Applied Design and Methodology of Delivery Robots Based on Human–Robot Interaction in Smart Cities

Wing Ting Law¹*, Kam Wah Fan ², Ki Sing Li³, and Tiande Mo⁴

¹²³⁴ Hong Kong Productivity Council, Hong Kong

Abstract

This paper proposes an optimised design of an autonomous delivery robot while adopting the latest technologies from the different branching fields of robotics, artificial intelligence, and tele-communication. As a prospective representation of a user-centric robot design, the proposal is design with the major focus on maximizing users’ satisfaction throughout every human–robot interaction (HRI) touchpoints. By the use of sensor fusion techniques along with the deployment of an image-detection-based technique accompanying the point-cloud-detection-based path-planning methodology, the robot delivery would be optimised with effective path-planning and obstacle avoidance capability. With the extension of 5G connectivity, it is proposed that the real-time status update and video stream would enable greater efficiency in terms of remote monitoring and centralised robot administration.

Keywords: Smart Cities, Smart Mobility, Robotics, Robot, Human-Robot Interaction; HRI; Human-centric computing; Interaction design; User centric design; Robot design; Robot Methodology.

Received on 30 August 2022, accepted on 02 April 2023, published on 26 June 2023

Copyright © 2023 Law et al., licensed to EAI. This is an open access article distributed under the terms of the CC BY-NC-SA 4.0, which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/etesc.2649

1. Introduction

Adapting to the shift of the modern world with increasing traffic of goods and a severe labour shortage, it is foreseeable that the trend of robot delivery would be much popular in order to serve as a versatile alternative for human labour in cities [1][2][3][4]. By having robots take part in labour-intensive tasks, the automation in various industries allows the precious labour resources to re-allocated on the important value-added tasks.

In recent years, on-demand delivery robots in buildings have been a welcomed substitution over human delivery in most business. This paper, therefore, aims to demonstrate a user-centric idea on Human-Robot Interaction [5] (hereafter called HRI) using various design and methodologies that adopt the latest technology among the fields of robotics, artificial intelligence, and tele-communication in real-life applications. Throughout the adoption of the different types of technology, it is realised that various HRI considerations could be taken into account in order to enhance the overall user experience by revamping human-robot touchpoints.

With careful considerations on the planned usage and optimised user experience, hardware design, and software design, the research team proposes the development of an advanced delivery robot that grants the advantages of “user-friendly” system flow, well-compartmentalised hardware architecture, advanced sensor-fusion that enables the deployment of an image-detection-based trajectories prediction technique accompanying the point-cloud-
2. General HRI Considerations in Delivery Robots

According to the research from Patric R. Spence, negative feelings would be easily expressed if the robot is perceived as useless, such that the users would then be encouraged to isolate themselves from the use of the robot [6]. As a matter of fact, it is perfectly common for users to avoid solutions that are perceived to be less efficient or effective when the users believe they could outperform those solutions themselves. This observation gives the insight to the research team that it is crucial to enlighten users with the perception that the robot has to be cognitively effective and efficient in terms of the quality of the robot performance both tangibly and intangibly. This very insight originates the research team’s initial understanding to HRI.

The rawest concepts of general HRI originated from the novel Robot of Isaac Asimov in 1941, which stated that a robot shall not injure a man being, or through inaction, allow a human being to come to harm. It is one of the later interceptions in 1994 that if a robot-based economy develops without equitable adjustments, the backlash could be considerable [7]. Therefore, it is the interpretation of the research team that, it is important to create a robot paradigm that balance between the economic or technological advantages provided by the robots and the overall drawbacks led by the deployment of the robots.

It is in the consensus of later HRI considerations that the goal of HRI is to allow comfortable and acceptable interactions [5][8]. It is common that users might have uncertain feelings when considering the interaction with robots in contrary to the face-to-face communication among actual human beings. As such, this paper focuses on enabling user-centric HRI by optimised designs and methodologies. In the proposed use-case, there would be numerous touchpoints between humans and robots during the user journey of the delivery robot.

3. Planned Usage and Optimised User Experience

For general delivery tasks, it is commonly believed that the system performance in terms of stability and reliability are the major focuses from an engineering standpoint. However, in the perspectives of HRI, the importance of having an intuitive and fool-proof system design that could optimise user experience outweighs every quantitative study on system performance. In short, the common users would always appreciate a simple robot that could be interacted and controlled easily and intuitively, over a flawless but complicated robot with rigid and exhaustive control measures. As suggested by Valeria Villani, the effective physical and cognitive interaction on robotic solution not only promotes efficiency, but also grant safety and intuitive ways to program and interact with robots [9].

It is therefore the sense of intuitiveness in user experience that drives the research team to develop a “user-friendly” user journey as the key foundation toward an optimised user journey, as shown in Figure 1 which lays out all touchpoints between the users and the robot.

From the perspective of general users, the journey of the delivery task always starts from the need for having something delivered, which is a very typical scenario in in-building use-cases. Then, as an automated delivery service, the item sender would call the robot for delivery service, which should be at its best availability to response to the sender’s call. Once the sender drops off the item to the robot, the robot would then begin its delivery in a secure, careful and promising fashion. Eventually, the robot should arrive at the destination, so that the item receiver could collect the delivered item in its perfect condition. Simple as the below logic flow might sound, the research team believes that the best user experience that could ever be provided, must be a well-planned and comfortable experience, i.e. “user-friendly”. It is therefore obvious that there would be numerous touchpoints between the robot and the users, such as the user interface, the storage cabinet,
the visual appearance and attraction, as well as the audible alerts for notifications.

In the expected use-case of the delivery robot, it is assumed that the robot would have to be able to freely travel within the different floors of the building with reasonable passage space and crowdedness of obstacles. In terms of free travel, the access towards elevators and door would also have to be interactable with the robot, such that the robot would be able to communicate and interact with the elevators and doors in order to get access towards the designated areas inside the building.

4. Hardware Design

With the aim to design a robust delivery robot as a potential alternative for human labour that facilitates and encourages smooth and comfortable user experience regarding user perception, interaction and engagement, the development of a versatile robot would be necessary, such that the robot would have to have an industrial design that emphasises on the robot’s physical appearance and functionality, let alone well-designed compartmentalisation.

In general, there are proposed to be six different units in the delivery robot, including power unit, Internet unit, sensor unit, I/O unit, storage unit and the driving unit as illustrated in Figure 2.

The power unit that lives up all parts, like the heart of a human, contains the rechargeable lithium-ion battery for providing dense and efficient energy, and the battery management system (BMS) for protecting the battery from operating outside its safe operating status, monitoring and reporting the usage and current information. To ensure safety, the emergency stop (e-stop) is also included as a part of the power unit to sever all power immediately if needed.

The sensor suite, as depicted as the eyes of the robot, helps gathering all environment data that enables the robot to have the correct moving directives when performing its task. The inertial measurement unit (IMU) measures and reports the orientation of the robot by using a combination of accelerometers and gyroscopes. The ultrasonic sensors and 2D LiDAR are used to support collision avoidance by detecting nearby objects. Similarly, the 3D LiDAR is used to detect objects in a greater range to achieve accurate localization and navigation. Moreover, the camera captures image and videos for further visual analytics.

The driving unit, as depicted as the feet of the robot, includes the wheels, motor drivers and bumper, which allow the robot to perform physical movements. To enhance its mobility, the robot is configured to be front-wheel-drive with omni wheels so that omni-directional manoeuvring through crowds could be achieved.

The I/O unit, as depicted as the mouthful and facial expression of the robot, serves as a communication pathway with other road-users, enabling multiple touchpoints. The on-board touchscreen with intuitive and relatable UI, two-way communication system and the LED visual indicator visualise the current status of the robot with the road users with messages like, “I am idle” or “I am in a task”.

The storage unit is a value-added unit, acting like the weight-bearing hands of a human, allowing users to deliver goods in a secure way by adopting the use of NFC readers and cards, and an electric lock.

The Internet unit, as the metaphorical representation of the human mind, provides access to the Internet and further expands the touchpoints of the robot.

Last but not least, the centre processing unit, as depicted as the brain of the robot, links up all the above units to
centralise the control over all the computational processes and coordinates all components to perform tasks.

In terms of visual and audio design, to increase touchpoints and communicate with users throughout the delivery task, the robot would provide signals through visual and audible indicators, including LCD Display, LED lighting, and audible alerts and prompts through speaker.

It is believed that with the help of the visual outputs in the forms of graphics and texts that are shown on the LCD display as well as eye-catching LED lighting for status indication, along with comprehensive voice prompts and attractive audible alerts, users and passers-by could have a more intuitive understanding to the robot and its status to avoid accidents, and increase the sense of user-friendliness. Details of visual LED and audible indicators are as described in Table 1.

<table>
<thead>
<tr>
<th>Status</th>
<th>Visual</th>
<th>Audible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>LED flashes green</td>
<td>-</td>
</tr>
<tr>
<td>Tasks running</td>
<td>LED flashes blue</td>
<td>Soft Melody</td>
</tr>
<tr>
<td>Tasks in pause</td>
<td>LED flashes amber</td>
<td>-</td>
</tr>
<tr>
<td>Manual control</td>
<td>LED flashes magenta</td>
<td>Vehicular Melody</td>
</tr>
<tr>
<td>Error detected</td>
<td>LED flashing red</td>
<td>Sound Alerts</td>
</tr>
</tbody>
</table>

5. Software Design

As an autonomous delivery robot with the primary duty of in-building delivery, successful navigation among in-building structures is the most vital part for the performance of the delivery robot, in terms of navigation robustness and simultaneous localization and mapping (SLAM) [10][11]. Therefore, it is crucial to organise all the related processes in a comprehensive manner. Figure 3 above illustrates the entire navigation process from sensor information inputs to hardware actuations, such that the robot could fully utilise the technology advancement in sensor fusion, SLAM, and 5G.

5.1 Sensor Fusion – Obstacle Detection and Path Planning

As mentioned, this delivery robot would be equipped with various kinds of sensors, including Inertial measurement unit (IMU) for acceleration measurements, motor drivers for odometry measurements, 2D and 3D LiDAR for real-time simultaneous localization and mapping (SLAM), camera for object recognition, and ultrasonic sensors for collision detection.

Whenever a navigation task is initiated, the inputs from all the sensors would be firstly combined in the sensor fusion process to enhance the pose estimation accuracy of the robot. After that, the more complete and accurate data would be used to localise the actual location of the robot in the environment through the localization process, through the use of different localization algorithms [12][13]. At the same time, the data from the camera would also be adopted in the object recognition process to identify the category of the filmed objects from the visual feed, so that the later obstacle avoidance process could make reasonable reflections for optimal path planning. Likewise, the data from the ultrasonic sensors would be used for the real-time collision detection process, such that imminent collisions
Figure 5. Sensor fusion – Object Recognition in Distance

would be detected and address corresponding response, as illustrated in Figure 4 and Figure 5.

Figure 3 and Figure 4 are the expected visual representation of the navigation and localization methodology. It should be the core value of this sensor-fused integration to be able to understand the environment in terms of the existence of surrounding obstacles. As depicted in Figure 4, if the point cloud data from LiDAR is obtained as the sole input, accurate size and distance information would be complete for only items that exist in the most effective region of the LiDAR range.

On the other hand, with the use of sensor fusion techniques, as depicted in Figure 4, from the image feed from the front camera, the use of the detection models allows the robot to clearly identify every visibly detected obstacle in terms of size and distance. This enhancement in obstacle detection facilitates further decision-making on the path-planning logics. For instance, for static obstacles like chairs and desk, the delivery robot should plan for alternative paths as these types of objects would always stay in their existing position; while if non-static obstacles like pedestrians are detected, the delivery robot could try to eliminate the chance of collision by playing voice prompt to alert those pedestrians or just wait a while in a safe position until the obstacle is clear from the path. It is believed that such soft handling would make the delivery robot more like a human and enable dynamic and case-based decision.

5.2 Motion Planning Flow of an In-Building Delivery Task

As the core component of an in-building delivery robot, the functionalities of navigation and localization have been given special focuses in terms of collision avoidance, path-planning, and cross-floor travel handling.

In the traditional approach for collision avoidance, the algorithm simply stops the robot until the obstacle is out of the detection range. Alternatively, it is a more advanced method to perform path-planning based on the result of model-based trajectory predictions on the movement of moving obstacles like walking pedestrian. While the traditional approaches are less desirable in densely populated environments such as streets and parks given the constant crowdedness, the latter approach performs much effectively on those situations with reasonable traffics. As depicted in Figure 6 and Figure 7, the trajectory prediction model uses the gathered information from the sensor fusion process, such as current distance, and current linear and angular velocity to serve the purpose of object tracking.

With the aid of sensor fusion technology, the research team has adopted the image-detection-based trajectories prediction technique along with the point-cloud-detection-based path-planning methodology enabled by the use of 3D and 2D LiDAR, such that the object trajectories information on each of the individually tracked obstacle would be imported to the prediction algorithm to perform movement prediction with a fair confidence level, which would then be combined with the point-cloud-based navigation algorithms to plan for an optimal and obstacle-free path.

To be specific on the main motion planning flow of the delivery robot, each delivery task the robot handles involve three major iterative steps, namely the “handshakes”, the “walk”, and the “ride”, as illustrated in Figure 8. Every delivery task starts with the first simple “handshake”, where the users would entrust items to the robot’s cabinet.
using the NFC card reader and initiate the delivery task by pressing one single button on the robot, which would then acknowledge the delivery task by audible alerts. This action-and-reaction-based “handshake” aims to work as the initial basic contact that would establish a sense of trustworthiness towards the users.

When the delivery task is initiated, the “walk” is begun, where a smooth melody would be played to indicate the robot’s presence, and a vivid LED lighting and display would be used to notify near pedestrians, just like taking a good walk together in music and glares. In details, the “walk” is actually a check-point-based path where the robot proceeds to the next check-point after reaching one. The process of path-planning, as mentioned beforehand, consists of two major parts, including the image-based obstacle detection and the point-cloud-based navigation, such that the obstacle detection model would provide trajectory prediction upon moving obstacle while the navigation stack would provide distancing information about static obstacles. To facilitate effective path-planning, the robot takes into concern information from both ends and output the optimal obstacle-free path for the robot to travel.

Then, the “ride” begins when the robot has to travel in floors with elevator, such that the robot would communicate with the elevator server wirelessly to acquire the current status of each elevator, so that the robot could call for elevators that are currently idle and allow the elevators to prioritise the needs of human passengers. Once, the elevator has been called and with the elevator door opened, another detection model would be used to search for passengers and pedestrians, just to ensure the robot would be able to enter the elevator safely. Right before entering the elevator, the robot would even notify pedestrians by clear audible alerts.

Once the destination floor is reached, the “walk” resumes and continues until the robot arrives at the final checkpoint for the delivery. The whole delivery task would then be completed by one last “handshake” where the users would retrieve the entrusted items from the locked cabinet using the NFC card reader and the robot would acknowledge its completion in delivery with another audible alert.

From an engineering perspective, this might have been one simple flow design for a basic delivery task. As a matter of fact, the major emphasis in this flow design, as well as the industrial design of the robot, has been the intractability with pedestrians, such that numerous touch points were provided throughout the “handshakes”, the “walk”, and the “ride” in terms of audio, visual and interface interactions as mentioned in previous sections.

5.3 Backend Server Connection and 5G Extensibility

Throughout the design of the delivery robot, various security, safety and administrative considerations were also involved, such that a backend server was developed for the purpose of centralised robot control, especially for those requiring low tolerance in latency for high-speed data transmission [14]. It is the experience of the research team that enlightens the importance of remote monitoring and control of robots. The backend server acts a remote data-logger that stores and reflects the current status of the robot, i.e. battery consumption, locational information, task status, task history, and any alert produced. Besides data-logging, for security and safety reasons, the backend server also acts as a remote control-panel that allows authorised personnel to control the robot in terms of check-point assignment, task initiation and termination, and real-time video feed from the robot’s front camera.

It is fully understood that the more important and powerful the backend server is, the higher the risk associated with the server-robot communication is intercepted, eavesdropped or even altered. Therefore, the communication architecture between the server and robot...
is reinforced in both software and hardware levels. In terms of software, the content for every outgoing communication would be encrypted with AES format, such that every status information and each frame from the video feed would be masked. However, with the increased processing time for the status update and video stream, it would be much difficult to establish a real-time connection. Thus, the 5G technology has been adopted to enhance the speed as well as the security level for the data transmission.

The use of 5G technology gives three major benefits to the communication architecture, including high data transmission speed, large maximum carrier bandwidth, and mass connectivity. In terms of the expected use-case, the high data transmission speed allows the robot to upload the loads of status messages and streams of video feed to the server with a very low latency, so that real-time monitoring of the robot could be achieved. To be specific, the 5G architecture uses Ultra-Reliable and Low Latency Communications (uRLLC) to allow the travelling network to be optimised for processing incredibly large amounts of data with minimal delay so as to support end-to-end latencies as low as 5ms [15]. This is achieved by implementing a new approach to handle radio frame slots. Instead of fixed radio slots of 1ms in 4G, 5G uses a multi-numerology approach to allow flexible definition of radio slots, so that radio slots with shorter timespan would be allows, in returns of a lower latency.

In addition, the large maximum carrier bandwidth benefited from the 5G technology makes the streaming of 4K resolution video possible in real-time. With the use of enhanced Mobile Broadband (eMBB), a greater data-bandwidth complemented by moderate latency improvements on both 5G NR and 4G LTE could be provided, such that the actual latency between image capturing and video presentation on the server for high resolution videos, like 4K or omnidirectional VR video, could be less than 0.3 seconds even after the implementation of the required data compression, and data encryption and decryption as mentioned. Last but not least, the advantage of mass connectivity provided by the adoption of Massive MachineType Communication (mMTC) allows extremely high connection density of online devices, and thus enables the deployment of robot fleets and the potential migration of sensors or processes through cloud computing in the future, where the large number of robots and independent sensors would no longer be the limitation for real-time data transmission and processing, even in conditions of limited communication resource [16].

It has been one of the biggest challenges for traditional indoor-deploying robots to face Internet instability due to the complex indoor environment. Therefore, to encounter this challenge, numerous 5G-enabling antennas and transceivers have been deployed throughout the building, such that the robot would have guaranteed full Internet connection while traveling among the designated deployment areas of the delivery tasks. As a matter of fact, the adoption of the 3.5GHz (sub-6) spectrum has enabled a cost-efficient deployment of the large number of transceivers such that the 3.5GHz spectrum allows a satisfying transmission rate from 193 to 430 Mbit/s down, while having a reasonable sensing range of in the building complex, when compared to the two other expensive and range-limiting spectrums of 26-28GHz and 4.9GHz [17].

6. Conclusion

It is obvious that the business model of modern companies is experiencing a huge change towards automation in recent years. It is inevitable for business to accompany this shift in paradigm by adjusting their business strategy and even business model to introduce a higher level of business automation. As a result, citizens would have a greater and more frequent exposure towards the different types of robots in workspace. It is therefore a crucial issue to design new robot morphologies, appearances, behaviour paradigms, interaction techniques to encourage a smooth and reassuring interaction between human and robots, let alone the design of systems, algorithms, interface technologies, and computational methods that supports such human-robot interaction.

As a prospective representation, the research team proposes the development of a user-centric design of a delivery robot that could optimise user experience. By adopting the latest technologies from the different branching fields of robotics, artificial intelligence, and telecommunication, the proposed delivery robot is intended to maximise users’ satisfaction throughout every HRI touchpoints.

As a delivery robot, the proposed hardware design follows the foundational design concept of a last-mile autonomous delivery robot, such that this robust robot would be able to provide a welcoming visual and audio appearance, as well as a reliable architectural framework with the six core units of power unit, Internet unit, sensor unit, I/O unit, storage unit and the driving unit, to metaphorically serve as the heart, mind, eyes, mouth, hands, and feet of a human. As the major attraction, the proposed robot adopts the latest techniques in artificial intelligence to perform real-time SLAM with both image-detection-based trajectories prediction technique and the point-cloud-detection-based path-planning methodology for effective navigation and obstacle avoidance. As the major duty of the delivery robot, a cross-floor travelling flow has also been proposed to enable and facilitate effective path-planning, communication between the robot and elevator/doors. It is another design attraction that the Internet unit is proposed to be deployed with 5G compatibility in the 3.5GHz (sub-6) spectrum, so that the server-robot connection would be
of higher speed and security, enabling stable status monitoring and video streaming.

It is believed that the proposed robot design would be able to serve as a viable proposal for a new generation autonomous delivery robot, with the major software design focus on sensor fusion, advanced motion planning, 5G-based high-speed backend server communication and streaming. It is the research team’s next action to build the actual prototype that employs the proposed design and methodologies. It is hoped that with such design specification, the user-friendly delivery flow and comfortable user experience provided would be able to encourage a closer and sophisticated human-robot interaction.

7. Further Development

While serving with a vital proposal, the research team believes that, as a further enhancement or a design re-deployment with customised use-cases, more reinforcement and alternative design could have been made with the latest technological advancement.

The first enhancement proposal would be the adoption and integration of Natural Language Processing (NLP) technology, like ChatGPT [18], which could be applied to allow natural and human-esque interaction so that verbal human-to-human interaction could be extended to a human-to-robot level. The touchpoints between the delivery robot would have been evolved to a whole different level, by expanding communication capabilities with contextual understanding and abstract command processing, and dynamic command adaption, such that causal verbal command could be actually used to control the flow of the delivery task, e.g. specifying the destination with abstract verbal description, pausing the robot movement by saying a simple “stop”, and re-defining checkpoints or directions by tipping the robot “turn left” or “turn right”, or even “take this package to the café near the entrance”.

Another interesting enhancement would be the implementation of multi-robot-to-robot communication, for instance the concept of collaborative multi-robotic tasks in the complex development [19], and hence the development of multi agent systems [20]. It is like the sociological paradigm of social collaboration, where multiple people, groups, or departments could interact with each other to achieve common goals, a fleet of robots could apply the same paradigm to accomplish a much intensive task. For example, in terms of items delivery, when the items to be delivered is way too many or too big for one single robot, it might be possible for the fleet of robots to collaboratively deliver parts of the big group of items or move in synchronous movement to deliver the single big item. If a reliable and self-sustaining robot-to-robot communication framework could be developed for general robot collaborations, it is believed that the level of automation could be vastly improved, which could thus further enhance the economic and technological advancement for the realm of HRI.

References


