UNIVERSITI PUTRA MALAYSIA

ONLINE TELEOPERATION OF WRITING MANIPULATOR THROUGH GRAPHICS PROCESSING UNIT BASED ACCELERATED STEREO VISION

FADI IMAD OSMAN ABU RAID

FK 2022 30

ONLINE TELEOPERATION OF WRITING MANIPULATOR THROUGH GRAPHICS PROCESSING UNIT BASED ACCELERATED STEREO VISION

By

FADI IMAD OSMAN ABU RAID

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

December 2021

DEDICATION

This thesis work is wholeheartedly dedicated to my beloved parents, who have been a source of inspiration and strength, who always provided their encouragement as well as emotional and financial support.

To my wife Leena, who has been a constant source of support, who put up with me through all the late nights and early mornings.

To my beloved daughter Joodi.

To my beloved sisters Fajer, Farah and Fida.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

ONLINE TELEOPERATION OF WRITING MANIPULATOR THROUGH GRAPHICS PROCESSING UNIT BASED ACCELERATED STEREO VISION

By

FADI IMAD OSMAN ABU RAID

December 2021

Chair: Associate Professor Sharifah Mumtazah bt Syed Ahmad Abdul Rahman, PhD
Faculty: Engineering

Remotely operated robots are of great use in unstructured environments where their autonomous counterparts are challenged. These robots are commonly operated using knobs, joysticks or wearable motion sensors. These modes of operation either require a skilled operator or hinder the natural movement. Therefore, contactless motion trackers such as stereo vision, structured light and time of flight systems were invented to allow natural control with minimum training. Stereo vision systems have the advantage over other systems of being cheaper while having scalable accuracy and range. These benefits however are challenged by the high computational requirements of the algorithms used. In this thesis, a framework is developed to enable the use of stereo vision in real-time teleoperation of a manipulator robot for the task of writing. The proposed algorithms aim to accelerate the processes of distortion removal and rectification using precomputed combined maps. Furthermore, the algorithms avoid computationally intensive correspondence problem by matching only the points of interest of a labeled pen to compute its position and orientation. These algorithms are then parallelized using Compute Unified Device Architecture CUDA C language to run on Graphics Processing Unit GPU for hardware acceleration.

Like most sensors stereo camera readings are susceptible to noise due to lighting conditions and detection errors. Most techniques used in filtering noise such as the simplest moving average or polynomial spline fitting either eliminates some original important data along with the noise or are not optimized for real-time filtering. Thus, Kalman filter which is popular in tracking applications is proposed. Kalman filter is applied to readings along with the Savitzky-Golay and Moving Average filters. The performance of filtering methods is compared in term of processing speed and Root Mean Square Error (RMSE) with ground truth data

collected from a high accuracy digitizing tablet. The proposed technique has demonstrated a significant improvement by reducing processing time and having low RMSE that is nearly equivalent to Savitzky-Golay filter and lower than Moving Average filter. Furthermore, it does not require a window that contains future data samples.

Cameras capture images at constant frame rates coupled with hand movement speed variations, this can result in the trajectory points generated have varying movement distance and exceeding the robot speed constraints. Furthermore, most manipulator robots in use are designed with the application in mind being pick and place. Therefore, they follow a speed profile between the pick and place points that begins with acceleration, cruise and finally deceleration. When sending a stream of points to robot controller the speed profile results in the robot having a jittery behavior. To deal with these issues, adaptive sampling algorithm is applied to the data points to adhere to the manipulators speed constraints. Then the response of the internal low-level PID controller was examined and gains were tuned to reduce the rise time to meet real-time requirements. To find the joint angles from the position and orientation of end effector the inverse kinematics equations of the arm are computed beforehand. Then joint angles are streamed at 0.02 seconds interval to ensure that the end effector glides smoothly along the drawn trajectory.

The performance of the whole system is evaluated by taking the end effector pose readings from internal sensors through Real-Time Data Exchange (RTDE) port and comparing those against the data from the digitizing tablet. The results demonstrated that the movement path is traced by the manipulator while maintaining minimum error of 0.95 to 1.81mm and minimum delay compared to other implementations. The positional error is reduced by 52% compared to the best-known implementation with captured frames being processed in real-time.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

TELEOPERASI MANIPULATOR PENULISAN SECARA ONLINE MELALUI UNIT PEMPROSESAN GRAFIK BERASASKAN PEMECUTAN PENGLIHATAN STEREO

Oleh

FADI IMAD OSMAN ABU RAID

Disember 2021

Pengerusi: Profesor Madya Sharifah Mumtazah bt Syed Ahmed Abdul Rahman, PhD
Fakulti : Kejuruteraan

Robot kawalan jauh kebanyakan digunakan di dalam persekitaran penstrukturan tidak sekata, dimana bahagian-bahagian autonomi mereka dicabar. Robot ini biasanya dikendalikan dengan menggunakan knob, kayu ria atau penderia pengesan pergerakan. Mod operasi seperti ini memerlukan pengendali yang mahir atau kawalan semulajadi. Oleh itu, pengesan pergerakan tanpa sentuhan seperti penglihatan stereo, cahaya berstruktur dan masa sistem masa penerbangan dicipta untuk membolehkan kawalan semula jadi dilakukan dengan latihan yang minimum. Sistem penglihatan stereo mempunyai kelebihan berbanding sistem lain kerana ia lebih murah, disamping mempunyai ketepatan dan julat berskala. Namun begitu, sistem ini memerlukan pengiraan algoritma yang komplek secara berkomputer. Dalam tesis ini, satu rangka kerja telah dibangunkan untuk membolehkan penggunaan penglihatan stereo pada masa sebenar untuk teleoperasi tugasan menulis melalui manipulasi robotic. Algoritma yang dicadangkan adalah bertujuan untuk mempercepatkan proses penghapusan herotan dan pembetulan peta gabungan secara berkomputer. Selain itu, algoritma mengelakkan masalah korespondesi intensif secara komputasi dengan hanya menanda titik yang dikehendaki yang dilabel pen untuk mengira kedudukan dan orientasinya. Algoritma ini kemudiannya diselari bersama bahasa mengira seni bina peranti bersatu *(CUDA C)* pada unit pemprosesan grafik *(GPU)* untuk pecutan perkakasan.

Seperti kebanyakan kamera penderia berstereo ia terdedah kepada bunyi bising kerana masalah pencahayaan dan ralat pengesanan. Kebanyakan teknik digunakan dalam menapis bunyi ialah seperti purata pergerakan mudah atau pemasangan spline polinomial dengan cara menghapuskan beberapa data penting asal bersama-sama dengan bunyi bising atau tidak dioptimumkan penapisan masa nyata. Oleh itu, penapis Kalman yang popular dicadangkan

dalam aplikasi penjejakan ini. Penapis Kalman digunakan untuk bacaan bersama-sama dengan penapis Savitzky-Golay dan penapis Purata Bergerak *(Moving Average filter)*. Kaedah pretasi penapisan dibanding dari segi kelajuan pemprosesan dan punca min ralat kuasa dua *(RMSE)* dengan data permukaan sebenar yang dikumpulkan dari tablet pendigitalan ketepatan tinggi. Teknik yang dicadangkan telah menunjukkan peningkatan yang ketara dengan mengurangkan masa pemprosesan dan mempunyai *RMSE* yang rendah yang hampir sama dengan penapis Savitzky-Golay dan lebih rendah daripada penapis Purata Bergerak *(Moving Average filter)*. Selain itu, ia tidak memerlukan tetingkap yang mengandungi sampel data masa depan.

Kamera menangkap imej bingkai pada kadar malar ditambah dengan variasi kelajuan pergerakan tangan, ini boleh menyebabkan mata trajektori yang dihasilkan mempunyai jarak pergerakan yang berbeza-beza dan melebihi kekangan kelajuan robot. Tambahan pula, kebanyakan robot manipulator yang digunakan direka dengan aplikasi dalam fikiran yang dipilih dan tempat. Oleh itu, mereka mengikuti profil kelajuan antara titik pilihan dan tempat yang bermula dengan pecutan, pelayaran dan akhirnya perlambatan. Apabila menghantar aliran titik kepada robot pengawal kelajuan profil menyebabkan robot yang mempunyai tingkah laku goyah. Untuk menangani isu-isu ini, algoritma persampelan adaptif digunakan pada titik data untuk mematuhi kekangan kelajuan manipulators. Kemudian tindak balas tahap rendah dalaman *PID* telah diperiksa dan gain di kawal untuk mengurangkan masa kenaikan dan memenuhi keperluan masa nyata. Untuk mencari posisi sudut bersama dan orientasi kesan akhir persamaan kinematics songsang lengan dikira terlebih dahulu. Kemudian sudut bersama distrim pada selang 0.02 saat untuk memastikan bahawa kesan akhir efektor meluncur di sepanjang garis trajektori adalah lancar.

Prestasi keseluruhan sistem dinilai dengan mengambil bacaan kesan akhir dari penderia dalaman melalui Pertukaran Data Masa Nyata *(RTDE)* dan membandingkannya dengan data dari tablet pendigitalan. Keputusan menunjukkan bahawa laluan pergerakan dikesan oleh manipulator sambil meminumum kadar kesilapan dari 0.95 hingga 1.81mm dan kelewatan minimum berbanding dengan pelaksanaan lain. Kesalahan kedudukan dikurangkan sebanyak 52% berbanding pelaksanaan yang paling terkenal dengan bingkai yang ditangkap diproses dalam masa nyata.

ACKNOWLEDGEMENTS

Throughout my journey in conducting this research and writing this thesis I have received a great deal of support and assistance.

First and foremost, I would like to thank my supervisor, Associate Professor Sharifah Mumtazah Binti Syed Ahmad Abdul Rahman, whose wealth of knowledge and meticulous editing was invaluable in formulating and directing this research. Her advice and encouragement pushed me to refine my work to bring it to a higher level.

Also, thanks to my supervisory committee members, Associate Professor Shaiful Jahari Bin Hashim and Dr. Khairulmizam Bin Samsudin, who offered guidance, support and valuable suggestions.

I would also like to thank Mr. Sallehuddin Bin Zainal and Mr. Mohamed Hisham Bin Ali, who were always there with me during the experimental lab work. Their company in the laboratory and the help they offered will be always remembered. I would like to specially acknowledge Mr. Sallehuddin Bin Zainal for his care in the laboratory and his help in taking photos and videos to document the work.

Finally, I would like to thank Mr. Zulkhairi Bin Ayop who helped with the translation of the abstract through his multilingual and technical knowledge.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Sharifah Mumtazah binti Syed Ahmad Abdul Rahman, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Shaiful Jahari bin Hashim, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Khairulmizam bin Samsudin, PhD
Senior Lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Member)

____
ZALILAH MOHD SHARIFF , PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 19 May 2022

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

API	Application Programming Interface
CAD	Computer Aided Design
CCD	Charge Coupled Device
COBOT	Collaborative Robot
CPU	Central Processing Unit
CUDA	Compute Unified Device Architecture
DH	Denavit Hartenberg
DoF	Degrees of Freedom
FIR	Finite Impulse Response
FLC	Fuzzy Logic Controller
GAPDs	Geiger-mode Avalanche Photodiodes
GPU	Graphical Processing Unit
HMI	Human Machine Interface
HSV	Hue, Saturation and Value
MA	Moving Average filter
MIMO	Multiple Inputs Multiple Outputs
MPC	Model Predictive Control
PID	Proportional Integral Derivative
PLA	Polylactic acid
PLA	Polylactic Acid Memory
RAM	Random Access
RGB	Red, Green and Blue
RTDE	Real-time Data Exchange
SG	Savitzky-Golay filter

SPADs	Single Photon Avalanche Diodes
SURF	Speeded Up Robust Features
TCP	Tool Centre Point
ToF	Time of Flight
ZN	Ziegler Nichols

CHAPTER 1

INTRODUCTION

1.1 Background

Over the time robots have become popular in industrial applications such as material handling, welding assembly and dispensing. They have demonstrated high efficiency and speed while performing preprogramed repetitive tasks that are often tedious and physically strenuous to human workers and may cause work injuries. But the future requires more than just the automation and execution of tasks in structured and static environments to realize and reap the real benefits of automation. However, the deployment of autonomous robots in dynamic and unstructured environments poses several challenges because of the complexity and inherent uncertainty. Recently, with the advancement in the field of artificial intelligence and the exponential increase in CPU processing power researchers have been successful in developing robots that can handle the complexity of unstructured environments to some extent. Nevertheless, there are still many applications where the tasks cannot be executed autonomously by robots due to the presence of unexpected scenarios or highly unstructured environments (Cui, Tosunoglu, Roberts, Moore, & Repperger, 2003). In such applications robots that are controlled remotely or teleoperated are used to combine the strength and endurance of the robot with the rational capabilities of the human operator. Remotely operated robots are beneficial when operation area is small, and human cannot fit such as in pipelines or in cases when the desired movement is very small as in medical surgery to minimize the incision. Furthermore, they are used to keep human operator away from hazardous areas by performing dangerous tasks such as operations in war zones, bomb defusal, handling of radioactive materials or mining. Moreover, they are used in tasks where the human operator would use a lot of effort to reach the area of operation this includes underwater exploration and inspection, in-space robotics and surveillance. Additionally, tele-operated robots are used in long-distance surgery which allows people to access world-leading expertise without the need to travel.

Most modes of control for robot teleoperation utilize equipment such as buttons, joysticks (Wang, Sun, Hu, Li, & Liu, 2009), joint sensors electromagnetic tracking devices, inertial sensors, data gloves or skin electrodes (Hu, et al., 2005). However, these devices require high precision motion control. It means that the operator needs to be experienced and have prior knowledge in order to manipulate the robot effectively. Besides other devices which can track the operator’s hand position and orientation in real time such as inertial sensors and data gloves are wearable sensors and may hinder the natural human-limb motion. On the other hand, vision and speech based control techniques are contactless thus allow for more natural unrestricted control.

Many studies were carried on robot control relied on hand gestures and speech recognition (Waldherr, Romero, & Thrun, 2000) (Rogalla, Ehrenmann, Zollner, Becher, & Dillmann, 2002). These allowed the operator to give out commands more naturally and more intuitively. However, in this method, the motion of the manipulated robot is constituted by series of simple commands or gestures, such as rotation, translating up and down, grabbing, etc. Therefore, it does not allow for high complexity, precise movements, as it is not easy translate them into a series of simple tasks. Furthermore, these commands and gestures require a trained and experienced operator to achieve a successful and efficient execution of tasks.

The advancements in the depth sensors technology for examples Kinect and the leap motion sensors can solve the problem of tracking the position and orientation to relatively high accuracy, speed and they can be interfaced easily. Thus, they have been used in manipulator control by some researchers (Qian, Niu, & Yang, 2013) (Du G. , Zhang, Mai, & Li, 2012) (Du & Zhang, 2014). However, the fundamental technology these sensors rely on has some drawbacks. These systems use infrared source along with infrared camera to map the surrounding. The downside of using infrared systems is the significant degrade in accuracy when operating under the interference of ambient light or other infrared sources. Additionally, since they rely on structured light projection, they suffer from motion artifacts. More advanced time-of-flight cameras such as Kinect 2 suffer from a low spatial resolution due to the extra processing that happens at the sensor level. These are also costly at the time being due to them being an exclusive technology of Microsoft.

Since cameras has become more affordable and the processing power of computers had increased, systems that rely of stereo correspondence are becoming increasingly viable. With better processing algorithms stereo cameras can offer the user more natural, contactless and user-friendly experience. This natural way to communicate with the robot allows the operator to focus on the task instead of thinking and remembering limited separate commands that the human-robot interface can understand like gesture-based approaches. Furthermore, since it is non-contact, stereo camera tracking avoids the problem that physical sensors, cables and other contacting interfaces may hinder natural motions.

The study proposes the use of stereo vision for robot teleoperation taking into consideration the advancements in hardware and software that makes it a viable option nowadays compared to other motion tracking methods. Handwriting is the task selected to be teleoperated due to its high precision requirements and the traces left behind by the pen which enables evaluating the accuracy of the system.

1.2 Problem Statement

As stated previously, stereo camera visual control has great advantage of being non-contact and does not require bulky sensors attached to limbs and joints of operator. Nevertheless, accomplishing real-time control with high accuracy which is a requirement for some applications is challenged by several issues.

The first challenge is the computational requirement to achieve precise measurement of operators moves using stereo cameras. The precision of the stereo camera system relies mainly on two factors the calibration of the cameras used and the resolution of images captured. The more accurate the calibration parameters and the higher the image resolution the better accuracy of the overall system. Researchers tend to avoid the tedious calibration step by ignoring the errors associated with distortion and misalignment or by relying on commercial infrared based visual sensors (Nowak, Białek, Lindner, & Wyrwał, 2020) (Du G. , Zhang, Mai, & Li, 2012) (Xi, et al., 2017). These decisions however affect the accuracy of tracking due to lack of calibration or because commercial sensors are not designed for the general use in every application. Commercial infrared sensors have a fixed field of view and resolution that meet the requirements they were designed for. For example, Microsoft Kinect was designed to track body and limbs for gaming which doesn't require the highest accuracy. Therefore, when uncalibrated cameras or commercial infrared visual sensors are used by researchers they achieve limited accuracy which doesn't meet the requirements of a task like writing (Nowak, Białek, Lindner, & Wyrwał, 2020) (Du G. , Zhang, Mai, & Li, 2012) (Xi, et al., 2017). As stereo camera calibration is essential it was applied in a research work by Moriya et al., (Moriya, Hayashi, Tominaga, & Yamasaki, 2005) and it proved that submillimetre accuracy can be achieved from images at the resolution of 640 × 480 pixels. However, this is only valid when the object tracked is at a short distance, which is a limiting factor when controlling a robot and trying to make use of the full working space. This calls for increasing the resolution to allow for more accurate tracking at a far distance. However, increasing the image resolution will be at the expense of frame rate unless the processing power of the computer is increased considerably. This due to the increased number of pixels that has to be processed to find a solution for correspondence problem which causes the process in general to be more computationally intensive.

Secondly, camera measurements are highly affected by lighting conditions. Lighting variations, reflections and refractions can result in unstable measurements and uncertainty errors. These errors can produce measurement fluctuations that if relayed to robotic system can cause much undesired jerks and jolts. This can be visualised in the results achieved by Du G et al. and Nowak et al (Du G. , Zhang, Mai, & Li, 2012) (Nowak, Białek, Lindner, & Wyrwał, 2020). Therefore, smoothing of the noisy measurements is essential for the smooth operation of the robot. A common candidate among scientists for signal filtering is moving average filter because of its simplicity and convenient for online real time processing (Xu, Zhong, Ding, Chen, & Chen, 2013) (Mitsantisuk, Piriyanont, & Ohishi, 2018) (Yang, Chang, Liang, Li, & Su, 2015). This is due to it being recursive so that new measurements can be processed as they arrive. However, simple moving average filter does not preserve data on some features such as peak height and width. Other filter type which is commonly used is polynomial fitting filter or spline interpolation filters. It has been demonstrated that it provides smooth filtered trajectory for robot manipulation (Biagiotti & Melchiorri, 2010) (Biagiotti & Melchiorri, 2013) (Mendoza, Kagami, & Hashimoto, 2010) (Feng G. , 1998). Nevertheless, it is not suitable for real-time applications due to the high computational cost as the entire set of via-points must be interpolated. Another filtering technique is Savitzky-Golay filter that is based on least squares fitting with a moving window (Schafer, 2011). It operates by replacing each value in the window with a new value which is obtained from a low order polynomial fit using a simple convolution. It was proven to be effective in noise filtering in real time as demonstrated be (Small, Nyarko, & Scott, 2008). Few researchers employed this technique and their results demonstrated its effectiveness in filtering 3D tracking data (Carreras, Ridao, García, & Nicosevici, 2003).

Finally, although manipulator robots come with internal controllers, however, these controllers do not fulfil the required specification of our applications. As these manipulators are commonly used for pick and place tasks in industry, they follow a certain speed profile. The speed profile is set so that Tool Centre Point (TCP) accelerates from the starting point until it reaches the set speed then moves with constant speed, after that it starts to decelerate until it reaches the set point where it stops. Maximum speed and acceleration can be set but robots will always try to follow this speed profile and therefore stop at every setpoint. Since our tracking application relies on feeding the robot with a series of trajectory points in real-time, this results in the robot having a very jittery movement and will induce lag and inaccuracies. Various types of controllers are typically applied to manipulator robots but the most common and effective is proportional integral derivative controller (PID). Yet, the main challenge with the application of PID controller is obtaining the optimal parameters that fit the application. Copot et al., (Copot, Muresan, Ionescu, Vanlanduit, & De Keyser, 2018) applied auto tuning technique which was proven to be effective in finding the optimal constants for a similar manipulator.

1.3 Objective of the Study

The main objective of this thesis is to design a framework that enables the use of stereo vision to achieve high accuracy and smooth real-time robot teleoperation.

The main objective can be broken down into few sub-objectives as follows:

i. To construct a stereo vision system using HD cameras and applying supervised self-calibration algorithm to achieve accurate real-time tracking of a tool in 3D space.
ii. To develop GPU based algorithms to achieve real-time visual control of a robotic manipulator by using transformation maps for rectification and grayscale segmentation for tracking.

iii. To evaluate experimentally the effectiveness of different smoothening filters when applied to handwriting tracking data and select the most effective filter to be applied.
iv. To optimize the closed loop robot joints controller, through tunning and providing uniformly spaced trajectory points to overcome the jittery behavior and achieve high accuracy as well as fast response.

1.4 Scope of the Study

The techniques proposed and evaluation methods performed during the study are summarized as follow:

i. The proposed capturing system consists of two consumer grade HD webcams. The webcams are set up and calibrated to form a stereo vision system. The accuracy of camera parameters was evaluated using qualitative measures of average reprojection error.
ii. The tracking algorithm proposed works by tracking the tool rather than the hand of the operator since teleoperation mainly aims on transferring skill of using a tool. And writing was used as an application since pens leave traces which facilitates the process of evaluating the accuracy of the system. The system accuracy was compared to Wacom digitizing tablet output as a standard. Furthermore, the framerate and processing time were measured.
iii. The Root Mean Squared Error (RMSE) between the proposed Savitzky-Golay smoothening technique and the ground truth reading from the tablet was evaluated. Then it was compared with error values attained by Moving Average and Kalman filters.
iv. The proposed system uses UR5 collaborative robot because of the inherent safety features. The behavior of the internal controller was evaluated. Then the PID controller is tuned to meet the real-time requirements. The controller performance was then evaluated and compared with the original controller parameters in term of rise time.

1.5 Thesis Layout

The thesis outline consists of five chapters. Chapter one gives a general overview about the significance of study, followed by the problem statement, objectives and scope of the study. Chapter two delivers a review of related concepts and principles on 3D tracking, stereo vision processing and robot control. It also covers an inclusive review of literature relevant to the present study including the advantages and disadvantages of methods followed. Chapter three presents the details of the experimental setup and explains the implemented methodologies and evaluation procedures. Chapter four articulates the important experimental results, discusses and analyze the achieved results. Chapter five reports major research findings along with recommendations for future improvements.

REFERENCES

Aksoy, Y., Aydin, T. O., Pollefeys, M., & Smolić, A. (2016). Interactive high-quality green-screen keying via color unmixing. *ACM Transactions on Graphics (TOG), 36*, 1.

Andersen, R. S. (2018). Kinematics of a UR5. *Aalborg University: Aalborg, Denmark*.

Araki, M. (2010). Control Systems, Robotics and Automation—vol II—PID Control. *Kyoto University, Japan*, 58-79.

Arango, C. A., Martínez, J. R., & Pérez, V. Z. (2013). Master-slave system using Kinect and an industrial robot for teleoperations. *2013 Pan American Health Care Exchanges (PAHCE)*, (pp. 1–6).

Bandera, J. P., Marfil, R., Molina-Tanco, L., Rodríguez, J. A., Bandera, A., & Sandoval, F. (2006). Robot learning of upper-body human motion by active imitation. *2006 6th IEEE-RAS International Conference on Humanoid Robots*, (pp. 314–320).

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). Surf: Speeded up robust features. *European conference on computer vision*, (pp. 404–417).

Benois-Pineau, J., & Le Callet, P. (2017). Querying Multiple Simultaneous Video Streams with 3D Interest Maps. In A. Carlier, L. Calvet, P. Gurdjos, V. Charvillat, & W. Ooi, *Visual Content Indexing and Retrieval with Psycho-Visual Models* (pp. 125–144). Springer.

Bergmann, D. (1995). New approach for automatic surface reconstruction with coded light. *Remote sensing and Reconstruction for three-dimensional objects and scenes*, *2572*, pp. 2–9.

Biagiotti, L., & Melchiorri, C. (2010). B-spline based filters for multi-point trajectories planning. *2010 IEEE International Conference on Robotics and Automation*, (pp. 3065–3070).

Biagiotti, L., & Melchiorri, C. (2013). Online trajectory planning and filtering for robotic applications via b-spline smoothing filters. *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, (pp. 5668–5673).

Bishop, G., Welch, G., & others. (2001). An introduction to the kalman filter. *Proc of SIGGRAPH, Course, 8*, 41.

Boyer, K. L., & Kak, A. C. (1987). Color-encoded structured light for rapid active ranging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14–28.

Bradski, G., & Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV library.* " O'Reilly Media, Inc.".

Bunke, H., Von Siebenthal, T., Yamasaki, T., & Schenkel, M. (1999). Online handwriting data acquisition using a video camera. *Proceedings of the Fifth International Conference on Document Analysis and Recognition. ICDAR'99 (Cat. No. PR00318)*, (pp. 573–576).

Burger, W. (2016). Zhang's camera calibration algorithm: in-depth tutorial and implementation. *HGB16-05*, 1–6.

Carreras, M., Ridao, P., García, R., & Nicosevici, T. (2003). Vision-based localization of an underwater robot in a structured environment. *2003 IEEE International Conference on Robotics and Automation (Cat. No. 03CH37422)*, *1*, pp. 971–976.

Carrihill, B., & Hummel, R. (1985). Experiments with the intensity ratio depth sensor. *Computer vision, graphics, and image processing, 32*, 337–358.

Charbon, E., Fishburn, M., Walker, R., Henderson, R. K., & Niclass, C. (2013). SPAD-based sensors. In *TOF range-imaging cameras* (pp. 11–38). Springer.

Chen, C.-S., Hung, Y.-P., Chiang, C.-C., & Wu, J.-L. (1997). Range data acquisition using color structured lighting and stereo vision. *Image and vision computing, 15*, 445–456.

Chui, C. K. (2017). *Kalman filtering.* Springer International Publishing.

Cohen, G. H., & Coon, G. (1953). Theoretical Considerations of Retarded Control" jul. 1953. *Transactions of the ASME*, 827–834.

Copot, C., Muresan, C., Ionescu, C.-M., Vanlanduit, S., & De Keyser, R. (2018). Calibration of UR10 robot controller through simple auto-tuning approach. *Robotics, 7*, 35.

Craig, J. J. (2009). *Introduction to robotics: mechanics and control, 3/E.* Pearson Education India.

Cui, J., Tosunoglu, S., Roberts, R., Moore, C., & Repperger, D. W. (2003). A review of teleoperation system control. *Proceedings of the Florida Conference on Recent Advances in Robotics*, (pp. 1–12).

De Azevedo, H. R., Branodao, S. F., & Alves, J. D. (1993). A fuzzy logic controller for DC motor position control. *Proceedings of IEEE 2nd International Workshop on Emerging Technologies and Factory Automation (ETFA'93)*, (pp. 18–27).

De Villiers, J. P., Leuschner, F. W., & Geldenhuys, R. (2008). Centi-pixel accurate real-time inverse distortion correction. *Optomechatronic Technologies 2008*, *7266*, p. 726611.

Dimond, T. L. (1957). Devices for reading handwritten characters. *Papers and discussions presented at the December 9-13, 1957, eastern joint computer conference: Computers with deadlines to meet*, (pp. 232–237).

Du, G., & Zhang, P. (2014). A markerless human–robot interface using particle filter and Kalman filter for dual robots. *IEEE Transactions on Industrial Electronics, 62*, 2257–2264.

Du, G., Zhang, P., Mai, J., & Li, Z. (2012). Markerless kinect-based hand tracking for robot teleoperation. *International Journal of Advanced Robotic Systems, 9*, 36.

Du, G.-L., Zhang, P., Yang, L.-Y., & Su, Y.-B. (2010). Robot teleoperation using a vision-based manipulation method. *2010 International Conference on Audio, Language and Image Processing*, (pp. 945–949).

Feng, G. (1998). Data smoothing by cubic spline filters. *IEEE Transactions on Signal Processing, 46*, 2790–2796.

Feng, X., Cao, M., Wang, H., & Collier, M. (2008). The comparison of camera calibration methods based on structured-light measurement. *2008 Congress on Image and Signal Processing*, *2*, pp. 155–160.

Fink, G. A., Wienecke, M., & Sagerer, G. (2001). Video-based on-line handwriting recognition. *Proceedings of Sixth International Conference on Document Analysis and Recognition*, (pp. 226–230).

Gee, T., Li, W., Delmas, P., Friedrich, H., & Gimel'farb, G. (2017). Distortion estimation to improve tsai calibration for cameras with large optical distortion. *2017 International Conference on Image and Vision Computing New Zealand (IVCNZ)* (pp. 1-6). IEEE.

Govaers, F. (2019). *Introduction and Implementations of the Kalman Filter.* IntechOpen. doi:https://doi.org/10.5772/intechopen.80600

Guiñón, J. L., Ortega, E., García-Antón, J., & Pérez-Herranz, V. (2007). Moving average and Savitzki-Golay smoothing filters using Mathcad. *Papers ICEE, 2007*.

Hansard, M., Lee, S., Choi, O., & Horaud, R. P. (2012). Characterization of Time-of-Flight Data. In M. Hansard, S. Lee, O. Choi, & R. P. Horaud, *Time-of-flight cameras: principles, methods and applications* (pp. 1-28). Springer Science & Business Media.

Hawkins, J. K., & Munsey, C. J. (1963). A parallel computer organization and mechanizations. *IEEE Transactions on Electronic Computers*, 251–262.

Hawkins, K. P. (2013). *Analytic inverse kinematics for the universal robots UR-5/UR-10 arms.* Tech. rep., Georgia Institute of Technology.

Hu, H., Li, J., Xie, Z., Wang, B., Liu, H., & Hirzinger, G. (2005). A robot arm/hand teleoperation system with telepresence and shared control. *Proceedings, 2005 IEEE/ASME International Conference on Advanced Intelligent Mechatronics.*, (pp. 1312–1317).

Hung, P. N., & Yoshimi, T. (2016). An approach to learn hand movements for robot actions from human demonstrations. *2016 IEEE/SICE International Symposium on System Integration (SII)*, (pp. 711–716).

Hwang, S., Chou, J., & others. (1994). Comparison on fuzzy logic and PID controls for a DC motor position controller. *Proceedings of 1994 IEEE Industry Applications Society Annual Meeting*, *2*, pp. 1930–1935.

Incremona, G., Ferrara, A., & Magni, L. (2017). MPC for robot manipulators with integral sliding modes generation. *IEEE/ASME Transactions on Mechatronics. 22*, pp. 1299-1307. IEEE.

Jain, R., Kasturi, R., & Schunck, B. G. (1995). *Machine vision* (Vol. 5). McGraw-hill New York.

Jędrasiak, K., Janik, Ł., Polański, A., & Wojciechowski, K. (2011). Vicon Motion Capture and HD 1080 Standard Video Data Fusion Based on Minimized Markers Reprojection Error. In *Image Processing and Communications Challenges 3* (pp. 209–216). Springer.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME--Journal of Basic Engineering, 82*, 35--45.

Kamberova, G., & Bajcsy, R. (1998). Sensor errors and the uncertainties in stereo reconstruction. *Empirical Evaluation Techniques in Computer Vision*, 96-16.

Karbasi, M., Bilal, S., Aghababaeyan, R., Rad, A. E., Bhatti, Z., & Shah, A. (2016). Analysis and Enhancement of the Denoising Depth Data Using Kinect Through Iterative Technique. *Jurnal Teknologi, 78*.

Keating, R. (2016). Ur5 inverse kinematics. *Slideshare. net. https://www. slideshare. net/RyanKeating13/ur5-ik[Accessed 13 Dec. 2018]*.

Kebria, P. M., Al-Wais, S., Abdi, H., & Nahavandi, S. (2016). Kinematic and dynamic modelling of UR5 manipulator. *2016 IEEE international conference on systems, man, and cybernetics (SMC)*, (pp. 004229–004234).

Khoshelham, K., & Elberink, S. O. (2012). Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors, 12*, 1437–1454.

Kim, C., Chiu, P., & Oda, H. (2017). Capturing Handwritten Ink Strokes with a Fast Video Camera. *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, *1*, pp. 1269–1274.

Kofman, J., Verma, S., & Wu, X. (2007). Robot-manipulator teleoperation by markerless vision-based hand-arm tracking. *International Journal of Optomechatronics, 1*, 331–357.

Kofman, J., Wu, X., Luu, T. J., & Verma, S. (2005). Teleoperation of a robot manipulator using a vision-based human-robot interface. *IEEE transactions on industrial electronics, 52*, 1206–1219.

Kuo, Y.-L., & Liu, S.-M. (2015). Position Control of a Serial Manipulator Using Fuzzy-PID Controllers. *International Journal of Automation and Smart Technology, 5*, 18–26.

Kurpjuhn, T. P., Nickels, K., Hauck, A., & Hutchinson, S. (1999). Development of a visual space-mouse. *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*, *4*, pp. 2527–2532.

Kuznetsov, Y., Sdobnov, A., Meglinski, I., Harmelin, A., & Kalchenko, V. (2019). Evaluation of handwriting peculiarities utilizing laser speckle contrast imaging. *Laser Physics Letters, 16*, 115601.

Li, L. (2014). Time-of-flight camera–an introduction. *Technical white paper*.

Loop, C., & Zhang, Z. (1999). Computing rectifying homographies for stereo vision. *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, *1*, pp. 125–131.

Mathworks. (1 6, 2021). *What Is Camera Calibration? - MATLAB & Simulink*. Retrieved from Mathworks : https://www.mathworks.com/help/vision/ug/camera-calibration.html

Medina, L., & Wykes, C. (2001). Multiple target 3D location airborne ultrasonic system. *Ultrasonics, 39*, 19–25.

Mendoza, R. H., Kagami, S., & Hashimoto, K. (2010). Online 3-d trajectory estimation of a flying object from a monocular image sequence for catching. *Visual Servoing*, 121.

Mishra, M. K., Kothari, A. G., Kothari, D. P., & Ghosh, A. (1998). Development of a fuzzy logic controller for servo systems. *Proceedings of IEEE TENCON'98. IEEE Region 10 International Conference on Global Connectivity in Energy, Computer, Communication and Control (Cat. No. 98CH36229)*, *1*, pp. 204–207.

Mitsantisuk, C., Piriyanont, B., & Ohishi, K. (2018). Haptic signal processing for human-robot collaboration system using moving average filter. *2018 IEEE 15th International Workshop on Advanced Motion Control (AMC)*, (pp. 675–680).

Moriya, M., Hayashi, T., Tominaga, H., & Yamasaki, T. (2005). Video Tablet Based on Stereo Camera-Human-friendly handwritten capturing system for educational use. *Fifth IEEE International Conference on Advanced Learning Technologies (ICALT'05)*, (pp. 909–911).

Munich, M. E., & Perona, P. (2002). Visual input for pen-based computers. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*, 313–328.

Naidu, V. P. (2008). Evaluation of Acceleration and Jerk Models in Radar and IRST Data Fusion for Tracking Evasive Maneuvering Target. *46th AIAA Aerospace Sciences Meeting and Exhibit*, (p. 877).

Nowak, P., Białek, M., Lindner, T., & Wyrwał, D. (2020). Manipulator teleoperation by using the position of the hand registered by vision system. *2020 International Conference Mechatronic Systems and Materials (MSM)*, (pp. 1–5).

Orfanidis, S. J. (2010). Introduction to signal processing, Sophocles J. *Orfanidis, Rutgers University, 13*, 795.

Park, S., Yu, S., Kim, J., Kim, S., & Lee, S. (2012). 3D hand tracking using Kalman filter in depth space. *EURASIP Journal on Advances in Signal Processing, 2012*, 1–18.

Persson, P.-O., & Strang, G. (2003). Smoothing by savitzky-golay and legendre filters. In *Mathematical systems theory in biology, communications, computation, and finance* (pp. 301–315). Springer.

Poroykov, A., Kalugin, P., Shitov, S., & Lapitskaya, I. (2020). Modeling ArUco markers images for accuracy analysis of their 3D pose estimation. *30th International Conference on Computer Graphics and Machine Vision (GraphiCon 2020). 2.* CEUR Workshop Proceedings.

Qian, K., Niu, J., & Yang, H. (2013). Developing a gesture based remote human-robot interaction system using kinect. *International Journal of Smart Home, 7*, 203–208.

Ramos, E. (2012). Kinect Basics. In *Arduino and Kinect Projects* (pp. 23–34). Springer.

Richard Hartley, A. Z. (2005). Multiple view geometry in computer vision. *Multiple view geometry in computer vision*. (C. U. Press, Ed.) Cambridge University Press.

Rogalla, O., Ehrenmann, M., Zollner, R., Becher, R., & Dillmann, R. (2002). Using gesture and speech control for commanding a robot assistant. *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication*, (pp. 454–459).

Salvi, J., Armangué, X., & Batlle, J. (2002). A comparative review of camera calibrating methods with accuracy evaluation. *Pattern recognition, 35*, 1617–1635.

Sanders, J., & Kandrot, E. (2010). *CUDA by example: an introduction to general-purpose GPU programming.* Addison-Wesley Professional.

Sato, K., & Inokuchi, S. (1984). Video Tablet-2D Coordinate Input Device with CCD Camera. *Trans. IEICE Japan, 67*, 726–727.

Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry, 36*, 1627–1639.

Schafer, R. W. (2011). What is a Savitzky-Golay filter?[lecture notes]. *IEEE Signal processing magazine, 28*, 111–117.

Sebastian, P., Voon, Y. V., & Comley, R. (2010). Colour space effect on tracking in video surveillance. *International Journal on Electrical Engineering and Informatics, 2*, 298-312.

Seok, J.-H., Levasseur, S., Kim, K.-E., & Kim, J. (2008). Tracing handwriting on paper document under video camera. *International Conference on Frontiers in Handwriting Recognition*, (pp. 109-114).

Small, S., Nyarko, K., & Scott, C. (2008). An Analysis of Smoothing Algorithms on Signal Jitter in Optical Tracking Data. *AIAA Guidance, Navigation and Control Conference and Exhibit* (pp. 6640-6649). Honolulu: Aerospace Research Central.

Teke, B., Lanz, M., Kämäräinen, J.-K., & Hietanen, A. (2018). Real-time and Robust Collaborative Robot Motion Control with Microsoft Kinect® v2. *2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, (pp. 1–6).

Thacker, N. A., Clark, A. F., Barron, J. L., Beveridge, J. R., Courtney, P., Crum, W. R., Clark, C. (2008). Performance characterization in computer vision: A guide to best practices. *Computer vision and image understanding, 109*(3), 305-334.

Tsai, R. (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal on Robotics and Automation, 3*, 323–344.

Universal Robots. (8 1, 2013). *Calibration manual.*

Universal Robots. (2014). *User Manual - UR5/CB3.*

Universal Robots. (15 4, 2016). *URScript API Reference.*

Universal Robots. (10 June, 2021). *Forward and inverse kinematics*. Retrieved from compas_fab - 3.3. Forward and inverse kinematics: https://gramaziokohler.github.io/compas_fab/0.15.0/examples/03_backends_ros/03_forward_and_inverse_kinematics.html

Universal robots. (1 6, 2021). *What is a singularity?* Retrieved from Universal robots: https://www.universal-robots.com/articles/ur/application-installation/what-is-a-singularity/

Waldherr, S., Romero, R., & Thrun, S. (2000). A gesture based interface for human-robot interaction. *Autonomous Robots, 9*, 151–173.

Wang, H., Zhang, N., Créput, J.-C., Ruichek, Y., & Moreau, J. (2016). Massively parallel GPU computing for fast stereo correspondence algorithms. *Journal of Systems Architecture, 65*, 46–58.

Wang, Y., Sun, F., Hu, L., Li, Z., & Liu, H. (2009). Space teleoperation with large time delay based on vision feedback and virtual reality. *2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, (pp. 1200–1205).

Xi, B., Liu, N., Lu, T., Lu, J., Cai, Y., & Wang, S. (2017). Robot teleoperation system based on SVDD. *2017 Chinese Automation Congress (CAC)*, (pp. 2830–2835).

Xu, Y., Zhong, J., Ding, M., Chen, H., & Chen, S. (2013). The acquisition and processing of real-time information for height tracking of robotic GTAW process by arc sensor. *The International Journal of Advanced Manufacturing Technology, 65*, 1031–1043.

Yamasaki, T., & Hattori, T. (1996). A new data tablet system for handwriting characters and drawing based on the image processing. *1996 IEEE International Conference on Systems, Man and Cybernetics. Information Intelligence and Systems (Cat. No. 96CH35929)*, *1*, pp. 428–431.

Yang, C., Chang, S., Liang, P., Li, Z., & Su, C.-Y. (2015). Teleoperated robot writing using EMG signals. *2015 IEEE International Conference on Information and Automation*, (pp. 2264–2269).

Yasin, S. Y. (2002). Systematic methods for the design of a class of fuzzy logic controllers. *Dissertation*. Western Michigan University.

Zanuttigh, P., Marin, G., Dal Mutto, C., Dominio, F., Minto, L., & Cortelazzo, G. M. (2016). Operating principles of structured light depth cameras. In *Time-of-flight and structured light depth cameras* (pp. 43–79). Springer.

Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence, 22*, 1330–1334.

Ziegler, J. G., Nichols, N. B., & others. (1993). Optimum settings for automatic controllers. *Journal of Dynamic Systems, Measurement, and Control, 115*(2B), 220-222.