

## Percibiendo el entorno en los robots sociales del RoboticsLab

Fernando Alonso-Martín University Carlos III of Madrid famartin@ing.uc3m.es	José Carlos Castillo University Carlos III of Madrid jocastil@ing.uc3m.es	Álvaro Castro-González University Carlos III of Madrid acgonzal@ing.uc3m.es	Juan José Gamboa University Carlos III of Madrid jgamboa@ing.uc3m.es
Marcos Maroto University Carlos III of Madrid marmarot@pa.uc3m.es	Sara Marqués Villaroya University Carlos III of Madrid smarques@ing.uc3m.es	Antonio J. Pérez Vidal University Carlos III of Madrid 100355616@alumnos.uc3m.es	Miguel Ángel Salichs University Carlos III of Madrid salichs@ing.uc3m.es

### Resumen

La percepción del entorno es un aspecto clave para cualquier robot que tenga que interactuar con el mundo, objetos y personas de su entorno. Por este motivo, el campo de la percepción sensorial tiene una gran importancia. Este artículo describe los mecanismos empleados para dotar a los robots sociales del RoboticsLab, de la Universidad Carlos III de Madrid, de la capacidad de percibir el entorno. Desarrollos recientes en tecnologías sensoriales y algoritmos han permitido en los últimos años alcanzar un nivel de percepción cercano al de los humanos en algunas situaciones. La capacidad de percibir el entorno comprende áreas de investigación tan diversas como: la detección y reconocimiento de objetos, lugares, y personas; el procesamiento del lenguaje natural; la detección, localización, y reconocimiento de los toques producidos por los usuarios con el robot; el análisis del nivel de “enganche” en la interacción del usuario con el robot, etc. En este artículo incluimos una descripción de los tipos de sensores equipados en los robots, los módulos software desarrollados, y el gestor de la percepción que logra empaquetar y unificar temporalmente la información suministrada por cada uno de estos en diferentes niveles de abstracción de la información percibida.

**Keywords**—Robótica social, percepción automática, gestor de la percepción, interacción humano-robot

### 1. Introducción

El éxito de la Interacción Humano-Robot (HRI, por sus siglas en inglés) depende en gran parte de la capacidad que tengan los robots sociales de percibir el entorno que les rodea. Es por ello, el interés en desarrollar y mejorar las técnicas que permiten llevarla a cabo. Es en este ámbito de la percepción donde se están produciendo numerosos avances que han permitido, entre otras cosas, conseguir vehículos con auto-pilotaaje, automatizar completamente cadenas de producción, etc. Sin embargo, su aplicación en robótica social resulta complejo, debido a la enorme cantidad de campos de investigación que involucra, tales como: reconocimiento

de voz, procesamiento del lenguaje natural, visión artificial, balizas electrónicas, y un largo etcétera.

El trabajo en este campo se centra en desarrollar nuevos sensores, nuevos algoritmos que usen estos nuevos sensores o los ya existentes, y nuevas arquitecturas de fusión de la información multisensorial (también conocida información multimodal). La amalgama de sensores que se pueden aplicar en este campo es extensa: cámaras de color, de temperatura, de profundidad; micrófonos de aire o de contacto; sensores táctiles como son los capacitivos, resistivos, piezoeléctricos; acelerómetros, giróscopos, sensores de temperatura y luminosidad; etc. Esta diversidad de sensores y métodos, unido al creciente interés en su desarrollo para su aplicación en todo tipo de dispositivos electrónicos y por supuesto en robots, convierte a la percepción automática en campo de investigación en constante crecimiento.

En este artículo se expone de manera concisa los mecanismos métodos que se aplican en los robots sociales del RoboticsLab de la Universidad Carlos III de Madrid para dotarles de la capacidad de percibir el entorno que les rodea. Reseñar que los métodos empleados están en consonancia con el estado del arte en esta área. El artículo continúa con un recorrido por la literatura relacionada en la sección 2. En la sección 3 se describen los robots del grupo, los sensores, los métodos para procesar la información sensorial obtenida, y la arquitectura de fusión empleados en nuestros robots sociales. Finalmente, en la sección 4 se presentan las conclusiones del trabajo desarrollado.

### 2. Trabajos relacionados

Los sistemas de percepción en robótica social tienen gran importancia dado que las capacidades de interacción y el conocimiento del entorno del robot dependen en gran medida de ellos. Existen numerosos trabajos que combinan información de uno o más sensores para conseguir este objetivo. Entre los robots con una percepción sencilla encontramos el robot Aibo [1] integra una serie de botones en el cuerpo que utilizaba para conocer la presencia de usuarios o el robot RHINO [2], que

también interactúa con las personas a través de una serie de botones.

Existen otros ejemplos de robots sociales que integran sistemas de percepción más sofisticados capaces de detectar automáticamente los usuarios y objetos a su alrededor. En esta línea, el robot Kismet [3] es capaz de detectar la cara y los ojos de las personas frente a él. Además, este robot está equipado con un sistema de reconocimiento de voz. El robot ALIAS [4] posee capacidades similares al anterior, detectando la cara de los usuarios y realizando análisis de voz. Adicionalmente, este robot incorpora una tableta para HRI.

En el caso de los robots presentados en el proyecto MOnarCH [5], estos disponen de mecanismos de percepción tanto para navegación por el entorno como para interacción. En el primer caso, los robots utilizan información de sensores láser y odometría para desplazarse. La percepción para interacción se realiza a través de sensores de tacto colocados en el robot, una pantalla táctil, análisis de voz y cámaras externas.

Aparte de realizar simplemente detección de los usuarios alrededor del robot, hay sistemas de percepción que permiten el reconocimiento y seguimiento de las personas. Este es el caso de plataformas como Valerie [6], un robot recepcionista que utiliza tarjetas magnéticas para reconocer a los usuarios. En el caso del robot JiboJibo<sup>1</sup>, se realiza un reconocimiento de rostros para diferenciar a los usuarios, así como análisis de voz.

En Springer Handbook of Robotics'16 [7] se presenta un extenso meta-análisis del estado del arte en cuanto a métodos de percepción en robots. En él, se describen y analizan los siguientes mecanismos de percepción: sensores para percibir el tacto, sensores de odometría y posicionamiento, cámaras de profundidad para navegación, manipulación, reconocimiento de objetos. Finalmente, se dedica un capítulo a los mecanismos de fusión de la información sensorial percibida en instantes temporales, tanto métodos, como arquitecturas, y aplicaciones de las mismas. En cuanto a métodos para realizar la fusión de la información sensorial se propone el uso de teoría bayesiana, mallas probabilísticas, filtro de Kalman, métodos basados en Monte Carlos. Por otro lado, se presentan varias arquitecturas de fusión como: Distributed Field Robot Architecture (DFRA) [8], Sensor Fusion Effects (SFX) [9], SEPIA [10], MESSIE (Multi Expert System for Scene Interpretation and Evaluation) [11], Active Sensor Network (ASN) [12].

<sup>1</sup>Sitio web de Jibo: <https://www.jibo.com>

### 3. Nuestro sistema de percepción

#### 3.1. Los robots sociales

El grupo de robots sociales de la Universidad Carlos III de Madrid ha desarrollado tres plataformas robóticas de investigación en Interacción Humano-Robot (HRI). El primero de ellos, el robot Maggie [13], es el más antiguo de los tres y su principal objetivo ha sido interactuar con niños de entre 4 a 14 años. El robot Mini [14] sirve como plataforma de investigación en interacciones con ancianos con problemas cognitivos. Por último el robot Mbot [15], desarrollado durante el proyecto MOnarCH<sup>2</sup>, se ha destinado a la interacción con niños en la planta pediátrica del hospital oncológico de Lisboa (ver figura 1).



Figura 1: Los robots sociales del RoboticsLab. Arriba, robot Maggie. Centro, robot Mini. Abajo, robot Mbot.

<sup>2</sup>Sitio web del proyecto MOnarCH: <http://monarch-fp7.eu/>

### 3.2. Sensores empleados

Los robots sociales del RoboticsLab están dotados de numerosos sensores que son empleados por los módulos de percepción para “conocer” el entorno que les rodea. A continuación se presenta una enumeración y breve descripción del conjunto de sensores empleados:

Los **micrófonos “de aire”** son los más extensamente usados en robótica y son empleados tradicionalmente para tareas fundamentalmente de procesamiento del lenguaje natural (NLP, por sus siglas en inglés). Por tanto, los principales usos son el reconocimiento de voz, identificación del hablante, identificación de la emoción del hablante. Adicionalmente, se pueden emplear micrófonos para tareas de localización de la fuente de sonido y para analizar el nivel de excitación sonora (*arousal*) del entorno. El emplazamiento y la elección del modelo concreto de micrófono debe ser analizado en profundidad para cada caso. En este sentido, existen multitud de tipos de micrófonos en función del empleo que se quiera hacer del mismo. En el caso de aplicarse fundamentalmente para tareas de reconocimiento automático del habla conviene el uso de micrófonos especialmente diseñados para esta tarea. Además, se deberán integrar en el robot en la posición más ventajosa para evitar, en la medida de lo posible, los ruidos producidos por el propio robot (altavoces, motores, etc.) y al mismo tiempo captar la voz de los posibles usuarios con claridad.

Los **micrófonos “de contacto”** en nuestro caso son empleados para dotar a los robots de “sentido del tacto”, es decir para detectar y reconocer que tipo de contactos produce el usuario sobre la superficie del robot. Estos micrófonos son muy sensibles a captar cualquier vibración sonora propagada por un cuerpo sólido. Aprovechando esta propiedad es posible analizar el tipo de contacto que se produce cuando los usuarios tocan al robot. Este análisis, usando técnicas de aprendizaje automático, permite clasificar el tipo de contacto realizado, así como localizar en que zona del robot se ha producido (ver figura 2).

Los **sensores táctiles capacitivos** proporcionan información menos rica que los micrófonos de contacto en el sentido de que su señal es normalmente binaria (toque vs no toque). Aún así, por su sencillez y bajo coste, se integra esta tecnología en nuestros robots para dotar a los robots de la posibilidad de detectar toques donde los micrófonos de contacto no son factibles (p. ej. en partes del cuerpo del robot con fuentes de ruido próximas como motores).

Las **cámaras** son empleadas por los algoritmos de



Figura 2: Principales sensores de nuestros robots: Arriba, izquierda, micrófono de contacto usado como sensor de tacto; Arriba, derecha, cámara Creative Sens3D; En el medio, izquierda, LeapMotion; En el medio, derecha, telémetro láser Hokuyo. Abajo, las balizas electrónicas.

visión artificial en tareas como el reconocimiento de personas, objetos, lugares, emociones, y poses. En nuestros robots se integran cámaras RGB tradicionales y cámaras de profundidad como la *Kinect* (I y II), *Creative Sens3D* (ver figura 2), y *Softkinetic DepthSense*. Además de las cámaras incorporadas en los propios robots, suele ser habitual el uso de cámaras posicionadas directamente en el entorno de trabajo del robot, en nuestro caso empleamos cámaras omnidireccionales tanto en el laboratorio como en las áreas donde se suelen mover los robots.

El dispositivo **LeapMotion**<sup>3</sup> (ver figura 2) es un tipo especial de cámara empleada para interacciones a menos de 1.5 metros. Normalmente, y en nuestro caso también, es empleado para detectar los gestos realizados por los usuarios con las manos. Basándonos en su uso hemos desarrollado el tradicional juego de “piedra, papel, o tijera” entre los usuarios y nuestros robots.

El uso del **magnetómetro** o **brújula** es llevado a cabo para detectar variaciones en el campo magnético. Esto puede permitir al robot localizarse en el entorno. Esta tecnología hace uso de lo que se conoce como “huella magnética del edificio”. Los **acelerómetros** y **giróscopos** permite detectar si

<sup>3</sup>[www.leapmotion.com](http://www.leapmotion.com)

el robot ha sido levantado o si ha sufrido una caída no programada.

Los **telémetros** como el láser, sonar, o infrarrojos son empleados para medir distancias del robot respecto a los objetos y/o personas que le rodean (ver figura 2).

Las **balizas electrónicas** proporcionan información acerca de los usuarios alrededor del robot (ver figura 2). Estos elementos no son capaces de dar una posición exacta, sino que aportan una estimación de la distancia de la baliza al receptor, siendo posible establecer la posición de los usuarios dentro de zonas concéntricas. Asimismo, las balizas permiten la identificación de los usuarios [16].

Al igual que los sensores de tacto, las **pantallas táctiles** proporcionan información acerca de la interacción de los usuarios con una pantalla, mitigando los problemas de reconocimiento asociados a los sistemas NLP a la hora de proporcionar información. Además, estos dispositivos proporcionan un soporte gráfico que aumentan las posibilidades de interacción del robot, pudiendo mostrar contenido multimedia, información variada o realizar juegos con el usuario.

Podrían emplearse sensores adicionales como barómetros, cámaras de temperatura, de huella dactilar, de luz, etc. Sin embargo, para los módulos que se han desarrollado y que a continuación presentamos, no han resultado necesarios.

### 3.3. Módulos desarrollados para percibir el entorno

Los robots sociales del grupo están dotados de técnicas que combinan desarrollos software con hardware para percibir el entorno. En concreto los módulos que se han desarrollado para dicha percepción son los que enumeramos a continuación:

1. Detección y reconocimiento de personas [17]: en la literatura podemos encontrar el empleo de diferentes sensores y métodos para esta tarea, las más habituales son mediante el empleo de técnicas basadas en visión artificial, a través del análisis del tono de voz del usuario, o mediante el uso de balizas electrónicas. En nuestro caso, es precisamente mediante la combinación de estos tres modos como llevamos a cabo esta tarea.
2. Detección y reconocimiento de objetos: de igual forma que en el caso anterior, mediante el empleo de algoritmos basados en el análisis de las imágenes obtenidas por las cámaras del robot, tanto de color como de profundidad, es posible la detección y el reconocimiento de los

objetos, previamente conocidos, que rodean al robot.

3. Detección y reconocimiento de lugares: mediante técnicas, normalmente aplicadas a tareas de navegación, es posible localizar el robot en el entorno, identificando la sala en la que se encuentra.
4. Detección y reconocimiento de emociones en los usuarios [18]: mediante técnicas basadas en el análisis facial de los rostros de los usuarios, así como otras técnicas basadas en el análisis de la voz, e incluso imágenes de temperatura, es posible clasificar la emoción de los usuarios.
5. Detección y reconocimiento de poses/gestos de los usuarios: es posible reconocer gestos realizados por los usuarios como saludos, despedidas, etc. Para ello se suele emplear técnicas basadas en el análisis del esqueleto del usuario en secuencias de imágenes continuas [19]. Asimismo, dispositivos como el LeapMotion proporcionan información acerca de los movimientos de las manos a través de técnicas de clasificación, que se pueden explotar para interacción [20].
6. Evaluación del nivel de “enganche” (*engagement*) de los usuarios con el robot: es de gran interés conocer cuando los usuarios prestan atención al robot durante la interacción o por contra la interacción para ellos carece de interés. Para analizar el enganche se suelen tener en cuenta aspectos como el número de veces que el usuario mira al robot, el número de turnos de diálogo, el grado de consecución de los objetivos marcados en la interacción, etc.
7. Detección y reconocimiento de “toques” de los usuarios en el robot [21]: los robots sociales suelen estar dotados de la capacidad de detectar cuando el usuario toca, al menos, una parte determinada del robot, por ello existe multitud de sensores y algoritmos para llevar a cabo esta tarea. En recientes trabajos hemos presentado un sistema para detectar, reconocer, y localizar el tipo de toque realizado en cualquier parte del robot. Para ello, nos hemos valido de micrófonos y técnicas de análisis de audio y aprendizaje automático.
8. Procesamiento del lenguaje natural (detección de actividad de voz [22], reconocimiento de voz. [23], enriquecimiento semántico [24]): es deseable que el robot sea capaz de comprender y generar el lenguaje natural, para interactuar por voz con el usuario de manera satisfactoria.

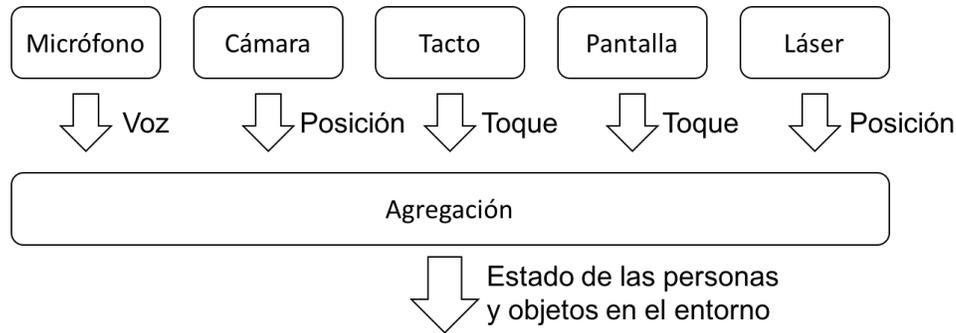


Figura 3: Módulo de agregación. La información proveniente de los módulos de procesamiento de los sensores se unifica para proporcionar información de más alto nivel acerca del estado de las personas y objetos en el entorno.

Cada uno de estos módulos emplea sensores y algoritmos que han sido descritos en profundidad en los artículos citados.

### 3.4. Agregación

El módulo de agregación se encarga de recibir los datos procesados por los módulos sensoriales y relacionarla para proporcionar la información relativa al entorno en cada momento. Recordemos que esta información es relativa a los usuarios y objetos que rodean al robot, la posición del robot, la emoción de los usuarios, etc. El objetivo de este módulo es empaquetar y entregar esta información en paquetes temporales con cierta coherencia temporal y fusionar fuentes de información complementarias. Este tipo de procesamiento involucra procesar información procedente de varios modos o módulos perceptivos simultáneamente a lo largo de un largo periodo de tiempo. La figura 3 muestra un ejemplo de la arquitectura de percepción con algunos de los módulos descritos en la sección 3.3.

La implementación de cada nivel de fusión de información se puede realizar fundamentalmente siguiendo dos posibles aproximaciones. La primera de ella consiste en establecer reglas de manera manual, normalmente mediante experimentos y pruebas de campo. La segunda aproximación podría establecer reglas automáticamente usando técnicas de aprendizaje automática e inteligencia artificial usando un conjunto de datos de entrenamiento y validación. Una aproximación híbrida entre la manual y la automática también se puede plantear como una alternativa válida.

En nuestro caso hemos desarrollado ambas aproximaciones. Reseñar que ciertos módulos sensoriales anteriormente mencionados internamente también incorporan mecanismos para fusionar información obtenida por varios sensores diferentes o por varios clasificadores. A modo de ejemplo el sistema de de-

tección de usuarios incorpora un Filtro de Kalman Extendido. El sistema de detección de emociones implementa un algoritmo de fusión bayesiano que le permite mezclar la información que suministran los algoritmos basados en visión con los basados en sonido.

## 4. Conclusiones

En este artículo hacemos una revisión completa del conjunto de técnicas empleadas para dotar a los robots sociales del RoboticsLab de ciertas capacidades sensoriales que les permiten percibir su entorno. Esta revisión incluye enumerar y describir los principales sensores empleados en los robots sociales, el conjunto de módulos perceptivos, así como el gestor de la percepción que permite mezclar la información suministrada por estos módulos perceptivos.

El objetivo final es que los robots sociales sean capaces de interactuar de manera natural con los humanos, y para ello es fundamental lograr que nuestros robots tengan capacidades sensoriales similares a las de los humanos. El conjunto de capacidades humanas, que de alguna forma se tratan de imitar, es amplio. Entre estas capacidades se encuentran las de detectar y reconocer con que humanos interactúan, que objetos les rodean, y en que lugar se encuentran. La de procesar el lenguaje natural, reconociendo y entendiendo el mensaje verbalmente transmitido por los usuarios. La de detectar y comprender los toques que realizan los usuarios sobre la superficie del robot. Detectar y clasificar las emociones de los usuarios que se generan durante la interacción, así como evaluar el grado de enganche del usuario con el robot.

El uso de estas capacidades sensoriales, inspiradas en las propias de los humanos, permite interacciones multimodales que usan varios modos/canales de interacción, como son: la voz, el tacto, el uso

de gestos, expresiones faciales, etc.

### Agradecimientos

La investigación desarrollada ha recibido financiación de dos proyectos: “Development of social robots to help seniors with cognitive impairment” (ROBSEN), financiado por el Ministerio de Economía y Competitividad; y “RoboCity2030-III-CM”, financiado por la Comunidad de Madrid y cofinanciado por los Fondos Estructurales de la Unión Europea.

### Referencias

- [1] Y. Moon, S. Dutta, and S. Oundhakar, *Sony AIBO: The world's first entertainment robot*. Harvard Business School Pub, 2005.
- [2] W. Burgard, A. Cremers, and D. Fox, “The interactive museum tour-guide robot,” in *Proceedings of the 15th national/10th conference on Artificial intelligence/Innovative applications of artificial intelligence*, pp. 11–18, 1998.
- [3] C. Breazeal and B. Scassellati, “A context-dependent attention system for a social robot,” in *International Joint Conference on Artificial Intelligence*, vol. 2, pp. 1146–1151, 1999.
- [4] J. Geiger, T. Leykauf, and T. Rehrl, “The Robot ALIAS as a Gaming Platform for Elderly Persons,” *Lebensqualität im Wandel von Demografie und Technik - 6. Deutscher AAL-Kongress mit Ausstellung*, 2013.
- [5] M. I. Aldinhas and J. Sequeira, “Designing a robotic interface for children: The monarch robot example,” in *Advances in Cooperative Robotics*, pp. 652–659, World Scientific, 2017.
- [6] R. Gockley, A. Bruce, J. Forlizzi, M. Michalowski, A. Mundell, S. Rosenthal, B. Sellner, R. Simmons, K. Snipes, A. C. Schultz, and J. Wang, “Designing robots for long-term social interaction,” in *2005 IEEE/RSJ IROS*, pp. 2199–2204, 2005.
- [7] M. Hägele, K. Nilsson, J. N. Pires, and R. Bischoff, *Industrial Robotics*. Cham: Springer International Publishing, 2016.
- [8] M. Long, A. Gage, R. Murphy, and K. Valavanis, “Application of the distributed field robot architecture to a simulated demining task,” in *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2005, pp. 3193–3200, IEEE, 2005.
- [9] R. Murphy and R. Arkin, “SFX: An Architecture For Action-oriented Sensor Fusion.,” in *IROS*, pp. 1079–1086, 1992.
- [10] L. Tolbert, H. Q. H. Qi, and F. Peng, “Scalable multi-agent system for real-time electric power management,” *2001 Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No.01CH37262)*, vol. 3, pp. 1–4, 2001.
- [11] P. Giraudon, G., Garnesson, P., y ONTESINOS, “MESÍAS : un multi-especialista en sistema de visión. ApplicationA interpretación de las imágenes aéreas,” *Procesamiento de Señales*, vol. 9, no. 5, pp. 403–419, 1992.
- [12] A. Makarenko and H. Durrant-Whyte, “Decentralized data fusion and control in active sensor networks,” in *Proceedings of the Seventh International Conference on Information Fusion*, vol. 1, pp. 479–486, 2004.
- [13] M. A. Salichs, R. Barber, A. M. Khamis, M. Malfaz, J. F. Gorostiza, R. Pacheco, R. Rivas, A. Corrales, E. Delgado, and D. Garcia, “Maggie: A robotic platform for human-robot social interaction,” in *2006 IEEE Conference on Robotics, Automation and Mechatronics*, pp. 1–7, June 2006.
- [14] Á. Castro-González, J. C. Castillo, F. Alonso-Martín, O. V. Olortegui-Ortega, V. González-Pacheco, M. Malfaz, and M. A. Salichs, “The effects of an impolite vs. a polite robot playing rock-paper-scissors,” in *International Conference on Social Robotics*, pp. 306–316, Springer, 2016.
- [15] V. González-Pacheco, Á. Castro-González, M. Malfaz, and M. A. Salichs, “Human robot interaction in the monarch project,” in *Proc. 13th Workshop Robocity2030*, pp. 1–8, 2015.
- [16] F. Alonso-Martín, A. Castro-González, M. Malfaz, J. C. Castillo, and M. A. Salichs, “Identification and distance estimation of users and objects by means of electronic beacons in social robotics,” *Expert Systems with Applications*, 2017.
- [17] F. Alonso-Martín, A. Ramey, and M. Á. Salichs, “Speaker identification using three signal voice domains during human-robot interaction,” in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction - HRI '14*, (Bielefeld (Germany)), pp. 114–115, ACM Press, mar 2014.

- [18] F. Alonso-Martin, M. Malfaz, J. Sequeira, J. Gorostiza, and M. A. Salichs, “A Multimodal Emotion Detection System during Human-Robot Interaction,” *Sensors*, vol. 13, no. 11, pp. 15549–15581, 2013.
- [19] V. Gonzalez-Pacheco, M. Malfaz, J. C. Castillo, A. Castro-Gonzalez, F. Alonso-Martín, and M. A. Salichs, “How much should a robot trust the user feedback? analyzing the impact of verbal answers in active learning,” in *International Conference on Social Robotics*, pp. 190–199, Springer, 2016.
- [20] Á. Castro-González, J. C. Castillo, F. Alonso-Martín, O. V. Olortegui-Ortega, V. González-Pacheco, M. Malfaz, and M. A. Salichs, “The effects of an impolite vs. a polite robot playing rock-paper-scissors,” in *International Conference on Social Robotics*, pp. 306–316, Springer International Publishing, 2016.
- [21] F. Alonso-martín, J. C. Castillo, J. J. Gamboa, and M. Á. Salichs, “Acoustic Sensing for Touch Recognition in a Social Robot,” in *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (ACM, ed.), (Viena), pp. 65–66, 2017.
- [22] F. Alonso-Martin, Á. Castro-González, J. Gorostiza, and M. A. Salichs, “Multidomain Voice Activity Detection during Human-Robot Interaction,” in *International Conference on Social Robotics (ICSR 2013)*, pp. 64–73, Bristol: Springer International Publishing, 2013.
- [23] F. Alonso-Martin and M. Salichs, “Integration of a voice recognition system in a social robot,” *Cybernetics and Systems*, vol. 42, pp. 215–245, may 2011.
- [24] F. Alonso-Martín, A. Castro-González, F. Luengo, and M. Salichs, “Augmented Robotics Dialog System for Enhancing Human–Robot Interaction,” *Sensors*, vol. 15, pp. 15799–15829, jul 2015.