

# Uma Arquitetura Conceitual de uma Interface Adaptativa Inteligente para um Robô de Assistência Doméstica

José R. A. dos Santos<sup>12</sup>, Thiago S. B. Meyer<sup>1</sup>, Plínio T. Aquino Junior<sup>1</sup>

<sup>1</sup>Centro Universitário FEI – Fundação Educacional Pe. Sabóia de Medeiros - Programa de Pós-Graduação em Engenharia Elétrica e Equipe RoboFEI@Home

<sup>2</sup>Instituto Federal de Educação, Ciência e Tecnologia do Pará, IFPA, campus Itaituba  
{ribamarz, thiago.spilborghs}@gmail.com, plinio.aquino@fei.edu.br

**Abstract.** *Adaptive Interfaces, also known as Intelligent Interfaces, are devices capable of adapting to different types of users and contexts. These allow your interaction rules to be adapted in order to provide a better communication experience between human and robot. In this sense, the present work describes a conceptual architecture of an intelligent adaptive interface that will be incorporated into a home care robot. The core of this architecture is a method of identifying and reacting to human emotions, which adapts to the environment and context of human-robot interaction.*

**Resumo.** *As Interfaces Adaptativas, também conhecidas como Interfaces Inteligentes, são dispositivos capazes de se adaptar a diferentes tipos de usuário e contexto. Estas permitem que suas regras de interação sejam adaptadas visando fornecer uma melhor experiência de comunicação entre humano e robô. Neste sentido, o presente trabalho descreve uma arquitetura conceitual de uma interface adaptativa inteligente que será incorporada a um robô de assistência doméstica. O núcleo dessa arquitetura é um método de identificação e reação a emoções humanas, que se adapta ao ambiente e contexto da interação humano-robô.*

## 1. Introdução

Segundo Findlater e McGrenere (2004), as interfaces adaptativas são aquelas em que o sistema administra as mudanças de modo que a adaptação não dependa da interação direta do usuário. Dessa maneira, pode-se entender que as interfaces adaptativas são estruturas que admitem, em tempo de execução e sem intervenção, a adequação de seu conteúdo e regras, às demandas e predileções dos usuários finais.

Neste contexto, este trabalho tem como objetivo apresentar a proposta de uma arquitetura conceitual de uma interface adaptativa inteligente para ser incorporada a um robô serviço doméstico, responsável por realizar atividades como, por exemplo, pegar, guardar ou limpar um utensílio de cozinha. Essa arquitetura deve ser capaz de expressar sentimentos (tristeza, alegria, raiva, entre outros) de acordo com a interação homem-robô, buscando-se desse modo alcançar um processo de comunicação entre homem-robô mais próximo da linguagem natural humana.

Para Soni *et al* (1995) uma arquitetura conceitual é uma estrutura de alto nível que descreve o sistema em termo de seus principais elementos de *design* e os relacionamentos entre eles. Nesse sentido, a arquitetura conceitual projetada neste trabalho descreve em alto nível os módulos necessários à implementação da interface planejada.

A área de Robótica Inteligente, da qual faz parte este trabalho, tem passado por um forte ciclo de crescimento, exigindo que os sistemas funcionem em ambientes complexos e não estruturados, como residências. Para operar nestes lugares, os robôs precisam ter capacidades similares as humanas. Nesse contexto, o uso de uma interface adaptativa inteligente integrada a um robô é apresentado como uma forma de conferi-lo inteligência. As demais partes do trabalho estão organizadas da seguinte maneira: a Seção 2 apresenta os Trabalhos Relacionados. Na Seção 3, discute-se a Arquitetura Proposta. Por fim, na Seção 4, apresenta-se as Considerações e Trabalhos Futuros.

## **2. Trabalhos Relacionados**

Alguns dos trabalhos analisados abordam aspectos específicos de interfaces inteligentes adaptativas. Por exemplo, o trabalho de Huang *et al* (2019) concentra-se na interação homem-robô por meio de uma interface de percepção multimodal que combina a percepção visual e demais formas de comunicação humana não-verbal envolvendo toque. Para Huang *et al* (2019), a modalidade visual é considerada a característica mais importante para a assimilação humana, porém a modalidade tátil não é descartada. No entanto, diferente de Huang *et al* (2019) este trabalho tem como particularidade o desenvolvimento de uma arquitetura conceitual para uma interface adaptativa que contará com um método de classificação da interação multimodal baseado em informações obtidas das modalidades áudio (fala) e vídeo (expressões faciais, gestos e movimentos corporais) visando prever a emoção humana com razoável precisão.

O trabalho de Wang *et al.* (2019) apresenta semelhanças com este por ter introduzido interação social humano-robô e utilizar uma interface de comunicação para auxílio na execução das tarefas provenientes da interação com o humano. Porém, diferente de Ning Wang *et al.* (2019) esta pesquisa buscou apresentar uma arquitetura conceitual adaptativa que não se restringe a idosos, tendo em vista que a solução foi projetada para atender indivíduos com diferentes perfis e faixas etárias. Além disso, ainda não foi possível a realização de testes aprofundados para efetiva validação da proposta.

A pesquisa de Fortmann e Lüdtkke (2013) trata apenas um ponto de interfaces adaptáveis inteligentes que é a conscientização da situação do usuário. A interface de Fortmann e Lüdtkke (2013) controla a percepção da situação do usuário por meio do rastreamento visual e guia a atenção do usuário para informações relevantes, que não foram atendidas. Porém, diferente de Fortmann e Lüdtkke (2013), que tem seu trabalho direcionado apenas para a percepção da situação do usuário por meio do rastreamento visual, este trabalho foca em uma arquitetura conceitual que leva em consideração mais de um aspecto da interação ao abranger uma interface multimodal baseada em dados de fala, gestos, expressões faciais, entre outros, oriundos da comunicação humano-robô.

### 3. Arquitetura Proposta

A contribuição principal deste trabalho é uma arquitetura conceitual que está organizada em quatro módulos. O módulo Analisador de Sentimentos e/ou Intenções (ASI), abrange funções para percepção de sentimentos e/ou emoções humanos, bem como objetos e o contexto da interação. Além disso, identifica a intenção e/ou emoção humana provenientes da comunicação, como por exemplo, gestos, fala e movimentos faciais.

No módulo Conversor de Áudio em Texto (CAT), os dados são obtidos por sensors durante um ciclo de interação. Nessa etapa, o áudio recebido da comunicação é convertido em texto para ser interpretado no modulo ASI, após esta análise o modulo CAT fornece uma resposta em áudio ou texto para a pessoa. Logo, se o questionamento da pessoa for bem compreendido a resposta será dada em áudio, caso contrário em texto.

O Módulo Conversor de Vídeo em Áudio/Texto (CVAT), compreende sensors. câmeras para captura de gestos, movimentos corporais e de face em formato de vídeo. Esses dados são analisados e interpretados no Modulo ASI, de posse dessa análise o Módulo CVAT gera uma resposta ao usuário em áudio ou texto. Se a intenção ou emoção do humano for bem compreendida uma resposta em áudio será fornecida, caso contrário em formato de texto.

Por fim, têm-se a Base de Dados de Respostas (BDR), responsável por armazenar as respostas classificadas como de alta confiança, estas são identificadas no Módulo ASI. A base de dados contribui para melhorar o tempo de resposta ao usuário, uma vez que o Módulo ASI fará menos processamento para obter uma resposta, pois o armazenamento de um volume expressivo de dados propicia a geração de resposta mais rápidas e eficientes a um determinado questionamento do humano. Essa arquitetura conceitual por ser projetada na forma modular diminui sua complexidade ao decompor o processo de interação da interface em módulos mais simples. Nesse sentido, a abstração é empregada como fundamento de organização dos níveis dessa estrutura hierárquica.

#### 3.1. Percepção de Pessoas

Um requisito essencial para um robô de serviço doméstico é a identificação de pessoas. Isso, inclui a identificação, o rastreio, a percepção e interpretação dos gestos humanos, o diálogo multimodal e o reconhecimento de fala. Nesse sentido, a arquitetura conceitual projetada neste trabalho leva em consideração tais características:

**Identificação da Pessoa:** Para detecção de pessoas, um *script* de identificação e registro de rosto será utilizado. A identificação do rosto será predita com base nos dados contidos no banco de dados “Módulo BDR” e suas atualizações, haja vista que se o robô não reconhecer o descritor este é armazenado na base de dados, facilitando futuras interações com esse descritor.

**Reconhecimento de gestos:** Os gestos são um mecanismo natural de comunicação das pessoas. Um gesto de acenar por exemplo, pode ser empregado para chamar a atenção de um robô para execução de uma determinada ação. Para detectar gestos a arquitetura contará com um sensor de profundidade e câmeras 3D.

**Sistema de Diálogo Multimodal:** Para se comunicar com uma pessoa a arquitetura da interface precisa está próxima de uma pessoa, quando a proximidade acontece o robô de serviço aponta seus sensores (câmera e microfone) para o rosto da

pessoa. Os dados em formato de áudio e vídeo são então processados nos Módulos da interface (ASI, CAT, CVAT, BDR) para que uma resposta seja fornecida ao usuário.

Reconhecimento e Síntese de Fala: A interface proposta gera fala de acordo com a tarefa a ser realizada, por exemplo, solicitar um comando da pessoa ou comunicar ao usuário seu atual objetivo. O reconhecimento adequado da fala vai depender do conjunto gramatical fornecido e dos rótulos semânticos para identificação da linguagem natural humana. Se a fala da pessoa for bem compreendida a comunicação humano-robô ocorrerá por fala, caso contrário uma resposta em formato de texto é fornecida.

#### **4. Considerações e Trabalhos Futuros**

Neste trabalho, é proposto uma arquitetura conceitual para uma interface adaptativa inteligente que será incorporada a um robô de serviço doméstico. A estrutura é baseada no reconhecimento de emoções e do contexto da interação humano-robô. O método da arquitetura delineada mostra que a solução apresentada pode se adequar habilmente as necessidades dos usuários e fornecer uma nova maneira de comunicação entre pessoas e robôs de serviço. Destaca-se como pontos fortes deste trabalho a forma modular como a arquitetura foi planejada e a descrição dos aspectos essenciais para um robô de serviço interagir com humanos. Em contrapartida, como pontos a serem melhorados aponta-se a definição de como os módulos da arquitetura serão obtidos, se por intermédio de soluções prontas de código aberto ou desenvolvimento de recursos próprios. Nesta etapa do trabalho não realizou-se um estudo minucioso sobre o estado da arte dos temas abordados nas pesquisas investigadas. Entretanto, a continuidade da pesquisa demandará um estudo mais pormenorizado. Logo, essa proposta inicial demonstrou por meio de uma arquitetura conceitual o procedimento necessário para que uma interface adaptável inteligente receba as informações de uma pessoa e forneça a esta um retorno o mais aceitável possível. Para trabalhos futuros pretende-se desenvolver testes e implementações em cada uma das camadas da arquitetura.

#### **Referências**

- Findlater, L., and McGrenere, J. (2004). A comparison of static, adaptive, and adaptable menus. *Proc. of ACM CHI 2004*, 6(1), pp. 89–96.
- Fortmann, F. and Lüdtko, A. (2013) uma interface adaptável a SA inteligente para auxiliar o controle de supervisão de um enxame de UAV, *IEEE Int. Conf. Ind. Informat.*, Bochum, Alemanha, 29 a 31 de julho de 2013, pp. 768 - 773.
- Huang, S., Ishikawa, M. and Yamakawa, Y. (2019) Human-robot interaction and collaborative manipulation with multimodal perception interface for human, *Proceedings of the 7th International Conference on Human-Agent Interaction (New York, NY, USA)*, HAI 19, Association for Computing Machinery, 2019, p. 289291.
- Soni, D., Nord, R. L. and Hofmeister, C. (1995) "Software Architecture in Industrial Applications," *17th International Conference on Software Engineering*, Seattle, Washington, USA, 1995, pp. 196-196.
- Wang, N., Di Nuovo, A., Cangelosi, A. and Jones, R. (2019) Temporal patterns in multi-modal social interaction between elderly users and service robot, *Interaction Studies* 20 (2019), no. 1, 4–24.