

Public Abstract

First Name:Alex

Middle Name:

Last Name:Yang

Adviser's First Name:Grant

Adviser's Last Name:Scott

Co-Adviser's First Name:

Co-Adviser's Last Name:

Graduation Term:FS 2017

Department:Computer Science

Degree:MS

Title:RELATIVE DEPTH ESTIMATION FROM SINGLE MONOCULAR IMAGES WITH DEEP CONVOLUTIONAL NETWORK

Depth estimation from single monocular images is a theoretical challenge in computer vision as well as a computational challenge in practice. This thesis addresses the problem of depth estimation from single monocular images using a deep convolutional neural fields framework; which consists of convolutional feature extraction, superpixel dimensionality reduction, and depth inference. Data was collected using a stereo vision camera, which generated depth maps through triangulation that are paired with visual images. The visual image (input) and computed depth map (desired output) are used to train the model, which has achieved 83% test accuracy at the standard 25% tolerance. The problem has been formulated as depth regression for superpixels and our technique is superior to existing state-of-the-art approaches based on its demonstrated its generalization ability, high prediction accuracy, and real-time processing capability. The proposed method identifies convolutional feature extraction, superpixel dimensionality reduction and depth inference as three key components. We utilize the VGG-16 deep convolutional network as feature extractor and conditional random fields depth inference. We have leveraged a multi-phase training protocol that includes transfer learning and network fine-tuning lead to high performance accuracy. Our framework has a robust modular nature with capability of replacing each component with different implementations for maximum extensibility. Additionally, our GPU-accelerated implementation of superpixel pooling has further facilitated this extensibility by allowing incorporation of feature tensors with flexible shapes and has provided both space and time optimization. Based on our novel contributions and high-performance computing methodologies, the model achieves a minimal and optimized design. It is capable of operating at 30 fps; which is a critical step towards empowering real-world applications such as autonomous vehicle with passive relative depth perception using single camera vision-based obstacle avoidance, environment mapping, etc.