DEPTH ESTIMATION: A MACHINE LEARNING BASED APPROACH

NIDHI CHAHAL

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

SEPTEMBER 2017

DEPTH ESTIMATION: A MACHINE LEARNING BASED APPROACH

by

NIDHI CHAHAL

DEPARTMENT OF ELECTRICAL ENGINEERING

Submitted

in fulfillment of the requirements of the degree of Doctor of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

SEPTEMBER 2017

Certificate

This is to certify that the thesis titled **Depth Estimation: A Machine Learning Based Approach** being submitted by **Ms. Nidhi Chahal** to the **Department of Electrical Engineering**, Indian Institute of Technology Delhi, for the award of **Doctor of Philosophy** is a record of bona-fide research work carried out by her under our guidance and supervision. In our opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree. The work presented in this thesis has not been submitted elsewhere, either in part or full, for the award of any other degree or diploma.

Dr. Mona Mathur
Adjunct Faculty
Department of Electrical Engineering
Indian Institute of Technology Delhi
New Delhi - 110 016

Professor Santanu Chaudhury
Department of Electrical Engineering
Indian Institute of Technology Delhi
New Delhi - 110 016

To my family

Acknowledgments

Firstly, I would like to express my sincere gratitude to my supervisors Professor Santanu Chaudhury and Dr. Mona Mathur for their continuous support and the stream of ideas that kept me occupied. Their expertise, vast knowledge, skills and patience added considerably to my doctoral experience. My regards to Prof. Santanu for his constant supervision which encouraged me in accomplishing my research work through all these years. I doubt that I will ever be able to convey my appreciation fully, but I owe him my eternal gratitude. Dr. Mona provided me with direction, technical support and became more of a friend and mentor, than a faculty.

I would like to thank rest of my thesis committee: Prof. S.D. Joshi, Dr. Sumantra Dutta Roy and Prof. Subhashis Banerjee for their appreciation, insightful comments and guidance. I would also like to thank my colleagues and friends for their support and encouragement through out my research work. All the discussions with them, exchange of knowledge, ideas and technical skills have been very helpful and motivational.

Finally, I must express profound gratitude to my parents who taught me good values that really matter in life and for their continuous mental support.

A very special thanks goes to my son for being such a sincere boy and always cheering me up. His love motivated me to work hard for providing him better future. Last, but not least, I am thankful to my husband for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of my entire research and writing this thesis. This accomplishment would not have been possible without support of my family. Thank you very much, everyone.

Abstract

The machine learning approach is explored in this thesis for depth estimation of two dimensional images involving various features and concepts. The depths are generated for stereo pairs using fixed point model which is obtained by learning ground truth depths of training data. For the combination of multiple cues, the depths from stereo and monocular cues are used as input features for the generation of prediction function which is the specialty of our approach for depth estimation. The accurate and reliable depths are predicted from proposed learning framework for new stereo pairs.

A novel approach of manifold learning is proposed for reliable depth estimation of single images. The deep learning process is applied which extracts CNN features from Caffe model which is available online. The features generated from this process are more authentic as compared to hand crafted features such as direct intensities of the image pixels. The complete model is trained using initial depths obtained from manifold and ground truth depths of the training data in a fixed point learning framework which gives refined and dense depth maps. The results are evaluated using various quantitative and visual measures and compared with ground truth depths of test images, if available.

For the test images of different feature distributions than training images, transfer learning is used in our work which is novel approach in the field of depth estimation. For this, manifold alignment is applied in which manifolds of source and target domains are projected to new space so that the difference in structures of both domains is reduced. After manifold mapping, the depths are generated from manifold learning and finally fixed point supervised learning process is used for prediction of refined depth maps.

सार

द्वि-आयामी के गहराई अनुमान के लिए इस शोध में मशीन सीखने की प्रक्रिया का पता लगाया गया है विभिन्न विशेषताओं और अवधारणाओं को शामिल करने वाली छवियां। गहराई स्टीरियो के लिए उत्पन्न होती है नियत बिन्दु मॉडल का उपयोग करके जोड़े जो प्रशिक्षण के आधार सच्चे गहराई सीखकर प्राप्त की जाती हैं जानकारी। कई संकेतों के संयोजन के लिए, स्टीरियो और मोनोकुलर संकेतों की गहराई होती है भविष्यवाणी समारोह की पीढ़ी के लिए इनपुट विशेषताओं के रूप में उपयोग किया जाता है जो कि हमारी विशेषता है। नए स्टीरियो जोड़े के गहराई अनुमान के लिए दृष्टिकोण प्रस्तावित से सटीक और विश्वसनीय गहराई का अनुमान लगाया गया है।

एकल के विश्वसनीय गहराई अनुमान के लिए कई गुना सीखने का एक नया दृष्टिकोण प्रस्तावित है इमेजिस। गहरी सीखने की प्रक्रिया को लागू किया जाता है जो कैफ मॉडल से सीएनएन सुविधाओं को निकालता है जो ऑनलाइन उपलब्ध है। इस प्रक्रिया से उत्पन्न सुविधाओं की तुलना में अधिक प्रामाणिक हैं जैसे कि छवि पिक्सेल की सीधा तीव्रता। तैयार की गई सुविधाओं को हाथ में लेना पूरा मॉडल प्रशिक्षण के कई गुना और जमीन सत्य गहराई से प्राप्त प्रारंभिक गहराई का उपयोग करके प्रशिक्षित किया गया है एक निश्चित बिन्दु सीखने के ढांचे में डेटा जो परिष्कृत और घने गहराई नक्शे देता है। परिणाम विभिन्न मात्रात्मक और दृश्य उपायों का उपयोग करके मूल्यांकन किया जाता है और जमीन की सच्चाई के साथ तुलना की जाती है परीक्षण छवियों की गहराई, यदि उपलब्ध हो।

प्रशिक्षण छवियों, हस्तांतरण सीखने की तुलना में विभिन्न फीचर वितरण की परीक्षण छवियों के लिए हमारे काम में प्रयोग किया जाता है जो गहराई अनुमान के क्षेत्र में उपन्यास है। इसके लिए, कई गुना संरेखण लागू होता है जिसमें स्रोत और लक्ष्य डोमेन के कई गुना नए प्रोजेक्ट किए जाते हैं अंतरिक्ष ताकि दोनों डोमेन के ढांचे में अंतर कम हो। मैनिफ़ोल्ड मैपिंग के बाद, गहराई कई गुना सीखने से उत्पन्न होती है और आखिरकार निश्चित बिंदु पर्यवेक्षण शिक्षण होता है प्रक्रिया परिष्कृत गहराई नक्शे के पूर्वानुमान के लिए प्रयोग किया जाता है।

Contents

List of Figures

List of Tables