Master Thesis

Evaluation of an efficient web-based video conferencing approach using image reconstruction from face-meshes

Author:

Felix Bublitz

(Matr. Nr.: 376304)

Adviser:

Prof. Dr. Manfred Hauswirth

Dr. Louay Bassbouss

Open Distibuted Systems

TU Berlin

April 29, 2022

Declaration of Authenticity

I declare that I completed the Bachelor thesis independently and used only these materials that are listed. All materials used, from published as well as unpublished sources, whether directly quoted or paraphrased, are duly reported. Furthermore I declare that the Bachelor thesis, or any abridgment of it, was not used for any other degree seeking purpose.

Place, Time	Signature

Abstract

Lorem Impsum

A – B – C

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The worldwide Corona pandemic led to extensive changes in many areas of life. In private and professional contexts, online conferences and video calls gained in importance. Even lectures and seminars in schools and universities are increasingly held online. The resulting rise of digital image transmissions leads to an increase of global traffic. While the annual global data volume in 2018 was 33 zettabytes, forecasts predict an increase to 175 zettabytes by 2025 [2]. Although sufficient network capacity is available in Germany for the moment [5], some precautionary steps have already been taken due to the Corona pandemic: Streaming providers such as Netflix and YouTube decided to reduce their transmission quality temporary in order to counteract possible network congestion due to the increased usage [4]. Especially in light of the high data volumes in video transmission, new possibilities are continuously evaluated to reduce the data while maintaining the same quality. Apart from video conferencing, there are efforts to enable even more encounters online in the future. In this regard, Meta recently introduced the Metaverse, in which people can meet as avatars in a virtual space [1]. An important basic requirement for this technology is the recognition of postures and gestures of real persons in order to transfer them into the virtual world. For this purpose, new developments in feature extraction can be used, which enable the recognition of postures, hand gestures and facial expressions via machine learning based on the camera image [3].

1.2 Problem Statement

In the special field of video conferencing, the transmitted image is usually a composition of a static background and a person in the foreground. Here, the change of the image over time is mainly limited to the movement of the head and shoulder area. These movements can be detected by body feature extraction, which produces memory-efficient geometry data of the face expression, as well as translation and rotation. By using such a real-time feature extraction method on the sender side and a suitable reconstruction method on the receiver side, an alternative video transmission can be realized. Considering the platform

independence, web technologies are particularly suitable for this purpose. Regarding this new approach, it is unclear if special requirements must be met, what performance and image quality can be achieved, and whether such a solution is appropriate for practice use.

1.3 Methodology

This thesis will examine, whether a data-saving, web-based video conferencing approach can be realized using feature extraction and a suitable reconstruction method. The quality of the solution will be evaluated on the basis of different parameters (e.g.: latency, bandwidth, CPU usage and image quality) and potential restrictions compared with traditional video transmission should be pointed out. For this purpose, the following objectives are defined:
Existing methods for face feature extraction and reconstruction are analyzed, and suitable methods are selected for the realization of a proof-of-concept. This will be designed and developed subsequent. The individual components of the solution are evaluated independently, and the general feasibility is verified. For this purpose, suitable metrics for determining quality and performance are developed and defined.

Chapter 2

Related Work

Chapter 3

Foundations

3.1 Feature extraction

3.1.1 mediapipe

3.1.2 open pose

3.2 face reconstruction

3.2.1 ...

Chapter 4

Implementation

Chapter 5

Evaluation

Chapter 6

Conclusion and future work

Bibliography

[1] CNET: *Everything Facebook revealed about the Metaverse in 11 minutes [Video]*. 2021. https://www.youtube.com/watch?v=gElfIo6uw4g29.04.2022.

[2] F. Tenzer: *Prognose zum Volumen der jährlich generierten digitalen Datenmenge weltweit in den Jahren 2018 und 2025*. 2018. https://de.statista.com/statistik/daten/studie/267974/umfrage/prognose-zum-weltweit-generierten-datenvolumen29.04.2022.

[3] Google: *MediaPipe*. 2020. https://google.github.io/mediapipe/29.04.2022.

[4] ZDF Heute: *Entlastung fürs Netz – Corona-Krise: Youtube reduziert Bildqualität*. 2020. https://www.zdf.de/nachrichten/digitales/coronavirus-youtube-bildqualitaet-100.html29.04.2022.

[5] ZDF Heute: *Telekom zu Corona-Krise – Netze halten dem Ansturm stand*. 2020. https://www.zdf.de/nachrichten/wirtschaft/telekom-zu-corona-krise-netze-halten-dem-ansturm-stand-100.html29.04.2022.

Anhang