UNIVERSITI TEKNIKAL MALAYSIA MELAKA

3D INTRINSIC SCENE CHARACTERISTICS EXTRACTION FRAMEWORK FOR A SINGLE IMAGE

HABIBULLAH AKBAR

DOCTOR OF PHILOSOPHY

2016

HABIBULLAH AKBAR

DOCTOR OF PHILOSOPHY

2016

Faculty of Information and Communication Technology

3D INTRINSIC SCENE CHARACTERISTICS EXTRACTION FRAMEWORK FOR A SINGLE IMAGE

Habibullah Akbar

Doctor of Philosophy

2016

3D INTRINSIC SCENE CHARACTERISTICS EXTRACTION FRAMEWORK FOR A SINGLE IMAGE

HABIBULLAH AKBAR

A thesis submitted
in fulfillment of the requirements for the degree of Doctor of Philosophy

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2016

DECLARATION

I declare that this thesis entitled "3D Intrinsic Scene Characteristics Extraction Framework for a Single Image" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature :

Name : Habibullah Akbar

Date :

APPROVAL

I hereby declare that I have read this thesis and, in my opinion, this thesis is sufficient in terms of scope and quality for the award of Doctor of Philosophy.

Signature	:	
Supervisor Name	:	Prof. Dr. Nanna Suryana Herman
Date	:	

DEDICATION

To my beloved mother, father, wife, daughter and sons

ABSTRACT

Three-Dimensional (3D) shape reconstruction is an important area of computer vision research because it has numerous potential applications from entertainment production to industrial inspection and clinical analysis. Existing 3D Intrinsic Scene Characteristics (3D-ISCs) extraction methods for a single image have focused solely on estimating diffuse characteristics, i.e. 3D shape, illumination, and reflectance models, of an object. As a result, they have neglected the specular characteristic, the shiny areas of a glossy surface. In reality, many real-world objects emit both specular and diffuse reflections, and thus the specular component may decrease the performance of the 3D-ISCs methods. This study has developed a framework to extract all of these characteristics. The framework combines a Specular Removal (SR) method and a Shape, Illumination, and Reflectance From Shading (SIRFS) method under a Bidirectional Reflectance Distribution Function (BRDF) model. Since the previous SR methods suffered from hue-saturation ambiguity, they are not suitable for this framework. To solve this problem, two SR methods were developed, evaluated, and compared with the standard SR methods. The proposed SR methods are referred as Chaotic Segmentation (CS) and Sparse Coding (SC) methods. To combine the SR and SIRFS methods, two BRDF models were also developed, evaluated, and compared. These models are referred as Modified Dichromatic Reflectance (MDR) and Modified Blinn-Phong (MBP) models. The performances of the proposed SR methods and the BRDF models for extracting 3D-ISCs were evaluated based on public datasets. The results showed that the SC method was more satisfactory compared to the CS and the benchmark method (iterative method). The accuracies of the diffuse and specular characteristics were improved by 7.6% and 53.5% respectively. Moreover, the combination of SC method and MDR model was capable of outperforming the SIRFS method. The computational speed was 19.2% faster. Meanwhile, the average accuracies of depth, surface normal, illumination, shading, and reflectance were improved by 11.4%, 6.5%, 50.5%, 35.2%, and 5.1% respectively. This study indicates that the specular reflection is an important aspect of 3D reconstruction from a single image. The proposed framework has also made considerable improvements in terms of accuracy and computational time of extracting 3D-ISCs.

ABSTRAK

Pembinaan semula bentuk Tiga-Dimensi (3D) merupakan bidang penting dalam penyelidikan visi komputer kerana ia mempunyai banyak aplikasi yang berpotensi daripada pengeluaran hiburan kepada pemeriksaan industri dan analisis klinikal. Kewujudan kaedah-kaedah pengekstrakan Ciri-ciri Paparan Instrinsik 3D (CPI-3D) untuk imej tunggal hanya tertumpu kepada ciri-ciri peresapan seperti model bentuk 3D, pencahayaan dan pembalikan sesuatu objek. Kaedah-kaedah ini mengabaikan ciri spekular (kawasan permukaan objek yang berkilat). Dalam keadaan realiti, kebanyakan objek sebenar memancarkan kedua-dua pantulan spekular dan peresapan, dan ini menjadikan komponen spekular mungkin mengurangkan prestasi kaedah CPI-3D. Kajian ini telah membangunkan kerangka kerja untuk mengekstrak CPI-3D. Kerangka kerja ini menggunakan kaedah Penyingkiran Spekular (PS) dan kaedah Bentuk, Pencahayaan serta Kepantulan dari Pembayangan (BPKP) di bawah model Fungsi Taburan Kepantulan Dwiarah (FTKD). Oleh kerana kaedah-kaedah PS sebelum ini mengalami kesamaran ketepuan warna, ia tidak sesuai untuk kerangka kerja ini. Bagi menyelesaikan masalah tersebut, dua kaedah PS telah dibangunkan, dinilai dan dibandingkan dengan kaedah standard PS. Kedua-duanya merujuk kepada Pensegmenan Camuk (PC) dan Pengkodan Bersela (PB). Bagi menggabungkan kaedah PS dan BPKP, dua model FTKD telah dibangunkan, dinilai dan dibandingkan. Kedua-duanya merujuk kepada Pengubahsuaian Kepantulan Dikromatik (PKD) dan Pengubahsuaian Blinn-Phong (PBP). Prestasi kaedah PS dan model FTKD yang dicadangkan untuk mengekstrak CPI-3D telah dinilai berdasarkan kepada set data awam. Keputusan menunjukkan kaedah PB lebih memuaskan berbanding PC dan kaedah penanda aras (kaedah lelaran). Ketepatan peresapan dan specular telah meningkat masing-masing sebanyak 7.3% dan 53.5%. Gabungan kaedah PB dan model PKD mampu menyaingi kaedah BPKP. Kelajuan pengiraan telah meningkat sebanyak 19.2% manakala ketepatan purata kedalaman, permukaan normal, pencahayaan, pembayangan serta kepantulan telah meningkat masing-masing sebanyak 11.4%, 6.5%, 50.5%, 35.2% dan 5.1%. Kajian ini menandakan bahawa kerangka kerja pengekstrakan CPI-3D cadangan telah menunjukkan kemajuan besar dari segi ketepatan dan masa pengiraan.

ACKNOWLEDGEMENTS

In the name of Allah, The Most Gracious and The Most Merciful. I wish to express my sincerest appreciation and gratitude to Prof. Dr. Nanna Suryana Herman for his valuable advice, patience, and support during my Ph.D. journey at Universiti Teknikal Malaysia Melaka. With his vision and guidance, this thesis may never have been completed on time. My deepest appreciation also to Prof. Datuk Dr. Shahrin Sahib for his help in developing motivation, ideas, and for willingly sharing his expertise and in-depth knowledge. I appreciate their available time for solid discussion despite their tight schedule. I wish to thank all my fellow postgraduate students for their discussions, support, and friendship. I also thank the staff, lecturers, and technicians in the Faculty of Information and Communication Technology for their help, cooperation, and encouragement. My deepest appreciation also to my parents, Tasman Abbas and Djasminar Anwar, for their prayer, love, and care. I pray their efforts will be rewarded by Allah in this world and the hereafter. I would like to thank my brother, Abdullah Denovan, and my sister, Nila Novari. I would also like to thank my wife (Nancy Kuspriyati), my daughter (Aisyah Syarafana), and my sons (Izdihar Zidan and Ashraf Ruwaifi) for their love, support, and patience. This work was supported by Universiti Teknikal Malaysia Melaka under Zamalah Scheme.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES