2D Tracking in Rock Climbing with Temporal Smoothing

André Oskar Andersen
`wpr684@alumni.ku.dk`

1 Introduction

The following paper is the description of my upcomming master thesis in computer science offered at the University of Copenhagen in 2023. It covers the motivation and the most important background behind the thesis, as well as the the aim and learning objectives of the thesis. The thesis will be done in collaboration with ClimbAlong [1], which will be described further in Section 3.

2 Background

Video analysis in sports have throughout the last decade become more and more common, as these recording contains a lot of important information. By analyzing such a video recording of people engaging in sports, we can for instance help a referre make the correct calls or help the people engaged in the sport develop their technique. However, most of these analyses requires the system to know where the relevant people and objects are in the video recordings. The models that perform such a task are already developed for the most popular sports, such as football or basketball, and tend to deliver very accurate results. On the other hand, for the less popular sports, such as rock climbing, such models are not as common.

To perform video analysis **convolutional neural networks** are often used. Convolutional neural networks are a type of neural networks that are commonly used for capturing the spatial information of an image or video. These models have a wide range of applications. For finding the people or objects in a video, the most relevant application is the act of **tracking**, where the model estimates the position of humans or objects in a given video.

---

[1] `https://climbalong.com/`

However, these convolutional neural networks most oftenly process the frames of an input video independently, leading to suboptimal results. As the individual frames of a video contains data that correlates across the frames, one can make use of **temporal smoothing** techniques to optimize this performance. There are multiple methods to exploit this temporal smoothing. For instance one could use a running average to quickly implement temporal smoothing. Another more sophisticated way of implementing temporal smoothing is by using a variation of a **recurrent neural network**. Recurrent neural networks are another type of neural networks, that are commonly used for capturing the correlating information across a sequence of input data.

3 Collaboration

The thesis will be developed in collaboration with ClimbAlong. ClimbAlong will be providing an annotated dataset of video recordings of people climbing a bouldering wall. This dataset is not a synthetic dataset but instead a real dataset, making the developed results with this dataset yield realistic results. ClimbAlong has also a developed and trained convolutional neural network for tracking the humans and climbing holds/grips in 2 dimensions in a given video recording, which processes each frame independently of the other frames of the video recording which will be provided for the thesis. A coorperation agreement with ClimbAlong is currently in development.

4 Problem Statement

The aim of this thesis is to incorporate temporal smoothing into a model for 2 dimensional tracking of rock climbers and climbing holds/grips. This may either be done by extending the already 2 dimensional tracking algorithm provided by ClimbAlong or by implementing a whole new model from scratch. Our developed model will thus either be processing features of the input video or be processing the predictions of the ClimbAlong algorithm. We wish to develop such a model by experimenting with various setups and network architectures, such that we can optimize the results. The development of the various setups and network architectures will be based on state-of-the-art methods, either by letting us be inspired by or completely implement these methods.

The development of the experiments will be done by making use of an annotated dataset of video recordings of rock climbers provided by ClimbAlong. This dataset may not be of sufficient size, so we may extend it by either making use of data

augmentation or by finding a similar dataset which can be used in addition to the ClimbAlong dataset. A coorperation agreement with ClimbAlong is currently in development.

Once we have developed a sufficient model for 2 dimensional tracking, we will be discussing our approach and the performance of this model. This will include a discussion about what we could have done differently or better, as well as some suggestions for possible future work for our developed model.

5 Learning Objectives

1. Present the theory behind convolutional neural networks and recurrent neural networks.

2. Survey state-of-the-art methods for 2 dimensional tracking with temporal smoothing.

3. Design, implement and evaluate models for 2 dimensional tracking of climbers and climbing holds/grips with temporal smoothing.

4. Discuss our approach and the performance of the best-performing developed model.