Intuitive Controls for Swerve Drive Robots

Christopher Xu
christopher.y.x@gmail.com
Palo Alto High School, Palo Alto, CA, United States

Abstract

Swerve drive robots are highly maneuverable robots suited for operating indoors, but are difficult to control due to their large number of actuators: 8 independently controlled motors where certain combinations of inputs cause the motors to oppose each other and drastically decrease efficiency. This study aims to gain new understandings about tackling the problem of overconstraint, having more inputs than the robot's degrees of freedom, while being intuitive for humans to control and capable of agile movements. A 2D simulation of a robot in a maze was created with different control modes with varying numbers of inputs, and participants were asked to submit their fastest time and rate each mode. The results show that many inputs slowed down operation considerably while few little inputs, although fast to learn, were too restrictive. These results agree with the existing body of knowledge that too many degrees of freedom in the controls decrease operator efficiency and suggest that there is an optimal number of inputs in between that is the easiest to learn, while reaping the full benefits of a maneuverable drivetrain.

Keywords

survey, simulation, swerve drive, overconstrained, intuitive controls

Introduction

Mobile robots with independently powered and steered wheels (swerve drive robots), can move in any direction and at any angle on a flat surface. Swerve drive wheels increase the agility of a robot, allowing them to maneuver in tight spaces. To compare, a car-like robot would not be able to quickly shift sideways to allow others to pass by. Therefore, swerve drive robots are well suited for indoors, working in corridors and navigating around dynamic obstacles such as people and other robots. They have seen use in hospitals as medical assistants, and as devices to manipulate radioactive objects in nuclear plants. (Tzafestas, 2014). Often, instead of operating completely autonomously, robots must be remotely operated by a human to ensure safe decisions and reactions to changing circumstances. However, the swerve drive system is difficult to control because it has many inputs (Siegwart, 2004). Swerve drive robots usually have 4 wheels in a rectangular configuration, with each wheel controlled with 2 motors: one motor for rotating the wheel on a horizontal axis propelling the robot forward, and one motor for rotating the wheel on a vertical axis to change the direction of propulsion. Each of these wheels is an input, a possible variable to be controlled. With 8 independent motors to manage speed and direction, it becomes very difficult for humans to operate the robot directly without additional software aid. According to studies in human-computer interfaces, humans tend to find more than input devices with many degrees of freedom, like a panel with too many buttons, to be confusing and difficult to learn (Swink, 2019). For remote operation of swerve drive robots, different control schemes have been developed to decrease the inputs for the human operator, making controls more intuitive. By improving the controls of swerve drive robots in particular, we can make indoor robot operation faster, safer, and reduce the cost of training human operators.

LITERATURE REVIEW

Transparency and Predictability

In robot teleoperation, also known as remote control, a common metric for the ease of controlling something is transparency (Fong et al., 2001). In this context, transparency is how well the human operator can create an accurate mental model of a robot and its surroundings so that they can predict what happens when a certain input is given to the robot (Sanders, 2014). This generates a sense of trust for the machine, so the operator is able to give commands with a clear expectation of how the robot will react (Wortham et al., 2017). For example, in a different study by C. Tzafestas et al. in 2008, participants were tracking a path using a joystick while watching a simulated robot in space on a monitor. The participants experienced more difficulty learning to control the robot with greater time delays, or lag, because the expected result was for the robot to always act according to the input immediately. As part of the study, the participants generated much worse mental models of the robot when the control system is ineffective. Applying this concept to specifically control a swerve drivetrain, we must design and program controls that novices can understand and know what to expect.

Connection to Game Controls

Although there have been many studies published about the stability and transparency of the swerve drive robot system based on physics models, there is little research done on what exactly makes a teleoperated robot easy to learn and model for a novice operator. To answer this, we turn to previous work from the design of game controls, which often deals with making novel characters and avatars controllable and intuitive. Teleoperation is very similar to game controls because the human operator acts for another being, be it a virtual character or physical robot miles away in a nuclear plant. By applying concepts from initiative game controls to robots, we can better design easy-to-learn controls for mobile robots. In his classic book about game design, Steve Swink explains two of the most common methods to make game controls more intuitive: let the player relate to the avatar, and bar users from giving inputs that do not do anything or are damaging (2009).

Relating to the Robot

To address the first point in robot controls, the human operator should be able to relate to the robot’s movements. In video games, humans have an extended proprioceptive sense (the sense that humans have that tell us where our body parts are), which lets the player's body relate to an avatar that can react to events that are happening in-game (Giddings, 2017). This means that for a more intuitive control of the video game characters, the artificial proprioceptive senses given by the game should match our proprioceptive sense in the real world, much like how physical robot controls strive for transparency, designing the output to match human expectations as much as possible. Often, this leads to the almost obvious design choice that pressing a key higher than the others in a computer game makes the avatar move forward or up, because the user is facing forwards towards the keyboard and screen. By this logic, robot operators should find it easier to remotely control a robot when a button to make the robot move in a certain direction is in the same direction that the robot will react, as shown on a screen or other feedback device. For indoor robot teleoperation, this feedback device is often a webcam mounted in a fixed location pointed towards the robot's area of operation, making it easy to locate the robot in the room (Shim et al., 2016). With this system, the robot's "forward" may not be the same direction as the operator's "forward". When designing the controls, there is a choice to let the "forward" button on the controls make the robot move left when the robot is facing towards to the left, or make the robot move forward as seen through the webcam. This is a choice between operating in the "robot reference frame" or "global reference frame", respectively (Siegwart, 2011). Since swerve drivetrains can move in any direction on demand, we can apply concepts from the design of game controls to hypothesize that a swerve robot will likely be easier to operate in the global reference frame.

Avoiding Invalid Inputs and Overconstrained Nature of Swerve

To address Swink's second point of making game controls easier to learn, swerve robot controls should likely avoid any possibility of invalid inputs that stall or damage the robot (2019). Swerve drive is unique because it is overconstrained, which is the mechanical property of a system to have more degrees of freedom in the input than degrees of freedom in the movement, making some combinations of inputs possibly damaging (Tzafestas, 2014). In the case of teleoperated mobile robots, degrees of freedom in the input refers to the number of variables an operator can modify to control the robot, and degrees of freedom in the movement refers to dimensions in which the robot can translate and rotate. Robots that have independently steerable and driven wheels, like swerve drive, have more controllable motors than degrees of freedom in movement, so are considered overconstrained. With swerve drive, there are 3 degrees of freedom: translation in 2 dimensions and rotation in 1 dimension, controlled by at least 6 motors. This means that there are damaging combinations of actuations, such as when motors are opposing each other, and operation may become difficult if the operator has full control of the individual motors. For example, the operator may accidentally steer the wheels in opposite directions, and regardless of how much power they send to drive the robot, the wheels will slip or the motors will stall.

In Wada (2000), the authors design a novel type of drivetrain that only uses 3 motors rather than 8, therefore eliminating the overconstraint and making wheel slipping mechanically impossible. However, the authors state that their drivetrain suffered from lower traction and balance due to there being 2 driven wheels instead of 4 driven wheels for the conventional swerve drive. If swerve drives are to be used in teleoperation, an intuitive control system will likely need to overcome the overconstraint problem without compromising traction or balance.

An existing solution to overconstraint is to use motion planning to make the controls become higher level commands (Fong et al., 2001). Higher level commands could mean that instead of a key being “move front left wheel in this direction”, it would be “move the entire robot in this direction and let the software calculate what the front left wheel should do.” In this way, the software is in charge of keeping the motor inputs away from ever putting the robot in a state where any of the wheels oppose each other. For example, in a study done by Dietrich et al. in 2012 at Stanford University, the control scheme implemented a potential field similar to that of magnets or electrical fields, which repel the inputs away from regions where they might command the robot to move in unrealistic or undesired ways. Human operators were free to control the robot in any way they wanted, but software quickly corrected for any damaging behavior. In another study by Oftadeh et al. in 2013, a swerve drive robot was controlled by always keeping the axes of the steered wheels towards a common point known as the instantaneous turn center of the robot, which makes sure none of the wheels apply forces that oppose each other. In this model, human operators command the entire robot to move in arcs (Oftadeh et al., 2013; Sorour, 2019). This was a higher level command than in Dietrich et al. and was more restrictive in the range of possible robot movements. Due to its higher level commands, the operator also loses some control over the exact movements of the robot, making it harder to react quickly to disturbances (Tzafestas, 2014). This is because there are a reduced number of degrees of freedom in the input, meaning there are fewer options.

Current Gap

There exists a large body of knowledge exploring the technical challenges of controlling an overconstrained robot and its implementations. However, little research has been done to test how these controls should be designed with the human operator in mind. Therefore, this study aims to explore the question: How can the controls of a teleoperated swerve drive mobile robot be designed to optimize the speed that novices can operate it in simulated indoor environments? This study will consider the effects of varying levels of degrees of freedom on transparency for human operators.

METHODOLOGY

Aim and Research Design

The overall purpose is to explore how controls of an over constrained swerve drive robot be designed to maximize operator satisfaction in the intuitiveness of operation in indoor environments. Many factors play a role in intuition, many of which cannot be described numerically, such as any previous experiences with similar controls and how connected they feel to the robot's movement. Quantitative data such as how fast the participant is able to complete the course is also useful to objectively compare different modes. Therefore, a case study was conducted because it allowed close collaboration with a just few participants to understand their motivations and opinions as an individual, rather than only viewing a broad range of opinions statistically. In this method, we can observe connections between qualitative and quantitative information within a single participant.

Consent and Ethical Issues

The online invitation asked each potential participant to complete a consent form to enter the study rather than replying to the invitation. In this way, other potential participants do not know whether someone chose to participate. The only identifying information about the participant is their signature on the consent form, which will not be shown to anyone else. The survey results are in a different anonymous form, so there is no way to link a participant's signature to their responses. All participants reported an age over 18, so no guardian consent was required.

Tools Used

A webapp in Javascript, CSS, and HTML is hosted on the web with Heroku. The survey was created and distributed with Google Forms.

Data Collection Procedure

First, a simple 2-dimensional robot simulation was created and hosted online. The physics simulation takes in the acceleration for each of the eight motors on the swerve drivetrain and outputs the velocities and pose of the whole robot using forward kinematics. Forward kinematics applies physics concepts to predict how the entire robot will react given the inputs to each of the 8 motors. A virtual controller was also implemented, which uses inverse kinematics to map keyboard inputs to desired speeds for each of the eight motors, then running a feedback loop to drive the current motor speeds to those desired speeds. By feeding the output of the virtual controller into the physics simulation at 50 times per second, the webpage is able to show the user a controllable model of a swerve drivetrain in real time.

To compare varying degrees of freedom, there were four control modes implemented by changing the key mappings and inverse kinematics in the virtual controller:

1. Each motor's steering was controlled directly by 2 keys on the keyboard, one for

forward and one for backward. A second pair of keys made all of the wheels roll forwards or backwards to move the robot. There are 5 degrees of freedom: one for each of the wheel steering angles, and one for the propulsion of all the wheels. This mode is in the "robot reference frame", meaning that forward and backward movements are relative to the direction the robot is currently facing (Siegwart, 2011).

2. A set of 4 keys controlled the front of the robot to move in the 4 cardinal directions. Another set of 4 keys controlled the back of the robot in the same manner. There are 4 degrees of freedom: translation on 2 axes for 2 sides of the robot. This mode is in the "global reference frame", meaning that the cardinal directions stay fixed regardless of the direction of the robot (Siegwart, 2011). Pressing the "right" key makes that part of the robot move towards the right side of the screen.
3. A set of 4 keys controlled the overall translation of the robot in the 4 cardinal directions. 2 more keys controlled the rotation of the robot. There are 3 degrees of freedom: translation on 2 axes for the entire robot and rotation on 1 axis. This mode is in the global reference frame.
4. A pair of keys controlled moving the robot forwards and backwards, and another pair for rotating the entire robot, similar to a car. There are 2 degrees of freedom. This mode is in the local reference frame.

Participants were challenged to move the robot from a "Start" point from the left side of the screen to the "Target" point on the right side. In between, there are static obstacles shown in dark gray representing indoor walls and furniture, and dynamic orange obstacles representing humans, animals, or other robots. The dynamic obstacles followed consistent paths but appeared at random times to simulate the flow of pedestrian traffic. This format of studying participants controlling a simulated robot in a maze was similar to (Green, 2008), another study that tested for robot teleoperation, though Green did not focus on swerve drivetrains or varying degrees of freedom.

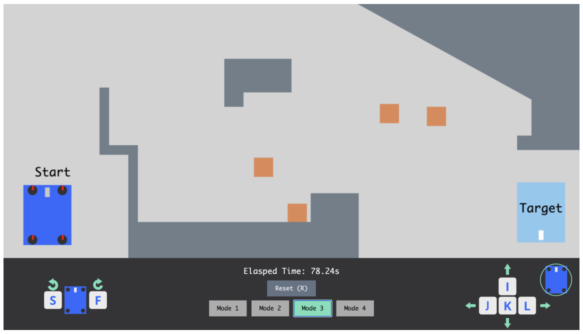


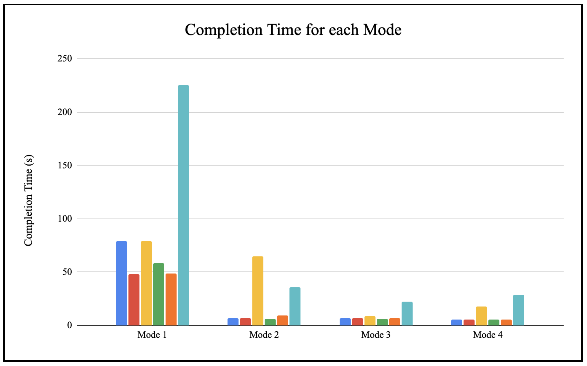
Figure 1. The robot simulation webpage at swervesim.herokuapp.com.

After participants completed the task for all four control modes, they answered a short survey about their experience, with a mix of quantitative and qualitative questions:

1. Fastest time for each control mode (Number)
2. Whether any controls in the simulation are similar to something experienced previously, and how (Open-ended)
3. Difficulty of learning the robot controls (Rank each mode 1-5)
4. Overall confidence in controlling the robot (Rank each mode 1-5)
5. Difficulty of reacting to random events (Rank each mode 1-5)
6. Preferred control mode (Select 1-4)
7. If they were to change any of the controls, how would they modify them? (Open-ended)
8. What control mode would you prefer, after your proposed modifications? (Select 1-4)

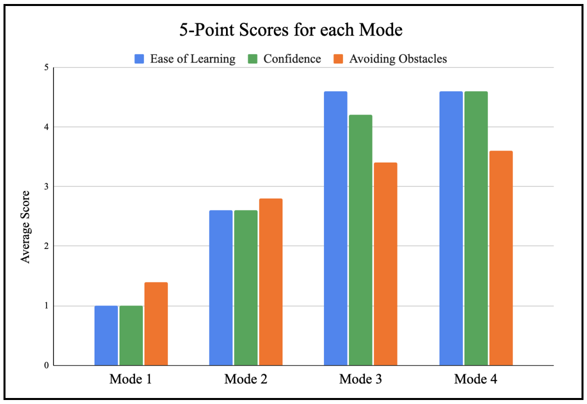
The survey asked for their ratings of various factors on a scale of 1-5 and a few short answers. This questionnaire is similar to Janabi-Sharifi et al.'s study in 2010, which measured the effect of lag on human operator performance on an obstacle avoiding robot.

Results

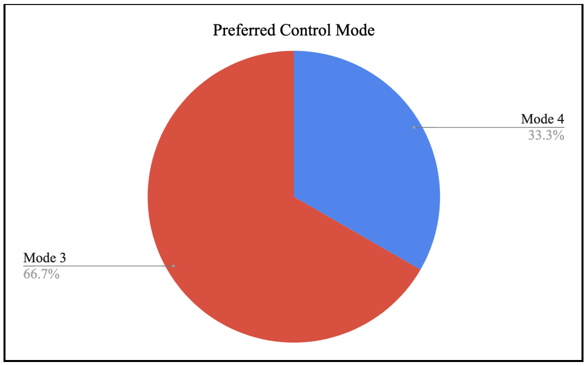


In comparisons of average time it took to complete the maze, Mode 1 was the slowest at 63 seconds, Mode 2 third at 19 seconds, Mode 4 second at 8 seconds, and Mode 3 fastest at 7 seconds. However, there was one outlier that took approximately 3 times longer than the other participants for all modes and another outlier who took 9 times as long for Mode 2. For this reason the median time will be used for comparison.

The median time for Mode 1 was the slowest at 58.04 seconds, Mode 2 third at 6.78 seconds, Mode 3 second at 6.58 seconds, and Mode 4 fastest at 5.78 seconds. All participants finished all modes, but Mode 3 had the least variation: the standard deviation was less than a second, while the rest had standard deviation above 5 seconds.



For the questions rated on a 5-point scale, Modes 3 and 4 had generally the same average score, with Mode 4 slightly better in difficulty learning and avoiding obstacles. They had the same results for confidence. The ease of learning, confidence, and ability to avoid obstacles are correlated with the completion time, with control modes being easier in ascending order, meaning that control modes with fewer degrees of freedom performed faster and were more well-received. In quantitative responses, participants conveyed that Mode 1 was difficult: 6 out of 6 said “very difficult” to learn how to control, 5 out of 6 said “not confident at all” controlling the robot in general, 4 out of 6 said “very difficult” reacting to dynamic obstacles. In miscellaneous comments, participants’ quotes included that “I will always crash things mode 1”, “Mode 1 seems nigh unsalvageable”, “Mode 1 was a bit painful”, “I actually didn't do Mode 1 again since I couldn't be pained to”, and “Mode 1 was too difficult to control”.



Participants saw modes 3 and 4 to be familiar, and rated those highly. 4 out of 6 participants compared mode 4 to be similar to “tank controls” in games, and 3 out of 5 responded that Mode 3 felt similar to the controls of physical vehicles. Even though Mode 4 was rated slightly better than Mode 3 using the 5-point scale, two-thirds of participants preferred mode 3 over mode 4 both with and without their proposed changes. One participant reasoned that Mode 4 felt more restrictive because the robot could not move sideways, diminishing the maneuverability of the drivetrain.

Discussion

In all aspects of this study, there was a trend that control modes with more degrees of freedom were more difficult to control. Modes with fewer keys to press completed faster, were easier to learn, instilled more confidence, and were better at avoiding obstacles, according to these survey responses. This concurs with Siegwart (2004) that machines with many inputs are difficult to control in general simply because of all the buttons and dials the operator must keep track of. Knowing the difficulty of excessive degrees of freedom, Wada’s mechanical solution in 2000 is rationalized because limiting the mechanical degrees of freedom to three avoids overconstraint entirely and becomes more intuitive to control.

The results that also relates with a theory about human nervous system development put forth by Bernstein (1966) that the human body is able to learn motor tasks with fewer degrees of freedom more quickly than those with more degrees of freedom, and relies on a large amount of practice to master the same task with higher degrees of freedom. Later, a study by

Newell (1996) showed that this is true in novice skiers, where at first the person learns to move the minimum number of joints to complete a motion, but as they practiced more, they were able to increase their range of motion by rotating more joints in their body. In this study, participants were not given much time to practice and master each control mode, so they were able to learn only the modes with fewer degrees of freedom. Curiously, one participant felt that they could master Mode 2 with more practice, which agrees with Newell (1996) because Mode 2 is very similar to Mode 3 but with more degrees of freedom and a larger range of motion, similar to how the skiers could master larger ranges of motion in their own joints with practice.

In addition, this study supports existing research in that familiar controls tend to perform better. In Brown et al. (2010) and Swink (2019), the success of a game controller design is heavily influenced by its similarity to previous game controllers, even when the design is not inherently good. Humans are able to learn more effectively if we make connections to past experiences. Although just a correlation in this study, survey results show that those who found Modes 3 and 4 familiar to video game controls preferred those two modes and had faster completion times.

However, this study does not fully align with the existing body of knowledge. According to Tzafestas (2014), fewer degrees of freedom and a smaller range of motion make robots less maneuverable and therefore less adept to reacting to changing circumstances. In this study, this idea should have shown a worse ability to avoid dynamic obstacles in Mode 4, which has only 2 degrees of freedom compared to 3 degrees of freedom in Mode 3. However, on a 5-point scale, participants responded with approximately equal scores, with Mode 4 being slightly better at avoiding obstacles.

Results of this study can raise questions beyond controls of indoor swerve drive robots. Designing any human-computer interface involves the decision of the degrees of freedom desired. For example, an interview with Apple designer Abraham Farag reveals a conflict within the company during the release of Macintosh where there was controversy over the mouse that contained only a single button (Farag, 2014). Critics argued that the single button was too restrictive and did not provide enough control for the user, while proponents argued that the single button was simple and intuitive. Further research can be done to explore the many different design clashes between few and many degrees of freedom and when each has been successful.

Conclusion

This study focuses on designing effective and easy to learn controls for indoor swerve drive robots and found that control modes with fewer degrees of freedom were beneficial. However, there was a small sample size, and the study had a limited scope of local students, who could have different experiences with controls than the global workforce population, which has a median age of around 40 (ILO, 2019). In addition, the study was limited to only 4 control modes on a 2-dimensional simulation all using keyboard input due to the short online survey format. Therefore, further research could be done to support these results by sampling a more representative sample of the workforce that may operate swerve drive robots and provide more tools for input such as joysticks, game controllers, or VR sets in an in-person setting for realism.

Limitations

The study uses convenience sampling. Due to the 2021 pandemic, all contact was online. An online invitation was posted to participate in the study through the free messaging application Discord. The invitation was intentionally not sent to robotics-related servers to avoid overrepresenting those already with experience operating robots. However, all respondents were under the age of 18 and only 6 completed the survey with appropriate parental consent (in the form of a digital signature).

The modes with fewer controlled degrees of freedom were numbered higher and it is likely that the participants went through the modes in numerical order, so they would have more practice with the robot simulation in general by the time they reached mode 4. This is a confounding variable that might encourage the result that fewer degrees of freedom causes easier controlling.

Only a keyboard was used, while in the real world it is likely that these robots are controlled with touchscreen, mouse, joysticks, VR sets, and various other controllers with analog inputs. With a keyboard, one can only send inputs as forward, backwards, or off.

This study only focused on four different control modes designed by the author, while there are infinitely many more possible control systems. So, we cannot confidently generalize results from one example of a control mode to an entire group of control schemes. Also, participants did not actively contribute to the design of the controls, so the survey responses for the possible improvement of the control modes were not implemented.

While the author tries to be as unbiased as possible when analyzing the responses, there may still be some positionality issues. The author designed and programmed the four robot control modes in the simulation, and some took more effort to finish. This may cause bias because the author may expect that a more complex system is more effective.

Acknowledgements

This paper would not have been possible without the guidance of my AP research teacher and all the students who gave valuable feedback throughout the writing process.

References

Bernstein, Nikolai. "The co-ordination and regulation of movements." The co-ordination and regulation of movements (1966).

Brown, Michael, et al. "Beyond the gamepad: HCI and game controller design and evaluation." Evaluating user experience in games. *Springer*, London, 2010. 209-219.

Dietrich, Alexander, et al. "Reactive whole-body control: Dynamic mobile manipulation using a large number of actuated degrees of freedom." *IEEE Robotics & Automation Magazine* 19.2 (2012): 20-33.

Farag, Abraham. (2014, March 9). Steve Jobs Hated The Idea Of A Multi-Button Mouse, Designer Claims. *The New Yorker.* https://www.cultofmac.com/269222/steve-jobs-hated-idea-multi-button-mouse-designer-claims/

Fong, Terrence, and Charles Thorpe. "Vehicle teleoperation interfaces." *Autonomous robots* 11.1 (2001): 9-18.

Giddings, S. (2017). The phenomenology of Angry Birds: Virtual gravity and distributed proprioception in video game worlds. *Journal of Gaming & Virtual Worlds*, 9(3), 207-224. https://doi.org/10.1386/jgvw.9.3.207_1.

Green, Scott A., et al. "Human-robot collaboration: A literature review and augmented reality approach in design." *International journal of advanced robotic systems* 5.1 (2008): 1.

ILO. (July 31, 2019). Median age of the global labor force from 1990 to 2025 (in years) [Graph]. In *Statista*. Retrieved March 30, 2021, from https://www-statista-com.ez.pausd.org/statistics/996530/median-age-global-labor-force-years/

Janabi-Sharifi, Farrokh, and Iraj Hassanzadeh. "Experimental analysis of mobile-robot teleoperation via shared impedance control." *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 41.2 (2010): 591-606.

Oftadeh, R., Ghabcheloo, R., & Mattila, J. (2013). A novel time optimal path following controller with bounded velocities for mobile robots with independently steerable wheels. *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4845-4851, https://www.doi.org/10.1109/IROS.2013.6697055.

Sanders, Tracy L., et al. "The influence of modality and transparency on trust in human-robot interaction." *2014 IEEE International Inter-Disciplinary*

Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA). IEEE, 2014.

Shim, Jae Hong, and Young Im Cho. "A mobile robot localization via indoor fixed remote surveillance cameras." *Sensors* 16.2 (2016): 195.

Siegwart, R. (2004). Introduction to Autonomous Mobile Robots. MIT Press.

Sorour, M., Cherubini, A., & Fraisse, P. (2019). Motion Control for Steerable Wheeled Mobile Manipulation. *2019 European Conference on Mobile Robots (ECMR)*, 1-7, https://www.doi.org/10.1109/ECMR.2019.8870958.

Swink, Steve. (2009). *Game Feel: A Game Designer's Guide to Virtual Sensation.* Morgan Kaufmann.

Tzafestas, S. (2014). Introduction to Mobile Robot Control. Elsevier.

Tzafestas, Costas, Spyros Velanas, and George Fakiridis. "Adaptive impedance control in haptic teleoperation to improve transparency under time-delay." *2008 IEEE International Conference on Robotics and Automation*. IEEE, 2008.

Wada, M., Takagi, A., & Mori, S. (2000). Caster Drive Mechanisms for Holonomic and Omnidirectional Mobile Platforms with no Over Constraint. *Proceedings of the 2000 IEEE International Conference on Robotics & Automation*, 1531-1538. https://www.doi.org/10.1109/robot.2000.844814

Wortham, Robert H., and Andreas Theodorou. "Robot transparency, trust and utility." *Connection Science* 29.3 (2017): 242-248.