

VISÃO COMPUTACIONAL SUPERVISIONADA POR VOZ EM VEÍCULO TERRESTRE NÃO-TRIPULADO PARA APLICAÇÕES URBANAS, RURAIS E INDUSTRIAIS

Bruna Fernandes Medeiros¹;
Prof. Dr. Edio Roberto Manfio²

¹Aluna da Faculdade de Tecnologia de Garça

²Professor da Faculdade de Tecnologia de Garça

Área do Conhecimento: Linguagens de Programação

Palavras-chave: Visão Computacional; OpenCV; PLN; Robô de conversação.

INTRODUÇÃO

As aplicações gerais da Visão Computacional incluem Construção de modelos 3D, Padrões de Movimento, Segurança Automotiva, Reconhecimento Ótico de Caracteres, Medicina, Realidade Aumentada, Visão de Máquina entre outros (SZELISKI, 2010). Neste trabalho, algumas bibliotecas voltadas ao reconhecimento de formas básicas presentes em ambiente controlado como pés de mesas e cadeiras, bebedouros e outros móveis ou em campo como troncos de árvores, postes, arbustos e automóveis talvez sejam suficientes, tendo em vista que a implementação será feita junto a Veículo Terrestre não Tripulado - VTNT - cuja finalidade deslocar-se enquanto efetua monitoramento remoto.

A identificação de obstáculos, entretanto, será semissupervisionada, ou seja, em alguns casos, o VTNT dependerá do comando do operador para atuar. Os comandos serão feitos por voz ou por texto escrito utilizando uma Interface Humano-Computador - IHC - com Processamento de Linguagem Natural - PLN - e intermediada pelo robô de conversação Solar, pertencente ao Projeto Solar-Sima. Detectada a presença de obstáculos, o VTNT pode seguir uma rotina prevista ou solicitar ao operador a instrução mais adequada: virar, voltar, mudar a rota entre outros. Nesses casos, o PLN permite que a máquina atenda aos comandos usando o idioma do operador (MANFIO, 2016; MORENO, MANFIO, BARBOSA, 2015; RICH, 1993).

Válido salientar que esta implementação está alinhada às necessidades do Projeto Fapesp 2017/19136-0 outorgado em 2018, em desenvolvimento na própria IES e no IPMet de Bauru, e que tem por finalidade coletar dados de chuvas de granizo em áreas do Estado de São Paulo e Paraná, com ênfase na região de Garça-SP.

OBJETIVOS

Este trabalho tem como objetivo a implementação de um sistema de Visão Computacional supervisionada por voz em VTNT para aplicações urbanas, rurais e industriais associada a recursos IoT. Como objetivos específicos pretende-se que o VTNT amplie sua capacidade em desviar-se de obstáculos (i) em ambientes controlados - pés de mesas e cadeiras, bebedouros etc, (ii) em campo - troncos de árvores, postes, arbustos, automóveis etc e (iii) em condições adversas como a na presença de cabos, alambrados entre outros.

METODOLOGIA

Para a realização do presente projeto foram utilizados uma câmera, um computador compacto, um microcontrolador e um VTNT. A câmera é o modelo Pi e o microcontrolador foi o Arduino. O protótipo é um VTNT com tração integral e estrutura em madeira. O controle por PLN permitirá que o

deslocamento seja feito com segurança caso o algoritmo não resolva um problema inesperado como queda de barreiras

Atualmente, o operador consulta os dados ou envia os comandos ao VTNT por voz ou texto escrito a partir do robô Solar que está instalado no servidor. O robô, por sua vez, solicita os dados aos protótipos - VTNT ou buscador solar - que retornam imediatamente, seja via USB em ambiente controlado, seja via rádio quando à distância.

O Solar é um robô de conversação desenvolvido em C# que utiliza o reconhecedor o Coruja (FALA BRASIL, 2019). A busca de palavras-chave é feita por meio de tabelas *hash* (MORENO, 2017) e há um recurso de aprendizado de máquina - *Machine Learning* - em implementação usando Python (SILVA; MANFIO; GUERRA, 2018) por alunos da Fatec. Os recursos de PLN utilizados no Solar permitem que possa atender a comandos nas linguagens escrita e falada e responder por texto escrito ou por síntese de voz. Ele constitui-se em uma IHC bastante natural ao operador (MORENO et al., 2017) e estará presente em versões mais avançadas.

RESULTADOS OBTIDOS E DISCUSSÃO

Após instalação do Raspbian, foi possível utilizar a interface do mini-PC e da IDE para a programação da detecção. A IDE é a Thonny, já instalada no Raspbian, compatível com Python e com ela todo o projeto foi desenvolvido com a biblioteca aberta OpenCV, adequada para finalidades acadêmicas na área de Visão Computacional.

A Visão Computacional funcionará por meio da detecção de face, manipulando o código para reconhecer rostos em fotos específicas para o teste da funcionalidade. Chegou-se, então à etapa de fazer o treinamento do software com formas mais simples que rostos humanos para reduzir o tempo de processamento. Uma das formas escolhidas foi a ‘caneca’, tendo em vista que testes com o VTNT não necessitariam reconhecer semblantes de pessoas.

A etapa seguinte testou diferentes canecas físicas – e não mais imagens de treinamento – posicionadas diante da câmera. Os resultados foram ficando paulatinamente melhores à medida em que o software era treinado e os falsos positivos ficaram bastante reduzidos.

A partir de então, foram acrescentadas linhas de código que permitissem alterar o estado de apenas uma porta lógica do Raspberry sempre que ocorresse o reconhecimento. Isso permitiu fazer com que software e hardware agissem em conjunto toda vez que algo era colocado diante da câmera. Toda vez que o objeto – a caneca – é reconhecido, o VTNT desvia – à direita, por exemplo. Toda vez que não ocorre o reconhecimento, o VTNT retrocede e toma um novo caminho.

Embora esse método seja muito simplificado e dependa de programação extra para o microprocessador que controla os motores do VTNT, a resposta foi aceitável, pois em condições reais, se o obstáculo é reconhecido e não oferece perigo, basta desviar. Em situação contrária, como não é identificável, melhor afastar-se e passar longe.

CONSIDERAÇÕES FINAIS

Até esse ponto da pesquisa foi possível verificar que a aplicação de um projeto dessa categoria oferece grau significativo de viabilidade, tendo em vista que a Visão Computacional supervisionada por voz em VTNT para aplicações urbanas, rurais e industriais associada a recursos IoT pode proporcionar grande capacidade de autonomia a processos que envolvem dispositivos automotores não tripulados. Mais pesquisas utilizando os recursos aqui adotados, entretanto, podem prover maior precisão no processo.

REFERÊNCIAS BIBLIOGRÁFICAS

FALA BRASIL. Disponível em: <<http://labvis.ufpa.br/falabrasil/>>. Acesso em: 02 jan. 2019.

MANFIO, Edio Roberto. **Avaliação de dispositivos acionados por voz e texto para o Português Brasileiro**. 2016. 121 f. Tese (Doutorado em Estudos da Linguagem) Programa de Pós-Graduação em Estudos da Linguagem – PPGEL. Universidade Estadual de Londrina – UEL, Londrina, 2016.

MANFIO, Edio Roberto; MORENO, Fabio Carlos; BARBOSA, Cinthyan Renata Sachs Camerlengo de. Tecnologia Interativa Conversacional sobre Assuntos Linguísticos - Tical: Linguagem e Significação. In: IX Seminário de Estudos sobre Linguagem e Significação e X Simpósio de Leitura da UEL "Convenções e Ousadias da Linguagem". **Caderno de Resumos**. Londrina: UEL, 2014a, p. 54-55.

MANFIO, Edio Roberto; MORAIS, Marcos Vinícius Bueno de; GUERRA, Marcos Paulo Guimarães; MORENO, Fabio Carlos. Processamento de Linguagem Natural aplicado a Veículo Terrestre Não Tripulado para monitoramento remoto e Gestão de Energia. In **Revista Regrad** (no prelo). 2019.

MORENO, Fabio Carlos; GUERRA, Marcos Paulo Guimarães; MANFIO, Edio Roberto; MORAIS, Marcos Vinícius Bueno de. IHCs dedicadas a energias sustentáveis. In **Revista ef@tec**.. ISSN: 2317-451X, vol. 7. n. 1, 2017. Garça, 2017.

MORENO, F. C. **Visual Taks**: ferramenta para analisar a eficácia de buscas das funções hash em um Léxico para Língua Natural. Dissertação (Mestrado em Ciência da Computação). Universidade Estadual de Londrina, 2017.

RICH, Elaine. **Inteligência Artificial**. Tradução Maria Cláudia Santos Ribeiro Ratto. São Paulo: Makron Books, 1993.

SZELISKI, Richard. **Computer Vision: Algorithms and Applications**. Springer: August 18, 2010.