附件 1

國立清華大學培育優秀博士生獎學金申請表

申請流程：申請人填妥資料→送指導教授審核→系所學位學程審核→學院審核→學院送綜合教務組提送委員會審議

<table>
<tr><td>姓名</td><td colspan="3">洪澤厚</td><td>學號</td><td>108065534</td></tr>
<tr><td>系所</td><td colspan="3">資訊系統與應用研究所</td><td>指導教授</td><td>徐正炘</td></tr>
<tr><td>研究領域</td><td colspan="5">Multimedia Networking, Augmented Reality and Virtual Reality.</td></tr>
<tr><td rowspan="3">學歷</td><td>學位</td><td colspan="3">校名(全稱)</td><td>科系(全稱)</td></tr>
<tr><td>碩士</td><td colspan="3">國立清華大學（逕博）</td><td>資訊系統與應用研究所</td></tr>
<tr><td>學士</td><td colspan="3">中山醫學大學</td><td>醫療產業科技管理學系</td></tr>
<tr><td>畢業(歷年)
成績</td><td colspan="2">總學分數</td><td colspan="2">成績(GPA)</td><td>系、班排名(%)</td></tr>
<tr><td>學士班</td><td colspan="2">130</td><td colspan="2">總平均：72.76</td><td>84.62</td></tr>
<tr><td>碩士班
(畢業)</td><td colspan="2"></td><td colspan="2"></td><td></td></tr>
<tr><td>碩士班
(逕讀)</td><td colspan="2">第 1 學期 10 學分
第 2 學期 9 學分
第 3 學期____學分
…</td><td colspan="2">第 1 學期 GPA：4.06
第 2 學期 GPA：4
第 3 學期 GPA：____
…</td><td></td></tr>
<tr><td>特殊表現</td><td colspan="5">1. ACM Multimedia Conference 2020: We submitted our paper to the conference, which is under review now.
2. Participate in two research project: a MOST project on multimedia and an UMC joint project on machine learning</td></tr>
<tr><td>應備文件</td><td colspan="5">☑本獎學金申請表
☑歷年成績單
☑指導教授推薦函
☑個人學習計畫書（包含個人分年度學習規劃，最多 3 頁）
☑其他有利審查文件</td></tr>
<tr><td>簽名欄</td><td colspan="5">本人同意配合本校執行科技部補助博士卓越提升試辦方案（參考資料詳培育機制表）
☑ 同意　☐不同意
申請人：洪澤厚　　109 年 7月 20 日
==
指導教授：（尚未確定指導教授者請送系、所、學位學程主管審核）
單位主管：
學院院長：</td></tr>
</table>

博士卓越提升培育機制表

培育面向	執行方式	輔導機制
專業研究素養提升	國際學者專家邀請開授課程	邀請國際專家學者(含本校配合國家重大政策發展方向之核心領域：綠色能源、巨量資料分析與服務創新、生物科技與醫療照護以及文創產業，延聘具國際化且有產業化潛力的菁英師資)每學期開設2門專業課程，學生可依個人學術專長志願參與修課。
	論文寫作與發表技巧培訓參與	本計畫以4年期為一基本規劃，期間**必須**參與清華學院語文中心規劃之論文寫作、發表技巧培訓等相關課程或工作坊至少3門/場次，並獲該中心認證。
多元知能建構培訓	精進教學與表達技能	本計畫以4年期為一基本規劃，期間**必須**參與教學發展中心規劃之教學原理、口語表達技巧、英語授課技巧工作坊等課程，至少3場次，並取得研習證書。
	團隊領導、組織管理、溝通訓練	1. 期間與次數：四年期為一完整規劃，每學期**必須**參與約2~3次的隔月團隊課程。 2. 團隊組成：以取得獎學金之博士生為跨領域的社群團隊，推薦數人為團隊幹部並建立團隊共識。 3. 輔導機制： (1) 各領域之資深業師課程，並進行個人的課程互動追蹤與日後該領域之就職場域的諮詢。 (2) 職涯發展師的定期輔導與個人諮商協助，尤其在第一年對自己能力盤點與博士學涯定錨之後，及第四年的職涯選擇時，更需要提供及時的協助與輔導。 4. 對於可能新加入的成員： 依其個別需求課程提出後，經由評估後提供相關加強課程，以便於盡快融入團隊之中。 5. 管理機制： (1) 出席率80% 以上：出席時數/課程或活動時數。 (2) 課程中之參與度，委請由業師於每一階段進行評量。 (3) 請假制度：若非個人事由，則由指導教授提出其公務證明，公務請假總次數不可高於 1/4。 (4) 每學期之出席與參與狀況，交由審議委員參考評估。
研發轉譯創新創業	工程轉譯團隊研發實作參與	依個人學術領域專長志願參與(GLORIA 協助鏈結)。
	創新創業生態圈資源鏈結	依個人學術領域專長與趣志願參與。
人才交流諮詢平台	知識匯產學研地圖脈絡掌握	依個人學術領域專長與趣按圖索驥。
	企業深度參訪與業師諮詢	本計畫以4年期為一基本規劃，**必須**於4年期間至少完成1屆綜合學務組規劃之「企業領航人才培育計畫」，並依該組訂定之修課規定辦理。
國際網絡拓展參與	國際學術網絡建立	本計畫以4年期為一基本規劃，**必須**於修讀期間至少參與1次國際研討會並以第一作者(或通訊作者)發表論文。
	國際產學網絡鏈結	依個人學術領域專長志願參與。

Tse-Hou Hung

Contact Information

Address Institute of Information Systems and Applications, National Tsing Hua University,
No. 101, Section 2, Kuang-Fu Road, Hsin Chu, Taiwan

Mobile +886-919-390136

Mail tsehou.nthu@gmail.com

Educations

2019 – Present **National Tsing Hua University (NTHU), Taiwan**, *Master program in Information Systems and Applications*
Thesis Topic: 6-DoF Immersive Video streaming to Head-Mounted Display
Advisor: Prof. Cheng-Hsin Hsu
Accumulated GPA: 4.03/4.3

2015 – 2019 **Chung Shan Medical University (CSMU), Taiwan**, *Bachelor degree in Health Policy and Management*

Research Interests

Multimedia Networking, Augmented Reality, and Virtual Reality

Publication

T. Hung, C. Hsu and C. Hsu Optimizing Immersive Video Streaming Using Deep Learning Approaches: A Case Study on TMIV Submitted to the ACM International Conference on Multimedia (MM'20), Seattle, United States, October 2020.

Working Experience

September 2019 – Present Research Assistant, Networking and Multimedia System Lab, Department of Computer Science, NTHU

March 2020 – Present Assistant System Administrator, Computer and Communication Center, NTHU

Research Experience

6-DoF Immersive Video Streaming, *(Supported by the MOST Project: Teleporting Through Space Across Time Using Head-Mounted Displays: A Case Study for Real Estate)*
Virtual Reality (VR) has become increasingly more popular in various business sectors. The modern VR systems that support six-degree-of-freedom (6-DoF) can provide more immersive experience, in which Head-Mounted-Display (HMD) user's viewport can be changed according to his/her position and orientation. However, because of the tremendous content size, 6-DoF immersive video streaming dictates too much bandwidth and computing resources. In this work, we propose a configuration optimizer that uses Reinforcement Learning (RL) and Convolutional Neural Network (CNN) to select the best configuration setting. Trough real experiments, we show that our solution reduces the bandwidth and computing resource consumption while delivering good video quality.

Machine Learning Platform, *(Supported by the UMC Project: Development for AI Related Edge and Infrastructure)*

Machine Learning (ML) has been around for decades and is now commonly used in many fields. In recent years, more and more people and companies try to use ML techniques to achieve or improve their productibility. However, capitalizing the potential of ML needs a lot of domain knowledge, along with tons of tuning for the best performance. Furthermore, ML applications are not done after a model is trained. This is because the trained models may become outdated in the future, due to the drifts of concepts. Therefore, after deploying an ML model, we still need to monitor its performance and retrain whenever necessary. To allow the ML developers to focus on their work, we need an ML platform that can automate the routine tasks. In this project, we build such an ML platform, which consists of various tools to speed up data preparation, model building, service serving and monitoring of multiple ML applications. We survey the existing platforms and generalize their components and functions. This leads to a general ML platform design that can be adopted in diverse scenarios. To demonstrate the practicality and efficiency of our design, we build a real testbed based on several open-source projects like Kubeflow. We use the testbed to conduct a case study, which results in a few new research problems, which were not solved in the literature. We are currently solving these problems jointly with UMC colleagues.

Research Proposal: 6-DoF Immersive Video Streaming to Head-Mounted Displays

Tse-Hou Hung
tsehou.nthu@gmail.com
National Tsing Hua University
Hsinchu, Taiwan

ABSTRACT

Six Degree-of-Freedom (6DoF) Immersive video streaming is very challenging due to various data representations, the strict network condition requirements and user experience. In my PhD study, we focus on three research problems to overcome the aforementioned challenges. First, we study different data representations to understand their pros, cons, and suitable usage scenarios. Second, we optimize the streaming system to reduce the bandwidth requirement of immersive video streaming. Last, we design and conduct a series of user studies to explore user experience. This research proposal shows the research plan and expected research outcome.

1 INTRODUCTION

Virtual Reality (VR) technology is thriving in various business sectors, including computer games, tourism industry, real estates, and occupational trainings. The VR scenes can either be generated with computer graphics or captured from nature scenes, which provide omnidirectional, a.k.a. 360°, viewing experience of virtual worlds. It is estimated that the global VR market size will reach 26.89 billion USD by 2022, with an annual growth rate of 54% from 2017 to 2022 [10]. As VR technology becomes more and more mature, researchers start to study its Quality of Experience (QoE) in order to deliver higher-quality scenes for better user experience.

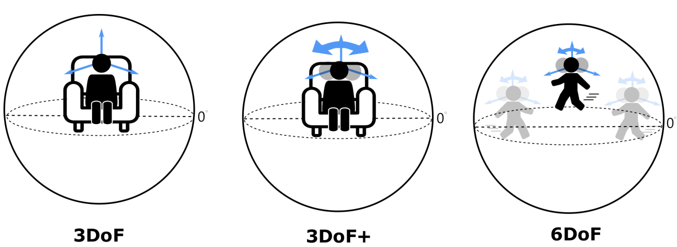

Figure 1: Three phases of VR technology development. The spheres represent the virtual worlds.

One key factor affecting the QoE of VR technology is the way users interact with the virtual worlds. MPEG-I group has defined three development phases of the VR technology: 3 Degree-of-Freedom (3DoF), 3DoF+, and 6DoF, which are illustrated in Fig. 1. 3DoF VR technology allows users to rotate their heads in three dimensions, which are represented by yaw ψ, pitch θ, and roll ϕ. However, with 3DoF VR, when users move their heads or stand up and walk, their position changes do *not* affect the views rendered in the Head-Mounted Displays (HMDs). To overcome this limitation, 3DoF+ VR is proposed to support *limited* movements. Namely, users can slightly move their heads to a certain extent when sitting on fixed chairs. 6DoF VR further enables users to freely move, e.g., walk, in virtual worlds. Users can not only rotate their heads in three dimensions but also move in the three additional dimensions, which are x, y, z coordinates.

A naive way to achieve 3DoF+ and 6DoF is to capture multiple 360° videos at different positions [2, 9], and only allow the HMD users to switch among these discrete positions in the scenes. We target more general 3DoF+ and 6DoF VR, where users can freely move in among the discrete positions. We refer to the streaming systems that support such 3DoF+ and 6DoF VR as *immersive video streaming* throughout this proposal.

Achieving real immersive video streaming is not an easy task, because there are many challenges need to be overcome:

- Different Data Representations: Various data types can be used to store the information of immersive video, e.g., 2D texture and depth image, light field, 3D point cloud, 3D mesh. Each of them has pros, cons, and usage scenarios, which hav not been comprehensively studied.
- Strict Network Condition Requirements: The data representations of immersive video have a huge data size. They need a lot of bandwidth to be streamed, while the realtime ness needs to be maintained. Take light field video as an example, streaming it consumes bandwidth in the range between 200 Gb/s and 1 Tb/s [1], which is much higher than the available bandwidth we consume today. Moreover, the interactive applications, e.g., 6DoF VR games and holographic conferences, need low latency to realize real-time communications. This inturn makes the network condition requirements more strict.
- User Experiences: User experience is the quality that perceived by users. It is an important performance metric for multimedia applications. However, the user experience of immersive videos is still not explored, and the challenges of measuring user experience include but not limited: (i) too many parameters may affect the user experience in immersive video streaming. and (ii) it's hard to build a immersive video streaming testbed.

In this research proposal, we propose three research directions to address the three challenges: (i) *Data Representations* (ii) *Optimal Streaming* (iii) *Quality of Experience*. We introduce them in details below.

2 RESEARCH PROBLEMS

2.1 Data Representations

Whiles data types mentioned in Sec. 1 could be suitable to immersive video streaming, Their pros and cons are unknown. In our research group, we have studied 3DoF VR streaming in the past few years [4, 5]. These works help us to understand the existing streaming techniques in 3DoF 360° video streaming. We are currently expand our research to different 6DoF data types to understand their pros and cons. We will also concretize their suitable usage scenarios, The final outcome of this work is several immersive video streaming testbeds that support heterogeneous data representations. Through real experiments, we hope to know more systems challenges in these emerging applications.

2.2 Optimal Streaming

To reduce the bandwidth requirement of immersive video streaming, the streaming systems must be optimized. In our recent work [7], we use Test Model for Immersive Video (TMIV) [11–13], which is a Depth Image Based Rendering (DIBR) codec for immersive video streaming from MPEG, to optimize the immersive video streaming. We develop algorithms to solve the configuration optimization problem based on deep learning, particularly the Neural Network (NN) approaches. The two proposed algorithms are: a Convolutional Neural Network (CNN) based algorithm and a Deep Reinforcement Learning (DRL) based algorithm [1]. Our CNN algorithm benefits from: (i) automatic extracting latent features and (ii) inferencing the prediction rapidly and directly according to the input, which allows the configuration optimizer to adapt to various video scenes and dynamic camera parameters. Our DRL algorithm systematically builds an *agent* that can adapt to dynamic *environments* [3, 6, 9, 14]. The trained agent learns how to quickly search through a large space for optimal configurations. The two proposed algorithms are trained to adaptively find the optimal configurations given the diverse video content, HMD user behaviors, and user-specified utility functions. We submitted our paper to ACM Multimedia Conference 2020, which is under review.

We are currently extending this work into a journal submission. In particularly, we plan to (i) design, implement, and evaluate an end-to-end DIBR-based immersive video streaming system, (ii) apply the state-of-the-art DRL algorithms and (iii) adopt larger datasets to train our optimizers. We expect the performance of our systems can be better after we apply these optimizations. Moreover, according to the research outcomes of Sec. 2.1, we will generalize our solutions to other data types.

2.3 Quality of Experience (QoE)

To measure QoE of users, we plan to design and conduct a series of user studies to quantify the relationship between each parameters and user experience in immersive video streaming sessions. In our previous work [7], we conduct a small-scale user study to evaluate our solution. The testbed used in that work can be extended and used in more comprehensive subjective experiments. The high-level architecture of the immersive video streaming system is shown in Fig. 2. Moreover, we plan to measure the Just-Noticeable Difference (JND) [8] bitrate of immersive video streaming. Using JND bitrates, we can intelligently save the network and computing resources without affecting the user experience. Combining the outcomes of the QoE study with the ones achieved in Sec. 2.1 and Sec. 2.2 gives a fully-optimized immersive 6DoF video streaming system.

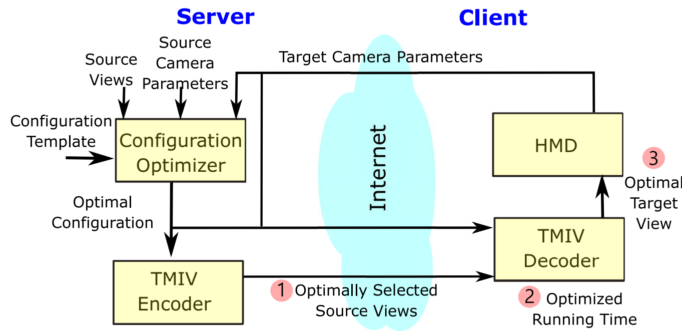

Figure 2: High-level architecture of immersive video streaming systems for our QoE study.

3 RESEARCH PLAN

Fig. 3 gives the Gantt chart of my research plan. The first problem is the data representation study. The outcome of this research problem is our better understanding on the immersive video streaming systems, which is also useful to the research community. Concurrently, We extend our recent work [7] to turn it into a journal paper. After that, we work on the QoE study. We plan to build an immersive video streaming testbed and conduct a series of user studies to understand the user experience. We will also continue optimizing the immersive video streaming testbed by applying different optimization algorithms developed by us. The results of QoE study can also help us better optimize the testbed. Last, we will design, build, and evaluate a real-time immersive video streaming system based on the results of our research to turn it into a journal paper.

4 EXPECTED IMPACTS

Currently, the techniques of immersive video streaming are still in their infancy. Most of the challenges identified in this proposal have not been rigorously studied. Existing VR/AR applications do not

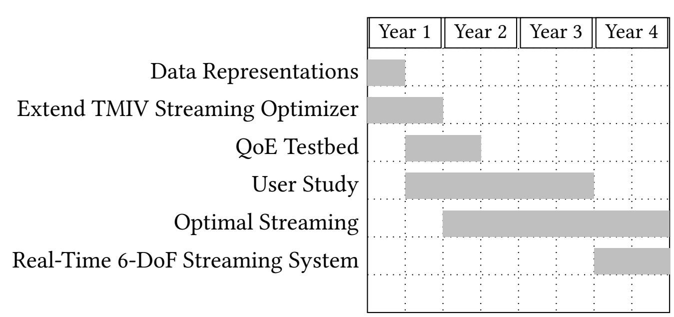

Figure 3: Gantt Chart

[1]We refer to them as CNN and DRL algorithms for short in the rest of this paper.

support real-time streaming; Even if they do, only 3DoF interactions are supported, which is not the *real* immersive experience.

In my PhD study, we will propose a series of solutions to overcome the challenges of immersive video streaming. As system researchers, these solutions through actual experiments. In data representation research, we will provide a comprehensive review, which can help the research community gain more knowledge on the features, challenges and opportunities of various 6DoF data types. Through user study and data analysis, we will find new discoveries on the user experience of immersive video streaming in QoE research. With these discoveries, the content creators can better meet users' needs when they create the immersive content. The researchers can also use the users' viewing behavior to develop more efficient streaming solutions. In optimal streaming, we will propose innovative algorithms to overcome various issues caused by scarce resources. The bandwidth and latency requirements of immersive video streaming will, therefore, fulfilled under diverse network conditions, and several 6DoF applications can be realized.

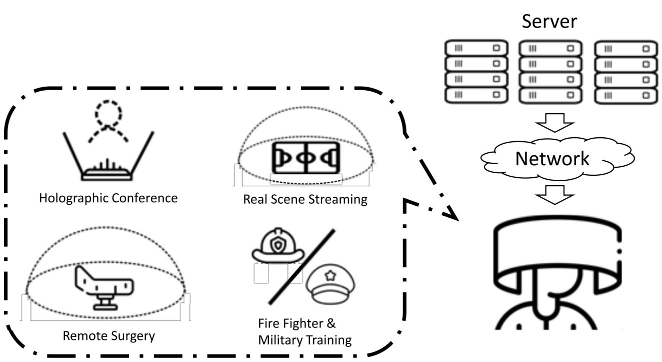

Figure 4: Usage scenario of immersive video streaming.

The techniques of immersive video streaming can make people's life more convenient and productive. As shown in Fig. 4, immersive video streaming can be used in various usage scenarios.

- **Holographic conferences** may become reality. Users can see the projection of remote people and communicate with one another naturally. The holographic conferences will provide exactly the same experience as face-by-face.
- **Real scene streaming** may provide immersive live streaming for sports, speech, and other events. For example, users can watch live sports without blind angles to enjoy the games.
- **Remote Surgery** can also be performed. Immersive video streaming can provide the situations of remote patients to the doctor, who can carry out the surgery remotely. This will improve the healthcare quality in rural areas.
- **Fire fighter and Military training** can be clone in a safer and more efficient way. The training sessions can simulate real situations to help learners experience danger environments without risks.

REFERENCES

[1] A. Clemm, M. Vega, H. Ravuri, T. Wauters, and F. Turck. 2020. Toward Truly Immersive Holographic-Type Communication: Challenges and Solutions. *IEEE Communications Magazine* 58, 1 (2020), 93–99.

[2] X. Corbillon, F. Simone, G. Simon, and P. Frossard. 2018. Dynamic Adaptive Streaming for Multi-viewpoint Omnidirectional Videos. In *Proc. of ACM Multimedia Systems Conference (MMSys'18)*. 237–249.

[3] L. Costero, A. Iranfar, M. Zapater, F. Igual, K. Olcoz, and D. Atienza. 2019. MAMUT: Multi-Agent Reinforcement Learning for Efficient Real-Time Multi-User Video Transcoding. In *Proc. of IEEE Design, Automation Test in Europe Conference Exhibition (DATE'19)*. 558–563.

[4] C. Fan, W. Lo, Y. Pai, and C. Hsu. 2019. A Survey on 360° Video Streaming: Acquisition, Transmission, and Display. *Comput. Surveys* 52, 4 (2019).

[5] C. Fan, S. Yen, C. Huang, and C. Hsu. 2020. Optimizing Fixation Prediction Using Recurrent Neural Networks for 360° Video Streaming in Head-Mounted Virtual Reality. *IEEE Transactions on Multimedia* 22, 3 (2020), 744–759.

[6] T. Huang, R. Zhang, C. Zhou, and L. Sun. 2018. QARC: Video Quality Aware Rate Control for Real-Time Video Streaming Based on Deep Reinforcement Learning. In *Proc. of ACM International Conference on Multimedia (MM'18)*. 1208–1216.

[7] T. Hung, C. Hsu, and C. Hsu. 2020. *Technical Report: Optimizing Immersive Video Streaming Using Deep Learning Approaches: A Case Study on TMIV.* Technical Report. https://reurl.cc/z8YVbe.

[8] Y. Jia, W. Lin, and A. A. Kassim. 2006. Estimating Just-Noticeable Distortion for Video. *IEEE Transactions on Circuits and Systems for Video Technology* 16, 7 (2006), 820–829.

[9] H. Pang, C. Zhang, F. Wang, J. Liu, and L. Sun. 2019. Towards Low Latency Multi-viewpoint 360° Interactive Video: A Multimodal Deep Reinforcement Learning Approach. In *Proc. of IEEE Conference on Computer Communications (INFOCOM'19)*. 991–999.

[10] ZION Market Research. 2018. Virtual Reality (VR) Market by Hardware and Software for (Consumer, Commercial, Enterprise, Medical, Aerospace and Defense, Automotive, Energy and Others): Global Industry Perspective, Comprehensive Analysis and Forecast, 2016–2022. Retrieved April 21, 2020 from https://www.zionmarketresearch.com/report/virtual-reality-market

[11] B. Salahieh, B. Kroon, J. Jung, and M. Domański. 2019. Test Model 2 for Immersive Video. International Organization for Standardization Meeting Document ISO/IEC JTC1/SC29/WG11 MPEG/N18577.

[12] B. Salahieh, B. Kroon, J. Jung, and M. Domański. 2019. Test Model 3 for Immersive Video. International Organization for Standardization Meeting Document ISO/IEC JTC1/SC29/WG11 MPEG/N18795.

[13] B. Salahieh, B. Kroon, J. Jung, and M. Domański. 2019. Test Model for Immersive Video. International Organization for Standardization Meeting Document ISO/IEC JTC1/SC29/WG11 MPEG/N18470.

[14] R. Sutton and A. Barto. 2018. *Reinforcement Learning: An Introduction* (2 ed.). A Bradford Book.

中山醫學大學學生歷年成績表

學系：健康管理學院醫療產業科技管理學系　學號：0448004　姓名：洪澤厚　修業總學分： 130

第一學年(104 年 9 月到 105 年 6 月)

科目	必選修	第一學期 學分	第一學期 成績	第二學期 學分	第二學期 成績
體育(一)(體育領域)	必	0	84		
套裝軟體－MS EXCEL	必	2	92		
經濟學(一)	必	2	90		
管理學	必	3	75		
微積分	必	2	82		
會計學	必	2	81		
英文(一)(大學英文)	必	2	74		
生理學	必			2	69
體育(二)(體育領域)	必			0	76
醫務管理概論	必			2	81
微積分	必			2	63
組織行為	必			3	69
會計學	必			2	72
大學之道(大學之道)	必			1	90
英文(二)(大學英文)	必			2	72
閱讀與書寫(大學國文)	必			2	75
大學之道(大學之道)	選	1	90		
線性代數	選	3	63		
毒物與生活(科技應用領域)	選	2	67		
全民國防教育軍事訓練-國防政策(國防通識領域)	選	1	80		
醫護與社會(倫理哲學領域)	選	2	87		
生活與意外(人文關懷領域)	選			2	50
物件導向程式設計(一)	選			3	63
經濟學(二)	選			2	63
健康管理概論	選			2	80
學業總分			1730		1736
學業平均成績			78.64		69.44
修習學分數		22		25	
實得學分數		22		23	
累計學分數		22		45	
軍訓成績					
體育成績			84		76
服務教育成績					
操行成績			88		85

第二學年(105 年 9 月到 106 年 6 月)

科目	必選修	第一學期 學分	第一學期 成績	第二學期 學分	第二學期 成績
體育(三)(體育領域)	必	0	48		
衛生行政	必	2	69		
醫療經濟學	必	2	34		
病歷與疾病管理	必	2	83		
電腦程式設計	必	2	81		
統計學	必	2	56		
人力資源管理	必	2	86		
體育(四)(體育領域)	必			0	77
臨床醫學概論(科技應用領域)	必			2	64
衛生醫療制度	必			3	83
醫事與衛生法規(倫理哲學領域)	必			2	93
醫療產業政策分析	必			3	84
統計學	必			2	71
保健與人生(人文關懷領域)	選	2	71		
普通生物學	選	2	23		
醫學英文	選	2	3		
普通化學	選	2	28		
影視史學(創意美學領域)	選			2	85
醫藥與保健(科技應用領域)	選			2	75
普通心理學(科技應用領域)	選			2	62
學業總分			1068		1401
學業平均成績			53.4		77.83
修習學分數		20		18	
實得學分數		10		18	
累計學分數		55		73	
軍訓成績					
體育成績			48		77
服務教育成績					
操行成績			86		86

第三學年(106 年 9 月到 107 年 6 月)

科目	必選修	第一學期 學分	第一學期 成績	第二學期 學分	第二學期 成績
社區照護管理	必	2	78		
醫療管理會計	必	2	66		
健康保險	必	2	86		
管理數學	必	2	84		
流行病學	必	2	49		
體育(三)(體育領域)	必	0	81		
醫療經濟學	必	2	90		
統計學	必	2	82		
醫院管理實務	必			1	86
作業管理	必			2	60
醫療行銷管理	必			2	82
醫療品質管理	必			2	69
資料庫管理	必			2	84
管理數學	必			2	80
研究方法與研究設計	必			2	73
大數據分析	選	2	80		
消費者行為	選	2	94		
閱讀與書店(語文溝通領域)	選	2	60		
大腦與人類行為(邏輯推理領域)	選	2	74		
生物資訊學	選	2	66		
研究方法與生醫資料分析	選			3	80
生物統計實務	選			2	50
進階程式語言(科技應用領域)	選			2	95
學業總分			1818		1512
學業平均成績			75.75		75.6
修習學分數		24		20	
實得學分數		22		18	
累計學分數		95		113	
軍訓成績					
體育成績					
服務教育成績					
操行成績			88		86

第四學年(107 年 9 月到 108 年 6 月)

科目	必選修	第一學期 學分	第一學期 成績	第二學期 學分	第二學期 成績
醫務策略管理	必	2	62		
論文研究	必	1	90		
科技管理	必	2	74		
醫務管理實習	必	2	85		
流行病學	必	2	75		
資材及供應鏈管理	必			2	
醫療資訊學	必			2	
論文研究	必			1	
西方文化與創意產業英文(語文溝通領域)	選	2	86		
資料分析	選	2	75		
管理資訊系統	選	2	88		
電子病歷實務	選	2	89		
R 語言	選			2	
健康資訊學	選			2	
流行病學實例討論	選			2	
醫學倫理(倫理哲學領域)	選			2	
學業總分			1358		0
學業平均成績			79.88		0
修習學分數		17		13	
實得學分數		17		0	
累計學分數		130		130	
軍訓成績					0
體育成績					0
服務教育成績					0
操行成績			86		

中山醫學大學
成績單專用章
教務處

備註:#抵免 *通識 @暑修 &免修 P 檢定通過 F 檢定未通過 ##校際選課　最低畢業學分數 128 學分(必修學分數 81 學分)　全系人數 39 人,該生 1 至 4 年級學業成績總平均 72.76,名次為第 33 名,名次佔全系 84.62%

中山醫能力百分百：P　(117)

國立清華大學
NATIONAL TSING HUA UNIVERSITY

國立清華大學學生歷年成績表

學號	108065534	姓名	洪澤厚		
學院	電機資訊學院	系所組別	資訊系統與應用研究所碩士班	等級制累計平均成績	4.06
入學年月	108年9月	學制	碩士班	畢業年月	---

科號	科目名稱	學分	成績
	108學年度第1學期(108年9月至109年1月)		
ISA 565100	資訊應用書報討論	1	#A-
	以下選修		
CS 135501	計算機程式設計一	3	#A
CS 522300	高等計算機網路	3	#A
CS 526200	多媒體網路與系統	3	#A+

等級制學期學業平均成績：4.06
對應百分制學期學業平均成績：91.20
實得學分數：10　累計學分數：10　操行成績：A

科號	科目名稱	學分	成績
	108學年度第2學期(109年2月至109年6月)		
ISA 565100	資訊應用書報討論	1	#*
	以下選修		
CS 411100	平行計算概論	3	*
EECS204001	資料結構	3	#*
ISA 590200	資訊系統實作二	2	#*

等級制學期學業平均成績：N/A
對應百分制學期學業平均成績：N/A
實得學分數：N/A　累計學分數：N/A

備註：

說明：1.TR：抵免　W：停修　@：不及格　*：成績未到　#：英語授課 2.暑修成績不列入當學期平均惟列入畢 業平均成績 3.教育學程通識語文(科號5字頭以上課程不在此限)軍訓體育不計入學期及學業平均成績 4.^重複修讀不計入學期及學業平均成績。 5.本校研究所學生學業成績不排名。 6.自107學年度起入學學生之成績單，亦提供修課相對成績，等級制成績數字越大，表示表現越好；修課相對成績數字越小，表示表現越好。 7.研究生學業成績以B-為及格，操行成績以C-為及格。

Grade	A+	A	A-	B+	B	B-	C+	C	C-	D	E	X	P	N
Score	100~90	89~85	84~80	79~77	76~73	72~70	69~67	66~63	62~60	59~50	49~1	0~0	通過	不通過
Grade point	4.3	4	3.7	3.3	3	2.7	2.3	2	1.7	1	0	0	PASS	FAIL

印製日期：109年05月29日

國立清華大學教務處註冊組 REGISTRAR'S OFFICE National Tsing Hua University

View Reviews

Print

Paper ID	899
Paper Title	Optimizing Immersive Video Streaming Using Deep Learning Approaches: A Case Study on TMIV
Track Name	Main Track

Reviewer #1

Not Submitted

Reviewer #2

Questions

1. [Paper Summary] What is the paper about? Please, be concise (3 to 5 sentences).
This paper attempts to generate better TMIV configurations, and focuses on two parameters of TMIV configurations: the number of passes and the number of view per pass.

The paper proposes two deep-learning based algorithms: a CNN based one and a deep reinforcement one.

The paper then describes an experiment where these algorithms are compared to the default TMIV configuration and to the optimal configuration found by exhaustive search: there is no statistically significant difference in the quality of the output, however, the configurations of the proposed approach lead to fewer views, and thus, better performance in terms of network and execution time.

The paper finally describes a subjective experiment where users are asked to rate the outputs quality, which confirms that there is no significant difference in terms of perceived quality.

2. [Relevance] Is this paper relevant to an audience to ACM Multimedia? Please check https://2020.acmmm.org/call-for-paper.html.
Of limited interest to an audience

3. [Significance] Are the results significant?
Significant

4. [Novelty] Are the problems or approaches or applications/systems novel?
Novel

5. [Evaluation] Is the idea proposed in this paper well supported by theoretical analysis or experimental results?
Sufficient

6. [Paper Strengths] Please discuss. Justifying your comments with the appropriate level of details about the strengths of the paper (i.e. novelty, theoretical approach and/or technical correctness, adequate evaluation, clarity, applications, etc.). For instance, a theoretical paper may need no experiments, while a paper with a new approach or application may require comparisons to existing methods.
The paper clearly states a novel problem it tries to solve and proposes two possible solutions.

The paper compares the proposed solutions to the default and optimal solutions: it provides an in-depth analysis of the many metrics of the algorithm (number of required views, video quality, running time, utility function value, inference time) as well as a subjective evaluation where 23 subjects were recruited to compare the algorithms.

The key message of the paper is that the proposed methods give similar outputs, but require less resources and does a good job at demonstrating it.

The two methods have different behaviours and the paper gives recommendation about when to use which method depending on the context.

7. [Paper Weaknesses] Please discuss. Justifying your comments with the appropriate level of details about the weaknesses of the paper (i.e. lack of novelty – given references to prior work, lack of novelty, technical errors, or/and insufficient evaluation, etc.). Note: If you think there is an error in the paper, please explain why it is an error. It is not appropriate to ask for comparisons with unpublished papers and papers published after the ACM Multimedia deadline. In all cases, please be polite and constructive.
My main concern is that the inference time of the system is described very quickly in this paper. The paper states that the CNN takes 37ms and DRL takes 51ms on average to compute the configuration. However, it is not clear how many times those algorithms are run during a session, and while I agree that the paper demonstrates that the systems allow for network resources saving, it is unclear on how much computing resources are saved by running an algorithm that consumes computing resources to infer a configuration that will save computing resources later.

8. [Preliminary Rating] Please rate the paper according to one of the following choices.
Poster

9. [Rebuttal Requests] Please pose questions you want to be answered in the rebuttal. Please do NOT ask the author(s) to include any new results (e.g., experiments and theorems) in the rebuttal.
I would like to know how many times the algorithms are run during a typical streaming session, and if there is an estimation of how much computing resources are spent during this step.

10. [Confidence]
Not Confident

Reviewer #3

Questions

1. [Paper Summary] What is the paper about? Please, be concise (3 to 5 sentences).
The paper proposes two NN algorithms to help find the best source videos for view synthesis for a target viewpoint. The algorithms can be used in free-viewpoint video applications, or TMIV (Test Model for Immersive Video) as named by MPEG and this paper. The evaluation shows the proposed algorithms delivering similar or better view quality compared to the default algorithm defined in MPEG spec while using much less time (fast inference).

2. [Relevance] Is this paper relevant to an audience to ACM Multimedia? Please check https://2020.acmmm.org/call-for-paper.html.
Relevant to researchers in subareas only

3. [Significance] Are the results significant?
Moderately significant

4. [Novelty] Are the problems or approaches or applications/systems novel?
Novel

5. [Evaluation] Is the idea proposed in this paper well supported by theoretical analysis or experimental results?
Somewhat weak

6. [Paper Strengths] Please discuss. Justifying your comments with the appropriate level of details about the strengths of the paper (i.e. novelty, theoretical approach and/or technical correctness, adequate evaluation, clarity, applications, etc.). For instance, a theoretical paper may need no experiments, while a paper with a new approach or application may require comparisons to existing methods.
The paper tries to propose a better algorithm to select reference views. It is probably a niche problem to the current audience but will be an important one if free-viewpoint video or volumetric video becomes popular in the future. The proposed algorithms use the similar idea from video coding and applies well. The algorithms are evaluated with both objective and subjective experiments. A lot of effort but there are some major issues with the evaluation as I will point out next.

7. [Paper Weaknesses] Please discuss. Justifying your comments with the appropriate level of details about the weaknesses of the paper (i.e. lack of novelty – given references to prior work, lack of novelty, technical errors, or/and insufficient evaluation, etc.). Note: If you think there is an error in the paper, please explain why it is an error. It is not appropriate to ask for comparisons with unpublished papers and papers published after the ACM Multimedia deadline. In all cases, please be polite and constructive.
The first section needs a better explanation. Maybe adding a drawing to illustrate the relationship between source views and synthesized views.

I originally thought the paper is about view synthesis for systems like free-viewpoint video. But the experiments in the paper using 360 video to simulate different source views puzzles me. Using different view angles in a 360 video doesn't address the real problems in free-viewpoint video, like occlusion exposure issue, or insufficient sampling.

Another problem is also with evaluation. It looks to me that you are using all three sample videos in your training and then use the same video in testing. Does that mean your system needs to train the test content first before being used on it?

8. [Preliminary Rating] Please rate the paper according to one of the following choices.
Borderline Reject

9. [Rebuttal Requests] Please pose questions you want to be answered in the rebuttal. Please do NOT ask the author(s) to include any new results (e.g., experiments and theorems) in the rebuttal.
Your paper referred in many places to the technical report that you submit together. I find it hard to get a full picture of your work if I don't read the technical report. Is your technical report published anywhere? or how do you plan to resolve this if your paper is actually accepted?

Also, what is "View FoV" in table 1 of your technical report? Is that the FoV of the your simulated camera view? But the FoV in Fig 10 is just a normal FoV.

I am also not entirely sure what does 256x256x(7+7+21+21) mean in Fig 4. 7 for texture, 7 for depth, 21 for position and 21 for orientation? bytes?

10. [Confidence]
Confident

Go Back