BACHELOR

Labeling markers & filtering noise for modeling flexible objects

Weemaes, Luka

Award date:
2022

Department of Mechanical Engineering
Section of Dynamics and Control

Labeling markers & filtering noise for modeling flexible objects

BSc. Graduation Project

Weemaes, L.
ID. 1371940

Project supervisors: Ir. Maarten Jongeneel
Dr. Ir. Alessandro Saccon

Eindhoven, Thursday 3rd February, 2022

Contents

1 Introduction

In the introduction the content of the research is presented, followed by a literature study on the subject.

1.1 Introduction to the problem

Over the past few years, the e-commerce market has grown and with the current state of the economy, it is only expected to grow even further in Europe [1]. The potential workforce cannot keep up with demands from the industry. Warehousing is an important factor in the e-commerce market and is expected to grow with this growing market. Productivity needs to increase and automation will help this growth. For these applications robots are used since they can reproduce a series of task identically to each other. However, to reach the desired productivity current robots are either too slow or are not able to deal with a variety of packages. A solution to both of these problems is instead of robots picking up and placing down parcels, the parcels will be tossed. This is exactly the goal of the I.AM. H2020 EU Project [2]. Together with Vanderlande [3] the goal of this project is to decrease the cycle time of pick-and-put operation by 10% [4]. In this project Vanderlande's Smart Item Robot (SIR) [5], shown in Figure 1.1, is used to experiment with.

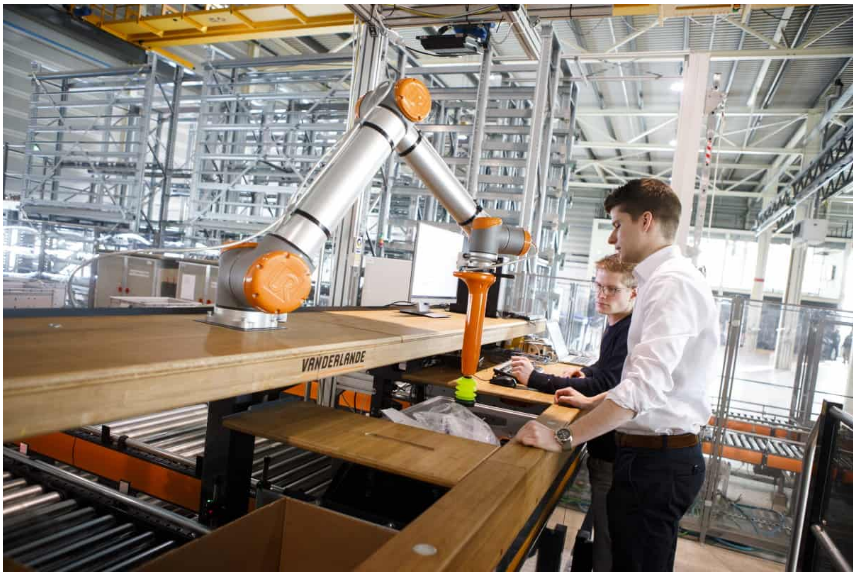

Figure 1.1: Vanderlande's Smart Item Robot (SIR) [5]

Currently, research is done on flexible objects, like bags of clothing, because rather than rigid bodies flexible objects can slightly change in shape. This deformation makes it harder for robots to predict the outcome after a toss. Dynamical models of this motion are made to predict the outcome better. These models are validated using real-life experiments with the actual objects. The motion of the objects is tracked using motion capture cameras. To capture the movement of these objects markers are placed on the edges of them. The markers are small stickers that the motion capture cameras can track. Individual tracking of markers placed on the flexible objects will help understand the motion and validate and improve dynamic models. However, validation is not yet possible since occlusions of the markers appear during the experiments. These occlusions appear because markers are covered during the experiments due to the shape of the bag or the angle with respect to the motion capture cameras. Besides spurious markers are also detected, these are markers that are seen by the motion capture cameras but are not on the object. Because of reflections of light on the flexible bag these spurious markers also appear in the data. This results in a lot of manual assigning the data point to the correct markers, which is time-consuming when doing multiple experiments.

The research will focus on reducing this post-processing time. First a literature research will be done, to find out more information about motion capture and the problems it still has. With this in mind a problem definition can be proposed from with the research goal of this project is determined. The data that is used for this project will be the data that is needed to model flexible objects. Experiments will be performed with flexible objects. The data from the experiment is used in the proposed method in this research which will be explained later on. The results of the method are shown and discussed at the end of the report.

1.1.1 Terminology

The terms used to describe various processes differ widely. What follows are brief definitions of terms as they are used in this paper:

- *Markers*: Dots on the object that are tracked by a motion capture system. A set of markers placed on an object makes it possible to record the motion of the object.
- *Labeling*: Assignment of the anonymous 3D positions to the labels of the defined markers. Manual labeling is usual done for an arbitrary instant of time.
- *Accuracy*: A quantity to indicate the amount of markers which are labeled correctly. It is always shown as a percentage between 0 and 100%. When all the markers are labeled correctly, all the markers correspond with the actual object and no data points are skipped, the accuracy is 100%.
- *Occlusion*: During experiments markers on objects can be covered, which make them disappear out of the data-set. This is called an occlusion.
- *Spurious markers*: data-points in the data-set that do not belong to any of the markers of the object. These spurious markers appear in the data-set due to reflections of light and are seen as markers by the motion capture system.

1.2 Literature review

In the field of motion capture already much research has been done. Different setups are used and various experiments are executed to test and validate different methods of cleaning up the motion capture data.

In [6] the authors addressed the problem of motion capture, which is the labeling afterwards, as a time consuming and labor-intensive post-processing part. Motion capture is used to track human body motion, by placing marker on a motion capture suit, which is a body suit with markers on it. In this research, the problem of consistently labeling the markers in spite of temporary occlusions and appearance of spurious markers is tackled. Different experiments were executed, in the experiments with a lot of occlusion the model has an accuracy of 94.9%. In experiments with no occlusion at all an accuracy of 98.7% is reached. In this research comparisons were made with different available methods, they will be discuss next. In [7] an auto labeling approach specifically designed for hands is proposed. The model presented uses an inverse kinematics approach, this means the actual shape of the object is used to assign data-point to markers. With the use of the shape of the object various orientations of the object could be omitted, which makes the labeling easier. In the research all the data points which could be assigned to the correct hand marker. In [6] this method was used, obtaining an accuracy of 83.2%. The methods of [6] and [7] are the same, the model in [7] is tuned to be used on hand motion, while the model in [6] is used on human body motion. The method proposed in [8] also uses human motion to label data-points to the corresponding markers. In [6] this method is also used, the accuracy achieved in this case is 88.2%. In [6], [9] was presented as an alternative technique of the labeling problem on hands presented in [7]. In [9] this method achieved an accuracy of 93.7%. However, in the experiments a lot of motion capture cameras were used to track the hand, so the accuracy may be lower when less cameras are available.
All these models are made to label the data-points the correct markers. Comparing these methods with each other, as done in [6], the method with the highest accuracy can be found. However, none of these methods is experimented on flexible objects. The shape of human hands and human bodies is more complicate than the shape of a simple bag, this will make labeling markers on a bag easier than labeling markers on a human body.
Therefore, the actual post-processing model for the bag data can be less complicate than the model for a human body.

In [10] an algorithm is presented that labels markers. The algorithm can label different markers to a defined body, as well as fill in gaps in the data-set. The model combined with the Hungarian algorithm [11] is used to find the global optimal matching marker. When comparing this model with other models, as done in the research, [10] obtained a very high accuracy compared to other available models on the market, it was almost 100% accurate in labeling and filling in gaps in the particular data. In [12] a model about data reconstruction was published. The algorithm presented can fill in missing values in data-sets of the human body. The reconstruction was done using principal component analysis (PCA) [13], combined with a weighting procedure. The model was able to fill in data-sets missing 70% of all the time frames. The algorithm was able to fill in the entire trajectory of the markers with an error of less than 1.5 cm of the original marker even if the last half of the trajectory was completely missing. In [14] an automated method is proposed. The method allows for gap-filling on a data-set that consists of human body motion. The algorithm uses inverse kinematics to fill the rows of data for each frame of

time. An 80% reduction of completion time is observed compared to manual gap-filling. Also a higher accuracy was observed compared to manual gap-filling, The mean of the error in the inverse kinematics decreased by 21%.
All these researches present algorithms which can label markers of objects and also fill in empty data-sets. This feature of the algorithm is useful if the actual data consists of many missing data-points. The necessity of this feature can be determined after experiments are performed and the data is analyzed. For now this research will focus on only labeling the data=points that are present in the data and not filling in the missing values, but it is a nice option for further research.

In [15] a rotation matrix is derived. For this rotation matrix only one angle θ is needed, instead of three angles α, β and γ. This matrix is used in the method that will be described in Subsection 4.1.

2 Problem definition, research goal & constrains

In this section, the problem statement and corresponding research goal are explained. The following sections of the report will try to fulfill this goal and show some good results.

2.1 Problem definition & research goal

The experiments were done at Vanderlande's Innovation Lab at the TU/e campus. With the available setup used at the lab at the TU/e campus different types of objects can be tracked. On these objects markers are placed so the motion capture system can track them. Different types of markers are shown in Figure 2.1. Figure 2.1.a shows a rigid body with spherical pins as markers, whereas Figure 2.1.b shows a flexible object with round stickers as markers. Both these objects can be tracked by the motion capture system. For the rigid body an algorithm can be set up in the software environment Motive [16], [17], which is used to convert the images of the motion capture cameras into a 3D spatial view. This means that the software knows this object is present when a certain set of markers is detected, which makes tracking of individual markers possible. Also, if a marker disappears for any reason, the software is able to label the new appearing data-point to the correct marker because of the shape of the object that stays the same. For the flexible objects however, such an algorithm is not incorporated in the software since the distance between the markers will vary as the object moves. This means if a marker disappears the new appearing data-point is labeled as a different marker. In order to track individual markers manual labeling is needed, which is very time-consuming and also very sensitive to errors. The main problem is given by:

Labeling software is not able to label the data-points to the correct markers on a flexible object.

From the literature review in Subsection 1.2 the lowest accuracy of the presented methods is 83.2%. This means that if the accuracy of my model is the same or higher, it can be used for labeling markers on objects. With this desired accuracy the research goal for this project can be described by:

Create a method that automatically labels the data-points to markers on a flexible object, which makes tracking of individual markers possible. The method should have an accuracy of at least 83.2%.

(a) Rigid body with 3D markers

(b) flexible object with 2D markers

Figure 2.1: The different types of objects

2.2 Constrains for the experiments

In the following section, the experiments will be performed. In order to extract useful data from these experiments, constrains are established. If the experiments are executed according to the constrains, the data derived from the experiments will be useful to validate the model.

- The markers on the bag will not move outside the view of all the cameras. All markers will be present and within the view of at least two cameras during the entire experiment. When a marker is seen by less than two cameras the motion capture system is no longer able to compute the location of the bag since it requires at least two cameras to create a 3D spatial view of the bag.

- The markers on the bag will be visible at some point during the experiment. The cameras should be able to track all the markers of the bag. Markers can disappear because of occlusions or a change of angle of the bag so the reflection does not reach the camera, but at some point in time, a marker should be visible for at least two cameras. As explained before at least two cameras are required for a 3D spatial view.

- The bag will be held in the same way as during transportation with the robotic arm. The robot will hold the bag in the middle with its suction cup. This means that during the experiments the bag will be hold with one hand in the middle of the bag. The robot basically swings from one point to another, so in the experiments also a swinging motion will be used to derive data. In this way, realistic experiments will be conducted and reliable results will be extracted from the experiments.

3 Available experimental setup

In this section, the available experimental setup that is used to perform experiment is explained. The experiments that are performed to extract the data which is used in the remainder of this report will be shown and

3.1 Setup

The experiments were done at Vanderlande's Innovation Lab at the TU/e campus, as explained in Section 2. A picture of the available setup in the lab is shown in Figure 3.1. The setup consists of a conveyor belt with two robots placed next to it. Above the robots and the conveyor belt motion capture cameras are mounted to the wall or to some other tubing frame. These cameras are strategically placed so that together they can capture the entire workspace of the robots.

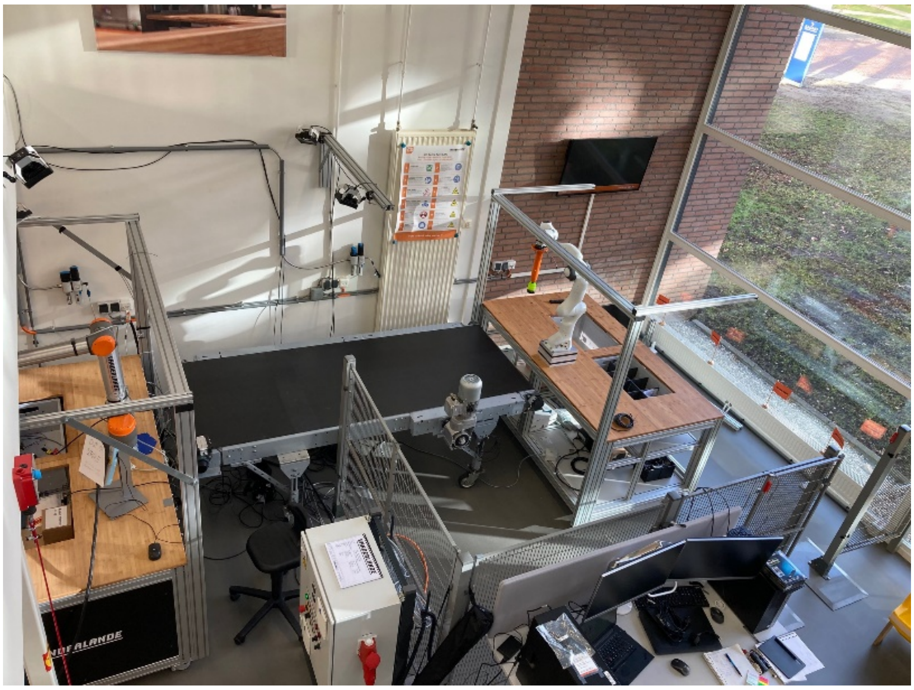

Figure 3.1: Overview of the setup in the lab

The cameras that are used are Optitrack [18] motion cameras. In total there are six cameras in the lab which detect and track objects. Two types of cameras are used, five *Prime* 17*W* cameras [19] are used in the setup as well as one *Prime*^{x} 22 camera [20].

3.2 Conducting the experiments

The experiments use the flexible bag shown in Figure 2.1.b The bag is placed on the conveyor belt so that all the markers are seen by the cameras, this can be checked in the Motive environment [16], which is the software used to read the images from the cameras and convert it to a 3D spatial view. To start the experiment the bag is picked up and held in the middle, just like the robotic arm would hold it. Thereafter the bag is swung around in a circular path, this motion is shown in Figure 3.2. After a few swings, the bag is tossed on the conveyor belt and the experiment is completed. The swings of the object during the experiment represent the motion of the flexible object when it is hold during real-life transportation with the robotic arm, this is why it is chosen to use this motion in the experiments. In this way the experiments represent the motion of the flexible objects as good as possible.

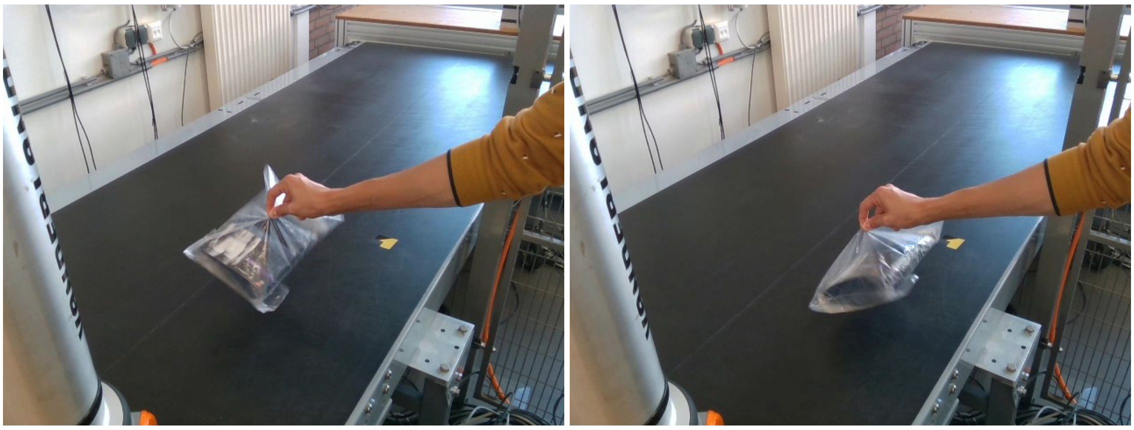

Figure 3.2: These figures show the motion of the flexible object during the experiment

3.3 The acquired data

The acquired data is shown as snapshots from during the experiments in Figure 3.3. As can be seen in this figures, sometimes all the five markers are present. Sometimes a few markers are covered and cannot be seen in the figures, these are the occlusions. Also at other points in time more than five markers can be observed, the extra markers are the spurious markers and do not belong to any of the five markers of the object.

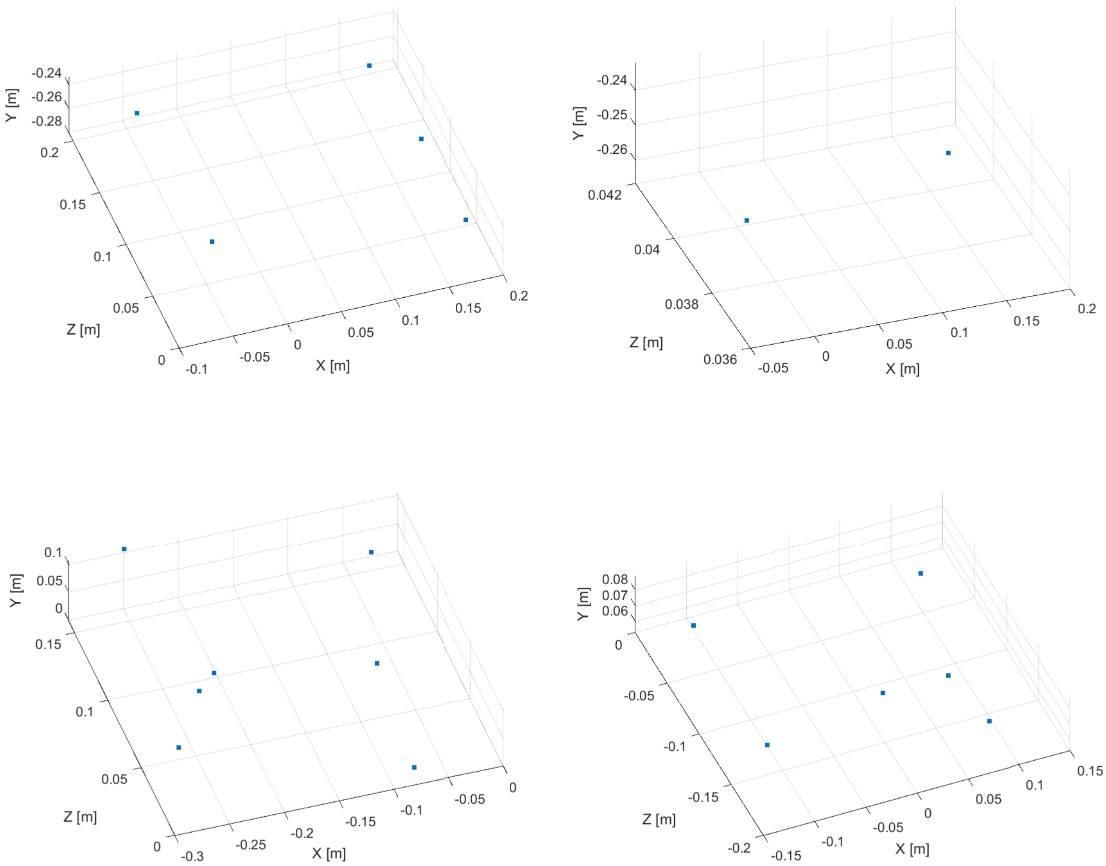

Figure 3.3: Various snapshots from the experiment

4 Method and results

4.1 Method

After the experiments are executed, as presented in Section 3, the data can be post-processed. This chapter will give an explanation of which steps are taken and the mathematics used in it. In Subsection 4.2 the results of this model are shown.

4.1.1 Data Preprocessing

The bag is represented as seen in Figure 4.1. The markers are assigned randomly to the markers the bag on appearance in the data-set. as

$$M_i,$$

where i is the corresponding marker.

$$\text{For } i \in N,$$

where N is the amount of markers, the markers are given by

$$M_1, M_2, M_3, M_4, M_5.$$

At the start of the algorithm, the centroid of the data points for each frame is calculated and subtracted from the data to have the data centred around the origin, this makes the markers invariant to translation. The next step is to rotate the data points for each time frame. Due to the rotation, the frame of reference is changed from an observational reference frame to a marker position reference frame. With the use of this marker position frame, the data points will be viewed as the observer is traveling with the bag. From this reference frame, it will be easier to distinguish between the different markers of the bag.

Figure 4.1: Representation of the flexible object with markers as shown in Figure 2.1.b

4.1.2 Shifting the reference frame

To determine this marker position reference frame a couple of steps are needed. The approximated orientation of the markers in the reference frame at each time-step needs to be calculated to rotate all the data-points in a certain direction. This is achieved by using two points of the bag and calculating the direction from one point to another as

$$V_{1\to 2} = V_{M_2} - V_{M_1}. \tag{4.1}$$

This vector is used in the following equation to calculate angle θ between $V_{1\to 2}$ and the x-axis, V_x as

$$\cos(\theta) = \frac{V_{1\to 2} \cdot V_x}{|V_{1\to 2}||V_x|}. \tag{4.2}$$

With the angle and axis of rotation known, the rotation matrix is derived. This matrix, derived in [15], is defined as a matrix of rotation R by angle θ around the axis $\mathbf{u} = (u_x, u_y, u_z)$ with a unit vector $u_x^2 + u_y^2 + u_z^2 = 1$. It is given in the following equation by

$$R = \begin{bmatrix} \cos\theta + u_x^2(1-\cos\theta) & u_x u_y(1-\cos\theta) - u_z\sin\theta & u_x u_z(1-\cos\theta) + u_y\sin\theta \\ u_y u_x(1-\cos\theta) + u_z\sin\theta & \cos\theta + u_y^2(1-\cos\theta) & u_y u_z(1-\cos\theta) - u_x\sin\theta \\ u_z u_x(1-\cos\theta) - u_y\sin\theta & u_z u_y(1-\cos\theta) + u_x\sin\theta & \cos\theta + u_z^2(1-\cos\theta) \end{bmatrix}. \tag{4.3}$$

This results in the plot shown in Figure 4.2, in which the data is presented in a marker position reference frame. The data is all plotted in the same color, because it has not yet been labeled to the corresponding marker. The markers of the bag can be seen as lines and curves of data points because the bag has internal degrees of freedom the markers will move with respect to each other. Another observation that can be done is the amount of noise present in this data. Most of the noise is located in the middle of the bag. This is the noise that is expected beforehand since the reflections from the lights in the lab on the bag are often seen by the cameras as markers. However, in the figure, also marker outside the surface of the bag are shown. These spurious markers could be reflections from the light on the conveyor belt or the arm of the person swinging the bag, like a watch or a ring.

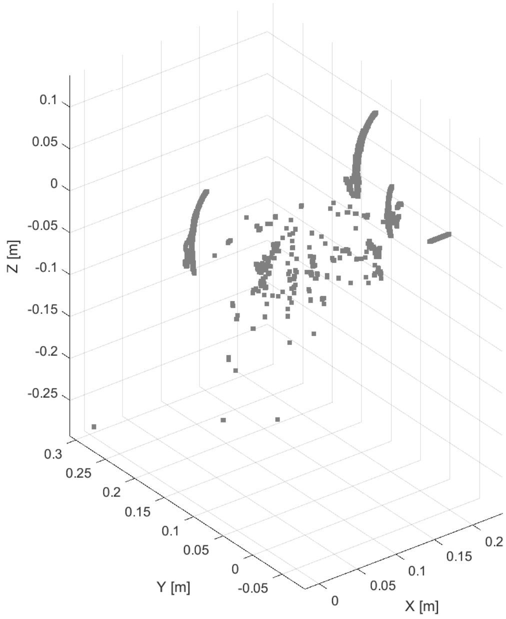

Figure 4.2: Data shifted to a marker position reference frame

4.1.3 Labeling the data points

From Figure 4.2, the markers of the bag can be distinguished from the noise, but to do this by hand is still a time-consuming process. The data points will be assigned to the corresponding markers using an allowed deviation from an arbitrary point, this approach is visualized in In Figure 4.3. In this case, the first point is chosen as a reference point since in this pose the markers all lie approximately in the same plane and the bag is in its original position.

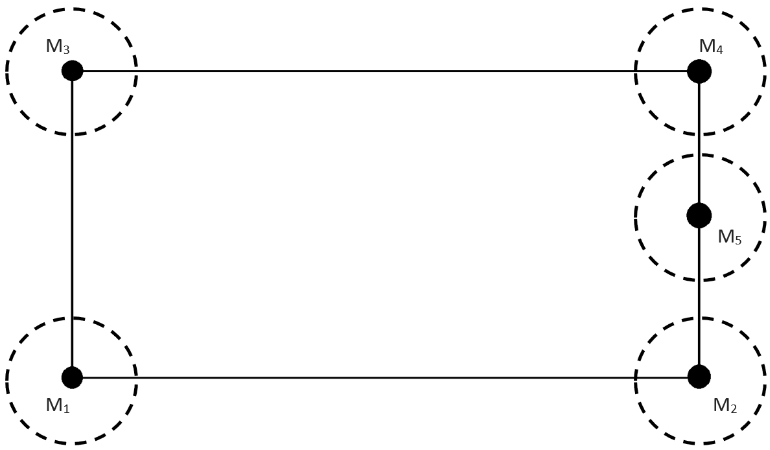

Figure 4.3: Representation of the flexible object with markers as shown in Figure 2.1.b with the allowed areas where markers can be

The allowed deviation factor is chosen by visual inspection of Figure 4.2. After trial and error with different values for the allowed deviation, the correct data points are assigned to the markers. All the noise that is present within the allowed deviation will also be assigned to the marker, however, as can be seen in Figure 4.2 most of the noise data points are located in the middle of the bag.

4.2 Results

In the previous subsection, the method of post-processing is discussed. As a result the data is filtered and stored into different marker data sets. In this section, the results of the model will be shown and discussed.

After the data set is post-processed using the method described in Subsection 4.1 the result can be plotted. In **??** the result of the post-processing of one of the data sets is shown.

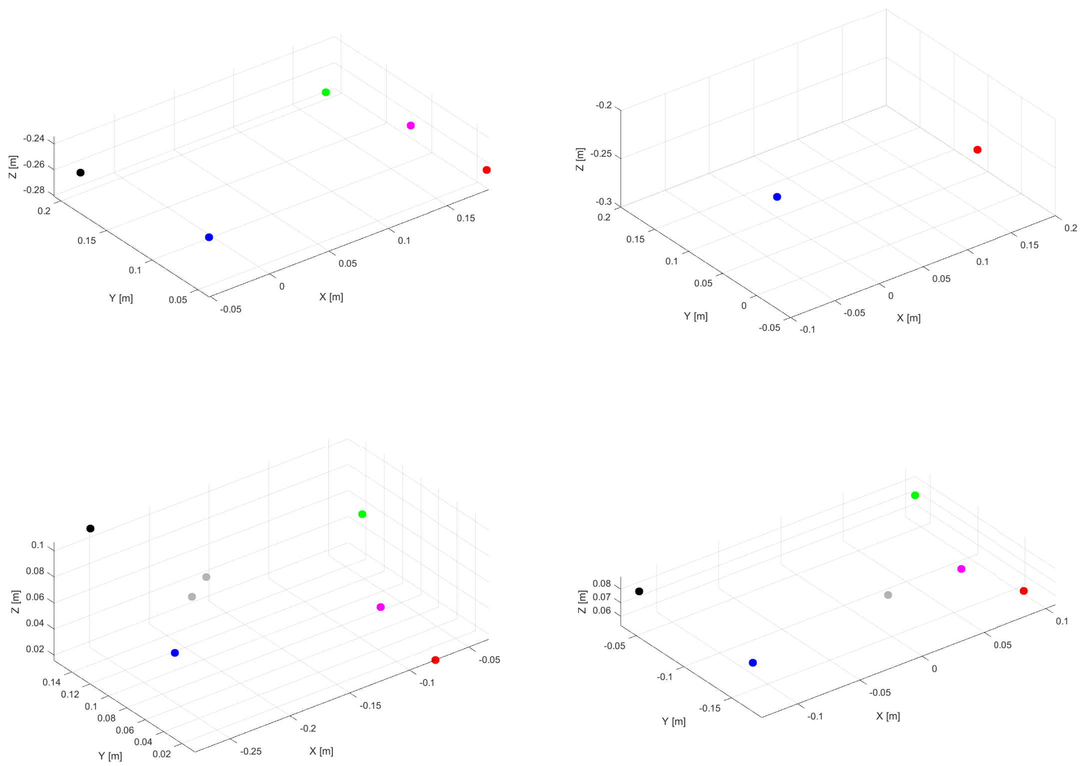

Figure 4.4: Various snapshots from the experiments after post-processing

In these figure shown in Figure 4.4 the markers as colored with each a different color, representing the corresponding marker. The noise in the plots is colored light gray. The same snapshot location are taken as in Figure 3.3, to show the same markers, but now they are either labeled or discarded as noise.

To evaluate the accuracy of the method manual labeling is done to prove accuracy. In Table 4.1 the evaluation of the three different experiments is shown.

Table 4.1: Accuracy of the model in different experiments

Experiment	*Time frames*	*Total unlabeled markers*	*Accuracy*
Swing 1	5703	98	96.5%
Swing 2	4047	53	99.7%
Swing 3	5390	153	91.5%
Average	**5047**	**101**	**95.9%**

In Table 4.1 the amount of time frames, total unlabeled markers and accuracy is shown. From these values you can clearly see a relation between them. The longer the experiment takes, the lower the accuracy is. Also the total unlabeled markers are affecting the accuracy, the higher the number, the lower the accuracy. Averaged over three experiments in which all the same constrains were taken into account, the accuracy of this method is 95.9%.

5 Discussion & conclusion

After the results are shown in the previous section they will be discuss in this section, also a conclusion will be determine if the research goal is reached.

5.1 Discussion

The method created uses a simplified version of the principal component algorithm (PCA), if the real version is implemented correctly the accuracy might be even higher than right now. However, with an accuracy of 95.9%, the model is very good at labeling the individual data points to the correct markers. Besides, the computing time for this simplified version of PCA is only a few seconds using MATLAB.

The experiments used in this research take around ten to fifteen seconds, at which time approximately 5000 data frames are created. The time it will take for a robot to pick up an object and toss it to its destination will not take more time than this, so the collected data seem to be a valid representation of the real application.

Whenever the object that is tracked makes some unexpected movements instead of smooth a transition from A to B, the cameras will likely lose track of the markers on the object. An improvement on this tracking could be made by adding more cameras to the setup. Obviously more cameras will contribute to easier tacking the object. Moreover, the method presented in this research will still be sufficient when more unexpected movement will occur with more cameras added to the current setup.

Another suggestion would be to add more markers to the object so the cameras have more markers to track and will possibly keep track of some during unexpected movement. However tests with an object shown in Figure 5.1 showed that the cameras are not able to track markers whenever they are close to each other, more research could be done on the exact distance markers can be from each other until they are not distinguishable anymore for the cameras.

Figure 5.1: Sheet of markers

5.2 Conclusion

With an accuracy of 95.9%, the research goal of an accuracy 83.2% is reached with quite some room for margin in the method. With the simplified PCA computation, motion capture cameras can track a flexible object like a bag presented in this research. The model allows to filter noise from the acquired data and consequently allows to track individual markers placed on the flexible object.

Improvements could be made on making the model more usable in different situation of motion capture. In the current data-sets, always two nearly complete marker-sets were recorded by the motion capture system, the method is based on this feature. When only one or none full marker-sets are present, the model might need some improvement. Also future work could be done on filling in the missing values in the data-sets. When this is done the entire motion of the object can be determined, not only the movement of some individual markers.

References

[1] S. ; Lone, N. ; Harboul, and J. W. J. Weltevreden, "European E-commerce Report," Amsterdam University of Applied Sciences, Amsterdam, Tech. Rep., 2021, p. 111.

[2] *I AM Project.* [Online]. Available: `https://i-am-project.eu/`.

[3] *Vanderlande.* [Online]. Available: `https://www.vanderlande.com/`.

[4] *Future packing robots exploit collisions, instead of fearing them*, Mar. 2020. [Online]. Available: `https://www.cursor.tue.nl/en/news/2020/maart/week-1/future-packing-robots-exploit-collisions-instead-of-fearing-them/`.

[5] *Smart Item Robotics (SIR) - Material Handling 24/7.* [Online]. Available: `https://www.materialhandling247.com/product/innovative_order_and_batch_picking_systems`.

[6] S. Ghorbani, A. Etemad, and N. F. Troje, "Auto-labelling of Markers in Optical Motion Capture by Permutation Learning," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11542 LNCS, pp. 167–178, 2019. DOI: `10.1007/978-3-030-22514-8{\_}14`.

[7] J. Maycock, T. Röhlig, M. Schröder, M. Botsch, and H. Ritter, "Fully automatic optical motion tracking using an inverse kinematics approach," *IEEE-RAS International Conference on Humanoid Robots*, vol. 2015-Decem, pp. 461–466, Dec. 2015. DOI: `10.1109/HUMANOIDS.2015.7363590`.

[8] S. Holzreiter, "Autolabeling 3D tracks using neural networks," *Clinical Biomechanics*, vol. 20, no. 1, pp. 1–8, 2005, ISSN: 02680033. DOI: `10.1016/J.CLINBIOMECH.2004.04.006`.

[9] S. Han, B. Liu, R. Wang, Y. Ye, C. D. Twigg, and K. Kin, "Online optical marker-based hand tracking with deep labels," *ACM Transactions on Graphics*, vol. 37, no. 4, 2018, ISSN: 15577368. DOI: `10.1145/3197517.3201399`.

[10] S. Xia, L. Su, X. Fei, and H. Wang, "Toward accurate real-time marker labeling for live optical motion capture," *The Visual Computer : International Journal of Computer Graphics*, vol. 33, no. 6-8, pp. 993–1003, Jun. 2017, ISSN: 0178-2789. DOI: `10.1007/s00371-017-1400-y`.

[11] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, vol. 2, no. 1-2, pp. 83–97, Mar. 1955, ISSN: 00281441. DOI: `10.1002/nav.3800020109`.

[12] Ø. Gløersen and P. Federolf, "Predicting missing marker trajectories in human motion data using marker intercorrelations," *PLoS ONE*, vol. 11, no. 3, Mar. 2016. DOI: `10.1371/JOURNAL.PONE.0152616`.

[13] Pearson K., "Pearson, K. 1901. On lines and planes of closest fit to systems of points in space.," *Philosophical Magazine*, vol. 2, pp. 559–572, 1901.

[14] J. Camargo, A. Ramanathan, N. Csomay-Shanklin, and A. Young, "Automated gap-filling for marker-based biomechanical motion capture data," *Computer Methods in Biomechanics and Biomedical Engineering*, pp. 1180–1189, 2020. DOI: `10.1080/10255842.2020.1789971`.

[15] C. J. Taylor and D. J. Kriegman, "Minimization on the Lie Group SO(3) and Related Manifolds Minimization on the Lie Group SO(3) and Related Manifolds 1," Yale University, Tech. Rep., 1994, p. 8.

[16] *OptiTrack - Motive - In Depth.* [Online]. Available: `https://www.optitrack.com/software/motive/`.

[17] *OptiTrack Documentation Wiki - NaturalPoint Product Documentation Ver 2.2.* [Online]. Available: https://v22.wiki.optitrack.com/index.php?title=OptiTrack_Documentation_Wiki.

[18] *OptiTrack - Motion Capture Systems.* [Online]. Available: https://optitrack.com/.

[19] *OptiTrack - Prime 17W - Wide angle coverage for large volumes in smaller spaces.* [Online]. Available: https://optitrack.com/cameras/prime-17w/.

[20] *OptiTrack - Primex 22 - In Depth.* [Online]. Available: https://optitrack.com/cameras/primex-22/.