

Gabriel Takaoka Nishimura

SLAM 3D e RGB-D: Uma abordagem com FPGA

Plano de Pesquisa apresentado ao Departamento de Engenharia de Computação e Sistemas Digitais da Escola Politécnica da Universidade de São Paulo para processo de seleção de Mestrado do segundo semestre de 2017

Universidade de São Paulo - USP

Escola Politécnica

Engenharia Elétrica com ênfase em Computação

Orientador: Prof. Dr. Bruno Carvalho de Albertini

São Paulo

2016

Resumo

Apresentação tema, popularidade SoC com duas pilhas, apesar disso (apenas avaliações comparativas, ao invés de cooperativas), mecanismo de cooperação, análises numéricas e testes para avaliar

SLAM 3D tem várias aplicações, algumas mais baratas, como o RGB-D com Kinect. Em sistemas **Palavras-chave:** Nuvem de pontos, Localização e Mapeamento Simultâneos,

RGB-D, FPGA

Introdução

Redes neurais convolucionais são redes neurais artificiais em que se aplica convolução em pelo menos uma de suas camadas. Normalmente a arquitetura de uma rede neural convolucional é composta pela camada convolucional e pela camada de sub amostragem.

A camada convolucional realiza a operação de convolução, nela, existem neurônios que estão conectados em uma região reduzida da camada de entrada. Esses neurônios ficam organizados em múltiplas subcamadas paralelas. Cada neurônio dentro de uma sub camada compartilha os mesmos parâmetros de pesos. O resultado dessa camada é a extração de diferentes características da camada de entrada. Como exemplo será levado em conta um classificador CNN de imagens. Na primeira camada de convolução - a entrada - é possível a extração de características como bordas, linhas e cantos.

A camada de sub amostragem tem a ideia de gerar características invariantes de tradução. Essas características a partir do cálculo de estatísticas efetuadas sobre as subcamadas gerada na camada convolucional. Considerando-se novamente um classificador CNN de imagens como exemplo, sua função é diminuir o tamanho da imagem para encontrar possíveis padrões que possam ficar mais evidentes. CNNs são utilizadas em várias aplicações visuais como classificação de imagens, detecção de objetos e estimação de idade.

A utilização de CNNs em reconhecimento de imagens deve-se a algumas vantagens que ela possui:

- A detecção de imagens usando uma CNN é resistente a distorções causadas na imagem devido a vários fatores como, por exemplo, a variação de lentes de câmeras, condições diferentes de iluminação, diferentes poses, oclusões parciais e mudanças de posicionamento horizontal e vertical.

- Menos requisitos de memória em relação a uma rede neural padrão.

- Maior facilidade em treinar a rede neural em comparação a uma rede neural padrão causada pela diferença de número de parâmetros, que no caso de uma CNN é menor.

É interessante disponibilização de softwares de reconhecimento de imagens dentro do âmbito de aparelhos celulares já que uma porcentagem expressiva, por exemplo, da população brasileira 77,9 (pessoas acima de 10 anos em grandes regiões) possui aparelho celular e é conveniente poder classificar fotos tiradas pelo próprio aparelho celular

Existe o aplicativo do ebird merlin por exemplo

Só que existem problemas críticos em aparelhos móveis em termos de poder

computacional, bateria e capacidade da memória. Então existe a necessidade de adaptar a rede neural convolucional para encaixar com esse problemas críticos do aparelho celular exemplos:

Objetivos

O objetivo desse projeto é a realização do estudo de técnicas e algoritmos para obter componentes relevantes a SLAM 3D (distância percorrida e ângulo de giro), analisando Núvens de Pontos geradas por sensores RGB-D em co-processador usando arquitetura FPGA. Por fim, o trabalho deve criar um sistema SLAM 3D com baixo consumo energético e computacional - utilizando as técnicas estudadas anteriormente - para implementação e testes em ambiente físico com robô.

1 Metodologia

Com o intuito de implementar um sistema de navegação robótica utilizando SLAM 3D, propõe-se um método que consiste em duas etapas. A primeira etapa

Uma placa FPGA será utilizada para

2 Cronograma

Referências

- BAO, G. *On Simultaneous Localization and Mapping inside the Human Body (Body-SLAM)*. Tese (Doutorado) — WORCESTER POLYTECHNIC INSTITUTE, 2014. Nenhuma citação no texto.
- NIKOLIC, J. et al. A synchronized visual-inertial sensor system with fpga pre-processing for accurate real-time slam. In: IEEE. *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. [S.l.], 2014. p. 431–437. Nenhuma citação no texto.
- RUSU, R. B.; COUSINS, S. 3D is here: Point Cloud Library (PCL). In: *IEEE International Conference on Robotics and Automation (ICRA)*. Shanghai, China: [s.n.], 2011. Nenhuma citação no texto.
- TERTEI, D. T.; PIAT, J.; DEVY, M. Fpga design and implementation of a matrix multiplier based accelerator for 3d ekf slam. In: IEEE. *ReConFigurable Computing and FPGAs (ReConFig), 2014 International Conference on*. [S.l.], 2014. p. 1–6. Nenhuma citação no texto.
- WOLCOTT, R. W.; EUSTICE, R. M. Robust lidar localization using multiresolution gaussian mixture maps for autonomous driving. *The International Journal of Robotics Research*, SAGE Publications Sage UK: London, England, p. 0278364917696568, 2017. Nenhuma citação no texto.