FAST DEPTH CODING IN 3D-HEVC USING DEEP LEARNING

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ELECTRONIC AND INFORMATION ENGINEERING
OF THE HONG KONG POLYTECHNIC UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Zhen-xiang WANG
October 2017

CERTIFICATE OF ORIGINALITY

I hereby declare that this dissertation is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written nor material which has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

______________________________(Signed)

______________________________(Date)

Abstract

The 3D Extension of the High Efficiency Video Coding standard (3D-HEVC), which has been finalized by the Joint Collaborative Team on Video Coding (JCT-VC) in February 2015, is the new industry standard for 3D applications. The 3D-HEVC provides plenty of advanced coding tools specifically for addressing the coding of auto-stereoscopic videos which have the format of multiple texture views along with the depth maps which are responsible for synthesising intermediate views with sufficient quality for auto-stereoscopic display. The provided tools take advantage of the statistical redundancies amongst texture views and depth maps in the video sequences, as well as the unique characteristics of depth maps to significantly shrink the bit-rate while preserving the objective visual quality of the 3D videos. However, those tools with high capability in terms of compression come with the high complexity of computation which has made the encoding time of the 3D video sequences much longer than ever by traversing a lot more candidates, calculating time-consuming RD Cost for each of them, especially in the wedgelet searching process for depth maps. While this full-search style method can promise to find the best candidate in depth intra mode decision, the time cost is expensive.

In this dissertation we address the time cost by presenting a new intra mode decision method for depth maps, leveraging the deep convolutional neural networks to predict the wedgelet angles for the depth blocks. The predictions from the learned models are capable of reducing the number of wedgelet candidates by half as well as the angular modes in depth map coding. The size of the neural network has been carefully designed to balance the trade-off between the time cost of model prediction and the model prediction accuracy. Confusion matrix is used to monitor the training process. Top-K criteria is employed for the prediction. We have integrated the learned models into the reference software of 3D-HEVC for the experiments. The compiled executable binaries are able to harness the power of the simultaneous computation of CPU, as well as the parrallel computation of GPU to accelerate the predictions. The simulation results show that the proposed algorithm provides 64.6% time reduction in average while the BD performance has a tiny decrease comparing with the state-of-the-art 3D-HEVC standard.

Acknowledgments

Allow me first to give sincere thanks to my supervisor, Dr.Yui-Lam Chan, for his extremely generous support, most insightful advices and innumerable yet constructive feedback. I learned from him to first identify a problem, by reading a vast amount of articles to know what people have achieved and what bottlenecks they have encountered. I learned how to read papers, how to organize them to become the inner comprehension. He guided me to use the machine learning approach to solve the problem that has been found in the first stage. Without his guidance I will not have the idea to learn the deep learning technology and apply it to optimize the video coding. His encyclopedic knowledge and charming personalities made him my mentor in both research and life. I wish to thank Dr.Sik-Ho Tsang, for our in-depth discussions from which I can always find useful clues to proceed to next step. His great expertise in video coding significantly benefits me during my intensive period of learning. Also I would like to thank my friends Alex and Jacky, for our extensive discussions about artificial intelligence and their applications. Finally thank you my parents, for the great love and constant encouragement which give me confidence to face and handle all the challenges at every moment.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

asfvgjasdlKJFHASKLDFHASLKDFHJAKSDJPFHASKLJD.FHASDF ASDFASDFASDFASDFADS-FASDF

1.1 Motivation

fasdfasdfasdfasdfasdfasdf If you are

1.2 Welcome and Thanku

Pi expression	Value
π	3.1416
π^π	36.46
$(\pi^\pi)^\pi$	80662.7

Table 1.1: The effects of treatments X and Y on the four groups studied.

Figure 1.1: An electron (artist's impression).

Chapter 2

Background

With the rising popularity of the high definition videos, the new standard termed High Efficiency Video Coding (HEVC) for compressing videos in a more efficient way comparing with previous standards, such as H.264/AVC, has emerged under the efforts from the Joint Collaborative Team on Video Coding (JCT-VC). In the meanwhile, five extensions of the HEVC standard, comprising Format Range Extension (RExt), Scalability Extension (SHVC), Multi-view Extension (MV-HEVC), 3D Extension (3D-HEVC), Screen Content Coding Extension (SCC), have been finalized from 2014 to 2016 to support fulfill extra requirements in various scenarios.

2.1 Video Coding

... ...