Control Model for Productivity and Safety in Human-Robot Collaborative Systems

Granata I.*, Faccio M.*, Boschetti G.*

**Department of Management and Engineering, University of Padova, Stradella San Nicola 3,*
36100 Vicenza – Italy (irene.granata@phd.unipd.it, maurizio.faccio@unipd.it,
giovanni.boschetti@unipd.it)

Abstract: Nowadays, there is a substantial increase in the use of automation, to be able to produce more and gain a wider share of the market. This is part of the fourth industrial revolution that is based on the integration of increasingly more flexible systems and the Internet of Things. Among this wide set, collaborative robots (Cobots) represent one of the technologies that modern production systems are trying to integrate. This paper focuses on the impact that these technologies have on different levels within a productive plant and on the improvement of the collaborative experience. At workstation level, the control methodologies are investigated and developed: technologies such as computer vision and augmented reality can be applied to aid and guide the activities of the cobot, in order to obtain the following results. The first is a safe workspace where collisions can be avoided in real time. The second is an increase of overall productivity generated by the reduction of idle times and safety stops. This can be achieved using real time multi-camera systems and skeleton tracking to constantly know where the operator is in the work cell. The system will offer the possibility of directing feedback based on the discrepancies between the physical world and the virtual models so the technology can be applied to sectors that require a constant process control. In this way, human operator and cobot are not merely two single resources working in the same cell, but they can achieve a real human robot collaboration. In this paper, it is presented a framework that allows to reach the two aforementioned goals.

Keywords: Cobot, Safety, Modern Production Systems, HRC, Productivity

I. INTRODUCTION

Modern production systems need to be able to answer to the trend of the market that is oriented towards an increasingly mass customization, (Tseng et al. 1996), since we are living in the so called fourth industrial revolution, also called Industry 4.0, where there is a growing integration between physical and digital systems that have changed production methods. This means more flexible systems are required, in order to guarantee high volume but also high variety of products, as required by the fourth industrial revolution (Industry 4.0), Bai et al. (2020). One of the technologies that are recently being integrated are collaborative robots (cobots): moreover, since the 2020, their installation increased by 12%, of Robotics (2020). This is due to the fact that this type of robots doesn’t work separately from the operator but together, and the resource can also do parallel tasks. The aim of this paper is to realize a framework where a real time control algorithm is proposed. This model aims to achieve both high productivity and safety on human-robot collaborative systems, through the implementation of a safe workspace where collisions can be avoided in real time, monitoring both operator’s and cobot positions and comparing them with the ones expected from the virtual model before created.

The paper is organized as follow: Section 2 presents a brief literature review on multi-cameras systems and their functioning, and safety problem in collaborative work cell. Section 3 introduces Motion Capture (MoCap) technologies, while Section 4 is for the framework itself. The proposed setup is here explained along with the dynamic task allocation approach. Lastly, Section 5 concludes the work.

II. LITERATURE REVIEW

In this section the state of the art of the three main fields, helpful for the development of the system, are analysed.

One aspect that has to be considered, in order to cover all the work area, is the number of cameras, or better RGB-D sensors, that are necessary. In Kim et al. (2017), an analysis on the importance of using Microsoft Kinect V1 cameras is made; in particular, it is proposed to use 8 cameras to form a square figure with 2 acquisition tools per side. A system of this type promises to extract a cloud of points, that is generated by P^k coordinate systems where k represents the number of Kinect used. Another issue of fundamental importance is the synchronization of the frames where the cameras can present a lag between the acquisitions of the various sensors, resolvable with a spline interpolation. The fastest camera sends the signal of the captured position, and the positions detect by the other cameras are estimated as interpolation between the previous time and the current time using a system of cubic splines like the one proposed by Hermite or Ferguson.

The framework proposed by Otto et al. (2019) provides the use of a series of cameras to amplify the tracking area. Using a series of Microsoft Kinect V2 cameras, each of these creates a projection cone that widens, according to the distance from the focal plane. The accuracy of the generated point cloud decreases as the subject's distance increases.

On the other side, Ye et al. (2013) propose a calculation pipeline based on the data recorded by a single camera and it compared them with the positions contained in a database.

So far, these are some of the solutions presented for motion capture analysis. Now we can focus on the safety problem in collaborative systems, in order to evaluate if these technologies have been already used to guarantee the safety in the cell.

In collaborative work cell, the resources must share the space and so, they have to work together without interference. This is usually possible because cobots don't require fences, but their absence leads to some concerns.

There are different regulations for cobots safety, in particular ISO/TS 15066 (2016), where the requirements for collaborative cells are described:

- Safety Rated Monitored Stop (SRMS): it stops the cobot if it is too close to the operator.
- Hand Guiding (HG): for manual guide.
- Speed Separation Monitoring (SSM): it maintains a separation between the operator and the cobot.
- Power and Force Limiting (PFL): it limits the force applied by the cobot.

Usually, these securities are not integrated in the cobot, but they have to be implemented.

One solution to implement SSM is offered by Byner et al. (2019), where a laser scanner is introduced into the work area to monitor in real time the distance between the resources: if the cobot and the operator are too close, the first one must reduce its speed to a fixed limit or to zero.

In Galin et al. (2020), the authors achieved the Speed Separation Monitoring with the tasks division: the tasks assigned to one resource have to be sufficiently far in space from the tasks assigned to the other, in this way safety is guaranteed but the total time required to complete the process can increase. Moreover, the work cell should be integrated with other security measures, both hardware or software.

There are also solutions that combine two of these specifications, i.e. Lucci et al. (2020), where a definition of a simple framework that combines SSM and PFL is proposed. This can lead to considerable improvements in productivity while preserving safety criteria; to a reduction of the problem thanks to an optimization algorithm that results in a closed-form solution without any conservative assumptions. Inclusion of the configuration-dependent inertial properties of the robot, which enables a more comprehensive treatment of safety constraints, is proposed.

Lastly, some task allocation solutions to minimize cycle time are investigated, since the aim of the following framework is to propose a safe task allocation, with the makespan minimization, in collaborative systems.

Pearce et al. (2018) proposed an optimization framework to reduce makespan considering the physical strain. In this paper consideration about ergonomics are made in order to improve it. Their focus was more on the time required and the stress induced into the operator than on the safety issue.

Another solution for the reduction of cycle time is presented by Weckenborg et al. (2019), where the use of a cobot improves the productivity of 12%. Their model decides which tasks are assigned to which resource, with a genetic algorithm, considering only robot flexibility and collaboration flexibility.

Heyadaryan et al. (2018) prove that the collaboration can be useful to increase ergonomics and to reduce risk of injury, but the production time can increase. Their work considers safety issue reducing the change of interference between the resources.

Up to now, these three fields seem to be very disconnected from each other. In fact, the solutions for the improvement of the motion capture don't consider its use in collaborative systems and safety and task allocation issues are treated separately.

This paper presents a framework to integrate them all together, proposing a solution that allows the minimization of the makespan but at the same time it guarantees the required safety through the use of motion capture techniques.

III. MOTION CAPTURE TECHNOLOGIES

Based on the necessity of having a real time control of the position of the cobot in the workspace, a system that integrates computer vision and augmented reality can be realized. This system is based on the concept of motion capture (MoCap), that is the “process of recording a live motion event and translating it into usable mathematical terms”, Menache (2000). The definition means that is possible to take as input the real movement of an operator and to obtain, as output, a quantitative mathematical description that can be used as input for manipulation and control systems. Thanks to this, such systems are useful in a variety of industrial applications, for example to analyse forces applied to operators and their postures during work activities, but also to control their positions. By exploiting MoCap technologies, in fact, it is possible to translate human movement into models that can be processed by machines or software, thus allowing man to interact, in real time, with any robots present in the same environment.

A. MoCap Classification

This section is helpful to understand the current technologies for motion capture and to clarify the choices made in the following sections.

Systems used for capture motion can be divided in two main categories: non optical systems and optical systems. The first group included electro-mechanical systems, like wearable tracksuit, Figure 1., where IMU (inertial measurement unit) sensors are installed. On the inside, there are magnetometers, accelerometers, and gyroscopes, to obtain cinematic data of who is wearing it. The data available are linear and angular accelerations, velocities, and positions. The suit, also, contains a control unit that has the task of receiving the information of all the sensors connected to it and to translate them into discrete signals to be sent wirelessly to the computer, where these signals are processed. Non optical systems have some advantages like, Bortolini et. al (2020):

- no occlusion and potentially unlimited capture space;
- real-time visualization without postprocessing;
- capture of multiple subjects.

Some of the disadvantages, instead, are:

- global position can’t be calculated, sensor fusion algorithms must be used; real-time visualization without postprocessing;
- IMU sensors can suffer from drifting in position measurement, for this reason the combination of magnetometers is preferred;
- capture space is limited by wireless connection range;
- IMU are very sensible to electromagnetic disturbances and the data can be scrambled.

The second group included marker-based and markerless systems. Marker-based devices can be passive, where reflective surfaces radiated with infrared light emitted by specific cameras are used. The weakness is the fact that the environment must be light controlled to avoid reflections. There are also active markers, that are infrared LED lights that are displayed by cameras that do the captures. Such systems are more reliable than passive markers because their characteristic of emitting the signal allows them to be used with pulsating light in such a way that the camera, knowing which are active and which are off for each time unit, can exclude some external and potentially misleading signals through algorithms. The greater advantage is that this type of systems releases the person from heavy mechanical structures that greatly limit the freedom of movement and they allow to track wide movements as a person walking.

Fig. 1. Example of IMU suit

Markerless optical systems, Figure 2., are quite new technologies. They allow the reconstruction of the motion from simply processing a captured video file without any object physically connected to the human operator. They include Artificial Intelligence algorithms, Deep Learning and Vision Systems. These lasts are able, depending on the component they "see", to make a recognition in their database, to communicate to a system, for example a cobot, the correct orientation required and then to compare the framed face with the geometric and aesthetic dimensional data concerning that specific object.

Fig. 2. Example of markerless MoCap

IV. DYNAMIC TASK ALLOCATION

A. Architecture Setup

In the workspace there are simultaneously the human operator and a cobot working in the same area, so a system able to provide information on operator’s position can be useful to dynamically define the task allocation.

The motion capture architecture proposed includes *Intel Real Sense D435* cameras, Figure 3., with optical sensors integrated.

These are cameras that use an RGB sensor and two sensors for stereophotogrammetry that can measure the distance of a point from the position of each camera. In a depth frame the various pixels that compose it show the distance of the point from the focal plane. The two information streams (depth and RGB) are then synchronized by the camera software that allows to extract information from both.

The cameras are particularly appreciated for their “low light qualities” and the lightness of the objects (only 71.8g) and for the possibility of automatic calibration.

Fig. 3. Intel RealSense D435

Four cameras are chosen to have the collaborative workspace recorded from different points of view. In this way it is possible to increase the monitored area, to enhance the accuracy and to avoid any occlusions in case there are objects in the recorded trajectory, Faccio et al. (2019).

Motion capture is done by OpenPose library, that is used for body joints position recognition in real time. OpenPose is the first real-time multi-person system to detect human body, hand, facial, and foot keypoints on single images, Cao et al. (2019) and it utilizes bottom-up approach.

OpenPose uses a convolutional neural network system (CNNs) that runs the images and the models that are provided to recognize people in the frame and, for the estimation of the position between the various machines, it runs learning algorithms present in the field of computer science. For that is one of the most accurate and complete software among the open-source ones. The tool looks for the anatomical parts of the person by highlighting the position of the joints of the body of several people simultaneously; when they enter the frame, in every position and size (scale), the interaction of the various subjects makes the association, between the various parts of their body, possible.

B. Hypotheses and Assumptions

The goals of this framework are:

- the minimization of the makespan, i.e. the total time required by the resources involved to complete all the tasks assigned to them;
- to guarantee the safety of the workcell, i.e. to maintain a minimum safety distance between the resources involved.

As resources a human operator and a cobot are used.

The collaboration can be of different types, Mathenson et al. (2019), and for this application coexistence is chosen, in order not to introduce constraints. This means the resources perform the assigned tasks simultaneously, but they don’t share any task.

To correctly define the task allocation scheduling it is necessary to consider:

- the layout of the tasks is arranged in a grid;
- the tasks position;
- there are only assembly tasks;
- there may be spatial interferences between the resources;
- the performances of the cobot are considered;
- some tasks can be done by only one resources, while some others can be done by both.

The makespan *ms* is defined as follow:

$$ms = max\{T_o, T_c\}$$

where T_o is the time required by the operator to complete all of his tasks, while T_c is the time required by the cobot to do all the tasks assigned to it. Safety is considered as constraint in the optimization problem.

C. *Real Time Control and Task Allocation*

The aim of the here presented framework is the definition of a dynamic task allocation that considers different variables. The workflow of the framework is illustrated in Figure 5 and in the following lines it is explained.

Both the operator and the cobot have some tasks respectively assigned and the time required to complete each task.

As mention before, there are some tasks that can be done only by the operator and others that can be done only by the cobot; the remainders can be done by both. The goal to be achieved is to parallelize the activities to be carried out to the maximum, dividing them between the resources, in order to minimize the makespan. This may be accomplished through a dynamic task allocation that reassigns the tasks to the cobot based on operator's position.

The first step is to define a static task allocation where the objective functions are determined. These information are forwarded both to the operator and the cobot.

If the operator's position is the same as the one expected from the virtual model, no actions are needed. Instead, if the position is different, it is necessary to verify if the operator is still inside the work area which is defined as a controlled volume in which the cameras are set. In this way it is possible to monitor the distance between the two resources: cobot position is known while operator's position is recorded by the cameras, Figure 4.

Now, two scenarios can occur: the first one is if the operator is outside the work area, causing a so called path error. In this case it is necessary to evaluate if there may be a danger due to interference in the paths of the two resources. In such situation, if the movement of the cobot is dangerous to the operator, its trajectory can be modified, anyways there is no need to generate a new task allocation.

The other scenario is when the operator is inside the defined area and it can be caused by an operator's error: the operator may be performing an unassigned task, or he may take longer than the predetermined time. If this happens, a visual alert should be sent to him in order to correct the error and, if the alarm isn't enough not to exceed the planned time, the task should be assigned dynamically to the cobot, modifying its task allocation schedule, considering the objective functions.

The proposed algorithm use safety as a constraint but it can be used also as an objective function: the task allocation balancing should assign the tasks to the resources basing the division on the estimated distance between them, to respect the safety distance as defined by ISO/TS 15066 (2016). This can lead to a decrease of the makespan because the cobot can reach higher speeds (always collaborative) and it doesn't need to keep them low for all the time. In fact, when the resources are far from each other the safety is guaranteed by the distance itself.

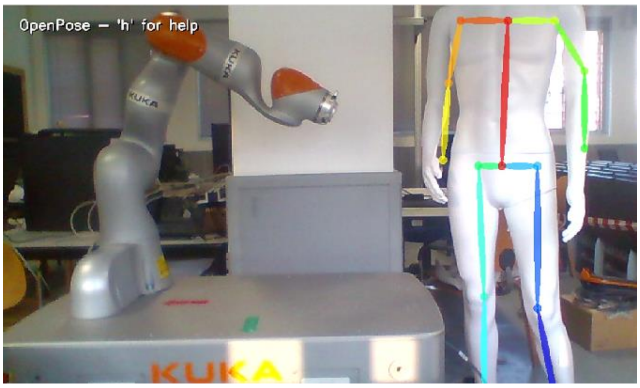

Fig. 4. Frame of the operator's position monitored through OpenPose network, realized in the Robotic Lab, Department of Management and Engineering, University of Padua

Moreover, cobot position is known since it is controlled and operator's position is recorded in real time by the cameras. So, another scenario can be introduced: if the real positions are too close the algorithm can generate a new scheduling for the cobot in order to ensure the safety distance and to avoid an increase of the makespan.

As other objective function, it is possible to include the minimization of energy expenditure by the operator, this because each task demands a certain amount of energy and it depends on the type and the number of movements to be made. If the task turns out to be too energetically demanding for the operator it is assigned to the cobot.

Also, other human factors can be used as objective functions. Physical Ergonomics (PE) and Mental Workload (MW) should be considered in the allocation of tasks, basing the evaluation of them with different indexes. These indexes can be combined together and they can provide a general evaluation of the burden of the analyzed task. As before, the more burdening tasks are assigned to the robot and the lighter ones are assigned to the operator.

V. CONCLUSIONS

This paper presented a framework based on the integration of a motion capture systems into a collaborative work cell to achieve high productivity but also to guarantee the safety required without the introduction of traditional security systems.

After a short description of MoCap techniques, the proposed one is explained. It includes four cameras to have the vision of all work area, which are interfaced with OpenPose to effectively recognise human body and body postures.

The main result of this paper is the definition of an approach to realize a new dynamic task allocation that can include different objective functions. That means, this is only the description of the methodology that represents the future agenda for the coming years, which involves the development of the necessary technologies and the problems optimization, consequently.

The first step is to define these functions, then to create this setup and to validate it, to verify how fast the software can work and if four cameras are enough to cover the controlled volume. The information is exchanged in real time with the cobot that is working in the same area, so, for that reason, one limitation that can be found is the need for high computational capacity.

It was preferred to change the trajectory to the cobot, rather than to the operator, as there is greater guarantee that this change will be followed.

All of this will be numerically and experimentally tested in a collaborative work cell developed in laboratory.

The aim is to avoid the introduction of any other devices, except these already described.

Lastly, this approach can be applied to different industrial case studies, including the supply of the parts that have to be assembled.

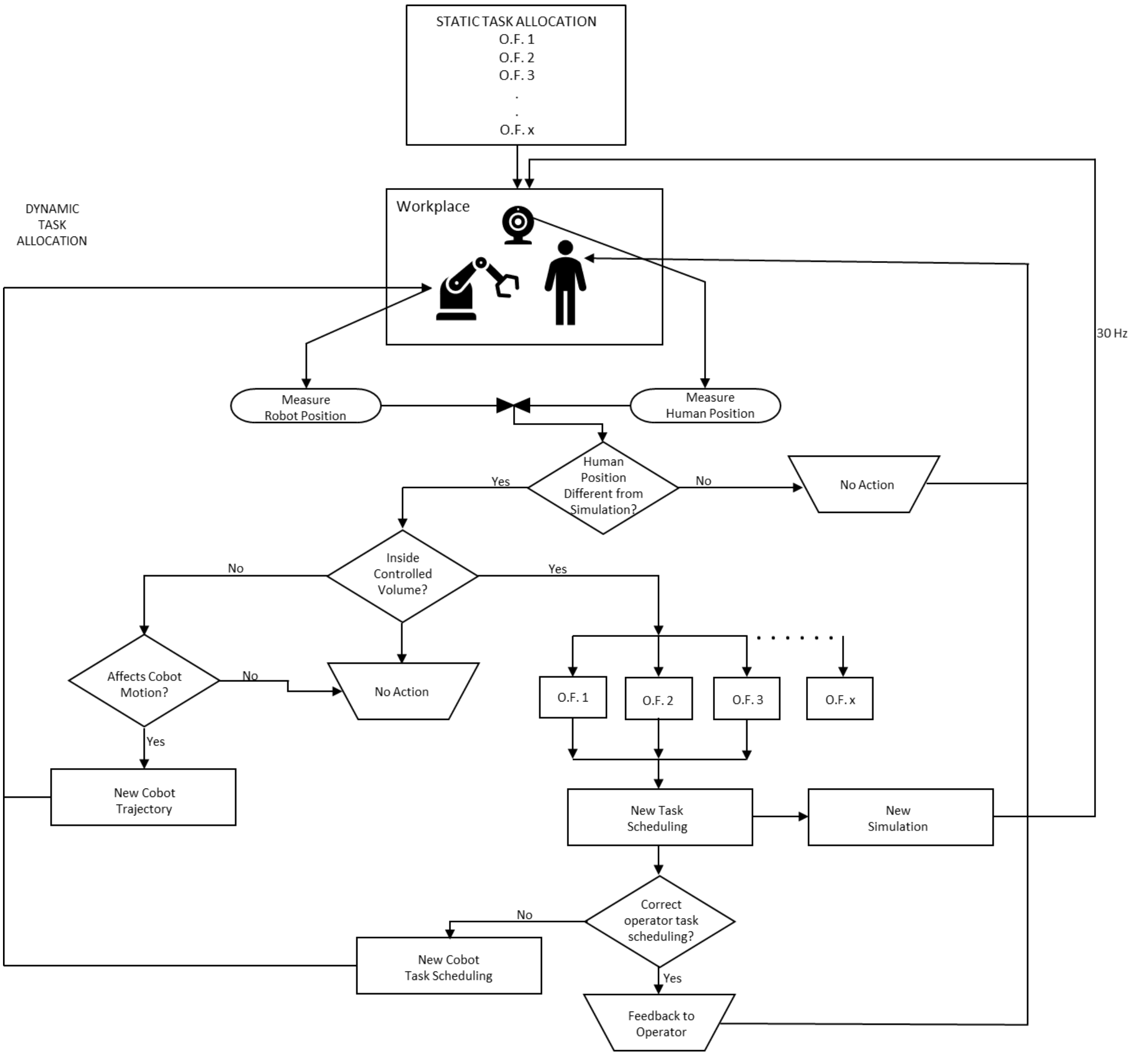

Fig. 5. Operating Algorithm Workflow

REFERENCES

[1] 2016. ISO/TS 15066:2016, Robots and Robotic Devices – Collaborative Robots. *International Organization for Standardization*.

[2] Bai, Chunguang, Patrick Dallasega, Guido Orzes, and Joseph Sarkis. (2020). Industry 4.0 technologies assessment: A sustainability perspective. *International journal of production economics* 229: 107776.

[3] Bortolini, Marco, Faccio, Maurizio, Gamberi, Mauro, and Pilati, Francesco. (2020). Motion Analysis System (MAS) for production and ergonomics assessment in the manufacturing processes. *Computers & Industrial Engineering*, 139: 105485.

[4] Boschetti, Giovanni, Maurizio Faccio, and Riccardo Minto. (2020). Control Model for Collaborative Manufacturing: An Integrated Opened Framework for Human-Robot Collaboration. In *The International Conference of IFToMM ITALY*, 403–413. Springer.

[5] Byner, Christoph, Björn Matthias, and Hao Ding. (2019). Dynamic speed and separation monitoring for collaborative robot applications–concepts and performance. *Robotics and Computer-Integrated Manufacturing* 58: 239–252.

[6] Cao, Z., G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. (2019). OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

[7] Galin, Rinat, Roman Meshcheryakov, Saniya Kamesheva, and Anna Samoshina. (2020). Cobots and the benefits of their implementation in intelligent manufacturing. In *IOP Conference Series: Materials Science and Engineering*, Vol. 862, 032075. IOP Publishing.

[8] Heydaryan, Sahar, Joel Suaza Bedolla, and Giovanni Belingardi. (2018). Safety design and development of a human-robot collab-oration assembly process in the automotive industry. *Applied Sciences* 8 (3): 344.

[9] Kim, Yejin, Seongmin Baek, and Byung-Chull Bae. (2017). Motion capture of the human body using multiple depth sensors. *Etri Journal* 39 (2): 181–190.

[10] Lucci, Niccolò, Bakir Lacevic, Andrea Maria Zanchettin, and Paolo Rocco. (2020). Combining Speed and Separation Monitoring With Power and Force Limiting for Safe Collaborative Robotics Applications. *IEEE Robotics and Automation Letters 5* (4): 6121–6128.

[11] Menache, Alberto. (2000). Understanding motion capture for computer animation and video games. *Morgan kaufmann*.

[12] of Robotics, International Federation.2020. "IFR Press Conference." url: https://ifr.org/downloads/press2018/Presentation_WR_2020.pdf, 9.

[13] Otto, Michael, Eva Lampen, Felix Auris, Felix Gaisbauer, and Enrico Rukzio. (2019). Applicability evaluation of kinect for EAWS ergonomic assessments. *Procedia CIRP* 81: 781–784.

[14] Pearce, Margaret, Bilge Mutlu, Julie Shah, and Robert Radwin. (2018). Optimizing makespan and ergonomics in integrating collaborative robots into manufacturing processes. *IEEE transactions on automation science and engineering* 15 (4): 1772–1784.

[15] Tseng, Mitchell M, Jianxin Jiao, and M Eugene Merchant. (1996). Design for mass customization. *CIRP annals* 45 (1): 153–156.

[16] Weckenborg, Christian, Karsten Kieckhäfer, Christoph Müller, Martin Grunewald, and Thomas S Spengler. (2020). Balancing of assembly lines with collaborative robots. *Business Research* 13 (1): 93–132.

[17] Ye, Mao, Qing Zhang, Liang Wang, Jiejie Zhu, Ruigang Yang, and Juergen Gall. (2013). A survey on human motion analysis from depth data. *In Time-of-flight and depth imaging, sensors, algorithms, and applications*, 149–187 Springer.