Saif, A. F.M.Saifuddin; Mahayuddin, Zainal Rasyid

Stereo Vision Based Localization of Handheld Controller in Virtual Reality for 3D Painting Using Inertial System

Published in:
Journal of Image and Graphics

DOI:
10.18178/joig.11.2.127-131

Published: 01/06/2023

Document Version
Publisher's PDF, also known as Version of record

Stereo Vision Based Localization of Handheld Controller in Virtual Reality for 3D Painting Using Inertial System

A F M Saifuddin Saif [1,*] and Zainal Rasyid Mahayuddin [2]

[1] School of Engineering, Aalto University, Espoo, Finland
[2] Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor, Malaysia; Email: zainalr@ukm.edu.my (Z.R.M.)
*Correspondence: a.f.saif@aalto.fi (A.F.M.S.S.)

Abstract—**Google Tilt Brush is expensive for virtual drawing which needs further improvement on the functionalities of mechanisms rather than implementation expects addressed in this research. Several issues are addressed by this research in this context, i.e., noise removal from sensor data, double integration-based drift issues and cost. Recently, available smart phones do not have the ability to perform drawing within artificial settings handling cardboard and daydream of google without purchasing Oculus Rift and HTC Vive (Virtual Reality Headset) because of expensiveness for large number of users. In addition, various extrinsic hardwares, i.e., satellite localization hardware and ultrasonic localization applications are not used for drawing in virtual reality. Proposed methodology implemented extended Kalman filter and Butterworth filter to perform positioning using six degree of freedom using Microelectromechanical Applications (MEMS) software data. A stereo visual method using Simultaneous Localization and Mapping (SLAM) is used to estimate the measurement for positioning implicating mobile phone (i.e., android platform) for the hardware system to estimate drift. This research implemented Google Virtual Reality application settings Kit with Unity3D engine. Experimentation validation states that proposed method can perform painting using virtual reality hardware integrated with controller software implicating smartphone mobile without using extrinsic controller device, i.e., Oculus Rift and HTC Vive with satisfactory accuracy.**

Keywords—**stereo vision, extended Kalman filter, 3D painting, virtual reality**

I. Introduction

With the recent advancement of technology transferable, strength, user friendliness and low expensive impacts hugely in modern day to day life. Recent progress of computer vision and machine learning makes it possible to imagine augmented and virtual reality technology with which individual can see more than others see, listen more than others can listen, and touch, smell in lieu with taste things that others get it difficult [1]. Augmented reality and virtual reality have made it possible to achieve computation elements and objects within actual earth flavor adding whole creates in lieu with structures to include everyday activities at the same time communicating without consciousness by using lack of gestures and verbal communication [2]. Various progress was done before within machine vision and image processing by using the platform of mobile and transferable systems based on android phones, tab, individual personal assistant [3]. Various unique techniques such as user friendliness, low expense of hardwares, number of platforms and effectiveness of operative systems, internet connectivity of firmware updates, comfortability of user interface are responsible for opting into these devices. Strengthen of these devices are accelerated with the advancement of half conductor materials and combined electronics materials.

Virtual Reality (VR) can be addressed as potential computer vision-based technology indicates a prominent process for communicating with growing modern community and is often considered as cumulative tools for upcoming fertile research to boost researcher for harvesting artificial experience a world beyond reality [4]. The usage of VR is to deliver data sources to our emotion, vision, listening and gestures. However, emotion and vision have received understandably high concentration with VR integrated with computer vision and machine learning, displaces for sign, hearing and touch received less attention by the researchers which needs to be addressed adequately to fulfil the demand of upcoming computer vision-based machine learning advancement [5]. For this reason, VR can be recognized as an emerging technology for today's smart phone to embrace very new and exciting tool for human computer interaction.

This research developed several applications in this regard using the advantage of accelerometer, gyroscope, magnetometer sensors of the smartphone. A controller application is developed which works by sending the real-world coordinates of the smartphone to the painting application through Wi-Fi using User Datagram Protocol (UDP). Besides, a panting application is developed to perform painting in open space by providing the

Manuscript received December 5, 2022; revised March 21, 2023; accepted May 8, 2023.

doi: 10.18178/joig.11.2.127-131

capability to be visible via virtual reality headset. In the overall methodology, noise is removed using signal processing. Overall methodology provides the novelty by giving prospective cost-effective approach in comparison to the already established ones rather than not using internal base station. Section II illustrates comprehensive and critical review of previous research, Section III depicts proposed methodology, Section IV demonstrates experimental results and lastly Section V illustrates presents final summary and comments.

II. Literature Review

Previously, various sensors were used for accomplishing various tasks. For tracking human motion, miniature intrinsic materials using biomechanical frameworks and sensor integration methodologies were used [1]. For this sensor, cost effective system did not require external camera, emitter, marker but rather was able to track all types of motion, i.e., running, jumping, crawling. However, miniature inertial sensor is hardware and software specific to process data. In addition, miniature inertial sensor cannot track any motion which is not compatible with the laws of physics. Miniature 6-DOF implicating Head Mounted Displays (HMDs) based head positioning application changes in direction and location merging the outputs of gyroscope and accelerometer sensors [2]. Several benefits were monitored in these sensors, i.e., low cost, low jitter, high update rates and low latency whereas these advantages could not give satisfactory performance due to reasonably small tracking area. Fusion of sensors using duration integrated architecture able to estimate deficiencies to perform manipulation of smooth position and direction data [3, 4]. Previously, sensor fusion was used for various purposes, i.e., indoor positioning in android devices [5], automotive vehicle positioning using Global Positioning System (GPS) [6], mobile robot positioning using accelerometer [7] etc. Although, sensor fusion can reduce random noise and serve low cost, random bias drift problems are common issue. Determination of position with orientation precisely, Tri-axial accelerometer can be useful in lieu of cost effective and provides impressive progress in tangential sensitivity [8]. However, tangential sensitivity in unexpected shape. For this, Gyroscope, Accelerometer, and magnetometer can be combined together for 3D knee kinematics which provides low error percentage results [9] and to develop error propagation equations [10]. Besides, combination of these sensors can form wrist worn application to recover 3D user hand pose [11]. However, combination of sensor's approach is expensive for implementation. For accurate navigation and orientation direction, approach such as integrated electronic devices containing accelerometers, gyroscopes and magnetometers incorporated Inertial Measurement Units (IMU) were previously implemented [12]. However, this approach provides drift error causes huge comparison for real position and application estimated localization. MEMS based low-cost system reduces drift error estimating towards gyroscope and accelerometer without magnetometer [13]. Previously, IMU was also used for gait analysis [14]. In this context, miniature inertial/magnetometer package was combined without wire to PDA for tracking pedestrian position depending on GPS signal [15]. Although, stride length independence without the requirement of calibration is the main advantage for that methodology, overall approach should be able to be impacted by low number of errors using effective architectures or constructions. For pedestrian localization in indoor environment, previously foot mounted inertial sensors were used where the advantage was less expense and unsignificant shape for the sensor required dedicated experimental space and walking range [16]. In that approach, high accuracy for position estimation was obtained using shift approach [17]. Kalman filtering is significant approach for smooth detection method which has several advantages [18, 19]. Kalman filtering needs smaller memory allocation for being appropriate for various situations and integrated systems and suitable form of real time processing. However, extended Kalman filter initializes final error level worse. In this context, one dimensional Kalman filter can describe actual structure for validation information through connecting with Wiener filter and GPS [20]. Accelerometer can be a good assistance for implementing Kalman filter for significant noise reduction in lieu of localization of weight using bonded loop structure [21]. For tracking human operators effectively, extended Kalman filter integration methodology in IMU/UWB should be able to be impacted option with high data rates and computation resources [22]. For global pose estimation, Kalman filtering was previously used implicating multi sensor fusion [23]. In this approach, correction of distortion for 3-axis magnetic compass assists accuracy and robustness of pose estimation. Kalman filter was also used by inertial head tracker sensor fusion [24]. Extended Kalman filter which is a nonlinear approach Kalman filter should be able to linearize the manipulation of present covariance and mean [25]. Although, extended Kalman filter provides low accurate measurement of covariance, extended Kalman filter assists to construct indoor localization implicating IMU/magnetometer [26] which provides faster computation time using accelerometer sensor.

III. Proposed Research Methodology

Raw sensor data was achieved from android sensor manage API by using controller device. Rotation matrix for the controller phone's rotation in arbitrary axis was received from virtual rotation sensor mentioned in Fig. 1. Android sensor manager was used for linear acceleration with invert rotation matrix vector in earth's coordinate system. Extended Kalman Filter (EKF) was used to achieve smooth three axial accelerations with constant bias for reduced noise. Optimum averaging factor for each consequent state from past states are found using Extended Kalman Filter. 3-axial acceleration with constant bias was received after filtering noisy data using extended Kalman filter. Bias can be estimated using long term average while the controller phone is in stationary position and subtract it from respective axes of acceleration. Linear acceleration was integrated to estimate velocity and integrated again over time to estimate displacement. Dead reckoning approach for accumulating positional information was followed because bias of accelerometer prolonged the usage of sensor's temperature errors and contributes to drift in position.

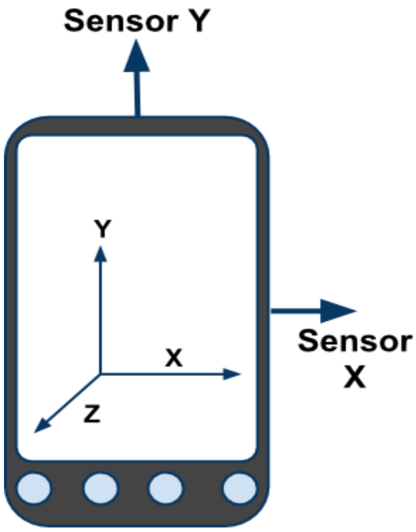

Figure 1. Virtual rotation axis from controller device.

This research implicated stereo SLAM [27, 28] for estimating artificial odometry of controlled materials from artificial virtual environment. Stereo SLAM builds and navigates map using images from more than one camera from tracking relative position and orientation [29, 30]. Good features between frames were tracked using Lucas Kanade sparse optical flow [31] for estimating essential matrix and rotation with translation information where dependency on external positioning system was reduced. Proposed controller uses dedicated button like mouse button or brush pressure where proposed application connects VR headset and controller smartphone using local Wi-Fi network. Protocol such as UDP socket protocol is used to transmit the location co-ordinate vector with timeframe followed by the collection of sensor data. Datagram was collected in constant frequency in order to paint in the air for experiencing in virtual reality. Proposed methodology used android Software Development Kit (SDK) to build controller application where TensorFlow was used for image processing, Unity3D for developing virtual reality environment.

IV. Experimental Results and Discussion

Proposed method involves several RGB cameras for several corners which are calibrated beforehand. Image resolution was to 700 × 500 pixels running at 20 frames/s. Gravity measurement is separately achieved from gravity hardware. Vision hardwares are integrated by implicating self-integration tools**.** Several reflective marks with diameter of 1.3 cm are placed in appropriate position and these indicators are not chosen as nonrigid body. Estimated pose estimate achievement summaries not for the non-rigid surrounding applications which contains axes for the accelerometer with offset from the center for camera coordinates corrected during comparison. For precise estimation, synchronized data was used with offline sequences recorded at fixed frame rate and accelerometer data. To enable synchronization, user holds the device horizontal and vertically at the beginning of the test, then moves it up and down several times. Performance of the proposed method is combined with adaptive thresholding approach using adaptive filtering to achieve precise accuracy and processing time. This research observed constant time domain for white noise propagation in raw linear acceleration shown in Fig. 2. In this context, controller device was lying in flat position.

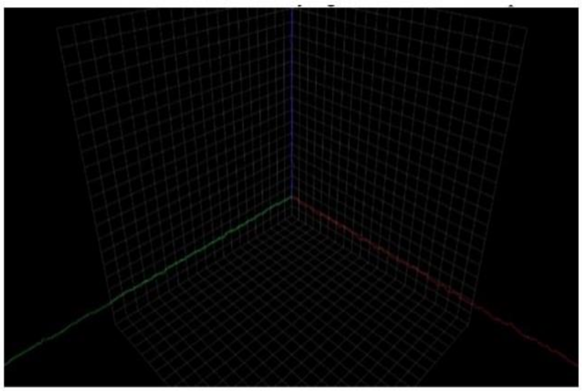

Figure 2. Raw accelerometer data.

Extended Kalman filter relatively smooth data shown in Fig. 3. Multiple values were set during construction of Extended Kalman filter, i.e., expectation value (0.005), process noise (1e-8) and measurement noise (0.001). linear acceleration of X, Y and Z- axis was delivered to extended Kalman filter from Data to transfer noise data for accurate information. Several biasness was found when accelerometer data were sent during filtration. Location of the mobile device was achieved by estimating the filter or smooth data by double integration. However, this strategy resulted rise in drift shown in Fig. 4. In addition, drift error was the resultant of biasness. Bias estimation was done before double integration for reducing drift.

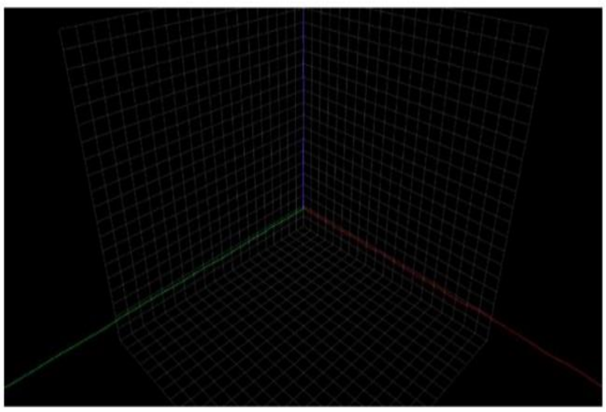

Figure 3. Kalmanfilter for noise reduction.

Flicker noise or 1/f noise was found based on MEMS inertial measurement sensors shown in Fig. 5. To avoid the contribution on velocity during the negligible movement of device, linear acceleration was passed through order 7 Butterworth high pass filter during the negligible movement of device. Small values of linear acceleration were attenuated by Butterworth filter and overall simulation drawing is shown in Fig. 6.

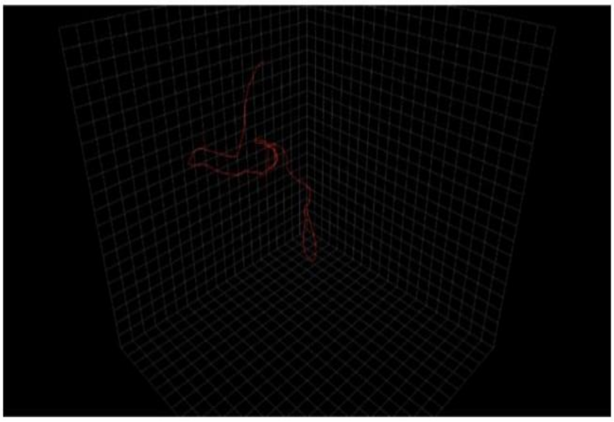

Figure 4. Drift of position in flat platform.

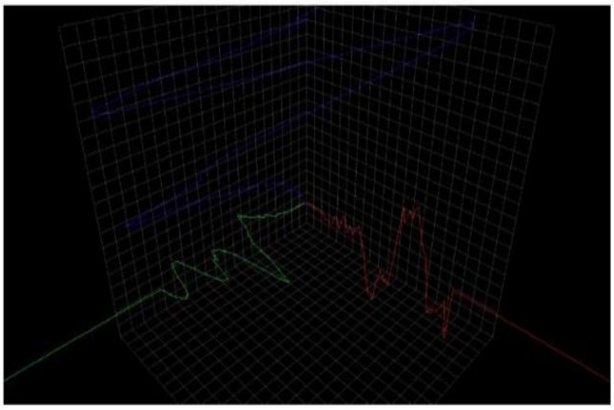

Figure 5. Accelerometer movement introduces flicker noise.

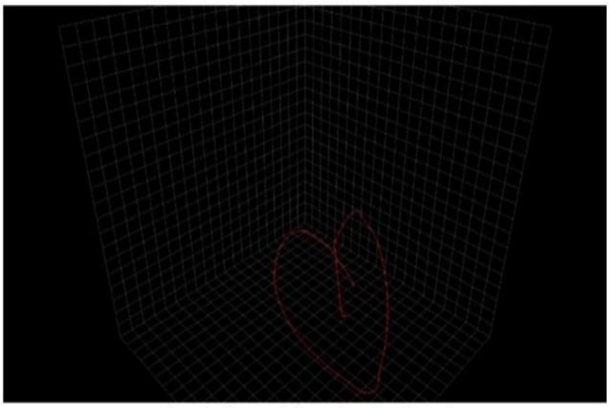

Figure 6. Simulation of drawing in the air.

In this research, stereo artificial intrinsic application was embedded with accelerometer for reducing flicker noise using significant velocity information. Features map was constructed using stereo images using PTAM or ORB-SLAM [28, 32]. Camera poses estimation was done using structure from motion in lieu of trajectory and achieved stable 3D coordinate for controller device. In addition, residual noises were also removed from IMU sensors [33–38].

V. Conclusion and Future Work

This research illustrates critical and comprehensive investigation to paint in virtual reality based on stereo vision instead of using extra controller devices. Validation was performed using Inertial Measurement Unit (IMU) where stage manipulation for acceleration, localization and shape were done using gyroscope, accelerometer and magnetometer applications in lieu of extended Kalman filter. Proposed methodology enables the ability to get room scale canvas by introducing validated experimental experience among users to perform drawing in open air using stereo vision. In future, this research will investigate stereo odometry further to estimate the coefficients for IMU given localization and artificial position in earth co-ordinate application. In this context, Unity with Google Virtual Reality kits based on stereo vision will be implicated to paint in virtual reality will be prime focus in future.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

A. F. M. Saifuddin Saif conducted the research; Zainal Rasyid Mahayuddin analyzed the data; A. F. M. Saifuddin Saif prepared the paper; all authors had approved the final version.

Funding

This work is supported by Universiti Kebangsaan Malaysia under the "Geran Universiti Penyelidikan" research grant, GUP-2020-064.

References

[1] D. Roetenberg, H. Luinge, and P. Slycke, "Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors," *Xsens Motion Technologies BV, Tech. Rep,* vol. 1, pp. 1–7, 2009.

[2] E. M. Foxlin, M. Harrington, and Y. Altshuler, "Miniature six-DOF inertial system for tracking HMDs," in *Helmet-and Head-Mounted Displays III*, SPIE, 1998, vol. 3362, pp. 214–228.

[3] W. Elmenreich and S. Pitzek, "Using sensor fusion in a time-triggered network," in *Proc. IECON'01. 27th Annual Conference of the IEEE Industrial Electronics Society (Cat. No. 37243)*, IEEE, 2001, vol. 1, pp. 369–374.

[4] W. Elmenreich, "Sensor fusion in time-triggered systems," *Citeseer*, 2002.

[5] U. Shala and A. Rodriguez. (2011). Indoor positioning using senso r-fusion in android devices. [Online]. Available: https://www.diva-portal.org/smash/get/diva2:475619/FULLTEXT02.pdf

[6] N. Magnusson and T. Odenman. (2012). Improving absolute position estimates of an automotive vehicle using GPS in sensor fusion. [Online]. Available: https://publications.lib.chalmers.se/rec ords/fulltext/159412.pdf

[7] H. H. Liu and G. K. Pang, "Accelerometer for mobile robot positioning," *IEEE Transactions on Industry Applications*, vol. 37, no. 3, pp. 812–819, 2001.

[8] P. Axelsson and M. Norrlöf, "Method to estimate the position and orientation of a triaxial accelerometer mounted to an industrial manipulator," in *Proc. IFAC*, 2012, vol. 45, no. 22, pp. 283–288.
[9] K. Kawano, S. Kobashi, M. Yagi, K. Kondo, S. Yoshiya, and Y. Hata, "Analyzing 3D knee kinematics using accelerometers, gyroscopes and magnetometers," in *Proc. 2007 IEEE International Conference on System of Systems Engineering*, IEEE, 2007, pp. 1–6.
[10] J. N. Sanders-Reed, "Error propagation in two-sensor three-dimensional position estimation," *Optical Engineering,* vol. 40, no. 4, pp. 627–636, 2001.
[11] D. Kim, *et al.*, "Digits: Freehand 3D interactions anywhere using a wrist-worn gloveless sensor," in *Proc. of the 25th Annual ACM Symposium on User Interface Software and Technology*, 2012, pp. 167–176.
[12] M. A. Brodie, A. Walmsley, and W. Page, "The static accuracy and calibration of inertial measurement units for 3D orientation," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 11, no. 6, pp. 641–648, 2008, doi: 10.1080/10255840802326736
[13] F. Montorsi, F. Pancaldi, and G. M. Vitetta, "Design and implementation of an inertial navigation system for pedestrians based on a low-cost MEMS IMU," in *Proc. 2013 IEEE International Conference on Communications Workshops (ICC)*, IEEE, 2013, pp. 57–61.
[14] T. Seel, J. Raisch, and T. Schauer, "IMU-based joint angle measurement for gait analysis," *Sensors,* vol. 14, no. 4, pp. 6891–6909, 2014.
[15] E. Foxlin, "Pedestrian tracking with shoe-mounted inertial sensors," *IEEE Computer Graphics and Applications*, vol. 25, no. 6, pp. 38–46, 2005.
[16] O. Woodman and R. Harle, "Pedestrian localisation for indoor environments," in *Proc. of the 10th International Conference on Ubiquitous Computing*, 2008, pp. 114–123.
[17] Y. Chen, S. Zhao, and J. A. Farrell, "Computationally efficient carrier integer ambiguity resolution in multi-epoch GPS/INS: A common-position-shift approach," *IEEE Transactions on Control Systems Technology,* vol. 24, no. 5, pp. 1541–1556, 2015.
[18] E. Brookner, *Tracking and Kalman Filtering Made Easy*, John Wiley & Sons, Inc., 1998.
[19] G. Welch and G. Bishop, *An Introduction to the Kalman Filter*, UNC-Chapel Hill, TR 95-041, July 24, 2006.
[20] R. Turner. The 1d Kalman Filter. [Online]. Available: http://www.gatsby.ucl.ac.uk/~turner/Notes/1DKalmanFilter/1d_kalman_filter.pdf
[21] T. Singhal, A. Harit, and D. Vishwakarma, "Kalman filter implementation on an accelerometer sensor data for three state estimation of a dynamic system," *International Journal of Research in Engineering and Technology*, vol. 1, no. 6, pp. 330–334, 2012.
[22] J. A. Corrales, F. A. Candelas, and F. Torres, "Hybrid tracking of human operators using IMU/UWB data fusion by a Kalman filter," in *Proc. 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2008, IEEE, pp. 193–200.
[23] G. Schall, *et al.*, "Global pose estimation using multi-sensor fusion for outdoor augmented reality," in *Proc. 2009 8th IEEE International Symposium on Mixed and Augmented Reality*, 2009, IEEE, pp. 153–162.
[24] E. Foxlin, "Inertial head-tracker sensor fusion by a complementary separate-bias Kalman filter," in *Proc. of the IEEE 1996 Virtual Reality Annual International Symposium*, IEEE, 1996, pp. 185–194.
[25] A. S. Saif and Z. R. Mahayuddin, *Augmented Reality-Based 3D Human Hands Tracking from Monocular True Images Using Convolutional Neural Network*, Handbook of Research on Artificial Intelligence and Knowledge Management in Asia's Digital Economy*,* 2021.
[26] H. Hellmers, A. Norrdine, J. Blankenbach, and A. Eichhorn, "An IMU/magnetometer-based indoor positioning system using Kalman filtering," in *Proc. International Conference on Indoor Positioning and Indoor Navigation*, IEEE, 2013, pp. 1–9.
[27] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics,* vol. 34, no. 4, pp. 1004–1020, 2018.
[28] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, "ORB-SLAM: A versatile and accurate monocular SLAM system," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.
[29] F. Bonin-Font, A. Ortiz, and G. Oliver, "Visual navigation for mobile robots: A survey," *Journal of Intelligent and Robotic Systems*, vol. 53, no. 3, pp. 263–296, 2008.
[30] C. Burbridge, L. Spacek, J. Condell, and U. Nehmzow, "Monocular omnidirectional vision-based robot localisation and mapping," in *Proc. TAROS 2008—Towards Autonomous Robotic Systems 2008, The University of Edinburgh, United Kingdom,* 2008, pp. 135–141.
[31] N. Nourani-Vatani, P. V. Borges, and J. M. Roberts, "A study of feature extraction algorithms for optical flow tracking," in *Proc. Australasian Conference on Robotics and Automation*, 2012.
[32] G. Klein and D. Murray, "Parallel tracking and mapping for small AR workspaces," in *Proc. 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, IEEE, 2007, pp. 225–234.
[33] X.-D. Nguyen, B.-J. You, and S.-R. Oh, "A simple framework for indoor monocular SLAM," *International Journal of Control, Automation, and Systems*, vol. 6, no. 1, pp. 62–75, 2008.
[34] A. S. Saif, Z. R. Mahayuddin, and H. Arshad. "Vision-based efficient collision avoidance model using distance measurement," in *Vision-Based Efficient Collision Avoidance Model Using Distance Measurement*, A. Ahmadian and S. Salahshour, Eds. Taylor & Francis, 2021, pp. 191–202, doi: https://doi.org/10.1201/9781003138341-12-12
[35] A. S. Saif and Z. R. Mahayuddin, "Moment features based violence action detection using optical flow," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 11, 2020.
[36] A. Singh and K. Venkatesh, "Monocular visual odometry," Undergraduate Project, vol. 2, 2015.
[37] A. S. Saif, Z. R. Mahayuddin, and A. Shapi'i, "Augmented reality based adaptive and collaborative learning methods for improved primary education towards fourth Industrial Revolution (IR 4.0)," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, 2021.
[38] Z. R. Mahayuddin and A. F. M. Saif, "Vision based 3D gesture tracking using augmented reality and virtual reality for improved learning applications," *International Journal of Advanced Computer Science and Applications*, 2021.