



WHAT WILL BE THE EFFECT OF DIFFERENT AMBIANCE CONDITIONS ON THE ACCURACY OF BALANCE RECOVERY MEASUREMENTS USING A CAMERA BASED SYSTEM?

Ewout van den Pol

ENGINEERING TECHNOLOGY/BIOMECHANICAL ENGINEERING EDWIN .H.F. VAN ASSELDONK

EXAMINATION COMMITTEE Aurora Ruiz Rodríguez G. Ruben H. Regterschot

DOCUMENT NUMBER BE - 815

UNIVERSITY OF TWENTE.

1 Abstract

Making use of an exergame using a camera-based system can result in less falling among stroke patients. Studies showed promising results so such a method could be used in a home environment. This study focuses on what the effect can be of light conditions, background conditions and clothing on balance recovery measurements using a camera-based system. The results of this study show that different ambience conditions will have an effect on balance recovery measurements and should be taken into account when designing a camera-based system that will be used in a home environment. Wearing casual clothing instead of black clothing and having items in the background will increase the accuracy of balance recovery measurements using a camera-based system. However, when the distance between the camera and the participants increases from about 3 meters to 3.3 meters, the accuracy of balance recovery measurements dramatically decreases.

Contents

1	Abstract					
2	2 Introduction					
3	Background 3.1 Noise measurement 3.2 Human posture recognition 3.3 Dark environment 3.4 Skeleton tracking 3.4.1 Skeleton Tracking SDK by Cubemos 3.4.2 Nuitrack 3.4.3 Openpose	2 2 2 2 2 3 3 3 3				
4	Protocol 4.1 Protocol 1	4 4 5 7 7 8				
5	Methods 5.1 Measurements overview 1 5.2 Coordinates system 1 5.3 Processing data 1 5.3.1 Missing joints 1 5.3.2 Positioning 1	9 .0 .2 .2				
6	Results 1 6.1 Recognized joints	4 .5 .8 21				
7	Discussion 2	7				
8	Conclusion 29					
A	Appendixes30A.1 Overview of percentage not recognized joints30					
в	Literature search 3 B.1 Useful information 3	1 33				

2 Introduction

Frequent falling is for stroke surviving people one of the most common medical complications. A geriatric unit reported that falling occurs by 15.9 out of 1000 patients per day [15]. Falling mostly occurs in stroke patients due to impairment of balance and self-care and also impairment of cognitive function. For stroke patients, falling is a leading cause of fractures, with a percentage of 23-50 %. Falling can also result in the reduction of daily independence activities, such as bathing, eating, dressing, and toileting of a patient [5]. A solution to this problem could be the use of an exergame, which means a game in which technology-driven physical activities are used to play the game. The participant will undergo different kinds of exercises, such as stepping, lifting a leg, etc. and receive feedback about his or her balance recovery abilities. This can help stroke survivors exercise stepping responses so falls could be prevented. For this reason, the HEROES [7] project was defined. One of the goals of the HEROES project is to play a video game while track movement from the patients, and for this a camera-based sensor will be used. To give useful feedback, it is important for the camera sensors that they can detect accurately where the patient is and how the patient is moving.

This will be possible by making use of the Intel RealSense D435 camera. This camera is an active stereo depth camera and is well suited for this occasion. The camera can make use of the infrared projector to improve the depth accuracy [2]. When combining this camera with the Intel Skeleton Tracking Software Development Kit (SDK), accurate tracking of the body is possible.

Other studies, such as Lai et al. [10] and Abrea et al. [1], already explored the usage of a camera-based system for rehabilitation therapy for post-stroke patients. They let patients do the exercises and this resulted in promising results for recovering balance function by stroke patients. Both studies made use of the Kinect system and concluded that such a system can be used as an alternative for complex and high-cost motion capture devices. However, with previous testing of the sensor we found out that such a system is not perfect when using different ambience conditions, for example, different lighting and different backgrounds, the system will fail to track the patient accurately. Most of the studies were done in the lab and not in a home environment, where ambience conditions are most of the time not optimal. Think of sunlight, lack of lighting, items in the background, etc.

Therefore, the research question of this paper is as follows: What will be the effect of different ambience conditions on the accuracy of balance recovery measurements using a camera-based system?

3 Background

3.1 Noise measurement

When capturing images using an image sensor, noise will always affect these images. Noise in images can be the result of the image formation process, the image recording or other processes which are needed for capturing an image. This will result in different kinds of sometimes random distortions which can make the captured image blurry. Noise in images can also be harmful to image processing, which could be the case for human posture recognition [13].

3.2 Human posture recognition

A lot of research about human posture recognition uses images that are made by an RGB (Red Green Blue) camera [8] [4] [12]. The limitations with only using RGB data was that the RGB data would be affected by environmental noise, such as the intensity of the light, the angle, etc. But when the Microsoft Kinect came out, a lot of researchers [11] [9] [17] were interested in the Kinect, because not only used it RGB data, but also depth data of the objects. Based on both data, the depth image data could be acquired, as well as the human body skeleton data [4]. Although the Kinect uses depth data, a light environment is still preferred for human detecting, because of the light-based RGB camera [3]. Another use of human recognition can be identifying hand gestures. It was still possible to recognize the hand and the gesture that it made, even with a complex background [16].

3.3 Dark environment

Most person detection libraries makes use of the RGB images and which are fed into it. Using an algorithm, body skeletons will be applied to the images. But in dark environments, typical RGB cameras can fail to generate sufficient visually understandable images. A solution to this problem can be the use of a thermal camera which can successfully operate in a dark environment [14]. However, in home environments an accessible camera based sensor is needed for an exergame and therefore this thermal camera is not an option.

3.4 Skeleton tracking

An important part of recognizing stepping responses is to be able to track the skeleton of an individual. The human body is virtually represented by a number of joints in skeleton tracking. Every joint has its own 3D positional coordinate. By making use of an RGB and an infrared camera, those 3D coordinates can be found. Using these joints, a skeleton body can be made of the person [17].

There are a variety of skeleton tracking systems available, e.g. Cubemos, Nuitrack, Open Pose, etc. This technique is quite new, which makes it hard to find good reliable information about the different skeleton tracking systems. The following information is obtained by the official websites from the software and forums.

3.4.1 Skeleton Tracking SDK by Cubemos

Skeleton Tracking SDK by Cubemos is useful for up to 5 people in a scene. Furthermore is it compatible with Intel RealSense cameras, so also compatible with the RealSense Depth Camera D435i which will be used during the assignment¹.

3.4.2 Nuitrack

The Nuitrack system can support the RealSense Depth Camera D435 i^2 . The free version has a time limit of a total of 3 minutes per session, then you will need to restart it³.

3.4.3 Openpose

Openpose should be able to work in 3D with the Intel RealSense d435i⁴. It can be more challenging with respect to Nuitrack and Cubemos, because it could require more knowledge than e.g. using the Skeleton Tracking SDK by Cubemos, which is built for RealSense cameras.

The Cubemos and Nuitrack options look the most promising, especially for the first time to set everything up. Openpose looks also promising because you have more freedom because it is open source. But the downside to it is that it probably is less user friendly, so for example for the first testing of the camera, the other options are more promising. The three software listed above seems like a good fit, however, for this assignment the Cubemos was select because of the synergy with the Intel RealSense camera which was used.

¹https://www.intelrealsense.com/skeleton-tracking/

 $^{^{2} \}rm https://community.nuitrack.com/t/can-nuitracksdk-support-intel-d435i-or-only-d435-d415/1403$

³https://nuitrack.com/faq

 $^{{}^{4}} https://github.com/CMU-Perceptual-Computing-Lab/openpose/issues/1625$

4 Protocol

In this section are the different types of protocols discussed. The pilot test protocol has been named protocol 1 and the later reformed protocol has been named protocol 2.

4.1 Protocol 1

First a pilot test was made, and the protocol goes as follow:



Figure 1: Overview of stepping directions.

The participant will move in a total of 8 directions, see figure 1. 5 times in each direction over a distance of a total 0.30 meters. The participant will step with both feet in a direction, 1 foot at a time, until he stands in the next position. The camera will be placed on a table at a height of 1.34 meters. The distance from the camera to the beginning position will be 2.5 meters, see figure 2.



Figure 2: Sketch of the experiment location of protocol 1.

The participant will do a step as follow:

- 1. Both legs will move in the up direction
- 2. Both legs will move in the up-left direction
- 3. Both legs will move in the left direction
- 4. Both legs will move in the down-left direction
- 5. Both legs will move in the down direction
- 6. Both legs will move in the down-right direction
- 7. Both legs will move in the right direction
- 8. Both legs will move in the up-right direction

After the pilot testing, I realized that this protocol could be improved. The distance between the participant to the camera has been increased, see figure 3, so the camera has a bigger field of view. Also, in Protocol 2, the participant will only step with one foot in a direction. Therefore the measured distances will be more accurate. That is why a new protocol was defined.

4.2 Protocol 2

The participant will move in a total of 8 positions, 5 times per position over a distance of a total of 0.30 meters. The directions in which the participant will

move are shown in figure 1. After each direction, the participant returns to the beginning position.

The participant will do a step as follow:

- 1. Left leg will move in the up direction
- 2. Left leg will move in the up-left direction
- 3. Left leg will move in the left direction
- 4. Left leg will move in the down-left direction
- 5. Right leg will move in the down direction
- 6. Right leg will move in the down-right direction
- 7. Right leg will move in the right direction
- 8. Right leg will move in the up-right direction

This approach was chosen based on Gosine et al. [6]. In their work, they showed that using a Kinect sensor they captured reasonably accurate measurements of the stepping displacements and velocities of a moving participant. Furthermore, the distance of the steps to the beginning point will be 0.3m. This is based on the results of Gosine et al. [6]. The distance between the camera and the participant is now increased to a total of 2.89 meters. The camera is again placed at a height of 1.34 meters. An overview of the situation can be seen in figure 3.



Figure 3: Sketch of the experiment location of protocol 2.

4.3 Different ambience conditions

For both protocols, a variety of ambience conditions were considered. Below is described which ambience conditions and what kind of measurements were used with the protocols.

4.3.1 Clothing types

Different clothing could have an effect on the results of the system. Therefore, the participant will use 2 different pairs of clothing: 1 pair of clothing with a coloured sweater and blue jeans. The other pair consists of a black long-sleeved shirt and black sweatpants. Only 2 sets of clothing are chosen because otherwise it would be very time consuming to take all the measures and a participant of the HEROES program would most likely wear casual clothing. To compare, a totally different colour of clothing has been chosen.

4.3.2 Lighting measurements

To analyse the lighting ambience in the accuracy of skeleton tracking some conditions were defined. These measurements are focused on what the impact will be of lighting. Furthermore, only the wall and the tiles on the floor will serve as background. Note: these measurements were made in the summer of The Netherlands.

- 1. Open curtains and at 16:00, when the sun is shining outside, lights off In home situations there is a lot of the time sunlight shining into the room. This measurement will indicate the effect of the incoming sunlight.
- 2. Open curtains and at 16:00, when the sun is shining outside, lights on. When a patient uses the rehabilitation training at home, this would likely be during the day when the sun is shining. Because only sunlight is not always sufficient for light levels inside, the lights in the room will be on.
- 3. Open curtains and at 16:00, and all the lights on. Also one lamp on which is shining almost directly into the lens. In home situations, it can occur that a lamp is at least partially aimed at the lens, which could have major consequences for the accuracy. Also, this measurement will be the measurement where the overall intensity of light will be the highest.
- 4. Closed curtains when it is dark outside, so at 22:30 or so and lights out. This will be the darkest measurement, with the lowest intensity of light. This situation will not most likely happen, but it can be interesting too see if even at very low light levels the system will be working.
- 5. Closed curtains when it is dark outside, so at 22:30 or so and lights on. This measurement will indicate what the influence of artificial light can be for accuracy.

4.3.3 Background measurements

Another important factor is that these measurements are taking in a homebased environment, that is why we need to consider the background noise due to background items. These measurements will be made during the day with the curtains open and the lights on.

- 6. Plain background: No items in the background, only the wall as background
- 7. Normal background: A background filled with random house items for a normal home ambience.

5 Methods

In this section will be explained what kind of measurements were made and how the data of these measurements were analysed.

5.1 Measurements overview

An overview of the measurements which were used with protocol 1 is shown in table 1. In this table, the different conditions of each measurement are showed.

Number	Daytime/night-time	Lights on	Light shining in lens	Background items	Type of clothing
1	Daytime	Lights on	х	Х	Casual
2	Daytime	Lights on	х	x	Black
3	Daytime	Lights on	Light shining in lens	x	Casual
4	Daytime	Lights on	Light shining in lens	x	Black
5	Daytime	Lights on	х	Background items	Casual
6	Daytime	Lights on	х	Background items	Black
7	Night-time	Lights on	х	x	Casual
8	Night-time	Lights on	x	х	Black
9	Night-time	х	x	х	Casual
10	Night-time	х	х	х	Black

Table 1: The first 10 measurements which made use of protocol 1 are shown with the specific conditions.

After the first ten measurements, additional measurements were needed to give some better results. This time, a measurement with only sunlight as a source of light was added, see the first two measurements in table 2. These new measurements made use of protocol 2.

Number	Daytime/night-time	Lights on	Light shining in lens	Background items	Type of clothing
11	Daytime	х	x	x	Casual
12	Daytime	х	x	x	Black
13	Daytime	Lights on	x	x	Casual
14	Daytime	Lights on	x	х	Black
15	Daytime	Lights on	Light shining in lens	x	Casual
16	Daytime	Lights on	Light shining in lens	x	Black
17	Daytime	Lights on	x	Background items	Casual
18	Daytime	Lights on	x	Background items	Black
19	Night-time	Lights on	x	x	Casual
20	Night-time	Lights on	х	х	Black

Table 2: The new measurements which made use of protocol 2 are shown with the specific conditions.

5.2 Coordinates system

The SDK makes use of a X, Y and Z coordinate for every joint. The position of the joints can be seen in figure 4 5 .

 $^{^5\}mathrm{This}$ picture is from the Cubemos documentation, which is stored locally when you download the SDK



Figure 4: Overview of joints which will be recognized by the Cubemos script

The coordinate [0,0,0] is referring to the centre of the physical image. The positive x-axis points to the right, the positive y-axis points down and the positive z-axis points forward. Based on these coordinates one can conclude if the system is working properly, or if the system is not so robust and therefore affected by ambience conditions⁶. An overview of the test setup can be seen in figure 5.

 $^{^{6} \}rm https://dev.intelrealsense.com/docs/projection-in-intel-realsense-sdk-20$



Figure 5: An overview of the test setup which was used

5.3 Processing data

In this section will be explained how missing joints in the data could be useful to analyse the data. Also will be explained how the coordinates provided by the Cubemos script will be used to get the distance of a certain position and how to evaluate this position.

5.3.1 Missing joints

An interesting part of the data will be where the system does not recognize the joints. When this happens, the system will not work properly and therefore this should happen at least as possible. The final system should provide accurate feedback, which will not be possible if joints are missing. That is why the number of missing joints can be useful to analyse the data. This was done by dividing the amount of not recognized joints by the total amount of joints.

5.3.2 Positioning

The Cubemos script which was used, also delivers coordinates, as explained above. These coordinates must be accurate to give the patient feedback about his or her balance recovery skills. If the accuracy of the coordinates is better, then the system can provide better feedback. The coordinates were saved to a .csv file, which was later used to analyse the data. This was done by calculating the velocity and using that find the coordinates of the positions where the patient should be standing. By making use of the known measured distance of the steps, which is 0.3 meters, this distance can be compared with the distance calculated based on the coordinates. This was visualized making use of a polar plot. To automatically get the data where the participant is standing in a certain position, the coordinates, as well as the velocity at that moment was used. The only data that was used for the polar plot needed to meet the following requirements: the velocity must be lower than 0.03 m/s and the location of the coordinates must be in a threshold of 0.2 meters around the desired standing position.

6 Results

In this section, the results of the measurements are shown. When executing the protocols, running the script resulted in a video of images like the image of figure 6.



Figure 6: Overview of the test setup fully working.

In figure 6 you can see the participant and the joint locating process of the script. All the joints are connected with a yellow line which goes in a straight line from joint to joint.

6.1 Recognized joints

To quantify the results, looking at the percentage of not recognized joints indicates how good a measurement has performed. The hip joints were most of the time recognized, with a total percentage of not recognized joints of between 1 and 2 % and not many significant differences. However, the ankle joints were most of the time not recognized with significant differences and therefore the ankle joints were used to analyse the data. If a joint is not recognized, this means that the system will be unable to give feedback about the balance recovery of the patient.

When comparing the number of recognized joints for the right ankle and left ankle, in almost every case the percentage of missing joints is higher for the left ankle than for the right ankle. This is probably because on the left side of the participant was a wardrobe located and therefore less light in comparison to the right side. With no lights on, the sunlight came in from the right side of the participant and therefore there was also a higher percentage of not recognized joints for the left ankle in comparison to the right ankle. A table with an overview can be found in section A, see table 3.

Furthermore, measurements 11-20 were made after measurements 1-10 and measurements 1-10 made use of protocol 1 and 11-20 made use of protocol 2. Overall the percentage of missing joints of the measurements making use of protocol 1 is higher compared to the measurements making use of protocol 2. This could be explained by setting up the camera further away from the participant, which increased the field of view of the camera and therefore also the number of recognized joints.

6.1.1 Type of clothing

As mentioned before, this experiment used two types of clothing. These types are viewed in figure 7.



Figure 7: Comparison of both types of clothing.

The difference between black clothing and casual clothing is that casual clothing has a lower percentage of not recognizing the joints when compared to black clothing in almost every case, as shown in figure 8.



Figure 8: Percentage of missing joints of both types of clothing.

However, when looking at the results of the very dark measurements in figure 9, this is not the case. This is most likely due to the flashlight which was used making these measurements and this flashlight was held by hand and not fastened. This could have resulted in different conditions compared to the other measurements. Furthermore, in most cases, the amount of not recognized joints for casual clothing is lower when compared with black clothing.



Figure 9: Percentage of missing joints of lights on and almost no light

6.1.2 Light condition

As mentioned before, when trying to measure in almost no light, it became clear that the system couldn't work properly.



Figure 10: Percentage of missing joints of the two protocols with a light shining into the lens.

Furthermore, when looking at the measurements of protocol 1 with a light shining into the lens, see figure 10, compared to daytime measurements with only the lights on of figure 8, the measurements with a light shining into the lens of the camera resulted in a higher failure percentage for casual clothing, with a mean percentage increase of 15.96 %. However, it resulted in a lower failure percentage for black clothing with a mean percentage decrease of 12.71 %.

But when taking new measurements using protocol 2, the measurements with a light shining into the lens, compared to the control measurements with only the lights on, for casual clothing, the mean percentage decreases by 0.47 % and the mean percentage for black clothing increases with 3.35 %. Compared to the percentages of the measurements using protocol 1 and the same conditions, these percentages are very low.



Figure 11: Percentage of missing joints of only daylight and daylight as well as lights on.

When comparing only daylight with daylight as well as lights on in figure 11, for casual clothing the percentage of also lights on increases with a mean percentage of 4.09 %. For black clothing, the percentage of also lights decrease with a mean percentage of 3.69 % when compared to only daylight. When looking at all the percentages, it seems to be that for casual clothing, only daylight is the best lighting condition.



Figure 12: Percentage of missing joints of daylight and lights on compared to night-time and lights on.

Lastly, for the difference between filming during night-time and daytime, see figure 12, when looking at measurements filmed during night-time with lights on compared to the daytime measurements with light on, for casual clothing the percentage of night-time compared to daytime increased with a mean percentage of 2.96 %. For black clothing, the percentage of night-time compared to daytime decrease with a mean percentage of 3.71 %.

6.1.3 Background items

To compare the effect of background items, the measurements with background items and the measurements without background items must be compared, as can be seen in figure 13



Figure 13: Percentage of missing joints of daytime and night-time with background items.

So when looking at the measurements making use of protocol 1 in figure 13 and comparing the measurements with only lights on with the background measurements, for casual clothing the percentage increased with a mean of 6.77 %. However, the percentage for black clothing decreases with a mean of 16.27 %. When from the new measurements which make use of protocol 2, comparing the measurements with only lights on with the measurements with background items, the mean difference of percentage not recognized joints for casual clothing is 1.27 % and the mean difference of percentage not recognized joints for black clothing. These differences are small when compared to the differences in the measurements of protocol 1. However, the difference in percentages of also the other measurements which make use of protocol 2 is already small when compared to the differences in the percentages of measurements making use of protocol 1. This can be seen in table 3 in the Appendix section.

6.2 Accuracy of joints

To display the space in between the means of each direction, a polar plot was made, see figure 14.



Figure 14: Polar plot of only daylight measurement, down direction points to camera. Blue star indicates the mean location, black bar indicates the standard deviation.

In the middle is the centre point and in the other 8 directions are the other points visible. From the middle to 0° points to the left side of the participant and from the middle to 270° points to the front of the participant. Ideally, the points should be on the 0.3 meters circle and also on 0° , then 45° etc. However, as can be seen in figure 14 the points pointing to the back are further away from the 0.3 meters circle than the points pointing to the front, within the backwards direction a total difference of 7.74 cm. This difference is a lot bigger than when looking at the other distances, the second biggest difference is just 3.64 cm. The mean of the difference in distance in every direction of the 0.3 meters circle is 2.61 cm. The mean of the difference in distance in every direction except the backwards direction to the 0.3 m circle is 1.89 cm. Along the z-axis, the standard deviation from every standing position varied from 0.75 cm to 2.70 cm. Only the 3 most backward positions resulted in a standard deviation of more than 2 cm.



Figure 15: Error of setup when executing protocol 2

Also, for most points, it looks like they are not quite on the 45° line, but the points should move counterclockwise to align with the 45° line. This is most likely because of the angle of the camera, which was set up with a slight angle when taking the measurements using protocol 2. This angle can be seen in figure 15.



Figure 16: Polar plot of the measurement with background items of protocol 2, down direction points to camera. Blue star indicates the mean location, black bar indicates the standard deviation.

When comparing figure 16 with figure 14, what stands out is that a 0.4 m circle is needed because the point in the backwards position is located at 0.428 m from the central point. Also, in contrast to figure 14, there does not seem to be a pattern in the size of the standard deviations. The highest standard deviation is now 3.60 cm. The mean of the difference in distance in every direction of the 0.3 m circle is 3.67 cm. The mean of the difference in distance in every direction except the backwards direction to the 0.3 m circle is 2.22 cm. The other measurements gave similar results as figure 16, with also coordinates for the background-position which was close to the 0.4 m circle line. However, for the measurement making use of protocol 2, with only sunlight and casual clothing and the measurement with a light shining into the lens with black clothing,

a full polar plot could not be constructed, due to missing valid coordinates. These were in the backward, left-backwards and right backward positions for the measurement with only sunlight and the forward and right-forward position for the measurement with a light shining into the lens.

7 Discussion

In this section will the results be discussed as well as compared with current literature.

Work done by Khoshelham and Elberink[9] stated that the random error for a distance of 3 meters should be about 1.4 cm. Comparing this to the found range of standard deviations of 0.75 cm to 2.70 cm, the found standard deviations seems reasonable, with only 3 standard deviations higher than 2 cm. However, this was the best result when looking both at the position of the coordinates and the standard deviation. Khoshelham and Elberink[9] also stated in their paper that the depth resolution, which refers to the minimum depth difference that can be measured, should be about 2.5 cm for a distance of 3 meters. Comparing this 2.5 cm to the found mean distance difference of 2.61 cm and 3.67 cm, it seems that the system performed as expected for the first value, and a bit poorly for the second value. However, if the step backwards, which differs significantly in the distance when compared to other stepping directions, is not included, the mean distance becomes 1.89 cm and 2.22 cm. Now both values are even more accurate than expected.

Making use of the velocity and the expected coordinates resulted in polar plots. However, not for every measurement, a polar plot could be constructed. This was due to the limits for velocity and position, which a lot of the time differed significantly. Therefore, only a few valid coordinates were found. However, if more valid coordinates are wanted, the velocity needed to be drastically increased. This was due to the system which especially for the ankle joints had difficulties when trying to identify the exact location of a joint and jumped between different values. That is why using velocity to identify the exact coordinates was not as effective as was thought beforehand.

The results of measurement 2 are odd in comparison to the other measurements. This is because the percentages of missing joints are so high that when comparing the percentages of other measurements with black clothing, the other measurements will result in a lower failure percentage. This seems especially odd when comparing this to the measurements where casual clothing was used. Therefore it seems to be likely that something went wrong when taking this measurement. However, looking at the footage, nothing peculiar can be identified why this measurement could be invalid, aside from that the system almost never identifies the ankle joints.

When executing Protocol 2, the percentages of missing joints of the measurements greatly decreased when compared to the results of Protocol 1. This is probably because the camera was placed back further from the participant. Therefore the field of view increased and with it the percentage of not recognized joints decreased. A flaw in the system which was identified early when doing this study was that the system sometimes identified the participant looking at the wall behind, instead of in reality facing the camera. This resulted in a lot of complications because now the left side became the right side and vice versa. When this happened, it resulted in a lot of times in exchanged coordinates from for example the right and the left ankle. A problem was that the participant needed to look down where to place his feet, instead of knowing exactly where to stand. Therefore, when looking down, the system did not identify his face anymore, so a solution to this problem could be that the participant does not have to look down to prevent this from happening.

This study showed the following about how to make optimal use of the system and a sensor in a home environment. It seems to be that it is most important to have the whole body of the participant inside the field of view of the camera.

Furthermore, enough light seems to be key, because especially with almost no light, the system didn't seem to work. Casual clothing in comparison to black clothing will also improve the working of the system.

However, looking at accuracy, when stepping back, the accuracy of the system decreases significantly. This could be due to a greater distance between the camera and the patient. But this phenomenon by itself does not explain the sudden significant increase in total distance when comparing the coordinates of the system with the measured distance.

To solve a lot of the problems with the system failing to identify or not accurately identifying the joints, a solution could be to make use of retro-reflective markers and place them on the joints. However, the method used now is not more patient-friendly, because no reflective markers have to be applied and no suit needs to be worn, normal clothing suffices. For future work, it would be interesting to look at what the effect would be of a person in the background or someone passing by on the ability of the system to identify the patient. Furthermore, only one participant was used for this study and therefore it would be interesting to see what effect a different physical type of patient will have. This study was done by making use of only one skeleton tracker, but it could be interesting to see if other skeleton trackers would perform better than the used skeleton tracker.

8 Conclusion

This study showed that wearing casual clothing instead of black clothing will increase the accuracy of balance recovery measurements. Furthermore, the use of background items will increase the accuracy of balance recovery measurements for only black clothing. Also, when the distance between the camera and the participant is about 3.3 meters, the accuracy of the balance recovery measurements dramatically decreases.

When wearing black clothing, a bigger percentage of joints were recognized when taking measurements during the night in comparison to taking measures during the day. Wearing casual clothing resulted in a higher percentage of recognized joints when filming during the day in comparison to when filming during the night. When comparing casual clothing with black clothing, in most situations casual clothing resulted in a bigger percentage of recognized joints and therefore performed better. This means that wearing casual clothing in comparison to black clothing will result in more accurate feedback about balance recovery measurements when using a camera-based system.

Using background items resulted in a higher percentage of recognized joints with respect to no background items. Furthermore, filming during night-time when no light was on, resulted in zero joints that were recognized. Only when providing light using a flashlight, the system could identify the participant. Furthermore, using a light directly shining into the lens of the camera resulted in similar results as the same conditions without a light shining directly into the lens.

When looking at the coordinates provided by the system, it seems to be that the system is about as accurate as expected by the other literature, with the exception of stepping to the backwards position. Different ambience conditions resulted in a not significant difference in the distance when comparing the coordinates provided by the system with the measured distance.

A Appendixes

A.1 Overview of percentage not recognized joints

To quantify the results, looking at the percentage of not recognized joints gives an indication of how good a measurement has performed.

Number	Daytime/night- time	Lights on	Light shining in lens	Background items	Type of clothing	Left ankle not recog- nized%	Right ankle not recog-
							nized%
1	Daytime	Yes	х	х	Casual	34.45~%	29.22~%
2	Daytime	Yes	х	х	Black	74.97~%	61.92~%
3	Daytime	Yes	Yes	х	Casual	50.47~%	45.11%
4	Daytime	Yes	Yes	х	Black	61.02~%	50.45~%
5	Daytime	Yes	х	Yes	Casual	43.45~%	33.75~%
6	Daytime	Yes	х	Yes	Black	55.64~%	48.71~%
7	Night-time	Yes	х	х	Casual	36.38~%	28.25~%
8	Night-time	Yes	х	х	Black	48.56~%	40.63~%
9	Night-time	х	х	х	Casual	57.26~%	23.86~%
10	Night-time	х	х	х	Black	32.96~%	18.44~%
11	Daytime	х	х	х	Casual	11.65~%	10.12~%
12	Daytime	x	х	х	Black	36.31~%	29.22~%
13	Daytime	Yes	х	х	Casual	16.19~%	13.75~%
14	Daytime	Yes	х	х	Black	32.48~%	25.68~%
15	Daytime	Yes	Yes	х	Casual	16.46~%	12.54~%
16	Daytime	Yes	Yes	х	Black	36.65~%	28.21~%
17	Daytime	Yes	х	Yes	Casual	18.64~%	13.84~%
18	Daytime	Yes	х	Yes	Black	26.46~%	24.21~%
19	Night-time	Yes	х	х	Casual	21.49~%	14.36~%
20	Night-time	Yes	Х	х	Black	23.86~%	26.89~%

Table 3: Overview of all the different measurements. Measurements 1-10 make use of protocol 1, 11-20 make use of protocol 2.

When looking at table 3 at the columns of the right ankle and left ankle, in almost every case the percentage of missing joints is higher for the left ankle than for the right ankle. This is probably because on the left side of the participant was a wardrobe located and therefore less light in comparison to the right side. With no lights on, the sunlight came in from the right side of the participant and therefore there was also a higher percentage of not recognized joints for the left ankle in comparison to the right ankle.

B Literature search

In this section will be an overview of the terms that were used during the literature search progress.

The table below gives an overview of all the search terms and in B.1 is a description associated to each search number.

Search number	Query	Database	Results	Useful amount	$\begin{array}{c} {\rm Relevant} \\ {\rm (y/n)} \end{array}$	
1	intel realsense d435i skeleton tracking	Google	14.000	Too much	У	
2	lighting AND background AND skeleton tracking	Scopus	4	Not enough	n	
3	lighting AND background AND image quality	Scopus	340	Too much	у	
4	lighting AND background AND camera	Scopus	876	Too much	v	
5	lighting AND background AND camera AND tracking	Scopus	181	Too much	v	
6	effect of background and lighting on skeleton tracking	Google	8.490.000	Too much	v	
7	(intel AND skeleton AND tracking)	Scopus	5	sufficient	J V	
8	(intel AND realsense) AND (image AND noise)	Scopus	6	sufficient	y y	
9	(data AND collection) AND processing AND (image AND quality) AND tracking	Scopus	151	Too much	У	
10	effect image noise on skeleton tracking	Google	13.800.000	Too much	У	
11	("Effect of background" OR "environmental noise" OR "ligthing" OR "Shadows") AND ("Kinect" OR "Real Sensor" OR "Camera based sensor" OR "Camera")	Scopus	17.540	Too much	У	
12	Accuracy and resolution of kinect depth data for indoor mapping applications	Scopus	4	Good amount	У	
13	("Effect of background" OR "environmental noise") AND ("Kinect" OR "Real Sensor" OR "Camera based sensor" OR "Camera")	Scopus	662	Too much	n	
14	TITLE-ABS-KEY ("Effect of background" OR "environmental noise") AND TITLE-ABS-KEY ("Kinect" OR "Real Sensor" OR "Camera based sensor"	Scopus	119	Enough	n	
15	OR "Camera") TITLE-ABS-KEY("Effect of background" OR "environmental noise" OR "ligthing" OR "Shadows") AND TITLE-ABS-KEY("Kinect" OR "Real Sensor" OR "Camera based sensor" OR "Camera") AND TITLE ABS KEY("poice")	Scopus	368	Too much	n	
16	("Effect of background" OR "environmental noise" OR "ligthing") AND ("Real Sensor" OR "Camera based sensor" OR "Camera")	Scopus	643	Too much	n	
17	("effect") AND ("background" OR "environmental noise" OR "ligth*" OR "shadow") AND ("kinect" OR "Real Sensor" OR "Camera based sensor" OR "Camera") AND ("balance recovery")	Scopus	41	Good amount	у	
18	("effect") AND ("background" OR "environmental noise") AND ("kinect" OR "Real Sensor" OR "Camera based sensor" OR "Camera" OR "image quality") AND ("object recognition")	Scopus	869	Too much	n	
19	(skeleton AND tracking) AND (("cubemos") OR ("nuittrack") OR ("openpose"))	Scopus	169	Bit too much	У	
20	("low ligthing" OR "environmental noise") AND ("kinect" OR "Real Sensor" OR "Camera based sensor"	Scopus	442	Too much	у	
21	(accuracy) AND (skeleton AND tracking) AND (rehabilitation) 32	Scopus	765	Too much	у	
22	((accuracy) AND (skeleton AND tracking) AND (rehabilitation) AND (protocol))	Scopus	116	Enough	У	

B.1 Useful information

1. Some general info about the skeleton tracking system and an example of how it works.

2. Only Ebooks which were not accessible and not really relevant information.

3. Not exactly the information I am looking for.

4. Too much results, not exactly the information I am looking for.

5. Good results, but they are mostly technical about tracking and how the system works, instead of some more information about lighting and background.

6. Practical information about the skeletal tracking and settings of the program and camera on the following website, although it uses a different skeleton tracking system. Found it on the following website:

https://community.nuitrack.com/t/skeletal-tracking-and-depth/2043.

7. The first article had some interesting information about tracking and it compared the first and the second edition of the Xbox Kinect.

8. In the article: Analysis and Noise Modeling of the Intel RealSense D435 for Mobile Robots is described how errors from the intel realsense d435 are measured in depth.

9. Mostly information about tracking for supervision systems.

10. The information from the articles I found are very useful, some articles about noise measurements using the Xbox Kinect sensor and an other article about the camera which will be used for this project.

11. Mixed results, one article, "Kinect Shadow Detection and Classification", is a very useful article with information about shadow and the consequences for the Kinect. Not only has this article useful information, but it also refers to other similar articles.

12. Specific search because of search 11 where I found this article. Not only did I find the specific article, but also a similar article, which also could contain useful information. 17. This search concluded in a lot of results which are almost all about this particular problem

19. I tried to search for information about a comparison between nuitrack, openpose and cubemos, but found some other interesting articles about using a thermal camera and some overall information about OpenPose.

21. This search resulted in a lot of articles which are very close to what I want to investigate and are very useful.

22. Some useful information, but as with previous tries, most information was either too in debt or too superficial.

References

- João Abreu et al. "Assessment of Microsoft Kinect in the Monitoring and Rehabilitation of Stroke Patients". In: Advances in Intelligent Systems and Computing 570 (2017), pp. 167–174. DOI: 10.1007/978-3-319-56538-5_18. URL: https://link-springer-com.ezproxy2.utwente. nl/chapter/10.1007/978-3-319-56538-5_18.
- Min Sung Ahn et al. "Analysis and Noise Modeling of the Intel RealSense D435 for Mobile Robots". In: 2019 16th International Conference on Ubiquitous Robots, UR 2019. Institute of Electrical and Electronics Engineers Inc., June 2019, pp. 707–711. ISBN: 9781728132327. DOI: 10.1109/URAI. 2019.8768489.
- [3] Nor Asilah Saidin and S. A. Abdul Shukor. "An Analysis of Kinect-Based Human Fall Detection System". In: Proceeding - 2020 IEEE 8th Conference on Systems, Process and Control, ICSPC 2020 (Dec. 2020), pp. 220– 224. DOI: 10.1109/ICSPC50992.2020.9305797.
- [4] Kan Chen and Qiong Wang. "Human posture recognition based on skeleton data". In: Proceedings of 2015 IEEE International Conference on Progress in Informatics and Computing, PIC 2015. Institute of Electrical and Electronics Engineers Inc., June 2016, pp. 618–622. ISBN: 9781467380867. DOI: 10.1109/PIC.2015.7489922.
- [5] Robert G. Cumming et al. "Prospective Study of the Impact of Fear of Falling on Activities of Daily Living, SF-36 Scores, and Nursing Home Admission". In: *The Journals of Gerontology: Series A* 55.5 (May 2000), pp. M299-M305. ISSN: 1079-5006. DOI: 10.1093/GERONA/55.5.M299. URL: https://academic-oup-com.ezproxy2.utwente.nl/biomedgerontology/ article/55/5/M299/2948114.
- [6] Robbie R. Gosine, Harish Damodaran, and Judith E. Deutsch. "Formative evaluation and preliminary validation of kinect open source stepping game". In: *International Conference on Virtual Rehabilitation, ICVR.* Institute of Electrical and Electronics Engineers Inc., Dec. 2015, pp. 92–99. ISBN: 9781479989843. DOI: 10.1109/ICVR.2015.7358593.
- [7] HEROES: Home-based ExeRgaming fOr Enhancing resistance to falls after Stroke - ZonMw. URL: https://www.zonmw.nl/nl/over-zonmw/ehealth-en-ict-in-de-zorg/programmas/project-detail/imdi/ heroes-home-based-exergaming-for-enhancing-resistance-tofalls-after-stroke/ (visited on 07/21/2021).
- [8] Feifei Huo et al. "Markerless human motion capture and pose recognition". In: 2009 10th International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS 2009. 2009, pp. 13–16. ISBN: 9781424436101. DOI: 10.1109/WIAMIS.2009.5031420.

- Kourosh Khoshelham and Sander Oude Elberink. "Accuracy and resolution of kinect depth data for indoor mapping applications". In: Sensors 12.2 (Feb. 2012), pp. 1437–1454. ISSN: 14248220. DOI: 10.3390/s120201437.
- [10] Chung Liang Lai et al. "A Microsoft Kinect-Based Virtual Rehabilitation System to Train Balance Ability for Stroke Patients". In: Proceedings -2015 International Conference on Cyberworlds, CW 2015 (Feb. 2016), pp. 54-60. DOI: 10.1109/CW.2015.44.
- [11] Mark A. Livingston et al. "Performance measurements for the Microsoft Kinect skeleton". In: Institute of Electrical and Electronics Engineers (IEEE), Apr. 2012, pp. 119–120. DOI: 10.1109/vr.2012.6180911.
- [12] Dushyant Mehta et al. "VNect: Real-time 3D Human Pose Estimation with a Single RGB Camera". In: vol. 36. 4. July 2017. DOI: 10.1145/ 3072959.3073596. URL: http://gvv.mpi-inf.mpg.de/projects/ VNect/.
- [13] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. "Nonlinear total variation based noise removal algorithms". In: *Physica D: Nonlinear Phenomena* 60.1-4 (Nov. 1992), pp. 259–268. ISSN: 0167-2789. DOI: 10.1016/0167-2789(92)90242-F.
- [14] Md Zia Uddin and Jim Torresen. "A Deep Learning-Based Human Activity Recognition in Darkness". In: 2018 Colour and Visual Computing Symposium, CVCS 2018. Institute of Electrical and Electronics Engineers Inc., Oct. 2018. ISBN: 9781538656457. DOI: 10.1109/CVCS.2018.8496641.
- [15] Cahit Ugur et al. "Characteristics of falling in patients with stroke". In: Journal of Neurology, Neurosurgery & Psychiatry 69.5 (Nov. 2000), pp. 649-651. ISSN: 0022-3050. DOI: 10.1136/JNNP.69.5.649. URL: https: //jnnp-bmj-com.ezproxy2.utwente.nl/content/69/5/649%20https: //jnnp-bmj-com.ezproxy2.utwente.nl/content/69/5/649.abstract.
- [16] Yan Xi Zhang et al. "Gesture acquisition and tracking with kinect under complex background". In: Applied Mechanics and Materials 511-512 (2014), pp. 541–544. DOI: 10.4028/WWW.SCIENTIFIC.NET/AMM.511-512.541.
- [17] Zhengyou Zhang. Microsoft kinect sensor and its effect. 2012. DOI: 10. 1109/MMUL.2012.24.