

Interacción persona-robot: Interfaces

Juan Antonio Corrales Ramón

Departamento de Electrónica y Computación
Universidad de Santiago de Compostela

Dispositivos hardware para la interacción

- Los elementos hardware del robot definen sus capacidades de interacción a nivel de percepción y acción:
 - **Sensores visuales:** Detección y reconocimiento visual de humanos y de objetos en el entorno.
 - **Sensores auditivos:** Detección y reconocimiento de sonidos y del habla.
 - **Sensores táctiles:** Detección de contacto.
 - **Motores:** Movimiento de los miembros del robot para interaccionar físicamente o realizar gestos.
 - **Altavoces:** Generación del habla y de sonidos.
 - **Dispositivos específicos de interacción -HCI-:** integrados en el propio robot (pantallas, touchscreens, botones...) o externos (software simulador en PCs, gafas de realidad aumentada, EMGs, BCIs, dispositivos hápticos...).

Ejemplo de robot humanoide NAO

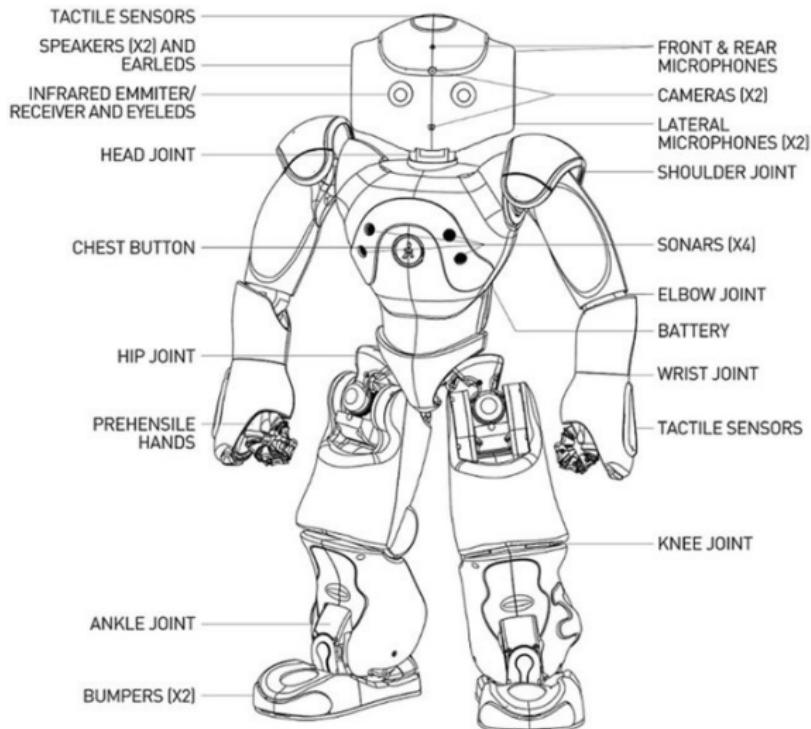


Figure 1: Componentes hardware del robot Nao

Ejemplo de manipulador móvil social Tiago

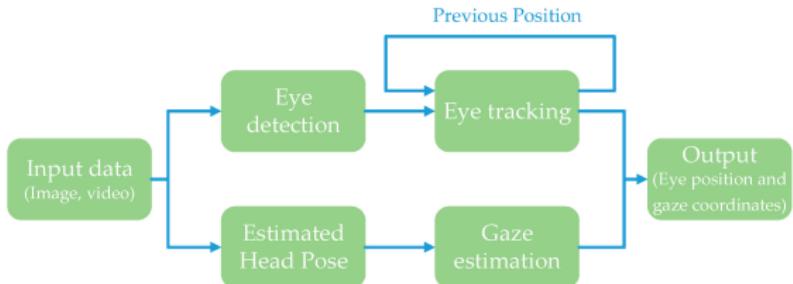


Figure 2: Componentes hardware del robot Tiago

- **Tecnologías:** cámaras RGB, cámaras RGBD, láseres, sónares...
- **Información:** imágenes, videos, mapas de distancias-profundidad...
- **Características de bajo nivel:** color, intensidad, forma, movimiento, textura...
- **Características de alto nivel:** caras, manos (gestos-gesture-), rasgos faciales (mirada-gaze), salientes visuales (salience), implicación (engagement)...
- **Algoritmos de tratamiento de las características:**
 - detección: segmentación-separación del resto de información
 - reconocimiento: identificación de una instancia concreta
 - seguimiento/tracking: evolución temporal de una característica específica
 - interpretación: comprensión de las causas de dicha evolución temporal y de sus implicaciones para la interacción

Gaze tracking: Objetivos

- La mirada proporciona información sobre: emociones, ocupación mental y movimientos oculares hacia necesidades.
- La mirada puede ser una interfaz con el robot:
 - Interfaz directa: El humano mueve de manera consciente sus ojos para enviar órdenes al robot. P. Ej. [Control de un dron](#).
 - Interfaz indirecta: El robot interpreta ciertos componentes del estado mental del humano para adaptar su comportamiento automáticamente. Se consigue una interacción más 'natural'. P. Ej. [Identificar el objeto al que el humano mira](#).
- Un 'gaze tracker' (dispositivo-algoritmo) realiza 2 tareas simultáneas: localización de los ojos y seguimiento de sus movimientos para estimar la línea 3D de la mirada.



Gaze tracking: Técnicas de localización de los ojos

- Basadas en formas:
 - Formas elípticas de los ojos
 - Formas complejas: templates deformables
- Basadas en características ('features'):
 - Features locales: borde de un ojo, área entre los dos ojos...
 - Detección del iris y las pupilas.
- Basadas en apariencia:
 - Basadas en templates con la forma y textura oculares.
 - Holísticas: Representaciones estadísticas del ojo mediante aprendizaje de un gran número de individuos.

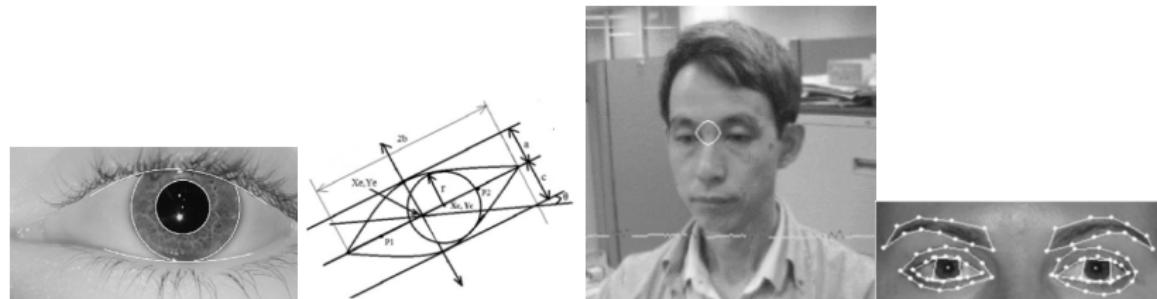


Figure 4: Forma elíptica, template deformable, área entre los ojos, template forma+textura. Fuente: DOI: 10.3390/s19245540

Gaze tracking: Técnicas de estimación de la mirada

- Basadas en modelos 3D geométricos: Estiman el centro de la córnea, los ejes visuales del ojo y estiman su intersección.
- Basadas en regresión: Mapean características visuales con coordenadas de la mirada mediante representaciones paramétricas (polinomios) o no-paramétricas (redes neurales).
- Basadas en proyección IR: Proyectan un patrón IR en el ojo y estiman la dirección de la mirada por proyección geométrica.

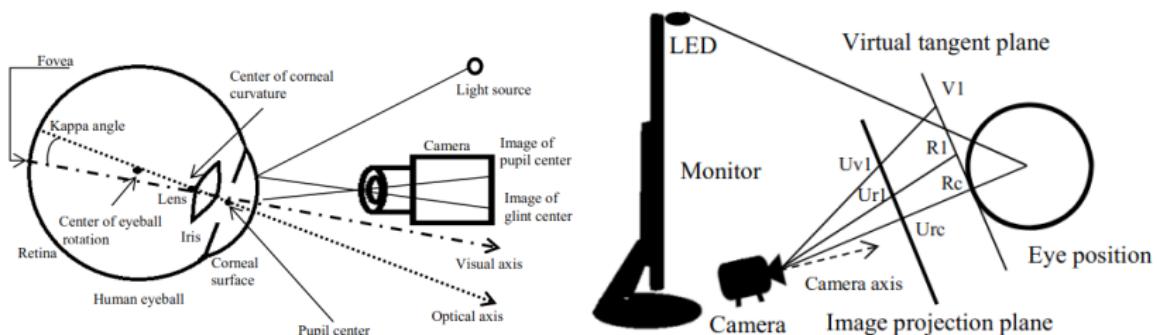


Figure 5: Modelo 3D, proyección IR. Fuente: DOI:
[DOI: 10.1109/ACCESS.2017.2735633](https://doi.org/10.1109/ACCESS.2017.2735633)

Gaze tracking: Interpretación y aplicación

Eye movement type	Movement rate	Latency/duration of occurrence	Functionality/Significance	Applications in Human Computer interaction
Fixation	< 15 – 100 deg/ms	180- 275 ms	Acquiring information, Cognitive processing, attention	Browsing information, reading, scene perception
Saccade	100- 700 deg/sec	Latency-200 ms, duration: 20–200 ms	Moving between targets	Visual search
Smooth pursuit	< 100 deg/sec based on target speed	100 ms	Following moving targets	Gaze based drawing, steering
Scanpath	--	--	Scanning, direct search	Assessing user behavior, user interface and layout quality
Gaze duration	--	--	Cognitive processing, conveying intent	Item selection, text/number entry
Blink	12 -15 per min	300 ms	Indicates behavioral states, stress	Eye liveliness detection, activate command/control
Pupil size change	4-7 mm/sec	140 ms	Cognitive effort, representing micro-emotions	Assessing cognitive workload, user fatigue, command/control

Figure 6: Posibilidades de movimientos oculares. Fuente:[DOI: 10.1109/ACCESS.2017.2735633](https://doi.org/10.1109/ACCESS.2017.2735633)

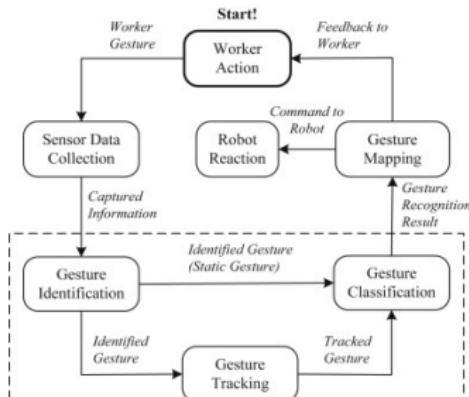
Gaze tracking: Dispositivos específicos

- Estabilizadores de cabeza: Alta precisión pero incómodos al restringir los movimientos de la cabeza.
- Trackers remotos: Compuestos por una cámara y una fuente IR. Tienen un área limitada y la luz solar crea interferencias.
- Trackers portables (wearables/head-mounted): Gafas con 2 cámaras integradas (una para la escena y otra para los ojos). Dificultades con movimientos oculares periféricos o luz solar.
- Trackers embebidos: Cascos de realidad virtual-aumentada que se usan como dispositivos de control.



Gesture Recognition: Objetivos

- Tres tipos de gestos para una interfaz H-R sin contacto:
 - Corporales: movimientos de todo el cuerpo. [Vídeo](#)
 - Con manos y brazos: señalar, ok, stop... [Vídeo](#)
 - Faciales y con la cabeza: asentir-negar, guiñar un ojo... [Vídeo](#)
- Pasos del 'gesture recognition': identificación (localización del gesto en la información sensorial), tracking (seguimiento del gesto durante el movimiento), clasificación (tipo de gesto) y mapping (asociación con un comando del robot).
- 'Emotion recognition': Asociar con estados mentales y emocionales (lenguaje corporal).



Gesture Recognition: Tecnologías

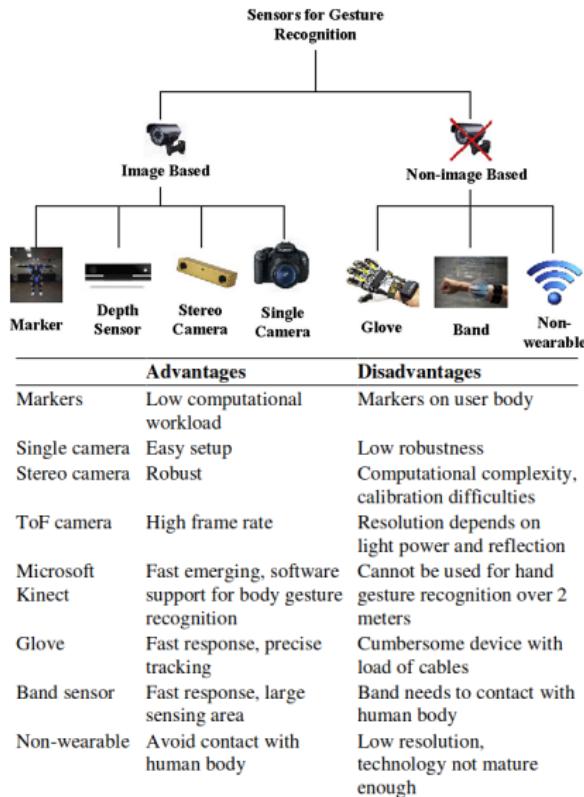


Figure 9: Tecnologías de gestos. DOI:10.1016/j.ergon.2017.02.004

Gesture Recognition: Técnicas

- Basados en modelos 3D:
 - Volumétricos: Superficies 3D (NURBS, meshes poligonales o super-cuádricas) con alto detalle y coste computacional.
 - Esqueletos: Modelo más simple con número de parámetros reducido (longitudes de segmentos + ángulos articulares).
- Basados en Apariencia:
 - Color/forma/silueta
 - Templates 2D deformables
 - Secuencias de imágenes: flujo óptico

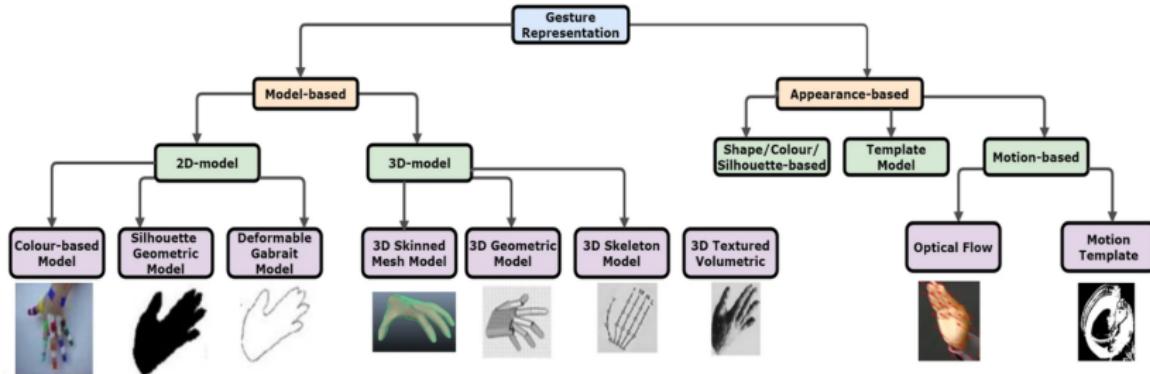


Figure 10: Algoritmos de gestos. DOI:10.1007/s42979-021-00827-x

Sensores táctiles

- **Tecnologías:** sensores de fuerza/par, matrices de sensores capacitivos/resistivos, acelerómetros-micrófonos (vibraciones), sensores con cámaras integradas para visualizar el contacto...
- **Información:** imágenes táctiles, vibraciones (texturas), temperatura...
- **Características de bajo nivel:** fuerzas normales, fuerzas tangenciales (fricción), pares de contacto, superficie de contacto...
- **Características de alto nivel:** características físicas del material (dureza), eventos de contacto (deslizamiento)...
- **Algoritmos de tratamiento de las características:**
 - control de las fuerzas de contacto y síntesis de agarres estables de objetos
 - detección y reconocimiento de deslizamiento (slippage)
 - reconocimiento de texturas
 - reconocimiento de la forma del contacto
 - control táctil (tactile servoing): pasar de un contacto inicial a un contacto final deseado

Sensores táctiles: Técnicas

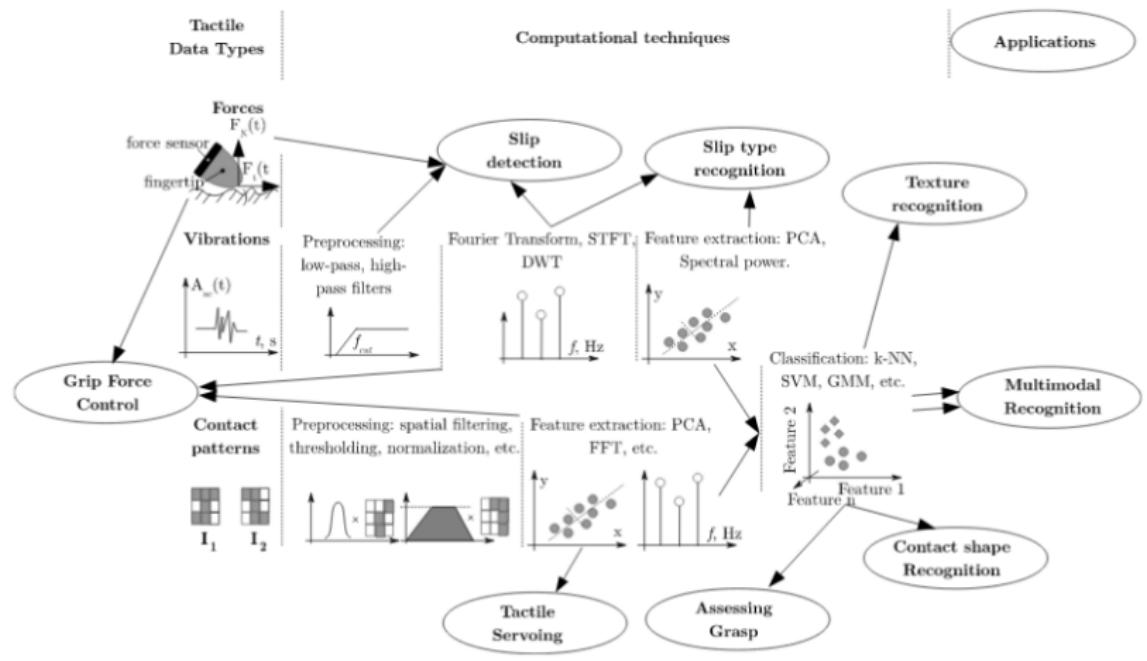


Figure 11: Algoritmos para tratar e interpretar la información táctil.
Fuente: DOI: 10.1016/j.robot.2015.07.015

Dispositivos hápticos: Principios

- Las interfaces hápticas son sistemas mecatrónicos que replican las sensaciones del tacto (interacción táctil y cinestésica -fuerza+posición-) en tareas de manipulación.
- Se pueden aplicar en entornos de simulación, para teleoperar robots ([vídeo](#)) o para mejorar la interacción real.
- El usuario desplaza el dispositivo háptico. El desplazamiento es aplicado al robot virtual/real. El contacto del robot con el entorno (fuerzas) es transmitido al dispositivo háptico (1KHz).

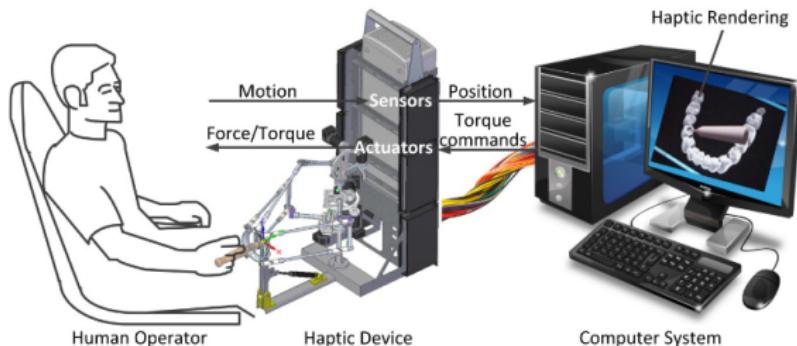
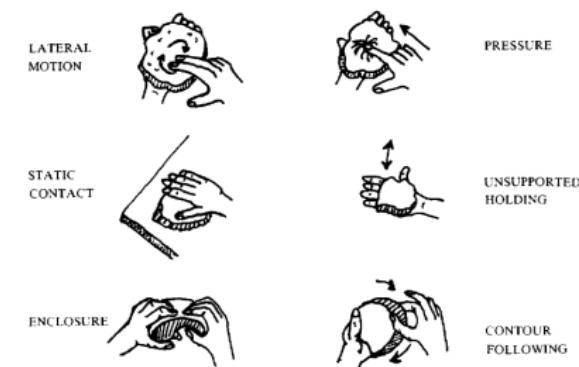


Figure 12: Haptic rendering. DOI: [10.3724/SP.J.2096-5796.2019.0008](https://doi.org/10.3724/SP.J.2096-5796.2019.0008)

Dispositivos hápticos: Exploración humana



Knowledge about object	Exploratory procedure
Substance-related properties	
Texture	Lateral motion
Hardness	Pressure
Temperature	Static contact
Weight	Unsupported holding
Structure-related properties	
Weight	Unsupported holding
Volume	Enclosure, contour following
Global shape	Enclosure
Exact shape	Contour following
Functional properties	
Part motion	Part motion test
Specific function	Function test

Figure 13: Exploración háptica. DOI:10.1016/0010-0285(87)90008-9

Dispositivos hápticos: Efectos

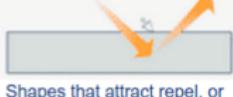
Barrier  <p>Sensation of hitting hard stop. indicates the end of a list or menu</p>	Damper  <p>Resistance force dependent on velocity. Creates the sensation of drag.</p>	Texture  <p>Sensation of roughness. Communicates the feel of a path or surface.</p>	Waves  <p>Vibration expressing sine, square or sawtooth waveshapes. Indicates intensity and velocity.</p>
Spring  <p>Opposing force increases/ decreases with controller position. communicates stiffness or rate level.</p>	Hill  <p>Increasing to decreasing force. Indicates transition to a new menu or list.</p>	Vector  <p>Directional force. Communicates preferred path.</p>	Detent  <p>Notches associated with selection position. Indicates position in a menu or list.</p>
Friction  <p>Force that resists relative motion. Provides the sensation of resistance.</p>	Grid  <p>Framework that attracts or repels. Indicates position in a two-dimensional space.</p>	Shapes  <p>Shapes that attract repel, or slide. Communicates a two-dimensional contour or form.</p>	Compound  <p>Combination of several basic effects. Detents can be combined with barriers or hills to construct lists with mid-points and end stops</p>

Figure 14: Efectos hápticos para interfaces de usuario. Fuente:
DOI:10.1016/B978-0-12-812800-8.00009-6

Dispositivos hápticos: Tecnologías

- Dispositivo 'graspable': Desplazar un joystick (p.ej. phantom) para obtener fuerzas/pares de reacción.
- Dispositivo 'wearable': Se instala directamente sobre los dedos (generalmente guantes) para detectar sus movimientos y aplicarles una respuesta específica (p.ej. fuerza de reacción y presión según el tamaño/forma/rigidez del objeto virtual).
- Dispositivo 'touchable': Se toca (generalmente pantallas) para obtener una sensación específica (p.ej. vibraciones, texturas, formas, fricción).

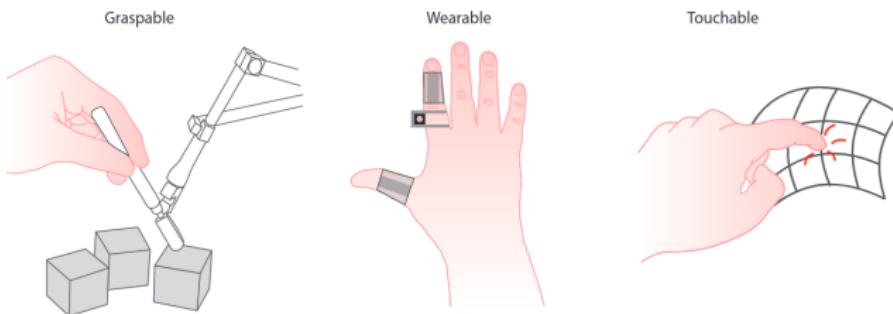
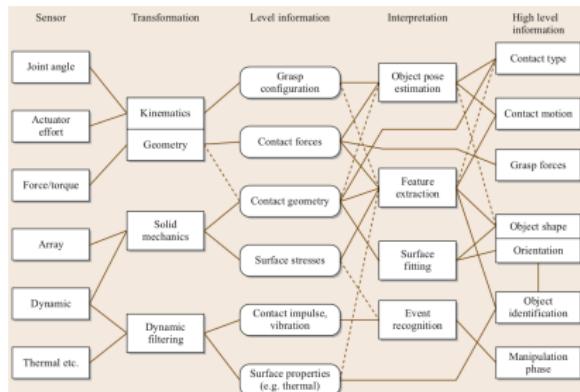
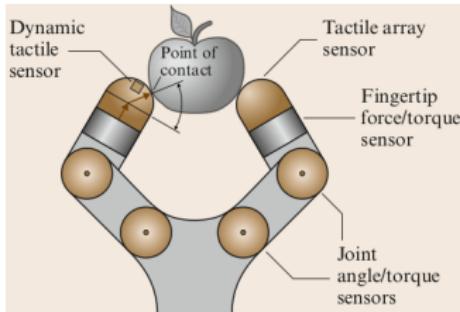


Figure 15: Tecnologías. DOI: [10.1146/annurev-control-060117-105043](https://doi.org/10.1146/annurev-control-060117-105043)

Dispositivos hápticos: Sensores

- Sensores propioceptivos (fuerza-par/posición): Instalados en las articulaciones del dispositivo haptico, detectan las posiciones de sus motores y las fuerzas/pares sobre ellos.
- Sensores táctiles matriciales: Se instalan en las superficies de contacto para conocer la distribución espacial de las fuerzas normales de contacto (presión= fuerza/área).
- Sensores táctiles dinámicos: Se instalan en las superficies de contacto para obtener información que varía rápidamente con el tiempo (vibraciones) y detectar deslizamientos o texturas.



Dispositivos hápticos: Técnicas de control en teleoperación

- Control directo: Sin autonomía del esclavo.
- Control compartido (shared): Combinación de control sensorial local y control directo del maestro.
- Control supervisor: El maestro envía comandos de alto nivel mientras que el esclavo tiene una autonomía local importante.

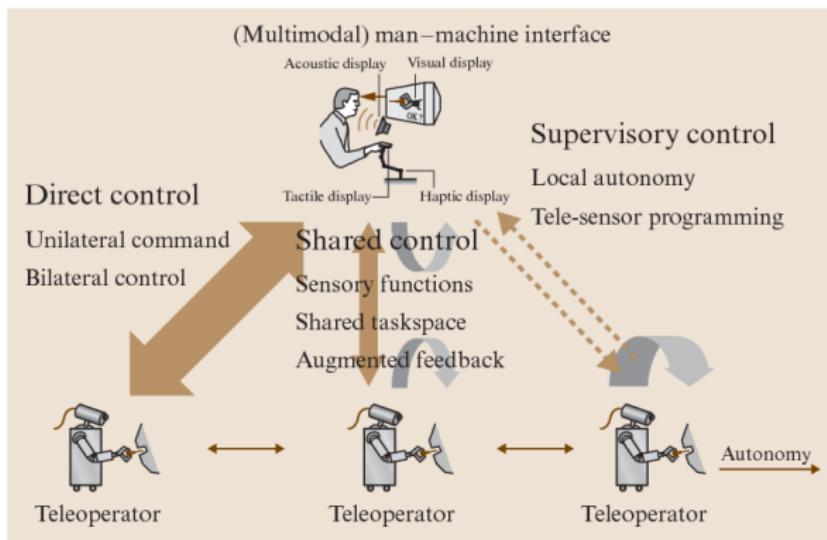
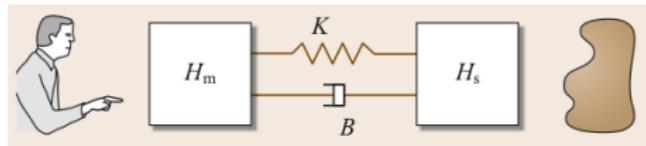


Figure 17: Teleoperación. DOI: [10.1007/978-3-540-30301-5_31](https://doi.org/10.1007/978-3-540-30301-5_31)

Dispositivos hápticos: Control bilateral

- Bilateral: El dispositivo registra los comandos del maestro (forward) y muestra la interacción del esclavo (feedback).
- Arquitectura 1 (Posición-Posición): Se relacionan la posición-velocidad del maestro (m) y del esclavo (s) con dos ganancias (K : elasticidad; B : amortiguamiento). Constantes iguales, fuerzas iguales. Constantes diferentes, escalado.

$$F_m = -K_m(x_m - x_{md}) - B_m(\dot{x}_m - \dot{x}_{md}),$$
$$F_s = -K_s(x_s - x_{sd}) - B_s(\dot{x}_s - \dot{x}_{sd}).$$



- Arquitectura 2 (Posición-Fuerza): Para evitar la fricción e inercia del esclavo, se realimenta el maestro con las fuerzas medidas en el esclavo. Inestable con superficies rígidas.

$$F_m = F_{\text{sensor}},$$

$$F_s = -K_s(x_s - x_{sd}) - B_s(\dot{x}_s - \dot{x}_{sd}).$$

Señales biológicas: Dispositivos EMG y EEG

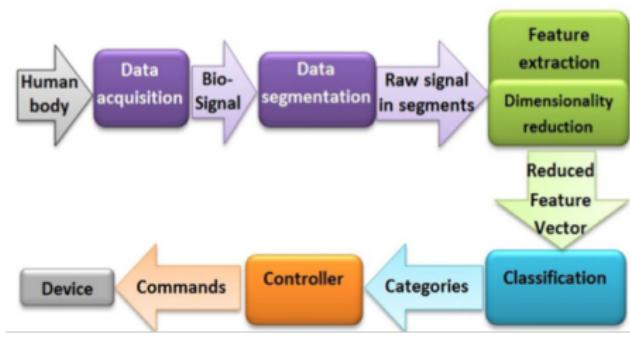
- EMG (Electromiografía): Actividad eléctrica de contracción de los músculos.
- EEG (Electroencefalografía): Actividad eléctrica cerebral.
- Estos dispositivos se basan en la instalación de electrodos sobre la piel (EMG) y el cuero cabelludo (EEG). Las señales eléctricas captadas son amplificadas y digitalizadas.
- BCI (Brain Computer Interfaces): Generalmente están basadas en dispositivos EEG no invasivos aunque también hay dispositivos invasivos (ECoG=electrocorticografía).



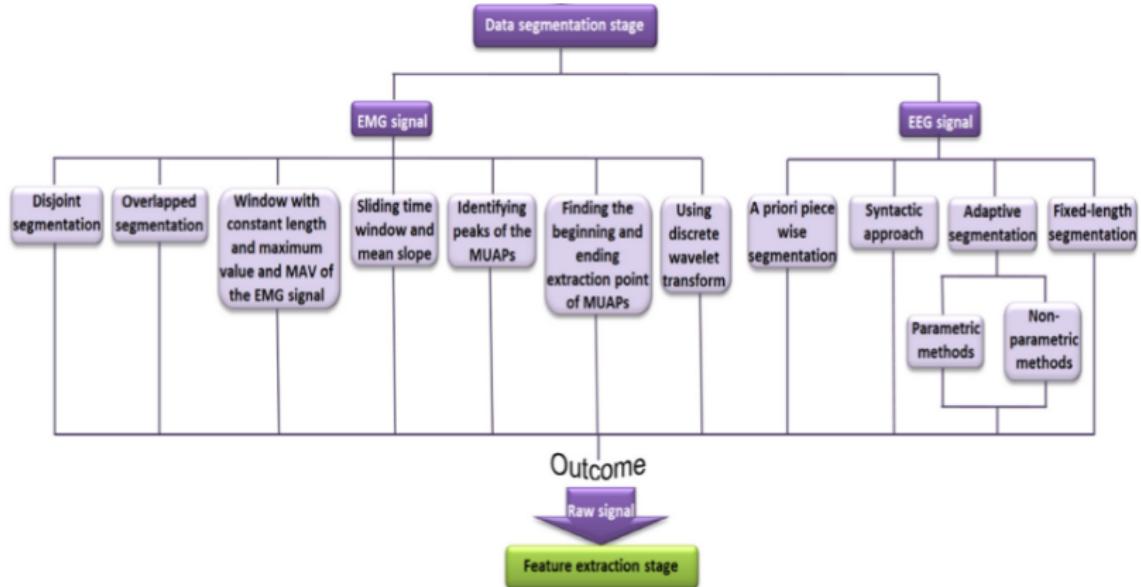
Figure 18: Dispositivos EMG (izq) y EEG (der).

Señales biológicas: Etapas de procesamiento

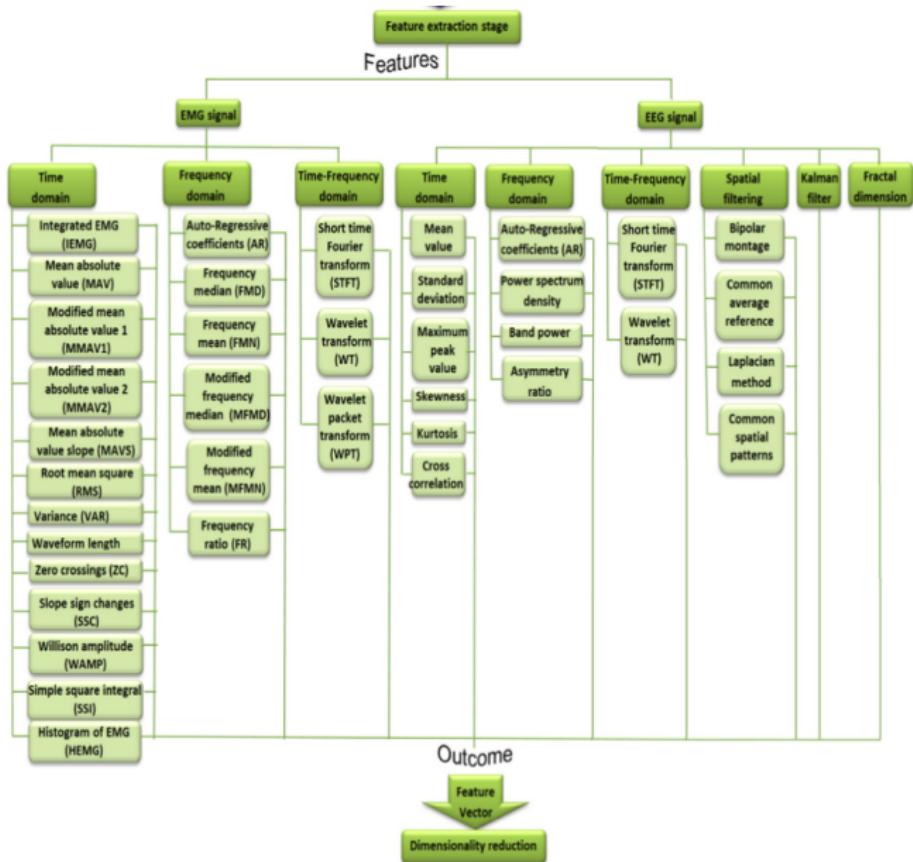
- Segmentación: Trocear la señal continua en segmentos.
- Extracción de características: Obtener el valor de un vector de características, que identifican la información importante. Hay 4 tipos: temporales (amplitud), frecuenciales (densidad espectral), temporales-frecuenciales (distribución de la energía), filtrado espacial (actividad en ciertos electrodos).
- Clasificación: Asociar valores de características con categorías.
- Controlador: Para cada categoría de las características, se manda un comando diferente al robot para diferentes aplicaciones (**teleoperación, asistencia, reconocimiento de emociones, evaluación ergonómica**).



Señales biológicas: Segmentación



Señales biológicas: Extracción de características



Señales biológicas: Clasificación

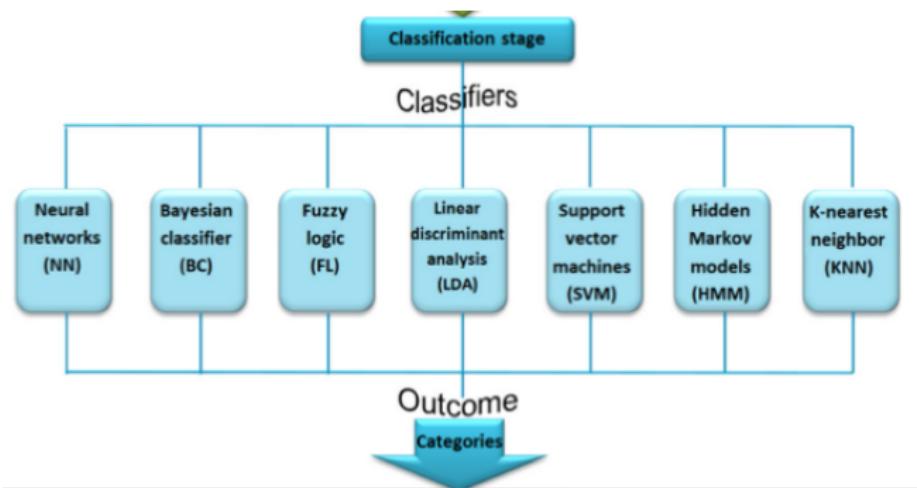


Figure 19: Clasificadores para EEG/EMG. DOI:
[10.1016/j.dcan.2015.02.004](https://doi.org/10.1016/j.dcan.2015.02.004)

Pantallas: Integración en robots



Pantallas: Diseño de la pantalla

- Elegancia y simplicidad: estilo unitario, refinado y adecuado.
- Escala, contraste y proporción: elementos claros. Armonía entre ellos y con su función.
- Organización y estructura visual: información agrupada, aprovechando jerarquías, y balanceada.
- Módulos: Enfocados a la resolución del problema global y organizados jerárquicamente.
- Imagen y presentación: Significado inmediato y generalidad.
- Estilo: Distintivo, integrado, comprensible y apropiado.

Pantallas: Ejemplo de Diseño

PÉREZ,MARIA23456789ZJUAN PEDRO DÍAZ
ANA230193TOMAS021090LUISA301299

PÉREZ, MARIA 23.456.789Z JUAN PEDRO
DÍAZ

TOMAS 02-10-90

ANA 23-01-93

LUISA 30-12-99

EMPLEADA: María Pérez DNI:23.456.789Z

CÓNYUGE: Juan Pedro Díaz

HIJOS:

NOMBRES FECHAS DE NACIMIENTO

Tomás 02-10-90

Ana 23-01-93

Luisa 30-12-99

EMPLEADA: María Pérez DNI: 23.456.789Z

CÓNYUGE: Juan Pedro Díaz

HIJOS:

Nº	NOMBRE	FECHA DE NACIMIENTO
----	--------	---------------------

1	Tomás	02-10-1990
---	-------	------------

2	Ana	23-01-1993
---	-----	------------

3	Luisa	30-12-1999
---	-------	------------

Pantallas: Mensajes de error

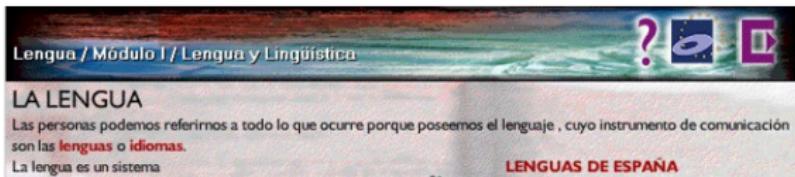
- Deben ayudar al usuario a detectar el problema y resolverlo.
- Mensajes poco adecuados pueden crear una mala experiencia en los usuarios.
- Incluir el diseño de los mensajes de error en el diseño de la interfaz.
- Consejos:
 - Mensajes específicos con consejos constructivos y tono positivo.
 - Redacción centrada en el usuario (ayudas contextuales, uso de sonidos...).

Error al utilizar un escáner

1. No se encuentra el dispositivo. Compruebe que está conectado al equipo.
2. La aplicación no encuentra el dispositivo. El error puede deberse a una mala conexión con el equipo.
Ver como resolver este problema.
3. No se encuentra el dispositivo. En la documentación podrá encontrar como resolver el problema.

Pantallas: Color

- Usar el color cuando es beneficioso (objetos 3D, diferenciación de contenidos...). No es necesario en aplicaciones alfanuméricas (BD, hojas de cálculo...).
- Usar el color de manera conservadora: diseñar primero en monocromo.
- Emparejar colores compatibles: negro+azul, azul+blanco. Evitar: magenta+verde, verde+blanco.
- Usar colores diferentes entre líneas de un listado.
- Mantener la consistencia de colores y evitar los fondos cargados que dificultan la lectura.



Prueba

Prueba

Prueba	Prueba	Prueba	Prueba
--------	--------	--------	--------

Pantallas: Elementos de interacción-Menús

- Selección por menús:
 - En menús binarios usar elementos opuestos (Aceptar/Cancelar)
 - Usar radio-buttons en menús múltiples
 - Usar posición fija para menús pull-down
 - Cada menú debe contener tareas relacionadas semánticamente. Emplear separadores en saltos de tareas.
 - Evitar una anidación excesiva: más de 2 niveles es complicado
 - Distinguir visualmente las opciones que se pueden usar
 - Nombres de funciones cortos y atajos de teclado relacionados
- Ejemplo: *Ver, Barra de herramientas, Suavizar, Barra de estado, Unidades de Medida, Perilar, Acuarela, Imagen, Centímetro, Pixel, Recortar, Cambiar el tamaño, Rotar, Regla, Efectos, Estampado.*

Ver	Imagen
Barra de herramientas	Recortar
Regla	Cambiar el tamaño
Barra de estado	Rotar
--	
Unidades de Medida	Efectos
Centímetro	Suavizar
Pixel	Perilar
--	
	Acuarela
	Estampado

Elementos de interacción-Huecos y Comandos

- Rellenado de huecos
 - Deben etiquetarse con nombres significativos
 - Su tamaño debe indicar su contenido
 - Usar una plantilla para alinear etiquetas/huecos
 - Indicar visualmente los campos obligatorios
 - Indicar los valores posibles (p.ej. formato de fechas/listas de valores)
 - Agrupamientos lógicos
- Lenguaje de comandos
 - Equilibrio entre potencia y simplicidad
 - Los comandos (léxico) deben ser familiares, específicos y congruentes y pueden sustituirse por abreviaturas
 - La sintaxis debe ser natural (fácil de recordar), evitar signos de puntuación y dar soporte de parámetros opcionales.
 - La línea de comandos de estar en un lugar visible, poder abrirse/cerrarse e incluir una ayuda interactiva.

Interfaces HRI: Sensores

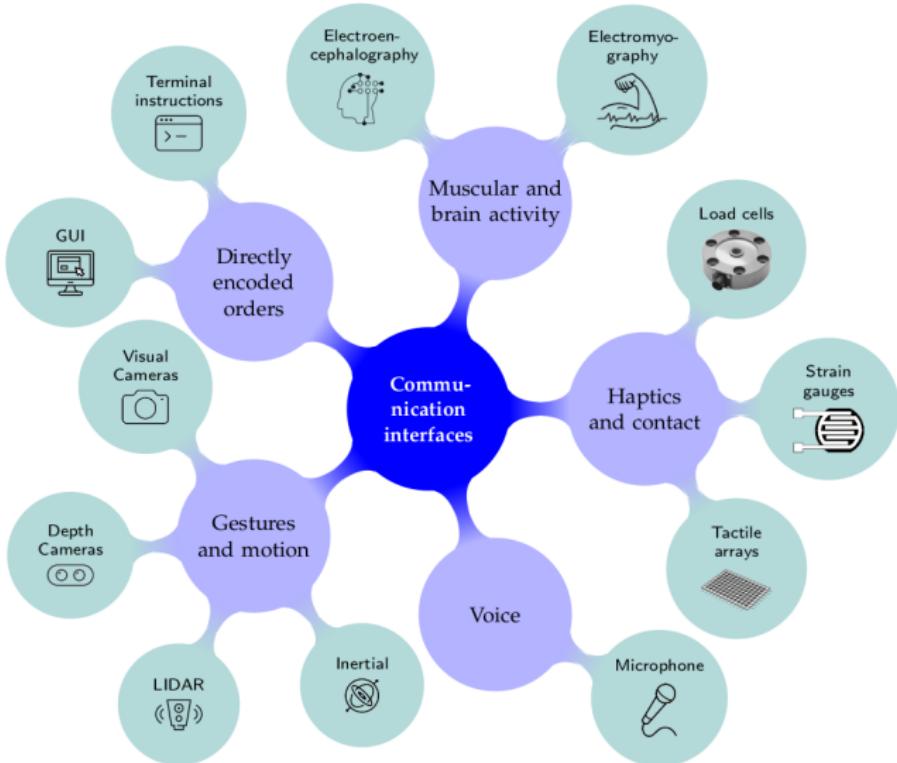


Figure 21: Sensores en interfaces HRI. DOI: [10.3390/s21124113](https://doi.org/10.3390/s21124113)

Percepción en robots sociales I

Social robot	Sensor	Feature	Semantic understanding task	Core algorithm
Kismet	camera	motion, color [20, 21]	object detection/recognition	nearest neighbor
	microphone	pitch, energy [22]	emotion recognition	Gaussian mixture model
Cog	camera	intensity [13]	face detection	template matching
	camera	motion [12]	object segmentation	correlation analysis
iCub	camera	intensity, shape [16]	human and object detection	mutual information
	camera	SIFT-based [17]	object recognition	hough transform + least-mean-squares mapping
GRACE	microphone	ITD, ILD, notches [18]	sound localization	
	laser reading	shape [27]	people in line perception	nearest neighbor
	camera	color, shape [28]	human tracking	Kalman filter
Robox	microphone	IBM's ViaVoice [27]	speech recognition	IBM's ViaVoice
	laser reading	raw data [30]	motion detection	nearest neighbor
	camera	shape [30]	object tracking	Kalman filter
	camera	color [31]	face detection	Heuristic filter
Recchham	microphone	Viterbi-based [9]	speech recognition	hidden Markov model
	camera	Harr-like [32]	face detection	boosting
	camera	learning-based [33]	face recognition	nearest neighbor
	camera	color, shape, motion [33]	human tracking	condensation
Robovie	camera	color, shape [33]	gesture classification	condensation
	camera	color, learning-based [34]	gesture recognition	nearest neighbor
	laser reading	human trajectories [35]	human behavior analysis	support vector machine + k -means
RUBI	laser reading	raw data [35]	human tracking	clustering and particle filter
	microphone	STBF-based [37]	emotion recognition	boosting
	camera	Gabor-based [38]	facial expression recognition	boosting + support vector machine

Figure 22: Sensores en robots sociales. DOI: 10.1007/s12369-013-0199-6

Ejemplos: [Kismet](#), [Cog](#), [iCub](#), [Robovie](#)

Percepción en robots sociales II

ARMARIII	camera	DCT-based [42]	face verification	nearest neighbor
	camera	intensity [43]	speaker tracking	particle filter
	microphone	time delay [43]	speaker tracking	particle filter
	microphone	learning-based [44]	sound classification	hidden Markov model
	camera	intensity, shape [45]	head pose estimation	Neural Network
	camera	color, intensity [46]	human gesture recognition	Neural Network + hidden Markov model
PaPeRo	camera	shape, 3D model [48]	face/eye detection, recognition	template match
	microphone	filter banks [49]	speech recognition	hidden Markov model
Huggable	tactile	calculation-based [50]	touch classification	Neural Network
MEXI	camera	color, intensity [13, 52]	face detection	template match
	microphone	frequency, energy [53, 54]	emotion recognition	fuzzy logic
	microphone	raw data [56]	sound localization	beam-forming
ROMAN	microphone	energy [56]	object tracking	particle filter
	camera	feature point [57]	emotion recognition	facial action coding system
	microphone	pitch, energy [59]	emotion recognition	bayes
BARTHOC	microphone	temporal shift [60]	voice detection	cross-power spectrum phase
	camera	haar-like [60]	human tracking	boosting
	laser reading	raw data [62]	human tracking	nearest neighbor
BIRON		color, haar-like [62, 63]	human tracking	boosting
microphone	time delay [63]	sound localization	cross-power spectrum phase	
Fritz	camera	haar-like [64]	face detection and tracking	boosting + Kalman filter
	microphone	time delay [64]	speaker localization	cross-power spectrum phase

Figure 23: Sensores en robots sociales. DOI: [10.1007/s12369-013-0199-6](https://doi.org/10.1007/s12369-013-0199-6)

Ejemplos: [ARMARIII](#) [PaPeRo](#) [Fritz](#)