

TALLINN UNIVERSITY OF TECHNOLOGY

Faculty of Information Technology

Department of Software Science

Allar Viinamäe, 163578 IAPM

**DATA DRIVEN GYMNASTICS SKILLS
RECOGNITION AND ANALYSIS**

Master's thesis

Supervisors: Sven Nõmm, PhD

Tallinn 2020

TALLINNA TEHNICAÜLIKOOL

Infotehnoloogia teaduskond

Allar Viinamäe, 163578 IAPM

**ANDMEPÖHINE VÕIMLEMISOSKUSTE
TUVASTUS JA ANALÜÜS**

Magistritöö

Juhendajad: Sven Nõmm, PhD

Tallinn 2020

Author's declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Allar Viinamäe

17.04.2020

Contents

List of Figures	3
List of Tables	4
1 Introduction	5
1.1 Human Activity Recognition Background	6
1.1.1 Human Activity Types	6
1.1.2 Recognition Types	6
1.2 Problem statement	6
1.3 Related work	6
2 Implementation	7
2.1 Implementation overview	7
2.2 Infrastructure and Tools	7
2.2.1 Client Infrastructure and Tools	7
2.2.2 Back-end Server Infrastructure	7
2.2.3 Research and development	7
2.3 Data acquisition	7
3 Data acquisition	8
3.1 Recording Human Actions	9
3.1.1 Choosing Activities To Record	9
3.1.2 Backflip	9
3.1.3 Back Handspring	11
3.1.4 Recording Process	11
3.2 Estimating and Collecting Action Poses	12

4 Data pre-processing	13
4.1 Data description	13
4.2 Drawing Entity	13
4.2.1 Outlier removal	13
4.2.2 JSON to Drawing entity Conversion	13
Bibliography	15

List of Figures

3.1	Example of a backflip	10
3.2	Example of a back handspring	10
3.3	Start and end markers for covering the full duration of a backflip . . .	11

List of Tables

Chapter 1

Introduction

Introduction...

1.1 Human Activity Recognition Background

1.1.1 Human Activity Types

1.1.2 Recognition Types

1.2 Problem statement

1.3 Related work

Chapter 2

Implementation

2.1 Implementation overview

2.2 Infrastructure and Tools

2.2.1 Client Infrastructure and Tools

Hardware

Software

2.2.2 Back-end Server Infrastructure

Software

2.2.3 Research and development

2.3 Data acquisition

Chapter 3

Data acquisition

As with most problems potentially solvable by implementing machine learning based solutions, the data to be explored and used to train machine learning models plays central part in our implementation. Although, machine learning promises to loosen the strictness of the data used in a system by trying to generalize and adapt the models themselves to the data, the quality of the initial data used to develop prototypes still directly influences the interpretation and value of the outcome. Interpretation of not only the outcome, but the entire solution is needed to have clear understanding of why certain outcomes exist, which in turn helps to spark discussion and spread knowledge gained during an experiment. The value of this work as a scientific paper, reusable by peers for new scientific experiments, is also directly influenced by the quality of the initial data. The author of this work is aware, that by choosing a very niche dataset he risks to reduce the social reach of this work and thereby stresses to point that he mainly wishes to contribute his finding to the human activity recognition domain and not necessarily to the specific sports domain of gymnastics.

The data acquisition chapter is split into two parts. First, the part explaining what kind and how much of data the author is collecting and the process behind it. The second part explains a more technical approach on how the initial data for the human activity recognition algorithm is acquired from recorded human motions and the tools used to do so.

3.1 Recording Human Actions

3.1.1 Choosing Activities To Record

When it comes to choosing difficult biomechanical activities performed by humans, gymnastics comes on the top of the list. Gymnastic feats don't require any additional equipment by humans, but physical strength, flexibility and kinesthetic awareness are a must to perform any of the skills demonstrated by elite athletes. Gymnastics is also special in its indisputable need for utilizing every part of the human body. All major muscle groups of an athlete need to be in superior condition to perform certain rotations, jumps and holds. So, many athletes start their career at a very young age.

Automatic recognition of human activity requires reference or training data to train our machine to recognize some gymnastic movements. The two activities chosen to prove the hypothesis of this paper are backflips and back handsprings. Figures 3.1 and 3.2 will hopefully give the reader a visual idea of backflips and back handsprings respectively.

For each activity there needs to be specific start and end markers, so it could be easily recognized when the activity starts and when does it end.

The following is a detailed description of each activity and its parameters.

3.1.2 Backflip

A backflip is a sequence of body movements in which a person leaps into the air and rotates backwards over the body's horizontal axis. For the backflip, we mark the start of a backflip as the frame when athlete's both arms pass the horizontal line at shoulder level moving downwards and generating momentum. We mark the end of a backflip as the frame when both heels of the feet touch the ground again. We choose the heels of feet so we can include the amortization part of the landing phase of the backflip in the recording. The red bar in figure 3.3 demonstrates visual markers for trimming the sample recording to include only the activity under investigation. These markers were chosen by the author to the best of his knowledge of the domain as the clearest points for a human observer to recognize a backflip.



Figure 3.1: Example of a backflip

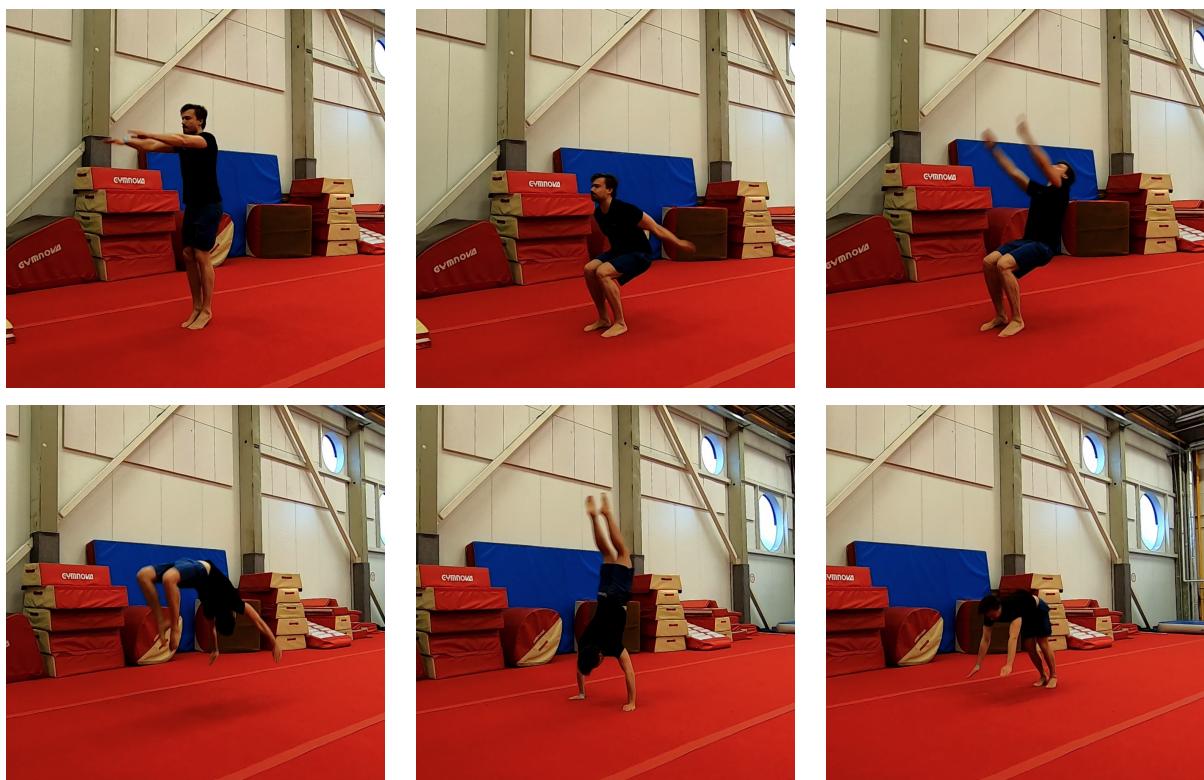


Figure 3.2: Example of a back handspring

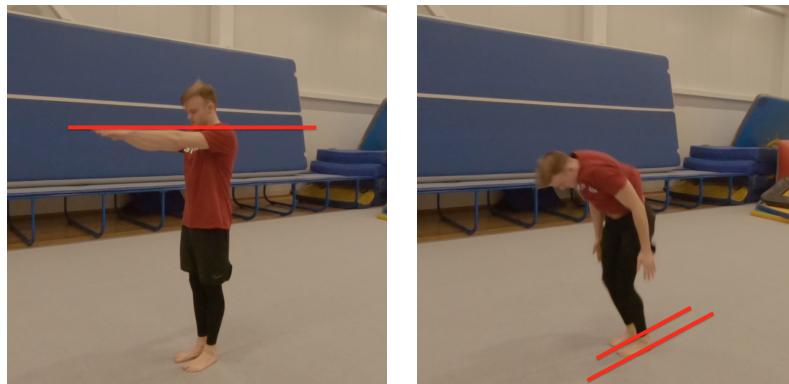


Figure 3.3: Start and end markers for covering the full duration of a backflip

Choosing different markers is a potential discussion topic for future improvements.

3.1.3 Back Handspring

TODO

3.1.4 Recording Process

Since these types of movements are performed for a duration of time, just still images of activities are not sufficient. We need the activities recorded in some kind of motion picture format. The device used to capture activities is a GoPro Hero7 Black, with the following basic settings:

- *RES* — 1080
- *FPS* — 60
- *FOV* — Linear
- *Low Light* — Auto
- *Stabilization* — Auto
- *Protune* — Off

Each activity is recorded as one still view from eye level angle fully showing the subjects body from head to toe.

3.2 Estimating and Collecting Action Poses

Chapter 4

Data pre-processing

4.1 Data description

4.2 Drawing Entity

4.2.1 Outlier removal

4.2.2 JSON to Drawing entity Conversion

Acknowledgments

Bibliography