

Plataforma ROSANA: Arquitetura de sistema robótico

Guilherme Bitencourt Nunes (CTI) gbnunes@cti.gov.br, Marcos Vinicius Cruz (CTI)
marcos.cruz@cti.gov.br

Supervisor Josué Junior Guimarães Ramos

Resumo

O artigo aborda a arquitetura do arcabouço de tecnologia de interação humano-robô ROSANA. Esse modelo de plataforma foi implementado através da arquitetura HBBA (Hybrid Behavior-Based Architecture), sendo essa uma arquitetura que combina recursos reativos e deliberativos com processamento em módulos distribuídos horizontalmente. A comunicação entre módulos é realizada através do middleware ROS através de um sistema Publish/Subscribe. A plataforma conta com ferramentas de visão computacional e processamento de linguagem natural, estando aberta a adição de novas funcionalidades através da adição de novos módulos à arquitetura.

Palavras-chave: HBBA, ROS, Plataforma robótica, microserviços

1. Introdução

O termo Indústria 4.0 foi utilizado pela primeira vez em 2011, durante a feira de Hannover (MALIK, 2021), e surge como forma de sintetizar as mudanças promovidas nos processos industriais através da informatização, que proporciona a conexão entre o mundo físico e digital. Esse novo conceito de modelo industrial é impulsionado por tecnologias como manufatura aditiva, computação em nuvem, internet das coisas, big data, robótica, blockchain, entre outras, com o objetivo principal de maximizar a produtividade e alcançar a produção em massa. (ECHCHAKOUI & BARKA, 2020)

Tendo o modelo de Indústria 4.0 como atual estado da arte, sua evolução, a Indústria 5.0 tem como objetivo a utilização da criatividade de especialistas humanos trabalhando em conjunto com máquinas eficientes, inteligentes e precisas, com foco na personalização em massa (ELFAR, 2021). Para uma ideal relação de trabalho em sinergia de robô e humanos, uma das tecnologias 5.0 chave é aquela que envolve o estudo da relação humano-robô (HRI). Nesse sentido, a comunidade de pesquisa propôs o desenvolvimento de plataformas robóticas, denominadas de robôs colaborativos, capazes de colaborar e compartilhar o espaço de trabalho com colaboradores humanos de forma a maximizar a eficiência dessa relação. (EL ZAATARI, 2019) Segundo Giansanti (2019), a robótica colaborativa está entre os principais desenvolvimentos tecnológicos dos últimos anos, com destaque para aqueles relacionados à robótica social. O autor também coloca a robótica social como uma evolução da robótica colaborativa, que antes limitada a ambientes de trabalho, agora se faz presente em ambientes sociais.

Robôs sociais, segundo Paiva (2020), em sua interação colaborativa devem ser capazes de:

- Estabelecer e manter relações sociais;
- Identificar modelos de comportamento;
- Saberem interpretar sinais naturais, como gestos e olhares;
- Expressar e perceber emoções;
- Utilização linguagem natural como forma de comunicação;
- Realizar navegação social;

Com foco nesses conceitos, o presente trabalho tem como objetivo apresentar a implementação de um robô socialmente interativo: a plataforma robótica ROSANA (RObot for Social interAction in uNstructured dynAmic environments), um robô social feito para interagir e auxiliar pessoas em um prédio público. Essa implementação é realizada em uma arquitetura robótica híbrida que combina dois paradigmas de controle de robôs: controle baseado em comportamento e controle híbrido. O modelo implementado foi baseado em serviços, distribuídos através do middleware ROS (Robot Operating System) de forma horizontal.

2. Plataforma Rosana

ROSANA é uma plataforma social interativa com capacidades de percepção de ambiente desenvolvida pela Divisão de Sistemas Ciber-Físicos. O modelo apresenta possui comunicação em nuvem de forma a aumentar suas capacidades de lidar com informações, reduzindo assim os requisitos de processamento on-board do sistema. ROSANA consiste de uma base móvel Pioneer 3DX, com dezesseis sonares, três prateleiras de suporte, um telêmetro a laser Hokuyo, um Microsoft Kinect V2, um notebook e um monitor com o avatar ANA (PAIVA, 2020).

O avatar ANA é um agente virtual interativo com capacidade de interação verbal e não verbal (gestos e expressões). O agente foi implementado como um serviços em nuvem, onde através do protocolo de comunicação SSH, os comandos que controlam expressões e falas são recebidos.

3. Arquitetura de Comportamento Híbrido (HBBA)

Para robôs interagirem de forma natural e eficaz com seres humanos, a integração de múltiplas modalidades de percepção e ação se fazem necessárias. Entretanto uma ampla gama de sensores e atuadores que requerem alto consumo de processamento, também podem acabar sobrecarregando o sistema embarcado de um robô autônomo. Em razão disso, optou-se pela incorporação da arquitetura HBBA com sistemas de serviços distribuídos e disponibilizados via rede com o intuito da redução do consumo de recursos.

A HBBA (figura 1) combina dois paradigmas de controle de robô, controle baseado em comportamento e controle híbrido (FERLAND; MICHAUD, 2014). O primeiro tem como base um conjunto de módulos distribuídos, chamados de comportamentos, que recebem entradas de sensores ou outros comportamentos do sistema, e fornecem saídas para os atuadores do robô ou outros comportamentos. Deste modo, as funcionalidades do robô se dão pela sua interação com o ambiente. Já o controle híbrido visa combinar a resposta em tempo real do reativo, e a racionalidade e otimização do deliberativo. Sistemas baseados em

comportamento têm se mostrado mais adequados para ambientes com mudanças dinâmicas significativas envolvendo respostas rápidas e adaptabilidade, enquanto o controle híbrido é apropriado em ambientes e tarefas que envolvem modelos internos e planejamento (FERLAND; CRUZ-MAYA; TAPUS, 2015).

O HBBA é composto por três camadas: a camada Organização, onde estão os módulos Motivação, a camada Coordenação, contendo o Espaço de Trabalho da Intenção (IW) e seus módulos associados, e a camada Comportamental, contendo os módulos Percepção e Comportamento (FERLAND, 2017). Os módulos são responsáveis por:

- Módulos de motivação: trabalham em cima dos dados e informações fornecidos pelos módulos de percepção, gerando desejos que podem recomendar ou inibir ações como ir a algum lugar ou expressar a fala, mas também tarefas de processamento de informações, como reconhecimento de pessoas. Nesse módulo é definido o que o robô deve fazer, mas não define como o irá fazê-lo. Aqui entra o conceito de engajamento, que está relacionado aos métodos de interação humano-robô.
- Módulos de percepção: responsável por captar as informações brutas dos sensores e os transformarem em percepções para os módulos de comportamento. Nele é onde, através de técnicas de programação de linguagem natural, os áudios captados pelo microfone são convertidos para sentenças que posteriormente servirão como base para extração de dados. Nesse módulo também é onde são realizadas tarefas relacionadas à visão computacional, como detecção de pessoas, onde as imagens brutas da câmera são processadas para produzir a localização de uma pessoa, se houver.
- Módulos de comportamento: através das percepções recebidas produzem comandos para os atuadores do robô. O acesso aos atuadores é arbitrado, geralmente em um esquema de subsunção baseado em prioridade.

Ainda há o Intention Workspace, módulo que reúne os desejos como por exemplo, o engajamento de pessoas e navegação pelo ambiente, para inferir uma intenção ao robô. Estes desejos possuem valores de intensidade que são utilizados para definir suas prioridades, pois esses competem pelos mesmos recursos do robô. O módulo ainda possui o Intention Translate, que transforma o conjunto selecionado de desejos não conflitantes em intenções com base em um banco de dados de estratégias. Cada estratégia descreve como uma classe específica de desejo pode ser cumprida em uma determinada plataforma de robô, o que inclui a ativação dos módulos de Comportamento e Percepção, e a transferência de parâmetros de configuração de Comportamento para esses módulos. Do intention Workspace também saem as crenças, que são informações compartilhadas entre os módulos de motivação sobre a dinâmica de interação interna e externa do robô, como a realização de desejos e conflitos. (FERLAND, 2017)

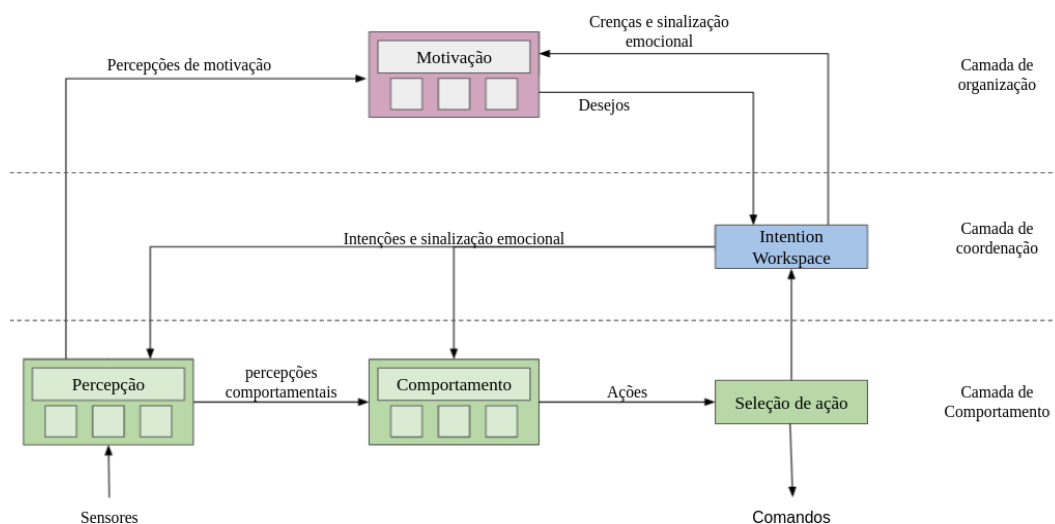


Figura 1: Hybrid Behavior-Based Architecture (HBBA).

3.1. ROS e arquitetura de microserviços

O HBBA é implementado em cima do ROS (Robot operating system), sendo esse, um conjunto de bibliotecas e ferramentas para auxílio na construção de robôs. O ROS é utilizado para a comunicação entre os módulos que funcionam de modo paralelo através do padrão publish/subscribe, podendo funcionar em um cluster heterogêneo de computadores.

Esse padrão consiste em nós que se comunicam através da publicação e subscrição em tópicos, ou seja, a comunicação é feita de forma indireta e independente. O publisher detecta eventos e os publica na forma de notificação, que faz o encapsulamento do conteúdo relacionado ao evento observado. Após publicado, a mensagem é entregue para todos os módulos e serviços com interesse na informação através do subscribe. Esse padrão facilita sua implementação em ambientes com diferentes elementos, linguagens de programação e tecnologias, pois se limita apenas aos dados publicados, permitindo a horizontalização e distribuição dos microserviços do sistema.

Microserviços é um estilo de arquitetura que enfatiza a divisão do sistema em serviços pequenos e leves que são propositadamente construídos para desempenhar uma função de negócios muito coesa, esta é uma evolução do estilo tradicional de arquitetura orientada a serviços (FOWLER; LEWIS, 2015). Também definido como "processo mínimo independente que interage via mensagens", sendo a arquitetura de microservices definida como "uma aplicação distribuída onde

todos os seus módulos são microserviços". Os benefícios comumente acordados desse estilo incluem: aumento da agilidade, produtividade do desenvolvedor, resiliência, escalabilidade, confiabilidade, capacidade de manutenção, separação de preocupações e facilidade de implantação (DRAGONI, 2017).

4. Caracterização do sistema

Como já detalhado, os microserviços compõem os módulos, que por sua vez compõem a arquitetura na qual a plataforma ROSANA foi implementada. Para sua operacionalização, o

sistema possui uma série de ferramentas relacionadas à visão computacional, processamento de linguagem natural e comunicação.

O sistema de visão tem como sistema de captura uma câmera OAK-D. Esse modelo de câmera possui módulo de Inteligência Artificial integrado voltado à visão computacional, que através da API DepthAI disponibiliza recursos para a identificação e reconhecimento de pessoas feitas pela rede neural OpenVino (MAHENDRAN, 2021). A OAK-D também disponibiliza um conjunto de câmeras estéreo que são utilizadas para estimativas espaciais da plataforma.

Para o sistema de reconhecimento de voz, o áudio capturado pelo microfone é transcrito através da API Google Speech-to-Text. Para a análise morfológica é utilizado a biblioteca Spacy. Através das informações obtidas, o robô pode fornecer uma resposta reativa baseada em um banco de dados local, ou uma resposta mais deliberativa através da rede neural Transformer rodando em um servidor GPU externo. Esse modelo de rede neural, Transformer, é o estado da arte se tratando de NLP, sendo melhor apresentado e descrito por Vaswani(2017).

Por meio dos demais pesquisadores do núcleo, outras tecnologias vêm sendo desenvolvidas para uma futura implementação na arquitetura implementada, criando assim um arcabouço de tecnologias voltadas à interação humano-robô. As tecnologias em desenvolvimento são:

- Navegação social: trata de aspectos relacionados às regras sociais que envolvem a locomoção e comportamento de robô frente a uma pessoa ou grupo. (BATISTA, 2021)
- Detecção de emoção: aspectos relacionados à percepção de sentimentos através da análise de expressões e comportamento do usuário, e suas influência na interação. (PAIVA, 2021)
- Detecção de emoção por voz: aspectos relacionados à percepção de sentimentos através da fala do usuário. (JOSHI, 2021)

3. Resultados

ROSANA se mostrou uma plataforma robótica altamente personalizável e de fácil acoplamento de dispositivos tecnológicos que venham a incrementar suas capacidades de interação. Através do modelo arquitetura com serviços distribuídos horizontalmente, a implementação desses dispositivos podem ser feitos em um grupo heterogêneo de computadores de forma a não sobrecarregar as capacidades de processamento do robô pois esse não estaria limitado ao seu sistema embarcado.

Testes de interação foram realizados de forma a validar as funcionalidades envolvendo processamento de linguagem natural e visão computacional, assim como a arquitetura fazendo o gerenciamento de recursos. Além do mais, testes de performance sugeriram que a utilização consumo de CPU em médio prazo foi constante e de menor intervalo de resposta em comparação ao sistema anterior. Em relação ao uso de memória, o HBBA a reserva enquanto cria variáveis, alocando recursos de acordo com o grau de intensidade para reduzir a fragmentação da memória.

Apesar da funcionalidade da plataforma ROSANA, podem haver dificuldades de interação. Pois por ela não possuir uma forma humanoide, estando limitada ao avatar de interação, pode haver uma dificuldades em captar a atenção e interesse do usuário.

4. Trabalhos futuros

Por meio de recursos cedidos pela Fapesp através do Projeto “Socially Interactive Robots Acting in Public Environments” no ano de 2022 a divisão (DISFC) adquiriu o robô Pepper da Softbank. Esse modelo é um modelo de robô altamente utilizado em pesquisas e com grande aceitação por parte de usuários. Assim, uma das propostas do projeto é a implementação de uma instância do arcabouço de software ROSANA ao robô Pepper e realização de testes em ambientes controlados e não controlados.

Referências

BATISTA, Murillo Rehder. Tracking e abordagem de pessoas para navegação socialmente aceitável. X Seminário em TI do PCI/CTI – 2021

RAGONI, Nicola et al. Microservices: yesterday, today, and tomorrow. Present and ulterior software engineering, p. 195-216, 2017.

ECHCHAKOUI, Said; BARKA, Noureddine. Industry 4.0 and its impact in plastics industry: A literature review. Journal of Industrial Information Integration, v. 20, p. 100172, 2020.

EL ZAATARI, Shirine et al. Cobot programming for collaborative industrial tasks: An overview. Robotics and Autonomous Systems, v. 116, p. 162-180, 2019.

ELFAR, Omar Ashraf et al. Prospects of Industry 5.0 in algae: Customization of production and new advance technology for clean bioenergy generation. Energy Conversion and Management: X, v. 10, p. 100048, 2021

FERLAND, François; MICHAUD, François. An Architecture for Integrated Design of Human-Robot Interaction Capabilities. AAAI Fall Symposium on AI in HRI, Arlington, VA. 2014

FERLAND, François et al. Coordination mechanism for integrated design of human-robot interaction scenarios. Paladyn, Journal of Behavioral Robotics, v. 8, n. 1, p. 100-111, 2017.

FERLAND, François; CRUZ-MAYA, Arturo; TAPUS, Adriana. Adapting an hybrid behavior-based architecture with episodic memory to different humanoid robots. In: 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN). IEEE, 2015. p. 797-802.

FOWLER, Martin; LEWIS, James. Microservices: Nur ein weiteres Konzept in der Softwarearchitektur oder mehr. Objektspektrum, v. 1, n. 2015, p. 14-20, 2015.

GIANSANTI, Daniele. The social robot in rehabilitation and assistance: what is the future?. In: Healthcare. MDPI, 2021. p. 244.

JOSHI, Neelakshi. Brazilian Portuguese emotional speech corpus analysis. X Seminário em TI do PCI/CTI – 2021

MAHENDRAN, Jagadish K. et al. Computer vision-based assistance system for the visually impaired using mobile edge artificial intelligence. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021. p. 2418-2427.

MALIK, Nishtha et al. Impact of artificial intelligence on employees working in industry 4.0 led organizations. International Journal of Manpower, 2021.

PAIVA, Pedro VV et al. ROSANA: Robot for Social Interaction in Unstructured Dynamic Environments. In: 2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE). IEEE, 2020. p. 1-6.

PAIVA, Pedro VV. Fusão de informações não-verbais no reconhecimento do estado afetivo via Aprendizado Profundo. X Seminário em TI do PCI/CTI – 2021

VASWANI, Ashish et al. Attention is all you need. Advances in neural information processing systems, v. 30, 2017.