Using a Dexterous Robotic Hand for Automotive Painting Quality Inspection

Bruno Santos, Gil Gonçalves, Vítor H. Pinto
SYSTEC (DIGI2), ARISE & ECE Dept.
Faculdade de Engenharia, Universidade do Porto
Porto, Portugal
{brunosantos, gil, vitorpinto}@fe.up.pt

Francisco M. Ribeiro
INESC TEC
Porto, Portugal
francisco.m.ribeiro@inesctec.pt

***Abstract*—The rapid evolution of industrial automation has revolutionised industries over the past few decades, with technology replacing manual labour through the integration of industrial robots. This paper delves into the pivotal role of industrial automation in bolstering productivity, curtailing labour costs, and elevating product quality, paramount pursuits for businesses striving for competitiveness. Situated within a "Digitisation of Human Skills" project, this research focuses on automating tasks in the automotive industry, specifically automotive painting quality inspection. The proposed approach combines a Shadow Dexterous Hand and a UR5 robotic arm, controlled through teleoperation with the BioIK kinematic retargeting algorithm. Two methods, OpenPose and MediaPipe, were assessed for acquiring human hand data for teleoperation. The study demonstrates the successful replication of human movements for automotive painting quality inspection utilising BioIK and MediaPipe, underscoring the potential of automation in this critical industrial domain.**

Keywords—Shadow Dexterous Hand; UR5; MediaPipe; BioIK; teleoperation.

I. INTRODUCTION

The automation of industrial processes has undergone significant evolution in recent decades, fundamentally transforming the operational landscape of the industries. Beginning in the 1960s, advancements in technology have led to the gradual replacement of manual labour with industrial robots, as noted in [1].

In fact, industrial automation has become imperative for companies striving to maintain competitiveness in the contemporary marketplace, owing to its inherent advantages [2]. Chief among these advantages is the substantial increase in productivity [3], attributable to the rapidity of robots compared to human operators, facilitating uninterrupted processes. Additionally, reduced reliance on human labour translates into decreased labour costs and fewer human errors, ultimately enhancing product quality.

As elucidated, this expansive domain of industrial automation holds immense potential and is poised to play a pivotal role in the future of various industries. Consequently, optimising the level of industrial automation should be a primary focus for companies.

This paper is situated within the context of the project "Digitalização da Arte Humana" (Cibertoque), translated as "Digitisation of Human Skills", which aims to develop a robotic system capable of replicating human movements for the purpose of automating certain tasks within the automotive industry, specifically at Stellantis company. The paper's particular focus lies in the realm of automotive painting quality inspection. The research presented constitutes the initial phase of a study whose ultimate goal is to fully automate this task by ideally replacing human operators with robotic systems.

For the accomplishment of the aforementioned task, the proposed approach involves controlling a Shadow Dexterous Hand [4] coupled with a UR5 [5], a collaborative robotic arm. The Shadow Hand consists of an anthropomorphic robotic hand that provides high flexibility of movements due to its 24 joints, specifically engineered to replicate, with the utmost fidelity, the kinematics and precision of the human hand. The control of this robotic set will be executed based on teleoperation, leveraging the kinematic retargeting algorithm BioIK. In order to acquire the human hand data for the teleoperation process, two methods will be explored, namely OpenPose and MediaPipe.

The paper is structured as follows: Section II presents a comprehensive literature review encompassing relevant topics; Section III elucidates the detailed implementation process of the algorithms employed; Section IV showcases the results achieved through the application of the methods discussed in Section III; and lastly, Section V offers conclusions drawn from the work developed in this study and the future work.

II. LITERATURE REVIEW

In this section, pertinent subjects will be addressed within the context of this paper. For each of the topics, the respective literature review will be presented.

In order to execute the teleoperation of an anthropomorphic robotic hand, various methods have been explored. Some studies have proposed approaches for human hand motion acquisition involving the use of data gloves, as demonstrated in [6] [7] [8], or gloves equipped with passive markers, as documented in [9] [10]. In addition, certain research endeavours have focused on methods for non-anthropomorphic robotic grippers, exemplified by [11] [12].

However, the present study is dedicated to implementing a teleoperation method tailored for an anthropomorphic robotic hand and leveraging vision-based techniques to capture human hand motion, offering distinct advantages in terms of flexibility and freedom. Notably, it eliminates the necessity of wearing

 ISBN: 978-1-68558-132-9

restrictive gloves. This choice was motivated by the alignment of these techniques with the specif ic requirements of automotive painting inspection and their cost-effectiveness when compared to the utilisation of data gloves.

Subsequently, the methods used will be presented and studied, organised into two sections: data acquisition and kinematic retargeting.

A. Data Acquisition

Within vision-based methods, there are two well-known algorithms, namely OpenPose [13] and MediaPipe [14].

OpenPose, founded on a convolutional neural network, stands out as the pioneering real-time, multi-person system designed to collectively detect 135 human body, hand, facial, and foot keypoints. Its selection is primarily attributed to its high accuracy, even in challenging scenarios characterised by occlusions and cluttered backgrounds. OpenPose's prowess is substantiated by its superior performance on two datasets, namely Max Planck Institute Informatik (MPII) Human Pose [15] and Common Objects in Context (COCO) [16], where it outperformed prior methods, thus attaining a state-of-the-art status. Additionally, OpenPose's real-time capabilities align seamlessly with the requirements of robotic hand teleoperation and it is an open-source method, meaning it is freely available for research utilisation. Furthermore, in [17], an evaluation of pose estimation accuracy was conducted utilising OpenPose in conjunction with a single RGB-D camera, this study reported a good performance by OpenPose, highlighting the standard deviation of the detected keypoints below 3 millimetres, indicating a high level of repeatability.

On the other hand, MediaPipe, developed by Google, also offers a high-fidelity solution for hand and finger tracking, leveraging machine learning algorithms to infer 21 hand keypoints from a single image. Unlike other methods, such as OpenPose, that rely primarily on powerful desktop environments, MediaPipe distinguishes itself by its low computational footprint, making it possible to achieve real-time performance even on a mobile phone and inclusively scaling to multiple hands. This real-time performance was, in fact, tested in three mobile devices, namely Google Pixel 3, Samsung S20 and iPhone 11, achieving impressing processing times, such as 16.1, 11.1 and 5.3 milliseconds, respectively, utilising the "Full" model [14]. The MediaPipe algorithm has demonstrated its efficacy in various robotic applications, such as in [18], where it effectively captured hand positions with good-quality results. This capability has translated into successful robot control by hand gestures, showcasing the algorithm's reliability and practical utility in real-world scenarios.

B. Kinematic Retargeting

Transitioning to the control of the robotic hand, this subsection undertakes a comprehensive exploration of existing literature on kinematic retargeting algorithms.

Notable among these is BioIK, first introduced in [19]. Then, a clean and high-performance C++ version of the same algorithm was presented in [20].

BioIK is an open-source software package for Robot Operating System (ROS) featuring a bio-inspired optimisation algorithm adept at solving complex optimisation problems, such as inverse kinematics. This algorithm is distinguished by its user-defined weighted goals, offering a high degree of flexibility and control encompassing 18 goal types. These goals include the fundamentals, such as position, orientation, and pose (position and orientation) goals, but also include more complex goals, such as:

- Minimal Displacement Goal: tries to keep each joint angle as close as possible to the previous one;
- Centre Joints Goal: tries to keep each joint centred at the respective joint limits;
- Avoid Joint Limits: similar to Centre Joints Goal;
- Look At Goal: tries to align the orientation of a specified link to a goal position;
- Maximum Distance Goal: tries to keep the position of a specific link within a maximum range from a goal position.

To elucidate the operational efficacy of BioIK, the authors conducted a series of illustrative demonstrations, including one involving the integration of the Shadow Hand with the KUKA LWR 4+ robotic arm. This particular demonstration focused on the application of BioIK to plan the robotic hand and arm motions for turning a wheel button on an audio mixer, as visually represented in Figure 1. For the execution of this task, a tripod grasp configuration was employed, with precise control exerted over the thumb, first finger, and middle finger through the BioIK algorithm. The final trajectory was achieved by combining a set of 200 solutions generated by the BioIK algorithm. This approach resulted in the Shadow Hand's successful manipulation of the wheel button, achieving the desired rotation without any incidence of slippage or unintended disengagement.

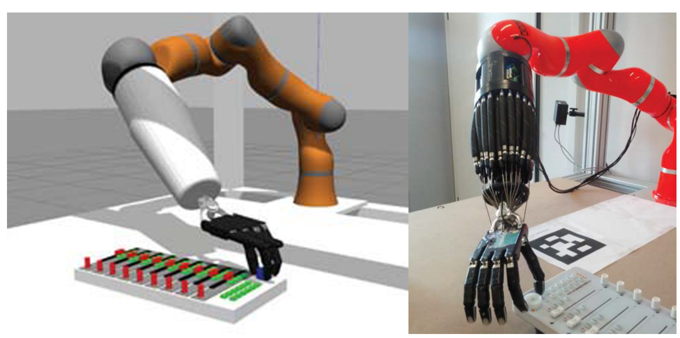

Figure 1. Shadow Hand experiment performed in Gazebo simulation and in reality [20]

In [21], Shuang Li et al. introduced a novel approach to address the complex challenge of kinematic retargeting, leveraging neural network techniques.

In this paper, the authors presented TeachNet, a novel end-to-end neural network architecture. This architecture's primary objective is the estimation of joint angles requisite for replicating the configuration of the human hand within the robotic domain, specifically with the utilisation of a Shadow Dexterous Hand.

The task of directly solving joint regression problems from human hand images is known to be a particularly challenging one. This difficulty primarily stems from the inherent distinction between the domains occupied by the robot hand and the human hand. In light of this, TeachNet comprises two distinctive branches: the robot hand branch, which assumes the role of a teacher, and the human hand branch, serving as the student. Both branches operate with depth images as input and jointly output the required joint angles. Between these two branches, there is a consistency loss strategically designed to align the corresponding features. To better understand the network architecture, it is illustrated in Figure 2.

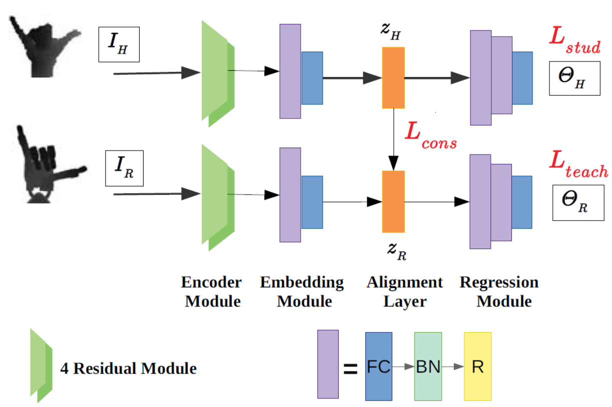

Figure 2. TeachNet architecture [21]

In order to train the developed network, the authors created a new dataset utilising the existing dataset BigHand2.2M [22] and the aforementioned BioIK solver. The process involved the extraction of depth images from the BigHand2.2M dataset and subsequently mapping the corresponding hand keypoints into the corresponding joint angles of the Shadow Hand with BioIK. The solver was set with the following goals: mapping of fingertip positions with a weight of 1, mapping of proximal interphalangeal joint positions with a weight of 0.2, and lastly, mapping proximal and distal phalanges with a weight of 0.2.

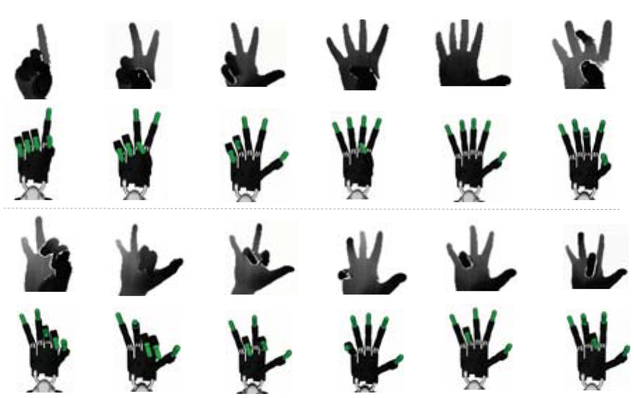

Figure 3. Successful teleoperation results [21]

Finally, to evaluate the efficacy of the proposed method, a series of hand gestures were executed by five novice teleoperators and replicated by Shadow Hand. This test resulted in a success rate of 78.26%, being some of the successful poses presented in Figure 3. Furthermore, manipulation experiments were also performed and compared regarding the necessary time to accomplish a specific task with the DeepPrior++ [23] method. The outcomes underscored the notable efficiency gains achieved through the application of TeachNet, with an average task execution time reduction of 57% compared to the utilisation of DeepPrior++.

III. Implementation

This section delves into the implementation of the algorithms previously mentioned in Section II.

A. Data Acquisition

Regarding the data acquisition phase, first, OpenPose was employed together with Microsoft Kinect v2, similarly as in [17]. The Kinect integrates a FullHD (1920 x 1080 pixels) RGB camera and a 512 x 424 pixels resolution depth camera. In addition, it utilises time-of-flight (ToF) technology to capture depth. This method measures distances based on, as the name implies, time-of-flight, i.e., the round trip time of a light signal emitted and then received by the device.

In the case of OpenPose, the Kinect had two primary purposes. Firstly, capturing 2D images of the human hand, these frames were processed by OpenPose to detect the corresponding 2D hand keypoints. Subsequently, the depth information associated with the detected keypoints was extracted also from Kinect, resorting to its depth sensor.

Afterwards, and resulting from the poor results achieved with OpenPose, detailed in Section IV, MediaPipe was tested in conjunction with a stereo-vision ZED camera. The ZED is a stereo camera developed by StereoLabs that captures high-definition images with depth in real-time. The camera utilises two sensors with a baseline of 12 centimetres to mimic the human stereoscopic vision, enabling it to generate a depth map for, ideally, each pixel in the image. The camera has a resolution of up to 2K (2208 x 1242 pixels) and a field of view of 110 degrees horizontally and 60 degrees vertically.

In this specific case, the ZED camera was utilised solely for capturing 2D raw images from both cameras, functioning as two independent cameras. However, the fixed and known baseline, i.e., the distance between the two cameras, provided an advantageous condition for applying stereo-vision techniques. The approach for 3D detection of human hand keypoints involved the independent 2D detection of keypoints for each camera with MediaPipe, followed by the conversion of these 2D positions from both cameras into depth values utilising stereo-vision techniques, following the formula:

$$Depth = \frac{f \cdot B}{d} \tag{1}$$

 ISBN: 978-1-68558-132-9

where f corresponds to the focal length, B to baseline and d to disparity, this is, the distance in pixels between corresponding keypoints in the left and right images.

B. Kinematic Retargeting

Moving to the kinematic retargeting algorithm, based on the analysis of the algorithm realised in Section II, the BioIK algorithm was selected for mapping human movements to the robotic system composed of the robotic arm and hand. BioIK was favoured due to its ability to simultaneously control the robotic arm and hand.

Prior to running the solver itself, some processing was applied to the keypoints; namely, the acquired human hand keypoints were transformed in a way that the links (distance between two keypoints, equivalent to bones distance) had the same size as the respective ones in the Shadow Hand. For this, an iterative process was performed, starting from the wrist and ending at each fingertip. This process consisted of maintaining the direction of each link and changing its size according to the respective link in the robotic hand. Being K a hand keypoint and d_{link} the Shadow Hand corresponding link distance between two keypoints, the mentioned iterative process can be described by:

$$K_{i+1} = K_i + d_{link} \cdot \frac{\overrightarrow{K_i\,K_{i+1}}}{||\overrightarrow{K_i\,K_{i+1}}||} \tag{2}$$

The final objective of applying this technique was to get better results since the BioIK solver is based on inverse kinematics. The result of the application of this technique is shown in Figure 4. Being Shadow Hand designed to be the same size as a human hand, the difference between keypoints is not substantial, except for the little finger since all the robotic hand primary fingers have the same size, but in reality, the little finger is smaller than the other primary fingers.

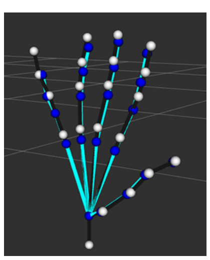

Figure 4. Keypoints before (in blue) and after (in white) the application of robotic hand mapping

Considering the implementation of BioIK itself, a total of 6 different goal types were utilised. The respective goals and their weights are detailed below:

- Position goals for fingertips mapping with weights of 1.0;
- Position goals for knuckles mapping with weights of 0.2;
- Position goals for wrist mapping with a weight of 0.25;
- Direction goals for proximal phalanges of the primary fingers with weights of 0.1;
- Direction goals for intermediate and distal phalanges of the thumb with weights of 0.1;
- Joint function goals for dealing with the coupled joints with a weight of 1.0;
- Centre joints goal with a weight of 0.1;
- Minimal displacement goal with a weight of 0.15;
- Joint function goal to ensure a maximum wrist extension with a weight of 0.5;
- Joint function goal to ensure a workspace above the base plane of UR5 with a weight of 0.15.

It should be noted that the penultimate goal mentioned aims to prevent collisions between the robotic hand's forearm and the bonnet where the paint quality inspection took place. Considering the physical properties of the Shadow Hand, if the hand and respective forearm were horizontally aligned, i.e., with null flexion of the wrist, the forearm would collide with the bonnet.

In summary, the procedure for performing this task involved capturing the keypoints of the human hand and subsequently providing them to the BioIK algorithm, which aims to determine the positions of each joint in the robotic arm and hand, thereby replicating the previously recorded movement, including the intrinsic hand movements and its spatial position. Lastly, these joints' positions were sent to Shadow Hand and UR5, resorting to SrRobotCommander, a high-level interface to control the robotic set.

The algorithms were implemented on a computer with the following specifications:

- Central Processing Unit (CPU): AMD Ryzen™ 9 7950X3D
- Random Access Memory (RAM): 32.0 GB
- Graphics Processing Unit (GPU): NVIDIA GeForce RTX 4090 24GB

IV. RESULTS

This section presents and discusses the outcomes obtained with the deployment of the algorithm described in Section III.

Commencing with the data acquisition, MediaPipe has shown a notable superior performance relative to OpenPose.

During testing with OpenPose algorithm, it was observed that specific keypoints, in particular the fingertips, were frequently detected incorrectly, leading to significant deviations from their actual positions. After further examination, it was found that these deviations were due to the integration with Microsoft Kinect. A slight difference between the actual fingertip and its detection by OpenPose was enough for the depth value extracted from the Kinect to deviate from the fingertip and correspond, for example, to the background, thus inevitably resulting in a keypoint position measurement quite dissimilar to the actual one. To minimise the impact of this issue, a median filter was applied to the acquired keypoints position. This filter was chosen due to the low influence of outliers on the filtered result. Figure 5 illustrates the effectiveness of the median filter in removing outliers, such as the wrongly detected index fingertip by OpenPose.

 ISBN: 978-1-68558-132-9

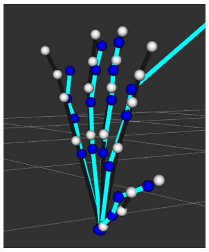

Figure 5. Hand keypoints before (in blue) and after (in white) the application of the median filter

In addition to this initial problem, OpenPose proved to be highly unstable. Frequently, this software incorrectly detected some of the fingers, as can be seen in Figure 6 by the change in the characteristic detection colour of each finger.

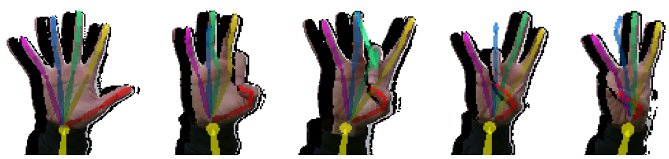

Figure 6. OpenPose misdetections (left image for reference)

Considering these results, in addiction to the fact that OpenPose is extremely computationally heavy compared with MediaPipe, this last method was chosen.

Moving on to the BioIK kinematic retargeting algorithm, its execution time has been experimentally verified as 200 milliseconds, with an end-to-end time delay of about 1 second. When practically applied, it was possible to infer that it behaved as initially expected, replicating with a fair degree of precision the movements made by the human hand.

Once both the acquisition method and kinematic retargeting algorithm were well defined, experimental tests were conducted to evaluate the developed algorithm.

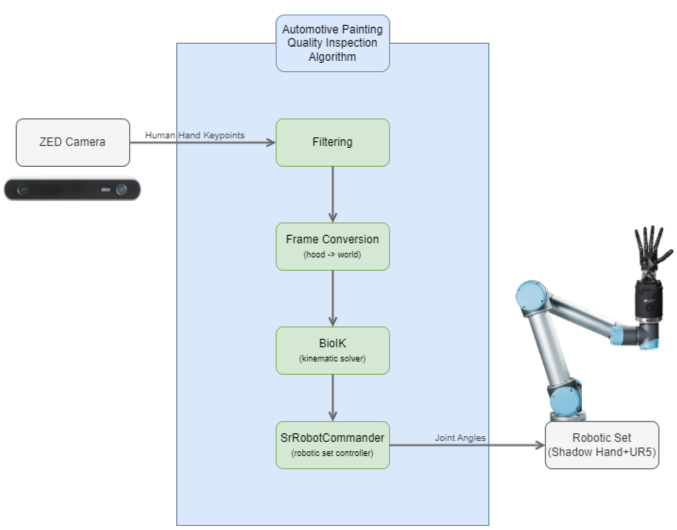

Figure 7. Automotive painting quality inspection algorithm diagram

The complete algorithm sequence, presented in Figure 7, for the accomplishment of the automotive painting quality inspection, includes the following steps:

1) Capture human hand images with a ZED camera;
2) Acquire 3D hand keypoints positions, utilising MediaPipe and stereo-vision techniques, according to a referential frame positioned on the car bonnet;
3) Change the referential frame from the car bonnet to the robot frame;
4) Calculate the robotic set joints' angles utilising BioIK;
5) Lastly, send the calculated joints to the robotic set.

Throughout the various iterations of tests carried out, slight differences (mostly less than a centimetre) were observed in the depth component, normal to the surface, of the acquired points. Consequently, in some instances, the robotic hand exhibited slight deviations from the surface.

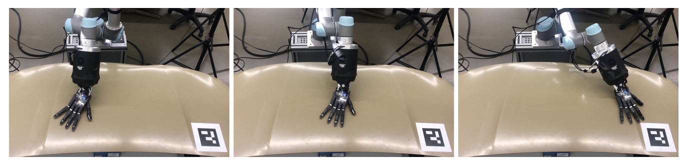

Figure 8. Automotive painting quality inspection process

Nonetheless, despite these minor discrepancies, the algorithm proved effective in replicating human movements, as evidenced by the series of frames presented in Figure 8, extracted from a demonstration video [24] showcasing the algorithm's execution.

V. Conclusion and Future Work

The research undertaken in this study focused on advancing the automation of quality inspection in automotive painting, specifically by replicating human movements through the integration of an anthropomorphic robotic hand, Shadow Dexterous Hand, and a UR5 robotic arm. Through the exploration of kinematic retargeting techniques, BioIK emerged as the preferred algorithm due to its efficient control over both the robotic arm and hand.

In the process, the integration of OpenPose and MediaPipe for data acquisition revealed notable insights. Despite initial considerations, OpenPose faced challenges related to precision and stability, leading to its exclusion from the final methodology. These challenges served as valuable lessons in the selection of appropriate technologies and highlighted the importance of robust data acquisition methods in achieving accurate task replication.

The exclusion of OpenPose prompted a reevaluation of alternative options for data acquisition. MediaPipe, with its more stable performance, was successfully integrated.

Regarding the automotive painting quality inspection task, thanks to the utilisation of the BioIK together with MediaPipe and stereo-vision techniques, successful replication of movements resembling those performed by a human operator during the paint quality inspection task was achieved.

 ISBN: 978-1-68558-132-9

As future work, it would be very interesting to integrate more sophisticated and especially more suitable tactile sensors into the robotic hand, particularly at the fingertips, capable of providing detailed and nuanced feedback. This integration would elevate the precision and effectiveness of the automated automotive painting quality inspection system.

Acknowledgement

This work was financially supported by *PPS 10: Neural Networks - Robotic Systems for Industry 4.0* from *Agenda GreenAuto: Green Innovation for the Automotive Industry*, no. C644867037-00000013, investment project no. 54, from the Incentive System to Mobilising Agendas for Business Innovation, funded by the Recovery and Resilience Plan and by European Funds NextGeneration EU. The authors are members of the ARISE Associated Laboratory (LA/P/0112/2020) and R&D Unit SYSTEC-Base (UIDB/00147/2020) and Programmatic (UIDP/00147/2020).

References

[1] Y. Zhao, "Current status and industrialization development of industrial robot technology," vol. 81, 2021, pp. 804–808, doi: 10.1007/978-3-030-79197-1_117.

[2] M. Bartoš *et al.*, "An overview of robot applications in automotive industry," *Transportation Research Procedia*, vol. 55, pp. 837–844, 2021, doi: 10.1016/j.trpro.2021.07.052.

[3] C. Rosa, F. Silva, and L. P. Ferreira, "Improving the quality and productivity of steel wire-rope assembly lines for the automotive industry," *Procedia Manufacturing*, vol. 11, pp. 1035–1042, 2017, doi: 10.1016/j.promfg.2017.07.214.

[4] Shadow Robot Company, London, GBR, *Shadow Dexterous Hand Technical Specification*, 2021, [Online]. Available from: https://www.shadowrobot.com/wp-content/uploads/2022/03/shadow_dexterous_hand_e_technical_specification.pdf 2024.02.05.

[5] Universal Robots, Odense, DNK, *UR5 Technical specifications*, 2016, [Online]. Available from: https://www.universal-robots.com/media/50588/ur5_en.pdf 2024.02.05.

[6] B. Fang, F. Sun, H. Liu, and D. Guo, "A novel data glove using inertial and magnetic sensors for motion capture and robotic arm-hand teleoperation," *Industrial Robot*, vol. 44, p. 155 – 165, 2017, doi: 10.1108/IR-07-2016-0179.

[7] J. Nassour *et al.*, "A robust data-driven soft sensory glove for human hand motions identification and replication," *IEEE Sensors Journal*, vol. 20, pp. 12 972–12 979, 2020, doi: 10.1109/JSEN.2020.3001982.

[8] Q. Zhu *et al.*, "Teleoperated grasping using a robotic hand and a haptic-feedback data glove," in *2020 IEEE International Systems Conference (SysCon)*, 2020, pp. 1–7, doi: 10.1109/SysCon47679.2020.9275927.

[9] A. Fabisch *et al.*, "A modular approach to the embodiment of hand motions from human demonstrations," in *2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids)*, 2022, pp. 801–808, doi: 10.1109/Humanoids53995.2022.10000165.

[10] Y. Lee *et al.*, "Visual-inertial hand motion tracking with robustness against occlusion, interference, and contact," *Science Robotics*, vol. 6, 2021, doi: 10.1126/scirobotics.abe1315.

[11] C. Meeker, M. Haas-Heger, and M. Ciocarlie, "A continuous teleoperation subspace with empirical and algorithmic mapping algorithms for nonanthropomorphic hands," *IEEE Transactions on Automation Science and Engineering*, vol. 19, pp. 373–386, 2022, doi: 10.1109/TASE.2020.3035156.

[12] X. Bai *et al.*, "Kinect-based hand tracking for first-person-perspective robotic arm teleoperation," in *2018 IEEE International Conference on Information and Automation (ICIA)*, 2018, pp. 684–691, doi: 10.1109/ICInfA.2018.8812561.

[13] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "Openpose: Realtime multi-person 2d pose estimation using part affinity fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, p. 172 – 186, 2021, doi: 10.1109/TPAMI.2019.2929257.

[14] F. Zhang *et al.*, "Mediapipe hands: On-device real-time hand tracking," 2020, doi: 10.48550/arXiv.2006.10214.

[15] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, "2d human pose estimation: New benchmark and state of the art analysis," 2014, p. 3686 – 3693, doi: 10.1109/CVPR.2014.471.

[16] T.-Y. Lin *et al.*, "Microsoft coco: Common objects in context," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8693 LNCS, p. 740 – 755, 2014, doi: 10.1007/978-3-319-10602-1_48.

[17] F. Lygerakis, A. C. Tsitos, M. Dagioglou, F. Makedon, and V. Karkaletsis, "Evaluation of 3d markerless pose estimation accuracy using openpose and depth information from a single rgb-d camera," 2020, p. 533 – 538, doi: 10.1145/3389189.3398005.

[18] S. Sreenath, D. I. Daniels, A. S. D. Ganesh, Y. S. Kuruganti, and R. G. Chittawadigi, "Monocular tracking of human hand on a smart phone camera using mediapipe and its application in robotics," 2021, doi: 10.1109/R10-HTC53172.2021.9641542.

[19] S. Starke, N. Hendrich, S. Magg, and J. Zhang, "An efficient hybridization of genetic algorithms and particle swarm optimization for inverse kinematics," 2016, p. 1782 – 1789, doi: 10.1109/ROBIO.2016.7866587.

[20] P. Ruppel, N. Hendrich, S. Starke, and J. Zhang, "Cost functions to specify full-body motion and multi-goal manipulation tasks," 2018, p. 3152 – 3159, doi: 10.1109/ICRA.2018.8460799.

[21] S. Li *et al.*, "Vision-based teleoperation of shadow dexterous hand using end-to-end deep neural network," vol. 2019-May, 2019, p. 416 – 422, doi: 10.1109/ICRA.2019.8794277.

[22] S. Yuan, Q. Ye, B. Stenger, S. Jain, and T.-K. Kim, "Bighand2.2m benchmark: Hand pose dataset and state of the art analysis," vol. 2017-January, 2017, p. 2605 – 2613, doi: 10.1109/CVPR.2017.279.

[23] M. Oberweger and V. Lepetit, "Deepprior++: Improving fast and accurate 3d hand pose estimation," vol. 2018-January, 2017, p. 585 – 594, doi: 10.1109/ICCVW.2017.75.

[24] https://youtu.be/lpxIPF_6WYc 2024.02.05.