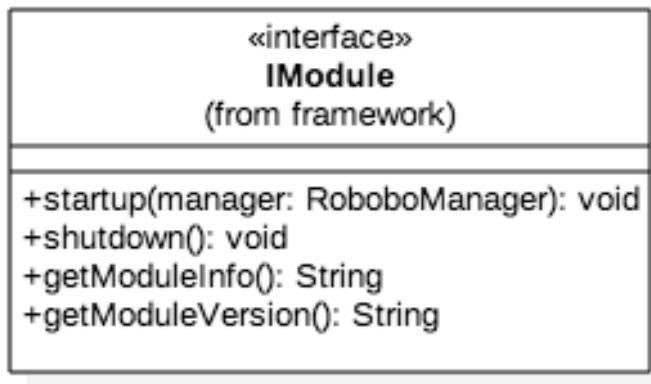


Figura 2.4: Interfaz IModule

No tengo muy claro como funciona la parte que va dentro del ROB

En el ROB se encuentra la *obo-interface*, cuya finalidad es mantener la comunicación con el smartphone, y una serie de módulos programados en el firmware de la plataforma como la odometría y la detección de caídas.

El ROBOBO! Framework sigue una filosofía modular, para ello proporciona una interfaz común, *IModule* (figura 2.4), para los módulos que permite al framework gestionar el arranque y la finalización de los mismos

2.1.1. Android

Android es un sistema operativo de código abierto desarrollado por Google enfocado a dispositivos móviles. Actualmente es el sistema operativo más extendido en cuanto a terminales móviles, ocupando un 82.8 % de la cuota de mercado en el segundo cuarto de 2015 (IDC, Aug 2015) y contando la tienda de aplicaciones de la plataforma, la Play Store, a Junio de 2016 con 2.2 millones de aplicaciones disponibles. Dados estos datos de extensión y con la capacidad de cómputo y sensorización de los teléfonos se escogió este sistema como base para el ROBOBO. Android está construido sobre un núcleo Linux sobre el cual corre el *Android Runtime* (ART) o en versiones anteriores del sistema operativo la máquina virtual *Dalvik*, siendo estos los entornos de ejecución de las propias aplicaciones Android.

2.1.2. ROS

El Sistema Operativo Robótico [2] , por sus siglas en inglés ROS, es un framework libre de desarrollo de software para robots que provee los servicios que se podrían esperar de un sistema operativo, tales como la abstracción del hardware, la gestión de dispositivos a bajo nivel, implementación de funcionalidades comunes, paso de mensajes entre procesos y gestión de paquetes, sin embargo, ROS no es un sistema operativo y debe ejecutarse sobre un entorno Linux. ROS esta organizado en forma de grafo en el cual los diferentes procesos procesan datos en red. En esta red podemos encontrar diferentes elementos:

- Nodos:procesos que realizan una computación. Un sistema de control robótico suele contar con multitud de ellos que procesan datos en paralelo.
- Master: provee servicios de registro de nombres y consulta al resto del grafo computacional, funciona como punto de encuentro entre varios nodos.
- Servidor de parámetros: permite almacenar datos bajo un indice en una ubicación central, es parte del Master.
- Mensajes: Sistema de comunicación entre nodos, estructuras de datos formadas por tipos simples.
- Temas(topics): Semejante a un bus de datos, permite la generación de mensajes y la suscripción al mismo por parte de un nodo para recibirlos.
- Servicios: Mecanismo que permite a un nodo recibir peticiones y responderlas de manera concreta.
- Bolsas(bags): formato para almacenar y reproducir mensajes. Permite la reutilización de datos como los de sensorización para ser empleados posteriormente en desarrollo o pruebas.

ROS cuenta con una amplia comunidad a lo largo del mundo que contribuye de manera activa al proyecto. El repositorio de software de ROS contiene multitud de paquetes que proporcionan funcionalidades de forma simple, de manera que el programador pueda centrarse en el desarrollo del control de su robot sin tener que preocuparse de implementar los algoritmos y técnicas más comunes.

La funcionalidad de utilizar ROS en el ROBOBO esta provista por el módulo del ROBOBO! Framework *ros-module* que emplea la librería **rosjava** para este cometido.

Capítulo 3

Antecedentes

En este capítulo se introduce el tema de la Interacción humano robot, y el diseño conceptual del sistema de interacción desarrollado para la plataforma ROBOBO.

3.1. Interacción humano robot

Desde finales de la década de los 90, la interacción entre humanos y robots ha cobrado importancia, creando un nuevo campo de investigación en la ciencia [3], la interacción humano-robot (HRI por sus siglas en inglés). El campo de la HRI busca entender, modelar y evaluar las diferentes modalidades de interacción entre las personas y los robots. La comunicación entre humanos y robots puede dividirse en dos categorías generales:

- Interacción remota
- Interacción próxima

La interacción remota suele ser referida como control supervisado o teleoperación, dependiendo de si el robot es autónomo con supervisión de un humano, que interviene en caso de necesidad, o si el robot es controlado por el humano directamente. Este tipo de interacción puede verse en robots de tipo industrial o en vehículos autónomos, como los llamados Drones del ejercito.

La interacción próxima es aquella en la que el robot interactúa directamente con el humano, llegando incluso a haber interacción física. Este tipo de interacción incluye elementos emotivos y sociales, y se puede encontrar en, por ejemplo, los

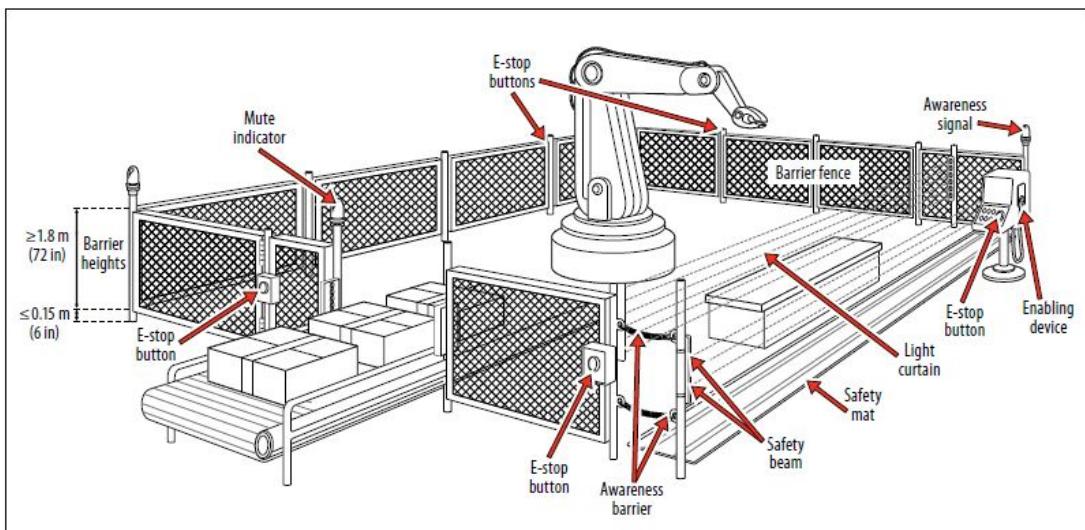


Figura 3.1: Esquema de una jaula de seguridad para un robot industrial

robots asistenciales o educativos. Este tipo de interacción es la que se tratará en este trabajo, en el cual se diseñarán diferentes sistemas de interacción para la plataforma ROBOBO.

3.1.1. HRI en la industria

Generalmente, en los procesos industriales, la interacción entre los robots y los operadores suele ser remota, los sistemas se programan para realizar una tarea y el humano solamente interviene en caso de necesidad, sin embargo pueden darse casos de interacción próxima con los robots. Uno de estos casos podría ser el aprendizaje de tareas mediante demostración, proceso mediante el cual, un operador humano realiza una tarea, por ejemplo moviendo manualmente el brazo de un robot, para que el controlador aprenda a realizar esa tarea. Este tipo de interacción permite que los robots aprendan comportamientos de alto nivel difícilmente programables.

Sin embargo a la hora de realizar tareas de forma cooperativa entre robots y humanos, la interacción cercana con robots industriales conlleva riesgos importantes, como pudo verse en el accidente del 2015 [4] en la planta de Volkswagen cerca de Kassel, Alemania, en el que un operario fué golpeado por un brazo industrial durante su instalación, resultando en la muerte del técnico. Para evitar esta clase de accidentes, se está buscando la conciencia del entorno en los manipula-

dores robóticos, para poder adaptar sus reacciones al contexto actual. Este tipo de conciencia no solo disminuye los riesgos de operación, sino que también disminuye los costes espaciales, ya que los robots no requerirían de jaulas de seguridad (figura 3.1) , también la productividad se vería afectada positivamente, ya que tareas imposibles de realizar para un robot y para un humano individualmente, pueden llevarse a cabo mediante la llamada robótica cooperativa. En *Cooperative Tasks between Humans and Robots in Industrial Environments* [5] presentan un sistema de robótica cooperativa en el que un operador y un robot colaboran de manera cercana para llevar a cabo diferentes tareas, de manera que el robot realiza las tareas repetitivas y peligrosas, mientras que el humano lleva a cabo las tareas que requieren de cierta precisión o inteligencia con la que no cuenta el robot. En este sistema el operador lleva un traje de posicionamiento, que permite al robot conocer su posición, pudiendo así adaptar sus movimientos de manera que el humano no corra riesgos.

3.1.2. HRI en robots asistenciales

Uno de los campos en los que la interacción entre humanos y robots cobra mucha importancia es en el nicho de los robots asistenciales. Los robots asistenciales, también llamados de servicio, son definidos por la federación internacional de robótica como *Robots que operan de forma total o semiautónoma para realizar servicios útiles para el bienestar de humanos y equipamiento, excluyendo las operaciones de manufactura* [6]. En esta definición se diferencia entre dos tipos de robots asistenciales:

- Robots personales
- Robots profesionales

Los robots personales son aquellos que se utilizan para labores no comerciales, generalmente por personas sin perfil técnico. Por ejemplo, sillas de ruedas eléctricas, robots de asistencia de movilidad, o aspiradoras automáticas.

Los robots profesionales son aquellos utilizados para realizar tareas de asistencia en un entorno comercial, generalmente manejados y supervisados por un personal especializado. Por ejemplo robots de limpieza automatizados para zonas públicas, robots de mensajería en oficinas u hospitales, robots anti-incendios, robots quirúrgicos y de rehabilitación en hospitales o los robots terapéuticos.



Figura 3.2: Robot terapeutico Paro

En los robots terapéuticos se pueden encontrar múltiples formas de interacción, por ejemplo, el robot Paro [7] (figura 3.2) es un robot terapéutico con forma de bebé foca, utilizado con éxito en terapias contra la Demencia, que busca una interacción emocional con el paciente, para ello cuenta con cinco tipos de sensores diferentes, táctiles, auditivos, de temperatura, de luz y posturales. Los pacientes realizan una interacción con el robot como la que tendrían con un animal, y el robot responde acorde a los estímulos que recibe. Esto, en conjunto con la forma física del robot, más semejante a un animal de peluche que a una máquina, permite al paciente desarrollar emociones. El robot NAO [8](figura 3.3) es otro de los robots que se están empleando con éxito en tareas asistenciales, tanto de manera terapéutica, como robot de relaciones publicas o para tareas de educación. El robot cuenta con una amplia variedad de sensores y actuadores que le permiten interactuar de diversas formas con el usuario:

- Cámaras: Permiten reconocimiento de caras y procesado de imagen.
- Sensores táctiles y de presión: Permiten una interacción física con el robot.
- Altavoces: Permiten al robot producir diversos sonidos y hablar.
- Micrófonos: Permiten reconocer habla y ubicar el origen de los sonidos espacialmente.
- Sensores de distancia: Permiten detectar la distancia a los objetos.

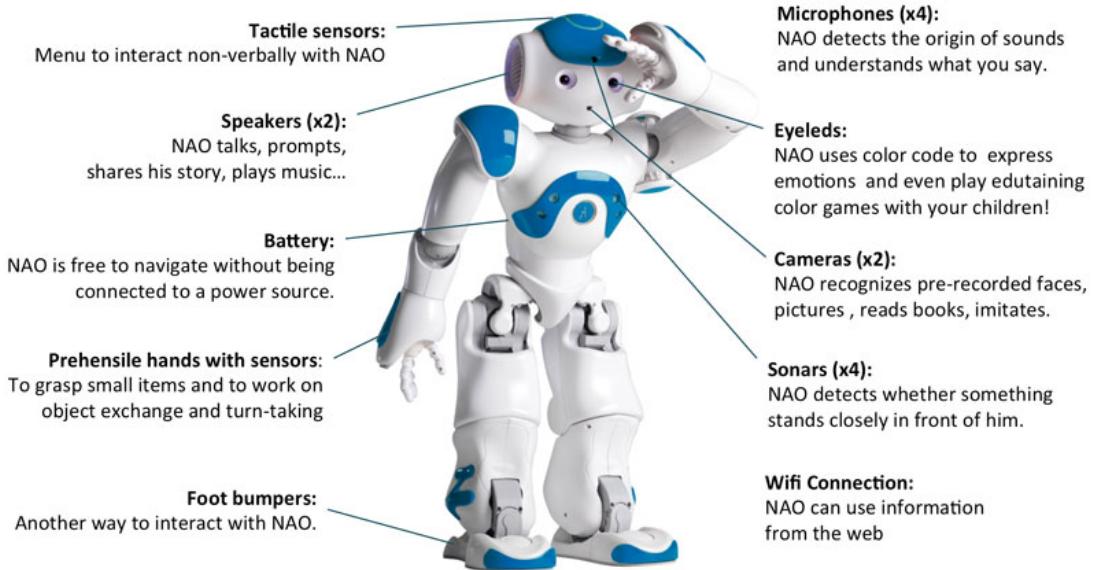


Figura 3.3: Esquema de características del robot NAO

- 25 Grados de libertad: Permiten al robot interactuar físicamente con su entorno y realizar comunicación no verbal.
- Unidad de medición inercial: Permite detectar aceleraciones y giros.

En el caso de los robots médicos, los quirúrgicos específicamente, la interacción no suele pasar de la teleoperación del robot, es decir, el robot se controla como si fuera una extensión del cirujano. El robot más conocido en este campo es el sistema quirúrgico Da Vinci (figura 3.4), utilizado en cirugía de precisión. En este robot el cirujano controla los diferentes brazos del aparato desde una consola que cuenta con controles con feedback háptico, es decir, el operador no solamente mueve el brazo, sino que siente lo que hay al final del mismo, la presión ejercida y la resistencia al movimiento, dando al cirujano el tacto necesario para realizar las diferentes tareas que se realizan en una operación, como coser o cortar, de forma natural y con un alto grado de precisión, ya que permite escalar los movimientos del cirujano.

3.1.3. HRI en robots de entretenimiento

Se entiende por robots de entretenimiento aquellos cuya finalidad no es más que divertir al usuario. Dentro de esta categoría podríamos incluir a los robots



Figura 3.4: Sistema quirúrgico daVinci

mascota, que generalmente realizan una interacción de alto nivel con el usuario. Por ejemplo, uno de los robots más relevantes en el ámbito de los robots mascota es el desarrollado por Sony, Aibo [9] (figura 3.5) , cuya finalidad era comportarse como un perro. En este robot podemos encontrar múltiples tipos de interacción, interacción física en forma de movimientos perrunos, reconocimiento facial a través de cámaras, a través de caricias utilizando los sensores táctiles, y en las últimas versiones del robot mediante una matriz de leds situada en la cara del aparato, que permite poner diferentes expresiones en función del "humor" del robot. Mediante todas estas capacidades motoras y sensoriales, se puede interactuar con una unidad Aibo de manera semejante a la que se tendría con un perro real. Otro ejemplo de mascota robótica es el Bandai SmartPet (figura 3.6) , un robot pensado para utilizar junto un smartphone IPhone, que es colocado en la cabeza del robot y provee múltiples formas de interacción, reconocimiento de gestos mediante la cámara frontal, reconocimiento de sonido utilizando el micrófono y reconocimiento de gestos táctiles en la pantalla.



Figura 3.5: Robot mascota Aibo de Sony



Figura 3.6: Robot mascota Smartpet de Bandai



Figura 3.7: Beebot

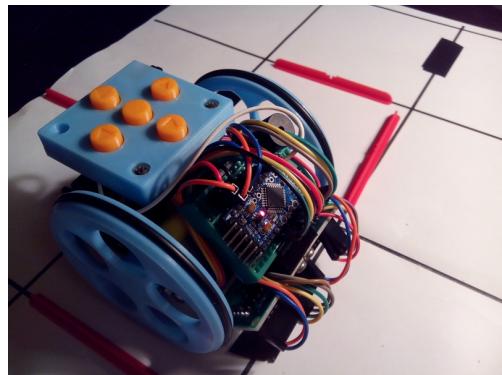


Figura 3.8: Escornabot

3.1.4. HRI en robots educativos

En los últimos años el uso de robots con fines educativos, si bien ya eran usados de manera educativa en enseñanza superior, ha tomado impulso en la educación primaria y secundaria, apareciendo múltiples robots enfocados a este tipo de mercado. La interacción con esta clase de robots puede darse de diferentes formas según el público objetivo, dependiendo principalmente del rango de edades del mismo.

Enfocados a la educación infantil nos podemos encontrar con robots como el BeeBot (figura 3.7) o su alternativa libre, el Escornabot (figura 3.8), la finalidad de estos robots es la introducción a los niños a la programación, en forma de pensamiento secuencial, y a la resolución de problemas. En esta clase de robots la interacción está limitada a la introducción de comandos en la botonera de la parte superior del robot, que serán traducidos a movimientos del robot posteriormente.

En la educación primaria y secundaria los robots utilizados ya adquieren una mayor complejidad y suelen emplearse para una introducción real a la programación, generalmente utilizando lenguajes gráficos de muy alto nivel como Scratch [10]. Uno de los robots más relevantes en este ámbito es el kit Mindstorms de Lego (figura 3.9), que no solo permite programar el robot, sino también construirlo. El Lego Mindstorms cuenta con diferentes sensores y actuadores que ofrecen diversas maneras de interacción con el robot:

- Motores: Permiten el movimiento del robot de manera relativamente precisa
- Sensores de distancia: Permiten medir distancias y actuar en consecuencia a los datos medidos



Figura 3.9: Kit básico Lego Mindstorms EV3

- Sensores de luz ambiente: Miden la intensidad de la luz del entorno
- Sensores de color: Permiten detectar los colores básicos
- Unidad inercial: Permite medir giros en el plano horizontal

En la llamada educación especial la robótica también está siendo empleada con éxito, por ejemplo el robot NAO (figura 3.3), del que se habló en la sección anterior, se utiliza para educar a niños con trastornos de espectro autista. El éxito de este tipo de educación viene dada debido a que la interacción con el robot, al ser programada, resulta predecible para el alumno, creando un entorno estable y pautado que resulta óptimo en este tipo de trastornos. La interacción en este caso se da en forma de movimientos preprogramados y mediante los leds de los ojos del robot, que cambian de color según las emociones que intenta expresar el robot, lo cual ayuda al alumno a ejercitarse uno de las mayores dificultades dadas por el autismo, la dificultad de establecer relaciones empáticas.

Entre otros ejemplos de robots educativos podemos encontrar el Zowi (figura 3.10) de la empresa española BQ. Zowi es un robot educativo enfocado a niños, cuenta con un sensor de ultrasonidos para medir distancias, un micrófono que le permite detectar sonidos, una pequeña matriz de leds a modo de boca, que permite comunicarse mediante expresiones faciales, y con cuatro servos posicionados

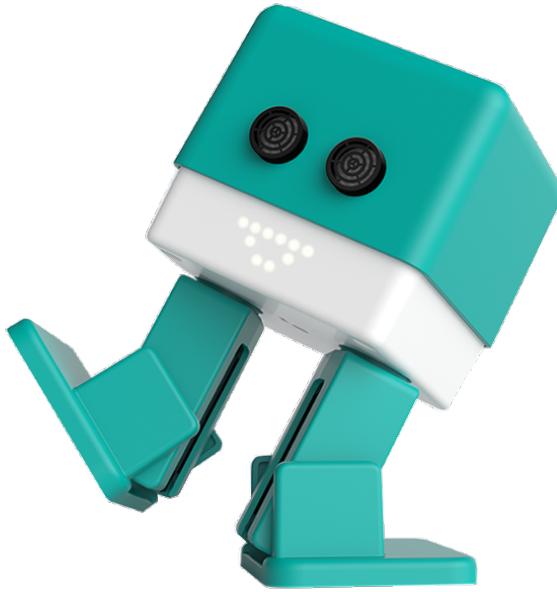


Figura 3.10: Robot Zowi de BQ

en forma de dos piernas que permiten moverse al robot. El robot viene con varios juegos para niños instalados de serie, y cuanta además con una aplicación de control para smartphones. Zowi puede ser empleado para enseñar programación a niños con un lenguaje gráfico similar a Scratch [10], Bitbloq [11].

Otro robot educativo es el Thymio (figura 3.11). Este robot cuenta con gran variedad de sensores, de distancia, acelerómetros, temperatura, control por infrarrojos, seguidores de líneas y micrófono, y con dos ruedas compatibles con las piezas de lego. La interacción con el usuario se realiza a través de los 39 leds RGB situados por todo el chasis, 5 teclas capacitivas que se encuentran en la parte superior del robot y los sensores citados anteriormente. El robot está pensado para enseñar programación a los niños de diferentes rangos de edad, para ello provee 3 lenguajes diferentes de programación, Visual Programming Language (VPL) para los niños a partir de 6 años, Blockly, un lenguaje de programación gráfico semejante a Scratch desarrollado por Google, para niños a partir de 9 y programación textual mediante el software Aseba Studio para niños a partir de 12 años.

Dash and Dot (figura 3.12) son una pareja de robots educativos fabricados por Wonder Workshop enfocados a niños de poca edad. Dash cuenta con una



Figura 3.11: Robot Thymio

plataforma motorizada equipada con sensores de distancia, «orejas» motorizadas, tres micrófonos, un altavoz, un «ojo circular» con 12 leds blancos, leds rgb en las «orejas» y «pecho». Dot en cambio, es una alternativa más económica que solo incluye la «cabeza» de Dash. Gracias a la gran cantidad de actuadores, Dash, y en menor medida Dot, cuentan con diversas formas de interacción con los niños, desde el movimiento del robot, habla a través del altavoz, gestos con los leds del ojo e incluso, mediante accesorios vendidos de forma separada, tocando un pequeño xilófono o lanzando pelotas con una pequeña catapulta. Dot además cuenta con un accesorio «orejas de conejo» que le permite suplir la falta de plataforma motorizada y moverse. Estos robots están pensados para enseñar programación a los niños, y cuentan con varias aplicaciones diseñadas para este cometido, *Path*, que permite a los niños dibujar la ruta que desean que tome el robot, *Wonder*, un lenguaje visual de programación y *Blockly*, el lenguaje de programación por bloques desarrollado por Google.

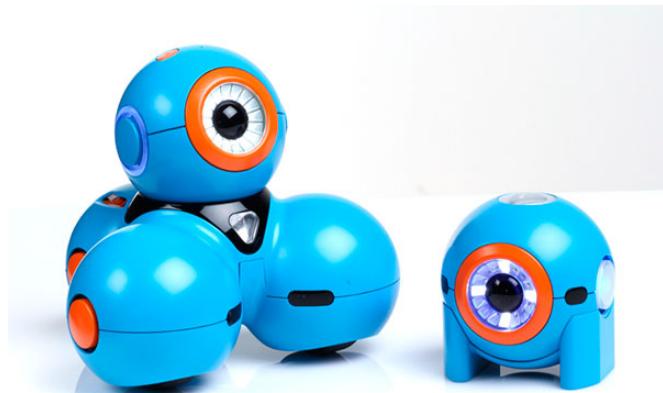


Figura 3.12: Robots Dash y Dot

Capítulo 4

Fundamentos tecnológicos

4.0.1. Sphinx

4.0.2. TarsosDSP

4.0.3. OpenCV

4.0.4. Gmail Background

Capítulo 5

Desarrollo

En este capítulo se desarrollarán los detalles de la implementación de los diferentes módulos de interacción que se han diseñado para el ROBOBO así como la metodología de trabajo seguida.

5.1. Modelo conceptual para el subsistema de interacción

La plataforma ROBOBO, cuenta con características que permiten realizar diversas maneras de interacción, el hecho de estar basado en un smartphone, pone todas las capacidades del mismo a disposición a la hora de interactuar con los usuarios. Partiendo del hardware con el que se cuenta, base motorizada *ROB* y smartphone *OBO*, se han definido, en una analogía a los sentidos, una serie de paquetes para dotar al ROBOBO de diversas funcionalidades de interacción.

El primero de estos paquetes sería el paquete de habla, que en la sección anterior se pudo ver su aplicación en robots como Nao, Dash y Dot. Siendo el habla, probablemente, la principal forma de interacción que se da entre los humanos, parece lógico que, teniendo la capacidad de proceso de un smartphone moderno, dotar al robot ya no solo de la capacidad de producir habla, sino de entender texto hablado es un concepto interesante. Este paquete daría al robot la posibilidad de comunicación bidireccional con los humanos, además de darle al robot cierta personalidad más «orgánica», haciéndolo más atractivo, por ejemplo, a los niños.

El segundo, de forma análoga al sentido del tacto, teniendo una «piel artificial»

cómo es la pantalla táctil del teléfono, dar la capacidad al robot de sentir toques y caricias parece una idea interesante a la hora de llevar una interacción cercana con el robot, este tipo de interacción mediante toques en la pantalla táctil es similar al del SmartPet, y de una manera más general, al sentido del tacto del robot Paro o el Aibo.

El tercero, siguiendo la analogía con los sentidos, sería el sentido del oído. La capacidad de producir y reconocer sonidos abre puertas a tipos de interacción interesantes. Por ejemplo, la capacidad de reaccionar ante sonidos fuertes como palmadas o la capacidad de reconocer notas musicales podrían permitir al robot comunicarse mediante el uso de la música, y estando el ROBOBO enfocado directamente a la educación, esta clase de interacción podría emplearse para la enseñanza musical en la educación infantil y primaria. También, la producción de sonidos, es una forma de comunicación que parece adecuada a la hora de interactuar con niños, ya sea mediante sonidos básicos (risas, alarmas, abucheos...) como a través de notas musicales, que podría usarse en conjunto con la capacidad de reconocimiento de las mismas a la hora de la enseñanza musical. Esta capacidad de producir música se puede ver en el robot Dash mediante el módulo del xilófono y la de detección de música y palmadas en el Zowi.

El cuarto paquete, teniendo hoy en día todos los smartphones como mínimo una cámara, busca dotar al ROBOBO de sentido de la vista. Con este sentido, el robot podría detectar la presencia de gente y a qué distancia se encuentran y de esta manera adaptar su comportamiento a su entorno, por ejemplo, alejándose si nota que alguien está muy cerca o siguiendo con la «cara» a la gente a su alrededor. La capacidad de discernir diferentes colores también es una habilidad interesante de cara a la educación de niños muy pequeños. La capacidad de detección de caras y colores se pudo encontrar en los robots más complejos vistos en la sección anterior, como Nao o Aibo, que cuentan con cámaras en su cabeza.

Por último, teniendo la capacidad de conectarse a internet del smartphone, parece interesante dotar al robot de la capacidad de comunicarse a través de la red. El correo electrónico es un sistema de comunicación muy extendido y casi obligatorio hoy en día, así que, dada la extensión del sistema, y la cantidad de usuarios con los que cuenta, parece la opción correcta para este tipo de comunicación.

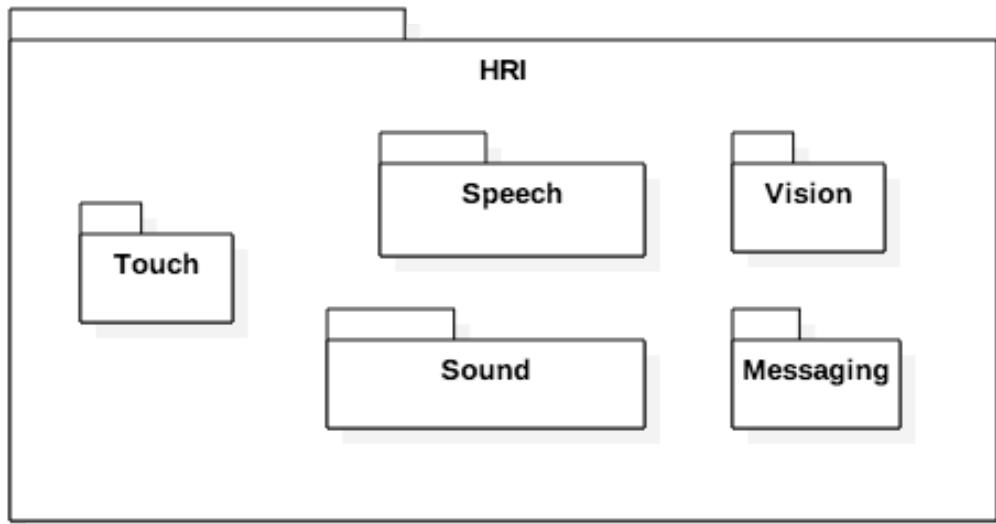


Figura 5.1: Estructura de paquetes dentro del paquete HRI del Robobo

5.2. Arquitectura Global

El sistema de interacción humana del sistema ROBOBO fue desarrollado de una forma modular, de manera que las diferentes funcionalidades puedan ser utilizadas de manera separada.

La gestión de los módulos la realiza el Framework Robobo, que proporciona una interfaz común que debe ser implementada

Cada «sentido» de los elaborados conceptualmente en la sección 5.1 se ha implementado en forma de módulo Android, y cada funcionalidad específica ha sido integrada en módulos que implementan IModule, interfaz común para los módulos del ROBOBO , para poder ser manejados por el ROBOBO! Framework [12].

- Paquete Speech: Gestiona las operaciones que involucran habla.
 - Módulo SpeechProduction: Permite utilizar la síntesis de voz.
 - Módulo SpeechRecognition: Permite el reconocimiento de voz, por palabras clave o por gramáticas.
- Paquete Touch: Gestiona el sentido del tacto del robot.

- Módulo Touch: Permite la detección de gestos en la pantalla táctil del móvil.
- Paquete Sound: Gestiona el sentido del oído del robot.
 - Módulo SoundDispatcher: Módulo auxiliar para los módulos que emplean la librería TarsosDSP [13]
 - Módulo ClapDetectionModule: Módulo que permite la detección de sonidos percusivos.
 - Módulo PitchDetection: Módulo que permite la estimación de la frecuencia de un sonido.
 - Módulo NoteDetection: Módulo que detecta notas musicales a partir de frecuencias.
- Paquete Vision: Gestiona el sentido de la vista del robot.
 - Módulo BasicCamera: Provee de un stream constante de frames capturados desde la cámara frontal del smartphone.
 - Módulo FaceDetection: Permite detectar caras en los frames capturados por la cámara.
 - Módulo ColorDetection: Permite detectar colores sobre fondos blancos en los frames.
- Paquete Messaging: Gestiona los métodos de mensajería del robot
 - Módulo Messaging: Permite enviar correos electrónicos, pudiendo adjuntar imágenes.

Cada paquete se distribuye en forma de librería Android.

5.2.1. Estructura de un módulo

Todos los módulos comparten una estructura de paquetes similar, un paquete con el nombre del módulo , que contiene la interfaz del módulo (de la forma *INombreModulo*) que extiende la clase *IModule*, una clase abstracta (*ANombreModulo*) que implementa la interfaz anterior y gestiona labores comunes como la suscripción a los Listeners, y los paquetes de las diferentes implementaciones

específicas. Opcionalmente, este paquete también podrá contener la interfaz de los listeners, y clases de utilidad como enumerados.

Las implementaciones de los módulos tienen el nombre *NombreImplementaciónNombreModulo*, y deben extender la clase abstracta citada anteriormente.

5.3. Metodología

La metodología seguida fue un proceso iterativo en el cual en cada iteración se fueron añadiendo nuevos módulos de interacción y realizando correcciones en los módulos ya implementados.

5.3.1. Primera iteración

En esta primera iteración el trabajo de desarrollo se centró en los paquetes de producción de habla (Módulo *Speech*) y el de tacto (Módulo *Touch*). Se implementaron las funcionalidades básicas de cada módulo:

- *SpeechProductionModule*: Se implementaron las funcionalidades de síntesis de voz, selección de voces y de localización.
- *SpeechRecognitionModule*: Se implementó el reconocimiento básico por palabras clave.
- *TouchModule*: Se implementaron el reconocimiento de toques largos(touch) y cortos(tap), así como el reconocimiento de caricias(caress) y deslizamientos rápidos(fling).

5.3.2. Segunda iteración

En la segunda iteración se desarrolló el paquete de interacción por sonidos y se realizaron correcciones tanto en el módulo de reconocimiento de voz como en el táctil.

Nuevas implementaciones en esta iteración:

- *SoundDispatcherModule*: Implementado para dar soporte al resto de módulos que usan la librería TarsosDSP [13].
-

- *ClapDetectionModule*: Implementado el módulo de detección de sonidos percusivos.
- *PitchDetectionModule*: Implementada el módulo de detección de frecuencias en hertzios.
- *NoteDetectionModule*: Implementada el módulo de detección de notas musicales. Implementada clase para representar las notas.

Actualizaciones de módulos de la anterior iteración:

- *SpeechRecognitionModule*: Implementada la búsqueda por gramáticas
- *TouchModule*: Mejorada la información producida por el evento *onFling*.

5.3.3. Tercera iteración

En esta iteración se desarrollaron los paquetes de visión y de mensajería. También se sacaron de dentro de los módulos los diferentes parámetros, para poder ser configurados desde el exterior mediante ficheros *properties*.

Nuevas implementaciones en esta iteración:

- *BasicCameraModule*: Implementado el notificador de frames.
- *FaceDetectionModule*: Implementado el modulo de detección de caras.
- *ColorDetectionModule*: Implementado el modulo de detección de colores.
- *MessagingModule*: Implementado el módulo de mensajería por gMail.

Actualizaciones de módulos:

- *SoundDispatcherModule*: Parametrización mediante properties.
- *ClapDetectionModule*: Parametrización mediante properties.
- *PitchDetectionModule*: Parametrización mediante properties.

5.3.4. Cuarta iteración

En esta iteración se añadieron dos módulos de producción de sonidos, y se eliminaron clases abstractas inútiles de los módulos ya existentes.

Nuevas implementaciones esta iteración:

- *EmotionSoundModule*: Implementada la reproducción de sonidos prefijados.
- *NoteGeneratorModule*: Implementada la generación de tonos.

Actualizaciones de módulos:

- *ColorDetectionModule*: Añadida una comprobación de la varianza del color para evitar falsos positivos.
- *SoundDispatcherModule*: Eliminada clase Abstracta vacía.

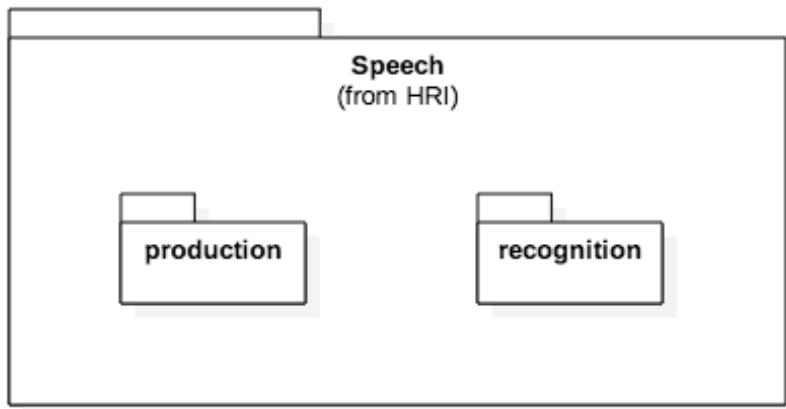


Figura 5.2: Paquete Speech

5.4. Librerías de interacción

A continuación se hablará sobre los detalles de implementación de las diferentes librerías de interacción desarrolladas.

5.4.1. Paquete Speech

Este subsistema gestiona todas las labores que tienen que ver con la generación o detención del habla. En su interior encapsula dos paquetes diferentes (figura 5.2) , Production y Recognition, que contienen respectivamente los módulos de producción y reconocimiento del habla.

5.4.1.1. Módulo recognition

El módulo recognition(figura 5.3) permite al usuario reconocer texto hablado de forma fácil. Provee las Interfaces *ISpeechRecognitionModule* y *ISpeechRecognitionListener*, así como la clase abstracta *ASpeechRecognitionModule*.

ISpeechRecognitionModule ofrece la declaración de métodos de configuración del reconocedor de voz y de suscripción de listeners.

Especificamente permite:

- Añadir y eliminar palabras reconocibles.

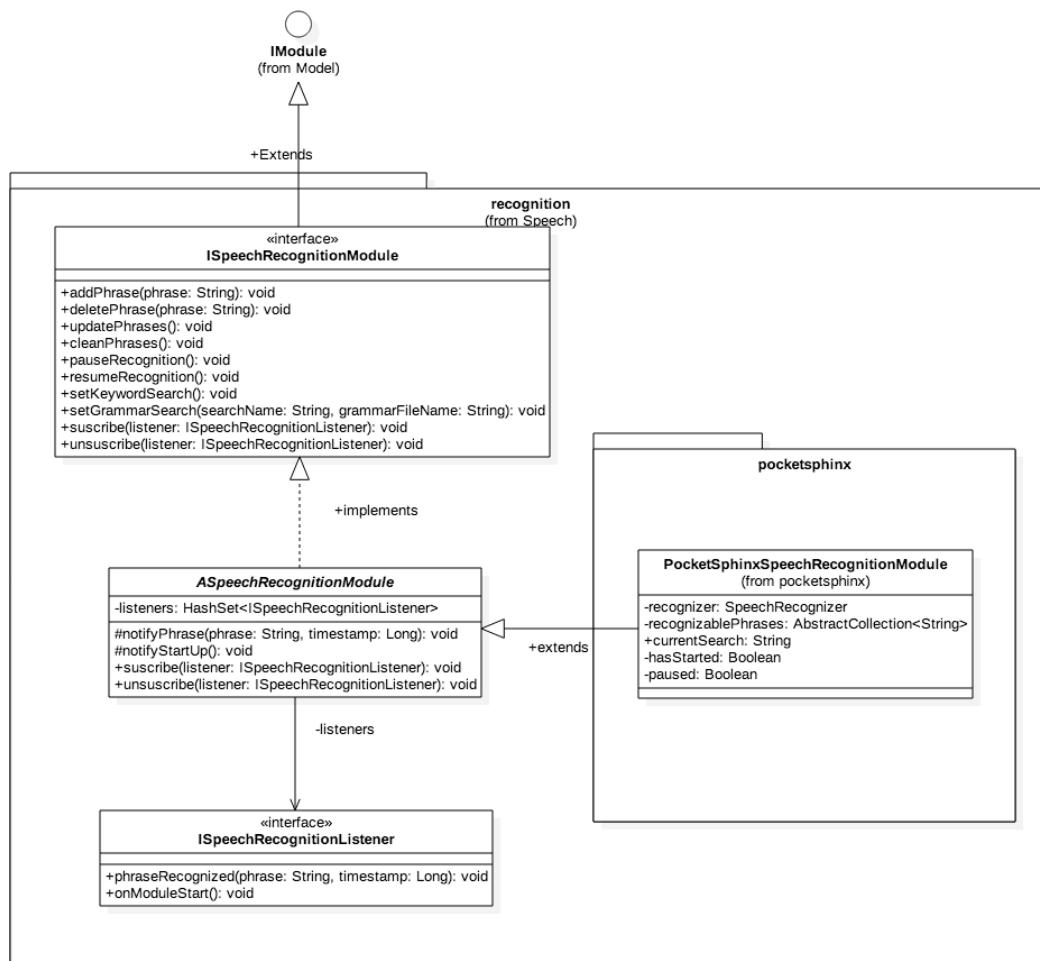


Figura 5.3: Módulo SpeechRecognition

- Actualizar el listado interno de frases
- Limpiar el listado interno
- Pausar y reanudar el reconocimiento
- Añadir búsquedas por gramática
- Utilizar búsqueda por palabras

ISpeechRecognitionListener es la interfaz que deben implementar las clases para ser notificados de los eventos que produce el módulo de reconocimiento.

Especificamente notifica cuando:

- El reconocedor ha sido inicializado
- Una frase ha sido detectada

ASpeechRecognitionModule es la clase abstracta que gestiona las tareas comunes a todas las implementaciones del modulo, a saber:

- Mantener un listado de los listeners
- Suscribir y desuscribir listeners
- Notificar a los listeners de los diferentes eventos

Pocketsphinx

Se ha decidido realizar la implementación del módulo utilizando la librería CMU Sphinx [14], desarrollada por la Universidad de Carnegie Mellon. Específicamente se utiliza PocketSphinx, una versión de CMU Sphinx ligera, diseñada para su uso en sistemas embebidos.

PocketSphinx permite varios tipos de búsqueda diferentes, pero al contrario que las versiones más completas de la librería, sólo permite tener una búsqueda activa al mismo tiempo. Entre los modos de búsqueda que ofrece se encuentran la búsqueda por Keywords y la búsqueda por Gramáticas, que son las que son ofrecidas por la interfaz del módulo. Adicionalmente PS también permite el uso de búsquedas por fonemas, pero no se ha contemplado su uso.

El módulo se llama *PocketSphinxSpeechRecognitionModule*, extiende *ASpeechRecognitionModule* e implementa la interfaz *RecognitionListener* que provee PocketSphinx.

La búsqueda por keywords permite al usuario definir, en tiempo de ejecución, una serie de palabras que podrán ser reconocidas, estas palabras deben estar contenidas en el diccionario del modelo fonético.

La búsqueda por gramática permite definir archivos de gramáticas en el formato JSFG 1.0 [15], lo cual permite detectar construcciones más complejas de palabras. Para funcionar correctamente deben colocarse en el directorio assets-sync/ de la aplicación tanto el modelo fonético del idioma, como los archivos de gramáticas que se vayan a utilizar, todo archivo dentro de este directorio debe contar con un fichero de mismo nombre y extensión .md5 conteniendo un hash del original en su interior. Los archivos deben ser declarados dentro del fichero assets.lst.

Dado que el inicio del reconocedor de voz se realiza de manera asíncrona, es necesario notificar a la aplicación principal de cuando el módulo esta preparado, de lo contrario, llamar a una operación del mismo podría ser inseguro. Le arranque del módulo se notifica a través del método *onModuleStart()* que debe ser implementado por cualquier clase que vaya a utilizar el módulo.

5.4.1.2. Módulo Production

```
public interface ISpeechProductionModule extends IModule {
    Integer PRIORITY_HIGH = 1;
    Integer PRIORITY_LOW = 0;
    void sayText(String text, Integer priority);
    void setLocale(Locale newloc);
    void selectVoice(String name) throws VoiceNotFoundException;
    void selectTtsVoice(ITtsVoice voice) throws VoiceNotFoundException;
    Collection<ITtsVoice> getVoices();
    Collection<String> getStringVoices();
}
```

El módulo production(figura 5.4) permite al usuario producir voz a partir de texto de manera simple. Provee la interfaces *ISpeechProductionModule* y *ITtsVoice*, así como la excepción *VoiceNotFoundException*. La primera contiene las

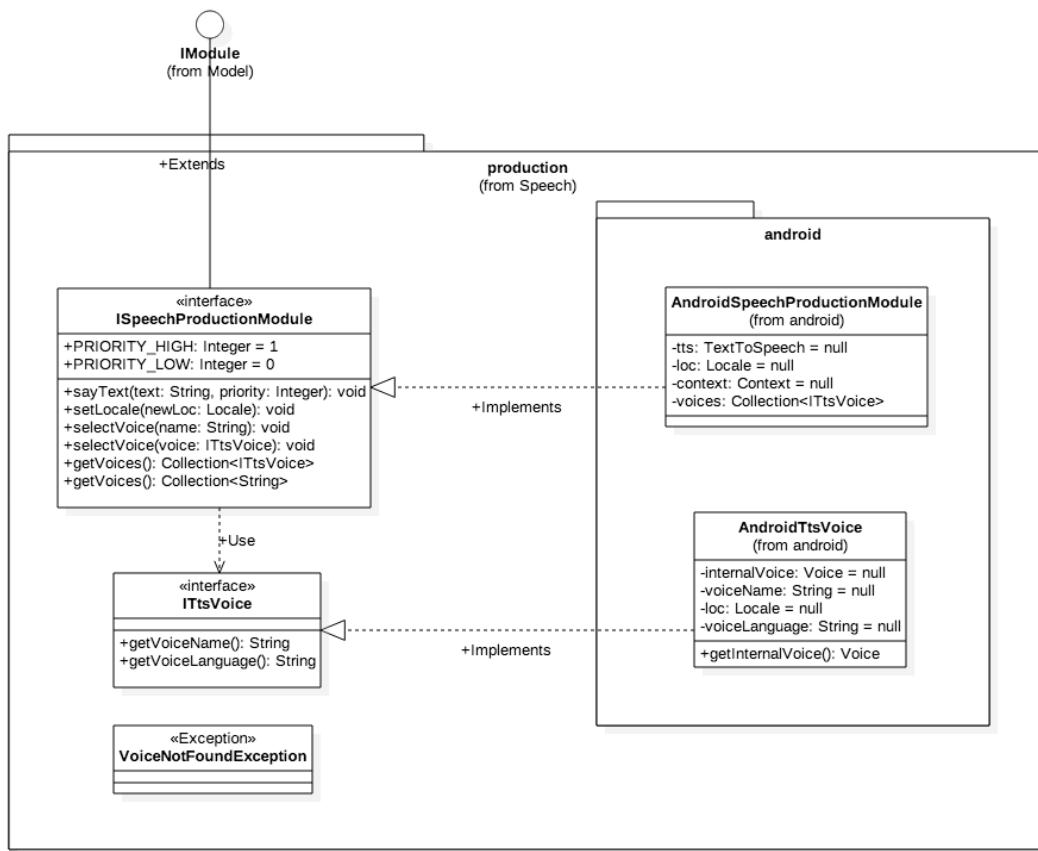


Figura 5.4: Módulo SpeechProduction

declaraciones de los métodos del módulo en sí, la segunda provee una forma de representar las diferentes voces a utilizar por el módulo.

Las funcionalidades que provee la interfaz son:

- Pronunciar un texto
- Cambiar la localización
- Seleccionar voces
- Recuperar las voces disponibles

Implementación específica:

La implementación de la interfaz se ha realizado con las librerías del propio sistema Android (**android.speech.tts.TextToSpeech**)

Ejemplos de desarrollo

Para la prueba de los módulos de voz se realizaron dos ejemplos básicos, una aplicación simple en la que el usuario introduce un texto y el teléfono lo pronuncia y un sistema de control mediante voz en el que el usuario le daba órdenes básicas al Robobo mediante la voz y este, al terminar la frase, ejecutaba los movimientos indicados.

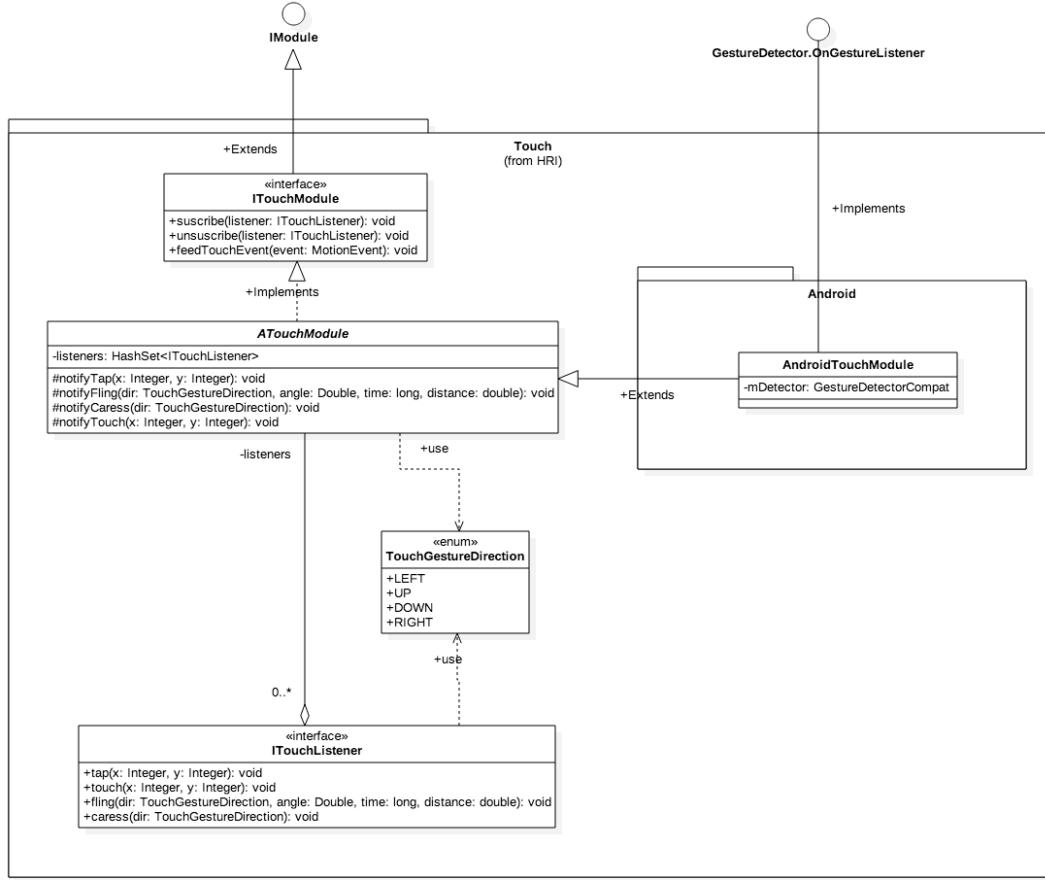


Figura 5.5: Módulo Touch

5.4.2. Paquete Touch

El subsistema Touch permite interacción con el robot en forma de gestos táctiles sobre la pantalla del teléfono. En su interior encontramos un único modulo, el módulo touch.

5.4.2.1. Módulo Touch

El módulo(figura 5.5) contiene las interfaces *ITouchModule* e *ITouchListener*, la clase abstracta *ATouchModule* y por la clase enumerado *TouchGestureDirection*.

ITouchListener es la interfaz que debe implementar cualquier clase que quiera ser informada de los gestos táctiles que ocurrán en la pantalla del dispositivo. Las

acciones que se detectan son:

- Tap: Toque simple y rápido en la pantalla
- Touch: Toque mantenido en la pantalla
- Caress: Deslizamiento sobre la pantalla
- Fling: Al terminar un deslizamiento de manera rápida

ITouchModule es la interfaz que debe cualquier implementación del módulo. El único método remarcable de la misma es *FeedTouchEvent*, en el cual se le pasan los eventos táctiles al módulo desde la aplicación principal, que debe sobrecargar el método *OnTouchEvent* de la clase *Activity*, esto es debido a que el módulo no puede extender la clase *Activity*.

Detección táctil de Android

Para la implementación del módulo se escogió usar la clase *GestureDetector* propia de Android, que provee de varios listeners para diferentes eventos táctiles. Se han simplificado dichos eventos, para producir los definidos en *ITouchListener*

Ejemplos de desarrollo

Se implementó una simple aplicación que mostraba por pantalla los diferentes eventos producidos por el módulo *TouchModule*.

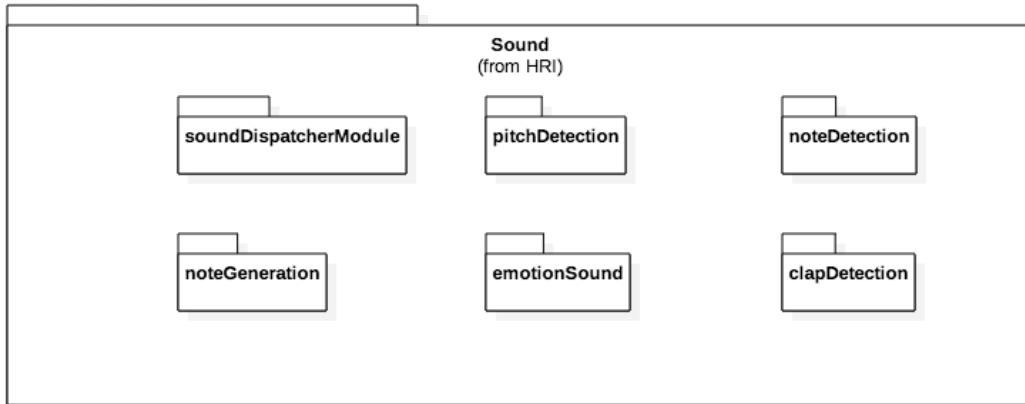


Figura 5.6: Paquete Sound

5.4.3. Paquete Sound

Este subsistema contiene los módulos (figura 5.6) que realizan diversas tareas de reconocimiento de audio. Han sido implementadas detección de tonos, notas musicales y palmadas.

5.4.3.1. Módulo Sound Dispatcher

SoundDispatcher (figura 5.7) es un módulo de utilidad, necesario para el uso de los módulos implementados a partir de la librería TarsosDSP [13]. Este módulo es el que se encarga de capturar el sonido del micrófono y segmentarlo en partes para su posterior procesado. Los distintos módulos se registran con sus procesadores en este módulo. El paquete contiene en su interior *ISoundDispatcherModule*, la interfaz del módulo, que permite

- Añadir procesadores de audio
- Eliminar procesadores de audio
- Iniciar el procesado de audio
- Parar el procesado de audio

ASoundDispatcherModule implementa *ISoundDispatcherModule*, pese a estar vacía se conserva para mantener la estructura común de los módulos.

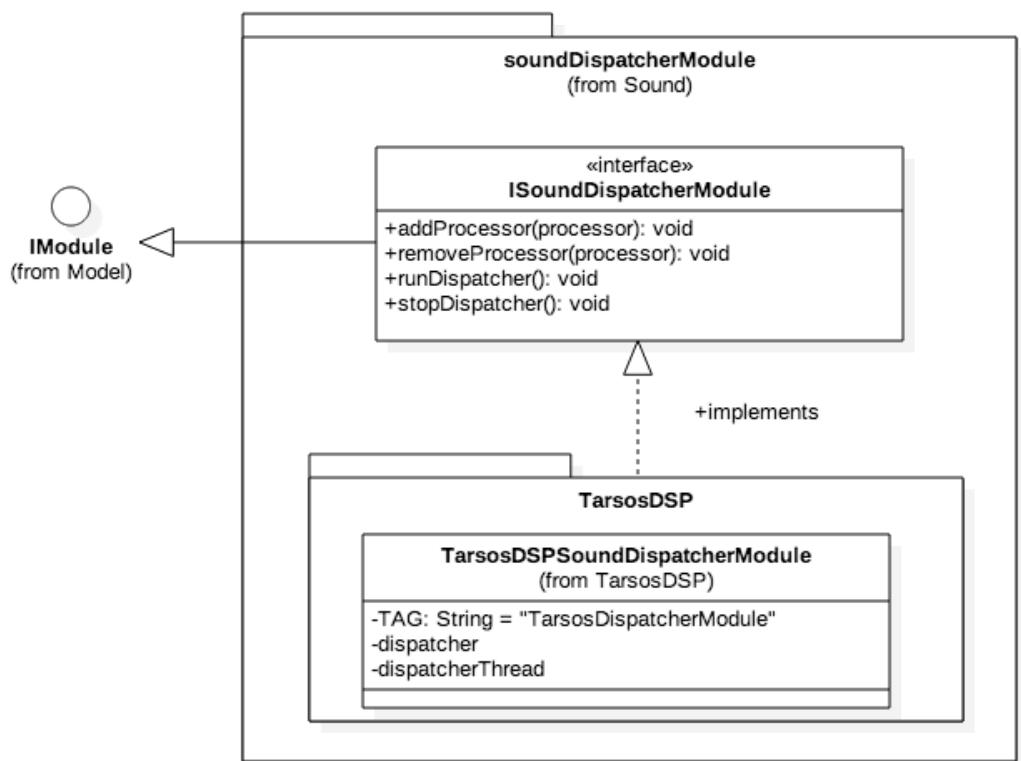


Figura 5.7: Módulo SoundDispatcher

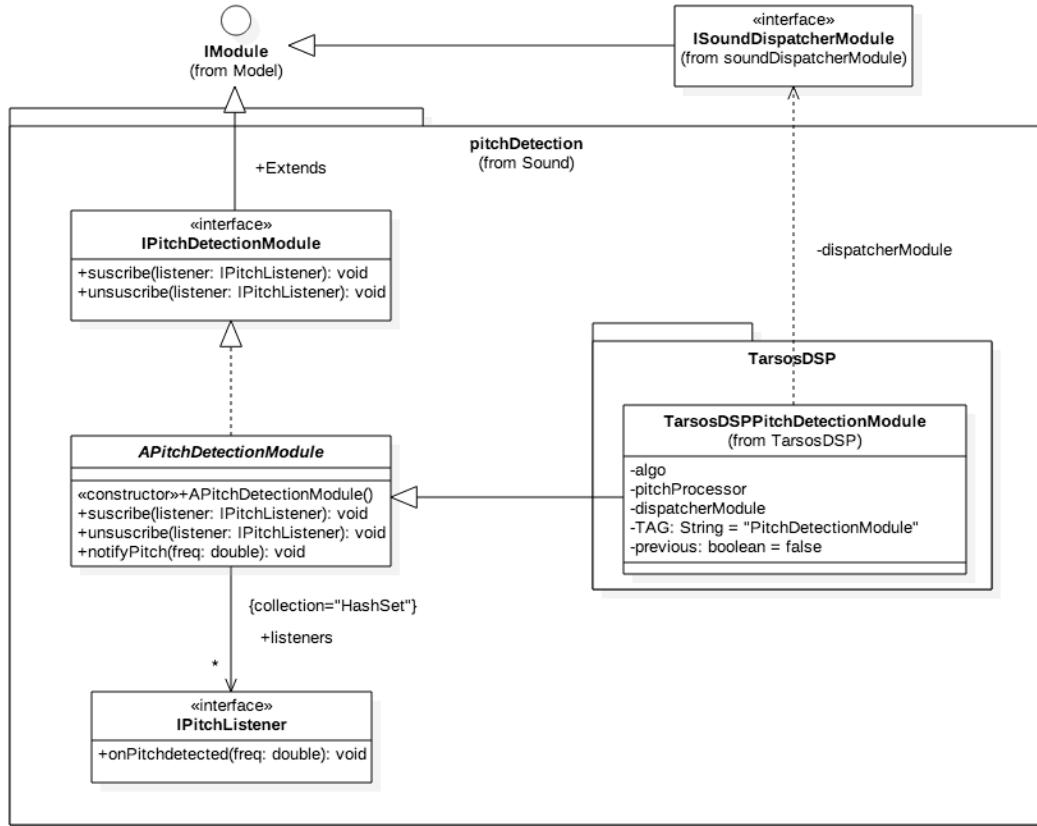


Figura 5.8: Módulo PitchDetection

TarsosDSP

La implementación se realiza en el módulo *TarsosDSPSoundDispatcherModule*, que extiende la clase *ASoundDispatcherModule*. En su interior se crea un objeto *AudioDispatcher*, que se encarga del procesado del audio, y al que pueden añadirse distintos procesadores de audio.

5.4.3.2. Módulo Pitch Detection

PitchDetection (figura 5.8) permite detectar la frecuencia en Hertzios de sonidos, es utilizado, por ejemplo en el módulo de detección de notas musicales. En su interior encontramos la interfaz *IPitchDetectionModule* y la clase abstracta *APitchDetectionModule*, que gestionan la suscripción de listeners a los eventos. También se encuentra *IPitchListener*, que debe ser implementada por cualquier clase que desee recibir notificaciones del módulo de detección de tonos.

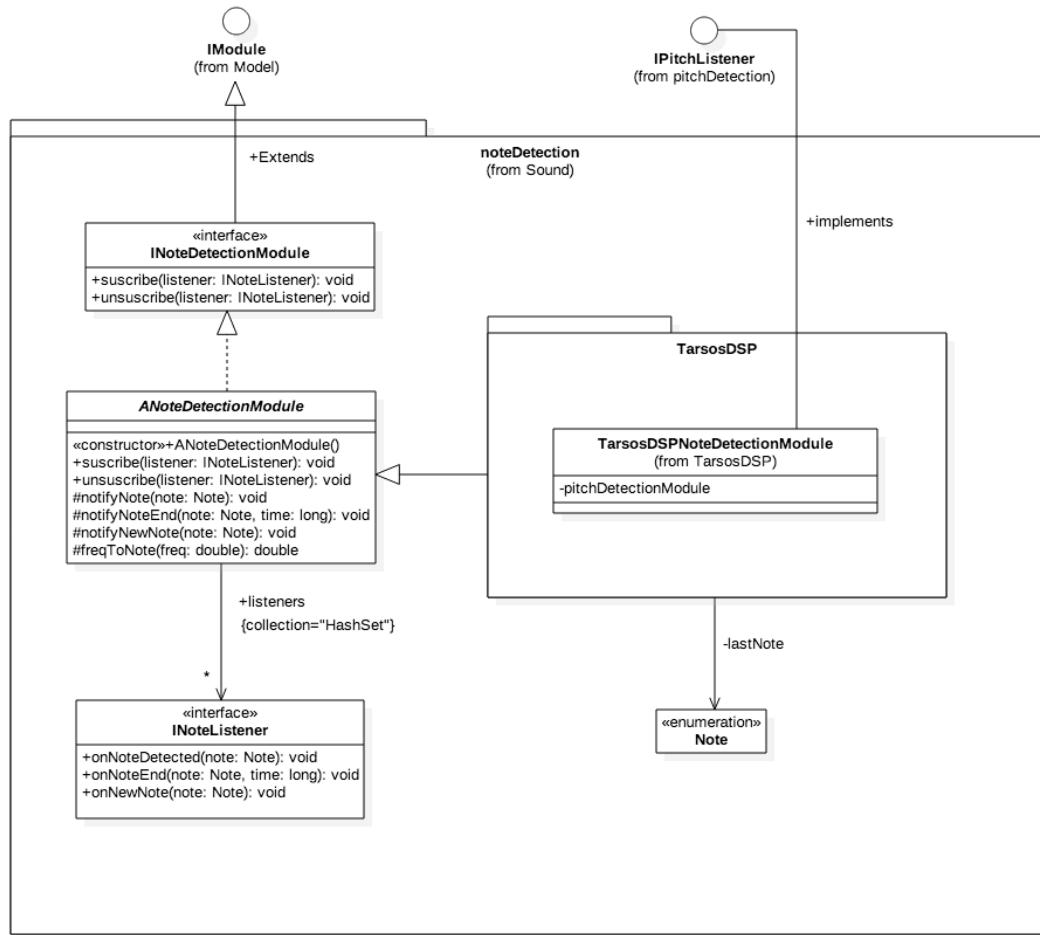


Figura 5.9: Módulo NoteDetection

TarvosDSP

La implementación del módulo se ha realizado en la clase *TarvosDSPPitch-DetectionModule*. Esta implementación depende del módulo SoundDispatcher, ya que en su inicialización se registra a si mismo en el susodicho modulo. Esta implementación utiliza el algoritmo YIN [16] para la detección de frecuencias. La implementación del algoritmo viene dada por la propia librería TarvosDSP. El módulo genera un evento con la frecuencia cuando un sonido es detectado, y al terminar la detección devuelve un -1.

5.4.3.3. Módulo Note Detection

El módulo NoteDetection(figura 5.9) busca proporcionar una manera simple de detectar notas musicales. Depende directamente del módulo PitchDetection, del cual obtiene la frecuencia que traduce en notas musicales. En su interior encontramos la interfaz del módulo *INoteDetectionModule*, la clase abstracta *AnoteDetectionModule*, que además de la gestión de listeners provee un método de conversión de frecuencias a notas, la interfaz *INoteListener* que recibe notificaciones al:

- Comenzar una nota
- Detectar una nota
- Terminar una nota

Toda clase que desee recibir los eventos producidos por el módulo debe implementar esta interfaz. Por último en el paquete encontramos una clase *Note* que se utiliza como representación de las notas musicales. Este enumerado contiene la información del índice de la conversión de notas, la nota musical en formato anglosajón, y la octava a la que pertenece la nota.

TarsosDSP

TarsosDSPNoteDetectionModule extiende *AnoteDetectionModule* e implementa *IPitchListener*. Simplemente convierte los tonos detectados a notas musicales, filtra las notas desafinadas y controla los inicios y finalizaciones de las notas, generando eventos de los cuales serán notificados los listeners.

5.4.3.4. Módulo Clap Detection

El módulo ClapDetection(figura 5.10) permite detectar ruidos percusivos tales como palmadas. Igual que el PitchDetection, depende directamente del módulo SoundDispatcher para su funcionamiento. En su interior se encuentra la interfaz del módulo, *IClapDetectionModule*, la clase abstracta *AClapDetectionModule* y la interfaz del listener, *IClapListener*. Las clases que implementen este listener recibirán una notificación cada vez que se detecte una palmada.

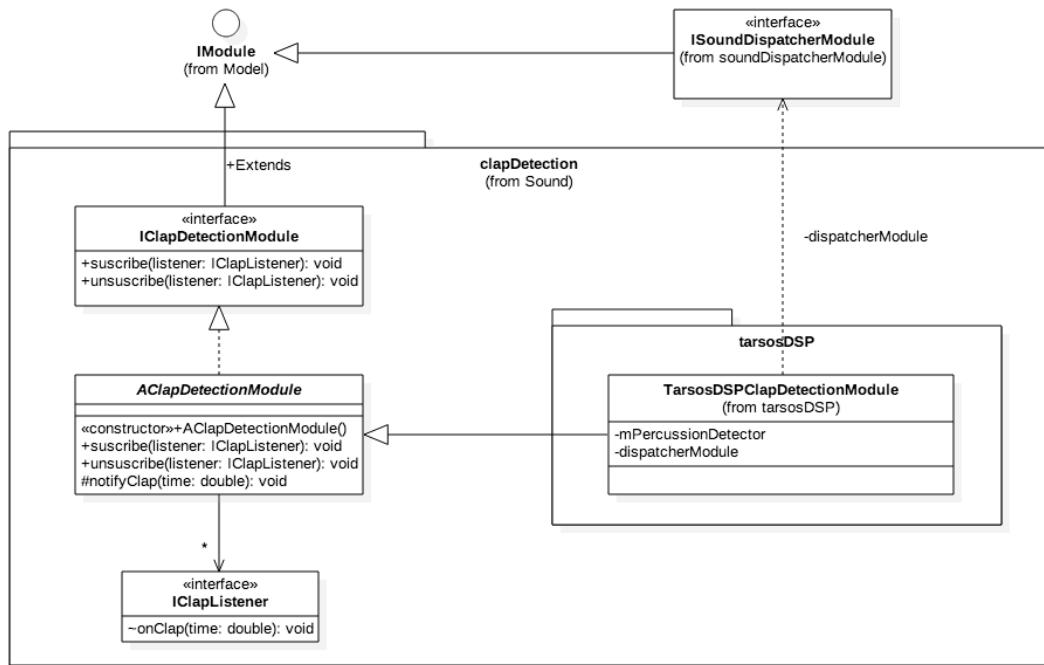


Figura 5.10: Módulo ClapDetection

TarsosDSP

La implementación del módulo, contenida en `TarsosDSPClapDetectionModule`, emplea el objeto provisto por la librería TarsosDSP `PercussionOnsetDetector`, que es registrado en el `DispatcherModule` activo en ese momento. Este objeto permite la detección de sonidos percusivos, como palmadas de una forma sencilla.

5.4.3.5. Módulo NoteGenerator

El módulo NoteGenerator (figura 5.11) provee una forma de sintetizar tonos musicales y hacer sonar secuencias de notas. En su interior se encuentra la interfaz del módulo `INoteGeneratorModule` que define las funcionalidades del módulo, la clase abstracta `ANoteGeneratorModule` que gestiona la suscripción a los listeners. La interfaz del listener se encuentra en `INotePlayListener` y su finalidad es notificar cuando una nota o una secuencia de notas termina de reproducirse. La clase `Note` es la representación de las notas musicales en el sistema, contiene la nota, el índice y la octava.

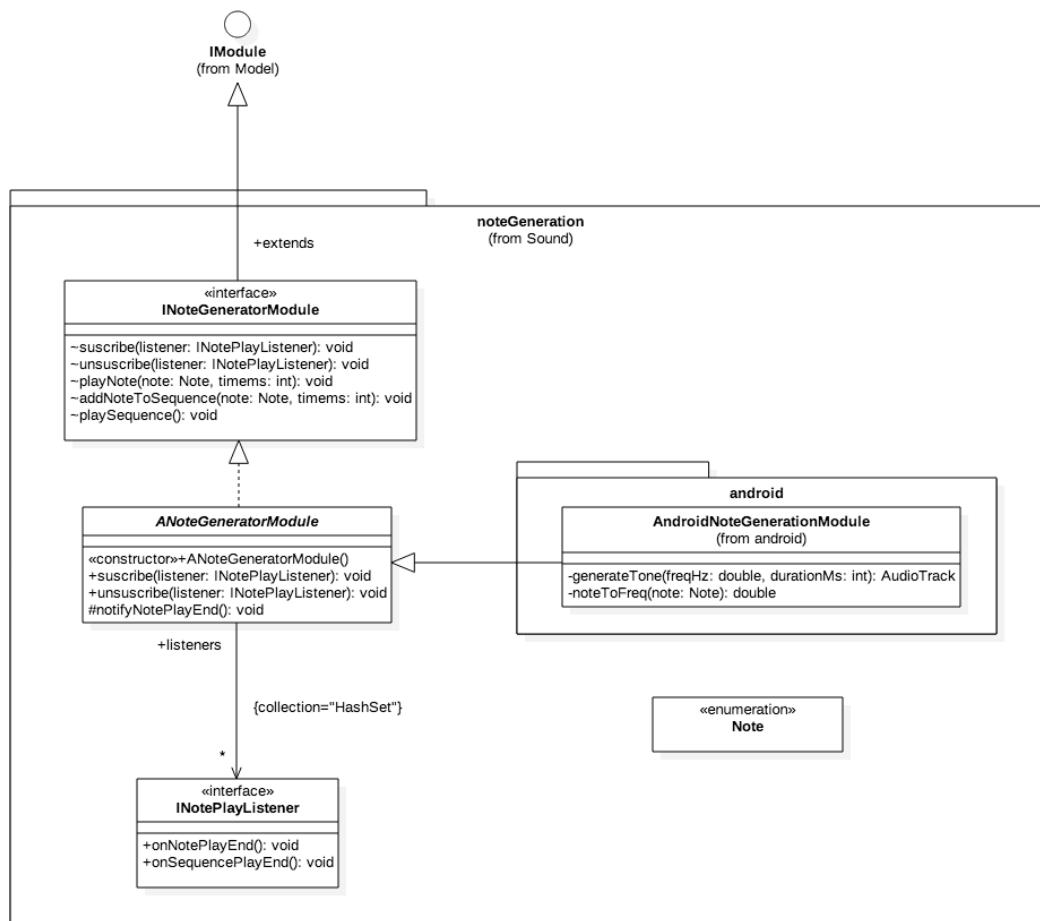


Figura 5.11: Módulo NoteGenerator

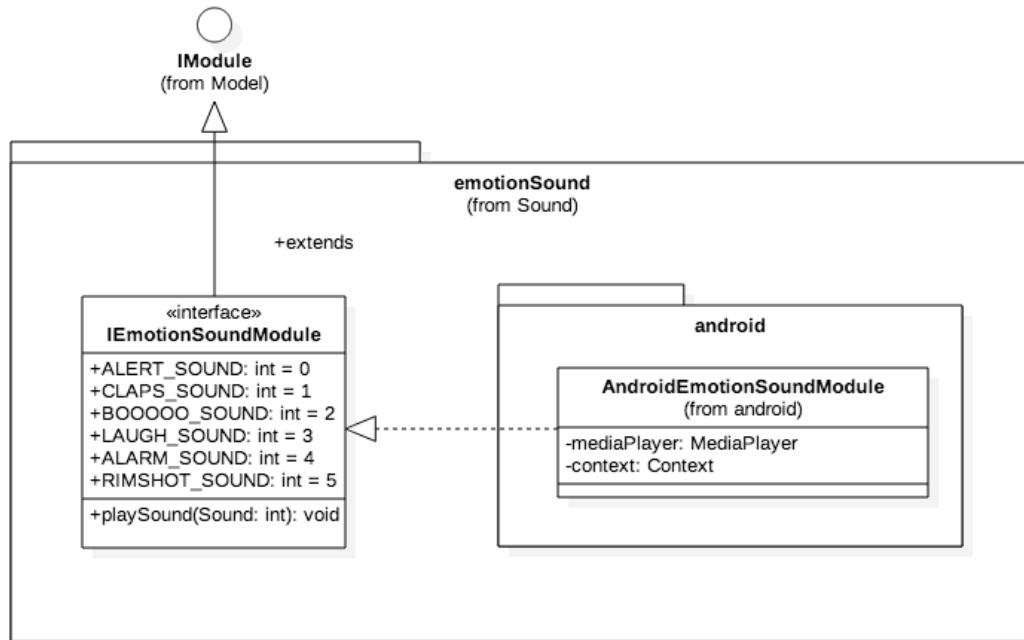


Figura 5.12: Módulo EmotionSound

Android

La implementación específica del módulo se encuentra en `AndroidNoteGeneratorModule` y se ha realizado mediante la clase `AudioTrack`. El objeto `AudioTrack` se crea sample a sample utilizando la ecuación de una onda senoidal de la frecuencia de la nota requerida. La reproducción de secuencias se implementa mediante el uso de las clases `Timer` y `TimerTask`

5.4.3.6. Módulo EmotionSound

El módulo EmotionSound (figura 5.12) busca proporcionar una manera de reproducir una serie de sonidos prefijados para expresar diferentes emociones. La interfaz del módulo está definida en `IEmotionSoundModule`, y en ella están definidos, en forma de constantes, los sonidos utilizables. Los sonidos se encuentran en el directorio raw de la carpeta res del paquete Sound.

Android

La implementación de este módulo se ha realizado mediante la clase `MediaPlayer` provista por la API de Android, que permite reproducir diferentes recursos

audiovisuales. El módulo está definido en la clase *AndroidEmotionSoundModule*.

Ejemplos de desarrollo

Para la prueba de los módulos de sonido se realizaron dos ejemplos básicos, una aplicación simple en la que el usuario introduce un texto y el teléfono lo pronuncia y un sistema de control mediante voz en el que el usuario le daba ordenes básicas al Robobo mediante la voz y este, al terminar la frase, ejecutaba los movimientos indicados.

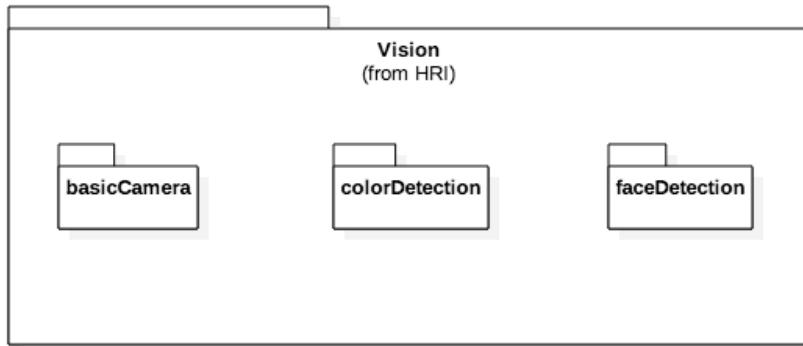


Figura 5.13: Paquete Vision

5.4.4. Paquete Vision

Este subsistema(figura 5.13) contiene los diferentes módulos de captura y procesado de imagen

5.4.4.1. Módulo Basic Camera

Este es el módulo básico de cámara(figura 5.14) , del cual hacen uso el resto de módulos de procesado de imagen. Este modulo produce un stream constante de imágenes, que no deben ser necesariamente mostradas, capturadas de la cámara frontal del dispositivo y notifica a los listeners suscritos. La interfaz del módulo se encuentra en `ICameraModule`, la clase abstracta que gestiona los listeners en `ACameraModule` y la interfaz del listener en el que se notifica la captura de las imágenes en `ICameraListener`. Además, el modulo proporciona una clase `Frame`, que representa a las imágenes capturadas con sus características.

Android Camera2

Para realizar la implementación del módulo se ha empleado la clase `Camera2` que proporciona Android. Esta implementación permite obtener el stream mencionado anteriormente sin mostrar las imágenes en la pantalla, pero produce una velocidad de captura de imágenes baja, de alrededor de dos cuadros por segundo en el smartphone de pruebas (BQ Aquaris M5) variando de teléfono a teléfono.

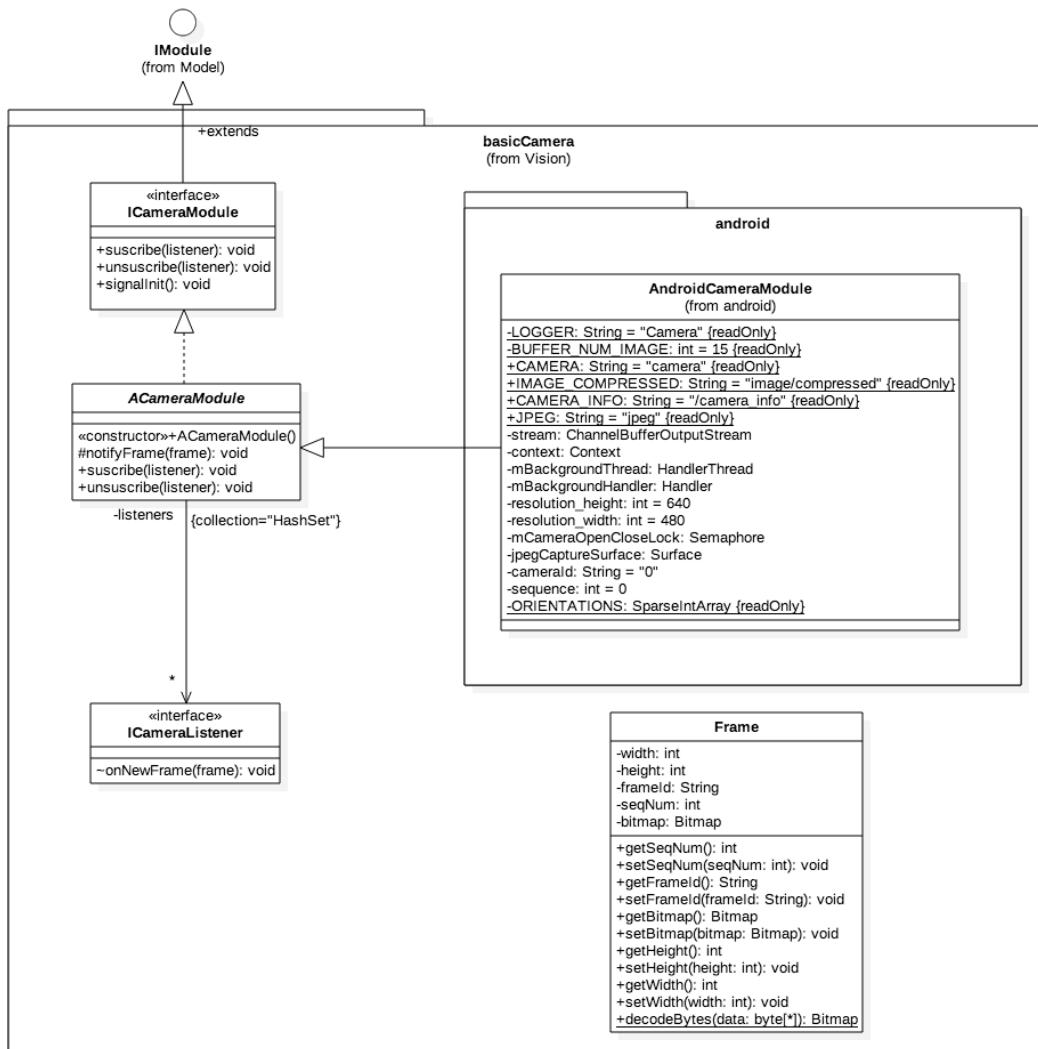


Figura 5.14: Módulo BasicCamera

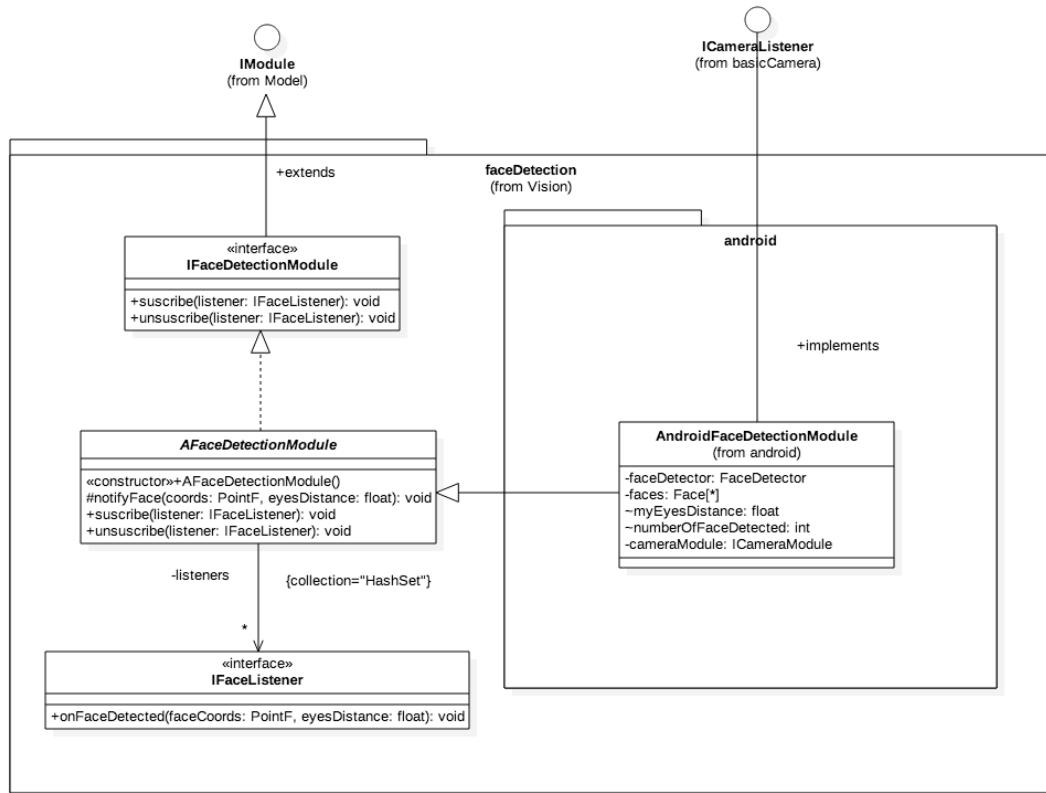


Figura 5.15: Módulo FaceDetection

5.4.4.2. Módulo Face Detection

El módulo FaceDetection (figura 5.15) permite detectar caras en los frames producidos por el módulo BasicCamera. La interfaz del módulo es *IFaceDetectionModule* y su clase abstracta *AFaceDetectionModule*. El listener que debe ser implementado por la clase que utilice el módulo es *IFaceListener*, que notifica de las coordenadas del centro de la cara y de la distancia entre ojos cuando una cara es detectada. Solo se contempla la detección de una cara al mismo tiempo.

Android FaceDetector

Para implementar este módulo se ha empleado la clase *FaceDetector* provista por el paquete Media de Android, el módulo se llama *AndroidFaceDetectionModule*.

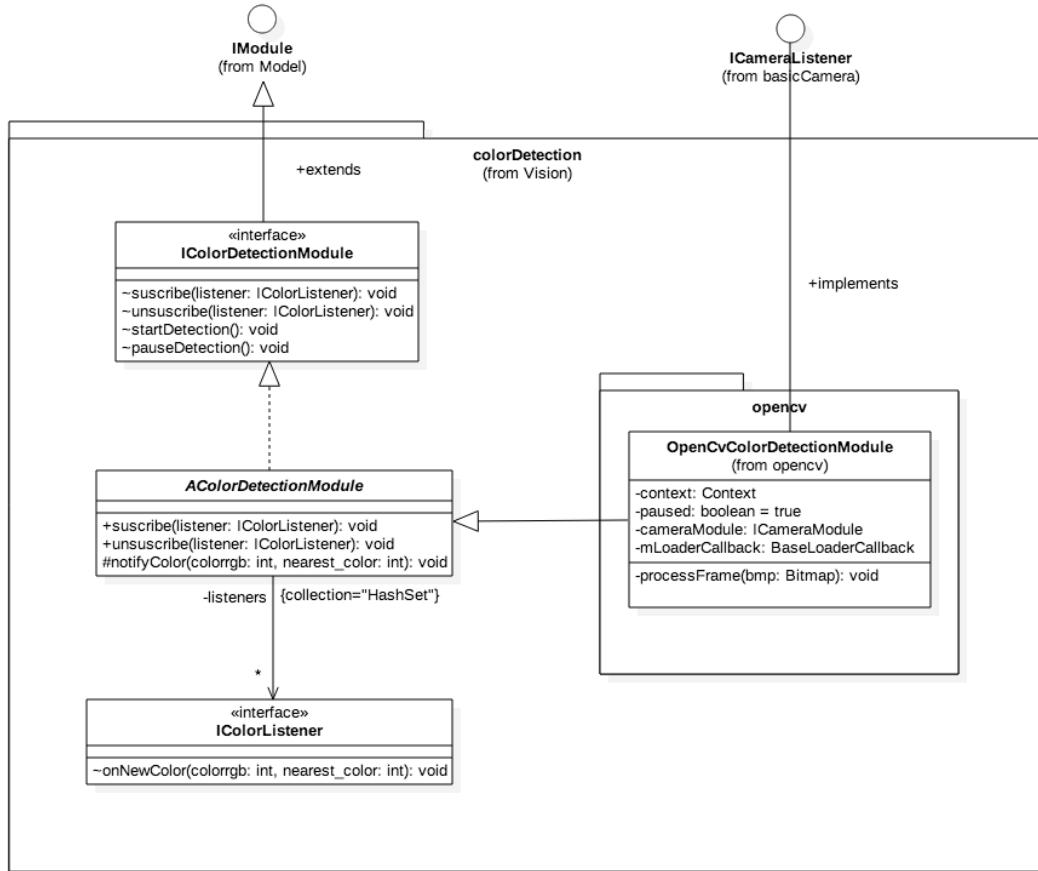


Figura 5.16: Módulo ColorDetection

5.4.4.3. Módulo ColorDetector

Este módulo(figura 5.16) provee la funcionalidad de detección de colores sobre fondos con alto contraste. La interfaz del módulo se puede encontrar en *IColorDetectionModule*, la clase abstracta *AColorDetectionModule* gestiona los listeners en los que será notificada la detección de los colores. La interfaz de dicho listener se encuentra en la clase *IColorListener* y debe ser implementada por toda clase que desee recibir las notificaciones de los colores.

5.4.4.3.1. OpenCV Para realizar la implementación de este módulo se ha empleado la librería OpenCV [?] una de las más usadas en visión por computador. Este módulo, contenido en *OpenCvColorDetectionModule* realiza los siguientes pasos para la detección de colores:

- Detección de bordes mediante el algoritmo de Canny, obteniendo el contorno con mayor area.
- Creación de una máscara binaria con el área detectada
- Conversión de la imagen original al espacio de color HSV
- Media de color en los pixeles no nulos en el canal H de la imagen resultante de un AND entre la máscara y la imagen convertida
- Filtrado por la varianza del color, si supera 1.1 es descartado
- Clasificación del color detectado

5.4.5. Paquete Messaging

El subsistema Messaging aglomera las diferentes opciones de comunicación por mensajería de texto.

5.4.5.1. Módulo email

El módulo Email(figura 5.17) permite al usuario la comunicación mediante correo electrónico, pudiendo mandar tanto texto como imágenes, por ejemplo, las capturadas con el módulo *BasicCamera*.

GmailBackground

Para implementar el módulo, se ha empleado la librería GmailBackGround [17], que permite el envío de mensajes de correo electrónico con una cuenta de Gmail. El módulo se encuentra en la clase *GmailBackgroundMessagingModule*.

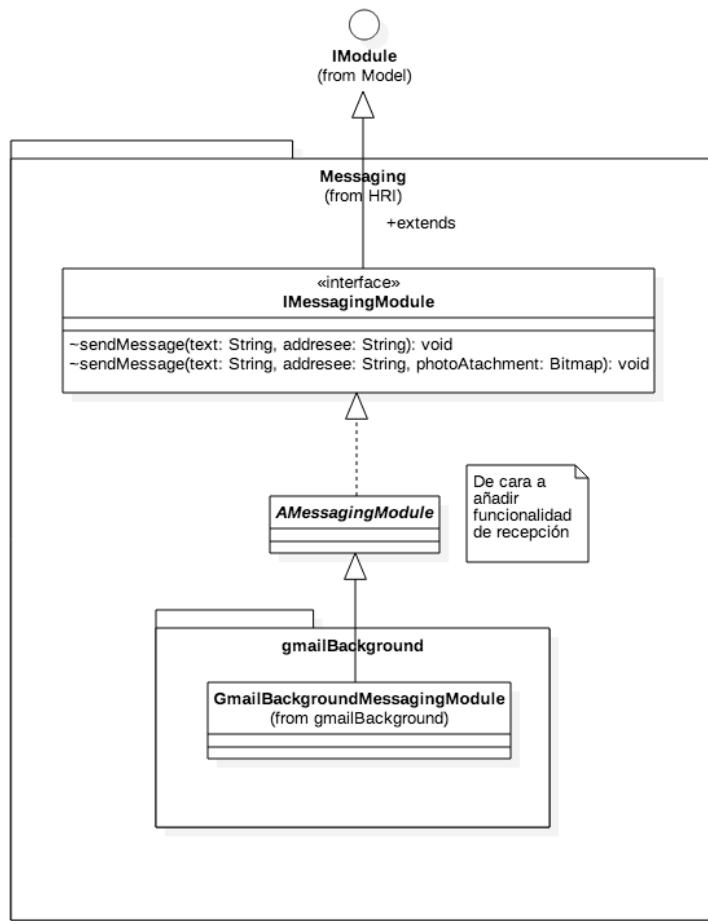


Figura 5.17: Módulo Email

Capítulo 6

Resultados y pruebas

En este capítulo se expondrán los resultados y el funcionamiento de los módulos desarrollados en este trabajo mediante tres aplicaciones de ejemplo.

6.1. Ejemplos de uso

Para probar el correcto funcionamiento de los módulos de interacción, a parte de los ejemplos de desarrollo que se comentaron en el capítulo 5.4, se desarrollaron tres aplicaciones Android que utilizan en conjunto los diferentes módulos de interacción.

6.1.1. Simon dice musical

El primero de los ejemplos desarrollados es un juego educativo musical, similar al juego Simón, en el que los participantes deben repetir una secuencia de colores que se va ampliando en cada iteración de la partida, pero utilizando notas musicales. El objetivo final de este juego es que el usuario consiga asociar los tonos escuchados con la nota musical correspondiente, además de ejercitarse la memoria. En este ejemplo se han utilizado los siguientes módulos:

- SoundDispatcherModule: Necesario para el funcionamiento de los módulos de sonido
- PitchDetectionModule: Usado indirectamente por el NoteDetectionModule

- NoteDetectionModule: Utilizado a la hora de detectar las notas musicales producidas por el usuario.
- ClapDetectionModule: Se utiliza para iniciar el juego, una primera palmada prepara al robot y la segunda inicia el juego.
- EmotionSoundModule: Diferentes sonidos son reproducidos en función del resultado del juego.
- NoteGeneratorModule: Se emplea para generar los diferentes tonos que debe reconocer el jugador.
- VoiceProductionModule: Se usa para dar información al usuario y dar una mayor sensación de interacción con el robot

Además de los módulos creados en este trabajo, se emplean varios módulos que provee el ROBOBO! Framework:

- EmotionModule: Provee la cara del robot, con la posibilidad de cambiar sus expresiones, es la interfaz gráfica que se le muestra al usuario.
- DefaultMovementModule: Permite el uso de la plataforma robótica de forma simple, permitiendo el movimiento del robot.
- IRobInterfaceModule: Provee un método para obtener la clase IRob, que se emplea para un control más avanzado del Rob, por ejemplo, para utilizar los leds de la base.

El jugador inicia la partida dando una palmada, momento en el que el robot pronunciará a través del modulo de voz «Ready», una segunda palmada comenzará la partida en si, el robot dirá una frase retando al usuario, por ejemplo «follow me if you can», y procederá a reproducir un tono musical. En el momento en el que termine el tono, cambiará la cara dando a entender que está escuchando al usuario, en el momento en el que el jugador termine de producir la nota se comprueba si corresponde al tono original, si coincide el robot realizará una pequeña celebración y volverá a producir tonos, añadiendo uno más a los reproducidos anteriormente. En caso de que el jugador falle, el robot pondrá cara de

Pentatonic A Minor Scale



Figura 6.1: Escala pentatonica menor de La

disgusto, producirá un sonido de abucheo y volverá al estado inicial, esperando a una palmada para iniciar el juego.

En esta implementación del juego, para que tenga una sonoridad agradable, solamente se producen tonos dentro de la escala pentatónica menor de La (figura 6.1)

6.1.2. Robobo Vigilante

6.1.3. Robobo Mascota

6.2. Problemas conocidos

- La tasa de refresco del módulo *BasicCamera* es baja y varía mucho entre terminales móviles.
- El *ColorDetectionModule* puede confundirse si el fondo no es homogéneo, se recomienda usar tarjetas con colores sobre fondo blanco.
- El *EmailModule* puede causar el bloqueo de la cuenta de Gmail si esta no es configurada previamente para usar mediante IMAP.
- El *TouchModule* requiere el paso explícito de los TouchEvents de la actividad en pantalla.

Capítulo 7

Conclusiones

7.1. Trabajo Futuro

Apéndice A

Manual de uso

Bibliografía

- [1] “D.r.e.a.m. project,” <http://www.robotsthatdream.eu/>, accessed: 2016-15-6.
- [2] “Robot operating system (ros),” <http://www.ros.org/>, accessed: 2016-08-14.
- [3] M. A. Goodrich and A. C. Schultz, “Human-robot interaction: a survey,” *Foundations and trends in human-computer interaction*, vol. 1, no. 3, pp. 203–275, 2007.
- [4] “Worker at volkswagen plant killed in robot accident,” <http://www.ft.com/cms/s/0/0c8034a6-200f-11e5-aa5a-398b2169cf79.html>, accessed: 2016-08-19.
- [5] J. A. Corrales Ramón, G. J. García Gómez, F. Torres Medina, and V. Perdereau, “Cooperative tasks between humans and robots in industrial environments,” 2012.
- [6] “Definition of service robots,” <http://www.ifr.org/service-robots/>, accessed: 2016-08-19.
- [7] “Paro robots,” <http://www.parorobots.com/index.asp>, accessed: 2016-08-19.
- [8] “Nao robot, aldebaran robotics,” <https://www.aldebaran.com/en/cool-robots/nao/find-out-more-about-nao>, accessed: 2016-08-19.
- [9] “Sony aibo — the history of the robotic dog,” <http://www.sony-aibo.com/>, accessed: 2016-08-21.
- [10] “Scratch language project main page,” <https://scratch.mit.edu/>, accessed: 2016-08-20.
- [11] “Lenguaje gráfico bitbloq,” accessed: 2016-08-28.
- [12] “Robobo framework,” <https://bitbucket.org/mytechia/robobo-framework>, accessed: 2016-08-13.

- [13] J. Six, O. Cornelis, and M. Leman, “TarsosDSP, a Real-Time Audio Processing Framework in Java,” in *Proceedings of the 53rd AES Conference (AES 53rd)*, 2014.
- [14] “Cmu sphinx open source speech recognition kit,” <http://cmusphinx.sourceforge.net/>, accessed: 2016-08-6.
- [15] “Jspeech grammar format,” <https://www.w3.org/TR/jsgf/>, accessed: 2016-08-6.
- [16] A. de Cheveigné and H. Kawahara, “Yin, a fundamental frequency estimator for speech and music,” *J Acoust Soc Am*, vol. 111, pp. 1917–1930, 2002. [Online]. Available: http://audition.ens.fr/adc/pdf/2002_JASA_YIN.pdf
- [17] “Gmail background library,” <https://github.com/yesidlazaro/GmailBackground>, accessed: 2016-08-4.
- [18] “Open source computer vision library,” <https://github.com/opencv>, 2015.