

PROJETO DE PESQUISA

Processos atencionais e aprendizado de máquina para sistemas robóticos

Aluno: Erik de Godoy Perillo
Orientadora: Esther Luna Colombini

Resumo

Entender o ambiente ao seu redor é uma tarefa fundamental para o desafio de se obter máquinas autônomas que interagem com o mundo de forma semelhante à nossa. A alta dimensionalidade dos dados captados por sensores usados para este fim é em geral problemática, pois muitas vezes há redundância e irrelevância de informação. Neste projeto, propomos o uso de processos atencionais em cascata com aprendizado de máquina a fim de se obter um sistema de identificação de classificação de entidades que seja eficiente, preciso e geral. Por fim, planeja-se implementar uma estrutura que permita o uso das técnicas por projetos robóticos em geral com o uso de GPUs embarcadas.

Universidade Estadual de Campinas

22 de abril de 2016

1 Introdução

Máquinas capazes de realizar tarefas complexas, perigosas ou maçantes são objeto de interesse e desejo do ser humano há tempos. Robôs industriais já dominam os mais variados setores de produção. Não apenas parados em linhas de montagem, sistemas atuais podem locomover-se em regiões uniformes e previsíveis – como galpões de estoques – trabalhando vinte e quatro horas por dia [14]. Um desafio ainda em aberto, entretanto, é a concepção de robôs que lidam com o imprevisível, reagindo de forma apropriada às mais diversas situações e ambientes do mundo real.

Sistemas que interagem com o ambiente, objetos e pessoas com uma variedade de maneiras semelhante à nossa têm o potencial de ser incrivelmente benéficas para a sociedade. Diversos desafios fascinantes estão associados ao processo de concepção de máquinas gerais desse tipo. Um deles, certamente, é a compreensão. Entender o que lhe cerca é fundamental para uma interação complexa com o resto do mundo.

Ter ciência de sua relação espacial com outros objetos e suas possíveis interações com cada um deles é essencial para a navegação de robôs por casas, ruas e até mesmo locais que seriam impossíveis para um ser humano estar. Essa é uma das tarefas principais que máquinas como carros autônomos, robôs caseiros e de resgate devem cumprir bem. Técnicas que utilizam sensores como câmeras [9] e LIDAR [7] têm sido usadas com sucesso para identificação e classificação de entidades em um ambiente com o qual o robô interage.

1.1 Motivação

Um problema de sensores usados para tarefas mais complexas de navegação em geral – como câmera e LIDAR – é que seus volumes de dados podem ser demasiadamente grandes. Em um momento qualquer, é improvável que toda a informação provida de um sensor seja necessária. Lidar com todo o conjunto de dados para depois determinar qual parte é relevante pode ser uma tarefa muito computacionalmente custosa. Isso é especialmente crítico para sistemas robóticos que precisam de respostas rápidas para interagir com o ambiente em que estão inseridos.

Além disso, o *hardware* utilizado em robôs autônomos tem sido por muitos anos limitado pelas restrições de em geral ter de ser embarcado e compacto. Avanços recentes têm mudado isso, entretanto. Além do aumento da eficiência e diminuição do tamanho de CPUs impulsionado pelo mercado *mobile*, esforços têm sido feitos para a concepção de computadores com pro-

cessadores gráficos (GPUs) embarcados. A *NVIDIA* lançou recentemente os modelos *Jetson TK1* e *Jetson TX1* [3], que são sistemas computacionais completos com GPUs poderosas e energeticamente eficientes neles inclusas, desenvolvidos especialmente para serem embarcados em projetos móveis. O uso de GPUs em robótica abre a possibilidade do uso de técnicas mais sofisticadas e computacionalmente caras que até então não podiam ser exploradas. Sua natureza paralela permite um grande benefício de operações que podem ser divididas em subproblemas e realizados ao mesmo tempo. Esse é o caso de diversas tarefas como processamento de sinais e técnicas de aprendizado de máquina [1], por exemplo.

1.2 Processos atencionais e aprendizado de máquina

O fato de sensores darem uma quantidade muito grande de informação que nem sempre é necessária motiva uma abordagem inspirada no ser humano. Não damos foco a todas as conversas que escutamos ou em tudo que vemos, mas sim apenas na parte que nos convém. Técnicas que fazem uso de processos atencionais têm sido usadas para a segmentação de áreas de relevância na percepção e entendimento do ambiente com sucesso [13] [2]. Com essas abordagens, pode-se filtrar muito do que seriam dados irrelevantes ou redundantes, aumentando consideravelmente a eficiência computacional dessas técnicas.

Não só o foco é necessário: em ambientes dos mais variados é preciso, além de identificar, classificar entidades a fim de interagir com elas de forma apropriada. O aprendizado de máquina tem sido uma área de grande sucesso nessa tarefa. Diversas técnicas podem ser utilizadas, como *SVMs* [8] e redes neurais [6] a fim de se classificar entidades em geral com alta taxa de acertos.

O que se busca, então, é uma mescla dessas duas técnicas: o uso dos processos atencionais em um estágio inicial de processamento pode atuar como um filtro na informação, passando adiante somente o que é relevante para o estágio de classificação com uso de técnicas robustas de aprendizado de máquina. O que se espera é que tal abordagem possibilite uma percepção eficiente, geral e complexa do ambiente para sistemas autônomos, o que os possibilitaria realizar tarefas mais sofisticadas e traria benefícios objetivos para a área da percepção em robótica.

2 Objetivos

Os objetivos deste trabalho são:

- Obter técnicas que permitam o entendimento de ambientes diversos para a navegação de sistemas robóticos autônomos por meio do uso de processos atencionais e aprendizado de máquina.
- Realizar uma avaliação de efetividade e eficiência de nossa abordagem e compará-la com outras técnicas utilizadas atualmente.
- Construção de uma plataforma para uso em GPUs embarcadas com ferramentas que permitam o uso de nossa técnica desenvolvida para aplicações variadas de robótica.

3 Materiais e Métodos

3.1 Caso de uso

O desafio *Trekking* [10] é uma categoria da série de competições de robótica realizada pela *Robocore* [11]. Nessa categoria, o robô deve completar um percurso em um campo de futebol passando por diversas bases em uma ordem pré-definida. Em uma certa região do campo, há obstáculos dos quais o robô deve desviar. O percurso deve ser feito no menor tempo possível.

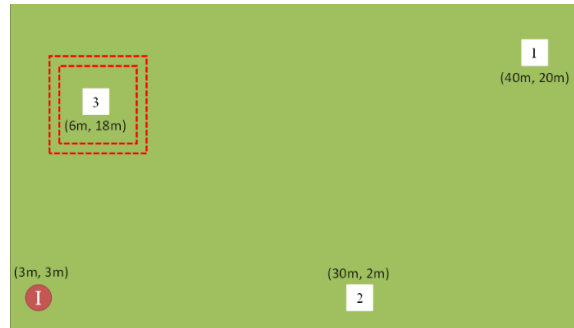


Figura 1: O ambiente de locomoção do robô na categoria Trekking

A *Equipe Phoenix de Robótica da Unicamp* [4] é formada por estudantes de graduação de diversos cursos de Engenharia e Computação da Universidade Estadual de Campinas. Na equipe, desenvolvem-se variados projetos de robótica, com foco principal em robôs para competir nas categorias das competições da *Robocore*. A equipe desenvolve atualmente um robô para participar na categoria *Trekking*, nomeado *Projeto Piranha*. O robô consiste em um carro elétrico com tração nas quatro rodas e locomoção por motores *brushless*. O motor é controlado por meio de um conjunto da *NXP* composto

da placa *FRDM-K64F* e da placa *FRDM-STBC-AGM01* [12], que contém diversos sensores inerciais como magnetômetro, giroscópio e acelerômetro. Há também um kit de desenvolvimento com GPU embarcada *NVIDIA Jetson TX1*. A *Jetson TX1* possui um processador *ARM Quad-Core Cortex-A57*, 4GB de memória *RAM* e uma GPU *NVIDIA Maxwell* com 256 núcleos de processamento [5]. Seu sistema operacional é o Ubuntu com *kernel Linux*. Duas câmeras *Logitech C270* são os sensores principais usados pela *Jetson TX1* para a ambientação do robô no campo.



Figura 2: O robô do projeto Piranha

A pesquisa usará o robô *Piranha* como caso de uso para avaliação da abordagem a ser desenvolvida, com o uso da *Jetson TX1* como plataforma para implementação das técnicas. O robô usa a câmera como principal sensor. Assim, a pesquisa foca em imagens como o sinal a ser processado e analisado.

A ideia é aplicar as técnicas a se obter para o reconhecimento de entidades na imagem, como terreno o qual o robô pode andar, obstáculos do qual deve desviar e bases às quais deve-se chegar. Soluções em processos atencionais serão desenvolvidas separadamente do estágio de classificação de objetos com aprendizado de máquina. Após isso, a união dos dois será feita.

3.2 Avaliação da abordagem proposta

Será feita uma comparação com três técnicas diferentes:

1. Uma abordagem específica é utilizada. Um simples *threshold* de cor pode ser capaz de identificar objetos imagem e classificá-los dado que diferentes classes de objetos têm diferentes cores.

2. Aprendizado de máquina puro para reconhecimento de objetos. Nessa abordagem, será dado foco igual a toda a informação. Assim, todo o conjunto deverá ser analisado pelos algoritmos.
3. Processos atencionais são utilizados em um estágio anterior ao de classificação com aprendizado de máquina, com a intenção de selecionar subconjuntos de toda informação dada pelo sensor que tenham maior relevância para serem analisados no estágio de classificação.

Com essa comparação, espera-se avaliar cada técnica pelos seguintes critérios:

- Desempenho: o custo computacional da técnica.
- Eficiência: quão corretamente a técnica identifica e classifica as entidades de interesse.
- Generalidade: possibilidade da técnica ser utilizadas em outros cenários.

3.3 Construção de plataforma para navegação usando GPUs

Com as técnicas obtidas pelo projeto, espera-se construir uma plataforma de uso geral para robôs autônomos que usem GPUs embarcadas. *frameworks* de aprendizado de máquina serão estudados para se decidir quais usar. Toda biblioteca será escolhida de forma que se possa utilizar o processamento da GPU. Por meio da linguagem de programação *C++* juntamente com *Python*, um ambiente de treinamento e rotinas de classificação para uso em imagens será desenvolvido.

4 Cronograma

Tarefa/mês	Jul	Ago	Set	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun
Revisão Bibliográfica	x											
Modelagem de processos atencionais		x	x									
Modelagem aprendizado de máquina			x	x								
Coleta de imagens para treinamento					x							
Integração dos sistemas						x						
Implementação de técnicas							x	x				
Testes comparativos									x			
Construção de plataforma para GPU										x	x	
Relatório final												x

Referências

- [1] Dan C. Ciresan et al. *Flexible, High Performance Convolutional Neural Networks for Image Classification*. 2011.
- [2] E. L. Colombini, A. da Silva Simões e C. H. Costa Ribeiro. “An Attentional Model for Autonomous Mobile Robots”. Em: *IEEE Systems Journal* PP.99 (2016), pp. 1–12. ISSN: 1932-8184. DOI: 10.1109/JSYST.2015.2499304.
- [3] *Embedded Systems*. 2016. URL: <http://www.nvidia.com/object/embedded-systems.html> (acesso em 19/04/2016).
- [4] *Equipe Phoenix de Robótica da Unicamp*. 2016. URL: <http://phoenixunicamp.com.br> (acesso em 19/04/2016).
- [5] *Jetson TX1 Development Kit*. 2016. URL: <https://developer.nvidia.com/embedded/buy/jetson-tx1-devkit> (acesso em 19/04/2016).

- [6] Alex Krizhevsky, Ilya Sutskever e Geoffrey E. Hinton. “ImageNet Classification with Deep Convolutional Neural Networks”. Em: *Advances in Neural Information Processing Systems 25*. Ed. por F. Pereira et al. Curran Associates, Inc., 2012, pp. 1097–1105. URL: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.
- [7] J. Levinson et al. “Towards fully autonomous driving: Systems and algorithms”. Em: *Intelligent Vehicles Symposium (IV), 2011 IEEE*. Jun. de 2011, pp. 163–168. DOI: 10.1109/IVS.2011.5940562.
- [8] Larry M. Manevitz e Malik Yousef. “One-class Svms for Document Classification”. Em: *J. Mach. Learn. Res.* 2 (mar. de 2002), pp. 139–154. ISSN: 1532-4435. URL: <http://dl.acm.org/citation.cfm?id=944790.944808>.
- [9] N. P. Papanikolopoulos, P. K. Khosla e T. Kanade. “Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision”. Em: *IEEE Transactions on Robotics and Automation* 9.1 (fev. de 1993), pp. 14–35. ISSN: 1042-296X. DOI: 10.1109/70.210792.
- [10] *Regras robô Trekking*. 2016. URL: https://www.robocore.net/upload/attachments/robocore__regras_robo_trekking_610.pdf (acesso em 19/04/2016).
- [11] *Robocore*. 2016. URL: https://www.robocore.net/modules.php?name=GR_Eventos (acesso em 19/04/2016).
- [12] *Sensor Toolbox Development Platform*. 2016. URL: <http://goo.gl/kIwBFn> (acesso em 19/04/2016).
- [13] C. Siagian e L. Itti. “Rapid Biologically-Inspired Scene Classification Using Features Shared with Visual Attention”. Em: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.2 (fev. de 2007), pp. 300–312. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2007.40.
- [14] Peter R. Wurman, Raffaello D’Andrea e Mick Mountz. “Coordinating Hundreds of Cooperative, Autonomous Vehicles in Warehouses”. Em: *Proceedings of the 19th National Conference on Innovative Applications of Artificial Intelligence - Volume 2*. IAAI’07. Vancouver, British Columbia, Canada: AAAI Press, 2007, pp. 1752–1759. ISBN: 9781577353232. URL: <http://dl.acm.org/citation.cfm?id=1620113.1620125>.