Proceedings of CONF-CDS 2025 Symposium: Application of Machine Learning in Engineering
DOI: 10.54254/2755-2721/2025.AST34382

Lidar SLAM and Visual SLAM for Unmanned Aerial Vehicles: A Review of Principles, Applications and Challenges

Jiayi Tang

Department of Mechanical Design, Manufacturing and Automation, Shanxi University of Science and Technology, Xi'an, China
1600690330@qq.com

Abstract. Simultaneous Localization and Mapping (SLAM) is the foundation for autonomous operations of unmanned aerial vehicles (UAV) in environments without GPS. This paper systematically analyzes two main SLAM paradigms - Lidar-based (gapping, Hector SLAM, Cartographer, LeGO-LOAM) and vision-based (monocular, stereo, RGB - D SLAM) - highlighting their principles, advantages and limitations. LiDAR SLAM has an advantage in geometric accuracy (for example, Cartographer achieves ≤0.12 mite in structural inspection), but it faces payload and computational burden. While visual SLAM is cost-effective, it has difficulties in terms of lighting variations and texture-free scenes. Practical applications in electrical inspection, agriculture and exploration have revealed key trade-offs: LiDAR can achieve millimetre-precision modelling of power lines 17, while visual-inertial SLAM (such as ORB-SLAM3) is suitable for real-time processing at 28.36 FPS 810 for crop monitoring. The current challenges include the latency of multi-drone collaboration (for example, c2virl-slam alleviates this problem through UWB-VIO fusion), and the real-time processing requirements of embedded systems (LeGO-LOAM operates at 10Hz on Jetson TX2 37). And robustness in degraded environments (adaptive sensor compensation reduces Lidar error by 97.9% 3). The future directions give priority to heterogeneous fusion (for example, radar-inertial vision systems 9), semantic integration (SLAM-LLM for context understanding 16), and end-to-end learning (reinforcement learning for point cloud navigation 7).

Keywords: Simultaneous Localization and Mapping (SLAM), Unmanned Aerial Vehicles (UAV), LiDAR SLAM, Visual SLAM, Sensor Fusion, Autonomous Navigation, Agricultural Inspection, Multi-UAV Collaboration

1. Introduction

Unmanned aerial vehicles (UAV) have significantly transformed various industries in society by achieving autonomous operation in complex and dynamic environments [1, 2]. A key technology for realizing the autonomy of UAV is Simultaneous Localization and Mapping (SLAM), which enables UAV to construct environmental maps without prior knowledge while determining their own pose [1]. The selection of an appropriate SLAM system directly affects the performance of UAV in terms of accuracy, efficiency, and robustness.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.AST34382

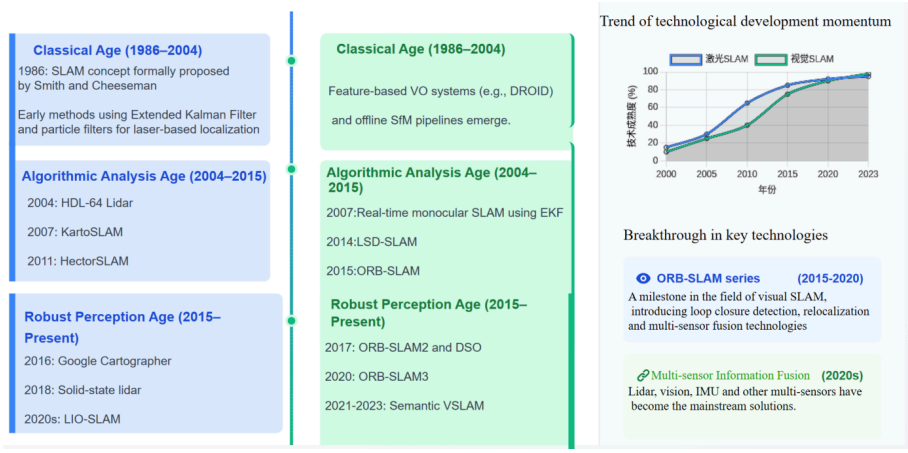


Figure 1. The timeline of the development of LiDAR SLAM and Visual SLAM technologies

This article classifies the SLAM technology adopted on unmanned aerial vehicles (UAV) into LiDAR-based and vision-based methods, and analyzes their core principles, advantages and disadvantages. In particular, by studying the representative algorithms of Gmapping, Hector SLAM, Cartographer and LeGO-LOAM for LiDAR SLAM, as well as monocular, stereo and RGB-D systems for visual SLAM. Practical applications in wire detection, agriculture and exploration are mentioned to show the performance trade-offs in practice. The current problems of multi-UAV cooperation, real-time computing and environmental robustness have been addressed, and future research will explore the integration of end-to-end learning and large language models [3-7].

2. Core principle of SLAM for UAV

2.1. LiDAR SLAM

LiDAR SLAM is based on laser rangefinders that send out laser pulses and determine how long it takes for these pulses to return after hitting objects nearby. Generate a dense point cloud and obtain the full-size 3D model of the scene. The above detailed descriptions are necessary for drone position and map construction under poor visibility or when visual features are sparse, such as in foggy weather or large featureless indoor spaces [8-10].

2.1.1. Gmapping: Rao-Blackwellized particle filter-based map generation for small- scale UAV operations

Gmapping is a well-known algorithm that uses Rao-Blackwellised particle filters [3]. A Particle Filter is used in this way. Each particle is a potential flight attitude of the drone, and each corresponds to an environment map. Laser range data and odometry (if available) are used to update the position and orientation of particles, and a map is optimised [4].

Gmapping's particle-based approach can be used to quickly obtain a relative map in a small area, such as an indoor environment or a small outdoor courtyard, [3]. Rao-Blackwellisation is used to reduce the computational complexity by factoring out the mapping problem from pose estimation. Lightweight drones with limited computing resources are relatively easy to use for GMapping as well because they can be run on low-power embedded systems [3]. However, when the drone is in a large open space and moves, the number of particles required to accurately represent the uncertainty of the posture will increase exponentially [4]. Therefore, there will be fewer particles, and they will

DOI: 10.54254/2755-2721/2025.AST34382

be forced to converge to a local optimum that has accumulated estimation errors. As a result, the accuracy of the generated map will decline with time [3].

2.1.2. Hector SLAM: odometry-free 2D mapping for UAVs with unstable motion models

Hector SLAM is a type that does not need odometry data, which can be noisy or unreliable in a UAV with an unstable motion model [5]. ICP is used to match laser scans, and no other techniques are employed [6].

Hector SLAM does not need odometry, so it has an advantage when a UAV is in an unknown environment with unpredictable motion, such as due to high winds or sudden altitude changes. It is very suitable for resource-constrained UAVs, and even in the event of extremely minimal hardware setup, it can still generate a 2D map efficiently. The attack-type drone has a sudden turn or a rapid rise and fall, and thus cannot operate. Due to insufficient overlap of the laser scanning data during these movements, the ICP algorithm fails to identify corresponding points accurately, and thus the map quality and errors in attitude estimation may decline [5]. Hector SLAM also has good performance in environments that are rich in objects. It may be that there are too few matching features in a sparse environment, and thus the system will not work properly.

2.1.3. Cartographer: Google's pose graph optimization-based 2D and 3D mapping for UAV

Google has developed the Cartographer advanced simultaneous localisation and mapping (SLAM) technology, which combines local sub-maps with loop closure detection [7]. A Branch and Bound Search Strategy was employed to optimise the attitude map. The Map shows the relationships among all the Positions of the Unmanned Aerial Vehicles during the survey.

Google has developed a high-end SLAM method called Cartographer that integrates loop closure detection with local sub-maps [8]. A pose graph can show the connections among the locations of all the unmanned aerial vehicles in the mapping process. The optimiser selects the branch-and-bound search method. However, the cost of this added function is a higher computational load. Some unmanned aerial vehicle platforms, especially those with stringent load and power restrictions, may not be able to support the relatively high-performance onboard computing system needed for real-time operation of Cartographer [7].

2.1.4. LeGO-LOAM: feature-segmented LiDAR SLAM for real - time UAV applications on embedded systems

LeGO-LOAM (Lightweight and Ground - Optimised Lidar Odometry and Mapping) solves the computational problems of LiDAR SLAM for UAVs by introducing feature segmentation [9]. It can classify points in a LiDAR point cloud as ground or non-ground.

Segmentation reduces the computational cost significantly by considering only the necessary non-ground features for mapping and localisation. For real-time UAV operation, such as in inspection tasks that require immediate feedback, the lightweight Design of LeGO-LOAM can run efficiently on embedded systems. Provide real-time pose estimation and map updates for safe navigation and task execution by the UAV [10]. LeGO-LOAM has strong performance advantages only when certain distinct environmental features are available. In areas of homogeneity, such as large open spaces or dense forests with sparse or uniform features, it may be difficult for the algorithm to extract features from scans for matching, and thus both map construction and localisation accuracy will be reduced [9].

DOI: 10.54254/2755-2721/2025.AST34382

2.1.5. Discussion of LiDAR SLAM

LiDAR-based SLAM systems generally have better performance in harsh environments than visual-based methods. For instance, when the light is too low and a camera cannot get a clear image, LiDAR can still work normally because it is an active-sensing technology [11, 12]. LIO-BIM is used to combine LiDAR-inertial odometry and building information models for sub-centimetre-level positioning accuracy of mobile robots. The above high-precision positioning accuracy is very suitable for the requirements of UAVs in industrial inspection that need to locate specific parts or structures precisely.

LiDAR SLAM is not perfect either, however [11]. LiDAR sensors are inherently relatively heavy and have a high power demand. Small UAVs have a small payload capacity and a short battery life; thus, they are not suitable for long-endurance and heavy-duty applications. Processing the large amount of point cloud data produced by LiDAR sensors also introduces a delay [9]. Although Cartographer has good map accuracy, it is relatively slow and therefore not suitable for real-time mapping by UAVs [7]. If the control system of the UAV needs to wait for the SLAM algorithm to update the map and pose information, it may be unstable in flight and potentially dangerous [10].

2.2. Visual SLAM

Visual SLAM uses camera images for localisation and mapping [13]. Cameras are relatively light and inexpensive, and they can provide rich semantic information of the environment for UAVs [14]. However, they are more susceptible to environmental factors than LiDAR-based systems.

2.2.1. Monocular SLAM: single-camera-based localization and mapping for lightweight UAV

Monocular SLAM systems are based on a single camera for pose estimation and map construction of the UAV [13]. The above systems employ feature extraction and matching to track the movement of the camera in relation to the environment.

The first few devices are typically simpler in design and less power-hungry. For very lightweight drones that are only needed to observe indoors or in areas outside, single-camera SLAM can be used because it adds little to the weight of the drone. State-of-the-art monocular SLAM algorithms such as ORB-SLAM3 have achieved good results by using feature matching and loop closure detection [14]. They can identify the features of edges and corners in the video streams of several consecutive frames. Monocular SLAM is thus limited by a size-ambiguity problem. The absolute scale of the surroundings cannot be obtained directly from the photos because a single camera only records 2D information. In a large-scale environment or over a long distance, the estimated pose and map will be inaccurate [13]. Monocular SLAM algorithms may also struggle to find enough matching features in low-texture environments where few distinctive features are available for tracking and perform poorly as a result.

2.2.2. Stereo SLAM: Triangulation-Based 3D mapping for UAV with moderate payload capacity

Two cameras are used in stereo SLAM systems, and they are usually positioned at a predetermined baseline distance from one another [15]. Stereo SLAM uses triangulation to recover the absolute scale of the environment by taking pictures from two different angles.

DOI: 10.54254/2755-2721/2025.AST34382

Monocular SLAM is affected by the scale ambiguity problem, but this ability to determine scale has been solved. Stereo SLAM can construct a high-precision three-dimensional map of an organised environment, such as a city with roads and buildings. Better localisation and mapping are enabled by the two-camera configuration to obtain more details of the scene's depth [16]. The two cameras need to be finely tuned for stereo SLAM to function correctly. Inaccurate Depth Estimation and Pose Calculation may be caused by any misalignment or calibration problems. A reasonable basal distance of the cameras should be chosen to obtain a large enough parallax for accurate triangulation [15]. A larger baseline distance can achieve more precise depth data, but it results in an increase in the size and weight of the camera array; thus, the payload capacity of the drone needs to be expanded.

2.2.3. RGB-D SLAM: combining visual and depth information for UAV mapping with limited range

RGB-D cameras combine the visual information of an RGB camera with depth data, and this depth information is usually obtained by means such as time-of-flight or structured light. Accurate 3D mapping can be realised through the above combination [17].

RGB-D SLAM is cost-effective and high-performance. Under certain circumstances, it can obtain higher-accuracy depth information than that provided by monocular and stereo SLAM. For example, if the measurement range of the depth sensor is large enough, then RGB-D SLAM can build a fine-grained three-dimensional map of an indoor scene quickly [18]. However, the depth range of RGB-D cameras is generally less than 10 meters and therefore not suitable for outdoor UAVs. Outdoor UAV may need to cover a large area at a long distance. They are more likely to be affected by changes in environmental light [17]. Strong reflection or a hard light source can reduce depth measurement accuracy; therefore, the result may be inaccurate [18].

2.2.4. Discussion of Visual SLAM

Visual SLAM has a low weight and cost compared with LiDAR SLAM [19]. Cameras are relatively light and inexpensive compared with laser radar sensors. Hier-SLAM is a recent improvement in visual SLAM that employs a hierarchical semantic representation method and has shown better generalization ability and efficiency in complex environments [20]. Hier-SLAM can learn the structure of the environment through multi-level classification of objects and thus obtain a high-accuracy map and position.

Visual SLAM is also quite sensitive to changes in the surrounding environment at the same time. A Change in the Light Will Also Affect the Operating Conditions of the Visual SLAM System. For example, moving from a dark place to a very bright one suddenly will reduce the ability to recognise features or match them [19]. When the drone is in rapid movement or quick rotation, motion blur will also occur and reduce the effectiveness of Visual SLAM. Balance the computational efficiency and robustness of the two. Although DROID-SLAM and other deep learning-based methods can achieve state-of-the-art performance in some cases, they require powerful onboard processing capabilities, and not all drone platforms have such capabilities, especially those with a small payload [20].

DOI: 10.54254/2755-2721/2025.AST34382

3. Applications

3.1. Electrical wire inspection

LeGO-LOAM is a Lidar-based system for power-line inspection that has shown good results. They can generate precise 3D models of the line network and carry out distance measurement to reduce the trouble of defect detection [21]. They can still work normally in different light environments and are suitable for various times of the day [22]. Visual SLAM has a relatively low cost, but it is difficult to address the problem of motion blur when a drone moves at a high speed near power lines [21].

3.2. Agriculture

RGB-D SLAM can be selected due to its low cost and high efficiency for crop monitoring and field mapping in agriculture. Cartographer's 2D map function helps to produce field borders and irrigation plans [7]. Lidar-based systems such as Cartographer are generally more robust in dense vegetation; otherwise, a vision-only system might fail due to occlusion [23].

3.3. Exploration

LiDAR SLAM is robust in exploration of unknown environments (such as disaster areas) and is therefore chosen here [24]. The LIO-BIM framework can be aligned with an architectural model, and therefore, urban search and rescue tasks have also been conducted using it [11]. Visual SLAM is a lighter-weight option for the initial observation task that employs semantic information [13].

4. Current challenges

4.1. Multi-UAV collaborative SLAM

Efficiently synchronize data and fuse maps for collaborative mapping of multiple drones [25]. Problems include the processing of asynchronous data, resolution of inconsistencies in map updates, and guarantee of stable network connections in an area without GPS [26]. There are no unified standards for optimising pose maps among unmanned aerial vehicles currently.

4.2. Real-time performance

Drones need high-frequency SLAM updates (≥10Hz) for stable control [27]. Although Hector SLAM has been implemented on embedded systems, complex three-dimensional mapping using LiDAR or deep learning-based vision methods is not feasible in real time due to limited computing power [20].

4.3. Robustness

Environmental Changes are still problems. LiDAR systems have the problem of point cloud sparsity at a distance, and visual methods fail in dynamic light conditions [13]. Bad weather, such as rain and fog, is also very damaging to visual SLAM and can restrict the use of drones in bad weather [28].

DOI: 10.54254/2755-2721/2025.AST34382

5. Conclusion

Analyze the strengths and weaknesses of the different categories of SLAM for unmanned aerial vehicles in this paper. Systems based on LiDAR have good accuracy and robustness, but they are relatively heavy and power-hungry. Visual SLAM is relatively light and inexpensive, but it is also affected by the environment [29, 30].

The above problems will be addressed in the next research. Combine the advantages of several SLAM technologies to enhance the overall performance through fusion of LiDAR and visual data [29]. In addition, research on the integration of end-to-end learning and large language models has shown promising results in expanding the generalisation and semantic recognition capabilities of SLAM for better operation in complex and changing environments for drones. Therefore, more areas will be added to the list of drone applications.

References

[1] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part I, " IEEE Robotics & Automation Magazine, vol. 13, no. 2, pp. 99–110, Jun. 2006, doi: 10.1109/mra.2006.1638022.
[2] J. P. Queralta et al., "Collaborative Multi-Robot Search and Rescue: planning, coordination, perception, and active vision, " IEEE Access, vol. 8, pp. 191617–191643, Jan. 2020, doi: 10.1109/access.2020.3030190.
[3] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with Rao-Blackwellized particle filters, " IEEE Transactions on Robotics, vol. 23, no. 1, pp. 34–46, Feb. 2007, doi: 10.1109/tro.2006.889486.
[4] D. R. Isaacson, A. V. Sadovsky, and D. Davis, "Tactical scheduling for precision air traffic operations: past research and current problems, " Journal of Aerospace Information Systems, vol. 11, no. 4, pp. 234–257, Apr. 2014, doi: 10.2514/1.i010119.
[5] S. Kohlbrecher, O. Von Stryk, J. Meyer, and U. Klingauf, "A flexible and scalable SLAM system with full 3D motion estimation, " Proc. IEEE Int. Conf. Intell. Robots Syst. (IROS), pp. 155–160, Nov. 2011, doi: 10.1109/ssrr.2011.6106777.
[6] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes, " IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14, no. 2, pp. 239–256, Feb. 1992, doi: 10.1109/34.121791.
[7] W. Hess, D. Kohler, H. Rapp, and D. Andor, "Real-time loop closure in 2D LIDAR SLAM, " May 2016, doi: 10.1109/icra.2016.7487258.
[8] F. Dellaert and M. Kaess, "Square Root SAM: Simultaneous localization and mapping via square root information smoothing, " The International Journal of Robotics Research, vol. 25, no. 12, pp. 1181–1203, Nov. 2006, doi: 10.1177/0278364906072768.
[9] T. Shan and B. Englot, "LEGO-LOAM: Lightweight and Ground-Optimized LIDAR odometry and mapping on variable terrain, " 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Oct. 2018, doi: 10.1109/iros.2018.8594299.
[10] J. Zhang and S. Singh, "LOAM: Lidar Odometry and Mapping in Real-time, " Proc. Robotics: Sci. Syst. (RSS), Jul. 2014, doi: 10.15607/rss.2014.x.007.
[11] D. T. Fasiolo, L. Scalera, and E. Maset, "Comparing LiDAR and IMU-based SLAM approaches for 3D robotic mapping, " Robotica, vol. 41, no. 9, pp. 2588–2604, Apr. 2023, doi: 10.1017/s026357472300053x.
[12] L. Wallace, A. Lucieer, Z. Malenovský, D. Turner, and P. Vopěnka, "Assessment of Forest Structure Using Two UAV Techniques: A Comparison of Airborne Laser Scanning and Structure from Motion (SfM) Point Clouds, " *Forests,* vol. 7, no. 3, p. 62, Mar. 2016, doi: 10.3390/f7030062.
[13] R. Mur-Artal and J. D. Tardos, "ORB-SLAM2: an Open-Source SLAM system for monocular, stereo, and RGB-D cameras, " IEEE Transactions on Robotics, vol. 33, no. 5, pp. 1255–1262, Jun. 2017, doi: 10.1109/tro.2017.2705103.
[14] C. Forster, M. Pizzoli, and D. Scaramuzza, "SVO: Fast semi-direct monocular visual odometry, " IEEE Transactions on Robotics, pp. 15–22, May 2014, doi: 10.1109/icra.2014.6906584.
[15] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite, " 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 3354–3361, Jun. 2012, doi: 10.1109/cvpr.2012.6248074.

DOI: 10.54254/2755-2721/2025.AST34382

[16] R. Gomez-Ojeda, F.-A. Moreno, D. Zuniga-Noel, D. Scaramuzza, and J. Gonzalez-Jimenez, "PL-SLAM: a stereo SLAM system through the combination of points and line segments, " *IEEE Transactions on Robotics*, vol. 35, no. 3, pp. 734–746, Apr. 2019, doi: 10.1109/tro.2019.2899783.
[17] S. Izadi et al., "KinectFusion, " Proc. ACM Symp. User Interface Softw. Technol. (UIST), Oct. 2011, doi: 10.1145/2047196.2047270.
[18] R. Du et al., "DepthLab: Real-time 3D Interaction with Depth Maps for Mobile Augmented Reality, " pp. 829–843, Oct. 2020, doi: 10.1145/3379337.3415881.
[19] S. A. Bahnam, C. De Wagter, and G. C. H. E. De Croon, "Improving the computational efficiency of ROVIO, " Unmanned Systems, vol. 12, no. 03, pp. 589–598, Mar. 2023, doi: 10.1142/s2301385024410012.
[20] J. Engel, V. Koltun, and D. Cremers, "Direct sparse odometry, " IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 3, pp. 611–625, Apr. 2017, doi: 10.1109/tpami.2017.2658577.
[21] W. Zhang et al., "The Application Research of UAV-based LiDAR System for Power Line Inspection, " Jan. 2017, doi: 10.2991/iccia-17.2017.174.
[22] H. Zhang, W. Yang, H. Yu, H. Zhang, and G.-S. Xia, "Detecting Power Lines in UAV Images with Convolutional Features and Structured Constraints, " *Remote Sensing*, vol. 11, no. 11, p. 1342, Jun. 2019, doi: 10.3390/rs11111342.
[23] J. Bendig et al., "Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley, " International Journal of Applied Earth Observation and Geoinformation, vol. 39, pp. 79–87, Mar. 2015, doi: 10.1016/j.jag.2015.02.012.
[24] A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann, "6D SLAM—3D mapping outdoor environments, " Journal of Field Robotics, vol. 24, no. 8–9, pp. 699–722, Aug. 2007, doi: 10.1002/rob.20209.
[25] C. Cadena et al., "Past, present, and future of simultaneous localization and mapping: toward the Robust-Perception age, " IEEE Transactions on Robotics, vol. 32, no. 6, pp. 1309–1332, Dec. 2016, doi: 10.1109/tro.2016.2624754.
[26] J. A. Placed et al., "A survey on Active Simultaneous Localization and Mapping: State of the Art and new frontiers, " IEEE Transactions on Robotics, vol. 39, no. 3, pp. 1686–1705, Mar. 2023, doi: 10.1109/tro.2023.3248510.
[27] Y. Xie et al., "RDC-SLAM: a Real-Time distributed cooperative SLAM system based on 3D LiDAR, " IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 9, pp. 14721–14730, Dec. 2021, doi: 10.1109/tits.2021.3132375.
[28] S. Hinterstoisser et al., "Model based training, detection and pose estimation of Texture-Less 3D objects in heavily cluttered scenes, " *in Lecture notes in computer science*, 2013, pp. 548–562. doi: 10.1007/978-3-642-37331-2_42.
[29] S. Thrun, "Simultaneous localization and mapping, " in Springer eBooks, 2007, pp. 13–41. doi: 10.1007/978-3-540-75388-9_3.
[30] D. Novak and R. Riener, "A survey of sensor fusion methods in wearable robotics, " Robotics and Autonomous Systems, vol. 73, pp. 155–170, Sep. 2014, doi: 10.1016/j.robot.2014.08.012.