

Navegação com restrições sociais com base em uma percepção robocêntrica *

Fábio Leite * Manoel Andrade * Gabriel Alves * Daniel **
José Díaz-Amado * Raphaell Sousa ** Dennis Barrios-Aranibar ***

* Engenharia Elétrica, Instituto Federal de Educação, Ciência e
Tecnologia da Bahia - IFBA, Vitória da Conquista, BA,
(fabioabdon97@gmail.com, manoelpedro@ifba.edu.br,
leirbaggabriel07@gmail.com, jose_diaz@ifba.edu.br).

** Instituto Federal de Educação, Ciência e Tecnologia da Paraíba -
IFPB, Cajazeiras, PB, (almeida.daniel@academico.ifpb.edu.br,
raphaell.sousa@ifpb.edu.br)

*** Departamento de Engenharia Elétrica e Eletrônica, Universidade
Católica San Pablo, Perú (dbarrios@ucsp.edu.pe)

Abstract: In the last few years, the application of robotics in environments common to human beings has been increasing. There exists several proposals of socially aware navigation frameworks for mobile robots, however all of them generally focused in a specific aspect of social relationships between humans and robots; so, there exist a lack of approaches that integrate all the aspects related to social navigation. This work aims to propose an autonomous navigation framework based on the integration of social perception elements (from a robocentric perspective) with proxemics modelling, considering the presence of human beings and the perception of their needs, feelings or intentions. We verified the feasibility of our approach by implementing it in ROS and Gazebo, and making a qualitative evaluation of its performance in two simulated scenarios where we included people with different feelings about robot presence, that triggered changes in the path planned by the robot in real time. So, it was concluded that this framework is feasible for implementing social navigation in mobile robots.

Keyword: Adjustment Social Robots; Emotion Analysis; Framework; Proxemic; Social Navigation;

Resumo: Nos últimos anos, as aplicações da robótica em ambientes comuns aos seres humanos tem aumentado. Existem várias propostas de *frameworks* de navegação socialmente conscientes para os robôs móveis, no entanto todas elas se centraram geralmente num aspecto específico das relações sociais entre seres humanos e robôs; por isso, existe uma falta de abordagens que integrem todos os aspectos relacionados com a navegação social. Este trabalho visa propor um *framework* de navegação autônoma baseado na integração de elementos de percepção social (de uma perspectiva robocêntrica) com modelagem proxêmica, considerando a presença de seres humanos e a percepção das suas necessidades, sentimentos ou intenções. Verificamos a viabilidade da nossa abordagem, implementando-a em ROS e Gazebo, e efetuando uma avaliação qualitativa do seu desempenho em dois cenários simulados, onde incluímos pessoas com diferentes emoções em relação à presença de robôs, que desencadearam mudanças no caminho planejado pelo robô em tempo real. Assim, concluiu-se que este *framework* é viável para a implementação da navegação social em robôs móveis.

Palavras-chaves: Análise de Emoções; Framework; Navegação Social; Proxêmica; Robôs Sociais de Ajuste;

1. INTRODUÇÃO

Nos últimos anos, a aplicação da robótica em ambientes comuns aos seres humanos tem aumentado. Seja para ajudar em atividades de rotina ou em aplicações industriais, os robôs têm sido inseridos em contextos de coexistência entre grupos de pessoas, o que tem estimulado o desenvolvimento de trabalhos sobre robôs sociais e as suas relações com os seres humanos. Para que esta integração ocorresse de uma forma mais natural, muitos trabalhos foram desenvolvidos em diferentes componentes sobre interação

* Esta pesquisa foi apoiada financeiramente pela Chamada Pública 05/2020, Seleção de projetos de apoio ao empreendedorismo inovador com foco na Economia 4.0, pelo FONDO NACIONAL DE DESARROLLO CIENTÍFICO, TECNOLÓGICO Y DE INNOVACIÓN TECNOLÓGICA - FONDECYT como entidade executora do CONCYTEC sob o convênio nº. 01-2019-FONDECYT-BM-INC.INV no projeto RUTAS: RUTAS: Robots for Urban Tourism Centers, Autonomous and Semantic-based e pelo PIBIC/PIBITI

social e a navegação; por exemplo, existem na literatura abordagens para a percepção da presença humana e das suas características sociais e costumes (Nin and Osório, 2011); (Cunhal, 2014); (Redmon et al., 2016); (Ginés et al., 2019), para a modelação de zonas proxémicas em humanos e robôs (Patompak et al., 2020); (Ginés et al., 2019) e para a navegação considerando restrições sociais (Daza et al., 2021); (Cunhal, 2014).

Em ambientes com muitas pessoas, a trajetória dos robôs é restringida por espaços ocupados por humanos (zonas proxémicas) ou interações entre grupos de pessoas. Assim, é fundamental considerar as restrições para a navegação autónoma em ambientes que contenham humanos (Daza et al., 2021). Assim, o primeiro passo da implementação de tal navegação está relacionado com o sensoriamento e percepção, onde os robôs devem ser capazes de identificar obstáculos (Nin and Osório, 2011) e classificá-los, tendo em conta a sua natureza (diferenciar objetos dos humanos) e prováveis interações com o robô e vice-versa (Redmon et al., 2016). Quando um “objeto” no espaço de trabalho do robô é reconhecido como sendo uma pessoa, são necessários novos tipos de processos, como a análise das emoções, que podem fornecer características que ajudam a dar contexto às decisões. Desta maneira, a informação extraída pode ser utilizada para traçar caminhos que tenham em conta este tipo de características (Cunhal, 2014). Por exemplo, se após a análise das emoções for identificado que uma certa pessoa, ou um grupo de pessoas está expressando medo devido à abordagem do robô, pode ser mais interessante utilizar um caminho que evite tanto quanto possível o espaço próximo dos indivíduos referidos; por outro lado, se as pessoas tiverem expressões mais felizes, podem ser mais receptivas a uma abordagem dentro dos limites do espaço pessoal ou do grupo (Ginés et al., 2019).

Algumas obras inseriram elementos de percepção integrados em técnicas autónomas de navegação (recorrentes na literatura), tais como em (Nin and Osório, 2011) onde os autores incluíram seres humanos e desvio de objectos para melhorar a navegação, ou em (Cunhal, 2014) onde o autor utiliza a detecção de expressão facial para definir as ações que o robô irá realizar (inserindo elementos culturais no planeamento da trajetória), para que o robô possa mover-se obedecendo às regras sociais do ambiente em que foi inserido. Além disso, essa integração deve ser baseada na modelagem de espaços de interação (zonas proxémicas) que são normalmente incluídos no melhoramento da navegação social. Por exemplo, alguns autores investigaram como o espaço individual de cada pessoa pode ser afectado em situações de interações sociais (Patompak et al., 2020). Também, em (Daza et al., 2021), os autores apresentaram um estudo sobre as zonas proxémicas (ZPs) nas interações entre robôs e humanos, de modo que é possível notar como as ZPs dependem de fatores externos, tais como o grau de intimidade entre as pessoas, ou mesmo a dependência do ambiente em que os indivíduos estão inseridos. Além disso, em (Ginés et al., 2019) os autores fizeram considerações sobre outros aspectos que afetam o espaço individual, como no caso dos sons emitidos pelos robôs e os seus efeitos no distanciamento social. Existe também uma relação entre a proxémica e a atividade que a pessoa está realizando no momento. Finalmente, é importante notar que as zonas proxémicas foram modeladas usando Gaussianas, que

podem ser simétricas ou assimétricas, como mostra (Pérez-Higueras et al., 2014). Em (Daza et al., 2021) os autores introduzem a definição de espaço social de grupos de pessoas por meio da intersecção entre gaussianas individuais.

Como vimos anteriormente, vários autores contribuíram com o desenvolvimento de técnicas que utilizavam a percepção do robô para interpretar algum elemento relacionado com o contexto social em que o robô está inserido, integrando estes dados aos algoritmos clássicos utilizados na robótica móvel autónoma. No entanto, existe uma falta de abordagens que integrem todos os aspectos relacionados com a navegação social. Neste sentido, este trabalho visa propor um sistema de navegação autónoma baseado na integração de elementos de percepção social. Para este fim, o *fremework* proposto utiliza a análise dos sentimentos como método para alterar os parâmetros de uma gaussiana assimétrica que é utilizada para descrever ZPs de cada indivíduo detectado pelo sistema de visão por computador. Propomos implementar a detecção de pessoas e a análise de emoções utilizando um modelo neural pré-treinado (Deep Learning). Esta informação é então utilizada por um algoritmo tradicional de planeamento de percursos (e.g. A*). Os resultados da simulação (utilizando ROS e Gazebo) do quadro proposto são apresentados para validar a sua viabilidade.

Este trabalho está organizado da seguinte forma: a Seção 2 apresenta trabalhos relacionados com o âmbito desta proposta, depois a Seção 3 apresenta uma descrição da proposta apresentada neste trabalho. Finalmente, as Seções 4 e 5 apresentam os resultados e conclusões obtidos, respectivamente.

2. ESTADO DA ARTE

Existem várias propostas de *fremework* de navegação socialmente conscientes para os robôs móveis, contudo todas elas se centraram geralmente num aspecto específico das relações sociais entre humanos e robôs. Por exemplo, Kose and Yavuz (2021) propuseram um *fremework* de navegação social que inclui a capacidade de ter em conta o movimento das pessoas (utilizando o chamado modelo de Força Social baseada na Previsão de Colisão), a fim de incluir o movimento das pessoas como uma restrição na navegação dos robôs. Além disso, Hurtado and Valada (2021) propuseram um *fremework* para os robôs aprenderem o comportamento social a partir de observações ou demonstrações humanas, com especial enfoque na diminuição de preconceitos no comportamento social dos robôs, uma vez que no comportamento social humano é possível existir injustiça social, tal como discriminação e segregação e os robôs poderiam replicá-las, promovê-las e amplificá-las. Além disso, Mavrogiannis and Knepper (2022) propuseram um quadro para a navegação social chamado Momentum Social que incorpora no planeamento de percursos a forma como os humanos podem perceber e reagir ao movimento dos robôs com o chamado valor do Momentum Social baseado na estimativa dos protocolos de evasão mais prováveis de outros com base nos seus comportamentos de movimento passados. Che and Sadigh (2020) propuseram uma estrutura com uma ideia semelhante ao Momentum Social que utiliza uma combinação de implícito (movimento robô) e explícito (*feedback* visual/áudio/háptico) a fim de apro-

ximar movimentos contínuos e modos de comportamento discretos na navegação humana. Finalmente, Pham et al. (2021) propuseram um enquadramento para uma navegação socialmente consciente em robôs baseada na análise da pose humana através da observação do esqueleto humano para prever atividades sociais humanas, incluindo corrida, caminhada, parado em pé, sentado e deitado.

Apesar de, na literatura, ser possível encontrar muitos *fremework* para a navegação social, a maioria deles baseia-se na detecção da presença/ausência de seres humanos ou grupos de seres humanos e a sua intenção de movimento. Assim, sendo importante incluir outros elementos como a compreensão do contexto geral em que a dinâmica social se desenrola. Neste sentido, alguns trabalhos utilizam a fusão de sensores para obter dados mais robustos e precisos. A fusão de sensores é uma técnica que tem sido comumente aplicada em automóveis autônomos para localizar ou detectar obstáculos (Kumar et al., 2020). Além disso, algumas obras utilizaram técnicas de aprendizagem de máquinas e especificamente algoritmos de Deep Learning para diferenciar humanos e objetos comuns (Kim et al., 2019); (Lu et al., 2020).

Outro aspecto esquecido nas estruturas de navegação social, mas amplamente estudado na área da interação homem-robô é a análise de emoções. Paul Ekman, um psicólogo pioneiro no estudo das expressões faciais, através da sua pesquisa descobriu que quando se sente uma das 7 emoções básicas como a felicidade, por exemplo, são sempre ativados os mesmos músculos faciais independentemente da cultura, etnia ou idade (Ekman, 1999). Desta forma, através da visão por computador, é possível verificar se os músculos faciais que se movem com uma determinada emoção são ou não ativados. Assim, ao detectar estes movimentos nas chamadas zonas de ação (partes do rosto que se movem com determinados músculos), é possível definir qual a emoção que a pessoa está sentindo. Após esta identificação, os dados da análise das emoções podem ser utilizados para modificar parâmetros relacionados com as dimensões das zonas proximicas. Isto significa, melhorar os algoritmos de navegação.

Características culturais, padrões espaciais pessoais (zonas proximicas), regras de etiqueta, entre outros, são alguns dos parâmetros relacionados com a interação entre os seres humanos e os robôs autônomos. Relativamente à proximica (que é o conceito mais utilizado nos *fremeworks* de navegação social), o espaço pessoal é basicamente estudado, de modo que os limites teóricos são considerados a partir das relações de distância entre indivíduos, criando assim zonas de conforto e espaços pessoais que devem ser intuitivamente respeitados. A grosso modo, verifica-se que as pessoas tendem a sentir-se mais confortáveis nas interações em que os limites pessoais são respeitados. Isto acontece de uma forma antagônica ao desconforto causado quando alguém se aproxima demasiado perto de outra pessoa.

A proximica aplicada à robótica visa melhorar a interação entre humanos e robôs. Para isso é necessário que os robôs respeitem as chamadas zonas proximicas de cada indivíduo no ambiente (Hall, 1963). Para manter este equilíbrio, existem várias formas de determinar as zonas proximicas de pessoas num ambiente. Aghaei et al. (2020) aplicaram

a visão computacional para estimar estes espaços teóricos. A técnica utilizada chama-se “homografia”, na qual a inclinação do terreno sob as pessoas presentes é identificada e as zonas proximicas com tamanhos pré-determinados são desenhadas na imagem.

Do ponto de vista de como as pessoas interpretam o espaço pessoal do robô, Adam Rojik e Matej Hoffmann (Lehmann et al., 2020) realizaram um estudo sobre como as pessoas compreendem as zonas de interação com os robôs tendo em conta as suas dimensões. Neste estudo, os autores discutem como o pequeno tamanho do robô não faz com que as pessoas compreendam intuitivamente que podem se aproximar dele. Com base nos resultados deste trabalho, de acordo com os autores, as pessoas mantiveram uma distância social padrão, então essa distância não é proporcional ao tamanho do robô.

As aplicações da proximica na determinação do espaço pessoal também tentam relacionar como as características emocionais das pessoas podem influenciar a determinação do espaço pessoal. (Ginés et al., 2019) desenvolveram um modelo em que as zonas proximicas são estimadas com base na expressão facial do indivíduo. Utilizando o algoritmo detector de emoções (Quiroz et al., 2022), é possível analisar as emoções das pessoas. Com base nesta informação, podem ser determinados espaços pessoais adaptáveis, dependendo dos elementos de percepção.

3. PROPOSTA

Este artigo tem como principal proposta o desenvolvimento da navegação social considerando a presença dos seres humanos e a percepção de suas necessidades, sentimentos ou intenções. Para alcançar esse objetivo, foi desenvolvido um diagrama de blocos, representado pela Figura 1, onde propomos um conjunto básico de etapas para definir uma navegação social autônoma. Para chegar a essa navegação, foi necessário implementar diferentes objetivos, como filtragem dos dados provenientes dos sensores, fusão de informações, representação das zonas proximicas e finalmente traçar uma rota utilizando a perspectiva visual robocêntrica. Em seguida, vamos explicar cada componente do quadro proposto.

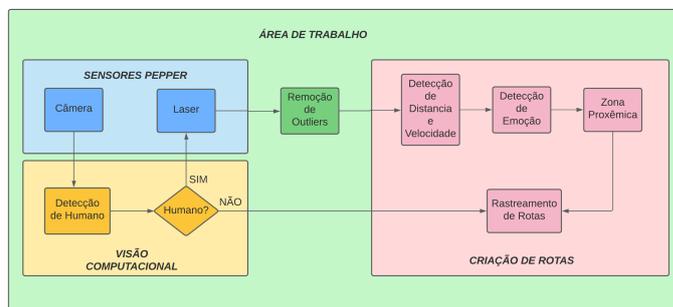


Figura 1. Diagrama de blocos das etapas necessárias para uma estrutura de navegação social e autônoma

3.1 Sensores do Pepper

Um elemento indispensável para o correto funcionamento da estrutura seria a qualidade das informações que ele

receberá através dos sensores do robô. Para esta aplicação estamos trabalhando com informações provenientes de três sensores Laser (que podem cobrir uma área de 270 graus ao redor do robô) e uma câmera RGB (com resolução de 640x480). Inicialmente optamos por implementar nossa proposta, usando o *Robot Operating System* (ROS) e o *software* de simulação Gazebo¹ para validar nossos algoritmos. Além disso, implementamos uma estrutura para o robô Pepper, mostrado na Figura 2, o qual será o robô utilizado na simulação. Para isto criamos um ambiente de simulação 3D, que nos permitiu trabalhar com o robô Pepper, pessoas e objetos diferentes no mesmo ambiente.

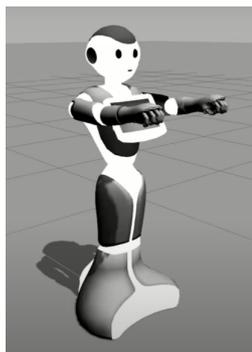


Figura 2. Robô Pepper utilizado na simulação

3.2 Detecção Humana

Para iniciar o processo desenvolvido no diagrama, primeiro precisamos obter as informações do ambiente, identificando e classificando pessoas e objetos existentes nele. É importante notar que o processo de identificação e classificação precisa ser feito utilizando uma percepção visual robocêntrica, ou seja, considerando que a perspectiva de um robô pode variar dependendo de vários aspectos, como deslocamento, vibração, agentes externos, movimentos circulares, angulares entre outros aspectos que fazem parte do processo natural do robô quando ele está se movendo para um determinado alvo. Para alcançar esse propósito, utilizamos o algoritmo de detecção de objetos YOLO v3 (Du, 2018), que nos permite diferenciar eficientemente pessoas e objetos no ambiente. Esta versão possui uma arquitetura subjacente de 106 camadas totalmente convolucional, o que torna o YOLO um pouco mais lento, mas com precisão superior em comparação com seu antecessor, YOLO v2.

3.3 Remoção de Outliers

Ao utilizar sensores de distância, podem surgir erros nos dados provenientes deles, por isso algumas estratégias foram implementadas para reduzir tais erros. O LIDAR, sensor laser utilizado, emite diversos feixes de laser em um ângulo de 360°, porém há um ângulo entre cada feixe.

A primeira fonte de erro identificada está intimamente relacionada com a resolução do sensor laser, pois o LIDAR, sensor laser utilizado, emite diversos feixes de laser em um ângulo de 360°, porém existe um ângulo relativamente grande entre cada feixe, o que significa que quanto maior

¹ simulador de robótica 3D de código aberto. <https://gazebo.org/home>

a distância entre o robô e o objeto, maior o espaço sem monitorar este sensor e, conseqüentemente, sem dados. Por exemplo, se o ângulo entre o laser central ($\theta = 0$) e o próximo laser for $\theta = 30^\circ$; e se um objeto for identificado pelo laser central a 5m de distância, o próximo laser só capturará outro valor aproximadamente 3m ao lado do ponto onde o laser central encontra o objeto. Ou seja, a resolução do LIDAR é um problema que pode causar muitos erros nas medições. Outro erro identificado devido à baixa resolução do dispositivo ocorre quando um dos lasers passa entre as pernas da pessoa na frente do robô, mostrado em vermelho na Figura 3. Esta distância lateral entre os lasers pode ser determinada por uma relação trigonométrica simples (Equação 1).

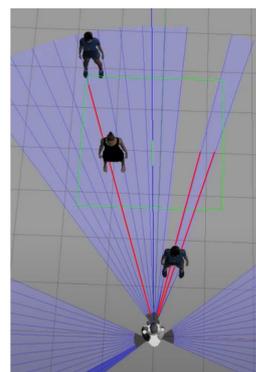


Figura 3. Feixe do laser passando entre as pernas das pessoas

$$D_{xn} = \tan \left(\theta \left| \frac{N+1}{2} - n \right| \right) D \quad (1)$$

A Equação 1 mostra como a distância horizontal entre cada laser está relacionada ao ângulo entre cada feixe e a distância D para o objeto onde N é o número total de lasers e n é o índice do laser analisado. Para resolver esse problema, é realizado um cálculo de média harmônica entre os lasers n_{min} até n_{max} . Esse tipo de média foi escolhida porque é menos sensível a medidas muito diferentes das demais. O cálculo da média harmônica é feito com a Equação 2.

$$M = \frac{n_{max} - n_{min}}{\frac{1}{d_{n_{min}}} + \frac{1}{d_{n_{min}+1}} + \dots + \frac{1}{d_{n_{max}}}} \quad (2)$$

Assim, mesmo que algum laser meça uma distância muito maior do que a distância real para a pessoa ($d_n \gg d_{n_{min}}, d_{n_{min}+1}, d_{n_{min}+2}, \dots, d_{n_{max}}$) o valor médio M não mudará muito, o que tende a mantê-lo como um valor próximo ao real, mesmo que haja algum erro como um laser que passa entre as pernas da pessoa.

3.4 Detecção de distância e velocidade

O uso de aprendizado de máquina em uma imagem RGB é suficiente para garantir a detecção e o reconhecimento de objetos. No entanto, isso não é suficiente para indicar a distância do robô para tal objeto. Para isso, é utilizado o sensor LIDAR chamado "Laser Scan", porém ele não dispara os lasers de forma tridimensional, mas mesmo assim é suficiente para a aplicação desejada neste trabalho.

A partir da estimativa de que as bordas medidas pelo laser estão alinhadas com as bordas da imagem da câmera, uma aproximação matemática foi feita considerando que se os lasers forem sobrepostos na imagem, eles seriam igualmente espaçados uns dos outros, com os lasers nas bordas correspondentes aos pontos $x = 1$ e $x = 640$ ou seja, eles estão no primeiro e último pixel da imagem.

3.5 Detecção de Emoções

Para que o robô seja capaz de detectar expressões faciais em um humano, é necessário que ele identifique as zonas de ativação facial que estão envolvidas em cada uma das 7 emoções básicas. Para a identificação das emoções através do algoritmo detector de emoções (Quiroz et al., 2022), o primeiro passo é fazer com que o robô identifique quais zonas de ação estão sendo ativadas na face do indivíduo que está sendo analisado. Para realizar a detecção, o software Make Human (Tang and Guo, 2016) foi usado para criar modelos humanos com expressões faciais. Estes modelos foram usados para testes. A Figura 4 mostra o processo de identificação da expressão facial.

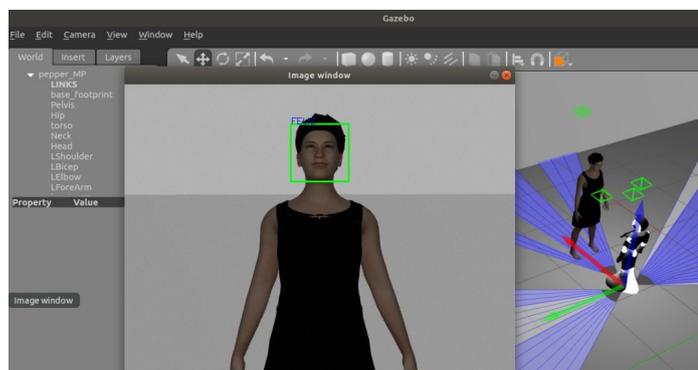


Figura 4. Identificação da Expressão Facial

3.6 Zonas Proxêmicas

A maneira como os robôs devem abordar os humanos no ambiente para alguma interação, ou mesmo a maneira como eles devem manter sua distância dos humanos estão diretamente ligadas às zonas proxêmicas de cada indivíduo. Uma área proxêmica é basicamente um limite espacial onde as pessoas tendem a se sentir desconfortáveis se alguém cruzá-la. A área de conforto de uma pessoa está ligada a vários fatores que podem mudar sua forma e tolerância. Neste artigo, a distribuição gaussiana assimétrica é usada para determinar as zonas proxêmicas (Lu and Smart, 2013) dos indivíduos no ambiente com base em suas expressões faciais. Esta forma de distribuição recebe como parâmetros principais as coordenadas dos limiares de seus lados, frente e trás.

Após a realização das etapas descritas acima, é necessária uma forma de visualização para os resultados obtidos na fusão dos sensores. Usamos como referência representar as zonas proxêmicas, onde podemos analisar a perspectiva do robô em relação às pessoas do ambiente. O algoritmo foi projetado de tal forma que dados como distância média (resultado médio harmônico), classe do objeto detectado e índices dos pixels em que a pessoa está situada aparecem em tempo real através do ROS.

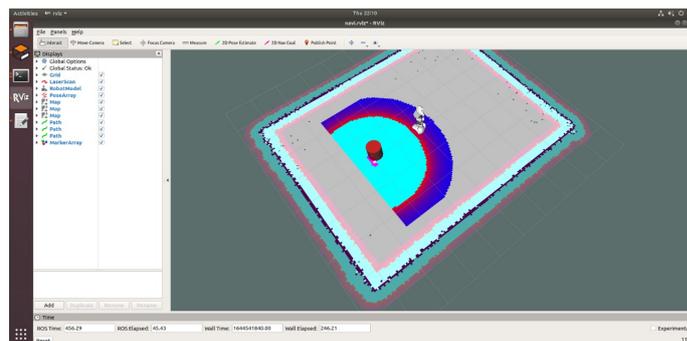


Figura 5. Zonas Proxêmicas Obtidas da Visão Robocêntrica

Na Figura 5 podemos ver as zonas proxêmicas calculadas, onde o azul escuro representa uma zona pública. Neste caso o robô pode invadir essa área, mas a zona social (azul claro) e a zona íntima (vermelho), o robô não será capaz de entrar nelas, levando a uma possível modificação do caminho de navegação do robô. A área da zona social é diretamente influenciada em tamanho e forma pela emoção detectada pelo robô na pessoa à sua frente.

3.7 Rastreamento de Rota

A navegação autônoma baseia-se no planejamento e execução de uma trajetória em que o robô percorre a menor distância possível, garantindo a integridade de zonas proxêmicas previamente estabelecidas e evitando colisões com outros objetos presentes no ambiente. Em um ambiente dinâmico, com objetos móveis inseridos, o planejamento do caminho deve considerar que após a criação da primeira rota, o robô deve ser capaz de calcular novas rotas de acordo com as mudanças no ambiente, já que existe a possibilidade de algum objeto ou alguma pessoa interferir. A definição das zonas proxêmicas das pessoas que estão na frente do robô determinará uma navegação dinâmica e em tempo real.

4. RESULTADO

O trabalho realizado tem, entre outros aspectos, uma forte relação da percepção dos robôs com os ambientes, incluindo a detecção, classificação dos objetos à volta e a decisão das zonas proxêmicas dos humanos. Além disso, a relação com a navegação autônoma das máquinas, que está intimamente ligada ao estágio de percepção e é responsável por realizar o movimento do robô de um ponto inicial a um ponto final pelo caminho mais curto possível sem ultrapassar os limites das zonas proxêmicas das pessoas presentes no local. Levando em conta essa subdivisão do trabalho nos dois tópicos mencionados, os resultados obtidos quanto à percepção do ambiente (detecção e classificação de objetos e determinação das zonas proxêmicas dos indivíduos detectados) e navegação autônoma (definição do caminho mais curto evitando a colisão com objetos e invasão de zonas proxêmicas estabelecidas) foram satisfatórios.

Na percepção robocêntrica do ambiente, pode-se observar na Figura 6 que o *fremework* é capaz de mostrar várias pessoas, usando cores diferentes, e o robô consegue identificar a distância e a emoção de cada uma delas. Para comprovar

a eficácia das zonas proxêmicas, foram realizados dois testes sendo que, no primeiro, o robô teve que evitar passar no meio de um grupo de duas pessoas, criando a partir da trajetória um caminho ao redor do grupo até seu destino final. Portanto, é possível visualizar na Figura 7 o primeiro cenário no Gazebo, no qual, a partir dos sensores do robô, um grupo de duas pessoas foi identificado e representado no RViz¹ com um cilindro vermelho, logo, a trajetória percorrida pelo robô tem relação com a gaussiana gerada em torno das pessoas. É importante notar que uma pessoa possui uma área de gaussiana maior que a outra. Isso se deve ao fato de o robô detectar nela uma emoção de tristeza, e na outra, com a área de gaussiana menor, um sentimento de felicidade.

Na Figura 8, é apresentado o segundo cenário no Gazebo, onde um grupo de três pessoas em fila foi identificado a partir dos sensores e representado no RViz através de cilindros vermelhos. Na simulação, o robô parte da posição (0,0) em direção à pessoa mais distante dela. Nota-se que, primeiramente, o robô opta por ir do lado direito da segunda pessoa. E à medida que os sensores detectam as outras pessoas na fila, a zona proxêmica se transforma de acordo com as emoções detectadas nas pessoas, o robô continua seu caminho no mesmo sentido. Com isso, fica claro que em ambos os testes o robô Pepper respeitou a proxêmica nos cenários propostos e, assim, o espaço pessoal dos humanos, de acordo com a emoção delas. Neste primeiro momento é representado apenas a zona proxêmica da emoção mais negativa detectada, de acordo com a Tabela 1, não sendo possível, ainda, representar cada obstáculo com uma zona proxêmica diferente ao mesmo tempo. Ou seja, caso fosse identificado as emoções feliz e triste ao mesmo tempo, a zona proxêmica gerada em ambas as pessoas seria a representada pela emoção triste. Seguindo a ordem feliz, sério, surpresa e triste.

Tabela 1. Relação entre as emoções e sua classificação

<i>Emoção</i>	<i>Classificação</i>
Feliz	Positiva
Sério	Neutra
Surpresa e Triste	Negativa

Nas Figuras 9-10 é possível observar o resultado de simulação para diferentes trajetórias realizadas pelo Pepper, considerando diferentes emoções. Na Figura 9 (a) pode-se observar o Pepper e a gaussiana ao redor da localização da pessoa. Nota-se, no referido caso, que a gaussiana possui uma dimensão maior, comparado-se com os demais casos, como a Figura 10 (d), na qual a pessoa está feliz, permitindo a aproximação máxima do Pepper. Na Figura 10 pode-se observar as trajetórias obtidas nos testes com diferentes emoções (Figura 9). É perceptível a adaptação das dimensões da gaussiana, bem como o planejamento de trajetória do Pepper, que se aproxima ou se distancia das pessoas em função da análise de emoções.

5. DISCUSSÕES

Nesta primeira versão verificamos a viabilidade e a eficiência dos algoritmos utilizados assim como dos sensores do

¹ Ferramenta de visualização 3D para ROS, obtido ao instalar o Gazebo

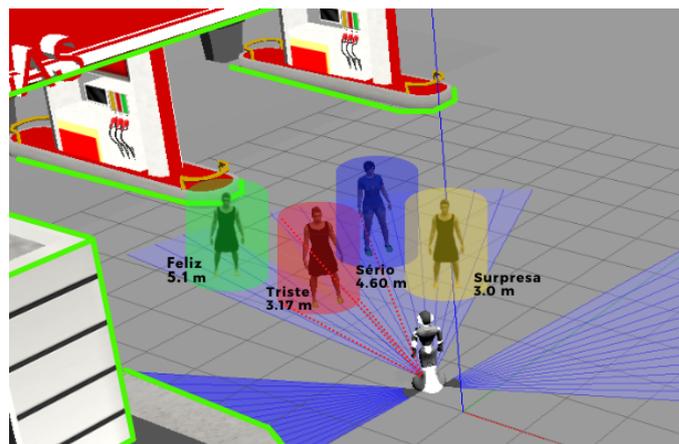


Figura 6. Sistema consegue identificar múltiplas pessoas individualmente, além disso, determina suas respectivas distâncias e emoções

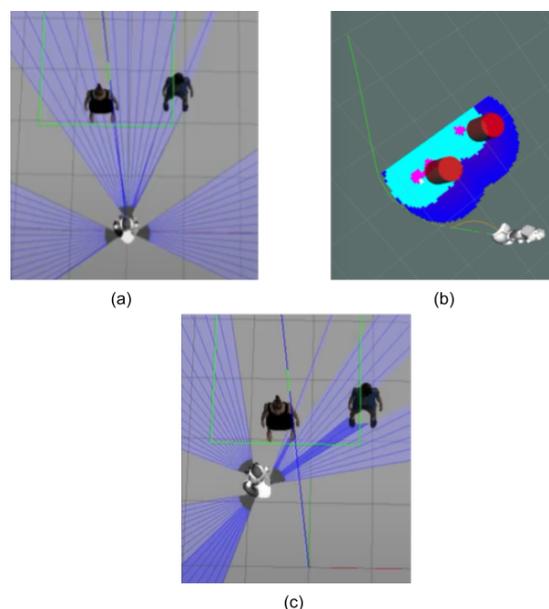


Figura 7. Primeiro teste: (a) simulação do cenário no Gazebo, (b) representação do cenário e visualização dos sensores no RViz, (c) trajetória percorrida pelo robô

robô em questão. Também foi possível extrair algumas das limitações existentes.

5.1 Perspectiva robocêntrica

A perspectiva do robô é crucial pois é ela que trará uma noção tridimensional do ambiente. Porém esta perspectiva pode variar a depender de vários aspectos, como vibrações, luminosidade, aspectos ambientais, locomoção, entre outros. Neste primeiro momento trabalhamos diretamente com a perspectiva visual e conseqüentemente com os movimentos da cabeça do robô. Percebemos que, por serem testes realizados virtualmente, com modelos humanos virtuais, as imagens obtidas pela câmera são de baixa resolução, o que dificulta a análise dos objetos. Porém acreditamos que quando implementados no robô real, estes resultados serão melhores.

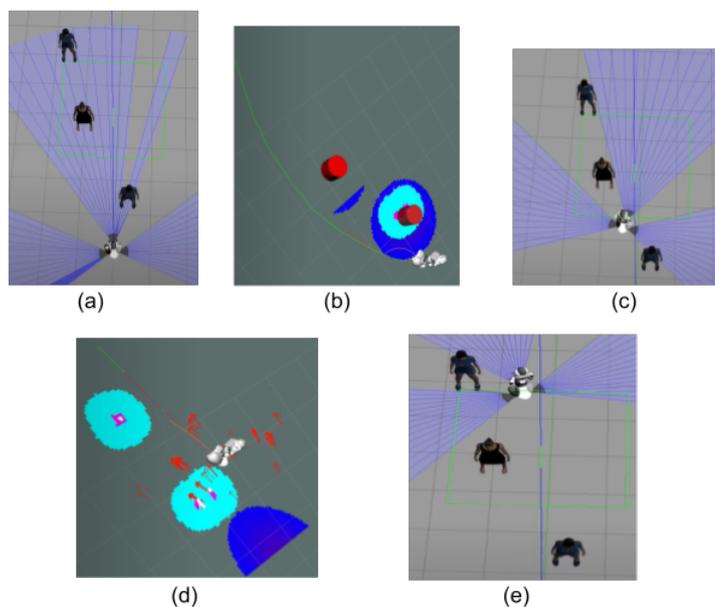


Figura 8. Segundo teste: (a) simulação do cenário no Gazebo, (b) representação das pessoas identificadas no cenário com um cilindro vermelho no RViz, (c) primeira decisão de trajetória tomada pelo robô em relação à Gaussiana formada ao redor as pessoas, (d) representação no RViz do desenvolvimento da Gaussiana de acordo com a trajetória do robô, (e) chegada ao destino.

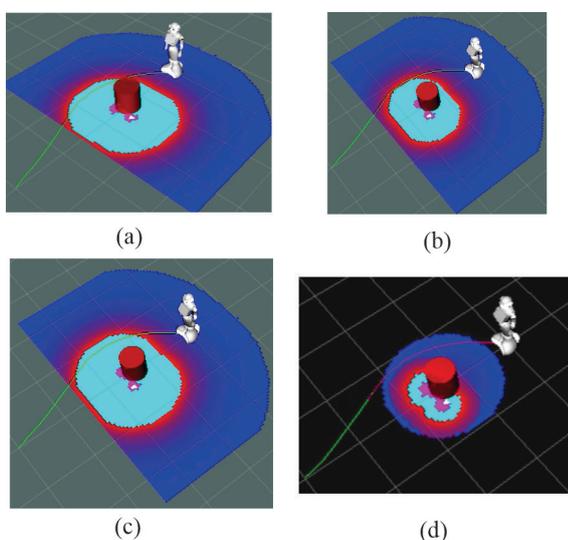


Figura 9. Planejamento de caminho realizado com diferentes emoções: (a) Raiva, (b) nojo, (c) medo (d), felicidade (e).

5.2 Representação das características nas zonas proxêmicas

A navegação de robôs em ambientes com a presença de humanos tem sido um assunto muito discutido e estudado pela comunidade científica. E para que esses robôs sejam aceitos trafegando no mesmo ambiente que os humanos é necessários que eles respeitem uma certa distância entre ele e o humano. Distância esta que pode ser definida por diversos aspectos como gênero, idade, etnia, emoção e até aspectos físicos. Este distanciamento caracterizamos

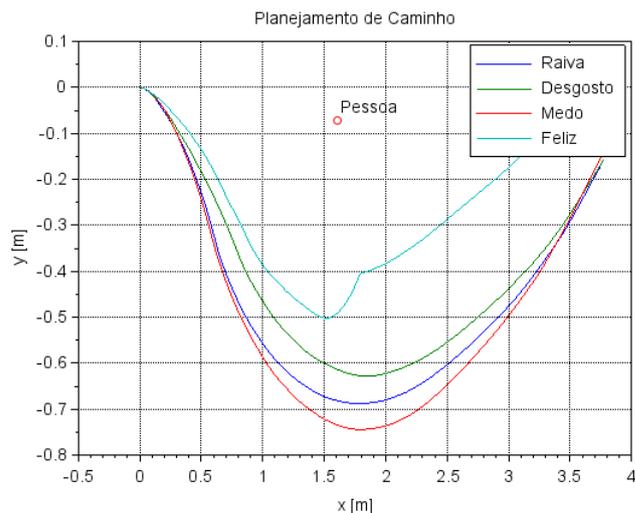


Figura 10. Trajetórias realizadas pelo Pepper em diferentes condições de detecção de emoções.

como zona proxêmica. Pensando nisso, desenvolvemos este trabalho para apresentar um sistema de navegação com diversos algoritmos capazes de oferecer uma navegação mais segura para os humanos respeitando, em primeiro lugar, o seu espaço social. Estes algoritmos foram testados em ambientes virtuais contendo humanos com expressões faciais distintas, onde o robô teria que analisar a expressão e determinar a que distância mínima ele poderia se locomover da pessoa e ao mesmo tempo que ele analisa, ele traça sua melhor trajetória. Os resultados obtidos são analisados qualitativamente conforme o comportamento do robô frente à emoção detectada. Os resultados indicam que respeitar o espaço pessoal pode melhora significativamente a aceitação de robôs autônomos no meio social.

6. CONCLUSÃO

Após a implementação e o teste de nossa proposta, verificamos sua viabilidade. É fato que a perspectiva de um robô é fundamental, porque trará uma noção tridimensional do ambiente, mas essa perspectiva pode variar de acordo com diversos aspectos, como vibrações, luminosidade, aspectos ambientais, locomoção, entre outros.

Para que robôs sejam aceitos viajando no mesmo ambiente que humanos é necessário que eles respeitem uma certa distância aos humanos. E essa distância pode ser definida por diversos aspectos, sendo o gênero, idade, etnia, emoção e até mesmo aspectos físicos levados em consideração. Foi caracterizada essa distância como zona proxêmica. A proposta foi testada em ambientes virtuais contendo humanos com diferentes expressões faciais, onde o robô teria que analisar a expressão e determinar o melhor cenário de distância que poderia se deslocar. Os resultados obtidos foram analisados qualitativamente de acordo com o comportamento do robô diante da emoção detectada e mostram que é possível respeitar o espaço pessoal das pessoas e espera-se com isso melhorar significativamente a aceitação de robôs autônomos no meio social.

AGRADECIMENTOS

Esta pesquisa foi apoiada pelo Grupo de Inovação e Pesquisa em Automação e Robótica (GIPAR) do Instituto Federal da Bahia (IFBA), Vitória da Conquista, Bahia e também pelo Grupo de pesquisa em controle, automação e robótica (GPCAR) do Instituto Federal da Paraíba (IFPB), Cajazeiras, Paraíba.

REFERÊNCIAS

- Aghaei, M., Bustreo, M., Wang, Y., Bailo, G., Morerio, P., and Bue, A.D. (2020). Single image human proxemics estimation for visual social distancing.
- Che, Yuhang; Okamura, A.M. and Sadigh, D. (2020). Efficient and trustworthy social navigation via explicit and implicit robot–human communication. *IEEE Transactions on Robotics*, 36(3), 692–707. doi:10.1109/TRO.2020.2964824.
- Cunhal, M.J.A. (2014). *Sistema de visão para a interação e colaboração humano-robô: reconhecimento de objetos, gestos e expressões faciais*. Ph.D. thesis.
- Daza, M., Barrios-Aranibar, D., Diaz-Amado, J., Cardinale, Y., and Vilasboas, J. (2021). An approach of social navigation based on proxemics for crowded environments of humans and robots. *Micromachines*, 12(2), 193.
- Du, J. (2018). Understanding of object detection based on cnn family and yolo. In *Journal of Physics: Conference Series*, volume 1004, 012029. IOP Publishing.
- Ekman, P. (1999). Basic emotions. *Handbook of cognition and emotion*, 98(45-60), 16.
- Ginés, J., Martín, F., Vargas, D., Rodríguez, F.J., and Matellán, V. (2019). Social navigation in a cognitive architecture using dynamic proxemic zones. *Sensors*, 19(23), 5189.
- Hall, E.T. (1963). A system for the notation of proxemic behavior. *American anthropologist*, 65(5), 1003–1026.
- Hurtado, Juana Valeria; Londoño, L. and Valada, A. (2021). From learning to relearning: A framework for diminishing bias in social robot navigation. *Frontiers in Robotics and AI*, 8. doi:10.3389/frobt.2021.650325. URL <https://www.frontiersin.org/article/10.3389/frobt.2021.650325>.
- Kim, J., Kim, J., and Cho, J. (2019). An advanced object classification strategy using yolo through camera and lidar sensor fusion. In *2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 1–5. IEEE.
- Kose, H.K.F.C.H. and Yavuz, S. (2021). Social navigation framework for assistive robots in human inhabited unknown environments. *Engineering Science and Technology, an International Journal*, 24(2), 284–298. doi:<https://doi.org/10.1016/j.jestch.2020.08.008>. URL <https://www.sciencedirect.com/science/article/pii/S2215098620308727>.
- Kumar, G.A., Lee, J.H., Hwang, J., Park, J., Youn, S.H., and Kwon, S. (2020). Lidar and camera fusion approach for object distance estimation in self-driving vehicles. *Symmetry*, 12(2). doi:10.3390/sym12020324. URL <https://www.mdpi.com/2073-8994/12/2/324>.
- Lehmann, H., Rojik, A., and Hoffmann, M. (2020). Should a small robot have a small personal space? investigating personal spatial zones and proxemic behavior in human-robot interaction.
- Lu, David V; Allan, D.B. and Smart, W.D. (2013). Tuning cost functions for social navigation. In *International Conference on Social Robotics*, 442–451. Springer.
- Lu, Y., Zhang, L., and Xie, W. (2020). Yolo-compact: An efficient yolo network for single category real-time object detection. In *2020 Chinese Control And Decision Conference (CCDC)*, 1931–1936. doi:10.1109/CCDC49329.2020.9164580.
- Mavrogiannis, Christoforos; Alves-Oliveira, P.T.W. and Knepper, R.A. (2022). Social momentum: Design and evaluation of a framework for socially competent robot navigation. *J. Hum.-Robot Interact.*, 11(2). doi:10.1145/3495244. URL <https://doi.org/10.1145/3495244>.
- Nin, M.C. and Osório, F. (2011). Navegação de robôs móveis autônomos e detecção de humanos baseada em sensor laser e câmera térmica.
- Patompak, P., Jeong, S., Nilkhamhang, I., and Chong, N.Y. (2020). Learning proxemics for personalized human–robot social interaction. *International Journal of Social Robotics*, 12. doi:10.1007/s12369-019-00560-9. URL <https://doi.org/10.1007/s12369-019-00560-9>.
- Pérez-Higueras, N., Ramón-Vigo, R., Merino, L., and Caballero, F. (2014). Navegacion local de un robot movil con costes sociales aprendidos por demostracion.
- Pham, N.A., Nguyen, L.A., and Truong, X.T. (2021). Socially aware robot navigation framework: Social activities recognition using deep learning techniques. In *2021 8th NAFOSTED Conference on Information and Computer Science (NICS)*, 381–385. doi:10.1109/NICS54270.2021.9701551.
- Quiroz, M., Patiño, R., Diaz-Amado, J., and Cardinale, Y. (2022). Group emotion detection based on social robot perception. *Sensors*, 22(10), 3749.
- Redmon, J., Divvala, S.K., Girshick, R.B., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788.
- Tang, Jin; Luo, J.T.T. and Guo, F. (2016). Robust arbitrary-view gait recognition based on 3d partial similarity matching. *IEEE Transactions on Image Processing*, 26(1), 7–22.