SCHOOL OF COMPUTATION,
INFORMATION AND TECHNOLOGY

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Robotics, Cognition, Intelligence

Evaluation of the BeamStellar tracking and pose estimation system for a Beaming Display

Jonas M. Weigand

SCHOOL OF COMPUTATION,
INFORMATION AND TECHNOLOGY

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Robotics, Cognition, Intelligence

Evaluation of the BeamStellar tracking and pose estimation system for a Beaming Display

Auswertung des BeamStellar Verfolgungs- und Positionsschätzungssystems für Beaming Displays.

Author:	Jonas M. Weigand
Supervisor:	Prof. David Plecher
Advisor:	Dr. Christian Eichhorn, Prof. Yuta Itoh
Submission Date:	29.09.2025

I confirm that this master's thesis in robotics, cognition, intelligence is my own work and I have documented all sources and material used.

Munich, 29.09.2025 Jonas M. Weigand

Acknowledgments

Abstract

Contents

1 Introduction

Virtual Reality Headsets and other Head Mounted Displays arrived at the consumer market in the year 2016. Since then many global mega corporations took their shoot to conquer this market and developed many HMDs. We have seen cable bounded models where computation can be done on a high power stationary PC and battery life is no limitation for the experience but cables limit the movement range of the user and might oppose a danger of tripping over it. Some approaches used gaming laptops in a backpack to avoid the cable issues but had some loss in image quality due to less computational resoureces available on the mobile platform and limited battery life to support everything. Also these backpacks were pretty heavy and exausting for users. A newer approach is to include highly efficient smartphone chips of custom made silicon into the headsets to improve battery life and reduce the physical size of the computational side. This approach also limits the apps and games that the headset is capable of running at a decent framerate. Another approach was a wireless connection to stream the video to the headset from a stationary PC but this requires a sophisticated highsped wireless bandwidth which might compete with WIFI and Bluetooth devices for time on the medium and thus might suffer from reduced throughput or instabilities. Altough all their efforts non of the models could achieve high adoption. REFERENCE STEAM HARDWARE SURVEY AND STATISTA SELL NUMBERS Manufacturers put a lot of effort to improve the Headset ergonomics and improve the user experience. Displays improved, headset Weight was reduced and better distributed, motion sickness was studied and prevention tactics strengthend. Anyways broad adoption by mainstream users still lacks behind. This raises the question if the current headset approaches for the VR headset technology can overcome the challenges that keep users from using this technology or if we need a new approach to give users a VR or AR experience. A key aspect for VR and AR experiences is that users percieves a virtual 3D enviornment or the real world is overlayd with additional graphical information that adapts and follows the users head and body movement. Current approaches do this by mounting displays right infront of the users eyes and constraint them to their head. Some devices in the AR field use semi transparent mirrors to redirect a displays light in the optical path of the eyes. All these approaches add a lot of technology to the user himself and try to optimize for some aspects in the weight, graphical result and battery life triangle. Beaming Displys are a new concept for head mounted displays that try to

remove as much technology as possible from the user. The Beaming Display concept is relativly new and foundations have to be layed there is no commercial available product yet. This concept uses a projector and a steering mirror that are mounted somewhere in a room. A user wears a minimalistic pair of glasses that allow for tracking the user and receiving the light from the projector without harming the users eyes. Any computation can be done stationary with access to a wall outlet and the hardware can be as capable and size demanding as desired. This probably boosts the users visual experience to unseen heights and details while preserving a natural presence in the real world. We know from previous research that the motion to photon latency is critical in any AR application for user acceptance. Thus a fast tracking and controlling system is nessessary for the Beaming Display concept and topic of this Thesis.

This thesis evaluates a fast 6-DoF Tracking and Pose estimation system for a Beaming Display, a special kind of Head Mounted Displays (HMD). Specifically XXX well established Perspective-n-Point (PnP) pose computation algorithms are implemented in the programmable logic (PL) of a Field Programmable Gate Array (FPGA) for highspeed HMD tracking and projection correction. The resulting pose estimations are validated against ground truth data provided by an outside in tracking sysetem. Also a fesible application volume for the Beaming Dispay station is narrowed down.

1.1 Section

Citation test [Lam94].

1.1.1 Subsection

2 Related Work

2.1 User acceptance of VR

Citation test [Lam94].

2.1.1 Subsection

2.2 Beaming Display

2.3 Sensing the Glasses

2.4 Perspective-n-Point (PnP) Pose Computation

The glasses are not parallel projection surface in relation to the projector in the Base Station. This means depending on the users position and head rotation a rectangle sent by the projector is recived as a croped trapeze in most cases as the user stands rarely parallel to the projectors plane. Also the mirrors angels affect the optical pipeline and add additional effects to the image. Such transformed images will defnitly break the users immersion and crash the accaptance and has to be avoided. We focus on digital methods to manipulate the image beforehand on the projector side in a way that the received image is rectangular again. High resolution projectors became broadly available in the past years and some of the resolution can be sacrificed to counter the cropping problem and still deliver decent resolution to the user. The Pose of the user relative to the projector has to be known to support any image rectification methods. In the BeamStellar concept we try to reduce the number of external devices as much as possible. Many VR Headsets needed external tracking devices for positioning systems and input device tracking which required preparing of a room for VR. These tracking devices had to be mounted on walls and needed to be kept stationary to avoid recalibration. We expect low user acceptance if they need a regular tracking system and additional units in the room to use a Beaming Display. Thus integrating the tracking in the same device as the projector would be really beneficial. Previous works have used a LED on the glasses that is tracked through the same optical path as the projector sends

the image. Some systems use a camera and others used 2D Position Sensitive Devices (PSD sensors) to track the LED point.

The Base Station resembles a classical pinhole camera no matter which type of camera sensor is used. Estimating a pose of an object in an image is a solved problem in computer vision but has some constraints that have to be accounted for. The concept is called Perspective-n-Point Pose Computation and requires n points that have to be visivle on the object. Knowing the real translations of the points on the real object is key to get good estimations. Diffrent methods to solve the PnP problem have be proposed in the past and optimised for different use cases. Some require more points some can cope with fever but might provide multiple possible pose solutions, other methods require a coplanar constraint that all points have to lay on the same plane on the obejct, One of the biggest obstacles in applying the PnP computation is a reliable feature tracking and mapping in the observed image. We want to track a pair of glasses which front facing part is often a plane so we can use algorithms that require a coplanar constraint. Altough we could add some kalman filter to our pose output and thus choose the right pose if a method proposes multiple it's easier to add an additional feature to the glasses and reducing the possible poses by this.

There are multiple implementations of PnP solvers in OpenCV implemented and offer a tested and transparent reference for our implementation.

From the list of available solvers we've selected the following methods.

2.4.1 IPPE

2.4.2

List of Figures

List of Tables

Bibliography

[Lam94] L. Lamport. *LaTeX : A Documentation Preparation System User's Guide and Reference Manual*. Addison-Wesley Professional, 1994.