We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Real-Time Sensory Information for Remote Supervision of Autonomous Agricultural Machines

Daniel D. Mann, Uduak Edet, Mitchell Green,

Olayinka Folorunsho, Avery Simundsson and Franklin Ogidi

Abstract

The concept of the driverless tractor has been discussed in the scientific literature for decades and several tractor manufacturers now have prototypes being field-tested. Although farmers will not be required to be physically present on these machines, it is envisioned that they will remain a part of the human-automation system. The overall efficiency and safety to be attained by autonomous agricultural machines (AAMs) will be correlated with the effectiveness of information sharing between the AAM and the farmer through what might be aptly called an automation interface. In this supervisory scenario, the farmer would be able to both receive status information and send instructions. In essence, supervisory control of an AAM is similar to the current scenario where farmers physically present on their machines obtain status information from displays integrated into the machine and from general sensory information that is available due to their proximity to the operating machine. Therefore, there is reason to expect that real-time sensory information would be valuable to the farmer when remotely supervising an AAM through an automation interface. This chapter will provide an overview of recent research that has been conducted on the role of real-time sensory information to the task of remotely supervising an AAM.

Keywords: autonomous agricultural machines, remote supervision, automation interface, visual information, ergonomics

1. Introduction

For several decades, university researchers have devoted time and effort to the pursuit of developing a driverless tractor. The scientific literature contains numerous articles describing various technologies that were evaluated, challenges that were encountered, conceptualizations of what future autonomous agricultural machines (AAMs) might look like, and issues (both technical and non-technical) in need of redress. In the late 1990s, the "agricultural ergonomics laboratory" was established at the University of Manitoba based on the hypothesis that engineers would continue to incorporate increasing levels of technology into agricultural

machines in pursuit of the ultimate goal of the fully autonomous machine. Based on the lessons learned when automation was introduced to other industry settings, the human operator would experience a changing role. Thus, there was a need to view agricultural guidance technologies from an ergonomic perspective.

Unlike two decades ago, it is now possible to find autonomous tractors that are either available for sale to farmers or are in final stages of field testing. Conceptually, there are still at least four distinct designs being promoted (**Table 1**). There are advantages associated with each of these four distinct types of AAMs. Those which retain the operator station provide flexibility to the farmer for those instances when it is desired that the human operator be physically present on the AAM; this is perhaps most critical in the early days when AAMs are being introduced to the market. AAMs that resemble current tractors (except for the operator cab) and attach to implements in the same manner as existing tractors will reduce the capital cost associated with transitioning to autonomous agricultural production because the farmer will be able to continue to use existing implements. The integrated tractor reflects the situation where the engineer will be able to optimize the design of the tractor-implement system; it potentially enables design opportunities not present with the current paradigm of a tractor pulling an implement (which is a hold-over from the early concept of a horse pulling an implement). The downside, of course, is that a whole new set of implements will be required as the farmer's existing implements will be incompatible with the integrated tractor. The swarm or fleet concept perhaps reflects the most radical concept, reversing the decades-long trend of building bigger agricultural machines. Perhaps inspired by the insect world, the concept is that a fleet of many small AAMs working in an organized manner can outperform a small number of large-sized AAMs. It is too early to predict whether a single concept will emerge as the industry standard, or whether all of these concepts will survive either in niche applications or in direct competition with one another.

Regardless of how the AAM industry evolves, it would be foolish for designers to neglect how these autonomous machines will interact within the larger human-autonomy system. It is inevitable that the AAM will need to interact with a human supervisor to receive instruction and to request assistance when problems cannot be self-corrected by the AAM. Appropriate principles from the discipline of human factors engineering will be essential to the successful integration of AAMs into production agriculture.

Concept	Example	Main Features
Retain operator station	Monarch Tractor	• Tractor unit attaches to implements in the traditional manner. • Small-sized tractor targeted towards orchard/vegetable production or livestock operations.
Eliminate operator station	CNH, John Deere, Kubota, Autonomous Tractor Corporation	• Tractor unit attaches to implements in the traditional manner. • Large-sized tractors to replace the large-sized tractors currently being used for cereal/oilseed production on the prairies.
Integrated tractor	DOT	• Tractor unit is integrated with the implements that it powers.
Swarm/Fleet	Fendt Xaver	• Tasks completed by a swarm or fleet of extremely small machines working in an organized manner.

Table 1.
Autonomous tractor concepts being developed by various companies world-wide.

2. Supervision of autonomous agricultural machines

Supervision is an activity that is undertaken for the purpose of ensuring that a task is done in such a way that it meets our approval (in terms of safety, in accordance with rules, etc.). We would find it absurd to hire a junior employee and not provide some means for supervision of their work. Even senior employees require supervision to ensure that they are held accountable for their performance. The same need for supervision applies to the AAMs currently being developed by engineers. Autonomous machines, though independent, still require human supervision [1, 2] to help minimize any catastrophe that may arise in case of unexpected situations such as system failure or malfunction that exceeds the capability of the machine [3]. Furthermore, since it is currently difficult to automate high-level reasoning and tasks, it is also beneficial that the human remains in the decision-making loop to assist with planning field operations, allocating resources, and coordinating the autonomous machines. Generally, involving the human (as a supervisor) in an autonomous system has been reported to increase the overall reliability and performance of the system [4].

Supervision can be carried out in proximity (where the supervisor and the system being supervised are collocated) or remotely (where the supervisor performs his/her roles from a distant location without being physically present in the work zone). Currently, supervision of agricultural field machines is mainly performed in proximity (i.e., with the operator seated in a cab on the machine), but it is envisioned that future AAMs will be supervised remotely due to farm labour shortages (i.e., enabling one person to supervise multiple AAMs) and to enhance the overall efficiency of the farmer (i.e., enabling the farmer to complete other farm management tasks while supervising AAMs in the field).

Remote supervision is not novel and has been practiced for decades in different sectors. There is evidence of remote supervision being used in non-agricultural sectors such as military, space exploration, marine, industrial applications, and rescue operations [5–8]. As an example, robots that are used to inspect pipelines for cracks are monitored remotely during operation [9] since these areas are not accessible by humans. Search and rescue robots and military drones have been monitored remotely [6, 10]. In both cases, supervisors make use of some type of interface (which may be portable or stationary) to monitor the robot and to receive status updates. In agriculture, remote supervision has been used in livestock husbandry, crop production, and crop storage [8]. Unmanned aerial vehicles (commonly known as drones) have been remotely monitored while using them to determine weed infested regions on the field. In a hog barn environment, the physiology of pigs (body temperature) and environmental conditions of the barn (air temperature, humidity, and concentration of carbon dioxide, hydrogen sulfide, and ammonia) have been remotely monitored to minimize contact with the pigs and to assist the farmer with decision-making from any location [11].

Several remote supervision concepts for autonomous agricultural field machines have been proposed by academic researchers and manufacturers alike. These concepts differ with respect to the type of human involvement, autonomy level, proximity of the remote supervisor to the autonomous machine, and number of autonomous machines being supervised simultaneously [12]. For example, [13] envisioned the human to only monitor autonomous machines, whereas [14] expected the farmer to both manually operate a field machine while supervising another autonomous machine (which may or may not be the same type of machine). A third supervision concept involved manually controlling the actual operation of a field machine remotely (i.e., teleoperation) [15]. Some researchers [2, 16] proposed

that the human would monitor just one machine while others [17, 18] envisioned the supervisor to monitor several machines simultaneously.

Edet and Mann [12] described four remote supervision concepts based on the location of the remote supervisor in relation to the AAM: 1) in-field supervision, 2) edge-of-field supervision, 3) supervision from the farm office, and 4) supervision from outside the farm site. A practical example of the 'in-field' supervision concept is the human-machine, master–slave interaction that involves having both an AAM and a human-driven machine working simultaneously on the same field. Supervision of the AAM would be done from an interface located in the cab of the human-driven machine. In the 'edge-of-field' remote supervision concept, the farmer is not operating any of the machines. This gives the farmer the opportunity to also be involved with the logistics of the operation such as bringing supplies to the field, making repairs, and responding to alerts. The 'supervision from the farm office' concept, on the other hand, makes it more challenging for the farmer to respond to in-field demands in a timely manner; the advantage is that the farmer can attend to other non-field related tasks rather than focusing on monitoring the AAM alone. Supervision from outside the farm site would theoretically allow a farmer to remain engaged in field operations while physically away from the farm for personal or vocational reasons, although it would be challenging to address system malfunctions. This role would need to be delegated to someone else, potentially contracted to an agency that would monitor and service AAMs for a fee.

Each remote supervision concept has corresponding benefits and shortcomings. For example, the 'in-field' concept would likely have the shortest response time, however, if the AAM breaks down or requires assistance, the entire field operation may come to a standstill since the manually-operated machine will also be stopped as its' operator handles the problem. In the 'edge-of-field' and 'from the farm office' remote supervision concepts, the farmer would not be controlling any of the AAMs in operation. Hence, the farmer is available to manage both i) malfunctions that are beyond the capability of the AAM and ii) other logistics associated with the operation without assistance from other farm workers. Of these two remote supervision concepts, 'edge-of-field' supervision may be preferred over 'from the farm office' supervision because of the closer physical presence to the AAM which, in theory, should allow for faster response to malfunctions that require human intervention. Remote supervision concepts that rely on servicing of AAMs being done by professional service technicians in a fee-for-service arrangement may not be accepted by many farmers due to a preference for self-sufficiency and the timely manner in which many farm operations need to be completed.

Generally, a suitable remote supervision concept should: i) require minimal labour to function, ii) enable the farmer to monitor and understand the status of the operating machine in the field, iii) not restrict the movement of the farmer, iv) allow the farmer to perform other farm tasks, iv) enable the farmer to attend to in-field problems in a timely manner, and vi) be cost effective. Other factors that may influence the choice of remote supervision for monitoring the operation of AAMs include the size of the farm, ease of use of the automation interface, type of field operation being conducted, business structure of the farm, the farmer's preference, and future legislation that might relate to the supervision of AAMs. Based on an unranked paired comparison analysis of the concepts, the 'edge-of-field' remote supervision concept was determined to be the most viable remote supervision concept for broadacre grain producers [19].

It can also be deduced that remote supervision of AAMs requires an automation interface since it is the communication link that enables the human supervisor to interact with the AAM. Edet [19] generated the following list of functional requirements for an automation interface; the remote supervisor should be able to:

1. Instruct the AAM to commence operation.

2. Monitor telemetrics of the AAM.

3. See key elements of the AAM in real-time.

4. Visualize the position of the AAM within the field.

5. Receive notifications of important events and anomalies from the AAM.

6. Query the AAM about planned actions.

7. Instruct the AAM to stop or shut down, or to alter plans.

In 2017, an important article entitled "From here to autonomy: lessons learned from human-automation research" was published by a leading expert in human-autonomy teams [20]. The key to a successful human-autonomy team is to assume that there will be instances where the autonomous system will require input from the human supervisor who is part of the human-autonomy team, and to ensure that the autonomous system is designed to most effectively share critical information with the human member of the human-autonomy team. In essence, [20] recommended that there should be shared situation awareness within the human-autonomy team where the human supervisor fully understands the actions being taken by the autonomous system so that appropriate actions can be taken by the human supervisor at any instant. Designing to support shared situation awareness is a non-trivial undertaking for the design engineer. Most autonomous systems require substantial complexity to fully automate the various tasks associated with the overall functioning of the machine. Passive monitoring of automation creates a high workload for the farmer [21] – this likely contradicts one of the reasons for using AAMs in production agriculture in the first place (i.e., to reduce the workload for the farmer). In her "human-autonomy oversight model", [20] recommended that a transparent automation interface be designed so that the human responsible for supervision of the automation will be able to successfully navigate from periods of passive supervision to periods requiring intervention. The next section of the chapter will focus specifically on the automation interface.

3. The automation Interface

3.1 The role of the automation interface

For decades, virtually all textbooks that have been written on the topic of ergonomics or human factors engineering have had chapters devoted to the design of displays and the design of controls. Displays must be designed well to clearly convey machine status information to the operator. The design and arrangement of controls is essential to allow efficient communication of instructions from the operator to the machine. When dealing with an autonomous machine, there is perhaps limited reason for the supervisor to need to communicate short-term actions to the autonomous machine. It is more reasonable to expect that communication in this direction will be reserved primarily to high-level management decisions. However, the flow of information from the autonomous machine to the human supervisor is anticipated to remain important.

The key to a successful system comprised of an autonomous machine and a human supervisor is a well-designed interface that allows for the exchange of information between the autonomous machine and the human supervisor. There are several papers published over the past two decades that touted the importance of an automation interface and have postulated on the features essential to an automation interface. In a paper published two decades ago, [22] explained their expectation that the farm manager would be responsible for overseeing the coordinating process from a computer located in the farm office. They proposed the term of 'tractor mimic display' for the automation interface that would be used to display telemetric data from the tractor unit, show the position of the tractor unit on a map, and display real-time video as seen through steerable cameras placed on the tractor unit. In the same year, [16] published a paper that investigated how humans can supervise AAMs. These authors discussed the challenges associated with designing an autonomous system that avoids both false positives and false negatives. Although they stated the desire to design such to err on the side of false positives (i.e., where a machine sees a problem where there are none), they further suggested the use of humans as 'remote troubleshooters' to classify positives as either true or false. Their system was designed to transmit images of the scene whenever the tractor detects an obstacle in its way; the images were presented to the remote troubleshooter using a 'remote operator interface'. In addition to sharing telemetric data and live video, the interface provided a warning when an obstacle was detected and explained what portion of the image was being classified as an obstacle. In a more recent paper, [23] described work completed to develop a team of robotic tractors for autonomous peat moss harvesting. In manual peat moss harvesting, a team leader supervises a team of three of more tractor operators using radio and/or hand signals. The autonomous peat moss harvesting system mimicked the manual harvesting system in that the human team leader communicated with the autonomous harvesters through a 'team leader user interface'. In this instance, the interface displayed telemetric information from each of a team of autonomous harvesters. Furthermore, a map was used to show the position of each harvester and to provide a visual representation of harvesting progress. Moorehead et al. [24] described a system of autonomous tractors for orchard maintenance. The autonomous system was comprised of tractors (equipped with perception systems and capable of driving autonomously) and a remote supervisor who assigns tasks, responds to requests when the perception system is unable to decide how to deal with a detected obstacle, and tracks the fleet of autonomous tractors. Although the tractors were equipped with cameras and the remote supervisor's interface was designed to display video, it was not intended that the supervisor should monitor the real-time video continuously. Rather, a warning message appears when an obstacle is detected and the tractor has stopped forward motion; the supervisor must then review the available video and decide whether a worker needs to be sent to remove the obstacle or if the warning is a false positive meaning that the tractor may proceed safely.

Based on this brief review of automation interfaces that have been reported in the published literature, there are several common elements that are envisioned for an effective 'automation interface'. First, it is anticipated that the automation interface will provide telemetric data related to the autonomous agricultural machine; such information is necessary to assure the human supervisor that the machine is functioning within normal operating parameters. Second, there is a need to show the location of the autonomous machine within the context of its operating environment (i.e., field, orchard, peat bog, etc). Third, it is envisioned that the autonomous machine will, at times, experience situations which will require human intervention. In these instances, a warning message will be

displayed for the supervisor. The autonomous machine will stop until the issue has been resolved by the supervisor and the machine is cleared to resume operation. To enable the supervisor to be able to see what is happening, cameras are necessary on the AAM to transmit real-time video which is viewed on the automation interface.

3.2 Identifying information to include on an automation interface

If the automation interface contains irrelevant information, this may result in overcrowding that could reduce the effectiveness of the interface. On the other hand, omitting essential information may impede the supervisor's ability to perform his or her role effectively. Thus, providing the supervisor with the right information is central in designing an effective user interface [25]. Identifying the right information can be achieved through the completion of a requirement analysis which involves identifying and understanding the goals of the task as well as the role of the user [26]. Endsley [27] noted that the supervisor should have a high situation awareness (i.e., "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future") of the automated machine to be able to take necessary action in a timely manner. This awareness includes: i) machine location awareness, ii) activity awareness, iii) status awareness, iv) surrounding awareness, and v) overall mission awareness [28].

The results of a requirement analysis that was conducted for the task of supervising an agricultural sprayer have been included to demonstrate this process. Both users and designers were consulted to inform the design of an automation interface for an autonomous agricultural sprayer. **Table 2** provides a summary of the information arising from a survey of the farming community [8] which was structured to determine the types of information that should be included on an automation interface in order to remotely supervise an autonomous agricultural sprayer. Parameters listed under the 'very useful' column were recommended by at least 75% of the respondents; [8] concluded that these pieces of information should be included in an automation interface. In further work towards the requirement analysis, [19] consulted with expert designers of AAMs. Parameters such as fuel level, tire pressure battery status, current location, global field (coverage map), tank level, spray pressure, application rate, nozzle status, and boom height were ranked as being essential information for an automation interface by the majority of designers interviewed.

A majority of the users and designers consulted in the completion of the requirement analysis indicated that live video footage of the autonomous sprayer should be made available to the remote supervisor [8, 19]. Panfilov and Mann [29] had previously investigated the importance of live video footage to the remote supervisor of an autonomous sprayer in an experimental study completed using a simulator in a lab environment. They reported that live video provided a sense of security to the supervisor, but was not typically used to detect malfunctions. They also noted that the supervisor spent only 30% of their time viewing the video. In their experimental study, participants spent the majority of their time monitoring telemetric data displayed using traditional display elements (i.e., gauges, dials, etc.). Finally, they suggested that it might be appropriate to provide real-time video on-demand. Edet [8] reported that some respondents felt that providing only one view of the sprayer was not enough to properly understand the entire spraying operation – they suggested having multiple views of the machine during operation available through an automation interface.

Information	Descriptions			
	Very useful	**Useful**	**Least useful**	**Other suggestions/ comments**
Machine status	Engine temperature, engine speed, fuel level, oil pressure, hose leakage, boom folding (open/close), and agitator	Tire pressure		Auto-steer status, GPS status, slippage
Spraying functions	Boom height, nozzle status, area covered, spray pressure, application rate, travel speed, wind speed, and wind direction	Daily temperature, skip/double application, delivery rate, current task	Humidity and altitude	Field condition, tank level, chemical mix of what is in the tank, sectional control, gallon per hour sprayed, acres covered, area sprayed per coverage, number of fills, and droplet size
Navigation features	Route taken, current location, and overhead view	Distance travel	Compass	Planned route, and coverage map
Warning and notifications	Plugged nozzle, machine breakdown, Obstacle detection, loss of GPS signal, and unexpected shutdown	Tank level drop, fuel level drop, and route change	Task completed, and skip/ double application	Emergency shutdown,

Table 2.
Interface requirements as determined by the farming community through survey. >75: Very useful, 51–75%: Useful, 25–50%: Least useful, and < 25: Not at all useful (adapted from [8]).

4. Practical considerations associated with incorporating real-time visual information in an automation Interface

4.1 Determining the appropriate 'look zones'

The previous section established the need to provide the remote supervisor of an AAM with real-time video showing the machine and its environment [8, 29]. Real-time video helps the supervisor to better understand abnormalities within the AAM [29]. Blackmore [22] also noted that the presence of live video will enable the remote supervisor to understand the machine's environment. These statements are evident in studies carried out by [2, 16, 28, 30, 31]. With these potential benefits, the challenge is to determine where the cameras should be placed to maximize the benefit of the real-time visual information.

Operators of conventional tractor-seeding machines visually monitor seven distinct areas of the machine and its environment (termed 'look zones') [32]. These areas included: i) forward, ii) right side, iii) planter, iv) planter edge, v)

display X (located at the top right corner from the operator's seating position), vi) display Y (located close to the arm rest of the operator), and vii) other (located close to the front-left tire of the tractor). Other researchers [33] identified four sectors: i) field ahead, ii) left boom, iii) right boom, and iv) the light bar while investigating the "workload associated with operating an agricultural sprayer equipped with a navigation device." Hence, it can be inferred from these studies that the visual information that is useful to operators can be derived primarily from i) displays located inside the machine's cab, ii) external field cues, and iii) the implement.

Although these studies identified the different regions of importance, they did not describe what information was gained by viewing those regions. Hence, a study was conducted to identify what visual information about the machine and its environment would assist the remote supervisor to make decisions [34]; the primary focus of this study was the high-clearance sprayer. GoPro cameras were mounted at difference locations on a sprayer to record the sprayer and its environment while in operation (**Figure 1**). After collection of this video footage, 29 experienced operators (defined as having at least two years experience as a sprayer operator) were recruited and presented with 10 distinct video clips of the high-clearance sprayer in operation (**Figure 2**). For each video clip, the operators were asked i) to describe what they saw in each video clip, ii) to describe the information gained from the viewing the video clip, and iii) to rank the importance of the visual information perceived using a 5-point Likert scale.

Not all video clips were equally effective at providing information that was relevant to the spraying operation. Among the 10 clips, clips 1 and 2 were considered very important and extremely important, respectively, by the operators while clips 3 through 7 did not provide much information that was considered relevant to the operators. The results of data analysis revealed that experienced operators generally preferred the view from the i) boom and nozzles, ii) the view ahead of the sprayer (front view), and iii) an aerial view of the sprayer. These regions enable them to determine the field/crop variability or conditions, upcoming field information (e.g., headland, obstacles) as well as assess if the sprayer is functioning properly and spraying effectively. Specifically, the information that was perceived from these views included the spray pattern, nozzle height and status (plugged or not), spray pattern/drift, obstacles in front and beside sprayer, poor areas in the field (i.e., crop condition), wet spots, approximate travel speed, headlands, type of crop being sprayed, weather (windy and sunny), location of sprayer in the field, overall picture of the field (aerial view), and if the sprayer was moving and following the right path (i.e., moving straight). This result was found to be independent of how
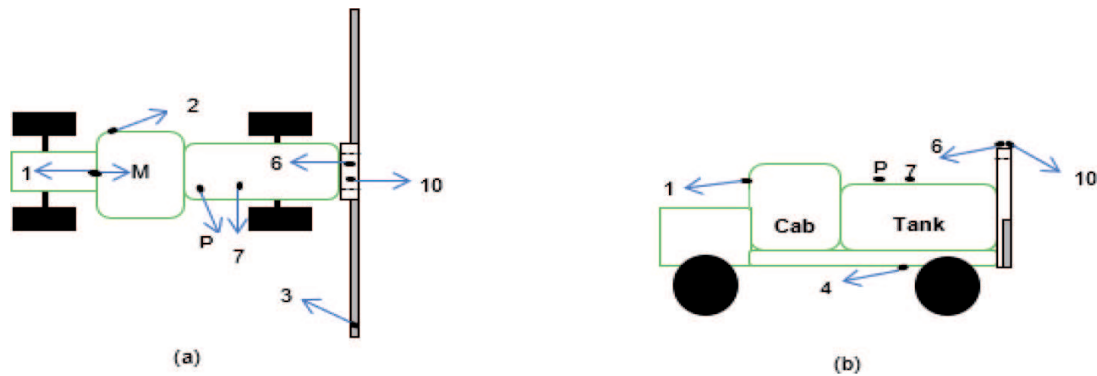

Figure 1.
A schematic showing the various camera positions on the high-clearance sprayer during the spraying operation (a) plan view (b) side view. This figure was originally published in [34].

Figure 2.
Screen shots of clips 1–10. This figure was originally published in [34].

frequently a particular view (or information) was presented. A visual representation of the information gained from the operator's 'look zones' is shown in **Figure 3**.

Camera angle and position influenced what information the operators perceived from the video clips [34]. Participants tended to describe features that were more prominent within the frame of the camera in comparison to less-prominent features. For example, many participants described features that were associated with the spray boom in clip 2 since the clip focused mainly on the right boom. Similarly, clip 9 emphasized the field and correspondingly most participants focused on the relevant information that was gained from the field. This finding suggests that designers can influence which features the user will perceive by positioning the camera such that those specific features are prominent in the camera's field of view.
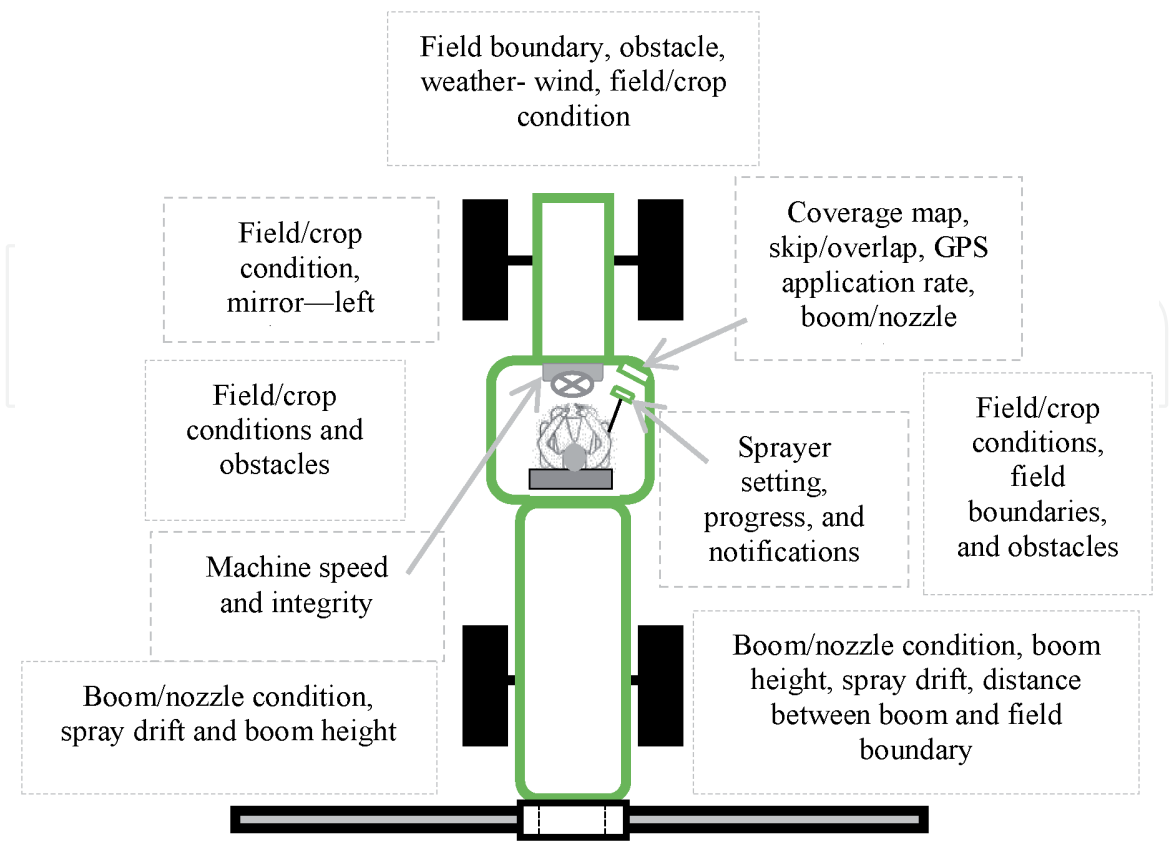

Figure 3.
Visual representation of the information gained from the operator's seated position in each of the look zones. This figure is adapted from a figure originally published in [34].

One of the most suggested views was to have a camera at one end of the boom facing forward. However, when analyzing the videos, this view was found to have minimal difference from the forward view for a sprayer with a long boom. Other views that were suggested included i) a close-up view of the nozzle and its tip, ii) a view from under the sprayer facing backward to see the spray pattern behind the sprayer and wheels, iii) dashboard/displays, and iv) a camera that would focus on the wheel to show how well the sprayer was either following old tire tracks or steering within the rows of a crop.

4.2 Effect of camera placement on the usability of look-ahead visual information

The importance of real-time visual information in the autonomous interface has been demonstrated in the previous sections, however, it is also necessary to consider how camera placement (i.e., camera height and camera tilt angle) influences the usefulness of this visual information. Previous research by [35] investigated the impact of camera placement on guidance performance for a manual guidance task in which the tractor operator relied on visual information provided by an implement-mounted camera that was displayed on a monitor close to the operator's seat. Tang and Mann [35] described a phenomenon that they called 'image velocity' which quantified the rate at which the visual information scrolled across the monitor from top to bottom as the tractor drove forward through the field. Image velocity is based on the camera's optical parameters, placement of the camera on the implement (height and tilt angle), and the tractor velocity. The reader is directed to [35] for a thorough description of how the parameter of image velocity was calculated.

The results published by [35] did not provide definitive evidence of a relationship between image velocity and lateral guidance error, however, trends were observed with lateral error increasing with image velocity. Test participants self-evaluated their performance following each trial; these results showed a decreasing linear trend with increasing image velocity. Participants preferred a tilt angle of 20° below horizontal as this gave them the best look-ahead view (i.e., the greatest look-ahead distance); however, the 30° tilt angle yielded the statistically smallest guidance error. It is unknown how the prior research by [35] will inform the current task of placing cameras on an AAM for the purpose of remotely supervising the machine on an automation interface because their research was focused primarily on trying to minimize lateral error associated with a manual guidance task. Nevertheless, their prior research inspired subsequent studies intended to determine the impact of camera placement on the usability of the visual information for the task of remotely supervising an AAM.

A lab experiment was conducted in which test participants were asked to watch pre-recorded video clips as a means of obtaining real-time visual information from the field (simulating the task of remotely supervising an AAM) [36]. Video footage was pre-recorded for nine unique combinations of camera placement, namely three camera tilt angles (20, 30, and 40°) and three camera heights (0.5, 1.0, and 1.5 m), to yield nine different look-ahead situations. Participants, some of whom were inexperienced agricultural machinery operators recruited from the university student population and some of whom were farmers experienced in operating agricultural machines, were asked to complete two distinct experimental tasks. First, they were asked to choose their preferred look-ahead position after watching two unique video clips playing side-by-side on the screen (**Figure 4**). Second, the participants responded to questions that would help determine the effect of camera placement on the difficulty of detecting and interpreting the randomly-placed frisbees in the video clips watched (**Figure 5**).

Figure 4.
Screenshot of look-ahead videos side by side (part one).

Figure 5.
Screenshot of video footage with frisbees appearing on the soil (part two).

An unranked pairwise comparison was used to analyze the data from part one of this study. This is a decision-making tool in which alternatives are compared to each other, one at a time, to arrive at the best choice. Each alternative is considered relative to other options available, with a value of one assigned to the more desirable option and a value of zero assigned to the less desirable option to arrive at alternative choice coefficients for each option being considered [37, 38]. Using this methodology, participants made pairwise comparisons for all nine look-ahead combinations. For both groups of participants (i.e., university students and experienced sprayer operators), look-ahead videos of 30° were the highest ranked of the nine combinations of height and tilt angle (**Figure 6**).

In the second part of the experimental study, participants were asked to rate each video clip based on i) the level of difficulty associated with detecting randomly placed frisbees and ii) the level of difficulty associated with interpreting randomly placed frisbees (each on a four-point Likert scale with one indicating low difficulty). Look-ahead views associated with a camera tilt angle of 30° were the look-ahead views perceived as creating the least degree of difficulty (**Figures 7** and **8**). Overall, the results of the experimental work completed by [36] suggest that forward-facing cameras on AAMs should be mounted such that they are 30° below horizontal to provide the most useful look-ahead visual information for remote supervision of AAMs.

4.3 Alerting the supervisor of a problem with the autonomous agricultural machine

With reference to an earlier section, the reader is reminded that previous researchers identified the need to warn the supervisor when the AAM experiences an abnormality which it cannot resolve itself. In such situations, there should be a means to communicate the problem to the remote supervisor immediately to increase the operational safety of the system. Different methods have been adopted in non-agricultural devices for similar purposes. They primarily make use of visual, auditory, and tactile (haptic) modalities [39]. Other modalities include olfactory (smell) and gustatory (taste) [39]. Visual, tactile, and auditory modalities have also been adopted in agriculture to inform operators about abnormalities in current agricultural machines. For example, both auditory and visual modalities have been
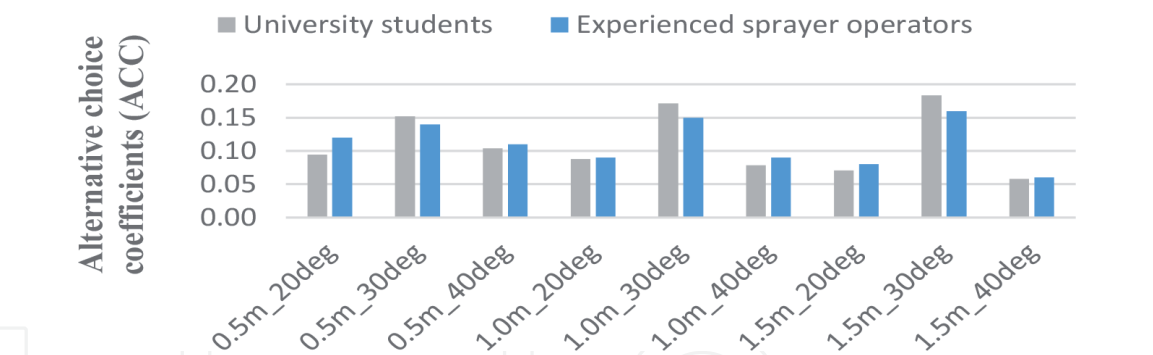
Figure 6.
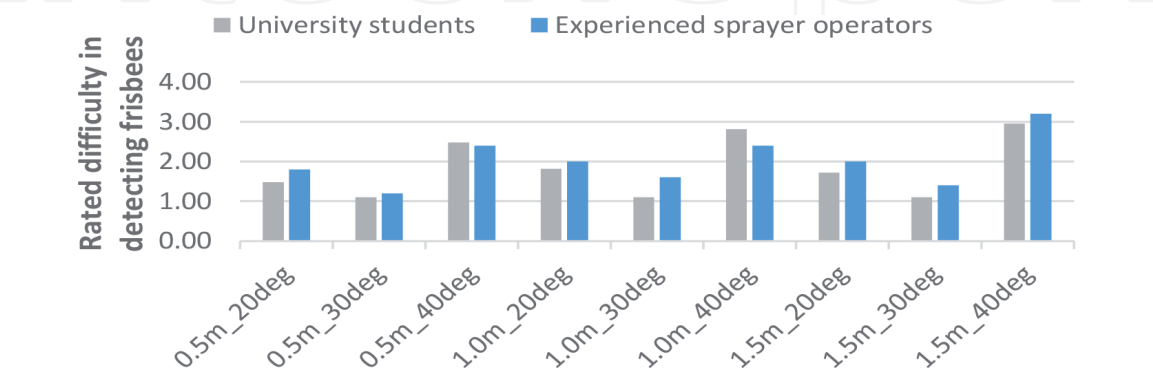*Alternative choice coefficients calculated based on the unranked paired comparisons of look-ahead videos.*

Figure 7.
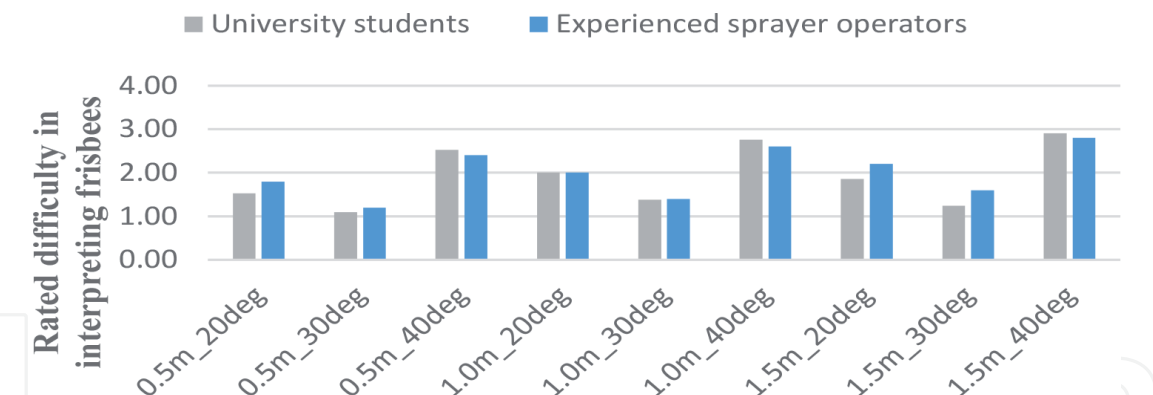*Rated difficulty in detecting frisbees.*

Figure 8.
Rated difficulty in interpreting frisbees.

used to notify operators about plugged nozzles while tactile and visual modality, respectively, have been used to inform operators about lateral deviation of the machine from its desired path.

Visual stimulus can be presented as text, graphics or flashing light [40] while auditory warning can be a continuous or periodic tone or tones (sounds), auditory icon (natural or symbolic), or verbal message [40, 41]. Tactile stimulus, on the other hand, communicates information through the skin (i.e., touch). Each modality has its benefits and shortcomings. For example, auditory modalities are omnidirectional unlike visual modalities that are more effective when the user is stationary. However, it may impede the user's ability to perceive the source of auditory warning. Tactile information is valuable in an environment where noise must be limited [42], but may be less effective if there is minimal contact between the tactile medium and the user's skin. One method that is widely used to assess the

effectiveness of these modalities to communicate information to the user is reaction (response) time [43, 44]. This is the time interval between when the warning is communicated (using one or more modalities) to when the user reacts to the warning. A shorter reaction time would imply that the warning is more effective than a longer reaction time. Reaction times have been reported to vary with age, gender, experience, education level, culture, personality types, and intelligence of the user [45].

Researchers [46–48] have also shown that there are benefits to using multiple modalities in comparison to single ones, especially in situations where the primary task or environmental condition overloads one sensory modality. For example, [49] found that drivers responded faster when presented with multimodal warnings in comparison to unimodal warnings when evaluating driver's response time under different situational urgency while [50] noted that unimodal warnings yielded longer reaction time responses in comparison to multiple modalities while investigating the effectiveness of seven warning methods (visual, auditory, tactile, visual and auditory, visual and tactile, auditory and tactile, and no warning) under three different types of interference (in-vehicle device, audio noise, and vibration of the vehicle). On the other hand, no significant differences were experienced between unimodal and bimodal warnings when informative tactile warning and audio-tactile warnings were compared [51] – suggesting that single warning methods can be as effective as multiple warnings, depending on how they are designed or presented to the supervisor. Hence, as agricultural machinery moves towards full automation, it would be useful to distinguish which of these modalities (single or multiple) would be the most effective in alerting the remote supervisor about a problem with the machine – since these modalities vary in their ability to draw the attention of the supervisor.

A study was conducted to assess which of the seven modalities (visual, auditory, tactile, audio-visual, audio-tactile, visual-tactile, visual–auditory-tactile) would be the most effective in providing feedback to the remote supervisor of an autonomous sprayer [52]. They modified an autonomous agricultural machine control interface (AAMCI) simulator that was designed by [29] to include the different warning methods. Their experiment involved participants playing a game on the secondary screen, monitoring the operation of the autonomous sprayer through the AAMCI simulator, and clicking an 'Alert Perceived' button when they are notified of any error. Response time was used to determine the effectiveness of each modality (single or multiple).

One of their sessions was conducted in a quiet environment without having participants play the game. The remaining sessions were conducted i) in a quiet environment, ii) with tractor background noise and iii) with office (call center) background noise, respectively, to replicate the various scenarios of the four remote supervision concepts that were described by [12]. The experimental setup is shown in **Figure 9**. Further details of the experimental procedure can be found in [52]. They noticed that all seven warning modalities were able to accurately warn the participants of the errors, but varied in their effectiveness (i.e., response time). Overall, the visual and tactile (visual-tactile) warning method was found to be the most effective warning among all the seven warning methods since it had the lowest response time regardless of the background noise or environment (**Figure 10**). However, this observation was only statistically significant for the tractor background noise (p < 0.05).

The response time obtained when participants were continuously monitoring the autonomous sprayer through the AAMCI was also compared with those obtained when participants had to play the game and monitor the simulation in a quiet environment. Their findings revealed that for all warning methods,

Figure 9.
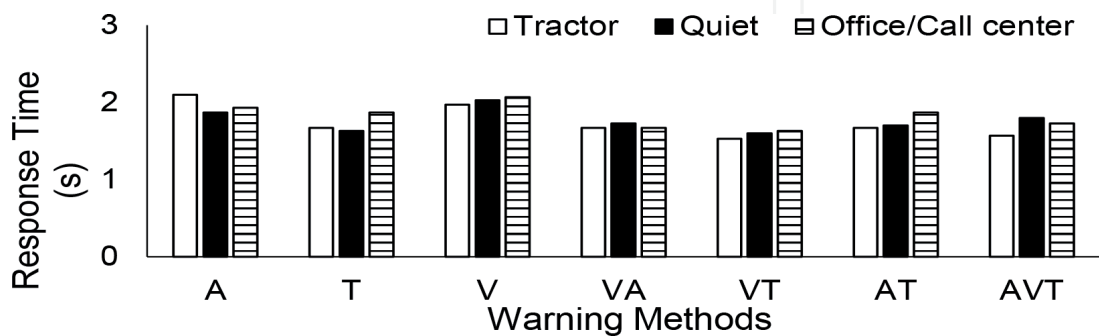*Experimental setup for the alert study.*

Figure 10.
Effectiveness of the warning methods as a function of response time for participants exposed to tractor noise, quiet environment, and office/call center background noise, respectively (a = audio, T = tactile, and V = visual).
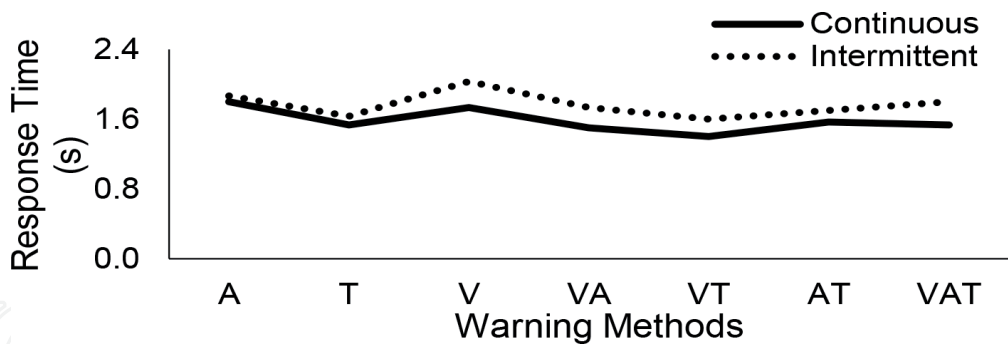

Figure 11.
Comparison of participants' response time in relation to the monitoring levels (a = audio, T = tactile, and V = visual).

participants responded faster when they were monitoring the simulated sprayer (i.e., the interface) continuously in comparison to intermittent monitoring (**Figure 11**). This result was found to be statistically significant (t-test, $\alpha = 0.05$). Despite this result, it was noted that most participants experienced boredom due to low mental workload during the 'No-Game' session (i.e., continuous monitoring) as demonstrated either through yawning, frequent eye blinking, and body posture adjustment.

Overall, the findings from the study may be biased by the fact that the simulation and game may have lacked the type of complexity and workload a remote supervisor may experience while monitoring an actual AAM. Hence, engineers must conduct further analysis during prototype testing to ensure that these results apply in an actual situation of remote supervision of an AAM.

4.4 Latency associated with transmission of real-time visual information

Real-time visual information originates from cameras mounted on AAMs and must be transmitted to the automation interface, perhaps located at the edge of the field, to enable 'edge-of-field' remote supervision. This will require the transference of data through some method of wireless transmission. Conversion of visual data into electronic signals and the time required for data to propagate incurs latency, or delay.

Latency can be described as the difference in time between an action and a response and in the context of autonomous vehicle surveillance can refer to several delay measurements. Glass-to-glass, or capture-to-display, latency is among the most typical to consider for a video being delivered to a user, and refers to the full latency from the occurrence of an event in front of a camera to the time the event can be recognized in the display used to monitor the machine [53]. Delays measured from the beginning of the encoding process to the end of decoding are also critical and are simpler to measure, as they require fewer external tools to evaluate than for glass-to-glass latency. Delays induced by the network, encoding and decoding, camera capture and video display, and the queueing of data packets can all be said to be important elements which comprise transmission latency [54].

The selected method of encoding and decoding is a significant source of latency for video. Networks tend to place restrictions on available bandwidth which must be mitigated to provide consistent video streams for a viewer. Coder-decoder (CODEC) formats, such as the widely used H.264 standard, compress and simplify video streams based on a range of algorithms and protocols. While this compression results in a significant reduction in the size of transmitted information, a trade-off is present where the computations required to reduce the size take varying amounts of time to complete. Compression can tend to be somewhat lossy such as in the case of H.264 [55], sacrificing what is considered an acceptable amount of visual information to produce a reasonably complete image for a viewer. Alternative transmission formats such as MJPEG instead send the video as a steady stream of captured JPEG video frames without such reductions due to compression, which correspondingly reduces image encoding time while resulting in significantly higher bitrates, resulting in demonstrably lower latencies in some comparative experimentation [56]. H.264 has been supplanted somewhat by H.265 video encoding [55], which promises faster rates of encoding and more efficient compression due to larger block sizes for the selection mechanisms used to simplify existing video frames, but still makes up a significant amount of the market today, used by 92% of developers in 2018 [57].

Transmission over the network is a key element of latency. If the required bitrate of video cannot be adequately accommodated by the allotted bandwidth it can result in increasing latency, as successive frames must wait for already queued frames to be received. This latency will theoretically approach infinity or some arbitrary limit, and frames of video will be dropped due to overflow of buffers used in the video stream to hold incoming video frames. The network path taken by video will inevitably introduce further delay, with longer paths with more frequent hops resulting in an increasing latency due to the travel time for data packets. Selection of the transmission level protocol used for the two devices to communicate will also have significant implications for overall latency. While the widely used TCP/IP communication protocol will eventually successfully transmit frames of video, the required two-way acknowledgement of data reception places significant time constraints. As frames are lost due to travel through the network this latency induced due to TCP can steadily increase as successive frames are forced to wait for complete transmission of earlier data [58]. In comparison, a transmission protocol

such as UDP makes no guarantees of successful reception due to lacking this same handshake mechanism, but delivers video frames with lower latency due to the reductions in time as a result of not requiring the handshaking process with each transmitted data packet.

Latency has been measured using a range of methods for various applications, which vary depending on the specific delay that is to be measured. Kaknjo [56] measured latency during transmission for a robot utilizing the common method of placing a pulse per second enabled LED in front of a camera to act as an event recognizable by the system. This direct test of the camera system was coupled with utilization of timestamps to measure latency when transmitting video over a larger network. A customized application for WebRTC communication [59] transmits specialized video frames containing a spinning object and a continuously counting timestamp to measure latency between users.

An investigation of transmission latency in an agricultural setting was undertaken at the University of Manitoba [60]. A Raspberry Pi 4 was configured using the GStreamer multimedia application to be able to selectively stream video over cellular internet and a direct radio connection. The Pi 4 was fitted with a cellular header and connected via Ethernet to the radio system and mounted to a riding mower. Open-source GStreamer libraries were then used to overlay timestamps into the video feed from the Raspberry Pi, which could then be decoded and extracted by a laptop acting as a receiver and compared against the laptop time to measure the latency experienced.

In a subsequent experimental study, video transmission latency was measured for three transmission distances (200, 400 & 600 m) and for three resolutions of video (480p20fps@400kbit, 480p25fps@500kbit & 576p20fps@600kbit) using two transmission modes (cellular and radio). Data were collected at four geographical locations within 1 h (driving time) from the university campus. Complete details of the experimental procedure and results can be found in [60].

For the relatively short transmission distances tested (which were selected as representative of the 'edge-of-field' remote supervision concept proposed by [12] there were no obvious differences or trends in transmission latency for either transmission mode (cellular or radio) with a couple of exceptions which can likely be explained by the presence of trees adjacent to one of the test sites which may have interfered with radio transmission. **Figure 12** shows the results of data collection at the Glenlea, Sanford, and university campus testing locations.

As was expected, transmission latency increased with increasing video resolution for both transmission modes (cellular and radio), but with the transmission times below 300 ms in most cases. Despite the more direct transmission path for
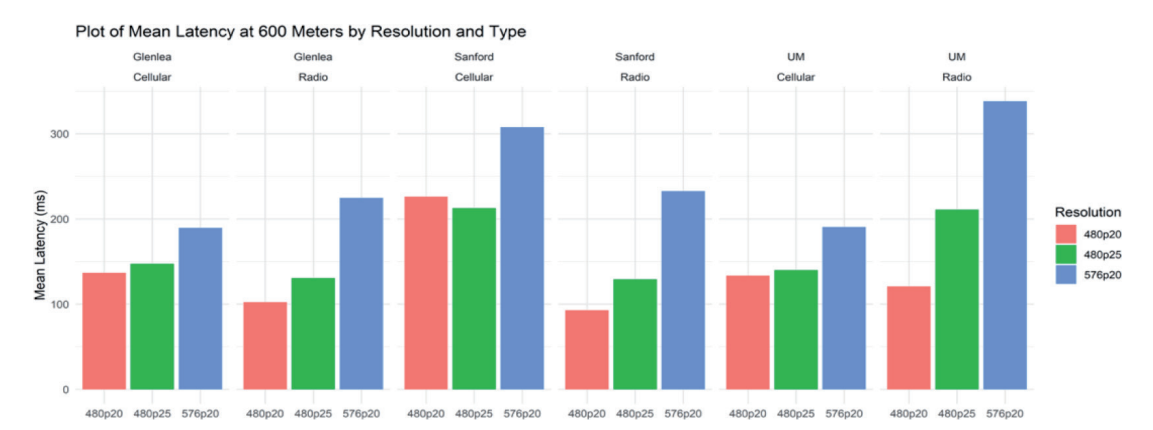

Figure 12.
Graphical depiction of experimental mean latency measurement results.

radio transmission, measured latencies were less for cellular transmission at test sites with strong signal strength. It was observed, however, that a couple of the test sites had poor cellular coverage. At one site, cellular transmission of video was not feasible with transmission latencies up to 86 s observed.

Overall, [60] concluded that it should be feasible to transmit real-time video from an AAM to an automation interface located at the edge of the field using either cellular or radio transmission. Latencies measured fell within acceptable international telecommunications union recommendations for acceptable one-way delay of less than 400 ms. These values were also in line with experiments for a telerobotic surgery simulator [61], where below 300 ms it was observed that performance of surgical tasks did not tend to degrade much with increasing latency. In locations where adequate cellular signal strength exists, cellular transmission is recommended as it causes less transmission latency and would give a greater overall range. Radio transmission of real-time video is recommended only in locations where there is poor cellular coverage.

Despite the promising results reported by [60] related to latency of real-time video transmission, there are several questions that warrant further investigation. First, research is warranted to determine the impact of transmission latency on the usability of the automation interface. Assuming constant latency, is there a magnitude beyond which it becomes impossible to remotely supervise an AAM? A related question is to determine the effect of varying latency on the usability of the automation interface. A second question worthy of further investigation is to determine the quality of video that is required for remote supervision of an AAM. [60] have reported that transmission latency increases with increasing video resolution, suggesting that it is beneficial to use low-resolution video for this application. The effect of video resolution on the usability of the automation interface must be determined. It is anticipated that the optimum video resolution for real-time supervision of AAMs will be a compromise between transmission latency and usability. A third issue is that the techniques and equipment used for transmitting video data have not yet been optimized. With the implementation of elements such as dedicated specific hardware, adaptive bitrate encoding and H.265 compression, it would likely be possible to further reduce latency by reducing the time required for encoding and decoding. Similar studies have been able to obtain latencies under 200 ms with Raspberry Pis [62] in different environments. It is important for developers and product manufacturers to consider these various aspects of video transmission to be able to provide low latency video feeds for end users. CODEC mechanisms should be selected to balance the requirements of bandwidth and latency and appropriate transmission protocols utilized to keep video streams loss tolerant while keeping latency low. Implementing the appropriate mechanisms in video streams will minimize delay for the AAM supervisor.

5. Case study: automation interface for an autonomous plot sprayer

This section presents a case study where knowledge gained from prior research activity related to the role of real-time visual information to the task of remote supervision has been applied to the design of an automation interface for an autonomous plot sprayer. The desire to design an automation interface for this specific machine was initiated by a group of undergraduate students interested in developing an AAM for the agBOT Challenge sponsored by Purdue University. To meet the objectives of the agBOT Challenge, the students would need to design and build an autonomous machine that was designed to autonomously navigate through a cornfield, detect and distinguish between weeds and corn plants, and

Figure 13.
CanAM ATV and setter plot sprayer used by a group of undergraduate University of Manitoba students to produce an autonomous sprayer for the agBOT challenge in 2019.

automatically spray the corn plants with fertilizer while spraying the weeds with herbicide (both in real-time). The students modified a CanAm ATV to navigate autonomously (**Figure 13**). Weed detection was achieved using a ground-facing camera mounted on the front of the ATV feeding data to image processing applications. Modifications were made to a Setter plot sprayer that would allow individual nozzles to be activated to apply herbicide to be applied when weeds were detected. The students desired an automation interface that could be used to remotely supervise the autonomous sprayer during the agBOT Challenge.

The autonomous agricultural sprayer is outfitted with four cameras (one of which is used for the weed and corn detection task) and a variety of sensors for the navigation and spraying tasks. The sensor input is processed by multiple onboard computers. One of these devices is dedicated to processing the visual input for the plant detection task, while the others process the remaining sensors and control actuators that allow the machine to move and spray the plants. The on-board computers communicate with each other through a middleware known as the Robot Operating System (ROS). The computers are also connected to a web server, through the internet, where they dump sensor data in real-time while the machine is running.

The automation interface for the autonomous agricultural sprayer was designed to display both sensor data and live video for a supervisor at a remote location. To enable remote supervision of the machine from anywhere in the world with internet connectivity, the interface is connected to the machine through a web server. The automation interface is shown in **Figure 14**.

From top to bottom, the interface is divided into three prominent sections: the toolbar, the video feeds, and the indicators (icons and graphical elements). The toolbar includes the start button and the emergency stop button. Since all other elements were designed for monitoring purposes, the start and stop buttons serve as the primary controls that the remote supervisor has over the machine. In the current iteration of the interface, the start button initiates the machine's autonomous operations, while the stop button terminates autonomous operations. The notification bar keeps the user informed about the status of the machine, its sensors, and its environment. The text-based notifications are enhanced by a color-coded status indicator – green, yellow, and red – to indicate the corresponding severity.
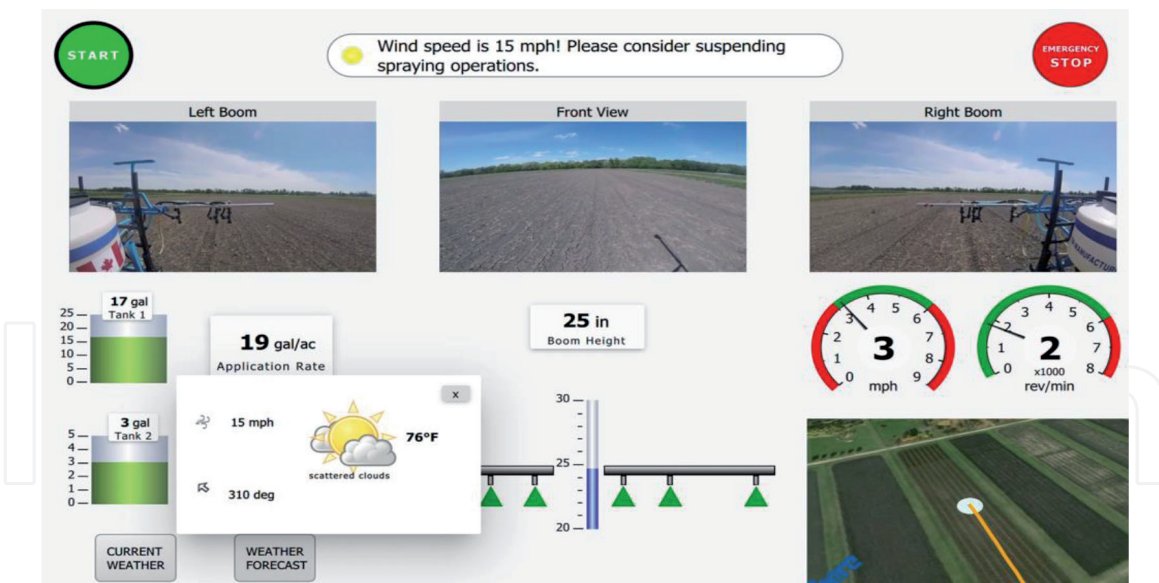
Figure 14.
The automation interface designed for an autonomous plot-sized agricultural sprayer.

Three video feeds were included in the automation interface following the recommendations made by [63]. The videos provide visual and auditory information, although at a lower fidelity than experienced when inside a tractor. The middle video provides a view ahead of the autonomous sprayer to show what is coming. The videos on the left and right sides are from rearward-facing cameras that show the left and right booms of the plot sprayer. These videos allow the remote supervisor to monitor both the machine and the spraying operation and take quick action in the case of an emergency.

Below the video feeds are the icons and graphical elements that display information regarding the state of the vehicle, sprayer, and the environment. The indicators were organized according to two main goals: i) monitoring the machine and ii) monitoring the sprayer, with the most important information placed towards the center of the display. Towards the far right of the interface is a group of indicators for monitoring the vehicle, including the vehicle speed, the engine speed, and a coverage map. In addition to providing up to level 3 situation awareness, the design of the speed indicators follows common design patterns for such indicators in most vehicles and is expected to fit the mental model of most users. The coverage map provides global situation awareness of the spraying operation.

The rest of the indicators, including the tank level, application rate and boom height, are related to the sprayer. The tank level indicates the amount of liquid that is currently in each tank, while the application rate provides information about how much liquid is being sprayed per area of the field from all the nozzles connected to the tank. While the application rate indicator supports only level 1 situation awareness, the tank level indicator was designed to support up to level 3 situation awareness by utilizing a digital display and color-coded value bar. Finally, towards the center of the display is an indicator, which was designed to provide an intuitive understanding of the state of the sprayer boom and the 6 nozzles attached to it. This indicator moves up and down in a similar fashion to the movement of the boom to indicate the height of the boom above the ground. The state of the nozzle is indicated by a green triangle (for an active nozzle) and a red square (for a blocked nozzle). Weather information is also included, which in addition to the information provided by other indicators, allows the remote supervisor to make judgments about the quality of the spraying operation, and project this judgment into the future (e.g., through available weather forecast) to take timely action.

6. Next steps in the design of an automation interface

6.1 Utilization of real-time auditory information

Auditory information can be extremely useful to a human operator, even one with minimal experience, and can provide information about changes in parameters being independently monitored via sensors (rpm, load, etc.) [64]. There is opportunity to consider what role auditory information might play in the task of remotely supervising an AAM through an automation interface.

Based on anecdotal information, it was recognized that machinery operators are often able to detect existing or impending problems from the changes in sound produced by the mechanical components of the machine. Karimi [65] reported that the addition of auditory cues did not improve steering performance (in a simulated agricultural vehicle) perhaps because steering is a purely visual task, however, auditory cues did improve the monitoring task. Donmez [64] investigated the use of sonifications (continuous auditory alerts) during the control of unmanned aerial vehicles and found that visual information supported by sonifications yielded faster reaction times than visual information supported by discrete auditory signals. Though an autonomous controller may not use or interpret sound in the same way as a human does, it is important to evaluate its potential use in control applications given its value in monitoring non-autonomous machine operation.

Though auditory information alone may not provide sufficient information for automated control, it can provide qualitative information on changing parameters, or be an indicator of a change in state. This information can be used directly to trigger certain responses, or can be used in a training set to become a single indicator of a specific state, replacing several other parameters that may have to be combined to glean the same information. Capturing high quality auditory information is generally simple and inexpensive with modern technology and can be captured from multiple locations within a machine or system, making it a good option for a variety of applications.

Classification of sounds with machine learning is already prevalent in music. There are a number of applications available to consumers to classify songs to both organize music and provide recommendations based on previous listening history. These applications use various classification algorithms (Fourier Transform, Mel Frequency Cepstrum Coefficients, etc.) to provide this service. The existence of these classification services implies that machine sounds could be classified in a similar manner to determine the state of operation, unexpected variations in parameters (i.e., malfunctions), and more. In the case where humans are controlling a machine with assistance from automation, sonifications have been shown to be very effective at helping the human operator predict the future state of the machine, and therefore react accordingly [64]. Thus, automation via audio feedback has the potential to not only improve human-machine interaction in semi-autonomous applications, but also to provide input to prompt automated responses in fully autonomous applications.

Classification is a pattern recognition problem. If a classifier can be built to recognize specific characteristics of an input signal that identify, within a certain level of confidence, what grouping or 'state' that input belongs to [66], it can then be classified, and this information can be used to produce an appropriate response. These specific characteristics are referred to as 'features' and can be comprised of any distinctive measurement or structural component of the signal that can be extracted. Multiple features may be needed for classification, but analysis can be performed to determine which features in which combinations produce the quickest classification algorithm with the highest level of confidence.

Two features that have been explored in experimentation are the spectral centroid and formant (dominant) frequency. The spectral centroid is used to detect the 'center of mass' of the spectrum (distribution of values) representing the frequency [67]. Sub band spectral centroids have been used successfully in speech recognition applications [68] and so are a good starting point for machinery audio classification. The formant frequency of a signal represents the concentration of acoustic energy (peak), and has also been used successfully in speech recognition, as well as biomedical signal analysis and musical instrumentation analysis [69].

A classification experiment was performed using video collected from the rear of an S680 John Deere combine harvester near the straw chopper [70]. The video was recorded using a GoPro Hero Session during harvest of canola in a Manitoba field during the 2017 harvest season. From this video, audio clips were extracted corresponding to the operational sounds of the machine. The audio was sampled at a rate of 48 kHz with AAC compression and automatic gain control and converted to .wav file format for analysis. Sound samples underwent a Fourier transform, and then eight features were extracted from each segment for analysis, all based around frequency characteristics. Features in each segment were analyzed to both build and then test a feedforward, pattern recognition neural network. Samples were divided into those used for training (70%), validation (15%), and testing (15%), and three operating modes were selected for classification: 1) Engine running with no threshing, 2) Engine running and threshing engaged, and 3) Engine running, and threshing engaged at 80% capacity.

By varying parameters such as segment size, accuracies of 88–100% were obtained with larger segment sizes (over 2048 segments) producing a consistent classification accuracy of 99%. This sample size allowed for a total of 1970 samples which is sufficient to declare a high degree of confidence in the result. The results of this experiment show that a relatively basic model with audio as a sole input can successfully be used to classify machinery operating modes in real-time. These results are promising enough to justify further study to better understand how to optimally apply this technique in a practical application.

The current study focuses on identifying three broad classes of operation based on a single audio input. However, it is possible that there are a number of operational modes, or even specific events, scenarios, or changes in conditions that can be classified through auditory input. Further research is required to understand what other audio inputs (location and type of sound recording), or combination of audio inputs can be used for classification of a broader range of machinery parameters. It is also critical to understand how various conditions (wind, crop type, machine parameters, etc.) impact classification and what types of calibrations or modifications may be necessary to account for variation in operating conditions. The previous study focused on a single crop type with all recordings taken under identical operating conditions. A robust field prototype would need to account for changing environmental conditions in order to be reliable. It would also be beneficial to investigate other methods of recording and processing sound along with how this input is used to build the classifier.

The current study used manual inspection for feature extraction from raw audio, but it is likely that efficiencies could be gained through automated feature extraction or some level of sound processing to enhance various aspects of the audio that may be more useful for classification. There are many ways to build and train a classification system, and it is likely that a practical system could be optimized with further investigation. One follow-up study that was conducted with the same data explored the use of a 7-layer convolutional neural network (CNN) as a classifier [71]. Through this method, greater accuracies could be achieved when lower numbers of audio samples were used compared with a conventional neural

network analysis. Using 5000 samples of audio segments resulted in an accuracy of 95% compared with 78% accuracy achieved with a neural network with the same samples.

It is likely that further investigation would provide insight into optimal audio sampling techniques and feature extraction and analysis to classify a greater variety of operating modes under more variable conditions.

7. Conclusions

The purpose of this chapter is to provide an overview of recent research that has been conducted to understand how to design an effective automation interface for the task of remotely supervising an autonomous agricultural machine (AAM). First of all, it has been assumed that the existence of an automation interface is essential because the owner of the AAM will always want to have some means of monitoring the status of the machine in the field and, in some instances, human input may be required to diagnose problems and/or to make management decisions. Secondly, it has been assumed that the automation interface needs to include real-time visual information showing the AAM within the field environment to complement telemetric data that is displayed using conventional means. Experimental data has supported the important role that is played by real-time visual information, and has provided insight on related issues such as i) where the cameras should be pointed to provide information that supports the supervisory task, ii) how the cameras should be positioned to yield useful look-ahead information, and iii) how to alert the supervisor of system problems, and iv) the latency associated with wireless transmission of live video. Early research results suggest that it may also be possible to use auditory information to provide additional information to the supervisor through the automation interface.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

Author details

Daniel D. Mann*, Uduak Edet, Mitchell Green, Olayinka Folorunsho,
Avery Simundsson and Franklin Ogidi
Department of Biosystems Engineering, University of Manitoba, Winnipeg,
MB R3T 5V6, Canada

*Address all correspondence to: Danny.Mann@umanitoba.ca

IntechOpen

References

[1] Scholtz J. Theory and evaluation of human robot interactions. In: Proceedings of the 36th Annual Hawaii International Conference IEEE; 2003.

[2] Berenstein R, Halevi IB, Edan Y. A remote interface for a human-robot cooperative vineyard sprayer. In: 11th International Conference on Precision Agriculture, Indianapolis, USA; 2012. p. 15-18

[3] Bechar A, Vigneault C. Agricultural robots for field operations: concepts and components. Biosystems Engineering. 2016;149:94-111.

[4] Adamides G, Katsanos C, Parmet Y, Christou G, Xenos M, Hadzilacos T, Edan Y. HRI usability evaluation of interaction modes for a teleoperated agricultural robotic sprayer. Applied Ergonomics. 2017;62:237-246.

[5] Taylor RM. Human automation integration for supervisory control of UAVs. In: Virtual Media for Military Applications Meeting Proceedings RTO-MP-HFM, Paper 12; Neuilly-sur-Seine, France. 2006. p. 12.1-12.10

[6] Cubber GD, Doroftei D, Rudin K, Berns K, Matos A, Serrano D, Sanchez J, Govindaraj S, Bedkowski J, Roda R, Silva E, Ourevitch S. Introduction to the use of robotic tools for search and rescue. In: Search and Rescue Robotics; IntechOpen: London, UK, 2017; p. 1—17. DOI: 10.5772/intechopen.69489

[7] Kulakov FM. Methods of supervisory remote control over space robots. Journal of Computer and Systems Sciences International. 2018;57(5):822-839.

[8] Edet U, Hawley E, Mann DD. Remote supervision of autonomous agricultural sprayers: the farmer's perspective. Canadian Biosystems Engineering. 2018;60:2.19-2.31.

[9] Wong C, Yang E, Yan XT, Gu D. An overview of robotics and autonomous systems for harsh environments. In: 23rd International Conference on Automation and Computing (ICAC); September 2017; IEEE; 2017. p.1-6

[10] Cummings ML. Human supervisory control of swarming networks. In: 2nd annual swarming: autonomous intelligent networked systems conference; June 2004; p.1-9

[11] Duan Y, Ma L, Zhang C, Liu G, Ma X, Li L. A remote monitoring system for physiology and environmental information in pig farming. In: 2014 American Society of Agricultural and Biological Engineers Annual International Conference; 13-16 July 2014; Montreal, Quebec; 2014.

[12] Edet U, Mann DD. Remote supervision of autonomous agricultural machines: concepts and feasibility. Submitted (November 2020) to Canadian Biosystems Engineering.

[13] Gonzalez-de-Santos P, Ribeiro A, Fernandez-Quintanilla C, Lopez-Granados F, Branstoetter M, Tomic S, Pedrazzi S, Peruzzi A, Pajares Kaplanis G, Perez-Ruiz M, Valero C, del Cerro J, Vieri M, Rabatel G, Debilde B. Fleets of robots for environmentally safe pest control in agriculture. Precision Agriculture. 2017;18:574-614.

[14] Zhang X, Geimer M, Noack PO, Grandl L. Development of an intelligent master-slave system between agricultural vehicles. In: 2010 IEEE Intelligent Vehicles Symposium; June 2010; IEEE; 2010. p.250-255

[15] Murakami N, Ito A, Will JD, Steffen M, Inoue K, Kita K, Miyaura S. Development of a teleoperation system for agricultural vehicles. Computers and Electronics in Agriculture. 2008;63(1):81-88.

[16] Stentz A, Dima C, Wellington C, Herman H, Stager D. A system for semi-autonomous tractor operations. Autonomous Robots. 2002;13(1): 87-104.

[17] Noguchi N, Will J, Reid J, Zhang Q. Development of a master–slave robot system for farm operations. Computers and Electronics in Agriculture. 2004;44(1):1-19.

[18] Emmi L, Gonzalez-de-Soto M, Pajares G, Gonzalez-de-Santos P. New trends in robotics for agriculture: integration and assessment of a real fleet of robots. The Scientific World Journal. 2014;Volume 2014, Article ID 404059:1-21. DOI: 10.1155/2014/404059

[19] Edet U. Design of an interface for remotely supervised autonomous agricultural sprayers [thesis]. Winnipeg: University of Manitoba; 2021.

[20] Endsley MR. From here to autonomy: lessons learned from human-automation research. Human Factors. 2017;59(1):5-27.

[21] Grubb PL, Miller LC, Nelson WT, Warm JS, Dember WN. Cognitive failure and perceived workload in vigilance performance. In: Mouloua M, Parasuraman R, editors. Human performance in automated systems: Current research and trends. Hillsdale, NJ: Lawrence Erlbaum; 1994. p.115-121

[22] Blackmore BS, Have H, Fountas S. A proposed system architecture to enable behavioural control of an autonomous tractor. In: Zhang Q, editor. Automation Technology for Off-Road Equipment. St. Joseph: ASABE;2002. p.13-23

[23] Johnson DA, Naffin DJ, Puhalla JS, Sanchez J, Wellington CK. Development and implementation of a team of robotic tractors for autonomous peat moss harvesting. Journal of Field Robotics. 2009;26(6-7):549-571.

[24] Moorehead SJ, Wellington CK, Gilmore BJ, Vallespi C. Automating orchards: a system of autonomous tractors for orchard maintenance. In: IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS) Workshop on Agricultural Robotics; 2012.

[25] Dorais G, Bonasso RP, Kortenkamp D, Pell B, Schreckenghost D, Adjustable autonomy for human-centered autonomous systems. In: Working notes of the Sixteenth International Joint Conference on Artificial Intelligence Workshop on Adjustable Autonomy Systems. 1999. p.16-35

[26] Endsley MR, Bolté B, Jones DG. Designing for situation awareness. An approach to user-centered design. New York: Taylor & Francis; 2003.

[27] Endsley MR. 1988. Design and evaluation for situation awareness enhancement. In: Proceedings of the Human Factors Society 32nd Annual Meeting. Santa Monica, CA: Human Factors Society; 1988. p.97-101

[28] Adamides G, Berenstein R, Ben-Halevi I, Hadzilacos T, Edan Y. User interface design principles for robotics in agriculture: the case of telerobotic navigation and target selection for spraying. In: Proceedings of the 8th Asian Conference for Information Technology in Agriculture (volume 36); September 2012; Taipei, Taiwan; 2012.

[29] Panfilov I, Mann DD. The importance of real-time visual information for the remote supervision of an autonomous agricultural machine. Canadian Biosystems Engineering. 2018;60:2.11-2.18. DOI: 10.7451/CBE.2018.60.2.11

[30] Lewis M. Human Interaction with multiple remote robots. Reviews of Human Factors and Ergonomics. 2013;9(1):131-174.

[31] Ishibashi M, Iida M, Suguri M, Masuda R. Remote monitoring of agricultural robot using web application. IFAC Proceedings. 2013;46(18):138-142.

[32] Sanchez J, Duncan JR. Operator-automation interaction in agricultural vehicles. Ergonomics in Design. 2009;17(1):14-19.

[33] Dey AK, Mann DD. A complete task analysis to measure the workload associated with operating an agricultural sprayer equipped with a navigation device. Applied Ergonomics. 2010;41(1):146-149.

[34] Edet U, Mann DD. Visual information requirements for remotely supervised autonomous agricultural machines. Applied Sciences. 2020;10,2794. DOI: 10.3390/app10082794

[35] Tang P, Mann DD. Factors contributing to guidance performance when using a camera-based guidance aid. Journal of Agricultural Safety and Health. 2003;9(1):47-60.

[36] Folorunsho O. The influence of image velocity on the situation awareness of the remote supervisor of an autonomous agricultural machine [thesis]. Winnipeg, University of Manitoba; 2021.

[37] Niec JH, Margheim W, Umble A. Applying an innovative equipment selection process using the unranked-paired comparison technique. Proceedings of the Water Environment Federation. 2010;15:2021-2032. DOI: 10.2175/193864711802713108

[38] Mulder P. Paired comparison method [Internet]. 2018. Available from: toolshero:https://www.toolshero.com/decision-making/paired-comparison-method/ [Accessed: 2021-02-04]

[39] Laughery KR, Wogalter MS. 2006. Designing effective warnings. Reviews of Human Factors and Ergonomics. 2006;2(1):241-271.

[40] Elbert KK, Kroemer HB, Hoffman ADK. Selection, design, and arrangement of controls and displays. In: Ergonomics: how to design for ease and efficiency. Academic Press; 2018. p.500- 557

[41] Stevens CJ, Brennan D, Petocz A, Howell C. Designing informative warning signals: Effects of indicator type, modality, and task demand on recognition speed and accuracy. Advances in Cognitive Psychology. 2009;5:84-90.

[42] Smith CAP, Clegg BA, Heggestad ED, Hopp-Levine PJ. Interruption management: A comparison of auditory and tactile cues for both alerting and orienting. International Journal of Human-Computer Studies. 2009;67(9):777-786. DOI: 10.1016/j.ijhcs.2009.05.002

[43] Wogalter MS, Conzola VC, Smith-Jackson TL. Research based guidelines for warning design and evaluation. Applied Ergonomics. 2002;33(3):219-230.

[44] Whelan R. Effective analysis of reaction time data. The Psychological Record. 2008;58(3):475-482.

[45] Jain A, Bansal R, Kumar A, Singh KD. A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. International Journal of Applied & Basic Medical Research. 2015;5(2):124-127.

[46] Burke JL, Prewett MS, Gray AA, Yang L, Stilson FR, Coovert MD, Elliot LR, Redden E. Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. In: Proceedings of the 8th

International Conference on Multimodal interfaces; November 2006; 2006; p.108-117

[47] Ho C, Reed N, Spence C. Multisensory in-car warning signals for collision avoidance. Human Factors. 2007;49(6):1107-1114.

[48] Baldwin CL, Spence C, Bliss JP, Brill JC, Wogalter MS, Mayhorn CB, Ferris TK. Multimodal cueing: the relative benefits of the auditory, visual, and tactile channels in complex environments. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting; September 2012; Los Angeles, CA: SAGE Publications; 2012. p. 1431-1435

[49] Politis I, Brewster SA, Pollick F. Evaluating multimodal driver displays under varying situational urgency. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 2014. p. 4067-4076

[50] Murata A, Kanbayashi M, Hayami T. Effectiveness of automotive warning system presented with multiple sensory modalities. In: International Conference on Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management; July 2013; Berlin: Springer; 2013. p. 88-97

[51] Ahtamad M, Robert G, Cristy H, Nick R, Charles S. Informative collision warnings: effect of modality and driver age. In: Proceedings of the Eighth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design 22-25 June 2015; Salt Lake City, Utah. Iowa City, IA: Public Policy Center, University of Iowa; 2015. p. 324-330.

[52] Edet U, Mann DD. Evaluating warning modalities for remotely supervised automated agricultural machines. Submitted (November 2020) to Journal of Agricultural Safety & Health.

[53] Bachhuber C, Steinbach E, Freundl M, Reisslein M. On the minimization of glass-to-glass and glass-to-algorithm delay in video communication. IEEE Transactions on Multimedia. 2018;20(1),238-252. DOI: 10.1109/TMM.2017.2726189

[54] Li M, Junior FEF, Sheng W, Bai H, Fan Z, Liu M. Measurement of latency on visual feedback in an immersive telepresence robotic system. In: Proceedings of the World Congress on Intelligent Control and Automation (WCICA); July 2018; Changsha, China. IEEE; 2018. p. 1757-1762. DOI: 10.1109/ WCICA.2018.8630495

[55] Anonymous. H.264: Advanced video coding for generic audiovisual services [Report]. International Telecommunication Union; 2003. Available from https://www.itu.int/ rec/T-REC-H.264-200305-S/en

[56] Kaknjo A, Rao M, Omerdic E, Robinson L, Toal D, Newe T. Real-time video latency measurement between a robot and its remote control station: causes and mitigation. Wireless Communications and Mobile Computing. 2018;Article ID 8638019: 1-19. DOI: 10.1155/2018/8638019

[57] Anonymous. Welcome to the 2018 video developer report [Internet]. 2018 Available from https:// go.bitmovin.com/hubfs/Bitmovin-Video-Developer-Report-2018.pdf [Accessed: 2021-02-18]

[58] Kumar S, Rai S. Survey on transport layer protocols: TCP & UDP. International Journal of Computer Applications. 2012;46(7): 975-8887.

[59] García B, Gortázar F, López-Fernández L, Gallego M. Analysis of video quality and end-to-end latency in WebRTC. 2016 IEEE Globecom Workshops, GC Wkshps 2016 - Proceedings, 2-7. 2016. DOI: 10.1109/ GLOCOMW.2016.7848838

[60] Green MK, Mann DD, Hossain E. Measurement of latency during real-time video transmission for remote supervision of agricultural machines. Submitted (February 2021) to Computers and Electronics in Agriculture.

[61] Perez M, Xu S, Chauhan S, Tanaka A, Simpson K, Abdul-Muhsin H, Smith R. Impact of delay on telesurgical performance: study on the robotic simulator dV-Trainer. International Journal of Computer Assisted Radiology and Surgery. 2016;11(4):581-587. DOI: 10.1007/s11548-015-1306-

[62] Jennehag U, Forsstrom S, Fiordigigli FV. Low delay video streaming on the internet of things using raspberry Pi. Electronics. 2016;5(3). DOI: 10.3390/electronics5030060

[63] Edet U, Mann DD. Visual information requirements for remotely supervised autonomous agricultural machines [Unpublished project report]. Winnipeg, University of Manitoba; 2018.

[64] Donmez B, Cummings ML, Graham HD. Auditory decision aiding in supervisory control of multiple unmanned aerial vehicles. Human Factors. 2009;51(5):718-729.

[65] Karimi D, Mondor TA, Mann DD. Application of auditory signals to the operation of an agricultural vehicle: results of pilot testing. Journal of Agricultural Safety and Health. 2008;14(1):71-78.

[66] Mahana P, Singh G. Comparative analysis of machine learning algorithms for audio signals classification. International Journal of Computer Science and Network Security. 2015;15(6):49-55.

[67] Peeters G. A large set of audio features for sound description (similarity and classification) in the CUIDADO project. Paris: Ircam, Analysis/Synthesis Team. 2004.

[68] Gajic B, Paliwal KK. Robust speech recognition in noisy environments based on subband spectral centroid histograms. IEEE Transactions on Audio, Speech and Language Processing. 2006;14(2):600-608. DOI: 10.1109/TSA.2005.855834

[69] Nisar S, Khan OU, Tariq M. An efficient adaptive window size selection method for improving spectrogram visualization. Computational Intelligence & Neuroscience. 2016; Article ID 6172453:1-13. DOI: 10.1155/2016/6172453

[70] Simundsson A, Mann DD, Thomas G. A neural network to classify auditory signals for use in autonomous harvester control systems. CSBE/SCGAB 2019 Annual Conference; 14-17 July 2019; Vancouver. Winnipeg: CSBE; 2019. Paper No. CSBE19-163.

[71] Thomas G, Simundsson A, Mann DD, Balocco S. Classification of auditory signals from a combine harvester based on mel-frequency cepstral coefficients and machine learning. Submitted (October 2020) to Canadian Biosystems Engineering.