Freie Universität Berlin
Fachbereich Mathematik und Informatik
Takustraße 9, 14195 Berlin

MASTER THESIS

USER POSITION PREDICTION IN 6-DOF MIXED REALITY APPLICATIONS USING ARTIFICIAL RECURRENT NEURAL NETWORK

VORHERSAGE DER BENUTZERPOSITION IN 6-DOF-MIXED-REALITY-ANWENDUNGEN UNTER VERWENDUNG EINES KÜNSTLICHEN REKURRENTEN NEURONALEN NETZWERKS

Oleksandra Baga

Freie Universität Berlin
Matrikelnummer 5480722
Master Computer Science
E-Mail: oleksandra.baga@gmail.com

Prof. Dr. First Supervisor
Fachbereich Mathematik und Informati
Freie Universität Berlin

Prof. Dr. Second Supervisor
Fachbereich Mathematik und Informati
Freie Universität Berlin

Statutory Declaration

I herewith formally declare that I have written the submitted master thesis independently. I did not use any outside support except for the quoted literature and other sources mentioned in the paper.

I clearly marked and separately listed all of the literature and all of the other sources which I employed when producing this academic work, either literally or in content.

I am aware that the violation of this regulation will lead to failure of the thesis.

29.06.2022.. Oleksandra Baga

Acknowledgments

This thesis was created in cooperation with the Fraunhofer Heinrich Hertz Institute.

First and foremost I would like to thank Prof FU Berlin, who supervised my thesis. Thank you very much for the helpful suggestions and constructive criticism.

A special thanks goes to the researcher Fraunhofer Heinrich Hertz Institute, Serhan Gül, who suggested an exciting topic for a research, which I was allowed to choose for my master thesis. I would like to express my sincere thanks for the commitment and the consultation during the preparation of this thesis.

Contents

List of Figures

Listings

List of Abbreviations

AR	Augmented Reality
CNN	Convolutional Neural Network
CPU	Central processing unit
DoF	Degree of freedome
DL	Deep Learning
KF	Kalman Filter
GRU	Gated Recurrent Unit
HMD	Head-Mounted-Display
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
LAT	Look ahead time
LSTM	Long-Short-Term Memory
M2P	Motion-to-Photon
MAE	Mean Absolut Error
MEC	Mobile Edge Computing
ML	Machine Learning
MR	Mixed Reality
NLP	Natural Language Processing
ReLu	Rectified Linear Unit
RNN	Recurrent Neural Network
SDG	Stochastic Gradient Descent
VR	Virtual Reality
3-DoF	Three degree of freedome
6-DoF	Six degree of freedome

1 Introduction

This thesis is focusing on designing and evaluation of the new approach for the predicting human head position within a 6-dimensional degree of freedom (6-DoF) of Extended Reality (XR) applications. At the beginning of the work the existing 3-DoF [1, 6] as well as 6-DoF [8] methods were analyzed, their similarities and differences [7] were taken into account when a model was designed. The presentation of a new developed approach, a neural network architecture and the way in which the data was collected and preprocessed is the main goal of this thesis. Proposed approach was evaluated and at the end of the work the obtained results were discussed and the suggestions for future work were done.

The correct and fast head movement prediction is a key to provide a smooth and comfortable user experience in VR environment during head-mounted display (HDM) usage. The recent improvements in computer graphics, connectivity and the computational power of mobile devices simplified the progress in Virtual Reality (VR) technology. The way users can interact with their devices changed dramatically. With new technologies of VR environment user becomes the main driving force in deciding which portion of media content is being displayed to them at any time of interaction with VR Applications [7]. Until recently the high-quality experiences with modern Augmented Reality (AR) and VR systems were not widely presented in home usage and were mainly used in research labs or commercial setups. The hardware for displaying the VR environment was once extremely expensive but recent years became more broadly accessible and the 6-DoF VR headset designed for the end-user were released[1]. It is possible now to experience virtual reality scenes and watch new type of volumetric media at home and the market interest for development VR and AR applications expected to be huge next years.

Although all mentioned above improvements, rendering of volumetric content remains very demanding task for existing devices. It is possible to decrease the computational load on the client device by sending the rendered 2D content instead of volumetric data [4]. The 2D view must correspond the current position and orientation of a user wearing HMD. The computational task can be offloaded to a server infrastructure [4]. Due to the added in this approach network latency and processing delays the rendered 2D image can appear even later on the display than with usage of local rendering system. The reducing the Motion-to-Photon (M2P) latency by prediction the future user position and orientation for a look-ahead time

[1]https://medium.com/@DAQRI/motion-to-photon-latency-in-mobile-ar-and-vr-99f82c480926

(LAT) at remote server and sending the corresponding rendered view to a client could be very effective for 6-DoF XR application with immersive media. Thus, the development and evaluation of approach for the position and orientation prediction for a look-ahead time (LAT) is the main goal ot this thesis.

1.1 Problem statement

The existing on this moment virtual environments can be divided into two main groups. Depending on position of the user and their ability to move inside the VR environment all the 3-DoF and 6-DoF.

1.2 Motivation for the research

The following chapter introduces the preprocessing pipeline, architecture, and evaluation process used to develop ahead motion prediction algorithm.

1.3 Structure of the thesis

The organization of this thesis is as follows. The literature review chapter introduces the concepts of XR technologies and principles of motion prediction algorithms. It follows an overview of previous research.

Chapter 1 - Introduction.
The following chapter introduces the preprocessing pipeline, architecture, and evaluation process used to develop ahead motion prediction algorithm.

Background

2

The length and complexity of your theoretical framework depends on your field and topic. Some studies have an obvious basis in a well-established theory, while others require more detailed explanation and justification.

2.1 Mixed reality

This chapter introduces [6, 2] the background of the thesis to investigate head motion prediction. First, the concept of extended reality (XR) related to the research problem is presented, followed by introducing different headmotion prediction approaches.

2.2 Six degrees of freedom

This chapter introduces the background of the thesis to investigate head motion prediction. First, the concept of extended reality (XR) related tothe research problem is presented, followed by introducing different head motion prediction approaches.

2.3 Motion-to-photon latency

Motion-to-photon latency describes the time gap between a user's physical motion and the resulting displayed content of an HM. Extended Reality enables 3D content to be experienced with six degrees of freedom with head-mounted displays (HMD).

2.4 Cloud-based volumetric video streaming

This chapter introduces the background of the thesis to investigate head motion prediction. First, the concept of extended reality (XR) related to the research problem is presented, followed by introducing different head motion prediction approaches.

2.5 Head motion prediction

In the last decade, Recurrent Neural Network (RNN) algorithms have been adopted for motion prediction of 3D sequences.

2.5.1 Challenges

This chapter introduces the background of the thesis to investigate head motion prediction.

2.5.2 Kalman Filter

This chapter introduces the background of the thesis to investigate head motion prediction.

2.5.3 Deep Learning Algorithms

ML refers to a range of techniques that aim to allow statistical models and algorithms to obtain information from given data.

3 Related work

This chapter introduces the background of the thesis to investigate head motion prediction. First, the concept of extended reality (XR) related to the research problem is presented, followed by introducing different head motion prediction approaches.

3.1 Predicting head position in 3-DoF

This chapter introduces the background of the thesis to investigate head motion prediction.

3.2 Previous methods in 6-DoF

Microsoft HoloLens, known under development as Project Baraboo, are a pair of mixed reality smartglasses developed and manufactured by Microsoft.

3.3 Deep Learning models in 6-DoF

This chapter introduces the background of the thesis to investigate head motion prediction.

Data and Model

4

It allows web pages to be updated asynchronously by exchanging data with a web server behind the scenes.

4.1 6-DoF Dataset

The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

4.1.1 Data collection from HMD

It allows web pages to be updated asynchronously by exchanging data with a web server behind the scenes.

4.1.2 Data preprocessing

It allows web pages to be updated asynchronously by exchanging data with a web server behind the scenes.

4.2 Network Architecture

The history of artificial neural networks is filled with colorful, creative individuals from a variety of fields, many of whom struggled for decades to develop concepts that we now take for granted.

4.3 Network input

The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

4.4 Training methods

The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

5 Implementation and experiments

Here some code for my super neural network. The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

```
class StudentFactory(DjangoModelFactory):
        class Meta:
                model = Student
        student_card = factory.SubFactory(StudentCardFactory)
        first_name = factory.Faker('first_name')
        second_name = factory.Faker('last_name')
```

Listing 5.1: StudentFactory

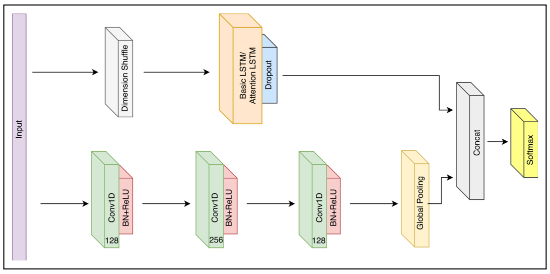

Figure 1: LSTM Fully Convolutional Networks for Time Series Classification

You might already know that you want to apply an established theory or set of theories to a specific context (for example, reading a literary text through the lens of critical race theory, or using social impact theory in a market research project).

5.1 Implementation

The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

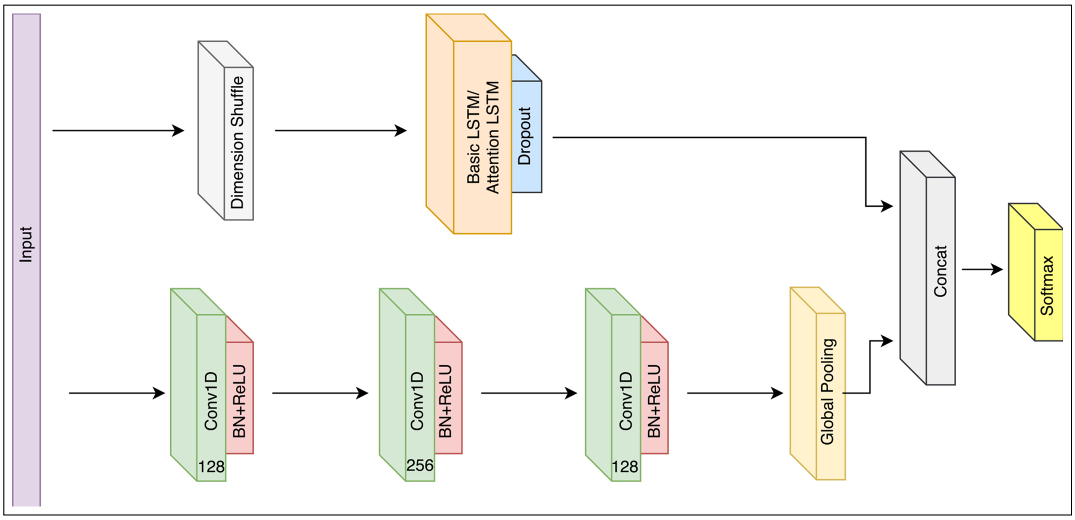

Figure 2: LSTM FCN WRAP

5.2 Experiments

The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

5.3 Evaluation metrics

The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

5.4 Results

The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

6 Analysis

The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

6.1 Limitations

The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

6.2 Conclusion

The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

6.3 Suggestions for future work

The artificial neural networks discussed in this text are only remotely related to their biological counterparts. In this section we will briefly describe those characteristics of brain function that have inspired the development of artificial neural networks.

Glossary

AJAX

AJAX (asynchrones Javascript und XML) ist der allgemeine Name für Technologien, mit denen asynchrone Anforderungen (ohne erneutes Laden von Seiten) an den Server gestellt und Daten ausgetauscht werden können. Da die Client- und Serverteile der Webanwendung in verschiedenen Programmiersprachen geschrieben sind, müssen zum Austausch von Informationen die Datenstrukturen (z. B. Listen und Wörterbücher), in denen sie gespeichert sind, in das JSON-Format konvertiert werden.

Bibliography

[1] Deniz Aladagli, Erhan Ekmekcioglu, Dmitri Jarnikov, and Ahmet Kondoz. *Predicting head trajectories in 360° virtual reality videos*. https://ieeexplore.ieee.org/document/8251913. 1-6. 2017 International Conference on 3D Immersion. Last access on 25.03.22.

[2] Serhan Guel, Sebastian Bosse, Dimitri Podborski, Thomas Schierl, and Cornelius Hellge. *Kalman Filter-based Head Motion Prediction for Cloud-based Mixed Reality*. https://arxiv.org/abs/2007.14084. Accepted at the ACM Multimedia Conference (ACMMM) 2020. 1-9. Last access on 05.04.22.

[3] Serhan Gül and Sebastian Bosse. *Kalman Filter-based Head Motion Prediction for Cloud-based Mixed Reality.* lulu.com London, 2020. ISBN: 978-1-291-61050-5.

[4] Serhan Gül, Dimitri Podborski, Thomas Buchholz, Thomas Schierl, and Cornelius Hellge. *Low-latency Cloud-based Volumetric Video Streaming Using Head Motion Prediction*. https://arxiv.org/abs/2001.06466. 30th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV) 2020. 1-7. Last access on 08.04.22.

[5] Fazle Karim, Somshubra Majumdar, Houshang Darabi, and Shun Chen. *LSTM Fully Convolutional Networks for Time Series Classification*. https://arxiv.org/abs/1709.05206. 1-7. 2017. Last access on 12.04.22.

[6] Anh Nguyen, Zhisheng Yan, and Klara Nahrstedt. *Your Attention is Unique: Detecting 360-Degree Video Saliency in Head-Mounted Display for Head Movement Prediction*. https://www.researchgate.net/publication/328370817. 1-9. 2018 ACM Multimedia Conference. Last access on 23.03.22.

[7] Silvia Rossi, Irene Viola, Laura Toni, and Pablo Cesar. *A New Challenge: Behavioural Analysis Of 6-DOF User When Consuming Immersive Media*. https://ieeexplore.ieee.org/document/9506525. 1-5. 2021 IEEE International Conference on Image Processing (ICIP). Last access on 13.04.22.

[8] Silvia Rossi, Irene Viola, Laura Toni, and Pablo Cesar. *From 3-DoF to 6-DoF: New Metrics to Analyse Users Behaviour in Immersive Applications*. https://www.researchgate.net/publication/357172010. 1-7. 2021. Last access on 13.04.22.

[9] Emin Zerman, Radhika Kulkarni, and Aljosa Smolic. *User Behaviour Analysis of Volumetric Video in Augmented Reality*. https://ieeexplore.ieee.org/document/9465456. 1-4. 2021 13th International Conference on Quality of Multimedia Experience (QoMEX). Last access on 27.03.22.