The Importance of Real-Time Visual Information for the Remote Supervision of an Autonomous Agricultural Machine

Ivan Panfilov¹ and Danny D. Mann¹

¹Department of Biosystems Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada
Corresponding author: Danny Mann (Danny.Mann@umanitoba.ca)

Submitted: 2018 February 22 Revised: 2018 April 17 Accepted: 2018 May 1 Published online: 2018 May 16

ABSTRACT
“Supervised autonomy” is a term that best describes the autonomous agricultural machines currently being envisioned. Such machines will be able to work autonomously, but require human supervision. The topic of interface design, however, has not received sufficient attention by designers. The goal of this study was to investigate the importance of live video for remote supervision of autonomous agricultural machines. The study was conducted using an autonomous agricultural machine control interface simulator, which provided information of machine status using graphical indicators (which interpreted and displayed sensor information) and live video (which was presented as seen by the camera). The participants of the study performed the role of the supervisor of an autonomous agricultural machine. The importance of live video was assessed by comparing the participant’s performance during trials with and without video. Information about the gaze direction was obtained with an eye-tracking system. The results showed that graphical indicators are the preferred source of machine status information, and live video is a secondary source. At the beginning of the experiment, participants split their attention evenly between the graphical indicators and the live video, but by the end of the experiment, their focus was on the graphical indicators 70% of the time. More than 75% of the participants indicated that the live video helped them to understand machine functions better and they felt more secure when the video footage was present. The participants suggested that live video should be available either full time or on demand. Control interface designers should consider including live video on the interface for autonomous agricultural machines to provide additional decision-making support to the supervisor.

KEYWORDS
Autonomous agricultural machines, visual information, automation interface, remote supervision.

CITATION

RÉSUMÉ
L’« autonomie supervisée » est le terme qui décrit le mieux les machines agricoles autonomes qui sont présentement envisagées. De telles machines seront en mesure de fonctionner de façon autonome, mais elles nécessiteront une supervision humaine. Toutefois les designers n’ont pas accordé suffisamment d’attention aux questions liées à la conception des interfaces de contrôle. Le but de cette étude était d’investiguer l’importance d’un lien vidéo direct pour la supervision à distance de machines agricoles autonomes. L’étude a été réalisée à l’aide d’un simulateur d’interface de contrôle pour machine agricole autonome qui, par ses indicateurs graphiques, fournissait des renseignements sur l’état de la machine (en présentant et interprétenant l’information recueillie par des capteurs) ainsi qu’une vidéo en direct (montrant les images tels que recueillies par une caméra). Les participants à l’étude ont joué le rôle de superviseur d’une machine agricole autonome. L’importance de la vidéo en direct a été déterminée en comparant les performances des participants durant les essais avec ou sans la vidéo. L’information sur la direction du regard des participants a été recueillie par un système de suivi oculaire. Les résultats ont démontré que les indicateurs graphiques constituaient la source préférée d’information sur l’état de la machine, tandis que la vidéo en direct venait au second rang. Au début des essais, les participants répartissaient également leur attention entre les indicateurs graphiques et le signal vidéo. Toutefois, vers la fin des essais, l’attention des participants était centrée sur les indicateurs graphiques durant 70 % du temps. Plus de 75 % des participants ont indiqué que la vidéo en direct les aidait à mieux comprendre le fonctionnement de la machine et qu’ils se sentaient plus à l’aise lorsqu’ils pouvaient voir la vidéo en direct. Les participants ont ainsi suggéré que la vidéo en direct devrait être disponible en tout temps ou sur demande. Les designers d’interfaces de contrôle pour l’opération de machine agricole autonome devraient inclure la vidéo en direct pour donner au superviseur un soutien additionnel à la prise de décisions.

MOTS CLÉS
Machines agricoles autonomes, information visuelle, interface d’automation, supervision à distance.
INTRODUCTION

Automation of routine agricultural processes is not a new concept. One of the first attempts to create an autonomous agricultural tractor happened in the 1940s when engineer Frank W. Andrew connected the tractor steering arm to a fixed wheel to make the tractor run spirals (Condon and Windsor 1940). Over the last two decades, automation of agricultural processes has received significant impetus because of the development of new technologies such as real time kinematic (RTK) global positioning system (GPS), computer vision, guidance sensors, and a general increase of computing power (Li et al. 2009; Lyon 2017). Currently, the development of automation technologies for different agricultural tasks, and autonomous tractors in particular, is a major focus of researchers (Billingsley et al. 2008).

Researchers working to automate routine agricultural tasks in orchards showed that it is possible to improve productivity of spraying and mowing by 30% over manual control (Moorehead et al. 2012). The improvement is achieved due to better path planning, elimination of rest breaks, operation at constant speed, and elimination of tractor reversing. The system requires human supervision to solve problems that automation cannot overcome; however, one operator can simultaneously supervise several tractors. An experiment called “Hands Free Hectare” was conducted in 2017 by a group of researchers from the UK (Hands Free Hectare 2017). The goal of the experiment was to prove that all of the technologies that are required to farm crops without direct human control have been developed. The experiment was finished in September 2017 with a successful harvest using only drones and autonomous agricultural machines. The autonomous tractor presented by New Holland in 2016 (New Holland 2016) is a transitional link between conventional tractors and fully autonomous tractors. The tractor can be operated manually or remotely using a tablet or a computer. The level of autonomy is described as “supervised automation”. The tractor has sensors to follow a predefined path and make turns in autonomous mode, but the supervisor decides how to avoid or bypass obstacles when they are detected.

Autonomous agricultural machines still have safety and reliability issues (Lyon 2017). Typical agricultural machines are big and heavy and may be very dangerous if safety systems fail (Conesa-Muñoz et al. 2015). Current automation does not exclude a human from the human-machine system. Autonomous and semi-autonomous agricultural machines still require human supervision to solve problems that automation cannot overcome – a concept described as “human supervisory control of robots in the performance of routine tasks” (Sheridan 2016). The robot performs a limited set of actions defined by software and has sensors for self-positioning and obstacle detection. At any time, the supervisor has an ability to step in to modify the task or to solve problems that the robot cannot overcome (Scholtz et al. 2004).

Analysis of an operator’s behavior during urban search and rescue operations revealed that an operator relies heavily on video footage due to limitations of sensors and the unstructured environment (Baker et al. 2004). Agricultural tasks are more structured and consistent and in most cases agricultural robots do not require full attention of the supervisor (Bechar and Vigneault 2016; Moorehead et al. 2012). The importance of live video for the supervisor is unknown. Some researchers suggest including live video in the control interface. A control interface of an autonomous tractor for orchard maintenance includes video from three cameras pointing forward, left, and right (Moorehead et al. 2012). A prototype mimic display of tractor status proposed by Blackmore et al. (2002) included video from an on-board steerable camera to allow better understanding of the tractor environment. The control interface of the CNH Industrial autonomous tractor includes video from four cameras mounted on the tractor. The cameras provide 360º view of the area around the tractor (New Holland 2016). There is also an example of a control interface based on a computer graphic without adding live video. The control interface for a tractor team in peat moss harvesting shows status of the tractors and the field scheme using only indicators and computer graphics (Johnson et al. 2009).

While many researchers have put their effort into improvement of different aspects of autonomous machines, development of a convenient and effective interface for the supervisor is also a key problem to overcome (Endsley 2017). The main objective of this study was to determine the importance of live video for remote supervision of an autonomous agricultural machine and to provide recommendations for interface designers.

MATERIAL AND METHODS

Simulators are widely adopted for training, design, and research purposes (Lehto and Landry 2013; Sandom and Harvey 2009). Usage of simulators ensures high flexibility, control over study variables, low cost, and safety of the experiment (Fitts 1951). Simulators are used for assessment of different parameters in human factors research, such as situation awareness (Plavšič et al. 2010), mental workload (Dey and Mann 2009; Pauzić 2008; Young and Stanton 2004), and effectiveness of user interface designs (Galitz 2007; Rakhra and Mann 2013). That is why an autonomous agricultural machine control interface (AAMCI) simulator was developed for this project.

The simulator was designed to create a realistic model of the supervisory task for an agricultural machine. An agricultural sprayer was selected as the machine for the AAMCI simulator. For the proposed research, the sprayer provides some advantages over other types of agricultural machines. The spraying process is simple and easy to explain to a person without farming experience. It is easy to show the details of the process using video cameras because all the operations occur above the soil surface. It
Fig. 1. Left screen of the AAMCI simulator: a) general view camera, b) front view camera, c) right boom camera, and d) left boom camera.

Fig. 2. Right screen of the AAMCI simulator: a) field scheme with tractor and sprayer icons, b) indicators related to the tractor, c) indicators related to the sprayer, d) deviation indicator, and e) pop-up window.
is relatively simple to stage sprayer malfunctions. The operator controls tractor speed, tractor deviation from the path, boom position, and nozzle valves.

A lawn tractor and a plot-size sprayer were used to record video footage for the experiment. Fifteen unique video clips with average length of 2 minutes were recorded. Three of them were used during training trials and twelve during experimental trials. Each video clip included 2 or 3 staged sprayer problems. The staged problems included a nozzle clogging, incorrect boom height, incorrect tractor speed, and tractor deviation from the path. Number, type, order, and timing of the problems were randomized for each video clip to ensure lack of patterns and equal distribution of the problems among the video clips. The videos were recorded on green space on the University of Manitoba Fort Garry campus.

The AAMCI simulator uses two computer screens to provide status information about an agricultural machine to the supervisor. The left screen showed video from four cameras mounted on the sprayer (Fig. 1). A general view camera was mounted on the tractor facing backward (Fig. 1a). The camera was intended to help the supervisor to estimate the sprayer status with a quick glance. A forward view camera was mounted on the right end of the right boom facing forward (Fig. 1b). The intent of this camera was to provide a view of the area in front of the tractor and to help the supervisor to detect obstacles. It was possible to detect tractor speed changes and the tractor deviation using this camera. Nozzle cameras were mounted on brackets approximately in the center of the booms. Each nozzle camera showed three nozzles and could be used to detect nozzle clogging (Fig. 1c, d).

The right screen of the simulator showed the information about the tractor and sprayer status via icons and indicators (Fig. 2). It included a schematic of the field with the animated tractor and sprayer icons (Fig. 2a). Indicators related to the functioning of the tractor (i.e., speed, engine RPM, fuel level, and engine temperature) were located in the top right corner (Fig. 2b). The indicators related to the sprayer (i.e., application rate, boom height, pump pressure, and tank level) and nozzle icons were grouped in the bottom right corner (Fig. 2c). The deviation indicator was located below the field schematic (Fig. 2d). Information on both screens was synchronized manually.

**Participant procedure**

Seventeen participants volunteered to participate in the study; fourteen were male and three were female. All participants were University of Manitoba students between 19 and 40 years of age. Only five participants had farming experience and two of them had experience operating an agricultural sprayer.

The participants were asked to perform a supervisory role. Their task was to watch two screens, detect the machine problems, and mark detected problems on a paper form. At the beginning of the experiment, the participants were given explanation of the experimental goals, the role of the participant, and the functioning of the AAMCI simulator. Next, the participants had three training trials and two experimental sessions with a 10-minute break between them. Each experimental session included six pairs of trials. Information about the machine status within a pair of trials was the same, but in Case 1 it was provided using indicators supplemented by live video and in Case 2 it was provided solely by graphical indicators. Order of the trials within each experimental session was randomized to minimize the influence of order on the results. During Case 1 trials, a problem could be presented using only an indicator (referred to as a Case 1a problem), using an indicator and a video footage simultaneously (referred to as a Case 1b problem), and using video footage solely (referred to as a Case 1c problem). The participants were informed about all possible ways the problems could be presented.

The participants were provided two paper forms, one for the screen with video footage and another for the indicators screen, and instructed to indicate detected malfunctions on the appropriate paper form. The Pupil Pro eye tracker was used for collection of gaze distribution data. It uses two cameras to track the pupil position and head movement. Both of the cameras are mounted on a frame similar to the frame of regular eyeglasses. The Pupil Pro software uses special markers to define surfaces of interest and calculates gaze distribution among them. The video and indicator screens were the two surfaces of interest in this experiment.

At the end of each experimental session each participant was asked to estimate level of confidence in his or her supervisory performance using a scale from 0 to 10 at the bottom of the paper forms. This assessment was done separately for the live video screen and for the indicators screen. At the end of the experiment, each participant answered a short questionnaire and was given the opportunity to provide comments about the experiment. The questionnaire included the following questions: i) how did the presence of video footage affect the level of understanding of machine functions, and level of confidence about machine status?; ii) was the video footage helpful in the detection of sprayer malfunctions?; iii) should video footage be provided to the supervisor fulltime or on demand?. Additionally, participants could give comments about the importance of live video for the supervisory task.

A conclusion about the importance of the live video for remote supervision was made based on analysis of the participant performance, gaze distribution, and subjective responses. The participant performance was assessed by calculating the ratio of correct problem detection to total number of problems. The participant performance compared for Case 1 (video + indicator) and Case 2 (indicator only) trials. The participant performance was compared for Case 1a (indicator only), Case 1b (indicator + video), and Case 1c (video only) problems. The
subjective level of confidence was analyzed in the same fashion as the participant performance. The level of confidence was compared for Case 1 and Case 2 trials, and for Case 1a, 1b, and 1c problems. The questionnaire answers were used for analysis of the subjective opinions about usefulness of the live video for the supervisory task.

RESULTS AND DISCUSSION

Participant performance

It was hypothesized that the presence of the live video would help the participants to detect problems and thus ensure better performance. Contrary to the expected results, the presence of the live video decreased the performance by 9.1% and the difference is statistically significant (t-test, \( \alpha = 0.05 \)). Most likely these results were observed because the live video did not provide sufficient clues to detect problems and at the same time forced the participant to share attention between two screens. Analysis of the different ways problems were presented during Case 1 trials supports this assumption (Fig. 3). Case 1c problems (presented solely by video) were detected in 14.7% of cases. Case 1b problems were correctly detected in 79.6% of cases using indicators, and only 10.4% of cases using video. For both Case 1a and Case 1b problems, the performance was poorer than for Case 2 problems and it was caused by the presence of the video footage. Additionally, there were about 0.2% false positive answers and all of them related to video. The participants of the study commented that even though it was possible to detect problems using video it required more effort than using indicators.

Fig. 3. Percentage of correct answers for the different ways problems were presented during Case 1 trials.

Gaze distribution

An average gaze distribution for each trial was calculated using the Pupil Pro software. The distribution between left (video) and right (indicators) screens is shown in Fig. 4. The graph includes data for three training trials and twelve experimental trials in chronological order. It was observed that the ratio of glances to the left and right screens shifted from 50/50 at the first training trial to a ratio of 30/70 toward the right (indicators) screen by the final trial of the experiment. Typical heat map gaze distribution for the first training (Fig. 5) and the last experimental trials (Fig. 6) show points of interest during the trials. Comparison of the heat maps show that at the beginning the participant attention was almost equally distributed over the two screens. At the end of the experiment, the participants focused on the nozzle camera images and paid much less attention to two other camera images on the left screen. General and forward view cameras provide information about general status of the sprayer, which can be assessed with a quick gaze. Even though it is possible to detect some of the sprayer problems (for example deviation or speed changes) using those cameras, they do not provide obvious clues for problem detection and participants preferred to use other sources of information. At the same time, the participants focused on the nozzle video to detect nozzle clogging.

On the right screen the participant attention was focused on the indicators. The nozzle icons demanded the attention of the participant although this demand ensured a very high level of malfunction detection (i.e. 96.1% and 97.1% for Case 1 and Case 2 trials, respectively). It might be explained by the design of the icons, which ensured Level 2 situation awareness (Endsley et al. 2003). As a supervisor, it is essential that the operator have an awareness of what is happening. This awareness of one’s surroundings is referred to as “situation awareness” (SA), which Endsley et al. (2003) defined as the perception of the elements in the environment (level 1 SA), the
comprehension of the current situation (level 2 SA), and the projection of future status (level 3 SA). The icon turned from a green circle to a red square when the problem occurred. The design of the remaining indicators enabled only Level 1 SA (Endsley et al. 2003) and the achievement of high supervisory performance required more attention.

Subjective responses
The participants provided two types of subjective responses: 1) subjective level of confidence in supervisory performance for camera and indicator screens separately, and 2) questionnaire answers. The comparison of the confidence level for Case 1 trials is provided in Fig. 7. The average level of confidence was 56.0% and 81.3% for the left (video) and the right (indicators) screens, respectively. The difference was statistically significant (t-test, α=0.05). The comparison of the confidence level for the right screen for Case 1 and Case 2 trials is shown in Fig. 8. The presence of live video during the Case 1 trials decreased the average level of confidence by 6.1% and the difference is statistically significant.

The analysis of the questionnaire answers revealed that all participants indicated that live video should be provided to the remote supervisor, although 82.4% of the participants would prefer to receive the video on demand. The remaining 17.6% of participants would prefer the video to be present all the time. The presence of the live video increased the level of understanding of machine status for 76.4% of the participants. Of the remaining participants, 11.8% reported no difference in level of understanding and 11.8% indicated that live video decreased level of understanding. With respect to level of confidence, 52.9% of participants felt more confident when the video footage was provided, 35.3% felt less confident, and 11.8% felt no different. Despite the actual results, 82.4% of the participants indicated that the video footage was helpful for problem detection.

Discussion
The study results showed that video is not suitable for detection of problems with an agricultural sprayer. Indicators linked to sensors can provide the same information in easy-to-understand format and ensure better performance. However, the participants indicated that they felt more secure when the video footage was provided. It is reasonable to deliver important information about machine status using indicators and to use video to provide a general view of the machine and the surrounding area. This conclusion is similar to the control interface designs provided by other researchers (Johnson et al. 2009; Moorehead et al. 2009), where cameras provide a general forward view or the camera is steerable.
From the literature, it has been observed that a supervisor of a rescue robot spends almost all of the time watching video (Baker et al. 2004). In this experiment involving a simulation of an agricultural machine, however, the participants spent only about 30% of their time watching video. A trend in Fig. 4 confirms that the video footage is a secondary source of information for the supervisor. As the participants got more familiar with the task, the percentage of the time they spent watching the video gradually decreased. The difference might be even greater if the design of the indicators had been optimized. Better indicators might help achieve Level 2 and Level 3 SA (Endsley et al. 2003) and ensure performance improvement. Even though the indicators worked well for the purposes of the study, a different design might be more effective.

CONCLUSION
The study results showed that live video is poorly suitable for detection of agricultural sprayer problems and participant performance was poorer during the trials with live video. Participant attention gradually shifted to the indicators screen from 50/50 at the beginning to 70/30 at the end of the experiment. However, the presence of live video helps the supervisor to feel more secure and the participants would like to see live video either full time or on demand. More than 75% of the participants indicated that the presence of live video helped them to understand machine functions better. It is recommended that an automation interface should include an optional window to display live video of the general view of the machine and the surrounding area. Live video may also be useful in situations when indicators do not provide enough information.

ACKNOWLEDGEMENTS
The authors acknowledge the financial assistance of the Natural Sciences and Engineering Research Council of Canada (NSERC) and the University of Manitoba Graduate Fellowship (UMGF) program. A special thanks to Eric Hawley, Dale Bourns and Matt McDonald for their technical assistance.

REFERENCES


https://doi.org/10.3390/s150305402


https://doi.org/10.1177/0018720816681350

https://doi.org/10.1002/hfm.20297


https://doi.org/10.1002/rob.20297


Lyon, N. 2017. Clearing the roadblocks to autonomous tractors.


https://doi.org/10.1002/hfm.20173

https://doi.org/10.13031/2013.42542


https://doi.org/10.1177/0018720816644364

https://doi.org/10.1080/00140130410001686348