



UNIVERSIDAD VERACRUZANA

Instituto de Investigaciones en Inteligencia Artificial

Una estrategia cortés para interacciones humano-robot locales

Que para obtener el título de
Maestro en Inteligencia Artificial

P R E S E N T A:

Jesús Eduardo Hermosilla Díaz

DIRECTOR:

Dr. Antonio Marín Hernández

Agradecimientos

Quiero agradecer a mis padres, Isabel Díaz Bolaños y Eduardo Herмосilla Rojas, por su amor incondicional y los sacrificios mediante los cuales siempre me han apoyado durante mi trayectoria académica. Así como a mis hermanas, Dania y Daniela, quienes siempre me han comprendido y han estado ahí para apoyarme. Las palabras de aliento de mi familia y la confianza que han tenido en mí han sido fundamentales para desarrollar este trabajo.

También quiero expresar mi gratitud a mis compañeros de generación. El trabajar en equipo y las valiosas discusiones que tuvimos han enriquecido mi aprendizaje y crecimiento profesional. Agradezco especialmente a Jennessi, Francisco, Nayeli, Isaac, Norma, Oscar y Jesús, porque a través de los momentos que compartimos en la cafetería o los pasillos del instituto me brindaron de la motivación y energía que necesitaba en mis momentos de mayor estrés. Sólo ustedes saben de los aciertos (y especialmente los desaciertos) que viví durante las clases, el laboratorio y en *Deutschland*.

Además, quiero expresar mi agradecimiento a mi asesor, el Dr. Antonio Marín Hernández y mis profesores, quienes han guiado mi trabajo y me han brindado una orientación académica invaluable. Así como al Dr. Stephan Pareigis, por recibirme en la HAW Hamburgo con las puertas abiertas y por mostrarme su apoyo durante toda mi estancia.

Por último, pero no menos importante, deseo agradecer a todas las personas que participaron como voluntarios. Sus deseos de participar han sido fundamentales para este trabajo. A todos aquellos que han contribuido de alguna manera a mi trayectoria académica y a la culminación de esta tesis, les estoy profundamente agradecido.

ÍNDICE GENERAL

1 | Capítulo 1 Introducción

- 1.1 Planteamiento del problema *13*
- 1.2 Justificación *15*
- 1.3 Pregunta de Investigación *17*
- 1.4 Hipótesis *17*
- 1.5 Objetivos *18*
 - 1.5.1 Objetivo general *18*
 - 1.5.2 Objetivos específicos *18*
- 1.6 Aportaciones *18*
- 1.7 Alcances y limitaciones *19*

2 | Capítulo 2 Antecedentes

3 | Capítulo 3 Marco teórico

- 3.1 Proxémica *35*
- 3.2 Regiones de interacción dinámicas *37*
 - 3.2.1 Transformación de coordenadas *37*
 - 3.2.2 Geometría de la región *38*
 - 3.2.3 Potencial de interacción *41*

- 3.3 Detección y localización de personas 42
- 3.3.1 Histograma de gradientes orientados 42
- 3.3.2 YOLO 44
- 3.3.3 Estimación de la posición 45
- 3.4 Planificación de trayectorias 46
- 3.4.1 Planificación global 46
- 3.4.2 Planificación local 47

4 | Capítulo 4

Diseño experimental y resultados

- 4.1 Implementación en ROS 49
- 4.1.1 Planificación de trayectorias con *movebase* 50
- 4.1.2 Detectores implementados 52
- 4.1.3 Simulación del ambiente 53
- 4.2 Estrategia empleada 55
- 4.3 Estructura del experimento 56
- 4.4 Proyecto integrador 57
- 4.5 Resultados 58
- 4.5.1 Gráficas 62

5 | Capítulo 5

Conclusiones y trabajo futuro

Apéndice A: Cuestionario 68

Apéndice B: Resultados de la prueba de *Shapiro-Wilk.* 69

Apéndice C: Análisis de diferencias significativas entre la configuración clásica y la propuesta. 70

ÍNDICE DE FIGURAS

1.1.	Comparativa entre robots.	10
1.2.	Vehículo de entrega de Starship	11
2.1.	Vista general del arreglo experimental.	22
2.2.	Arreglo experimental cuando el robot mantiene una mirada mutua.	26
2.3.	Trayectorias del robot y la persona en la simulación.	28
2.4.	Diagrama de fuerzas.	30
2.5.	Diagrama de una sub-meta durante la planificación.	31
2.6.	Diferentes tipos de espacios de interacción.	32
2.7.	Experimento con Pepper.	33
3.1.	Regiones del modelo original del espacio personal de Hall.	36
3.2.	Ejemplificación de los marcos de referencia.	39
3.3.	Dimensiones del espacio personal.	40
3.4.	Ejemplificación del potencial de interacción.	41
3.5.	Esquema general del histograma de bloques orientados.	44
3.6.	Arquitectura de YOLO V8.	45
4.1.	Configuración de los robots empleados.	50
4.2.	Mapa de celdas generado.	51
4.3.	Mapa de costos del mapa global.	51
4.4.	Mapa de costos local.	52
4.5.	Comparativa entre detectores.	53
4.6.	Visualización en tiempo real del marcador.	54
4.7.	Marcadores. (a) Velocidad de 1.3 m/s. (b) Caso estático.	54
4.8.	Simulación diseñada.	55
4.9.	Esquema general de la estrategia empleada.	56
4.10.	Diseño experimental del vertido.	57

4.11. Comparativa entre las trayectorias obtenidas I.	59
4.12. Comparativa entre las trayectorias obtenidas II.	60
4.13. Porcentaje de respuestas relacionadas con la confianza.	62
4.14. Porcentaje de respuestas relacionadas con la satisfacción.	63
4.15. Porcentaje de respuestas relacionadas con el disfrute.	64
4.16. Variaciones de la ecuación de control en diferentes casos.	65

Resumen

En esta tesis, se desarrolló una estrategia para las interacciones locales basada en la detección de personas, el cálculo de su posición y velocidad. Particularmente la generación de un novedoso diseño del espacio personal a partir de los parámetros mencionados (posición y velocidad). Estos espacios personales se incorporaron a la planificación de trayectorias del robot gracias a la creación de un mapa de costos, con el objetivo de que el robot respete los espacios personales a modo de cortesía. Durante el desarrollo, se aplicaron diferentes algoritmos de detección (HOG y YOLOv8) con el fin de identificar aquel con el mayor costo-beneficio al calcular la posición, orientación y velocidad de las personas. La característica de los espacios personales propuesta se basa en una descripción geométrica de las regiones de interacciones, en forma de elipses en donde la posición de las personas corresponde a uno de los focos [1]. La integración de esta estrategia en el *stack* de navegación para robots autónomos basados en ROS, permite su aplicación en diferentes sectores como lo son los robots de servicio y los vehículos autónomos. En general, los resultados muestran una mejora significativa en la capacidad del robot para interactuar de forma segura y eficiente con las personas.

INTRODUCCIÓN

Desde la antigüedad la humanidad ha soñado con robots; máquinas pensantes capaces de realizar tareas domésticas, labores de campo y actividades consideradas peligrosas, imposibles o extenuantes para los humanos. La palabra robot apareció por primera vez en 1920 en la obra teatral de ciencia ficción R.U.R. (*Rossum's Universal Robots*) del escritor checo *Karel Čapek*. El término se empleó originalmente para describir obreros artificiales; sin embargo, más de un siglo después y a pesar de múltiples debates, no existe una definición universal. Generalmente, el término se usa para referirse a máquinas autónomas capaces de replicar el comportamiento humano, esto se ve reflejado en la creciente atención hacia cuerpos humanoides como el planteado por *Čapek*, como se muestra en la figura 1.1.

Desde las culturas antiguas se han realizado representaciones primitivas de este concepto. Tal es el caso del mito del gigante de bronce llamado Talos, creado en la mitología griega y cuyo propósito era defender la isla de Creta de los invasores. En otras culturas se han empleado representaciones humanoides como los Golems o autómatas; compartiendo la idea de servir a los humanos a pesar de las diferencias mitológicas. Pese a la moderna invención de máquinas para realizar tareas domésticas, la ilusión constante de una máquina pensante y similar a nosotros refleja nuestro interés por estudiar y emular el comportamiento humano.

En la historia moderna, la popularización de los robots se debe en gran medida a las novelas escritas por el escritor estadounidense Isaac Asimov entre 1938 y 1942, donde también propuso las leyes (ficticias) que deben gobernar el comportamiento

de un robot [2]. Las leyes de la robótica son un conjunto de reglas descritas por Asimov en su cuento Runaway de 1942. Dichas leyes están diseñadas para garantizar que el comportamiento de los robots siga normas seguras y éticas, evitando que dañen a las personas o así mismos. De ellas se destaca la primera ley, ya que enfatiza la importancia de proteger la seguridad humana, siendo un principio central en el desarrollo de pautas éticas y modelos de interacción humano-robot modernos. A pesar de que, el material de origen es ficticio, hoy en día siguen influyendo en los debates sobre las implicaciones éticas de la robótica y la Inteligencia Artificial (IA) en nuestra sociedad.

Al mismo tiempo que ingenieros, escritores y artistas desarrollaron toda clase de ideas sobre autómatas y robots, en la filosofía y ciencias cognitivas se han explorado los límites de la inteligencia para comprender las implicaciones que una máquina de tal naturaleza tendría. En este contexto, la IA no se limita a la robótica, sino que se cuestiona la naturaleza de los pensamientos inteligentes y las acciones que conllevan a través de las computadoras como herramientas experimentales [3].

Si bien, esta área ha tenido algunos periodos de inactividad a lo largo de su historia, hoy en día está en crecimiento, particularmente debido a que ya se cuenta con el poder computacional para analizar una mayor cantidad de datos y por lo tanto, emplear algoritmos más potentes. Debido a que es un campo en crecimiento está teniendo un impacto significativo en una amplia gama de sectores, como la medicina, finanzas y la tecnología. Aplicaciones donde las capacidades del robot se utilizan para tareas de clasificación, detección y reconocimiento, etc. Por lo tanto, la importancia de la IA en la robótica radica en su capacidad para automatizar las tareas que realizan, lo que permite ahorrar tiempo y aumentar la eficiencia en la industria o en las actividades cotidianas donde se desempeñen.

Dentro de la robótica existen dos perspectivas directamente relacionadas con los conceptos fundamentales de la IA; los robots humanoides y los robots móviles dentro de los cuales podemos encontrar los vehículos autónomos (VAs). En la robótica se denomina vehículo autónomo a cualquier plataforma móvil con el nivel de autonomía suficiente para interactuar con su entorno y realizar tareas sin intervención humana. Aunque este término se popularizó con el incremento de los autos autónomos en la industria automotriz, sus aplicaciones se extienden a plataformas no humanoides como los vehículos de entrega u operadores de logística. Si bien, estos no poseen una estructura humana, no quiere decir que no deban comunicarse con las personas, ni que las tareas que llevan a cabo deban ignorar el comportamiento humano. Sino que, comparte retos similares a la robótica humanoide, como lo son el aprendizaje basado

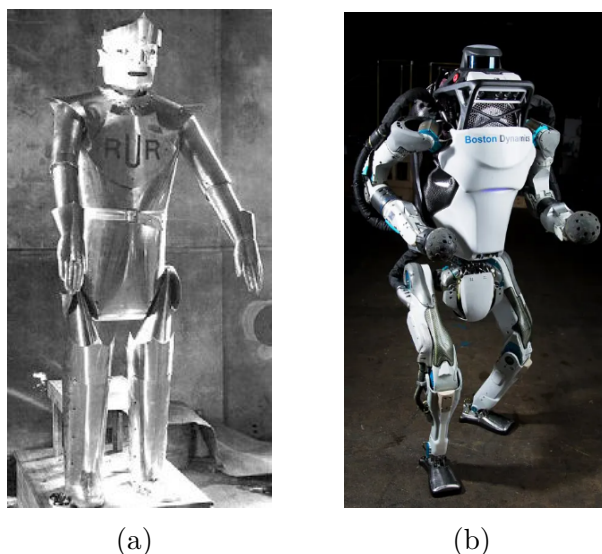


Figura 1.1: Comparativa entre robots, vestuario de la obra R.U.R. [4] de 1921 (de-
recha) y robot Atlas [5] moderno (izquierda).

en la experiencia, el razonamiento, el lenguaje natural, la percepción, el control, la exploración y cálculo de trayectorias, entre otros.

En el desarrollo de los VAs la tendencia en la industria se ha centrado en la navegación autónoma, sin embargo, en los últimos años se han explorado los retos relacionados con el ser humano. En los próximos años los VAs tendrán mayor presencia en la vida cotidiana, ya que las nuevas tendencias en el desarrollo de vehículos busca convivir de forma confiable y segura en los hogares, comunidades o lugares de trabajo, ofreciendo servicios de entretenimiento, educación, manufactura, sanidad, asistencia, etc.

Los retos que pueden surgir en estas situaciones sirven como fuentes de estímulos e inspiración para aportaciones al conocimiento en el campo de la robótica. Las soluciones a estos retos también presentan una oportunidad para fomentar la colaboración de diferentes disciplinas. A diferencia del enfoque industrial, la interacción entre humanos y VAs moderna concibe la comunicación a través de diferentes estímulos, como pueden ser sonidos, movimientos, lenguaje natural o expresiones faciales [6]. A diferencia de las computadoras, los VAs se distinguen por los procesos de percepción y acción con lo que interactúan con su entorno. De ahí que uno de los retos más desafiantes de la próxima generación de VAs sea la interacción física entre humanos y robots (pHRI por su siglas en inglés).

Hoy en día, existen muchas compañías ofreciendo diversos servicios relacionados con la interacción entre personas y VAs. Uno de los casos más populares debido a la simplicidad de la tarea que realiza es la plataforma *Roomba*, lanzado en el año 2002 por la compañía iRobot¹. Quienes ofrecen una versión para aspirar y otra para trapear el piso. Además de la gran variedad de empresas ofreciendo este tipo de servicios, existen áreas de oportunidad cuyo auge apenas está comenzando. Tal es el caso de la empresa *Starship Robots*², que lanzó un servicio sobre vehículos de entrega diseñados para navegar en espacios urbanos con el propósito de transportar comida, medicinas y paquetes pequeños en general. Actualmente, su uso está limitado a países desarrollados tecnológicamente; ya que requieren condiciones específicas en la infraestructura de las calles para realizar la conducción autónoma. Por otro lado, la empresa *Boston Dynamics*³ fundada en 1992 ha destacado por los robots humanoides que han desarrollado, siendo Atlas el modelo más avanzado. Equipados con movilidad, autonomía e inteligencia avanzada, sus prototipos son capaces de realizar tareas en espacios no estructurados o difíciles de recorrer como obras de construcción, almacenes o espacios recreativos al aire libre.



Figura 1.2: Vehículo de entrega de *Starship* [7].

La industria automotriz, tiene gran relevancia en el desarrollo de modelos de interacción, esto se ve reflejado en las ventas de vehículos autónomos por parte de empresas privadas. Tal es el caso de Tesla, una empresa de Estados Unidos que

¹Recuperado de <https://about.irobot.com> el 26 de febrero de 2023.

²Recuperado de www.starship.xyz/company/ el 26 de febrero de 2023.

³Recuperado de www.bostondynamics.com/atlas el 26 de febrero de 2023.

se dedica al desarrollo, producción y comercialización de vehículos semi-autónomos eléctricos e híbridos. Entre las ventajas que ofrecen frente a los vehículos clásicos está la mayor independencia del usuario durante la conducción, además de medidas reactivas para la prevención de colisiones con otros vehículos o peatones, sin embargo, este campo aún tiene mucho por mejorar.

A diferencia de los vehículos de uso exclusivo para ambientes de interior o en la industria, los modelos para describir la dinámica entre vehículos y peatones deben seguir los reglamentos de vialidad y tránsito vigentes en cada país. Adicionalmente, la presencia humana en la navegación de estos vehículos requiere considerar los efectos debidos al contexto social, de la misma forma que vehículos más pequeños diseñados para el servicio humano [8].

El sector industrial desarrolla VAs para tareas relacionadas con la producción en masa o la gestión de inventario. Tareas donde no sólo la presencia humana es mínima, sino que los escenarios son diferentes y el tipo de interacción se relaciona directamente con las medidas de seguridad acerca de los materiales y la zona en cuestión. Los vehículos de servicio se encuentran en una amplia variedad de entornos como hogares, oficinas, campus escolares, hospitales, casas de retiro, restaurantes, etc. Por lo que estos se encuentran en constante interacción con personas, no sólo con aquellos usuarios de su servicio sino con las que se encuentren en su camino. La investigación en este tipo de vehículos se enfoca en cómo mejorar las actividades que desempeñan al mismo tiempo que resguarden la seguridad de las personas a su alrededor.

Los VAs tienen aplicaciones prometedoras en situaciones de riesgo, como en escenarios posteriores a una catástrofe. En estos contextos, los VAs pueden enfrentar condiciones difíciles y únicas que pueden ser un desafío para los rescatistas. Estas tareas son peligrosas para los humanos y requieren una gran cantidad de esfuerzo y capacidad de adaptación para poder realizarlas de manera eficiente. Por lo que, los VAs pueden ser útiles para realizar tareas peligrosas y apoyar a los equipos de rescate en la búsqueda y rescate de supervivientes. En estos escenarios el motivo principal para usar VAs es aumentar la eficacia de las tareas de rescate al mismo tiempo que reducir los riesgos a los que se enfrentan las personas. Además de su capacidad para trabajar sin descanso durante largos períodos de tiempo en situaciones de emergencia. Siendo un área de oportunidad destacable debido a los requisitos de robustez, velocidad y movilidad que los vehículos deben desarrollar en el futuro. [9] Otra de las principales áreas de oportunidad se encuentra en la industria hotelera y las aplicaciones con vehículos de servicio. Haciendo énfasis en la interacción humano-robot para mejorar la calidad del servicio a los clientes. Hoy en día son pocos los hote-

les que pueden emplear estas tecnologías debido a los altos costos de adquisición y mantenimiento del equipo. A pesar de ello, muchos hoteles están interesados en desarrollar vehículos de servicio con un comportamiento más amigable hacia los usuarios debido al beneficio económico que el llamado "turismo inteligente" puede aportar. [10]

La ciencia ficción siempre ha sido una fuente de inspiración para la tecnología moderna, que continuamente se acerca a lo que antes parecía solo un sueño. Los AVs tienen cada vez más capacidades para realizar tareas que se asemejan a las capacidades humanas, ya sea conduciendo por sí mismos, tomando decisiones durante la conducción o adaptándose a diferentes situaciones de manera similar a como lo haría un ser humano.

Ahora, con el auge de la inteligencia artificial, los modelos de interacción entre VAs y humanos se ha consolidado como una rama fundamental en la robótica. Esta clase de modelos han demostrado ser una pieza clave en el desarrollo de sistemas más avanzados y capaces de realizar tareas de manera autónoma. Desde vehículos repartidores y turísticos hasta vehículos de exploración en otros planetas. Estos vehículos podrán ser utilizados para ayudar como guías, ser auxiliares en ciertas tareas o simplemente para hacer compañía a las personas. [11] Con la inteligencia artificial y robótica avanzando rápidamente, es probable que veamos a estos vehículos desempeñando un papel cada vez más importante en la vida cotidiana en un futuro cercano. Con el tiempo, es probable que seamos testigos de cómo la inteligencia artificial y la robótica trabajan juntas para lograr una mayor eficiencia y autonomía en una amplia gama de campos. Resulta asombroso presenciar como la tecnología está alcanzando niveles de sofisticación que antes parecían imposibles, y es probable que en el futuro veamos aún más innovaciones sorprendentes en el campo de la inteligencia artificial y los vehículos autónomos.

1.1 Planteamiento del problema

Los vehículos autónomos han revolucionado la forma en que se realizan muchas tareas, especialmente en el campo de la logística y el transporte. Estos han permitido una mayor eficiencia y seguridad en el manejo de cargas, así como una mayor flexibilidad en el transporte de personas o productos. Sin embargo, a medida que la tecnología avanza, es importante recordar que los algoritmos encargados de percibir, planificar e interactuar con el entorno continúan desarrollándose. A pesar de que actualmente se están implementando en el sector empresarial, existe un amplio margen de mejora para incrementar su capacidad de adaptarse y responder a entornos dinámicos en el futuro.

Clásicamente, la infraestructura de las ciudades se ha diseñado para minimizar el impacto del tráfico vehicular sobre el tránsito de las personas en la vía pública, permitiendo que coexistan sin mayores inconvenientes. Sin embargo, incluso en los espacios designados para la circulación de peatones, a menudo coexisten con plataformas de transporte alternativas. Ya que no fueron diseñadas para circular por las calles, sino que tienen la capacidad de invadir el espacio reservado para las personas, como las banquetas, pasillos o parques. En las ciudades modernas, la presencia de estos vehículos se debe a la demanda del tipo de servicio que realizan. Siendo especialmente problemático para las personas mayores, los niños y aquellos con discapacidad, que pueden sentirse inseguros o intimidados por la presencia de medios de transporte en espacios que deberían ser seguros para ellos. Su repentina introducción puede terminar generando conflictos debido a la nula comprensión del factor humano en la interacción.

En 2018, un peatón (*Eileen Herzberg*) perdió la vida al cruzar una autopista a causa de un vehículo autónomo propiedad de *Uber*. Este es el primer caso documentado de un accidente fatal entre un VA y un ser humano. Si bien, este incidente podría haber sido causado por diversos factores, vale la pena señalar que los sensores del vehículo detectaron la presencia del peatón, sin embargo, la reacción del automóvil fue incorrecta, causando el incidente [12]. Este incidente atentó directamente contra la vida de una persona, sin embargo, existen muchas situaciones donde la seguridad de uno o varios individuos se pone en peligro indirectamente debido a la presencia de VAs.

En la vía pública, los VAs no sólo son responsables de evadir cualquier colisión, sino que deben anticipar en alguna medida el comportamiento de las personas ante su presencia, para así prevenir accidentes causados por la sorpresa ante el comportamiento del vehículo. Este énfasis en la seguridad recuerda a la primera ley de la robótica mencionada previamente y además, añade una nueva dimensión al no sólo considerar su seguridad física sino también el nivel de cortesía durante la interacción. Se requieren modelos que describan correctamente las pautas que debe seguir los VAs al navegar en estos espacios, ya sea mediante la modificación de su trayectoria en función de posibles colisiones o mediante factores que no involucran contacto físico, como los gestos o el distanciamiento social [13]. Factores que pueden considerarse al estudiar las interacciones humano-robot e incorporar mecanismos en los VAs a través de la IA.

Si bien, se han estudiado ampliamente los modelos de tráfico vehicular (orientados principalmente hacia automóviles) y el comportamiento de multitudes, no se ha explorado en la misma medida el desarrollo de modelos que sean aplicables en interacciones locales. Esto es especialmente importante en el caso de los servicios ofrecidos por empresas, ya que algunas personas pueden tener comportamientos inesperados que pueden perjudicar o agredir a los vehículos. La problemática de no entender adecuadamente como se relacionan entre sí los VAs y las personas en diferentes contextos puede dificultar su uso efectivo en la sociedad. Esto se puede deber a que existe una falta de confianza, por parte de las personas sobre la capacidad de los VAs para realizar tareas de manera segura y eficiente. De igual forma, el no entender las interacciones puede conllevar a problemas de discriminación, ya que los robots pueden no ser capaces de comprender o respetar adecuadamente las preferencias y necesidades de las personas. Por lo tanto, la principal problemática se debe a la necesidad de realizar investigación en interacción humano-robot para garantizar que los VAs sean beneficiosos para todos.

1.2

Justificación

Hay varias motivos para considerar que los VAs serán necesarios en el futuro, uno de ellos es el cambio radical que está experimentando el sector empresarial. Con la diversificación de los servicios que pueden ofrecer y la creciente aceptación de los usuarios para acceder a ellos, las empresas se han adaptado para encontrar maneras de mejorar sus servicios. Los VAs pueden ser una herramienta valiosa para lograr esto, ya que pueden proporcionar asistencia y respuestas rápidas a algunas necesidades de los clientes, lo que permite a las empresas mejorar la productividad y satisfacción del cliente. También pueden mejorar la eficiencia y reducir los costos de mano de obra, ya que no necesitan ser remunerados ni recibir prestaciones. Sin embargo, es necesario establecer condiciones de seguridad mínimas para garantizar la seguridad al navegar en espacios públicos. Por lo que, es fundamental que los VAs cumplan con normas para minimizar el riesgo de accidentes y proteger a los usuarios a medida que enfrentan desafíos éticos y sociales durante la interacción.

Los VAs requieren una estrategia de procesamiento en tiempo real para funcionar adecuadamente en el mundo real. A diferencia de los experimentos realizados en un laboratorio, estos deben ser capaces de adaptarse y actuar de acuerdo a las convenciones sociales y a las normas establecidas en la proxémica, que se refiere al estudio de la distancia que se mantiene entre las personas durante la interacción social. Esto es esencial para garantizar la seguridad y la eficiencia en los servicios que ofrecen.

Solventar estas necesidades no solo implica crear computadoras más eficientes para tener el poder de cómputo necesario, sino desarrollar modelos que describan los fenómenos de interacción que ocurren debido a su exposición en las calles y la vida cotidiana.

Transitar a pie es una de las formas más comunes de transporte, ya sea caminando, trotando o corriendo. Por esta razón, desarrollar un modelo de interacción que incluya a peatones es un punto de partida adecuado. Además, este enfoque es apropiado debido a que los ciclistas y peatones ocupan los dos primeros niveles en la jerarquía de movilidad según el manual de calles mexicanas [14]. Esto significa que para garantizar la seguridad y eficiencia, los peatones y ciclistas deben tener prioridad sobre otros tipos de vehículos en la calle, así como, ser considerados en el diseño y la planificación del tráfico y la movilidad. Al enfocarse en interacciones locales es posible tomar ventaja del trabajo previo sobre la dinámica de las personas, realizado en áreas como la física, donde se tiene un registro de las principales medidas y variables involucradas en la interacción (e.g. velocidad, zancada y aceleración promedio). Por lo tanto, se pueden establecer métricas basadas en el comportamiento previamente reportado, que permitan comparar diferentes estrategias de interacción.

El enfoque principal para modelar los obstáculos que los VAs enfrentan se basa principalmente en obtener la cinemática de dichos objetos. Esto ha llevado a modelos rígidos que no tienen en cuenta la forma en que se mueven realmente los obstáculos ni el impacto que tiene el estado del vehículo en la interacción. Es necesario que los VAs usen su percepción del entorno para comprender el movimiento de las personas y reaccionar en tiempo real. Esto incluye la capacidad de detectar acciones comunes como caminar, correr o detenerse, así como, la habilidad de anticipar los movimientos futuros de las personas. Determinar qué cambios en la percepción y la interacción tienen un impacto directo en los accidentes o en la actitud de las personas hacia el vehículo es un problema abierto y no resuelto que contribuirá a mejorar el desempeño de los VAs en las ciudades del futuro. Por lo tanto, caracterizar el comportamiento de las personas durante las interacciones es esencial para desarrollar sistemas de conducción autónoma seguros y eficientes.

Una vez que se tenga una comprensión más profunda de los fenómenos que afectan la interacción entre personas y VAs, será posible implementar estrategias no invasivas en estos vehículos. Estas estrategias tendrían reglas similares a las leyes de la robótica, que establecen límites claros para la conducta de los VAs y protejan a las personas de posibles daños o agresiones. Estrategias de este tipo podrían ayudar a garantizar que

los VAs se comporten de manera responsable y segura en el mundo real, protegiendo a las personas y a otros usuarios en la vía. Originalmente, las leyes que planteó Isaac Asimov indican que:

- Un robot no puede herir a un ser humano ni, por inacción, permitir que un ser humano sufra daño.
- Un robot debe obedecer las órdenes que le den los seres humanos, excepto cuando dichas órdenes entren en conflicto con la primera Ley.
- Un robot debe proteger su propia existencia siempre que dicha protección no entre en conflicto con la Primera o la Segunda Ley.

Es importante seguir investigando y desarrollando estrategias con este enfoque para garantizar la seguridad y la eficiencia de los servicios que los vehículos podrían ofrecer en el futuro.

Si los vehículos autónomos se integran cada vez más en la vida cotidiana, es fundamental que puedan interactuar de forma natural e intuitiva con los humanos. En este sentido, la investigación sobre la interacción humano-robot puede ser de gran ayuda para mejorar el rendimiento y la funcionalidad de los VAs. Al estudiar cómo interactúan los humanos con estos vehículos, los investigadores pueden conocer las formas más eficaces de que se comuniquen, naveguen y realicen sus tareas de servicio. Esto puede ayudar a desarrollar vehículos que sean más fáciles de usar y comprender para los usuarios, lo que podría aumentar su aceptación en la sociedad. En general, la interacción humano-robot (IHR) es un campo importante de la investigación en robótica que se centra en cómo los humanos pueden interactuar de forma eficiente con los robots. Siendo un campo en constante evolución, con el potencial de mejorar la forma en que los vehículos y los humanos trabajan juntos y conviven en el mundo real.

1.3 Pregunta de Investigación

¿Modelar la interacción humano-robot puede contribuir a un comportamiento cordial en la navegación de vehículos autónomos?

1.4 Hipótesis

Es posible modelar la interacción entre vehículos autónomos y personas en espacios compartidos a través de la inteligencia artificial para mejorar la relación del individuo

con los servicios que éstos ofrecen.

1.5 Objetivos

1.5.1. Objetivo general

Modelar la interacción entre VAs y personas respetando los espacios personales como modo de cortesía al compartir espacio de navegación.

1.5.2. Objetivos específicos

- Analizar los principales modelos de interacción humano-robot en IHR y VAs.
- Implementar acorde al análisis previo, los algoritmos que permitan modelar la interacción
- Realizar experimentos simulados y reales de la estrategia en diferentes escenarios.
- Explorar el uso de métricas empleadas para la evaluación de la cortesía.
- Aplicar la estrategia en un VA para realizar una tarea de servicio.

1.6 Aportaciones

Como resultado de este trabajo, se realizaron las siguientes aportaciones:

- Una propuesta para modelar geoméricamente la región de interacción de una persona como una elipse dinámica, la cual considera la posición, velocidad y orientación de está.
- La creación de varios paquetes de ROS para la detección y estimación de la posición, velocidad y orientación de una persona en tiempo real, así como la generación de un marcador para visualizar la región de interacción correspondiente.
- La creación de un mapa de costos en ROS para integrar las regiones de interacción detectadas en la planificación de trayectorias.

- La publicación: Hermosilla-Diaz, J. E., Marin-Hernandez, A., & Gonzalez-Mendoza, M. (2022, November). A Proxemic Potential Field Approach for Modeling Interactions Between Autonomous Vehicles with Pedestrians and Cyclists. In 2022 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC) (Vol. 6, pp. 1-6). IEEE.
Ganador del premio a mejor artículo en el área de computación.

1.7 Alcances y limitaciones

La investigación que se llevará a cabo explorará los modelos propuestos en el área de robótica social, interacción humano-robot e inteligencia artificial. Al finalizar esta investigación se determinará una estrategia que permita auxiliar al vehículo durante la navegación. Esta estrategia servirá como guía para la toma de decisiones, así como para reaccionar a situaciones específicas al desempeñar sus actividades de servicio. La investigación se centrará en cómo los vehículos autónomos pueden interactuar con los humanos y cómo pueden comprender las señales humanas y reaccionar como modo de cortesía.

Para que este modelo funcione de manera eficiente, es necesario que los vehículos en los que se implemente cuenten con ROS (*Robot Operating System*), un sistema operativo de código abierto que permite el acceso a la odometría y, en general, a la codificación de la estrategia. Al contar con ROS, los vehículos podrán recopilar y procesar los datos necesarios para llevar a cabo la navegación de manera segura y eficiente. Además, ROS también permite el intercambio de datos y la comunicación entre diferentes componentes del sistema, lo que facilita la integración del modelo en un vehículo autónomo.

En este contexto se denominará como agentes a las personas que circulen sobre áreas públicas y no se van a tener en cuenta otros obstáculos que puedan estar presentes en el entorno. Además, no se van a considerar modificaciones o escenarios que puedan alterar el equilibrio del vehículo, como el uso de remolques o cargas no previstas. Adicionalmente, se asume que la reacción de los agentes es automática y que siguen las normas de comportamiento social convencionales. Esto significa que se espera que el comportamiento de los agentes sea estadísticamente consistente y que siga las reglas y normas establecidas en la sociedad. Las dimensiones del mapa, el vehículo y los agentes deben coincidir con escenarios típicos que se encuentran en entornos como escuelas, centros comerciales o zonas concurridas en general. Esto es importante para garantizar que la simulación sea lo más realista y precisa posible.

ANTECEDENTES

El comportamiento de un VA se rige por la tarea que realiza, de modo que su dinámica está diseñada únicamente para cumplir su objetivo. Este tipo de enfoque se puede aplicar exitosamente en ambientes controlados, donde el tipo de obstáculo es bien conocido. Desarrollar una estrategia cortés, por otro lado, requiere enfocarse en las personas como el agente principal con el que se interactúa y reconocer los efectos de la psicología en su comportamiento. Es posible estudiar este fenómeno en áreas que van desde la filosofía, psicología y artes, hasta las matemáticas y computación. En este capítulo, se abordarán los trabajos más relevantes en modelos de interacción humano-robot con énfasis en IHR e IA. Del mismo modo, se introducirán los antecedentes en las ciencias cognitivas que fundamentan estos modelos.

Cuando se trata de la interacción entre humanos y robots, es importante tener en cuenta que los retos que enfrenta un VA varían en función de la tarea que realice. Por ejemplo, desplazarse en una zona residencial puede presentar distintos desafíos que al hacerlo en un campus escolar. Sin embargo, en cualquier escenario, es fundamental considerar el factor humano y asumir que las acciones del robot pueden causar alteraciones medibles en los datos físicos detectados por el vehículo. Esto es especialmente importante cuando se diseñan modelos de interacción humana en estos escenarios, ya que es necesario tener en cuenta que pueden ocurrir interacciones imprevistas.

Con este objetivo en [15], Yue Hu et. al. analizaron la posibilidad de identificar

los factores de interacción medibles que podrían utilizarse en futuros controladores que involucren IHR. Los autores asumieron que las acciones del robot pueden causar alteraciones medibles en los datos físicos y fisiológicos de los usuarios y que estas mediciones podrían interpretarse a través de la personalidad y las percepciones de los usuarios. Se diseñó un experimento en el que los participantes usaron el robot como un *joystick* para jugar a un rompecabezas virtual, durante el cual el robot toma acciones imprevistas.

Para evaluar la interacción se recogieron datos fisiológicos y físicos, así como los resultados de dos cuestionarios sobre la percepción del robot y la personalidad de los usuarios. Los resultados del experimento sugieren que las personas extrovertidas parecen acercarse al robot desde una distancia más cercana y que las personas que perciben al robot como antropomórfico y animado podrían estar más estresadas o ansiosas al interactuar con él. Los autores destacaron que ninguno de los factores mostró correlaciones relevantes con el género. Estos hallazgos podrían servir para los investigadores interesados en desarrollar experimentos de IHR. Así como controladores para diversas aplicaciones, ya que las cantidades medibles son necesarias como entradas para los sistemas de control.

La seguridad en la interacción humano-robot puede dividirse en dos categorías, seguridad física y seguridad percibida, según se señala en [16]. La seguridad percibida está vinculada a factores como la incomodidad, la disminución de la seguridad percibida, la disminución del sentido de control y la desconfianza. Para investigar la relación entre estos factores y la seguridad percibida en la interacción humano-robot se llevó a cabo un experimento en el que los participantes interactuaron con un robot defectuoso. Se solicitó a los participantes que respondieran a cuestionarios que medían los factores relacionados con la seguridad, y se recogieron datos fisiológicos a través de una banda inteligente como se observa en la figura 2.1. Los resultados del cuestionario mostraron una correlación entre la comodidad, el sentido de control, la confianza y la seguridad percibida.

El trabajo concluye que es necesario centrar nuestra atención en comprender las condiciones en las que los seres humanos se sienten inseguros, ya que es en dichas condiciones donde se dan medidas cuantificables. Los factores clave que identificaron que influyen en la seguridad percibida son la comodidad, la experiencia/familiaridad, la previsibilidad, la sensación de control, la transparencia y la confianza. Estos factores deben tenerse en cuenta en las decisiones de diseño de las interacciones humano-robot para que éstas sean seguras. Además, los comportamientos de los robots deben ser

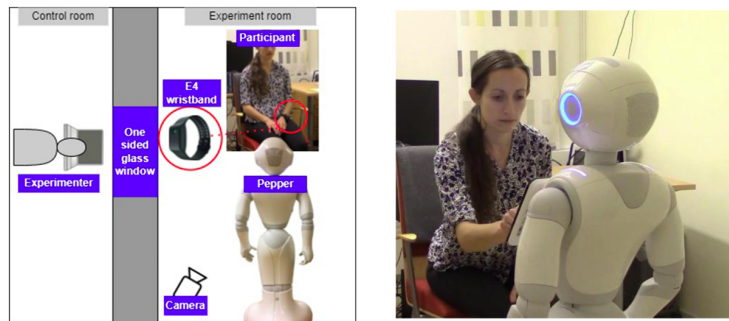


Figura 2.1: Vista general del arreglo experimental (izquierda) y una imagen en la que el participante contesta el cuestionario (derecha). [16]

familiares, predecibles y transparentes.

La falta de una representación humanoide en VAs ha incentivado el desarrollo de estrategias que compensen esta carencia. Tal es el caso de las interacciones basadas en cortesía, un concepto empleado en la comunicación verbal y no verbal con el usuario. En [17] por ejemplo, se llevó a cabo un estudio sobre la educación en la interacción humano-robot basado en la teoría de la educación de Lakoff (sociolingüista americano). Con el objetivo de evaluar si la aplicación de las reglas de Lakoff en el comportamiento de los robots sociales no humanoide es adecuada se evaluaron tres niveles diferentes de cortesía.

Los estudios incluyeron dos poblaciones diferentes (adultos mayores y jóvenes) y se realizaron en dos condiciones (vídeo y en directo). Los resultados mostraron que el comportamiento cortés del robot tenía un efecto positivo en la percepción que los usuarios tenían de la interacción con robots y que los participantes eran capaces de diferenciar entre los niveles de cortesía diseñados. El disfrute y la confianza en la interacción con el robot fueron mayores cuando los participantes en el estudio se sometieron a la condición en vivo en comparación con el vídeo, y los participantes se mostraron más satisfechos cuando interactuaron con un robot móvil en comparación con un manipulador. Los autores subrayan la importancia de evaluar los estudios de IHR en diversas condiciones y con diferentes poblaciones para poder generalizar los resultados.

Tal como Leimin y Sharon señalan en [18], las interacciones humano-robot están lejos de ser libres de errores debido a las limitaciones técnicas y la diversidad de usuarios. Esto se observa cuando un sistema es aplicado a condiciones del "mundo

real"después de haber sido probado en ambientes controlados dentro del laboratorio. Los errores de interpretación durante la interacción pueden degradar la percepción de la capacidad de un robot para realizar una tarea (definidos como errores de rendimiento en este trabajo) o degradar la percepción de la competencia socio-afectiva de un robot (definidos como errores sociales en este trabajo). El trabajo se centró en los errores sociales, identificando los principales atributos de la competencia socio-afectiva percibida mediante la revisión de estudios de IHR. Se propuso una taxonomía de errores sociales para IHR y se discutió el impacto de estos errores en diferentes escenarios. Su propuesta proporciona una base de datos para el análisis sistemático de la dimensión socio-emocional en IHR. Por otro lado, la propuesta de taxonomía de errores sociales fomenta el desarrollo de modelos centrados en el usuario.

Debido al incremento de servicios que un VA puede ofrecer, la seguridad es un factor crítico al modelar su interacción. Para sistematizar la bibliografía acerca de los niveles de seguridad necesarios durante la interacción humano-robot en [19] se proponen cinco categorías para clasificar los tipos de barreras sociales. Considerando únicamente las funciones básicas de los robots colaborativos al realizar tareas específicas. Revisando los métodos existentes para la seguridad psicológica durante la colaboración humano-robot y su impacto en el comportamiento del robot. De la misma forma, se analizaron los parámetros psicológicos de la incorporación de robots en entornos industriales y sociales.

Las categorías propuestas sobre las características de seguridad que minimizan el riesgo durante la interacción son: percepciones del robot para una IHR segura, control del robot habilitado por la cognición, planificación de acciones para una navegación segura alrededor de los humanos, características de seguridad del hardware y factores sociales y psicológicos. El estudio analizó las técnicas de evaluación de riesgos existentes para ofrecer seguridad, presentando así un análisis de la seguridad en los robots contemporáneos y proponiendo un marco con las características de seguridad mínimas, necesarias durante la interacción.

En [20], Leigh Clark analiza la existencia de barreras sociales entre la computadora y el usuario, destacando la importancia de una perspectiva cortés en el procesamiento del lenguaje natural. El autor afirma que puede haber limitaciones sociales en la comunicación oral apropiada durante la interacción. Se examinaron las funciones sociales que tiene la cortesía lingüística para determinar el uso apropiado del lenguaje que debe tener un robot en la interacción oral, así como la relación entre las voces y el lenguaje. Es evidente que una computadora tiene limitantes al ejecutar de forma adecuada ciertos conceptos lingüísticos relacionados con la socialización. Con

esta motivación, el trabajo de Clark utilizó la cortesía para investigar las posibles limitaciones sociales que rodean el habla en una computadora.

El trabajo sugiere que se deben llevar a cabo experimentos para comprender lo que los usuarios esperan de las interfaces de voz en términos de lenguaje y voz, y cómo esto se ve afectado por los diferentes contextos de interacción posibles. Además, propone investigar cómo afecta el habla sintética y humana al tipo de cortesía percibida, así como, evaluar las interacciones de los usuarios con una interfaz de voz que emplee estrategias de cortesía de este tipo. Estos experimentos ampliarían los conocimientos sobre la cortesía y la comunicación social con interfaces de voz, y también exploraron los límites entre la comunicación humana y la comunicación automática.

La existencia de factores directamente relacionados con la cortesía implica la posibilidad de cuantificar la distancia humano-robot para incrementar la seguridad de los usuarios. Como se menciona en [21], la adquisición de información en tiempo real sobre la distancia entre el robot y el humano permite al robot modificar su movimiento para garantizar la seguridad. En dicho trabajo, se propuso un marco de trabajo para calcular la distancia en tiempo real que permite una interacción humano-robot segura. En dicho marco el humano es detectado por un sensor RGB-D, y su posición se calcula mediante un algoritmo de detección y seguimiento del esqueleto. Para simplificar el problema, la representación del robot y la persona está representada por cápsulas, mientras que el cálculo de la distancia es realizado mediante el algoritmo GJK (*Gilbert-Johnson-Keerthi*). El algoritmo es validado mediante simulaciones y experimentos reales para demostrar su eficacia. Los resultados corroboraron la aplicabilidad del marco para aplicaciones seguras en IHR.

Clásicamente en un control de impedancia, el humano no tiene injerencia sobre la trayectoria deseada del robot. Sin embargo, en [22] Losey y O'Malley presentaron un algoritmo para el cálculo de trayectorias que, cuando se combina con el control de impedancia, permite al humano modular tanto las trayectorias actuales como las deseadas del robot.

A diferencia de trabajos relacionados, su método deforma explícitamente la trayectoria en función de las fuerzas aplicadas durante la interacción y no requiere de una guía humana constante. Los autores verificaron que su método es compatible con el control de impedancia tradicional y utilizaron una optimización restringida para derivar la forma de la deformación. El algoritmo se ejecutó en tiempo real y

se realizaron simulaciones para probar los parámetros involucrados. Los resultados experimentales mostraron que se obtiene una reducción del esfuerzo humano y una mejora en la calidad del movimiento en comparación con el control de impedancia clásico.

Es evidente que los principales factores para comprender la interacción humano-robot y mejorar la seguridad de los usuarios están relacionados con los procesos mentales presentes en la cognición. Uno de los conceptos más usados en IHR y estrategias de interacción cortés es la llamada proxémica. El término fue acuñado por primera vez por el antropólogo Edward T. Hall [23] para describir el estudio de la correlación entre las relaciones interpersonales con la distancia que guardan las personas al interactuar. Las variaciones en el comportamiento fueron asociadas a zonas del espacio representadas por circunferencias concéntricas que rodeaban al individuo. En su modelo, se clasificaron cuatro zonas en función del radio de cada circunferencia, la zona: íntima, personal, social y pública.

En robótica este marco teórico ha permitido obtener una interpretación más coherente de la realidad de la que se tendría con los modelos físicos usados clásicamente (e.g, el gas ideal, las analogías con fluidos o el movimiento browniano). Como se muestra en [24] donde se propone un marco empírico para la proxémica humano-robot. En dicho trabajo se muestra como la medición y control de las distancias interpersonales entre humanos y robots pueden usarse para interpretar, predecir y manipular el comportamiento proxémico en dichas interacciones.

A diferencia de los obstáculos no animados, las personas ajustan su posición, postura y demás factores en función el contexto donde se encuentren. Un análisis de las distancias interpersonales (humano-robot) y la influencia de la postura, sentada o de pie, del vehículo o persona al acercarse, fue realizado en [25]. El objetivo fue explicar cómo las personas se distancian física y psicológicamente de los robots a partir de la proxémica. Para ello realizaron un experimento con 60 personas en un ambiente controlado, los participantes interactuaron con el robot Wakamaru (figura 2.2). Realizando dos tareas, con el objetivo de examinar el distanciamiento físico y psicológico de los participantes. Donde la simpatía (agradable/desagradable) y el comportamiento de la mirada (mirada mutua/mirada invertida) del robot fueron manipulados.

Los resultados mostraron que los participantes a los que no les gusta el robot compensan el aumento de la mirada del robot manteniendo una mayor distancia fí-

sica con él, mientras que los participantes a los que les gusta el robot no difieren en su distanciamiento del robot en las distintas condiciones de mirada. Los resultados también sugieren que aquellos a los que no les gusta el robot revelan menos información personal. Estos resultados proporcionan un modelo teórico de la proxémica humano-robot que servirá de base para el diseño de comportamientos proxémicos para robots. Sin embargo, destacan que es necesario seguir trabajando para saber hasta qué punto estos resultados se pueden generalizar a otras situaciones, señales de comportamiento y diseños de robots.



Figura 2.2: Arreglo experimental cuando el robot mantiene una mirada mutua (izquierda) y cuando evita la mirada (derecha) con un participante. [25]

Al modificar la distancia con el vehículo la visión de las personas puede cambiar drásticamente su interacción. En este sentido, en [26] se analizó como perciben las personas a un robot sociable en función de su uso en distintas ubicaciones. El objetivo era comprender como es que los cambios en el habla y gestos al alejarse del vehículo modifican la interacción. Para ello se realizó un experimento en el que se atenuó artificialmente la capacidad del robot para comprender ciertas señales sociales en función de la distancia, variando su rendimiento. Los 160 participantes le dieron indicaciones al robot a través de la voz y gestos con el dedo, posteriormente respondieron a un cuestionario para proporcionar una evaluación del desempeño del robot. A diferencia de otros trabajos, no se encontraron pruebas que apoyen la idea de que la proxémica humano-robot sea relevante en la percepción de factores como la seguridad, la comodidad y la simpatía percibidas por el usuario.

Otros estudios exploran las características físicas a través de la proxémica para plantear el futuro diseño de robots sociables. En [27] Mohammad Obaid et al. analizaron las distancias interpersonales entre humanos y robots, así como la influencia de la postura, ya sea sentado o de pie en dichas distancias. Para llevar a cabo el estudio,

los datos de veintidós participantes fueron registrados y analizados para evaluar estos factores durante la interacción. Se diseñaron cuatro escenarios en las que el usuario podía desempeñar un papel activo o pasivo con respecto a un robot NAO. Además, el humano y robot podían estar de pie o sentados durante la aproximación.

Los resultados mostraron que los humanos permiten distancias interpersonales más cortas cuando un robot está sentado o en una posición más pasiva y dejan más espacio cuando se les acerca estando de pie. Por lo tanto, la postura del robot influyó significativamente en las distancias interpersonales en las interacciones humano-robot. A pesar de no presentar una correlación significativa, las mujeres dejaron menos espacio interpersonal entre ellas y el robot en todas las condiciones. Además, la estatura de los participantes no influyó en las distancias sociales. También destacaron que el uso del robot Nao afectó los resultados ya que su altura limitada y su expresión infantil influyeron en la cercanía entre el robot y los participantes.

En [28] se propone una estrategia de control que modela la interacción bilateral cuando un robot móvil no holonómico guía a una persona al mismo tiempo que mantiene una cierta distancia y sigue una trayectoria dada. Para modelar el comportamiento del ser humano en la tarea en cuestión, se propuso y evaluó un controlador cinemático no lineal. Los errores del control se calcularon a partir de las distancias entre el humano y el robot y entre el robot y la trayectoria. El análisis de estabilidad de los controladores y el sistema completo se realizó a partir de la teoría de estabilidad de Lyapunov. Es importante destacar que el trabajo incluyó la demostración teórica de la estabilidad no solo de los algoritmos de control independientes, sino también del sistema completo propuesto. Además, se propuso que en futuros trabajos se implementen zonas de interacción dinámicas en escenarios con múltiples humanos y múltiples robots.

En [29] se presenta un marco de seguridad humana en espacios de trabajo compartidos (entre humanos y robots). El marco general consta de diferentes etapas, en la primera se emplean algoritmos para la percepción del usuario. En la segunda etapa se estiman las señales sociales, en donde se recogen datos relacionados con la mecánica del usuario (posición, orientación y postura de las manos) que funcionan como parámetros de entrada. En la tercer etapa los estados humanos y sus señales sociales se integran en un modelo denominado zona social dinámica.

El objetivo de dicho marco fue integrar las zonas en el control de la navegación para hacer que los robots móviles se muevan con seguridad en espacios compartidos.

Posteriormente, en [30] se amplia este marco, permitiendo diferenciar la interacción al aproximarse a un ser humano y a un grupo de personas. Además, se incorpora al algoritmo de planificación D^* y el algoritmo de aproximación de ventana dinámica. A pesar de obtener resultados favorables, como se observa en la figura 2.3, se observaron limitaciones en la propuesta debido al método de detección y seguimiento de personas usado. Destacando que es necesario desarrollar técnicas de detección y seguimiento de personas robustas y en tiempo real. Adaptando las interacciones sociales a las características de diferentes culturas.

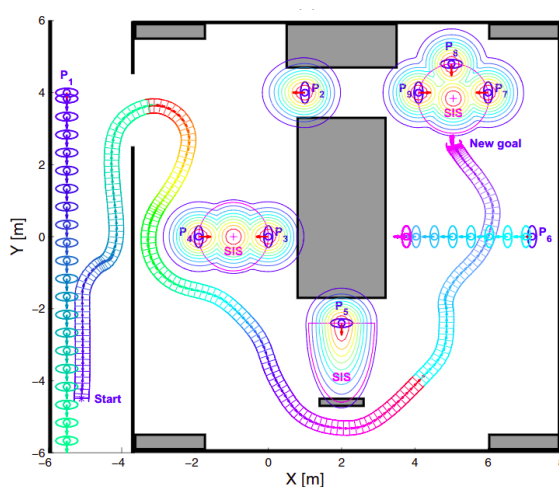


Figura 2.3: Trayectorias del robot y la persona en la simulación. [30]

En [31] se implementó un controlador reactivo basado en un mecanismo de toma de decisiones para guiar al robot hacia una determinada postura al mismo tiempo que mantiene una distancia con el usuario, denominado estado objetivo. Para lograrlo, se desarrolló un método de muestreo para estimar el estado objetivo proxímico del robot en relación con la distancia humano-robot y los parámetros de orientación. Así como un controlador activo basado en la proxémica para alcanzar dicho estado objetivo. Adicionalmente, se creó un planificador de trayectoria que maximiza los parámetros de interacción relacionados con la habla y los gestos.

En [32] se hace énfasis en las convenciones sociales, incorporando información sobre los estados humanos (posición, orientación y movimiento) y las interacciones sociales en el modelado del espacio de interacción social. A través de una función gaussiana asimétrica bidimensional se calculó el costo de cualquier punto en el espacio para modelar la interacción social (humano-robot). Posteriormente, se integró el

modelo en la planificación de trayectorias usando el algoritmo A*. El modelo busca minimizar la distancia recorrida hasta llegar a una meta, al mismo tiempo que mantiene una franja de seguridad alrededor de los obstáculos y mantiene una distancia cómoda con los humanos. Adicionalmente, tiende a permanecer a la izquierda de las personas (que es la convención social en China) y evita atravesar multitudes. Permitiendo que las personas se sientan seguras y cómodas cerca del robot.

El SFM (Social Force Model) propuesto por Dirk Helbin y Peter Molnar [33] plantea que las interacciones ocurridas en un entorno social pueden describirse mediante un sistema de fuerzas. El modelo se compone de fuerzas virtuales regidas por factores psicológicos, como la proxémica, junto a las fluctuaciones debidas a variables aleatorias que no están en nuestro control. En la simulación no se consideran los efectos debidos a fluctuaciones en el entorno ya que el propósito era enfocarse en la dinámica que tienen las multitudes en la vía pública.

En [34] se presenta un Modelo de Navegación Social (SNM) como modelo unificado de navegación e interacción que permite a un robot navegar en un entorno humano y responder a los humanos según sus intenciones. Partiendo de la posibilidad de querer evitar al robot, no evitarlo (mantener su rumbo) o acercarse (interactuar). El modelo propuesto se basa en la orientación de la cara y la superposición del espacio personal. Los parámetros se optimizaron sistemáticamente mediante un algoritmo genético. Las trayectorias de evasión, no evasión y aproximación de los humanos se clasifican en función de la orientación de la cara y la trayectoria prevista en un modelo de fuerzas sociales modificado. Además, se considera que el movimiento del robot es eficiente en términos de tiempo para el ser humano, ya que la trayectoria del robot no retrasa al ser humano cuando se desplaza hacia la meta.

En [35] se hace una extensión del SFM para modelar, calibrar y validar la interacción socio-psicológica del ser humano entre multitudes con el objetivo de integrar los VAs en la sociedad. Para el modelo se han asumido parámetros en función del grado de familiaridad del público con los VAs, el tipo de interacción humano-robot (e.g acercarse, enseñar, evitar al robot), la velocidad y altura, generando un diagrama de fuerzas como en la figura 2.4. Los resultados muestran que el rango de interacción de los humanos hacia a un robot de servicio es mayor en comparación con el rango de interacción hacia otro humano.

Una de los conceptos más utilizados para modelar la interacción son los campos potenciales, así como una de sus primeras implementaciones se le atribuye a Oussa-

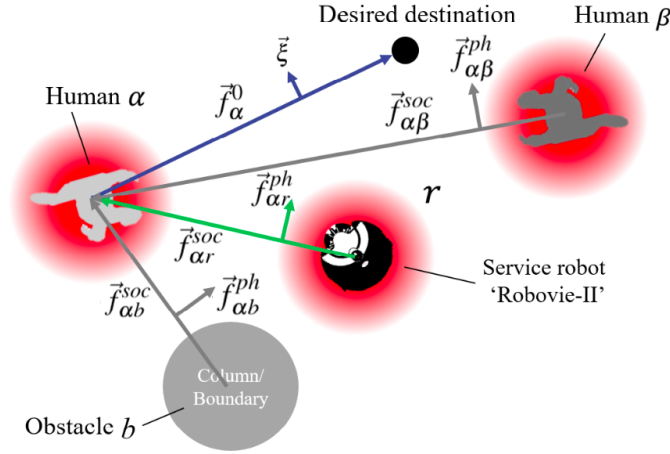


Figura 2.4: Diagrama de las fuerzas ejercidas sobre el humano α desde un humano β cercano, el robot r y el obstáculo b . [35]

ma Khatib en [36] donde se propuso un esquema de control basado en este concepto. El esquema tenía componentes de alto y bajo nivel para ser implementado en un brazo robótico PUMA¹. Tras implementarlo se demostró que el modelo era capaz de manipular y evadir ciertos objetos en un ambiente controlado. Su propuesta optimizaba el tiempo de ejecución al asignar el mecanismo de evasión de obstáculos a las componentes de bajo nivel mediante el concepto de campos potenciales. Llamó a su modelo FIRAS² ya que el lagrangiano mediante el cual se derivaron las ecuaciones de movimiento consideraba adicionalmente un campo potencial artificial.

Una propuesta moderna para considerar la dinámica en el modelo de interacción se encuentra en [37], donde se propuso un modelo dinámico que asumía que el espacio personal se podía representar mediante un óvalo. En dicho modelo se definió la longitud frontal del óvalo como una función directamente proporcional a las velocidades y aceleraciones del robot y la persona, así como el radio del espacio personal. A través de un experimento se evaluaron las zonas donde sentían mayor disconformidad ante la presencia del robot y la velocidad y aceleración asociadas. Los resultados del experimento corroboraban que se debía distribuir mayor superficie del espacio personal al frente de la persona, lo que justificó el uso de óvalos.

¹Programmable Universal Machine for Assembly, o Programmable Universal Manipulation Arm

²Force induisant une répulsion artificielle de la surface

En [38] se presenta un sistema de navegación para estimar las características pro-xémicas de las personas a su alrededor, y detectar el tipo de acción que están realizando en un entorno de una cafetería. En particular, los autores se enfocan en la acción de acercarse a la gente a la que consideran la acción que debe preceder al proceso de interacción. En este método, el robot no navega específicamente hacia un punto estático, sino que modifica su posición sub-meta dentro de un círculo de búsqueda dado unos puntos de referencia, como se muestra en la figura 2.5. El planificador de trayectorias propuesto está basado en la integración del modelado de acciones a partir de la detección por voz e imagen de las personas. Como modelo de la acción de una persona se usan los campos potenciales, siendo generados a partir de la posición estimada por la cámara del robot, así como la detección de voz de las personas.

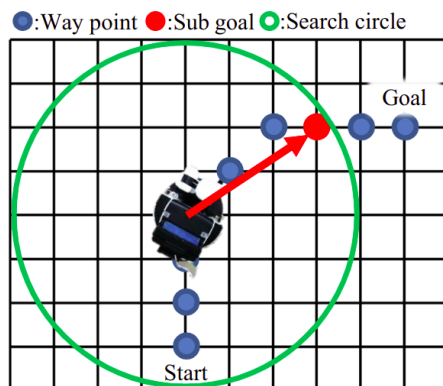


Figura 2.5: Diagrama de una sub-meta durante la planificación global y local. [38]

Los algoritmos actuales no tienen en cuenta la complejidad social de los entornos reales ni su relación con la hora del día o las actividades que se realizan en estos escenarios. Por lo que, en [39] se presenta un nuevo marco para la navegación social de robots en entornos humanos como centros de atención, introduciendo el concepto de mapeo social dependiente del tiempo. La ruta planificada por el robot considera variables que dependen de la hora y de las actividades programadas del centro. Ya que se describe cómo las áreas de interacción varían con el tiempo y cómo afectan a la navegación con conciencia humana.

El *stack* de navegación propuesto define una nueva función para el espacio de interacción social dependiente del tiempo que toma valores continuos y es configurable de acuerdo al usuario. El planificador global de rutas modela el espacio de interacción social mediante una función dependiente del tiempo. Mientras que los objetos

(obstáculos) se modelan mediante figuras geométricas, como se observa en la figura 2.6. Por último, se utiliza un optimizador de trayectoria de banda elástica como planificador local. Lo que le permite adaptar la navegación a posibles cambios durante la trayectoria.

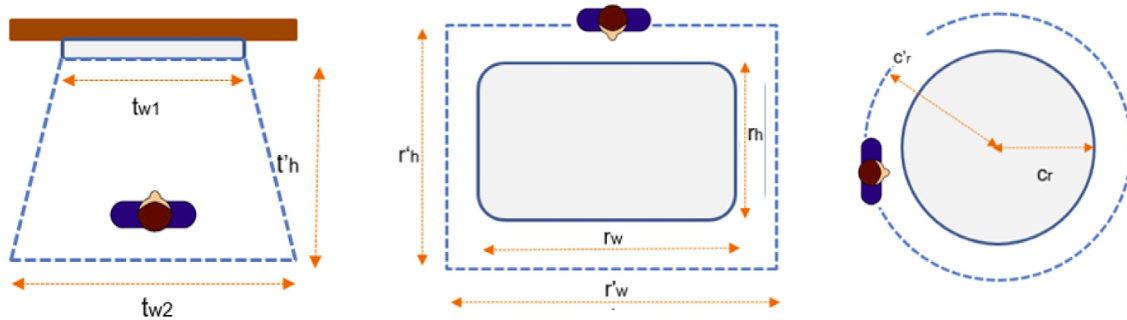


Figura 2.6: El espacio de interacción de los objetos se modela mediante un: trapecio isósceles (izquierda), rectángulo (centro) y formas circulares (derecha). [39]

Una estrategia de navegación para interactuar socialmente con las personas es propuesto en [40]. Donde se considera un área de interacción donde las personas pueden participar en interacciones de “alta calidad” con los robots, así como un área privada que no debe ser interferida por el habla o las acciones del robot. Para delimitar las dos áreas de cada persona se empleó el modelo de fuerzas sociales (SFM) con ayuda de un sistema de inferencia difusa. En el cual las funciones involucradas para la interacción se optimizan mediante un algoritmo de aprendizaje por refuerzo. El modelo propuesto se verificó mediante simulaciones y experimentos con un robot Pepper, permitiendo al robot mantener interactuar adecuadamente mientras mantiene su distancia como se observa en la figura 2.7.

En [41] se aplica aprendizaje por refuerzo para combinar el *stack* de navegación genérico de ROS con una política de control a alto nivel. La política adaptativa está diseñada para espacios de estados de baja dimensionalidad, por lo que no requiere procesos de entrenamiento prolongados y en cambio ofrece tiempos de reacción cortos, incluso cuando los recursos computacionales son limitados. Se comparó el método desarrollado con otros tres métodos de evaluación, obteniendo un porcentaje de colisiones del 3%. A diferencia de las estrategias tradicionales, el método desarrollado puede generalizarse fácilmente a nuevos entornos no considerados durante el entrenamiento.



(a)

Figura 2.7: Experimento con Pepper y un participante (izquierda) y su implementación en ROS (derecha). El área azul visualiza el área privada estimada. La línea verde y roja es el límite del área de interacción de calidad y privada respectivamente. [40]

Otro enfoque para modelar las interacciones humano-robot consiste en considerar que son procesos de decisión de Markov parcialmente observables (POMDP) indexados en el tiempo como el mostrado en [42]. Donde se obtuvieron mejores resultados que con otros modelos de Markov sin representación temporal explícita ni intenciones humanas como estados ocultos. El POMDP no dependiente del tiempo fue incapaz de representar con precisión las interacciones en las que el robot tenía la intención de tomar el Pittsburgh³ a la izquierda, y sus políticas optaron por saltarse el semáforo en rojo en lugar de arriesgarse a colisionar con el conductor humano. Las políticas POMDP dependientes del tiempo mostraron el comportamiento más consistente en todas las combinaciones de interacciones y fueron sensibles a las acciones de las personas. Este trabajo destaca la importancia de tomar las decisiones correctas a la hora de diseñar políticas de interacción eficaces para robots sociales.

Por último, se han usado metodologías tradicionales de análisis cognitivo, como la propuesta en [43], donde se presentan las redes de Petri como herramientas de modelado para la interacción humano-robot. En dicho trabajo, se construyen redes de operadores a partir de datos de operaciones de búsqueda y rescate con vehículos aéreos no tripulados (UAV). Permitiendo que la investigación futura puede utilizar la metodología no sólo de forma descriptiva, sino también preceptiva para construir modelos de interacción óptimos.

Después de revisar la literatura , se puede determinar que es importante recono-

³Término coloquial para referirse a la práctica de conducción en la que el primer vehículo que gira a la izquierda tiene preferencia sobre los vehículos que van de frente.

cer que las personas no deben ser consideradas simplemente como obstáculos inanimados. Adicionalmente, que existe una estrecha conexión entre los parámetros de movimiento como: la posición, velocidad y orientación de una persona, con su estado psicológico. También se observó que, aunque la seguridad de las personas es de suma importancia al interactuar con VAs, se destaca el enfoque de la cortesía al realizar investigaciones en el área de IHR. Por lo tanto, el comportamiento de un VA debe considerar la percepción psicológica de los individuos, tomando como base elementos como la proxémica, cortesía y parámetros de movimiento. Estos elementos formarán la base de la estrategia desarrollada más adelante.

MARCO TEÓRICO

Los VAs pueden utilizar diversas formas de detectar las señales o expresiones relacionadas con el comportamiento social de las personas. Estas pueden incluir gestos, expresiones faciales, movimientos con la mirada, la orientación de la cabeza, instrucciones verbales, comunicaciones basadas en el tacto, cambios en la dirección o velocidad, etc. Además, pueden ser voluntarias o involuntarias. Por lo tanto, los modelos de IHR requieren de la coordinación de diferentes áreas para ser implementados. En este capítulo, se presentarán los fundamentos teóricos para desarrollar una estrategia cortés para la interacción humano-robot, así como los métodos y algoritmos de detección utilizados en el desarrollo de esta estrategia.

3.1 Proxémica

Los animales territoriales, como los perros o gatos, utilizan marcas de olor para delimitar las zonas que habitan, de la misma forma que los humanos construyen muros y fronteras para marcar sus territorios. El sociólogo Edward T. Hall sugirió que los humanos siguen pautas de comportamiento como las seguidas por los animales, pero con interacciones más complejas. En lugar de atacar o escapar cuando alguien invade su espacio personal, los humanos utilizan señales para comunicar a los demás su estado de ánimo.

Hall observó la existencia de reglas de comportamiento dadas por la cultura y el contexto histórico que llevan a las personas a mantener determinadas distancias con

las personas que los rodean. El término proxémica, propuesto por Hall, se refiere al estudio del espacio que mantienen las personas entre sí y su relación con los factores psicológicos involucrados en la interacción [44]. Al extrapolar el comportamiento entre personas con interacciones entre personas y VAs, la proxémica establece un conjunto de pautas al modelar la interacción humano-robot. Inicialmente, se han identificado cuatro regiones estáticas de interacción, asociadas a diferentes distancias.

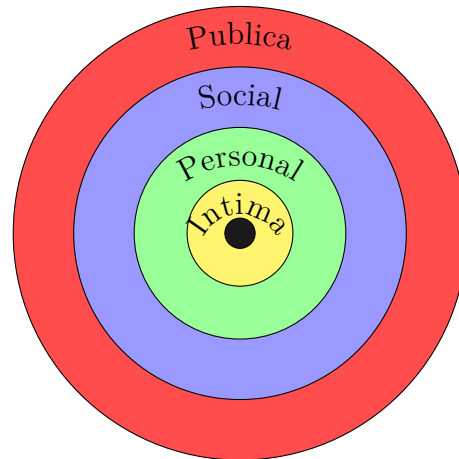


Figura 3.1: Regiones del modelo original del espacio personal de Hall, el individuo se encuentra en el centro.

En la Figura 3.1 se pueden observar las regiones proxémicas de Hall, representadas por círculos concéntricos. Su relación entre sí y con el espacio personal de un individuo se describen a continuación:

Región pública. Región donde se llevan a cabo las interacciones formales y se mantiene una distancia mayor. Suele utilizarse en contextos donde son necesarios los discursos formales, como las figuras públicas o por motivos de seguridad, como por ejemplo, durante la pandemia de Covid-19. Es la menos personal de las cuatro y se utiliza en contextos de distanciamiento social.

Región social. Tiende a ser un híbrido entre lo íntimo y lo público manteniendo una distancia intermedia. Suele utilizarse en contextos profesionales o casuales. En los espacios compartidos, la mayoría de las personas se encuentran en esta región ya que promueve una comunicación más efectiva.

Región personal. Está reservada para amigos o personas cercanas. Comúnmente, las personas se sienten incómodas cuando se invade esta región. Sin embargo, es

posible tocar a la persona mediante gestos o hacer aproximaciones paulatinas durante la conversación para promover sentimientos de cercanía.

Región íntima. Está reservada para los amigos más cercanos, la familia y parejas íntimas. Es imposible ignorar a las personas cuando invaden esta región. La violación de esta región puede ser reconfortante en algunos contextos y molesta en otros. Existen normas sociales que regulan el rango de estas interacciones en público, ya que algunas personas se sienten incómodas al ver a otras interactuando en la zona íntima.

3.2 Regiones de interacción dinámicas

Los movimientos descritos por las personas involucran cambios repentinos en su trayectoria. Por lo tanto, para desarrollar modelos interactivos es fundamental tener en cuenta características de movimiento como dirección y velocidad, así como características psicológicas como, por ejemplo, el espacio personal. La interacción de las personas con su entorno se puede modelar a través de regiones en el plano. Los peatones se modelan mediante regiones dinámicas elípticas. De modo que para tener interacciones seguras se requiere que los VAs no crucen estas regiones.

3.2.1. Transformación de coordenadas

En robótica, las transformaciones de coordenadas son fundamentales para describir y controlar el movimiento de los vehículos en un entorno tridimensional. Estas transformaciones permiten convertir las coordenadas de un sistema de referencia a otro, lo que facilita la planificación de movimientos y la interacción con objetos del entorno. La transformación de coordenadas homogéneas es una representación matemática ampliamente utilizada para describir la posición y orientación de un objeto en el espacio tridimensional. En esta representación, se utiliza una matriz 4×4 para combinar las coordenadas cartesianas (x, y, z) con una componente adicional.

Por lo que un vector en \mathbb{R}^n expresado en coordenadas homogéneas consta de $n + 1$ componentes. La adición de la nueva componente representa el factor de escala (w). Para volver a la representación en el espacio original basta con dividir cada componente del vector en \mathbb{R}^{n+1} por el factor de escala. Es conveniente asignarle un valor unitario a w para no alterar la interpretación física de los vectores posición. Las transformaciones homogéneas son utilizadas para establecer la configuración de un cuerpo rígido, así como para cambiar el marco de referencia desde el cual se describe

la configuración del vehículo.

El conjunto de matrices que representan la rotación de un cuerpo rígido en un espacio n dimensional alrededor del origen se conoce como grupo especial ortogonal $SO(n)$. En particular, las rotaciones de un ángulo θ en el plano tienen la forma:

$$R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}. \quad (3.1)$$

La matriz $R \in SO(2)$ representa una rotación en sentido antihorario. La traslación de un cuerpo rígido en \mathbb{R}^2 por otro lado, no se puede expresar como una matriz de 2×2 elementos. Sin embargo, al emplear la matriz de rotación R simultáneamente con una traslación hacia el punto p se define un grupo de transformaciones con las que es posible manipular un cuerpo rígido, conocido como $SE(n)$:

$$T := \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix} \in SE(n), \quad (3.2)$$

donde $R \in SO(n)$ y $p \in \mathbb{R}^n$.

Esto permite compactar las matrices de rotación y traslación sobre el plano utilizadas mas adelante ($SE(2)$) mediante una matriz cuadrada de 3×3 .

Un vector en coordenadas homogéneas descrito desde el marco de referencia $\{O\}$ se puede representar desde un marco de referencia $\{C\}$ a través de la transformación homogénea $T_{\{OC\}}$, como se observa en la figura 3.2.

$$T_{\{OC\}}\mathbf{p}_{\{O\}} = \mathbf{p}_{\{C\}}. \quad (3.3)$$

Dado que los VAs utilizan un sistema de posicionamiento basado en odometría y un mapa del entorno, estas matrices son útiles para realizar el cambio de coordenadas entre el marco de referencia global (que está fijo) y el local (que está en movimiento). En adelante, siempre que no se especifique lo contrario, se asume que la configuración (x, y, θ) de un VA está descrita desde el marco de referencia global.

3.2.2. Geometría de la región

En este trabajo se propone que la interacción de un peatón con un VA se modele a través de regiones elípticas en el plano. El perímetro de dicha región corresponde a los límites del espacio personal y como se mencionó tendrá la geometría de una elipse.

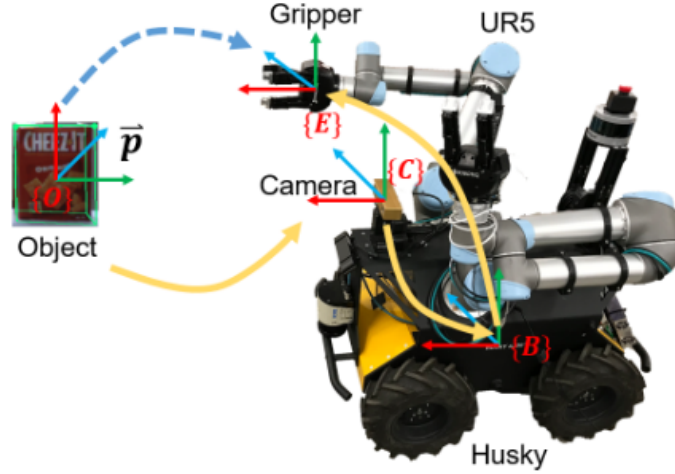


Figura 3.2: Ejemplificación de los marcos de referencia de la configuración empleada. [45]

Como se describe en la figura 3.3, para generar la región de interacción se considera que el peatón se posiciona en uno de los focos de la elipse, con el semi-eje mayor apuntando en la dirección del movimiento y donde el tamaño de este eje es relativo a su velocidad. Así, la distancia frontal $f_p(v)$ se define como una función escalar que depende de su velocidad:

$$f_p(v) = b_p + \lambda_p v \Delta t, \quad (3.4)$$

donde v es la velocidad del peatón, b_p es la longitud de la espalda, Δt es el tiempo transcurrido y λ_p es el factor de relación de velocidades. Este último se determina en función de la proporción deseada que tendrá la velocidad en la región alrededor del peatón. La longitud lateral l_p coincide con la distancia perpendicular al eje mayor.

A una velocidad $v = 0$, la región se convierte en un círculo de radio b_p , donde b_p es el tamaño del espacio personal considerado en la proxémica.

Resolviendo el sistema de ecuaciones formado, podemos expresar los elementos de la elipse en función de la longitud del espacio personal.

$$a(v) = \frac{f_t^2(v) + b^2}{2f_t(v)}, \quad (3.5)$$

$$c(v) = \frac{f_t^2(v) - b^2}{2f_t(v)}. \quad (3.6)$$

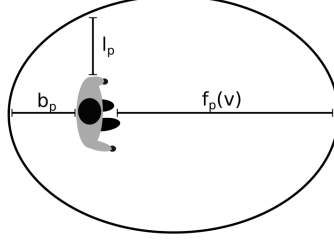


Figura 3.3: Dimensiones del espacio personal; donde b_p es la longitud trasera, l_p la longitud lateral y $f_p(v)$ la longitud frontal de una persona.

Considerando al robot como un marco de referencia local, la posición de un agente $\mathbf{r}(x_g, y_g, \theta_g)$ en el marco de referencia global se puede simplificar mediante una cadena de transformaciones homogéneas. En donde la rotación de \mathbf{r} alrededor del origen, se define como:

$$R_\theta \mathbf{r} = \begin{bmatrix} \cos(\theta_g) & -\sin(\theta_g) & 0 \\ \sin(\theta_g) & \cos(\theta_g) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.7)$$

La traslación de \mathbf{r} a uno de sus focos se expresa como:

$$T_c \mathbf{r} = \begin{bmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.8)$$

Y en general, la traslación a un punto P se calcula como:

$$T_p \mathbf{r} = \begin{bmatrix} 1 & 0 & p_{xg} \\ 0 & 1 & p_{yg} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.9)$$

Por lo tanto, la región interactiva de una persona respecto al robot (marco de referencia local), se obtiene a través de la traslación de la elipse hacia un foco, seguida de una rotación y una traslación a un punto P respecto al mapa (marco de referencia global).

$$T_p R_\theta T_c \mathbf{r}_{mapa} = \mathbf{r}_{robot} \quad (3.10)$$

La cadena de transformaciones que describe la elipse respecto al mapa se calcula mediante las transformaciones inversas de la ecuación anterior.

$$\mathbf{r}_{mapa} = T_c^{-1} R_\theta^{-1} T_p^{-1} \mathbf{r}_{robot} \quad (3.11)$$

La matriz de coeficientes A que define los elementos de la elipse, es decir, el eje semi-mayor y semi-menor, se define como:

$$A = \begin{bmatrix} 1/a^2 & 0 & 0 \\ 0 & 1/b^2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (3.12)$$

Haciendo $T = T_c^{-1}R_\theta^{-1}T_p^{-1}$, se define una transformación homogénea que genera una elipse con las características necesarias para que el vehículo establezca la región interactiva de cualquier agente. La siguiente ecuación:

$$E_i(\mathbf{r}(x_g, y_g, \theta_g)) = (T_i\mathbf{r})^{-1} AT_i\mathbf{r} \quad (3.13)$$

Genera la ecuación de una elipse dinámica ya que la longitud frontal del espacio personal es proporcional a la velocidad del agente. Por lo que la forma explícita de la ecuación de la elipse ubicada en $\mathbf{r}(x_g, y_g, \theta_g)$ es:

$$\frac{((y - y_g) \sin \theta_g + (x - x_g) \cos \theta_g - c)^2}{a^2} + \frac{((y - y_g) \cos \theta_g - (x - x_g) \sin \theta_g)^2}{b^2} = 1. \quad (3.14)$$

3.2.3. Potencial de interacción

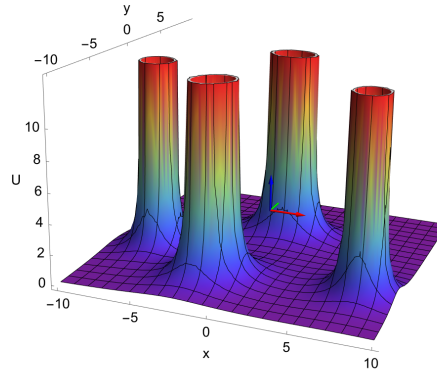


Figura 3.4: Ejemplificación del potencial total dado por cuatro agentes.

El comportamiento humano durante la interacción se puede modelar como una fuerza repulsiva que evita las colisiones entre personas y VAs. Si asumimos que esta

interacción actúa como una fuerza conservativa, podemos utilizar nuestro modelo como sistema auxiliar en la navegación autónoma a través del gradiente.

$$\mathbf{F} = -\nabla V^{agents}, \quad (3.15)$$

en donde el potencial de interacción se compone de las contribuciones debidas a la repulsión con n agentes.

$$V^{agents} = \sum_{i=0}^n v_{ti}(\mathbf{r}). \quad (3.16)$$

Nótese que no es necesario calcular el gradiente en todo el espacio de trabajo porque basta con que la función sea continua y diferenciable en la localidad de los agentes. En general, el potencial resultante se intensifica a medida que el vehículo se aproxima, haciendo que la fuerza disminuya a medida que se acercan. Por lo que la ecuación resultante es:

$$v_{ti}(\mathbf{r}) = \frac{\alpha_1}{E_{ti}(\mathbf{r})}, \quad (3.17)$$

donde la intensidad de la interacción está controlada por el coeficiente α_1 .

3.3 Detección y localización de personas

Los algoritmos de detección de personas desempeñan un papel crucial en el desarrollo de vehículos autónomos y la interacción humano-robot. A pesar de tratarse de una tarea específica, existe una amplia variedad de enfoques. A continuación se describirán los algoritmos empleados.

3.3.1. Histograma de gradientes orientados

El histograma de gradientes orientados (HOG, por sus siglas en inglés) es una técnica utilizada para extraer el vector de características de una imagen y así detectar objetos. El algoritmo divide la imagen en un conjunto de bloques y calcula un histograma a partir del gradiente de cada uno de los bloques. Estos histogramas representan la distribución de la orientación de los gradientes en cada bloque. Lo que permite utilizar un algoritmo de aprendizaje automático para detectar la forma de un objeto en la imagen.

Inicialmente se calcula el gradiente¹ vertical y horizontal (G_x y G_y) de la imagen I . El computo de la magnitud y orientación del gradiente se obtiene mediante:

$$|\Delta G| = \sqrt{G_x^2 + G_y^2} \quad (3.18)$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (3.19)$$

Posteriormente, la imagen es dividida en bloques conectados no superpuestos tal como se muestra en la figura 3.5. En cada bloque se computa el gradiente como se describió anteriormente. Se realiza un proceso de agrupación sobre las orientaciones del gradiente para generar los n intervalos de interés. En función de la manipulación de ángulos con o sin signo, se puede optar por un rango de $[-180, 180]$ o $[0, 180]$ respectivamente. El histograma se construye mediante un proceso de votación donde se asigna directamente el valor de la magnitud del gradiente. Aunque también existen técnicas de interpolación, donde se asigna la magnitud del gradiente a los dos intervalos más cercanos mediante un peso ω proporcional a la distancia de la orientación al centro de cada intervalo, como se describe en:

$$\omega_k = \max\left(0, 1 - \frac{\theta(x, y) - \theta_k}{\delta\theta}\right) \quad (3.20)$$

donde $\theta(x, y)$ representa el valor actual de la orientación del gradiente, θ_k el valor medio del k -ésimo intervalo y $\delta\theta$ el tamaño de los intervalos.

La frecuencia asignada al k -ésimo intervalo del histograma es proporcional a la magnitud del gradiente multiplicada por el peso correspondiente:

$$h(k) = \sum_{(x,y) \in I} \omega_k(x, y) |\Delta G(x, y)|. \quad (3.21)$$

Computacionalmente los histogramas de cada celda se representan mediante un arreglo de n elementos, donde cada elemento representa un intervalo del histograma. Cada bloque debe ser normalizado para prevenir el ruido debido a la iluminación y demás variaciones relacionadas con la luz. Esto se puede realizar mediante factores de normalización como la norma L_1 , L_1^2 , L_2 , etc. El vector de características de la imagen se forma al concatenar todos los histogramas normalizados. Al definir un conjunto de entrenamiento se puede emplear una máquina de soporte vectorial (SVM por sus siglas en inglés) para la detección de personas.

¹Alternativamente, se pueden usar aproximaciones como el detector de bordes Sobel por el bajo coste computacional.

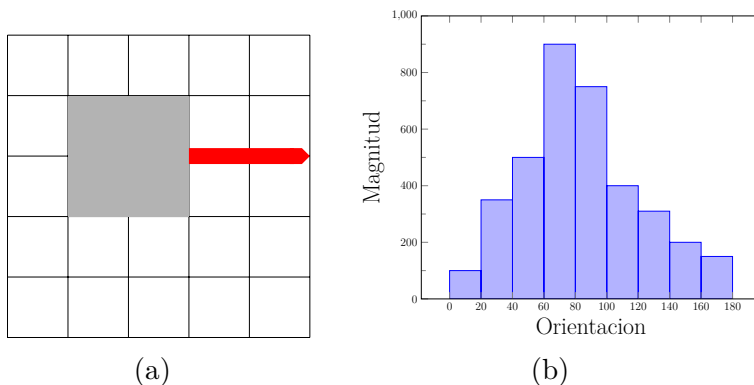


Figura 3.5: Esquema general del histograma de bloques orientados (sin signo) de una imagen: a) bloque a procesar, b) histograma generado.

3.3.2. YOLO

El acrónimo YOLO significa "You Only Look Once" (solo miras una vez), haciendo referencia a su capacidad de realizar la detección de objetos en una sola ejecución de la red neuronal. A diferencia de métodos previos que empleaban ventanas deslizantes seguidas de un clasificador que se ejecutaba múltiples veces por imagen, o enfoques más complejos que dividían la tarea en dos pasos, donde se proponían regiones candidatas y se aplicaba un clasificador a esas regiones, YOLO presenta una aproximación directa.

YOLO utiliza una salida de regresión para predecir las detecciones, en lugar de utilizar múltiples etapas o clasificadores adicionales. Esto se logra mediante el uso de dos salidas separadas: una para la clasificación de las probabilidades y otra para la regresión de las coordenadas de las cajas delimitadoras o englobantes (*bounding box*). Inicialmente, YOLO unificó los pasos de detección de objetos al predecir simultáneamente múltiples cajas delimitadoras.

Para esto, la imagen de entrada se dividió en una cuadrícula de tamaño $S \times S$, y se predijeron B cajas delimitadoras de la misma clase en cada elemento de la cuadrícula, junto con la confianza asociada a cada clase. Cada predicción de cada caja delimitadora incluye cinco valores: P_c , b_x , b_y , b_h , b_w . Donde P_c representa la puntuación de confianza del recuadro, reflejando la certeza del modelo de que el recuadro contiene un objeto y su nivel de precisión. Los valores b_x y b_y indican el centro de la caja en relación a la celda de la cuadrícula, mientras que b_h y b_w representan la altura y la anchura de la caja en relación a la imagen completa. El

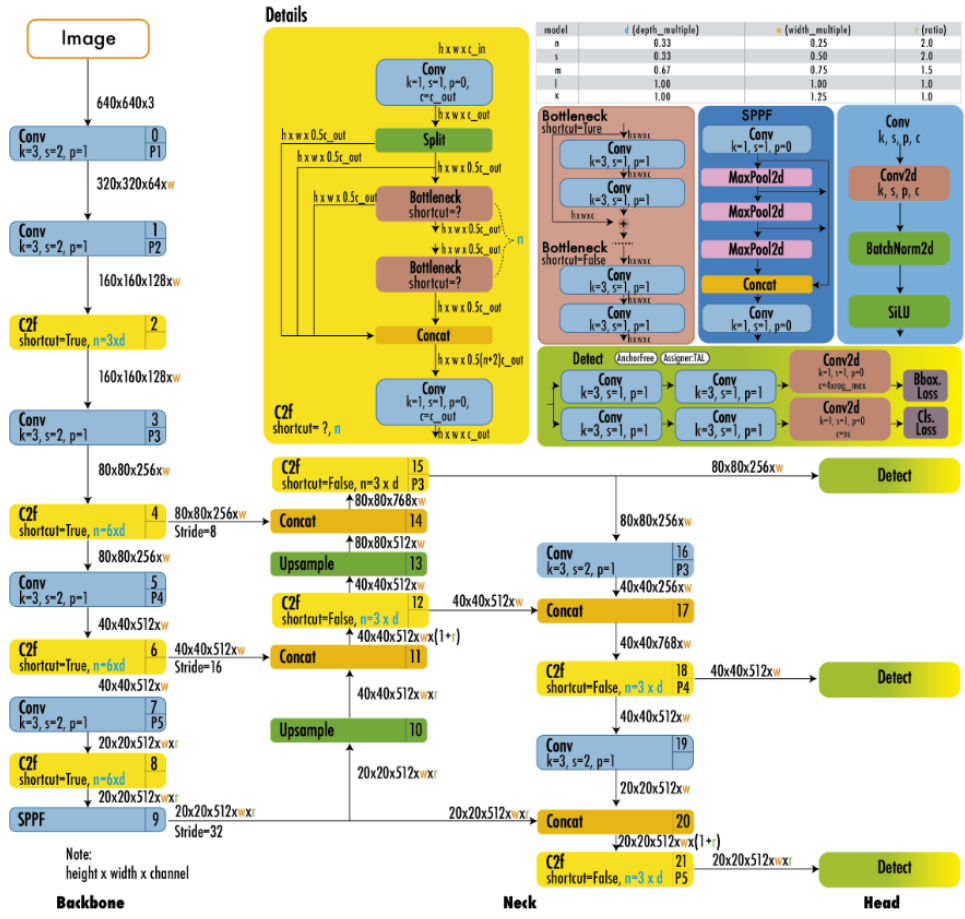


Figura 3.6: Arquitectura de YOLO V8. [46]

resultado de YOLO es un tensor de tamaño $S \times S \times (B \times 5 + C)$, al cual se le puede aplicar opcionalmente una supresión no máxima (NMS) para eliminar detecciones duplicadas.

3.3.3. Estimación de la posición

Tras aplicar la detección de personas, se obtiene una caja delimitadora, al calcular los píxeles (u, v) del centro se puede obtener una estimación de la posición en píxeles. Posteriormente, se utiliza la nube de puntos correspondiente para obtener la posición en el sistema de coordenadas de la cámara. Para obtener la nube de puntos a partir de una cámara RGB-D, se siguen los siguientes pasos:

- Se realiza una calibración de la cámara RGB y del sensor de profundidad para obtener los parámetros intrínsecos y extrínsecos de ambos dispositivos, obteniendo una matriz de parámetros k .
- Se capturan simultáneamente imágenes RGB y datos de profundidad.
- Se captura cada par de imágenes RGB y profundidad desde el mismo punto de vista, de forma que estén sincronizadas temporalmente.
- Se aplica la matriz de transformación k sobre cada píxel en la imagen RGB para generar la nube de puntos.
- Esto permite obtener la coordenadas (x, y, z) en el sistema de coordenadas de la cámara para cada píxel (u, v) de la imagen.

Para calcular la posición de una persona se realiza un procedimiento análogo, procesando únicamente el píxel correspondiente al centro de la caja delimitadora. En función de la capacidad de procesamiento es posible establecer un subconjunto de N posiciones consecutivas, desde las que se pueden calcular otros parámetros como la dirección de desplazamiento o la velocidad lineal. Debido a la existencia de falsas detecciones o de ruido, se puede optar por algoritmos de post-procesamiento con el fin de ajustar la trayectoria final como el uso de un filtro de Kalman.

3.4 Planificación de trayectorias

Para navegar por espacios compartidos, los VAs necesitan algoritmos de planificación de trayectorias. Y con el fin de evitar la explosión computacional, la planificación de trayectorias suele clasificarse según la perspectiva global y local. Estas son dos componentes fundamentales de los VAs para tomar decisiones inteligentes y ejecutar tareas con eficacia.

3.4.1. Planificación global

La planificación global consiste en determinar una estrategia de alto nivel o una secuencia de acciones para alcanzar un objetivo específico. Se ocupa de la toma de decisiones a largo plazo y opera sobre un mapa global o una representación del entorno del robot. Para ello, el vehículo utiliza sensores, como cámaras, Lidar² o sonares, percibiendo y recopilando información sobre el entorno. Posteriormente, los datos se procesan para construir una representación del mundo (e.g mundo de celdas,

²Acrónimo del inglés; Light Detection and Ranging

mapa de costos, grafo de visibilidad). Este modelo debe contener las características esenciales, incluidos los obstáculos, los puntos de referencia o las ubicaciones meta. Utilizando el modelo del mundo, el vehículo ejecuta algoritmos de planificación para encontrar una trayectoria óptima desde su ubicación actual hasta el objetivo deseado. Esto puede implicar algoritmos como A*, el algoritmo de Dijkstra o árboles aleatorios de exploración rápida (RRT). El resultado de la planificación global es un plan de alto nivel o una trayectoria que guía las acciones del robot hacia su objetivo.

3.4.2. Planificación local

La planificación local se ocupa de las acciones a corto plazo, así como evitar obstáculos en tiempo real. Funciona a un nivel inferior al de la planificación global y se encarga de garantizar que el robot pueda desplazarse con seguridad por su entorno mientras sigue el plan global.

De la misma forma que el planificador anterior, usa los sensores para identificar obstáculos o cambios en el entorno en tiempo real. El planificador local genera acciones o maniobras inmediatas para evitar obstáculos o navegar alrededor de ellos de forma segura. Esto puede implicar técnicas reactivas como los campos potenciales mencionados anteriormente o métodos como el Control Predictivo de Modelos (MPC). La planificación local se centra en la situación inmediata y garantiza que el robot pueda adaptarse a entornos dinámicos.

En conjunto, la planificación global establece la estrategia a largo plazo y genera un plan de alto nivel, mientras que la planificación local se ocupa de las acciones a corto plazo y la evasión de obstáculos para ejecutar el plan en tiempo real. La combinación de estos dos componentes de planificación permite a los VAs realizar tareas complejas de forma autónoma y segura en un entorno dinámico.

La combinación de transformaciones homogéneas, la detección y seguimiento de personas, y la planificación de trayectorias es fundamental para el desarrollo de estrategias de interacción humano-robot basadas en la cortesía. Al utilizar transformaciones homogéneas, los VAs pueden comprender su propia posición relativa a las personas, lo que les permite coordinar sus movimientos de manera segura y eficiente en un entorno conocido. La detección y seguimiento de personas brindan una base sólida para la interpretación de las acciones de estas, ya que permite extraer las características que se desea investigar. La planificación de trayectorias garantiza que los robots puedan moverse sin colisionar en entornos compartidos. Estos conceptos y técnicas permiten que los VAs reconozcan ciertos patrones de comportamiento y reaccionen a los movimientos de las personas, lo que facilita una interacción segura y

respetuosa. En conjunto, estos aspectos se combinarán para construir una estrategia de interacción humano-robot basada en la cortesía. Estrategias que en un futuro permitirán que un vehículo pueda adaptarse a la percepción de las personas, mostrando respeto, consideración y cortesía al interactuar.

DISEÑO EXPERIMENTAL Y RESULTADOS

4

En este capítulo, se abordará el diseño experimental de la estrategia. Con el fin de establecer una metodología sistemática para replicar y validar la propuesta en un futuro, así como para analizar la eficacia de los algoritmos implementados.

La implementación de la estrategia se llevó a cabo durante una estancia de investigación en colaboración con la *Hamburg University of Applied Sciences (Hochschule für Angewandte Wissenschaften Hamburg)*. El propósito adicional de la estancia fue proponer una solución para realizar tareas de servicio en una cafetería. El proyecto se estructura mediante dos tareas fundamentales. En la primera, se implementó la estrategia con el fin de realizar la planificación para llegar a la zona de vertido, así como la detección de personas. La segunda subtarea se colaboró en el desarrollo de un control en lazo abierto para la muñeca de un brazo robótico. En conjunto, se usó un robot móvil *Husky* con un brazo robótico modelo *UR5* disponible en la HAW, además de un robot de servicio denominado *UVerto* en la Universidad Veracruzana, como se observa en la figura 4.1.

4.1

Implementación en ROS

ROS (*Robotic operation system*) es un conjunto de bibliotecas y herramientas de software de código abierto ampliamente utilizado en la comunidad robótica ya que proporciona una plataforma flexible y modular para la investigación en robótica. En este contexto permite la comunicación entre diferentes componentes de hardware y

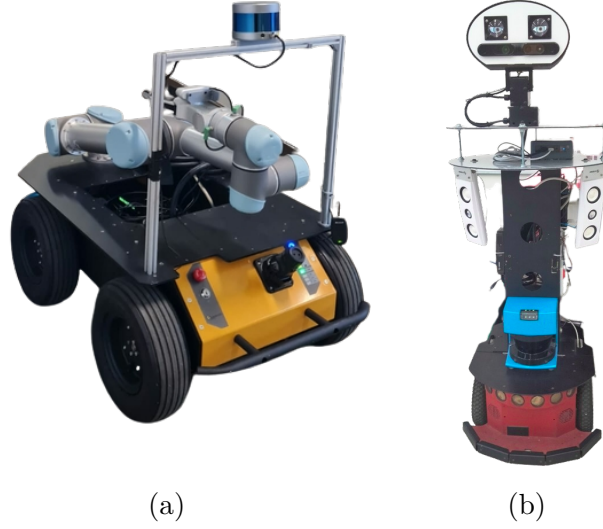


Figura 4.1: Configuración de los robots empleados, a) Husky y b)UVerTo.

software, la estandarización de paquetes (que contienen los algoritmos) y la capacidad de recopilar datos eficientemente, entre otros. La utilización de ROS en el diseño experimental permitió la colaboración y el intercambio de implementaciones con el grupo de investigación en Hamburgo.

4.1.1. Planificación de trayectorias con *movebase*

El paquete *movebase* proporciona una interfaz de ROS para configurar, ejecutar e interactuar con el *stack* de instrucciones para la navegación del vehículo. Dicho paquete enlaza el planificador global, local y genera un mapa de costo para representar las zonas accesibles o inaccesibles en el mundo real.

Para la realización del experimento se seleccionó un ambiente semi-controlado correspondiente a las instalaciones de la universidad ya que el vehículo tenía suficiente espacio para moverse. A través del paquete *gmapping* se creó un mapa de ocupación bidimensional del ambiente a partir de los datos láser y de odometría. Durante el proceso de escaneo se llevó a cabo una serie de recorridos en el entorno seleccionado, verificando cualitativamente la calidad del mapa generado y la correspondencia con el entorno real. El mapa generado y utilizado se observa en la figura 4.2. *Gmapping* utiliza un sistema de localización probabilístico para actualizar el mapa en tiempo real, a medida que se realiza la exploración.



Figura 4.2: Mapa de celdas generado. En color gris se muestran las escaneadas y en color negro los límites del mapa.

A partir del mapa de ocupación del entorno se genera un mapa de costos. En el cual, se asigna un valor numérico a cada celda de la imagen para indicar el costo o dificultad asociada a atravesar dicha región. Aunque cada celda puede tener 255 valores diferentes, el paquete sólo utiliza tres, definiendo así, cuando una celda puede estar: libre, ocupada o es desconocida. Por lo que se inflaron¹ los costos de los límites del mapa escaneado en función de un radio de inflación (equivalente a las dimensiones de Husky) como se muestra en la figura 4.3.

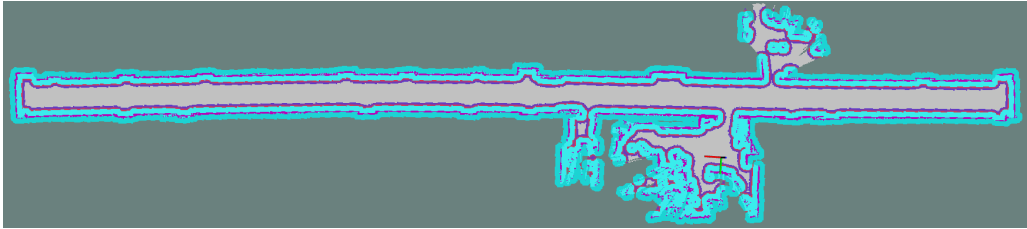


Figura 4.3: Mapa de costos del mapa global. En color gris se muestran las celdas no ocupadas, en violeta las ocupadas por obstáculos y en azul aquellas donde es probable una colisión.

Para delimitar la zona donde se realiza la detección de personas y se generan las regiones de interacción se definió un marco de referencia local con el robot en el centro. Utilizando los sensores del láser, se especificó el rango y en consecuencia un mapa de costos local. De modo que, a diferencia del mapa de costos global (estático), el mapa de costos local se actualiza en tiempo real como se observa en la figura 4.4.

La configuración final de *movebase* consta de los mapas de costos descritos anteriormente. Para generar una trayectoria global se usó el paquete *navfn* (*Navigation*

¹Operación referente al proceso de asignar celdas ocupadas.

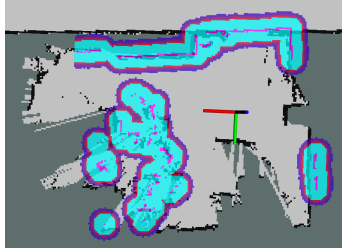


Figura 4.4: Mapa de costos local delimitado por el rango de sus sensores en un instante de tiempo. En color gris se muestran las celdas no ocupadas, en violeta las ocupadas por obstáculos y en azul aquellas donde es probable una colisión.

Function). Dicho planificador proporciona una función de navegación dada por una interpolación. Este asume que el vehículo es circular y opera sobre un mapa de costos, de modo que encuentra un plan de costo mínimo desde un punto inicial dado hasta un punto final en la cuadrícula. La función de navegación final se calculó usando el algoritmo de Dijkstra. Una vez que se ha calculado la trayectoria óptima, esta se publica en forma de *waypoints*.

Para generar una trayectoria local se usó el paquete de planificación conocido como DWA (*Dynamic Window Approach*). Dado el plan global anterior y el mapa de costos local, el proceso de planificación local permite desplazarse a través de *waypoints*. Este proceso utiliza la odometría para crear una función de valor y mediante una votación evaluar las posibles trayectorias. Lo que finalmente, permite establecer un control en velocidad que fue responsable de las modificaciones locales de la trayectoria.

4.1.2. Detectores implementados

Ya que se trata de un ambiente semi-controlado no fue posible controlar totalmente las condiciones relacionadas con la iluminación, distribución del espacio, etc. Sin embargo, se aseguró la ausencia de multitudes u objetos en movimiento, por lo que la interacción humano-robot se limitó a una sola persona.

Debido a las limitaciones del CPU del robot UVerto, con el que se desarrollaron los paquetes inicialmente, se optó por desarrollar un paquete en ROS que implemente el HOG (previamente entrenado) para la detección de personas. Debido al costo computacional de este algoritmo, se realizó un proceso de pre-procesamiento donde se convirtió la imagen a escala de grises, se aplicó un filtro gaussiano y se redujo de tamaño. El valor de los parámetros principales se encontró empíricamente, donde el

winStride (separación entre bloques) fue de 8×8 píxeles, el padding de 4×4 y la escala de 1.2. El objetivo del paquete fue publicar el centro de la caja delimitadora de la persona, i.e su posición en píxeles.



Figura 4.5: Comparativa entre detectores, HOG (izquierda) y YOLO (derecha).

Durante la estancia en la HAW, fue posible acceder a una GPU en la computadora de Husky, por lo que se optó por entrenar a *YOLOv8* con imágenes de los voluntarios. Este nuevo paquete realiza las mismas tareas, al suscribirse a la imagen RGB y publicar el centro de la caja delimitadora. La diferencia radica en el costo computacional y la precisión de las detecciones. Debido al uso de diferentes vehículos se emplearon ambos paquetes, el uso de ambos detectores se observa en la imagen 4.5.

Adicionalmente, se desarrolló un paquete para la transformación de coordenadas, este se suscribe al centro de la caja delimitadora (en píxeles) y publica la posición de la persona respecto al mundo aplicando la cadena de transformaciones descrita en el capítulo anterior. Una vez obtenida la posición de la persona respecto al mapa, se procesaron hasta cinco datos consecutivos para estimar la posición, velocidad y orientación. De esta forma se generó la elipse de acuerdo a la ecuación 3.13, para representarla se desarrolló un paquete que permite visualizar un marcador correspondiente a la región de interacción como se muestra en las figuras 4.6 y 4.7. Mediante la creación de un tercer paquete, fue posible establecer un mapa de costos que transformaba el marcador de la elipse generada en celdas de ocupación.

4.1.3. Simulación del ambiente

Previo a la realización de la estancia, se diseñó una simulación usando *Gazebo*, un *software* especializado en el uso de motores físicos para la simulación de cuerpos

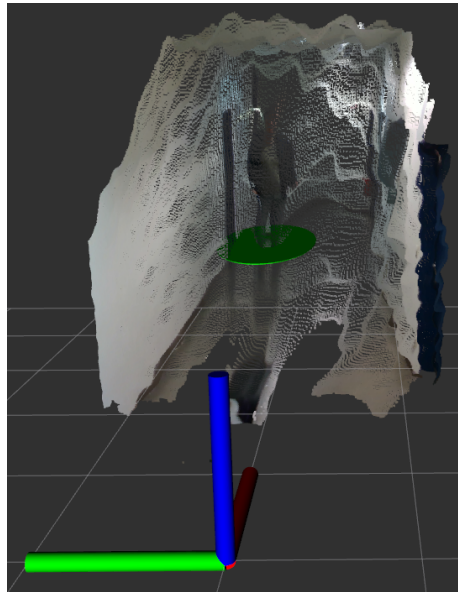


Figura 4.6: Visualización en tiempo real del marcador (en verde) que representa el espacio personal generado.

rígidos. Así como *Rviz*, un *software* enfocado en la visualización de mensajes generados por ROS. Ya que no se contaba con un mapa del lugar donde se realizarían los experimentos, se optó por diseñar un lugar alternativo donde se puedan probar los paquetes desarrollados como se muestra en la figura 4.8 (a). Dado que no fue posible añadir personas, los detectores se probaron en paralelo únicamente con ayuda de una cámara RGB-D. De modo que la simulación fue una zona de pruebas, donde se ejecutaron los paquetes principales (SLAM, planificación de trayectorias y transformación de coordenadas) antes de ser probados en el robot real.

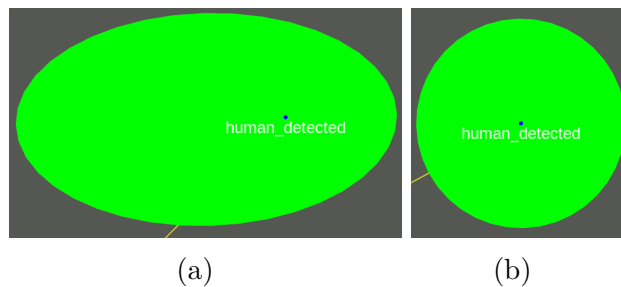
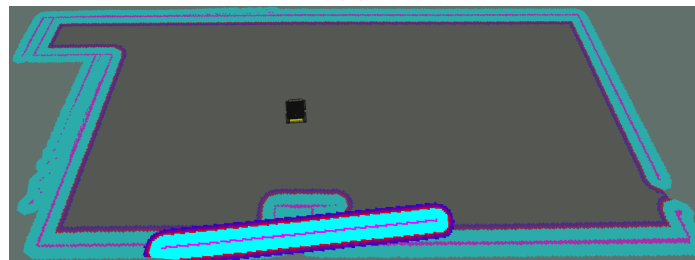


Figura 4.7: Marcadores. (a) Velocidad de 1.3 m/s. (b) Caso estático.



(a)



(b)

Figura 4.8: Simulación diseñada. Motor físico empleado (arriba) y Visualización de los sensores (abajo).

4.2 Estrategia empleada

Tras implementar los algoritmos para modelar la interacción y probar los paquetes en el simulador, se definió la estrategia que permite modificar la trayectoria del vehículo en función de los obstáculos detectados y la región de interacción correspondiente a la persona, como se muestra en la figura 4.9. La estrategia requiere de una meta global, correspondiente a la posición de vertido. Posteriormente, se genera el conjunto de *waypoints* que el vehículo debe visitar para llegar a la meta.

En todo momento, la cámara RGB-D del vehículo recibe retroalimentación del centro (u, v) de la caja delimitadora de la persona detectada, a partir de la cual se estima la posición $r(x, y, \theta)$ y velocidad. Desde la cual se genera el marcador correspondiente a la elipse $E(x, y, v, \theta)$ que representa la región de interacción. Finalmente el marcador se integra a un mapa de costos, por lo que la planificación no sólo toma en cuenta la geometría de los obstáculos dada por el láser del vehículo, sino que también añade y actualiza las regiones de interacción detectadas.

De modo que se propone usar las regiones elípticas dinámicas para modelar la interacción entre VAs y personas respetando su espacio personal al calcular la tra-

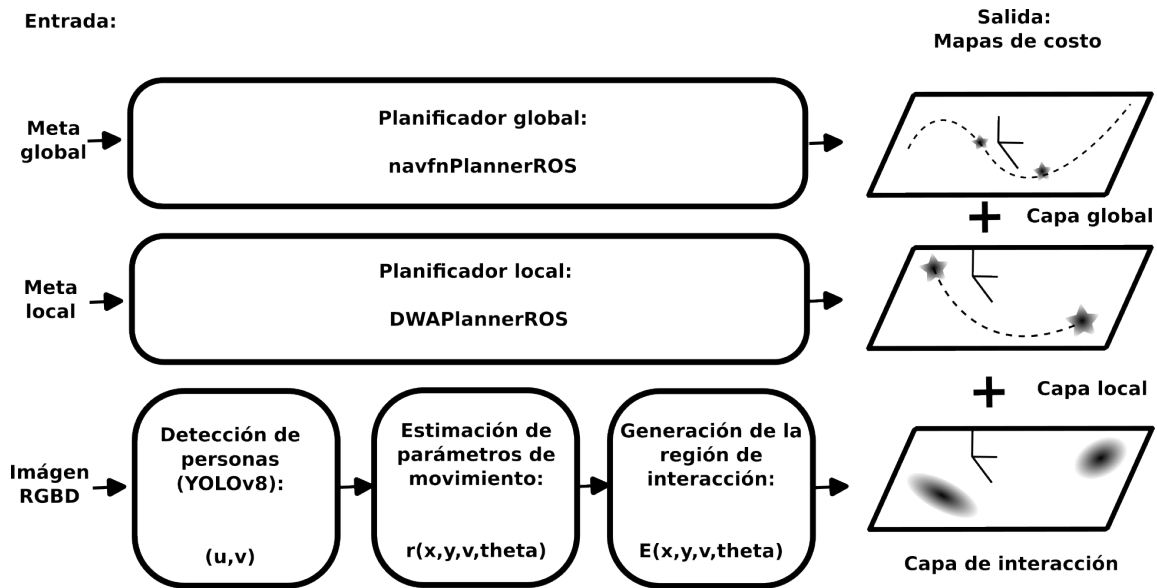


Figura 4.9: Esquema general de la estrategia empleada. Al combinar los mapas de costo, se obtiene un mapa total a partir del cual se calcula la trayectoria del vehículo.

yectoria en el mapa de costos como modo de cortesía.

4.3 Estructura del experimento

El experimento fue aprobado para realizarse en las instalaciones del departamento de cómputo de la HAW. Además, se obtuvo el consentimiento informado por parte de todas las personas involucradas. A partir de trabajos previos, se seleccionaron tres escenarios (A, B y C) con diferentes tipos de interacción. Se solicitó la participación de 5 voluntarios, en cada escenario se les indicó la meta hasta donde debían caminar, la cual daba fin al experimento. Cada voluntario realizó las tres pruebas utilizando un enfoque tradicional, así como la estrategia planteada que considera la región de interacción dinámica. Siguiendo la definición de cortesía de [17], se utilizó una encuesta (ver apéndice 5) que evalúa la cortesía a partir de tres métricas subjetivas; la confianza, satisfacción y el disfrute. Esta se aplicó al término de cada experimento.

Escenario A. El participante debía caminar en dirección contraria al robot, por lo que se enfoca en la interacción directa.

Escenario B. El participante debía caminar en la dirección frontal lateral del robot, sin obstruir su camino, hasta llegar a una meta. El enfoque era crear una

interacción indirecta.

Escenario C. El participante era libre de elegir el tipo de trayectoria con la que llegaba a la meta, pero debía excluir movimientos atípicos.

4.4 Proyecto integrador

Durante la estancia, fue posible colaborar en la tarea de vertido descrita al inicio de este capítulo. Si bien, este proceso no tiene una relación directa con la estrategia propuesta, fue un tarea en la que se colaboró en paralelo durante la estancia en la HAW. Partiendo de la ecuación de control 4.1 se desarrolló un paquete que controle la muñeca del robot hacia un ángulo específico mientras vierte agua de un contenedor cilíndrico. El objetivo principal fue encontrar experimentalmente el conjunto de parámetros que minimice o elimine por completo la generación de turbulencia durante el vertido.

$$\theta = \beta e^{-\lambda(t-\alpha)^2} \quad (4.1)$$

donde θ representa la posición de la muñeca en cada instante de tiempo, β representa el ángulo máximo de la muñeca, λ es la amplitud de la curva, α el desfase de la curva y t es el tiempo.

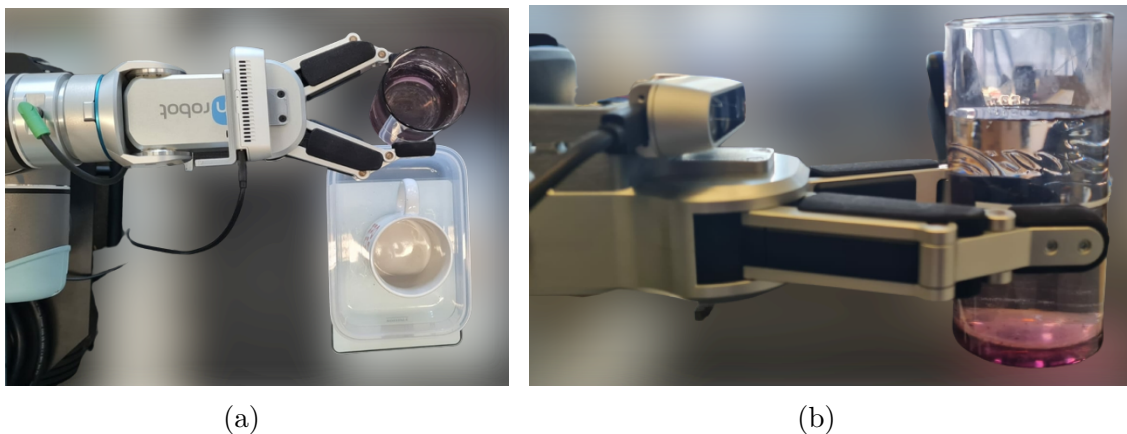


Figura 4.10: Diseño experimental del vertido. (a) Materiales utilizados para el vertido de líquido. (b) Posición inicial de la muñeca.

El proceso de vertido inicia cuando el robot llegó a la meta y se encuentra a una distancia estratégica frente a la mesa. Inicialmente, se manipula manualmente

el gripper para que sostenga un contenedor de base cilíndrica con características conocidas (dimensiones, volumen inicial, etc). Sobre la superficie de la mesa se colocó otro recipiente destinado a recibir el líquido vertido. Por último, se colocó una báscula debajo del recipiente receptor con el fin de registrar con precisión el peso del líquido vertido durante el proceso. Se diseñó el arreglo con el fin de tener una configuración estable y controlada que permita replicarse en el futuro.

Se propusieron los siguientes valores para cada uno de los parámetros de la ecuación 4.1

$\beta(rad)$	β°	λ	α
0.7853	45	0.2	2.0
1.0472	60	0.4	2.5
1.3090	75	0.6	3.0
1.5708	90	0.8	

Cuadro 4.1: Conjunto de valores para los parámetros de control.

Para cada experimento se considera, el mismo líquido con el mismo volumen. Después, se calcularon las 48 permutación posibles, resultando en 48 experimentos. Durante cada experimento, se registraron en una base de datos los parámetros θ , β , λ , α , t y el volumen vertido correspondiente a cada prueba realizada. Añadiendo una categoría para registrar si el conjunto de parámetros utilizados generó turbulencia durante el vertido.

4.5 Resultados

Con el fin de analizar las diferencias entre los detectores implementados, se marcaron trayectorias específicas, como líneas rectas, cuadrados y *zigzags*. Posteriormente, se realizaron pruebas, como se muestra en la figura 4.11 y 4.12, usando como *ground truth* las mediciones del piso registradas experimentalmente.

Dado que los robots usados tienen diferentes especificaciones y por lo tanto, distinto poder de cómputo, no es posible obtener una comparación cuantitativa precisa del procesamiento. Sin embargo, se observaron diferencias cualitativas significativas, donde:

- Considerando los parámetros usados, el modelo HOG consume hasta un 25% del CPU de UVerto pero ofrece una tasa de procesamiento de hasta 20Hz.,

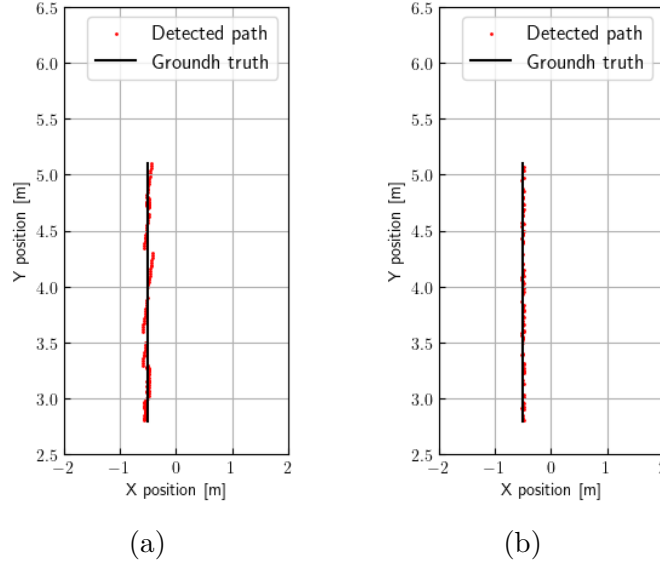


Figura 4.11: Comparativa entre la trayectoria obtenida por el modelo HOG (a) y YOLO (b).

- El uso del modelo YOLO requiere de una unidad GPU y ofrece tasas de procesamiento de hasta $10Hz$.
- A pesar de las consideraciones del HOG, las variaciones de iluminación (debido al contraste) son un factor determinante que en la práctica aumenta las detecciones erróneas.
- La localización dada por YOLO presenta menos ruido en comparación el HOG.

Por lo tanto, se seleccionó el modelo YOLOv8 como el algoritmo de detección que posteriormente permitió la localización de los participantes.

Tras realizar el experimento con el algoritmo de detección seleccionado, se aplicó el cuestionario, obteniendo una base de datos de todos los voluntarios. Dicha base contiene las respuestas de los voluntarios al probar los escenarios A, B y C con la configuración clásica y la propuesta. Posteriormente, se calculó el porcentaje obtenido para las 5 posibles respuestas, con el propósito de analizar las diferencias al usar la estrategia propuesta y al usar la configuración clásica, es decir, al usar el *stack* de navegación sin las regiones de interacción como se observa en las figuras 4.13 - 4.15.

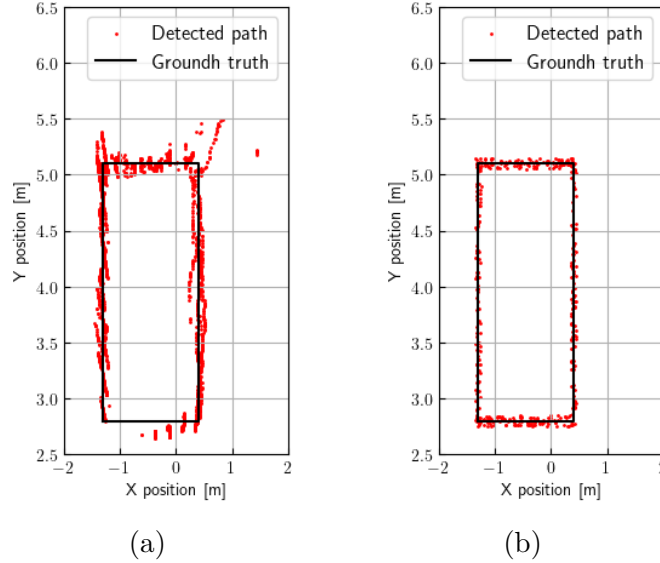


Figura 4.12: Comparativa entre la trayectoria obtenida por el modelo HOG (a) y YOLO (b).

Para analizar si existen diferencias significativas al usar la estrategia propuesta se aplicaron diferentes pruebas estadísticas. Con el fin de comparar la distribución de las respuestas al usar la configuración clásica y propuesta se aplicó la prueba de *Shapiro-Wilk* como se observa en el apéndice 5. De donde se determinó que:

- Al evaluar la confianza, los escenarios A y C de la configuración clásica siguen una distribución normal.
- Al evaluar la satisfacción, los escenarios C de la configuración clásica y C de la estrategia propuesta siguen una distribución normal.
- Al evaluar el disfrute, los escenarios B y C de la configuración clásica y C de la estrategia propuesta siguen una distribución normal.

Mientras que:

- Al evaluar la confianza, el escenario B de la configuración clásica y A,B y C de la estrategia propuesta no siguen una distribución normal.
- Al evaluar la satisfacción, los escenarios A y B de la configuración clásica y A de la estrategia propuesta no siguen una distribución normal.

- Al evaluar el disfrute, el escenario A de la configuración clásica y A y B de la estrategia propuesta no siguen una distribución normal.

Por lo que se aplicó la prueba de los rangos con signo de Wilcoxon para aquellos casos donde al comparar la configuración clásica y la estrategia propuesta al menos uno de los valores-p de la prueba de *Shapiro-Wilk* era menor a 0.05, i.e. no seguían una distribución normal. Así como, la prueba T para los casos donde ambos valores-p de la prueba de *Shapiro-Wilk* eran mayores a 0.05, i.e. siguen una distribución normal (ver apéndice 5).

Durante la implementación del cuestionario, se encontró que los participantes expresaban confusión sobre algunas de las preguntas empleadas. Además, debido al corto periodo en el que se realizó la estancia, no fue posible realizar el experimento en un mayor grupo de personas. Por lo que los resultados del experimento en los 5 participantes representan un estudio exploratorio sobre los efectos de la cortesía.

Al analizar las bases de datos generadas con el proyecto integrador se descartaron aquellas cuyos parámetros generen turbulencias, tengan un ángulo inicial lejano a cero o realicen el vertido en un periodo de tiempo ineficiente, como se observa en la figura 4.16. Al descartar este tipo de bases, se repitieron los experimentos de las bases restantes con el fin de observar si el sistema era estable o presentaba variaciones debido a la entropía. Se observó que, a pesar de emplear el mismo conjunto de parámetros y condiciones iniciales, existe una variación en el volumen de salida. Lo que sugiere la existencia de un intervalo de incertidumbre en el control de líquidos.

4.5.1. Gráficas

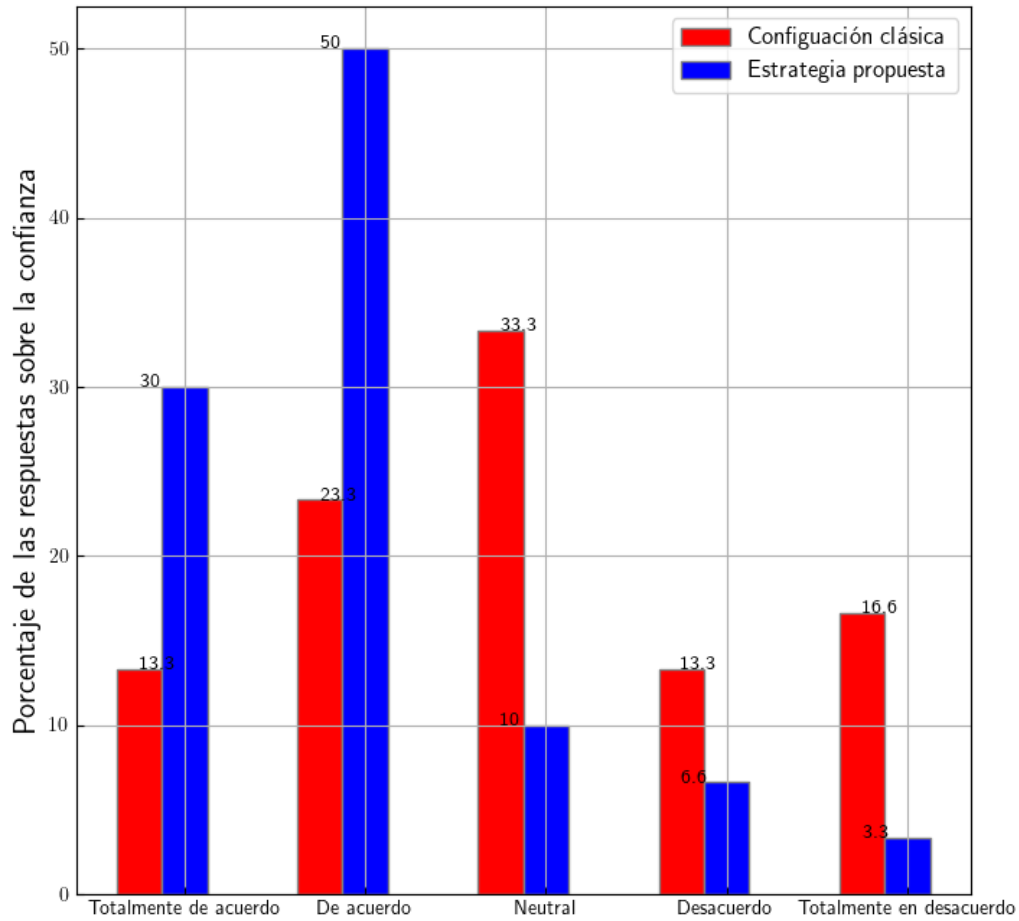


Figura 4.13: Porcentaje de respuestas relacionadas con la confianza al realizar las pruebas con el esquema clásico y con la estrategia implementada.

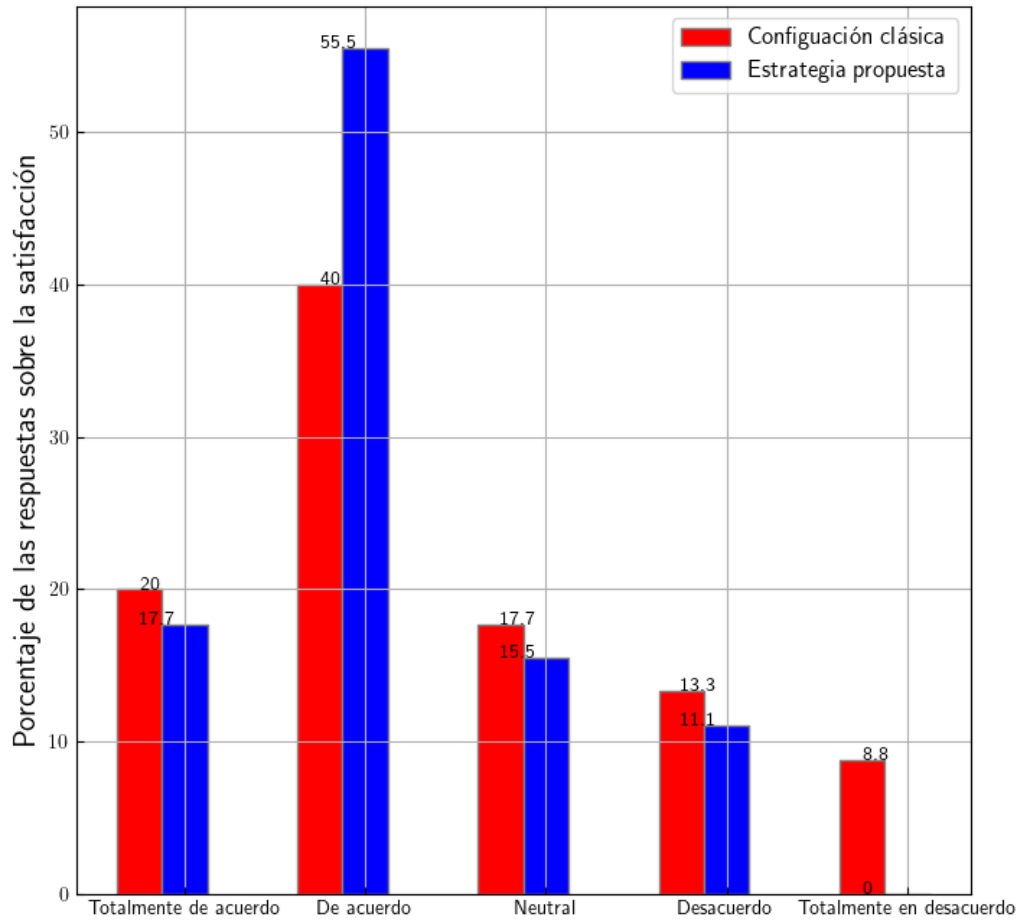


Figura 4.14: Porcentaje de respuestas relacionadas con la satisfacción de los participantes al realizar las pruebas con el esquema clásico y con la estrategia implementada.

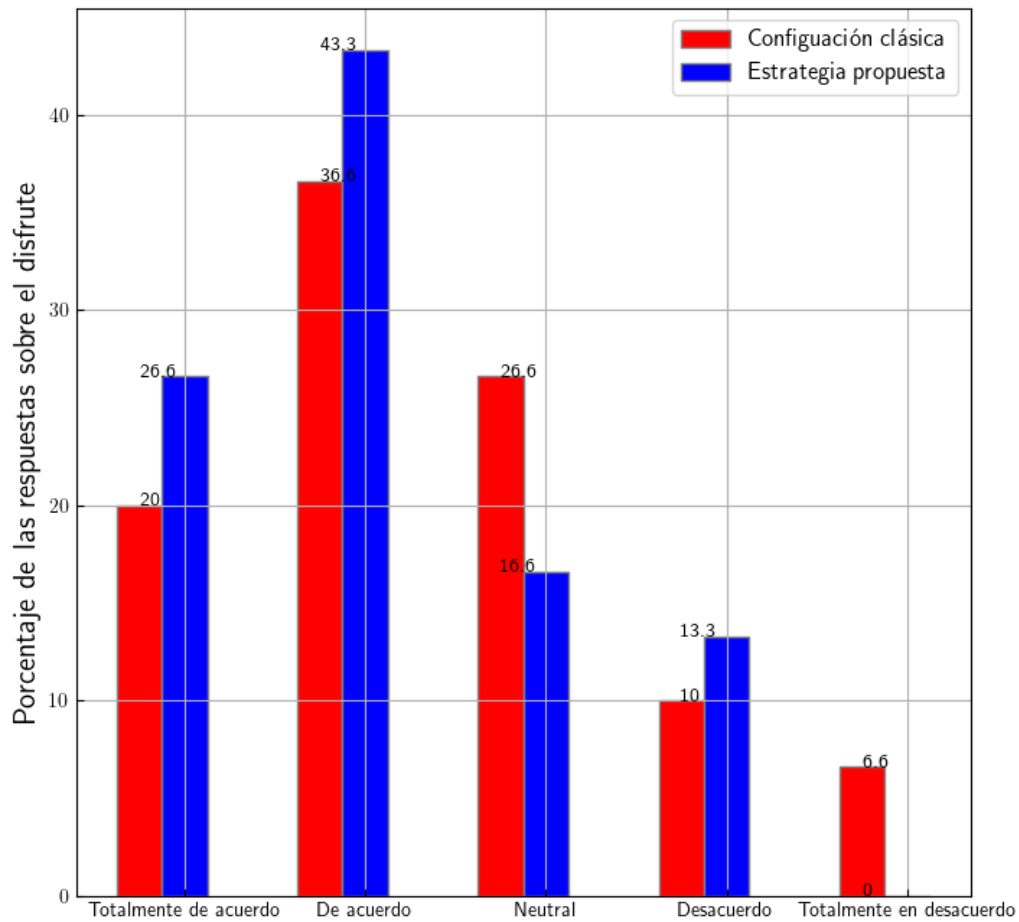
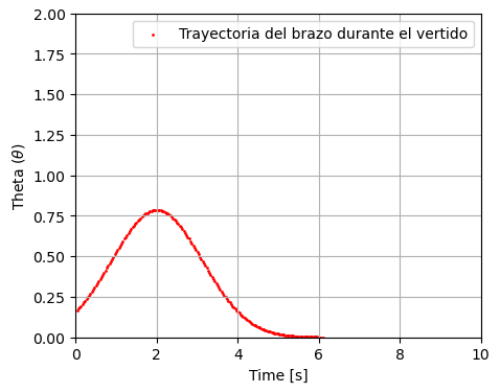
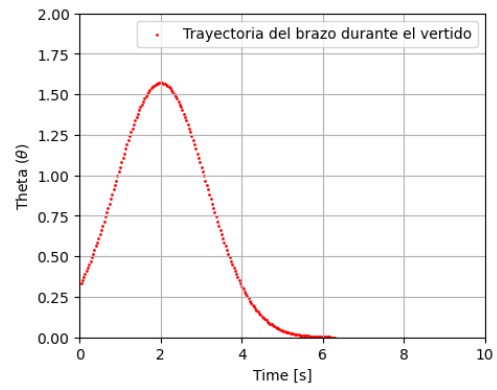


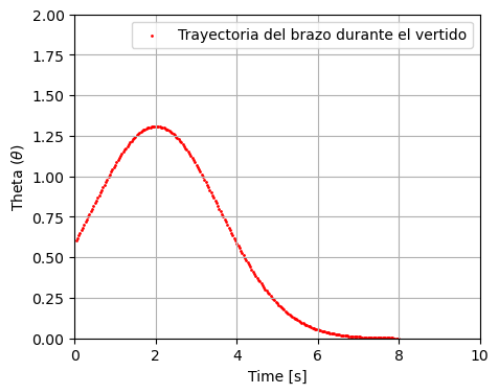
Figura 4.15: Porcentaje de respuestas relacionadas con el disfrute de los participantes al realizar las pruebas con el esquema clásico y con la estrategia implementada.



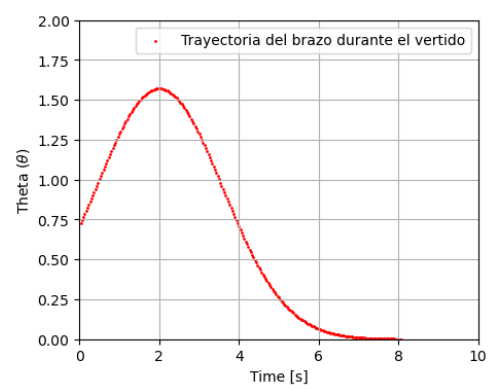
(a)



(b)



(c)



(d)

Figura 4.16: Variaciones de la ecuación de control en diferentes casos.

CONCLUSIONES Y TRABAJO FUTURO

5

En esta tesis, se desarrolló una estrategia de interacción basada en la detección de personas, el cálculo de su posición y velocidad, la generación de un espacio personal y la creación de un mapa de costos en el *framework* ROS. Los resultados obtenidos muestran la viabilidad y efectividad de esta estrategia en el contexto de la planificación de trayectorias. La información de la posición y velocidad, obtenida de los algoritmos de detección implementados, se utilizó para generar una región alrededor de cada persona detectada denominada como espacio personal.

Considerando las regiones mencionadas correspondientes a los espacios personales, las cuales se modifican en función de la velocidad y orientación de las personas, el robot puede considerar estas y modificar sus trayectorias para evitar invadir dichas regiones. Lo cual, se considera como un primer paso en la generación de estrategias de cortesía. La generación del mapa de costos en ROS permite integrar esta estrategia de interacción en el *stack* de navegación de cualquier vehículo autónomo basado en ROS.

Las pruebas estadísticas muestran que existe una diferencia significativa entre la configuración clásica y la estrategia propuesta en el escenarios B al evaluar la confianza (valor-p=0.047), así como en el escenario B y C al evaluar la satisfacción (valor-p=0.016 y valor-p=0.024 respectivamente). Estos sugiere que la estrategia empleada puede contribuir a mejorar la relación del individuo con el VA al navegar en espacios compartidos. Se espera que este enfoque tenga aplicaciones prácticas en entornos donde la interacción entre VAs y personas es cada vez más relevante.

Como trabajo futuro, se propone incrementar los ambientes, como podría ser dentro del hogar, en centros comerciales, campus escolares o parques. Así como incrementar las acciones de cortesía del robot, tales como la guía hacia sitios específicos, el dar instrucciones mediante comandos de voz, la disminución de la velocidad o el uso de luces *led*, con el fin de tener una evaluación mas significativa. También, se espera extender la interacción al estudio de multitudes, con el objetivo de realizar el experimento en ambientes o escenarios diversos.

Apéndice A: Cuestionario

¿En qué medida del uno al cinco estás de acuerdo con las siguientes afirmaciones sobre la interacción con el robot en este escenario? Donde 1 representa "totalmente en desacuerdo" y 5 representa "totalmente de acuerdo".

1. La interacción con el robot fue agradable. [Disfrute]
2. La interacción con el robot fue irritante. [Disfrute]
3. Durante el experimento, sentí que podía confiar en el robot. [Confianza]
4. Durante el experimento, me sentí cómodo con el robot. [Confianza]
5. Durante el experimento, sentí que el robot tenía en cuenta mis deseos. [Satisfacción]
6. Durante el experimento, sentí que el robot era sociable. [Satisfacción]
7. Estoy satisfecho con la forma en que el robot se ha comunicado conmigo [Satisfacción]

Apéndice B:

Resultados de la prueba de *Shapiro-Wilk*

Configuración clásica						
Métrica	Escenario A		Escenario B		Escenario C	
	Estadístico	Valor P	Estadístico	Valor P	Estadístico	Valor P
Confianza	0.960859	0.8139521	0.6840289	0.006470001	0.8282725	0.1350226
Satisfacción	0.5521817	0.0001309782	0.5521817	0.0001309782	0.8282725	0.1350226
Disfrute	0.5521817	0.0001309782	0.8208347	0.1185099	0.8208347	0.1185099

Cuadro 1: Resultados del valor estadístico y valor-p del experimento usando la configuración clásica.

Configuración propuesta						
Métrica	Escenario A		Escenario B		Escenario C	
	Estadístico	Valor P	Estadístico	Valor P	Estadístico	Valor P
Confianza	0.5521817	0.0001309782	0.6840289	0.006470001	0.7667168	0.04219866
Satisfacción	0.6840289	0.006470001	N/A	N/A	0.9140781	0.492481
Disfrute	0.6840289	0.006470001	0.5521817	0.0001309782	0.8810376	0.3140396

Cuadro 2: Resultados del valor estadístico y valor-p del experimento usando la configuración propuesta.

Apéndice C:

Análisis de diferencias significativas entre la configuración clásica y la propuesta.

Análisis usando Wilcoxon Signed Rank test						
Metrica	Escenario A		Escenario B		Escenario C	
	Estadístico	Valor P	Estadístico	Valor P	Estadístico	Valor P
Confianza	0	0.1814	0	0.04771	0	0.05676
Satisfacción	4	0.7728	N/A	N/A	N/A	N/A
Disfrute	0	0.3458	0	0.3711	N/A	N/A

Cuadro 3: Comparación entre el valor estadístico y el valor-p en los escenarios que no siguen una distribución normal.

Análisis usando T test						
Metrica	Escenario A		Escenario B		Escenario C	
	Estadístico	Valor P	Estadístico	Valor P	Estadístico	Valor P
Confianza	N/A	N/A	N/A	N/A	N/A	N/A
Satisfacción	N/A	N/A	-4	0.01613	-3.5	0.0249
Disfrute	N/A	N/A	N/A	N/A	-1.372	0.242

Cuadro 4: Comparación entre el valor estadístico y el valor-p en los escenarios que siguen una distribución normal.

BIBLIOGRAFÍA

- [1] Jesus Eduardo Hermosilla-Diaz, Antonio Marin-Hernandez, and Miguel Gonzalez-Mendoza. A proxemic potential field approach for modeling interactions between autonomous vehicles with pedestrians and cyclists. In *2022 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, volume 6, pages 1–6, 2022.
- [2] N. G. Hockstein, C. G. Gourin, R. A. Faust, and D. J. Terris. A history of robots: from science fiction to surgical robotics. *Journal of Robotic Surgery*, 1(2):113–118, Jul 2007.
- [3] Bruce G. Buchanan. A (very) brief history of artificial intelligence. *AI Magazine*, 26(4):53, Dec. 2005.
- [4] Lorissa Rinehart. Robot art and soccer, 2010. [Recuperado el 15 de febrero, 2023].
- [5] IEEE Spectrum. Atlas, 2016. [Recuperado el 15 de febrero, 2023].
- [6] C. Weaver. The new driver’s ed: Game developers teach cruise’s autonomous vehicles to understand gestures made by people on the street. *IEEE Spectrum*, 57(9):32–37, 2020.
- [7] IEEE Spectrum. Starship, 2015. [Recuperado el 16 de febrero, 2023].
- [8] Walter Brenner and Andreas Herrmann. *An Overview of Technology, Benefits and Impact of Automated and Autonomous Driving on the Automotive Industry*, pages 427–442. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018.
- [9] Jeffrey Delmerico, Stefano Mintchev, Alessandro Giusti, Boris Gromov, Kamilo Melo, Tomislav Horvat, Cesar Cadena, Marco Hutter, Auke Ijspeert, Dario Floreano, Luca M. Gambardella, Roland Siegwart, and Davide Scaramuzza. The

- current state and future outlook of rescue robotics. *Journal of Field Robotics*, 36(7):1171–1191, 2019.
- [10] Kichan Nam, Christopher S. Dutt, Prakash Chathoth, Abdelkader Daghfous, and M. Sajid Khan. The adoption of artificial intelligence and robotics in the hotel industry: prospects and challenges. *Electronic Markets*, 31(3):553–574, Sep 2021.
- [11] Tian-Miao Wang, Yong Tao, and Hui Liu. Current researches and future development trend of intelligent robot: A review. *International Journal of Automation and Computing*, 15(5):525–546, Oct 2018.
- [12] Udo Pesch. Who’s driving innovation? new technologies and the collaborative state, 2021.
- [13] Casey Weaver. The new driver’s ed: Game developers teach cruise’s autonomous vehicles to understand gestures made by people on the street. *IEEE Spectrum*, 57(9):32–37, 2020.
- [14] Territorial y Urbano Secretaría de Desarrollo Agrario. *Manual de calles: diseño vial para ciudades mexicanas*. Coordinadora Editorial SEDATU, México, D.F, 2019.
- [15] Yue Hu, Mehdi Benallegue, Gentiane Venture, and Eiichi Yoshida. Interact with me: An exploratory study on interaction factors for active physical human-robot interaction. *IEEE Robotics and Automation Letters*, 5(4):6764–6771, 2020.
- [16] Neziha Akalin, Annica Kristoffersson, and Amy Loutfi. Do you feel safe with your robot? factors influencing perceived safety in human-robot interaction based on subjective and objective measures. *International Journal of Human-Computer Studies*, 158:102744, 2022.
- [17] Shikhar Kumar, Eliran Itzhak, Yael Edan, Galit Nimrod, Vardit Sarne-Fleischmann, and Noam Tractinsky. Politeness in human-robot interaction: A multi-experiment study with non-humanoid robots. *International Journal of Social Robotics*, 14(8):1805–1820, Oct 2022.
- [18] Leimin Tian and Sharon Oviatt. A taxonomy of social errors in human-robot interaction. *J. Hum.-Robot Interact.*, 10(2), feb 2021.

- [19] Angeliki Zacharaki, Ioannis Kostavelis, Antonios Gasteratos, and Ioannis Dokas. Safety bounds in human robot interaction: A survey. *Safety Science*, 127:104667, 2020.
- [20] Leigh Clark. Social boundaries of appropriate speech in hci: A politeness perspective. *Bond, R. (ed.). Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI 2018)*, 2018.
- [21] Sezgin Secil and Metin Ozkan. Minimum distance calculation using skeletal tracking for safe human-robot interaction. *Robotics and Computer-Integrated Manufacturing*, 73:102253, 2022.
- [22] Dylan P. Losey and Marcia K. O’Malley. Trajectory deformations from physical human–robot interaction. *IEEE Transactions on Robotics*, 34(1):126–138, 2018.
- [23] Edward T. Hall, Ray L. Birdwhistell, Bernhard Bock, Paul Bohannon, A. Richard Diebold, Marshall Durbin, Munro S. Edmonson, J. L. Fischer, Dell Hymes, Solon T. Kimball, Weston La Barre, , J. E. McClellan, Donald S. Marshall, G. B. Milner, Harvey B. Sarles, George L Trager, and Andrew P. Vayda. Proxemics [and comments and replies]. *Current Anthropology*, 9(2/3):83–108, 1968.
- [24] Michael L Walters, Kerstin Dautenhahn, René Te Boekhorst, Kheng Lee Koay, Dag Sverre Syrdal, and Chrystopher L Nehaniv. An empirical framework for human-robot proxemics. *Procs of new frontiers in human-robot interaction*, 2009.
- [25] Jonathan Mumm and Bilge Mutlu. Human-robot proxemics: Physical and psychological distancing in human-robot interaction. In *Proceedings of the 6th International Conference on Human-Robot Interaction, HRI ’11*, page 331–338, New York, NY, USA, 2011. Association for Computing Machinery.
- [26] Ross Mead and Maja J Matarić. Proxemics and performance: Subjective human evaluations of autonomous sociable robot distance and social signal understanding. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5984–5991, 2015.
- [27] Mohammad Obaid, Eduardo B. Sandoval, Jakub Złotowski, Elena Moltchanova, Christina A. Basedow, and Christoph Bartneck. Stop! that is close enough. how body postures influence human-robot proximity. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 354–361, 2016.

- [28] Paulo Leica, Juan Marcos Toibero, Flavio Roberti, and Ricardo Carelli. Switched control to robot-human bilateral interaction for guiding people. *Journal of Intelligent & Robotic Systems*, 77(1):73–93, Jan 2015.
- [29] Xuan-Tung Truong, Voo Nyuk Yoong, and Trung-Dung Ngo. Dynamic social zone for human safety in human-robot shared workspaces. In *2014 11th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pages 391–396, 2014.
- [30] Xuan-Tung Truong and Trung-Dung Ngo. Dynamic social zone based mobile robot navigation for human comfortable safety in social environments. *International Journal of Social Robotics*, 8(5):663–684, Nov 2016.
- [31] Ross Mead and Maja J. Matarić. Autonomous human–robot proxemics: socially aware navigation based on interaction potential. *Autonomous Robots*, 41(5):1189–1201, Jun 2017.
- [32] Weihua Chen, Tie Zhang, and Yanbiao Zou. Mobile robot path planning based on social interaction space in social environment. *International Journal of Advanced Robotic Systems*, 15(3):1729881418776183, 2018.
- [33] Dirk Helbing and Péter Molnár. Social force model for pedestrian dynamics. *Phys. Rev. E*, 51:4282–4286, May 1995.
- [34] Photchara Ratsamee, Yasushi Mae, Kazuto Kamiyama, Mitsuhiro Horade, Masaru Kojima, and Tatsuo Arai. Social interactive robot navigation based on human intention analysis from face orientation and human path prediction. *ROBOMECH Journal*, 2(1):11, Aug 2015.
- [35] Bani Anvari and Helge A Wurdemann. Modelling social interaction between humans and service robots in large public spaces. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 11189–11196, 2020.
- [36] Oussama Khatib. *The Potential Field Approach And Operational Space Formulation In Robot Control*, pages 367–377. Springer US, Boston, MA, 1986.
- [37] Atsuki Toyoshima, Nozomi Nishino, Daisuke Chugo, Satoshi Muramatsu, Sho Yokota, and Hiroshi Hashimoto. Autonomous mobile robot navigation: Consideration of the pedestrian’s dynamic personal space. In *2018 IEEE 27th International Symposium on Industrial Electronics (ISIE)*, pages 1094–1099, 2018.

- [38] Yosuke Kawasaki, Ayanori Yorozu, Masaki Takahashi, and Enrico Pagello. A multimodal path planning approach to human robot interaction based on integrating action modeling. *Journal of Intelligent & Robotic Systems*, 100(3):955–972, Dec 2020.
- [39] L.V. Calderita, A. Vega, P. Bustos, and P. Núñez. A new human-aware robot navigation framework based on time-dependent social interaction spaces: An application to assistive robots in caregiving centers. *Robotics and Autonomous Systems*, 145:103873, 2021.
- [40] Pakpoom Patompak, Sungmoon Jeong, Itthisek Nilkhamhang, and Nak Young Chong. Learning proxemics for personalized human–robot social interaction. *International Journal of Social Robotics*, 12(1):267–280, Jan 2020.
- [41] Jin Soo Park, Brian Tsang, Harel Yedidsion, Garrett Warnell, Daehyun Kyoung, and Peter Stone. Learning to improve multi-robot hallway navigation. In Jens Kober, Fabio Ramos, and Claire Tomlin, editors, *Proceedings of the 2020 Conference on Robot Learning*, volume 155 of *Proceedings of Machine Learning Research*, pages 1883–1895. PMLR, 16–18 Nov 2021.
- [42] Frank Broz, Illah Nourbakhsh, and Reid Simmons. Planning for human–robot interaction in socially situated tasks. *International Journal of Social Robotics*, 5(2):193–214, Apr 2013.
- [43] Rosemarie E. Yagoda and Michael D. Covert. How to work and play with robots: An approach to modeling human–robot interaction. *Computers in Human Behavior*, 28(1):60–68, 2012.
- [44] Edward T. Hall, Ray L. Birdwhistell, Bernhard Bock, Paul Bohannon, A. Richard Diebold, Marshall Durbin, Munro S. Edmonson, J. L. Fischer, Dell Hymes, Solon T. Kimball, Weston La Barre, , J. E. McClellan, Donald S. Marshall, G. B. Milner, Harvey B. Sarles, George L Trager, and Andrew P. Vayda. Proxemics [and comments and replies]. *Current Anthropology*, 9(2/3):83–108, 1968.
- [45] Cong Wang, Qifeng Zhang, Qiyan Tian, Shuo Li, Xiaohui Wang, David Lane, Yvan Petillot, and Sen Wang. Learning mobile manipulation through deep reinforcement learning. *Sensors*, 20(3), 2020.
- [46] Juan Terven and Diana Cordova-Esparza. A comprehensive review of yolo: From yolov1 to yolov8 and beyond. *arXiv preprint arXiv:2304.00501*, 2023.

- [47] Agnivesh Pani, Sabya Mishra, Mihalis Goliass, and Miguel Figliozzi. Evaluating public acceptance of autonomous delivery robots during covid-19 pandemic. *Transportation Research Part D: Transport and Environment*, 89:102600, 2020.
- [48] Robert Bogue. The role of artificial intelligence in robotics. *Industrial Robot: An International Journal*, 41(2):119–123, Jan 2014.
- [49] J. Rios-Martinez, A. Spalanzani, and C. Laugier. From proxemics theory to socially-aware navigation: A survey. *International Journal of Social Robotics*, 7(2):137–153, Apr 2015.
- [50] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893 vol. 1, 2005.
- [51] Manas Kamal Bhuyan. *Computer vision and image processing: Fundamentals and applications*. CRC Press, 2019.