



THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Enhancing Awareness to Support Teleoperation of a Bulldozer

John James Dudley

B. E. (Mechanical) (Hons), B. A.

*A thesis submitted for the degree of Master of Philosophy at
The University of Queensland in 2014
School of Mechanical & Mining Engineering*

Abstract

Bulldozers are frequently used to execute work in hazardous environments. Teleoperation provides a viable method for allowing operators to perform work without being directly exposed to the hazards of these environments. However, removing the operator from the bulldozer presents many challenges associated with reduced task engagement and controllability. The relocation of the operator and placement within a remote teleoperation system may be simplistically modelled as placing a filter on the operator's inputs and outputs. The various cues that would excite the sensory systems of an operator on board the bulldozer must be remotely replicated. However, the fidelity and timeliness of cues provided to the teleoperator are unavoidably constrained by limitations in the systems that effect this replication. Likewise, there are similar implications for accurately achieving a machine response in accord with the command outputs of the operator.

This thesis focuses principally on the inputs to the teleoperator with the aim of identifying how perception enhancements might be applied to alter the characteristics of this input filter. The motivation for this research is to determine what factors are critical to achieving high levels of teleoperation performance and user acceptance.

To conduct this investigation, an enhanced perception cell capable of high fidelity replication of motion, visual and aural cues is integrated with an existing bulldozer teleoperation system. The cell enables targeted analysis of the influence of individual feedback cues on performance and user acceptance. Experiments have been conducted with the enhanced perception cell for a structured bulldozing task. Results indicate that visual quality is a dominant factor influencing operator performance. Motion feedback provides no additional benefit beyond that provided by enhanced visual quality. The value of task visualisation to support accuracy and planning is also highlighted.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the General Award Rules of The University of Queensland, immediately made available for research and study in accordance with the *Copyright Act 1968*.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

John Dudley, Jonathan Aw, and Ross McAree. Minimal perception requirements to support effective remote control of bulldozers. Technical Report C20021, ACARP, 2013.

Publications included in this thesis

John Dudley, Jonathan Aw, and Ross McAree. Minimal perception requirements to support effective remote control of bulldozers. Technical Report C20021, ACARP, 2013.

This report was prepared as an outcome of the wider research project from which this thesis stems. There are portions of text, considered directly relevant and authored by myself, and results from this report which are included in this thesis.

Contributor	Statement of contribution
John Dudley (Candidate)	Authorship and editing (70%)
Jonathan Aw	Authorship and editing (15%)
Ross McAree	Authorship and editing (15%)

Contributions by others to the thesis

The main study, of which the research presented in this thesis represents a component, was undertaken as part of a project funded by the Australian Coal Association Research Program (ACARP). This research project, entitled ‘Perception Requirements for Effective Teleoperation of Dozers’, was executed by the Smart Machines Group at the University of Queensland through the Cooperative Research Centre for Mining (CRCMining) with extensive support from Caterpillar Inc.

The design and construction of the experimental apparatus used within this study was a collaborative effort of the Smart Machine Group project team. The execution of the experiments as well as the continued development and maintenance of the experimental apparatus was also a collaborative exercise.

Statement of parts of the thesis submitted to qualify for the award of another degree

None.

Acknowledgements

This thesis has evolved from a collaborative research project executed by CRCMining and Caterpillar with funding from the Australian Coal Association Research Program (ACARP). I would like to acknowledge the contribution made by these groups in providing the funding and in kind support that allowed this project to be conducted and gave me the opportunity to write this thesis.

I owe much credit to the many people involved in the project. I would to thank those at Caterpillar for their kind and patient assistance. I feel the bulldozer operators who acted as subjects in this study deserve special recognition for their resilience and helpfulness despite the often frustrating experimental regime to which they were subjected.

I would like to thank my colleagues at CRCMining who were instrumental in originating the project and seeing it achieve its outcomes: Dr Anthony Reid who fathered the project and gave it direction; Dr Zane Smith who developed the software and sacrificed his life to testing; Jonathan Aw who developed the audio-visual system and endured my character flaws with his classic composure; as well as Lachlan Palmer and a number of others who at various times gave assistance. Lastly, I would like to thank my supervisor, Professor Ross McAree, who challenged, educated, prodded and reformed me while also giving me the time and freedom to explore my interests.

Keywords

teleoperation, bulldozer, perception, motion feedback, visual feedback

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 091302, Automation and Control Engineering, 100%

Fields of Research (FoR) Classification

FoR code: 0913, Mechanical Engineering, 100%

Table of Contents

1	Teleoperation, remote perception and operator performance	1
1.1	Teleoperation	1
1.2	The teleoperated bulldozer	3
1.3	Does perception influence performance?	4
1.4	Perception for teleoperation	5
1.4.1	Visual feedback	5
1.4.2	Audio feedback	8
1.4.3	Cue substitution	9
1.4.4	Latency	11
1.4.5	Motion feedback	12
1.4.6	Assessing perception requirements	17
1.5	Thesis overview	19
2	Examining perception requirements for teleoperation	21
2.1	Aim	21
2.2	Methodology	22
2.3	Perception cues relevant to bulldozer operation	24
2.4	Feedback factors for experimental evaluation	26
2.4.1	Vision	26
2.4.2	Motion	27
2.4.3	Audio	28
2.4.4	Contextual awareness	29
2.4.5	Summary of feedback factors to be evaluated	29
3	Experimental evaluation of perception requirements	31
3.1	Aim	31
3.2	The conventional teleoperation system	31
3.3	The enhanced perception cell	32

3.3.1	Audio-visual feedback system	33
3.3.2	Motion feedback system	35
3.3.3	Task visualisation	36
3.4	Experimental procedure	37
3.4.1	Design factors	38
3.4.2	Structured task	38
3.4.3	Nuisance factors	39
3.4.4	Experimental Controls	40
3.4.5	Performance metrics	40
3.4.6	Experimental plan	43
3.5	Results	44
3.5.1	Operator performance is degraded under teleoperation	45
3.5.2	Introducing motion feedback does not substantially influence task completion time	46
3.5.3	Depth of view dominates visual quality and audio	54
3.5.4	Task visualisation is critical to task completion	69
3.6	Discussion of key results	71
4	Task analysis for bulldozer operation	75
4.1	Aim	75
4.2	Linking perception to task behaviour	75
4.3	Modelling the human	76
4.4	Analysis of bulldozer operation and teleoperation	77
4.4.1	Task and cue surveys	77
4.4.2	Operator feedback	84
4.5	Modelling bulldozer control behaviour	86
4.6	Mapping perception to control	90
4.7	Overview of assembled model	94
5	Consolidating perception requirements	95
5.1	Aim	95
5.2	Perception groupings	95
5.2.1	Vision	95
5.2.2	Audio	97
5.2.3	Motion	97
5.2.4	Task visualisation	99

5.3	Ranking perception requirements	100
6	Conclusions and recommendations for further investigation	103
6.1	Conclusion	103
6.2	Original contributions of the thesis	105
6.3	Recommendations for future work	105
6.4	Concluding remarks	106
	Bibliography	107
A	Experimental Record	113

List of Figures

1.1	Afferent and efferent filters in teleoperation.	2
1.2	Elements of a bulldozer.	3
1.3	Degrees of freedom of a bulldozer blade.	4
1.4	Teleoperation of TALON robot using 3D interface.	7
1.5	Max Planck Institute motion simulator.	15
2.1	View from inside bulldozer cab.	25
2.2	Areas of significantly restricted visibility on a Caterpillar D8T.	25
3.1	Conventional teleoperation system.	32
3.2	Cropped segment of output feed at remote station illustrating difference in video quality.	33
3.3	Typical 3D frame prior to display interlacing.	34
3.4	Anaglyph representation of the 3D view.	34
3.5	Bulldozer frame and Platform frame.	35
3.6	Task visualisation display.	37
3.7	Structured test task.	38
3.8	On board versus non-line-of-sight teleoperation.	46
3.9	Completion under different motion feedback configurations	47
3.10	Motion cues - mean volume removed from within slot over time.	49
3.11	Motion cues - mean time taken to remove volume segment.	51
3.12	Motion cues - mean volume per push.	52
3.13	Operator A - mean workload total score.	53
3.14	Operator A - mean workload element scores.	54
3.15	Completion under different visual and audio quality configurations.	55
3.16	Completion with audio, visual and motion enhancements.	57
3.17	Combined vision, audio and motion cues - mean volume removed from within slot over time.	59

3.18	Combined vision, audio and motion cues - mean time taken to remove volume segment.	61
3.19	Combined vision, audio and motion cues - mean volume per push.	62
3.20	Left joystick control inputs over test duration.	63
3.21	Right joystick control inputs over test duration.	64
3.22	x-position of bulldozer within the slot over time.	65
3.23	Mean number of joystick corrections per minute.	66
3.24	Mean joystick inputs per minute on each control axis.	66
3.25	Operator B - mean workload total score.	67
3.26	Operator B - mean workload element scores.	68
3.27	Without task visualiser - percentage complete at termination of test.	69
3.28	Without task visualiser - mean volume removed from within slot over time.	70
3.29	Without task visualiser - mean workload total score.	72
3.30	Without task visualiser - mean workload element scores.	72
4.1	Situation awareness model.	78
4.2	State transition model.	87
4.3	Hierarchical Task Analysis for slot bulldozing.	88
4.4	Model of bulldozer operation.	91

List of Tables

3.1	Audio-visual feedback system cues.	35
3.2	Isolated motion cues.	36
3.3	Augmentations that can be applied to the existing remote station.	38
3.4	Nuisance factors.	39
3.5	NASA TLX rating scale definitions.	42
3.6	Data collection stage 1 - summary of test configurations.	44
3.7	Data collection stage 2 - summary of test configurations.	45
3.8	On board versus non-line-of-sight teleoperation descriptive statistics.	46
3.9	Motion configurations.	47
3.10	Descriptive statistics for motion feedback configurations.	48
3.11	Motion factors and levels.	48
3.12	Operator A - total workload score.	53
3.13	Audio-visual configurations.	55
3.14	Descriptive statistics for visual and audio quality test configurations.	56
3.15	Vision factors and levels.	56
3.16	Audio factors and levels.	56
3.17	Combined vision and motion configurations.	57
3.18	Descriptive statistics for configurations with audio visual and motion en- hancements.	58
3.19	Vision and motion factors and levels.	58
3.20	Joystick correction count for typical test.	65
3.21	Operator B - total workload score.	67
3.22	Without task visualiser - total workload score.	71
4.1	Motion focus - task focus statements.	79
4.2	Motion focus - cue focus statements.	79
4.3	Motion focus - task focus statement responses.	80
4.4	Motion focus - cue focus statement responses.	81

4.5	Vision, audio and motion focus - task focus statements.	82
4.6	Vision, audio and motion focus - cue focus statements.	82
4.7	Vision, audio and motion focus - task focus statement responses.	83
4.8	Vision, audio and motion focus - cue focus statement responses.	83
A.1	Completion time results - stage 1	114
A.2	Completion time results - stage 1 continued	115
A.3	Completion time results - stage 2	116
A.4	Completion time results - stage 2 continued	117

List of Abbreviations

NMF	Remote station with no motion feedback
CMF	Remote station with combined motion feedback
PMF	Remote station with pitch motion cues only
VMF	Remote station with vibration cues only
TMF	Remote station with translational acceleration cues only
NMF-TV	Remote station with no motion feedback and no visualiser
CMF-TV	Remote station with combined motion and no visualiser
2DSR	Remote station in standard audio-visual configuration
2DHR	Remote station with high resolution 2D video and stereo audio
3DHR	Remote station with high resolution 3D video and stereo audio
2DHR-A	Remote station with high resolution 2D video and no stereo audio
3DHR+PMF	Remote station with high resolution 3D video, stereo audio and pitch motion cues
3DHR+FMF	Remote station with high resolution 3D video, stereo audio and full motion feedback
OB	On board operation

Teleoperation, remote perception and operator performance

1.1 Teleoperation

As humans we are capable of acting as expert control systems. This capability is due in part to the robust feedback channels we are able to establish through our sensory processes. Our senses are involved in everything we do and we attach great value to the role they play in supporting interaction with our surroundings.

Teleoperation is *the operation of machines at a distance by remote control* (Oxford English Dictionary, 2013). An effective teleoperator¹ must feed the senses of the human controller; it must extend human sensing to the remote environment so that the operator receives sufficient information, presented in a suitable format for remote machine operation to be effective. It must also translate the operator commands so that they result in intended actions by the machine at the remote environment.

This thesis is concerned with the teleoperation of bulldozers in mining operations, the motivation for which is the desire to reduce the fatal risks that operators can be exposed to when on-board these machines. Teleoperation in this application fundamentally alters the conventional relationship between the operator and the machine. The teleoperator is interposed between the operator and machine: cues relevant to machine operation must be remotely replicated to the operator and must initiate appropriate operator commands

¹Any remote-controlled machine which mimics or responds to the actions of a human controller at a distance (Oxford English Dictionary, 2013)

to control the machine. Sheridan (1992) has abstracted the bilateral communication functions of teleoperators as filtered channels that he calls the *afferent* and *efferent* filters, see Figure 1.1.

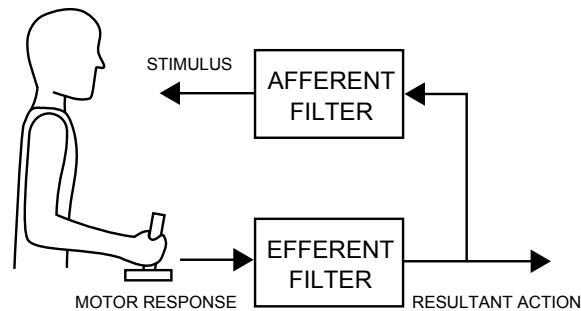


Figure 1.1: Afferent and efferent filters in teleoperation. After Sheridan (1992).

If one assumes that direct manual operation results in ideal closed-loop dynamics and that perfect fidelity and timeliness are unachievable for a teleoperation system then this ‘filtering’ must be expected to have a negative effect on closed-loop dynamics. Consistent with this theory, a deterioration in task performance has been observed in many applications of teleoperation.

When no practical alternative is available, reduced performance is accepted. However, when the risks associated with performing a task are conventionally tolerated then significantly reduced performance is unacceptable. For teleoperation to be embraced in these circumstances it is imperative that such systems receive careful design attention focused on maximising operator performance. This need provides motivation for obtaining a better understanding of the influence of teleoperation systems on operator perception, control behaviour, user acceptance and ultimately performance.

This thesis examines the role of perception within the theoretical context of the ‘filtering’ effect of teleoperation systems. The aim of this thesis is to evaluate the hypothesis that careful design of the ‘filtering’ characteristics of a teleoperation system can deliver improved operator performance and user acceptance. The ultimate goal is to identify attributes that are critical in a bulldozer teleoperator to achieve high levels of performance.

1.2 The teleoperated bulldozer

Bulldozers are highly versatile, multipurpose machines. The key elements of a bulldozer are shown in Figure 1.2. Their principle application is in cutting down into terrain and pushing material with their blade. On mining class bulldozers, the blade will typically have three degrees of freedom as shown in Figure 1.3. To disturb competent or semi-competent material, the ripper is lowered into the ground and pulled through the material. Tracked bulldozers are capable of traversing steep, rough and unstable terrain. The drive mechanism also allows for very tight turning circles.

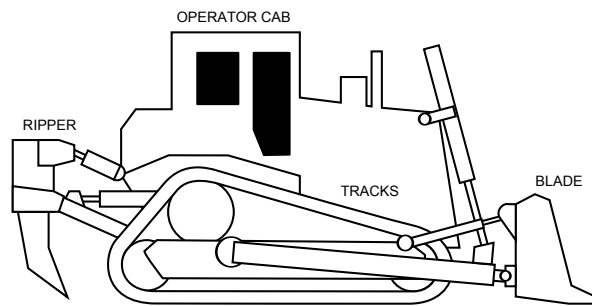


Figure 1.2: Elements of a bulldozer.

The versatility and manoeuvrability of bulldozers means that they are commonly used in many hazardous circumstances encountered in mining. Teleoperability is thus a desirable attribute for bulldozers. However, the bulldozer is also frequently used in production tasks where efficiency directly translates to throughput capability of the mining process. In these circumstances, the level of productivity that can be achieved with teleoperation is a major consideration in determining its uptake. This introduces the requirement for teleoperated bulldozers to be capable of levels of performance similar to on board operation. There is potential to advance the maturity of this technology through a better understanding of the influence of teleoperation on performance and where opportunities for improvement exist.

Teleoperation systems for bulldozers are commercially available. These systems can be broadly categorised into two types: line-of-sight and non-line-of-sight. Line-of-sight teleoperation requires the bulldozer to be within direct view of the operator. The control mechanism of such systems is typically hand held or slung from the shoulders and is not too dissimilar from a conventional hobby remote control. The implication of teleoperators of this type is that the operator must still be in direct line-of-sight to the bulldozer. Under

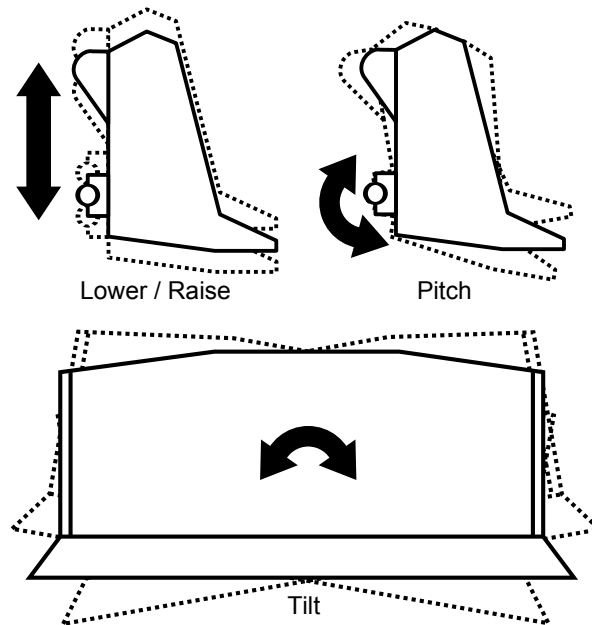


Figure 1.3: Degrees of freedom of a bulldozer blade.

certain operational conditions, this requirement may mean that the hazards to which the operator is exposed may be largely unaltered or even increased.

Non-line-of-sight teleoperation describes the condition where the bulldozer does not need to be within direct view of the operator. Instead, the operator is typically provided with vision from cameras installed on the bulldozer, and/or other cameras focussed on the work area. While non-line-of-sight teleoperation removes any maximum distance constraint, it does present significant challenges around transmitting and recreating the perception cues necessary to the effective operation of the bulldozer.

1.3 Does perception influence performance?

While the advantages that bulldozer teleoperation presents are well recognised, the limitations and opportunities for enhancement of these systems are not well known. The afferent-efferent filter model discussed earlier suggests that the introduction of a teleoperation system acts like a filter on the operator's inputs and outputs.

A hypothesis is proposed that performance in teleoperation can be enhanced by minimising the effect of this filter. This hypothesis is to be tested by investigating the influence of increased fidelity and timeliness of feedback cues provided to the operator performing

1.4 Perception for teleoperation

non-line-of-sight teleoperation. The research question this thesis seeks to answer is: *what feedback attributes are critical to maximising bulldozer teleoperation performance and user acceptance?*

1.4 Perception for teleoperation

Humans perceive their environment through interpreting sensory cues. This perception is critical to performing physical tasks. The relative significance of different sensory systems in perceiving an environment and performing a task varies depending on the nature of the environment and the nature of the task.

The operation of a vehicle is a relatively complex exercise where multiple sensory systems are called upon. The visual, aural, vestibular, tactile and proprioceptive systems all contribute to the operator's perception and ability to maintain control. The information provided by these sensory systems is fused by neural pathways before being further processed within different loops to derive an appropriate control response. The need to capture the information that would normally be obtained directly by the sensory systems and instead replicate it elsewhere so that it might be similarly interpreted is a significant technical challenge for teleoperation systems.

Teleoperation is not new and extensive research has been conducted into identifying factors that influence performance. Numerous studies have examined the role played by the different sensory systems and how feedback might best be provided.

Prominent avenues of exploration have been in visual feedback, audio feedback, cue substitution and the impact of latency. Motion feedback has emerged more recently as a relevant cue for teleoperation in performing certain activities.

1.4.1 Visual feedback

The human visual system is specialised for the perception of spatial structure. The ability to perceive and interpret space is critical in many teleoperation applications. Unfortunately, the capture and transmission of visual feedback for use in teleoperation can be data intensive. In addition, the natural capabilities of humans in an environment, i.e. binocular vision, field of view, and the ability to move the head to obtain different views, can be very difficult to accurately replicate remotely.

The challenge of providing appropriate visual information in teleoperation has motivated numerous studies. Chen et al. (2007) conducted an extensive review of the literature on vision-related human performance issues around teleoperation. The common vision-related issues encountered are lack of depth perception, poor situation awareness and cognitive tunnelling due to restricted field of view. The consolidation of the findings from the reviewed literature leads to the authors recommending wide field of view, stereoscopic displays and predictive displays in presence of high latency.

Murphy (1995) examined the potential for using a panospheric camera that is capable of capturing a full 360 degree field of view. Aside from the benefit of not requiring multiple or actuated cameras to capture appropriate field of view, the display produced by the panospheric camera was suspected of supporting a more immersive experience due to its seamless representation. Halme et al. (1999) identified that head tracking for adaptive visual display showed benefits when an unfamiliar or changeable task was encountered. However, there was no obvious benefit when a familiar task was being performed.

Depth information, as provided by stereoscopic video images, has been found to enhance certain aspects of performance but with its benefit highly dependent on the nature of the task. Drascic et al. (1989) examined the impact of 2D versus 3D visual feedback in telemanipulation of a bomb disposal robot application. The results of the study indicated that stereoscopic vision reduced completion time initially but with sufficient practice, similar completion times were achieved in the 2D visual configuration. This suggests that 3D depth cues are more intuitive, but similar information can be gained from the 2D video with sufficient practice.

Depth information has been found (Ferre et al., 2005) to marginally reduce task completion time in the execution of a peg-in-hole task performed using a telemanipulator. Lee and Kim (2008) also found an improvement in performance associated with the provision of stereoscopic images from a teleoperated robot performing an obstacle avoidance and navigation task. Edmondson et al. (2010) investigated the benefits of a 3D display over a 2D display for seven operational tasks encountered in teleoperation of a TALON bomb disposal robot (see Figure 1.4). 3D was found to improve performance in six of the seven tasks evaluated. 3D vision was also widely preferred by users but not by all.

Scribner and Gombash (1998) also examined the potential for stereoscopic vision to support teleoperation of a UGV. Results obtained indicate that stereo vision can reduce the rate of obstacle collisions but increases operator stress. However, the mean time taken to

1.4 Perception for teleoperation

navigate the test course was not found to be significantly influenced by viewing condition. Livatino et al. (2008) performed an evaluation comparing monoscopic and stereoscopic viewing in performance of a UGV navigation task. The task involved navigation of a UGV over a short distance (3.5m) through a field of obstacles. Results from experiments indicate that stereoscopic vision significantly reduced the number of obstacle collisions but no significant difference in completion time was detected. Importantly, users also reported no reduction in viewing comfort and a significant improvement in perceived realism.

A study by Vivéash et al. (2002) using the same stereoscopic visual implementation in experiments for both a driving and a manipulation task found that the relative benefit was less significant in the driving task compared with the manipulation task. This result is similar to that of Halme et al. (1999) in the sense that the benefits of high fidelity visual information are more significant when the task is changeable or when there is a clear need for depth information. The study performed by Lee and Kim (2008) placed subjects at a different starting position and set a different goal position in each run so that the task was always different. By contrast, the driving task in the study by Vivéash et al. (2002) involved subjects navigating the same circuit eight times. The lack of a clear benefit associated with stereoscopic images in the results of Vivéash et al. (2002) is thus likely a consequence of subjects becoming accustomed to the task and successfully interpreting monocular cues to the same effect.



Figure 1.4: Teleoperation of TALON robot using 3D interface (Edmondson et al., 2010).

The findings of Scribner and Gombash (1998); Halme et al. (1999); Vivéash et al. (2002); Lee and Kim (2008) generally appear to indicate that the benefits of high fidelity visual information are more significant when the task is changeable or when there is a clear need for depth information. McIntire et al. (2012) undertook a comprehensive review of the literature covering the performance influence of 3D versus 2D displays. Out of 71 studies

reviewed, 41 (58%) showed 3D to be better than 2D based on the performance metric selected in each study. When the studies were categorised into task types, the highest ratios of results showing positive impact of 3D was seen in spatial manipulation and spatial understanding tasks. However, McIntire et al. (2012) make an important consolidating observation that ‘3D helps little or sometimes not at all for tasks that are simple or well-learned, or for tasks that do not rely heavily on depth information for good performance.’ What the research into stereoscopic vision for task performance does show is that there are potential benefits that can be derived and their are implementations that are accepted by users. However, there appear to be many factors that dictate whether or not stereoscopic vision will be beneficial for a given teleoperation task.

1.4.2 Audio feedback

Audio can provide a powerful cue to support teleoperation through a variety of mechanisms. One key mechanism is the perception of spatial information by means of the audible signal emitted from an object. For example, the loudness of an audio signal can give an indication of the distance to the emitting object while direction can be inferred from differences in the audio signal perceived by the two ears. Another mechanism is the perception of information about an object’s state and behaviour based on the emitted acoustic response. Audio feedback also has strong potential for providing information that is readily interpreted without distraction through mechanisms such as voiced warnings and alarms. Furthermore, studies of the human sensory system indicate that human reactions to acoustic stimulus are faster than reactions to visual stimulus (Welch and Warren, 1986).

Nagai et al. (2002) investigated the potential benefit of audio feedback in assisting teleoperation of a manipulating arm in space. The audio feedback provided to operators consisted of three components. The first audio component generated a motor sound that was proportional to the magnitude of the force and torque imparted on the end-effector. The second audio component involved voiced announcement of key state changes. The third audio component voiced the command data issued as auditory confirmation. In limited experimental trials, a reduction in task completion time of between 30 and 50% was observed. The audio feedback also reduced the amount of time spent focused on the status display. An additional finding of the study was that distinctly different results were observed when tests were conducted in a virtual simulation of the task compared to actual teleoperation of the real satellite arm. This was suspected to be a consequence of

1.4 Perception for teleoperation

the added urgency and stress associated with the real task and the correspondingly real consequences of poor performance.

Liu and Meng (2005) examined the impact of audio feedback in supporting teleoperation of a mobile robot in an obstacle navigation and target finding task. The researches examined four conditions: vision only, audio only, vision and audio in consistent directions, and vision and audio in random directions. Due to observed learning effects in the experiments, the completion time results are not readily interpretable. Some useful general observations were made, however, including that the visual and audio feedback corroborate each other but when conflicting audio information is present, vision will dominate. This study was re-executed by Liu and Wang (2012) with a more precise audio feedback implementation. The results indicated no significant difference in completion time between vision-only and audio-only configurations suggesting that audio may be completely substituted for vision in a simple task with relevant audio cues.

Complicating any assessment of the influence of individual feedback cues, is the fact that audio feedback has been found to have an interaction effect on visual perception. This is most observable in the case of resolving ambiguous visual cues describing motion. Sekuler et al. (1997) conducted an experiment in which an ambiguous motion situation was presented to subjects. Two identical disks were presented on a digital display and shown to move directly towards each other, coincide, then continue in the same trajectory. The presence and timing of audio cues at and around the point of coincidence was investigated for its impact on whether subjects perceived the objects to have bounced off each other or to have passed through each other. The researchers observed that when an audible click was presented at the point of coincidence, approximately 60% of the subjects indicated that they perceived ‘bouncing’ as compared with only 20% when no click was presented. While this finding clearly shows that there are complex human intersensory interactions, it does suggest that audio cues may be useful in resolving ambiguous information or otherwise corroborating perception cues.

1.4.3 Cue substitution

The difficulty associated with providing useful visual information in certain applications has motivated research into cue substitution. The objective in these studies is to determine whether visual information can be substituted with alternate sensory information that is intuitive to environment perception. In 1936, de Florez demonstrated that pilots could successfully maintain control of an aeroplane in flight while blindfolded when two

instrument values were presented aurally. This indicates that vision can be completely substituted by audio in certain control tasks.

Massimino (1992) performed a comprehensive study of the potential for sensory substitution to assist in space teleoperation. Specifically, the focus of Massimino was on investigating the potential for auditory and vibrotactile displays to act as substitutes for force feedback. Experimental results were used to specify three distinct models for the effect of sensory substitution under different qualities of visual feedback. When visual conditions are ideal, simple sensory substitutions have no effect on performance. When visual conditions are degraded, simple sensory substitutions can improve performance. However, when visual conditions are ideal but sensory substitutions are complex, the result is distraction and degraded performance.

Lathan and Tracey (2002) experimented with a gesture based control glove for teleoperating a small UGV. In addition to conventional video and audio feedback, object proximity information was given by vibrotactile inserts in the sleeve of the control glove. The sonar array on the UGV was mapped to particular vibrotactile inserts in the sleeve to provide an indication of the direction of an obstacle. The number of task errors was found to be significantly reduced when the vibrotactile feedback was provided in addition to the video feedback. Spatial awareness was evaluated through four separate tests targeting recognition and manipulation aspects of spatial awareness. The results of Lathan and Tracey indicated that higher spatial awareness resulted in faster task completion times. However, this trend was not observed in the condition with the vibrotactile feedback. These findings have also been echoed in experiments conducted by Uusisalo and Huhtala (2009, 2011) in which a small hydraulic excavator was remotely operated by means of a hand-held gaming controller capable of providing vibrotactile feedback. Testing with inexperienced operators found that task completion time was reduced when using the remote control compared with on board operation. However, the impact on task completion time of adding vibrotactile feedback to the remote control was negligible. Despite this limited effect on task completion time, the subjects' sense of perception and feeling of control (as determined through a survey) was enhanced by the inclusion of the vibrotactile feedback. In discussion of their experimental results, Lathan and Tracey suggest that vibrotactile feedback may have been more useful to the operators with low spatial awareness. Lathan and Tracey conclude that an individual's spatial perception skills may dictate what feedback characteristics they require to effectively perform a given task.

1.4 Perception for teleoperation

These studies have shown that information provided through non conventional channels can provide some compensation for poor or even complete lack of direct visual feedback. The general philosophy that has evolved around teleoperation encourages the use of multiple channels of sensory information provided they are coordinated and intuitive. This suggests the value of determining what sensory information is essential to effectively performing the task that will be teleoperated and providing cues based on this assessment.

1.4.4 Latency

Examination of factors concerning a teleoperator's ability to make changes in the environment has revealed that time delay is a key determinant of performance. A number of early studies (Sheridan and Ferrell, 1963; Adams, 1962), motivated by the large transmission delays expected in earth-space teleoperation, were made to investigate the effect of delay on operator control. Experiments conducted by Sheridan and Ferrell (1963) demonstrated that when performing a remote manipulation task in the presence of delay, operators will follow a move-and-wait strategy. If this strategy is used, completion time can be approximately calculated as a function of the number of discontinuous moves required to achieve tolerance and the magnitude of the delay.

Adams (1962) conducted a comprehensive set of experiments with a remotely operated four wheeled ground vehicle. It was observed that a continuous path could be navigated with 98% accuracy at 2.96 km/h but if a three second delay was introduced to the control loop then a similar level of accuracy could only be achieved after the vehicle speed was reduced to 0.44 km/h (approximately 15% of the original speed). However, it has also been shown (Cunningham et al., 2001) that for certain tasks under certain circumstance, operators can adapt to delay and achieve comparable performance to executing the same task with no delay.

In the consolidated literature review performed by Chen et al. (2007), latency was widely found to be detrimental but the latency threshold past which performance is significantly degraded varies depending on the task performed (from 170 ms up to 500ms). Variation in latency is suggested to be particularly detrimental, and potentially more detrimental than larger but fixed latency.

1.4.5 Motion feedback

The concept of applying vestibular cues to assist teleoperation through direct motion feedback has received attention only recently. This is largely a result of the traditionally high cost of motion-bases and quality motion sensors. Nevertheless, significant literature exists around how vestibular cues influence human control behaviour and performance thanks to extensive research aimed at developing accurate pilot models to assist with the design of aircraft dynamics and simulators.

The relevance of vestibular cues to many vehicle operation tasks suggests that motion feedback can provide a useful channel of sensory information for teleoperation. Early studies (Shirley and Young, 1968; Stapleford et al., 1969; Ringland et al., 1971; Ringland and Stapleford, 1972; Levison, 1976) aimed at using motion feedback in conjunction with some vision based task to examine operator control behaviour. This early research was motivated by a desire to better understand operator control behaviour in order to assist with the design and testing of aircraft dynamics.

Shirley and Young (1968) examined the influence of roll-motion cues in isolation on a compensatory tracking task with a joystick. Their results indicated that adding motion feedback to the visual feedback had an effect of generating additional phase lead above frequencies of 3 rad/s and increasing gain more broadly. This resulted in a reduction in tracking error on the majority of vehicle dynamics evaluated with the effect more pronounced for lower-order systems.

Stapleford et al. (1969) made a similar study of the influence of motion feedback on pilot dynamics in a roll control task. The task involved simulated hovering in gusty air with roll and lateral translation motion imparted to the operator. The experimental results indicated that visual feedback dominates pilot dynamics at low frequencies while motion feedback dominates at high frequencies. Consistent with Shirley and Young, it was observed that motion feedback increased gain and reduced phase lag. There was a clear distinction between controlled elements requiring low frequency pilot lead compared to those that do not, with the former showing a greater improvement with motion feedback than the later. The authors also considered the design of simulators making the comment that washout less than 2 rad/s has only a minor effect. Consideration of vestibular sensory thresholds was also given with pitch and roll rate thresholds of 2.6°/s and 3.2°/s respectively and a linear acceleration threshold of 0.01g described. The small magnitude of the linear acceleration threshold places challenging demands on motion washout filter design for

1.4 Perception for teleoperation

travel limited simulators as observed in a later study by Ringland and Stapleford (1972).

Ringland et al. (1971) extended this original study by Stapleford et al. (1969) with a six degrees-of-freedom simulator performing a similar precision hovering task but with additional motion cues. Trials run with an angular motion only configuration revealed a clear distinction in performance and general operator preference compared with a combined linear and angular motion configuration. Test subjects commented that the linear motion was at times confusing and distracting. The researchers suggested that the angular motion only configuration was beneficial because the g-vector could be used more directly as an indicator of attitude. Investigation of operator instrument scanning behaviour revealed that the introduction of motion reduced dwell time and dwell fraction on the attitude display. In terms of motion fidelity, the authors noted the operators' comments about effects of simulator noise, vibration and motion limiting (hitting motion-base range limits) being particularly disconcerting. These same comments were reiterated in later experiments exploring a different control task (Ringland and Stapleford, 1972). A brief investigation of the operators' sensitivity to vision-motion disparity revealed that an effective time constant of 0.2s or greater was detrimental.

In a study similar to that of Shirley and Young (1968) and Stapleford et al. (1969), Levison (1976) obtained conflicting results on the effect of motion cues. In contrast with the former studies, Levison found that motion cues generated phase lag at high frequencies rather than generating phase lead. Also, a greater improvement was demonstrated with the introduction of motion cues for higher-order systems. The model developed by Levison also suggests that if operator attention is distracted by another task then motion cues will have a greater benefit.

Hosman and van der Vaart (1976) explored thresholds of motion perception again motivated by adding fidelity to the pilot model. It was found that distraction increases motion perception thresholds and also that person to person perception thresholds can differ significantly. However, if the operator has an internal model of how the system they are controlling is expected to behave, their motion perception is improved.

To gain a better understanding of how sensory systems interact, a number of studies have attempted to identify the interplay between the visual and vestibular systems. It has been shown that visual cues implemented in subtly different ways can have an effect similar to providing motion cues. For example, Hosman and van der Vaart (1981) demonstrated that performance of a roll disturbance rejection and tracking task can be improved by the

provision of motion and/or peripheral visual cues. For the results obtained, the improvement provided by motion cues compared to central monitor display only was 60% in the disturbance task and 33% in the following task. Analysis of the closed loop characteristics provides results consistent with established theory for the disturbance task (increased crossover frequency) but inconsistent results for the following task (decreased crossover frequency). Hosman and van der Vaart (1981) describe two alternate theories for the apparent positive interaction between vestibular and peripheral visual information: either redundant information improves accuracy or differences in sensors and perceptions combine to improve accuracy. A more recent study by Hosman and Stassen (1999) identified that motion stimuli are perceived earlier by the vestibular system than they are by the visual system.

Schroeder (1998) examined the requirements for helicopter flight simulation using a high displacement motion platform. The study examined the relative effects of roll, yaw, lateral translation and vertical motion. Lateral translation and vertical motion were found to be the dominant motion cues in improving performance, perceived fidelity and reducing workload. Yaw and roll were found to be of less importance. An additional test demonstrated that visual yaw cues could provide accurate sensation without actual yaw motion. The identified benefit of vertical platform motion conflicted with early comments from subjects who indicated that the vertical axis control task was primarily visual. On this point, Schroeder (1998, chap. 9) cautions that care should be taken when interpreting the subjective impressions of subjects relating to the benefits of particular motion cues.

Beykirch et al. (2007) employed a novel motion simulator based on a robotic arm (see Figure 1.5) to replicate aspects of the experiments of Schroeder (1998). Results obtained indicated that reducing from full to less motion scaling in roll and lateral motion increased the magnitude of control inputs, a consequence of the need to generate lead from the visual information rather than the motion cues. This finding was consistent with Schroeder (1998). However, in results contradictory to Schroeder (1998), the no motion baseline showed a decrease in joystick input magnitude and rate. For lateral hover stabilisation, the experimental results indicated that reducing motion scaling prevented the task from being performed effectively but when all motion was removed, the task performance was only marginally worse than with full motion feedback. The authors comment that the differences observed in the results from Schroeder (1998) may be due to the use of non-professional pilots being used in the study.

1.4 Perception for teleoperation

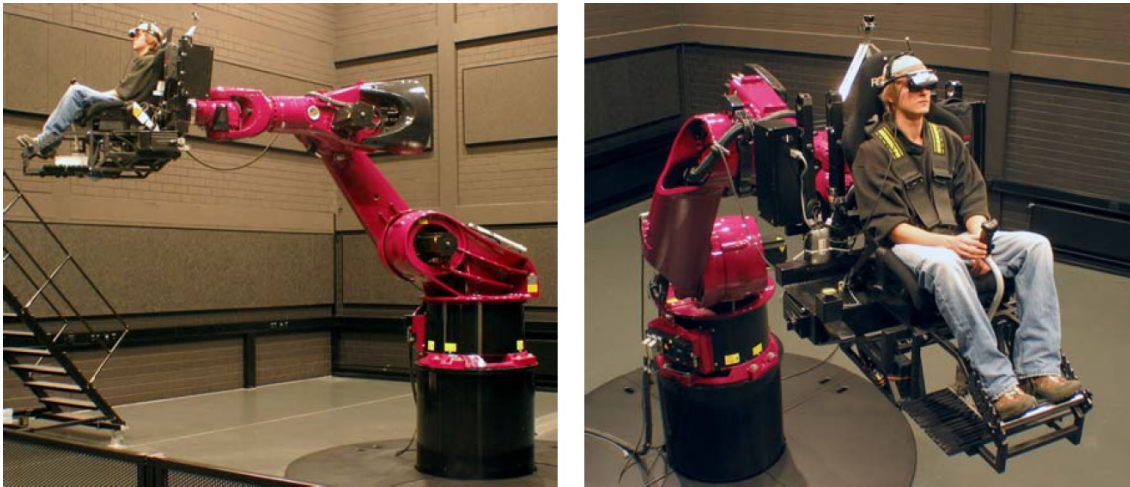


Figure 1.5: Max Planck Institute motion simulator (Beykirch et al., 2007).

Nieuwenhuizen et al. (2007) undertook a similar study to examine the effect of roll and lateral motion in a helicopter target-following and disturbance rejection task. Roll and lateral motion provided a significant improvement over the baseline condition without motion. The combined motion condition also delivered better performance than each individual motion cue. Examining control activity by means of the RMS of the pilot joystick input signal indicated that roll angle is the primary cue used for control but the addition of lateral motion can additionally reduce control effort while improving performance.

Karimi and Mann (2008) examined the applicability of motion feedback to a specific task in driving a tractor in a straight line. Experiments were conducted in a simulator with a motion and no-motion configuration. Results suggest that yaw is a useful motion feedback cue and reduced lateral error and supported a more relaxed driving style.

Researchers have found it useful to describe the relevance of motion cues to different tasks using Rasmussen's skill, rule, knowledge taxonomy (Rasmussen, 1983). It is argued (Bowen et al., 2006) that motion cues in vehicle operation are most relevant to executing skill-based behaviour (SBB). SBB relies on a 'very flexible and efficient dynamic internal world model' (Rasmussen, 1983) which has been built up over time through experience. This model provides subconscious awareness of how systems will respond to certain control actions and how the environment can be expected to change accordingly. To establish this model, it has been suggested (Bowen et al., 2006) that effective training of SBB is dependent on learning skills in realistic conditions. Conversely, it is a reasonable hypothesis that SBB can be leveraged in a teleoperation system that provides cues consistent with the

internal world model.

Despite extensive research over many decades, there still remains uncertainty regarding the precise behaviour and characteristics of the human vestibular system, particularly as excitation dynamics grow in complexity and other factors impose mental loading. The difficulty in accurately modeling the human has many facets. As Hosman and van der Vaart (1981) comment, investigation is difficult due to the fact that there is redundancy in sensory information as proven by the fact that sensory defective persons can learn to compensate with other sensors. Further complicating any assessment of the interaction between sensory information is the fact that there are neural processes that are responsible for resolving and fusing this information. For example, Merfeld et al. (1999) performed a study involving providing subjects with a false sense of linear acceleration by rotating and then tilting subjects. It was found that the eyes moved as would be expected if a true linear acceleration existed indicating that an internal model had been used to process the ambiguous sensory information. The fact that the vestibular system and its interaction with other sensory systems is not completely understood makes it difficult to identify requirements around motion feedback for a given task and vehicle.

Across the literature, there is general agreement that different tasks place different requirements on levels of fidelity and range of motion cues. Complicating any investigation, however, is the fact that different people potentially use cues in different ways. Findings related to how humans use internal models to resolve ambiguous information and improve their perception accuracy would appear to suggest that motion feedback applied to assist teleoperation should target maximum realism. Related to this also is the suggestion that the internal dynamic model is a key aspect of utilising SBB. However, the findings of Ringland et al. (1971), that a subset of motion cues provided higher performance than the full complement of motion cues, are at odds with the general hypothesis that situational realism should enable better utilisation of SBB. Even though this study was performed with a simulator and idealised visual cues, this result does provoke consideration in the context of a teleoperation system of whether cues that are experienced on board may actually have a negative influence on performance.

The application of motion feedback to assist specifically with teleoperation is a relatively new field of investigation. Ortiz et al. (2008) applied motion feedback to assist with operation of an Unmanned Ground Vehicle (UGV). Task completion time was not improved by motion feedback for a 'follow the path' exercise that included negotiation of a narrow

1.4 Perception for teleoperation

tabletop type ramp. Giordano et al. (2010) examined the potential for scaled feedback of the gravity vector to assist with a UAV movement and hovering task. Results based on task completion time indicated that motion feedback of this type provided no improvement against the task performed with visual information only. However, control effort as measured by the average magnitude of control input over all trials was lower with motion feedback. While the influence of motion feedback on performance in these circumstance permits an interesting assessment of teleoperator performance, it is fundamentally different from the objective of supporting teleoperation of equipment for which operators already have skill based behaviour learnt through on board operational experience. It is important to note that these experiments were performed with vehicles that cannot be operated on board. Consequently, it might be suggested that subjects responded to motion cues based on instinct and some learnt behaviour rather than according to true operational experience.

In the context of this study, it would appear that although accurate motion cues should be targeted, their benefit can only be assessed through experiments. The ability to modify motion feedback cues independently would thus appear to be a key one.

1.4.6 Assessing perception requirements

As the previous sections reveal, there is a well developed field of research around perception as well as its impact on teleoperation. An examination of the literature indicates that there is still significant unresolved understanding regarding what factors are critical to teleoperation and how perception characteristics might best be defined to improve performance. These gaps in understanding are both a result of the complexity and difficulty associated with examining the characteristics of human sensory systems and also the lack of studies performed on diverse vehicle types. The complex relationship between perception and performance makes it difficult to predict in a non-trivial and accurate way how altering the perception characteristics of a teleoperation system may influence task performance.

A review of the broad applications of teleoperation reveals that the factors that are critical to perception and performance vary depending on the machine under control and the nature of the task. This is perhaps intuitive but does not necessarily simplify the process of determining which factors are important for a given application and task without experimentation. Nevertheless, there are some clearly consistent findings across multiple studies. Time delay is a major determinant of performance, with minimum time delay de-

sirable. Different sensory systems can dominate at different stages in a task. It is possible to substitute information that would traditionally be obtained through one sensory system by exciting another sensory system in an intuitive way. While knowledge of these factors can assist in the design of a novel teleoperation systems it is insufficient to describe a generic implementation that will achieve optimal performance. Lathan and Tracey (2002) caution that ‘because of the differences between individual operators, designing an interface between the human operator and a teleoperated robot presents a unique challenge’. They go on to suggest that there may be a requirement to customise control and feedback interfaces to specific user needs. While this may be true, it is clearly desirable to be able to determine a base perception configuration that is useful to most operators.

Weir and McRuer (1968) posited that there are two alternate approaches to determining the closed-loop structure of a driver-vehicle system: from a perception basis and from a guidance and control theory basis. The issue with the perception approach is that it can be an involved and difficult exercise. The issue with the formulation as a guidance and control problem is that there is no unique solution. Weir and McRuer (1968) suggest that the best approach is to marry the two.

The methodology to be followed in this study is predominantly aligned to the perception approach. The control side of the teleoperation system is unaltered. This methodology takes significant inspiration from the approach of Ross et al. (2008). Ross et al. (2008) developed a high-fidelity teleoperation system for an unmanned ground vehicle to permit the independent variation of latency, resolution, field of view, frame-rate and motion. The high-fidelity teleoperation system provided an experimental apparatus that permitted broad testing of multiple perception factors on teleoperator performance. The experimental task involved the operator navigating the vehicle over unstructured terrain to reach a goal. Experiments revealed latency to be highly detrimental to performance and user acceptance, with latencies greater than 400-500 ms considered particularly severe. In a portion of testing where operators could select their own perception characteristics within certain bandwidth constraints, it was noticed that during regular driving, high frame-rate, wide field of view and medium resolution was preferred while during navigation of complex environments and route planning, there was a preference for high resolution, compromising on field of view and frame-rate. Motion and audio cues were found to support sensation of realism and presence.

The apparatus developed by Ross et al. (2008) permitted highly targeted testing of the

1.5 Thesis overview

role of perception. Using this methodology, the component of the system responsible for the provision of perception is capable of being modified independently, i.e. there is full control over the ‘filtering’ characteristics of the perception system. It is possible in principle to examine factors which influence teleoperation performance for any task, any remote control implementation and even any controlled element.

Experimentally, this study is believed to be novel in the sense that there is no simulation and the application of teleoperation uses a vehicle for which existing operators are expected to have extensive on board operating experience. This is in contrast to previous studies where the influence of different factors on performance have been examined by using a simulator or the vehicle being teleoperated is not designed for on board operation. While simulators greatly simplify experimentation, it is difficult to capture all of the disturbances that are introduced in a live system. Furthermore, as noted by Nagai et al. (2002), the performance of an operator can be significantly influenced by the urgency and perceived consequences of a real task as compared with a simulated task.

The focus on a vehicle that can be operated on board rather than a UAV or UGV allows for the exploitation of operators with operational experience and learnt SBB. It is believed that this will provide for a more direct evaluation of the influence of introducing a teleoperation system and how the changes it brings to the human-vehicle dynamics might be minimised.

1.5 Thesis overview

This thesis seeks to present a structured investigation of the perception requirements for effective bulldozer teleoperation. This chapter has highlighted the motivation for achieving high performance teleoperation of bulldozers based on safety and productivity concerns. Sheridan’s concept of perception ‘filtering’ in teleoperation was discussed and introduced as the guiding theoretical context for this study.

In Chapter 2, the aims and methodology of the study are introduced. A two part procedure is proposed involving experimental evaluation and task analysis.

Chapter 3 presents the experimental evaluation of the influence of different perception cues on the performance of the bulldozer teleoperator. The notion of performance is given resolution by considering multiple aspects including: task completion time, task progress and workload.

Chapter 4 presents a task analysis informed by the experimental results. The task analysis incorporates findings from the experimental evaluation, additional qualitative feedback obtained from operators and task observation.

An attempt is made in Chapter 5 to consolidate the results presented in Chapters 3 and 4. The influence of the various perception cues is discussed and ranked according to their influence on bulldozer teleoperation performance.

Chapter 6 concludes the thesis by summarising the contributions made and proposed areas of further investigation.

Examining perception requirements for teleoperation

2.1 Aim

The literature reviewed in Chapter 1 reveals that there is no clear generic method for determining perception requirements of a given mechanism under teleoperation. Nevertheless, the theoretical concept of the teleoperation system acting as a filtering element in the control loop of the operator and the equipment under control provides a sound context for experimental evaluation.

This thesis aims to determine: *What feedback attributes are critical to maximising bulldozer teleoperation performance and user acceptance?* To answer this question through experimental evaluation, there are five key steps:

1. *Evaluate the degree of performance degradation associated with bulldozer teleoperation compared to on board operation.*
2. *Evaluate the potential for motion feedback to improve performance and user acceptance in bulldozer teleoperation.*
3. *Evaluate the potential for enhanced visual and audio quality to improve performance and user acceptance in bulldozer teleoperation.*
4. *Evaluate the potential for task focused visualisation to improve performance and user acceptance in bulldozer teleoperation.*

5. *Determine when the various feedback modalities are most relevant to particular bulldozer control activities.*

Execution of the above steps provides a robust appreciation of the influence of feedback cues on bulldozer teleoperation in both a specific and general sense. Achieving the aim of this thesis makes a contribution to the state-of-the-art of bulldozer teleoperation systems by providing key knowledge necessary for effective perception system design. It is hoped that a secondary contribution of this thesis is in the advancement of methodologies for identifying and determining perception cues critical to teleoperation systems more generally.

2.2 Methodology

This thesis employs an experimental approach to test the hypothesis that enhanced perception can improve performance and user acceptance. The theoretical context for this approach involves framing the bulldozer teleoperation system as a filtering element in the feedback pathway.

The experimental approach relies on the specification of baseline perception characteristics which will provide the reference point against which the influence of perception enhancements can be evaluated. A logical baseline is provided by the default perception configuration of an existing non-line-of-sight bulldozer teleoperation system.

The baseline teleoperation system is subsequently modified to allow its perception characteristics to be altered. This aspect of the methodology takes strong influences from the approach employed by Ross et al. (2008).

The modification of the baseline system is achieved through the integration of an enhanced perception cell: a collection of additional hardware that provides experimental control over the perception characteristics but leaves the baseline perception configuration and the control mechanism of the system unaltered. This allows experimentation to solely concentrate on the influence of perception feedback and maintains a meaningful baseline reference configuration.

The enhanced perception cell provides the basis for experiments that are designed to isolate the influence of enhanced perception generally as well as the distinct influence of specific feedback cues.

2.2 Methodology

The key steps of the methodology are summarised below.

1. ***Develop and integrate the enhanced perception cell.***

The enhanced perception cell supports experimental evaluation by allowing for the simple addition or removal of specific feedback paths that are to be investigated. The feedback cues provided by the enhanced perception cell are selected based on their relevance to the tasks associated with bulldozer operation. The fidelity and timeliness of these cues can be modified independently, thus permitting controlled adjustment of the ‘filtering’ characteristics of the teleoperation system.

The enhanced perception cell is integrated with a conventional bulldozer teleoperation system. The perception characteristics of this conventional teleoperation system will provide the baseline against which perception enhancements are evaluated.

2. ***Establish the experimental procedure.***

A typical bulldozer operational task is selected as the basis for experimental evaluation using the enhanced perception cell. The task should be representative of the main elements of bulldozer operation.

A secondary consideration is the requirement that the task be sufficiently structured so as to minimise inconsistencies in performance and variability in results.

3. ***Execute the experimental plan.***

The enhanced perception cell provides the means to independently vary the perception characteristics of the teleoperation system. Experiments are conducted with the cell to evaluate the effect of different perception characteristics on task performance and user acceptance.

Multiple metrics are logged during the execution of experiments to enable a detailed analysis of performance and operator behaviour. In addition, qualitative feedback is obtained from the operator around perceived performance and workload.

4. ***Isolate influence of various perception cues.***

The results obtained from the experiments are analysed to first prove or disprove the hypothesis that enhanced perception can improve performance in teleoperation. Second, the relative contribution of individual factors on performance and user acceptance will be identified.

5. *Associate task aspects with perception requirements.*

The experimental results are consolidated with qualitative feedback and task analysis to associate perception requirements with the task elements and control aspects of operating a bulldozer.

The methodology summarised above is devised such that there is a degree of separation between the approach and the specific system under investigation. The intent is to not only facilitate investigation of the perception requirements for bulldozer teleoperation but also to validate the methodology as a more general approach to teleoperation perception requirements assessment.

2.3 Perception cues relevant to bulldozer operation

The on board operation of a bulldozer excites an extensive and varied set of sensory inputs. Bulldozers are a class of ground vehicle, however, there are several factors that make the act of operating a bulldozer fundamentally distinct from other ground vehicles such as road vehicles.

Mining class bulldozers are characterised by poor forward visibility: a consequence of the size of the engine housing and associated stacks as well as the size of the blade and its positioning. A typical field of view for an operator seated on board a mining class bulldozer is illustrated in Figure 2.1. The areas of significant restricted visibility on a Caterpillar D8T are shown in Figure 2.2. As Figure 2.1 shows, there is limited visibility directly in front of the bulldozer and no visibility in the region immediately in front of the blade.

A second characteristic of bulldozers is their use on steep and uneven terrain. Bulldozers frequently work on steep gradients and uneven contours which means that operators are accustomed to pronounced variation in orientation with respect to the gravity vector.

A third characteristic of mining class bulldozers is their typical use on rocky terrain. The ground conditions and lack of suspension exposes operators to a high level of vibration and jarring.

Bulldozers are commonly used to alter terrain geometry by shifting material around. The execution of a desired terrain transformation places unique demands on the operator in maintaining both a three dimensional mental map of the current and desired terrain as well

2.3 Perception cues relevant to bulldozer operation



Figure 2.1: View from inside bulldozer cab.

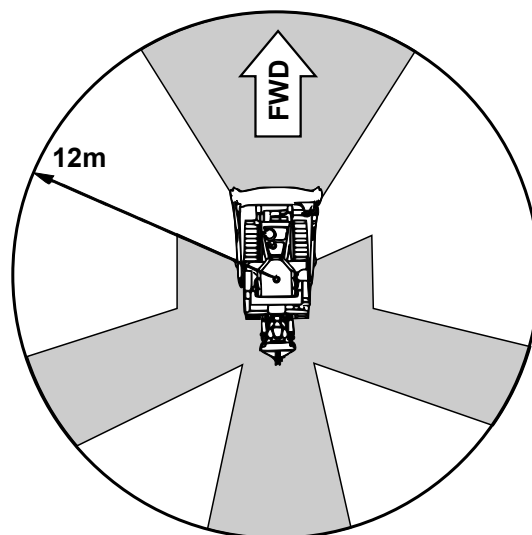


Figure 2.2: Shaded region indicates area of significantly restricted visibility on a Caterpillar D8T. (Caterpillar, 2009)

as their position and orientation within this reference frame. This requires a high degree of contextual awareness, loosely akin to the requirements on a pilot when skywriting.

The mining environment also introduces unique challenges around contour and object perception. Bulldozers in a mining context typically operate in expansive areas of either component or loosely disturbed dirt and/or rock. These conditions mean that there is very little contrast and terrain contours and individual terrain features can be very difficult to

perceive. However, the act of disturbing the ground can actually improve environment perception by exposing material with differing moisture content and hence colouration.

Bulldozer operation is also distinguished by the fact that the terrain may be actively altered as the bulldozer moves over it. There is thus a strong correlation between the control outputs of the operator and the motion response of the bulldozer.

In addition to this relationship between control decisions and the motion response, bulldozer operators are also exposed to the acoustic response of the drives which has a strong correlation with machine loading. The acoustic response can inform the operator of the need to make a control adjustment to modulate the loading on the machine. This is similar to making a gear or acceleration adjustment in a manual car based on the pitch of the engine whine.

These various characteristics of bulldozer operation reveal that multiple sensory systems play a role in achieving effective control. Removing the operator from the machine and placing them in a teleoperation system severs these direct feedback pathways. The teleoperation perception system must reconnect some or all of these pathways or otherwise meet feedback requirements to enable the operator to leverage their learnt skills when operating remotely.

Based on the unique characteristics of bulldozer operation, the cues that appear to be most relevant to effective operation are vision, motion, audio and contextual awareness.

2.4 Feedback factors for experimental evaluation

The cues theorised to be most relevant to on board bulldozer operation (vision, motion, audio and contextual awareness) are selected as the main feedback groupings for experimental evaluation. The specific factors that are explored within each of these perception groupings are justified below.

2.4.1 Vision

The key challenges in replicating the visual cues available to the operator on board are in providing adequate field of view, resolution, contrast, depth perception and latency. Complicating the provision of vision remotely are technology limitations that introduce certain constraints on the fidelity and timeliness of video streams over a wireless network.

2.4 Feedback factors for experimental evaluation

The limited forward visibility from within the cab forces operators to make use of a number of quite subtle visual cues. For example, due to the inability to directly observe how material is building up against the blade, operators look for signs of a fully loaded blade via the presence of material above the top edge of the blade. However, it can be difficult to distinguish the material appearing at the top edge of the blade from the background terrain. The amount of material observed flowing around the side of the blade also provides an indication of loading. The consistency and moisture content of the material can also be interpreted from the colouration and movement of material around the blade. Perception of this motion is reliant on the ability to distinguish between the flowing material and the stationary terrain.

The lack of contrast typical of the mining environment terrain is a challenge for the on board operator as well as for the teleoperator. A confounding factor is the typically poor performance of video capture in low contrast environments. Furthermore, the capture of a three dimensional terrain and features and its display on a two dimensional screen greatly reduces the ability of the teleoperator to perceive contours and geometry.

The subtlety of these various cues suggests potential value in high resolution video in the teleoperation station. Also, the potential benefits of depth perception in resolving these subtle cues also suggest value in capture and display in 3D.

2.4.2 Motion

A commonly held opinion among bulldozer operators is that the limited visibility available elevates the relevance and value of motion cues. The lack of forward visibility leads to the operator also relying on motion cues to detect variation in loading of the blade. For example, deep cutting will pull the bulldozer down. Uneven lateral cutting will pull one side of the bulldozer down. The operator's modulation of blade movements to achieve consistent fill and grade relies on these sensations.

Vision and motion both interact to support the operator's appreciation of terrain geometry. The roll and pitch orientation of the bulldozer as sensed at the operator seat provides an indication of the relative grade/slope. The motion of the bulldozer as it traverses over the terrain provides information that allows the operator to perceive the overall contour of the work area and update their internal model. This becomes a key component of contextual awareness, discussed later in this section.

2.4 Feedback factors for experimental evaluation

In traversing uneven terrain, experienced operators will employ special tactics to prevent rapid see-sawing and impacts on the bulldozer. At the top of a distinct rise, operators will typically command a direction change when they sense they have reached the tipping point of the bulldozer to help ease off the rise. A similar motion cue is provided in circumstance when the operator is pushing material over an edge. Change in balance of the bulldozer is sensed indicating the need to begin reversing.

Vibration is also constantly present throughout bulldozer operation. The vibration response of the bulldozer holds cues relating to machine loading as well as material response. A commonly encountered phenomenon in bulldozing is track slip. Track slip occurs when the bulldozer loses traction causing the tracks to dig into the ground instead of propelling the bulldozer forward. The occurrence of track slip is marked by a unique vibration response.

The vibration response of the bulldozer varies according to material conditions, providing information about expected diggability and material flow behaviour. For example, soft/wet material would dampen the vibrations felt in the cab, whereas rocky terrain would cause a greater magnitude of vibrations.

Translational acceleration is felt by the operator based on the direction and speed of travel. Slight deceleration while pushing/cutting may indicate bogging down and excessive loading of the blade. Slight acceleration while pushing/cutting may indicate undesired release of material. Such cues provide key feedback pathways for throttle and directional control.

This wide range of motion cues available to the operator on board indicates potential value in partial or full replication of bulldozer motion at the teleoperation station. The variety of motion cues would appear to demand the ability to provide a comprehensive range of motion replication for experimental evaluation. In addition, there is the necessary requirement to be able to isolate individual motion cues.

2.4.3 Audio

The sounds the operator hears on board the bulldozer provide corroborating evidence for other perceived cues. For example, the presence of track slip may be detected by sight and the vibratory response as well as by interpreting the sounds made by the tracks.

The loading on the engine and drive are also detectable through the audio signal. The pitch of the whine of the drive is proportional to the level of loading. For example, the pitch of

2.4 Feedback factors for experimental evaluation

the whine will be low when loading is high as more power is required. Conversely, the pitch of the whine will go up as the loading is released and less power is required.

The provision of audio is a suitable feedback factor for evaluation and comparatively simple to implement. Nevertheless, the various audio cues described are again quite subtle and require a relatively high quality audio stream to accurately replicate.

2.4.4 Contextual awareness

Many of the aforementioned cues are inputs to the operator's internal model/map of the work area. As discussed in Section 2.3, bulldozing tasks typically involve gross terrain transformation. This means that the work area is constantly changing and the ability of the operator to maintain an accurate internal map and orient and position themselves within this environment is paramount. Contextual awareness is also important for safety, particularly if the work area contains or is near to other pieces of equipment and plant.

The contextual awareness of the teleoperator is fundamentally challenged by the inability to stop and take stock of the environment by looking around in all directions. Alternative methods for providing contextual awareness are required. The proposed method for evaluation of the influence of contextual awareness is through the provision of a visualisation that displays terrain geometry and the pose of the bulldozer within a fixed reference frame.

2.4.5 Summary of feedback factors to be evaluated

The contextualisation of the teleoperation system as a 'filtering' element suggests that reducing the disparity between on board perception and remote perception may deliver improvements in terms of performance and user acceptance. The selection of feedback factors for experimental evaluation represents the pathways for which adjustment of the 'filtering' characteristics are to be explored.

Based on the perception cues hypothesised to be most relevant to bulldozer operation, the feedback factors selected for experimental evaluation are:

1. Visual resolution
2. Visual depth perception

2.4 Feedback factors for experimental evaluation

3. Presence of audio
4. Motion feedback incorporating gross orientation variation, vibration, and translational acceleration
5. Contextual awareness through visualisation of terrain geometry and bulldozer pose

The enhanced perception cell will allow the baseline perception configuration to be augmented in terms of these various factors. A key requirement for experimentation is to allow these cues to be tested with a degree of independence such that the individual effect of the various cues might be isolated.

Experimental evaluation of perception requirements

3.1 Aim

This chapter summarises the experimental evaluation of the influence on bulldozer teleoperation of the feedback factors identified in Section 2.4.5. The aim of these experiments was to determine what feedback attributes are critical to maximising performance and user acceptance.

The experimental approach is based on establishing control over the perception characteristics of the teleoperation system and examining the effect of varying these characteristics. A conventional non-line-of-sight teleoperation system was selected as a baseline perception configuration. Augmentations on this baseline perception configuration were examined through the application of the developed apparatus, termed the enhanced perception cell. The enhanced perception cell enables independent control over the targeted feedback factors.

3.2 The conventional teleoperation system

A Caterpillar D8T (see Figure 3.1a) was used as the remotely controlled bulldozer. The D8T was paired with a remote station configured for non-line-of-sight teleoperation.

The remote station provided visual feedback through a live stream of a combination of forward views and one rear view. An image of the typical operator interface at the remote

(a) Caterpillar D8T.



(b) Remote station interface.



Figure 3.1: Conventional teleoperation system.

station is presented in Figure 3.1b. The controls and layout of the cab are replicated with some slight differences in the remote station. The mechanism and characteristics of control over the bulldozer were unaltered for the experiments described. No operator assist technologies were utilised.

Under the default configuration, vision, audio and control data is exchanged between the bulldozer and remote station over an 802.11 link. The default audio-visual configuration for the bulldozer teleoperation system employs four NTSC cameras providing left, right, forward and reverse live video streams from the bulldozer. The camera sensor resolution is 512×582 pixels. Audio is captured in the cab and attached to one of the video streams. No motion feedback is provided.

The implementation relies on compression and transmission of all video feeds over the 802.11 link. Exacerbating the low capture resolution, the existing video quality suffers from compression and upscaling artefacts as well as latency (approximately 300 ms). The audio is similarly degraded.

The wireless link and on board processing of operator commands also results in additional latency in the response of implement and directional controls. This control latency was unaltered in all experiments conducted with the system.

3.3 The enhanced perception cell

The enhanced perception cell is the experimental apparatus used to augment the perception characteristics of the conventional non-line-of-sight teleoperation system. The enhanced perception cell consists of three distinct systems, each targeting specific feed-

3.3 The enhanced perception cell

back pathways. These are: (i) the audio-visual feedback system; (ii) the motion feedback system; and (iii) the task visualisation. Each of these elements is described in detail below.

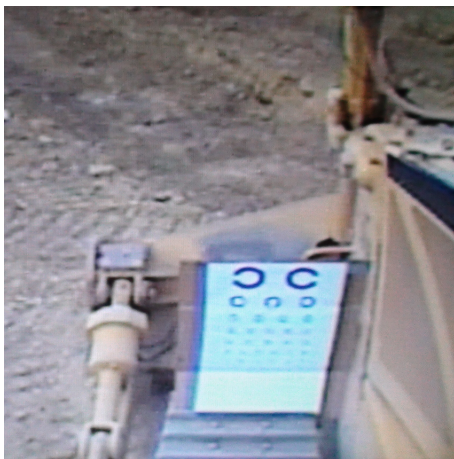
3.3.1 Audio-visual feedback system

The audio-visual feedback system allows the enhanced perception cell to augment the conventional teleoperation system through adjustment of visual resolution, depth perception and audio quality.

An independent audio-visual capture and transmission system was integrated in parallel with the default audio-visual path of the conventional teleoperation system. The audio-visual feedback system employs high definition cameras at the point of capture with output resolutions of 1920×1080 pixels. Video can also be captured in a stereoscopic mode through use of a paired camera arrangement.

Specialist broadcast hardware is used to transmit the high resolution camera feeds from the bulldozer to the remote station to reduce the need for compression. This maintains video quality and reduces capture to display latency to approximately 200 ms.

(a) Standard resolution.



(b) High resolution.



Figure 3.2: Cropped segment of output feed at remote station illustrating difference in video quality.

When video is captured in the stereoscopic mode, a 3D feed can be presented to the operator on a passive polarised 3D display. The standard forward left and forward right video feeds of the conventional teleoperation system can be substituted with either the 2D or 3D high resolution video feeds. Figure 3.2 provides a comparison of video quality between the standard and augmented arrangement. An example of the 3D video frame as

it appears prior to interlacing is shown in Figure 3.3.



Figure 3.3: Typical 3D frame prior to display interlacing.

The 3D camera pair is designed to provide depth information comparable to that which might be experienced by the on board operator. Figure 3.4 illustrates typical image disparity which the operator fuses to obtain depth information.

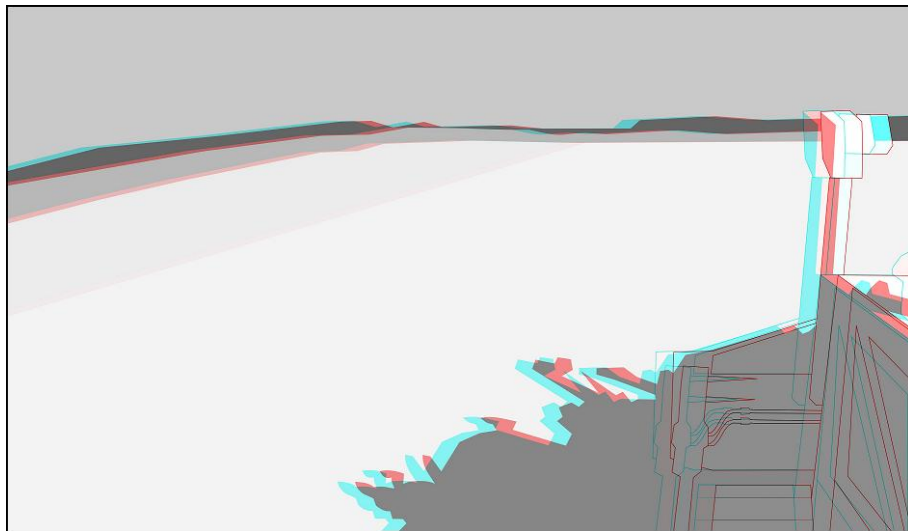


Figure 3.4: Anaglyph representation of the 3D view. Left and right component frames are interlaced and polarised in practice.

High quality stereo audio is captured in the bulldozer cab and embedded on one of the high resolution video streams. The operator is provided with this audio via stereo headphones at the remote station.

Summary of visual and aural cues

The audio-visual feedback system provides the capability to augment the default configuration to facilitate evaluation of the influence on performance and user acceptance of video resolution, depth perception and audio. The three independently applicable audio-visual cues are summarised in Table 3.1.

3.3 The enhanced perception cell

Table 3.1: Audio-visual feedback system cues.

Audio-Visual Cues	Description
High quality, low latency in-cab stereo audio	Stereo audio stream captured in the bulldozer cab and provided with low latency to the operator at the remote station.
High quality, low latency video	High definition video captured from the bulldozer and provided with low latency to the operator at the remote station.
High quality, low latency 3D video	High definition 3D video captured from the bulldozer and provided with low latency to the operator at the remote station.

3.3.2 Motion feedback system

The motion feedback system seeks to augment the standard remote station implementation by making realistic motion cues available to the teleoperator. To achieve this, the system must replicate the typical range of motion experienced by the operator when on board the bulldozer. Bulldozers are specifically designed to be able to traverse steep and rocky terrain. As such, the gross motion and vibration that must be capable of being replicated is extensive.

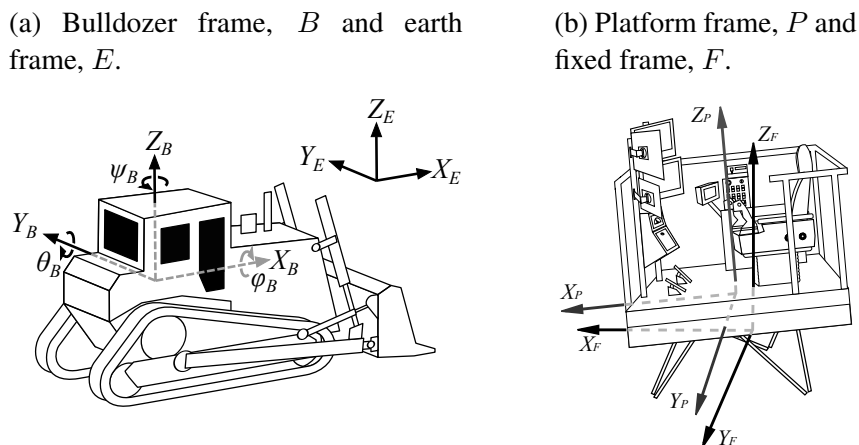


Figure 3.5: Bulldozer frame and Platform frame.

The conventional teleoperation interface was modified through the addition of a motion-

base. An inertial sensor suite installed on the bulldozer enables the motion-base and remote station to be slaved from the detected motion of the bulldozer. Motion data provides the location and orientation of the bulldozer frame, B in the earth frame, E as described in Figure 3.5a. Motion data was mapped with minor alteration to command motion of the motion-base frame, P as shown in Figure 3.5b.

Summary of motion cues

The dominant motions of the bulldozer were separated into individual cues that could be examined independently. This provides the capability to isolate the specific influence of different feedback cues. The design of the motion feedback system required that each of these cues could be replicated in isolation or in any combination. The separation of motion cues is summarised in Table 3.2. Each cue is described with reference to Figure 3.5a.

Table 3.2: Isolated motion cues.

Motion Cues	Description
Roll	Bulldozer roll, ϕ_B , represents rotational motion about axis, X_B .
Pitch	Bulldozer pitch, θ_B , represents rotational motion about axis, Y_B .
Yaw	Bulldozer yaw, ψ_B , represents rotational motion about axis, Z_B .
Translation	Translation represents motion sensations related to acceleration and deceleration along the X_B axis. Only translational effects along the X_B axis were considered given that this is the dominant direction of accelerations and decelerations in a bulldozer.
Vibration	Vibration represents accelerations in all axes appearing as frequencies greater than approximately 2Hz.

3.3.3 Task visualisation

The task visualisation depicts the location and orientation of the bulldozer within the confines of the test arena and with respect to fixed artificial boundary markers. The visualiser was updated based on current GPS and motion data. A typical screenshot of the task visualisation display is shown in Figure 3.6.

The display is composed of two primary views: a top-down view and a side profile view.

3.4 Experimental procedure

Terrain information is also depicted in each of the views: by a line representing the terrain surface in the side view; and by depth colouration in the top-down view.

The configuration shown in Figure 3.6 is customised for the slot bulldozing task, described in more detail in Section 3.4.2. The five rectangles that appear in the top-down view represent slot boundaries. Colouration of dark red through to yellow indicated cutting was required. Green colouration indicated target grade had been achieved within tolerance. Light blue through dark blue indicated overcutting below target depth.

The task visualisation display could be provided to the operator through the screen array on the remote station. It can be seen as the bottom right screen in Figure 3.1b. Note that during experiments the task visualisation was used as part of the default setup, as it provided feedback necessary for the operator to actually complete the structured task.

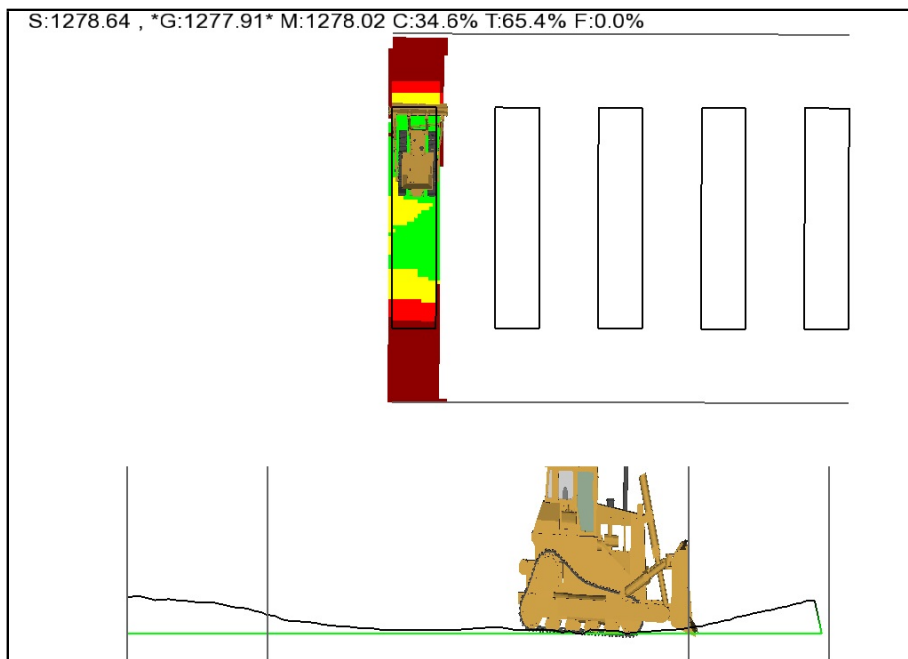


Figure 3.6: Task visualisation display.

3.4 Experimental procedure

The enhanced perception cell described permits independent testing of the influence of a range of perception cues. Experiments were conducted to explore the influence that the augmentation of teleoperator perception characteristics has on performance and user acceptance. The focus is on assessing the relative contribution made by the different per-

ception cues with reference to the conventional remote station perception configuration.

3.4.1 Design factors

The independent cues that can be applied to augment the perception characteristics of the default system are summarised in Table 3.3.

Table 3.3: Augmentations that can be applied to the existing remote station.

Default remote station can be augmented with:
AUG1. 2D high definition video stream
AUG2. 3D high definition video stream
AUG3. High quality stereo in-cab audio
AUG4. Roll motion feedback
AUG5. Pitch motion feedback
AUG6. Yaw motion feedback
AUG7. Vibration feedback
AUG8. Acceleration / deceleration motion cues (translation)
AUG9. Task visualisation

3.4.2 Structured task

An idealised slot bulldozing task was defined as a basis for factor investigation. The task required the removal of a volume of material over a designated area to achieve a prescribed elevation reduction. Figure 3.7 illustrates the experimental arrangement.

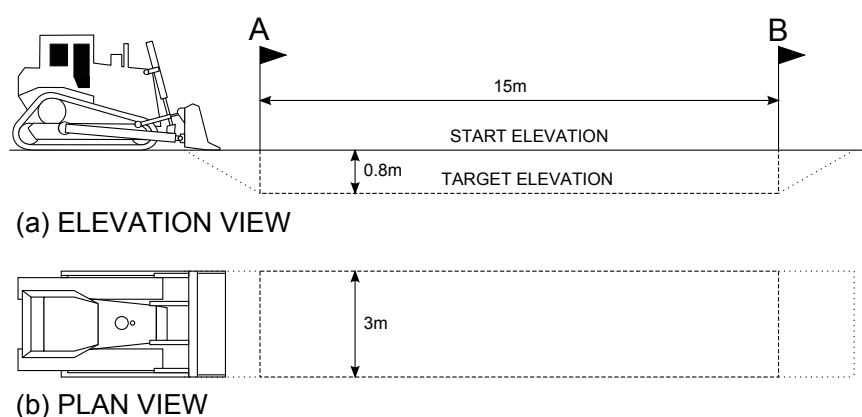


Figure 3.7: Structured test task.

The task requires repeated pushes over an initially flat surface 15 m long and 3 m wide.

3.4 Experimental procedure

Table 3.4: Nuisance factors.

Factor	Details
Operator skill level	<ul style="list-style-type: none">• The skill level of individual operators may differ and this may impact performance• There may be learning effects associated with the varied perception characteristics and subsequent configuration changes
Environmental conditions	<ul style="list-style-type: none">• Rain, dust or poor lighting may affect visibility and reduce performance
Material conditions	<ul style="list-style-type: none">• Conditions of the material moved as part of the experiment may differ depending on recent rainfall or ambient temperature and impact performance
Remote control fidelity	<ul style="list-style-type: none">• The performance of the remote control system may not be consistent across all tests and may impact performance
Bulldozer performance	<ul style="list-style-type: none">• The performance of the bulldozer may vary across the tests as a consequence of component failure or other equipment degradation

Pushes are repeated until the surface is lowered by 0.8 m and flatness at this new elevation is again achieved. The bulldozer always faces in the direction from A to B (see Figure 3.7) such that on pushing past B, the operator reverses the bulldozer back towards A before beginning another push. The operator is permitted to begin cutting into the earth before A and likewise continue cutting past B to assist with obtaining smooth transitions into and out of the slot. However, only the surface between A and B is considered when determining task completion. The geometry of the task is such that a range of bulldozer motion is excited and non-trivial control decisions are required.

3.4.3 Nuisance factors

In addition to the design factors introduced above, there are a number of nuisance factors that have the potential to introduce noise to the experiments and are controlled where possible. The most significant nuisance factors are summarised in Table 3.4.

3.4.4 Experimental Controls

During experimentation, effort was taken to maintain consistent material conditions. Variation in moisture content influences the traction of the bulldozer as well as the degree to which material will bind and roll when cut. Consistent moisture content is also important in terms of minimising dust that would in turn obstruct vision.

Pre-experimental trials also highlighted a significant learning effect associated with operating the teleoperation system and conducting the structured task. Effort was taken to remove this effect from results by conducting repeated trial runs prior to data collection until performance appeared to reach asymptotic levels.

Pre-experimental testing of the motion feedback element of the system revealed several additional sources of variability. Operator strategy varies slightly from test-to-test, sometimes with a significant influence on overall performance. The control on this source of variability was to encourage operators to decide on an appropriate strategy and maintain this as consistently as possible across all testing.

Experimental runs were performed in blocks of five of the same feedback configuration with a short break in between each run. The decision to execute experiments in blocks was a consequence of a suspected detrimental effect resulting from context switching. It is conjectured that switching alternatively between the static and motion configurations prevents effective immersion. Where possible, blocks of experimental runs of the same configuration were conducted on separate days in an attempt to account for day-to-day variability. Allocation of experimental blocks over test days was randomised to the extent possible.

3.4.5 Performance metrics

The principle aim of the experimental plan is to detect any statistically significant differences between the operator's performance under the default remote station configuration and the alternative perception configurations. It is useful to introduce as much resolution as possible to the analysis of aspects of performance. Greater resolution is likely to provide useful information about the precise influence of different feedback cues.

3.4 Experimental procedure

Task completion time

Task completion time, the time it takes to complete the designated test task, is selected as the primary performance metric. Task completion time provides a summary reflection of the multiple factors which dictate how well and how quickly the operator can complete the task. For the task described, completion time is calculated based on determining the time at which 90% of the terrain (discretised into a $0.25\text{m} \times 0.25\text{m}$ grid) is first brought within acceptable tolerance ($\pm 100\text{ mm}$) of the target elevation.

Task progress

In addition to examining the overview metric of task completion time, it is also useful to examine operator performance over the duration of the task. Task progress may be analysed quantitatively by examining volume removed over time. In the context of typical bulldozer operation, the rate of volume removal represents productivity. Productivity provides an indication of what capacity of full workload the bulldozer is operating at. For the task specified, productivity and task completion time are closely linked given that the amount of volume is consistent. However, productivity may change throughout the task and so it is useful to introduce some temporal aspect to the analysis of productivity.

Workload

Operator workload represents an important factor when considering how other performance metrics might extrapolate to extended durations of operation. Perceived and actual workload may influence the operator's fatigue and level of engagement. High workload is undesirable when work tasks must be performed constantly and over long durations.

Frequency of joystick corrections is selected as a means of capturing the operator's actual workload. Perceived workload is assessed through the NASA Task Load Index framework.

The NASA Task Load Index

In order to assess the impact of the different feedback configurations on operator workload, a NASA Task Load Index survey was conducted after completing each test. The operator was not made aware of their completion time before undertaking the survey.

The NASA Task Load Index (TLX) survey is a tool for assessing the workload perceived

Table 3.5: NASA TLX rating scale definitions.

Element	Description
Mental Demand	<i>How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?</i>
Physical Demand	<i>How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?</i>
Temporal Demand	<i>How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic? Performance How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?</i>
Effort	<i>How hard did you have to work (mentally and physically) to accomplish our level of performance?</i>
Frustration	<i>How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?</i>

by an operator in completing a given task. It was developed within the NASA Ames Research Centre and published in Hart and Staveland (1988) and has since been used in a wide range of applications.

The output of the NASA TLX survey is a perceived workload score, representing a weighted average of six underlying workload elements/subscales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration. The specific meaning of each of these workload elements in the context of the NASA TLX survey is summarised in Table 3.5.

The multidimensional nature of the survey allows the assessor to pick out the major contributing elements to the workload and also determine the sensitivity of these elements to a given change in system configuration.

3.4 Experimental procedure

The NASA TLX survey is separated into two parts. Part one of the procedure involves the subject rating the magnitude of each element experienced within the given task on a scale from 0 to 100. Part two of the procedure involves pairwise comparison between the elements, with the selected element in each pair deemed to be contributing more to the perceived workload by the subject. For the six elements, 15 possible comparisons are queried after which each element is assigned a weight. With the ratings and weights, the weighted average of the elements is calculated, giving the workload score. A lower score is indicative of a lighter perceived workload.

3.4.6 Experimental plan

The experimental plan was separated into two separate stages of data collection. This separation was a consequence of site availability, staffing constraints and continuing development of the enhanced perception cell. Unavoidably, a different operator was used as the subject in each stage. Both operators possessed extensive on board operational experience (>5 years) but limited line-of-sight and non-line-of-sight teleoperation experience. Note that the operator familiarisation, described in 3.4.3, was completed prior to conducting experiments.

In all cases, the unaugmented conventional non-line-of-sight bulldozer teleoperation system is considered the baseline against which the effect of perception enhancements is evaluated. However, the task visualisation is applied in all instances except where indicated. Tests were also run with the operator performing the task on board the bulldozer to provide a relative comparison.

The first stage focussed on motion feedback and task visualisation cues. Tests in stage one were executed by Operator A. The motion configurations examined include near-full replication of the motion the operator is accustomed to during on board operation, as well as separable motion cues provided in isolation. The isolated motion cues included: pitch motion only, vibration motion only and translational acceleration motion cues only.

The configurations where task visualisation was removed were designed to reveal how critical the task visualisation was to operator performance and whether the provision of motion feedback might supplement this. The eight configurations which constitute stage one are listed in Table 3.6. The specific augmentations applied in each configuration are listed as per Table 3.3.

Table 3.6: Data collection stage 1 - summary of test configurations.

Designation	Test Configurations	Augmentations
NMF	Remote station with no motion feedback (baseline in stage one of experiments)	AUG9
CMF	Remote station with combined motion feedback	AUG4, AUG5, AUG6, AUG7, AUG9
PMF	Remote station with pitch motion cues only	AUG5, AUG9
VMF	Remote station with vibration cues only	AUG7, AUG9
TMF	Remote station with translational acceleration cues only	AUG8, AUG9
NMF-TV	Remote station with no motion feedback and no visualiser	NONE
CMF-TV	Remote station with combined motion and no visualiser	AUG4, AUG5, AUG6, AUG7
OB	On board operation	AUG9

Stage two represents a combination of configurations examining aural and visual feedback as well as their combined effect with motion feedback. Tests in stage two were executed by Operator B. Visual feedback quality is examined at three levels: the standard vision arrangement, the high resolution 2D vision arrangement and the high resolution 3D vision arrangement. The influence of audio cues is examined by removing audio from the high resolution 2D vision configuration.

The configurations involving both visual and motion enhancements were selected to examine any interaction between motion feedback and enhanced audio-visual quality. The seven configurations which constitute stage two are listed in Table 3.7. The specific augmentations applied in each configuration are listed as per Table 3.3.

3.5 Results

The experimental results are presented and analysed with reference to the first four experimental objectives presented in Chapter 2. These are revisited below.

1. *Evaluate the degree of performance degradation associated with bulldozer teleoperation compared to on board operation.*

3.5 Results

Table 3.7: Data collection stage 2 - summary of test configurations.

Designation	Test Configurations	Augmentations
2DSR	Remote station in standard audio-visual configuration (baseline in stage two of experiments)	AUG9
2DHR	Remote station with high resolution 2D video and stereo audio	AUG1, AUG3, AUG9
3DHR	Remote station with high resolution 3D video and stereo audio	AUG2, AUG3, AUG9
2DHR-A	Remote station with high resolution 2D video and no stereo audio	AUG1, AUG9
3DHR+PMF	Remote station with high resolution 3D video, stereo audio and pitch motion cues	AUG1, AUG3, AUG5, AUG9
3DHR+FMF	Remote station with high resolution 3D video, stereo audio and full motion feedback	AUG1, AUG3, AUG4, AUG5, AUG6, AUG7, AUG8, AUG9
OB	On board operation	AUG9

2. *Evaluate the potential for motion feedback to improve performance and user acceptance in bulldozer teleoperation.*
3. *Evaluate the potential for enhanced visual and audio quality to improve performance and user acceptance in bulldozer teleoperation.*
4. *Evaluate the potential for task focused visualisation to improve performance and user acceptance in bulldozer teleoperation.*

The performance metrics introduced in 3.4.5 are evaluated in the following sections. Consistent sample sizes were sought but not always achieved due to site and equipment availability constraints. The results presented here represent the full set of valid data collected.

3.5.1 Operator performance is degraded under teleoperation

The performance degradation associated with the transition from on board operation to the conventional teleoperator is revealed in Figure 3.8. The plot shows the mean completion time which, as described in Section 3.4.5, represents the time taken to reduce the slot to the target depth of 0.8 m and establish grade tolerance (90% of the terrain as discretised into a $0.25\text{m} \times 0.25\text{m}$ rid is within ± 100 mm).

Table 3.8 presents the descriptive statistics for this result. The average completion time for Operator A in the conventional non-line-of-sight teleoperation configuration increased by 29.8% against their average on board performance. The degradation is less for operator B but still pronounced, with an increase in average completion time of 13.8%.

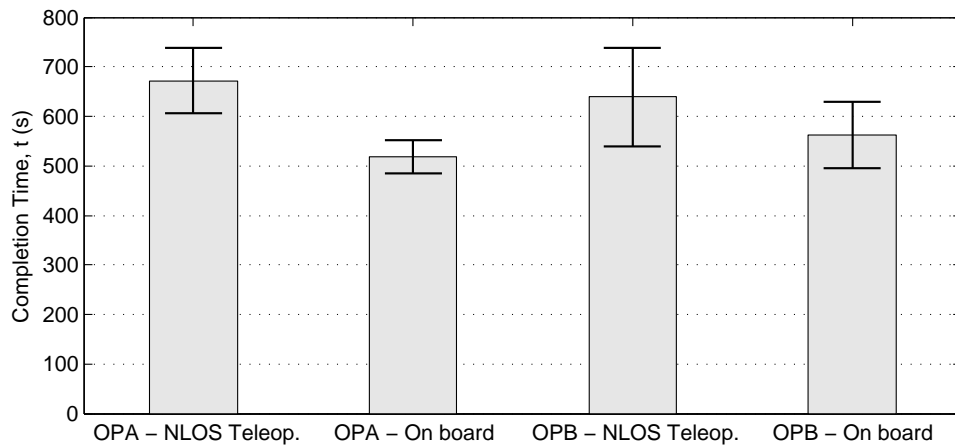


Figure 3.8: On board versus non-line-of-sight teleoperation. Error bars show standard deviation.

Table 3.8: On board versus non-line-of-sight teleoperation descriptive statistics.

Configurations	Completion Time (s)		Sample Size, n
	<i>mean</i>	<i>std</i>	
ES1 - NLOS Teleop.	672.2	65.2	15
ES1 - On board	517.8	33.1	4
ES2 - NLOS Teleop.	639.3	99.3	20
ES2 - On board	561.7	66.7	20

This result shows that degradation in performance is observable when the operators transition from on board operation to teleoperation under the conventional perception configuration.

3.5.2 Introducing motion feedback does not substantially influence task completion time

The influence of isolated and combined motion cues was examined through tests conducted in four distinct motion configurations. The baseline and the four alternate config-

3.5 Results

urations are described in Table 3.9. The specific augmentations applied in each configuration are listed as per Table 3.3.

Table 3.9: Motion configurations.

Designation	Test Configurations	Augmentations
NMF	Remote station with no motion feedback	AUG9
VMF	Remote station with vibration cues only	AUG7, AUG9
TMF	Remote station with translational acceleration cues only	AUG8, AUG9
PMF	Remote station with pitch motion cues only	AUG5, AUG9
CMF	Remote station with combined motion feedback	AUG4, AUG5, AUG6, AUG7, AUG9

Results from these experiments are compared to performance in the conventional non-line-of-sight teleoperation configuration (designated in this section as NMF). Figure 3.9 presents the average completion time in each configuration. Descriptive statistic are summarised in Table 3.10.

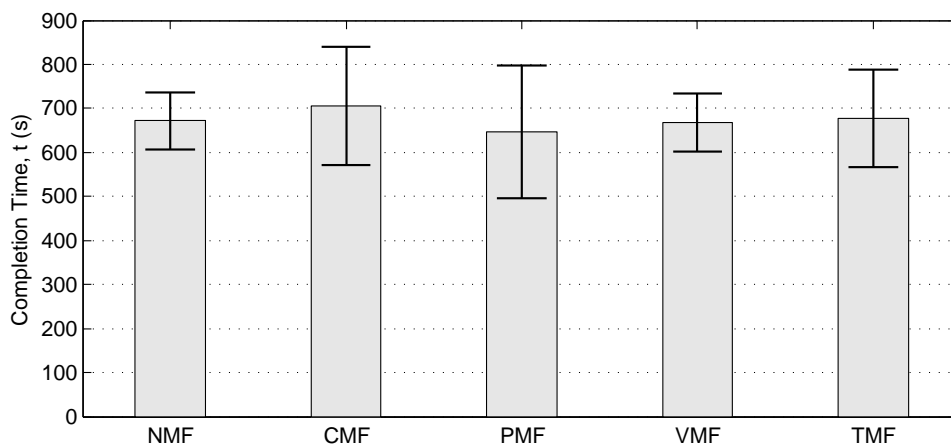


Figure 3.9: Completion under different motion feedback configurations. Error bars show standard deviation.

The fastest mean completion time among the motion configurations tested was achieved with pitch feedback only (PMF). Adding pitch feedback resulted in a mean completion

Table 3.10: Descriptive statistics for motion feedback configurations.

Configurations	Completion Time (s)		Sample Size, n
	<i>mean</i>	<i>std</i>	
No motion feedback, NMF	672.2	65.2	15
Vibration only motion feedback, VMF	668.4	65.3	9
Translation only motion feedback, TMF	677.4	111.4	10
Pitch only motion feedback, PMF	647.0	149.9	10
Combined motion feedback, CMF	706.0	134.3	15

time 3.7% faster than the conventional configuration (NMF). However, standard deviation is significantly elevated.

The vibration only and translation only configurations yielded average completion times within $\pm 1\%$ of the conventional configuration. The average completion time was slower with combined motion feedback (CMF). The standard deviation in the combined motion feedback configuration was also considerably higher than the conventional configuration which suggests reduced consistency in performance.

A four-way analysis of variance is applied using the factors and levels described in Table 3.11. No significant effect was detected with any factor. This result indicates that in terms of completion time, the presence of motion feedback does not significantly influence the operator's performance.

Table 3.11: Motion factors and levels.

Configurations	Vibration	Translation	Pitch	Other
No motion feedback, NMF	0	0	0	0
Vibration only motion feedback, VMF	1	0	0	0
Translation only motion feedback, TMF	0	1	0	0
Pitch only motion feedback, PMF	0	0	1	0
Combined motion feedback, CMF	1	0	1	1

Performance over task duration

The previous section has presented results aimed at examining the potential for the motion cues to improve performance in terms of task completion time. This section presents results that aim at revealing how the different feedback configurations influence the way the operator performs the task. This is achieved by examining operator activity over the duration of the task.

3.5 Results

Productivity, taken as the rate of volume removal, is a key performance metric for bulldozers. Progress based on volume removed over time provides an interesting point of examination of productivity throughout the different phases of the task.

Figure 3.10 plots the mean volume removed from the slot over time for all configurations test in stage one of data collection. The mean volume history for on board operation is also shown (designated OB). This plot is generated by averaging the volume removed from within the test slot boundary for each test in a configuration at each step along the time series. The time series includes only time spent moving forward within the boundaries of the slot, i.e. non-productive portions (such as reversing, pushing before or past the slot boundaries) of the test are removed. This is done to provide a direct comparison of the time spent performing productive work. The horizontal black line at $31.5m^3$ indicates the typical minimum volume removal required to complete the task within the tolerance requirements.

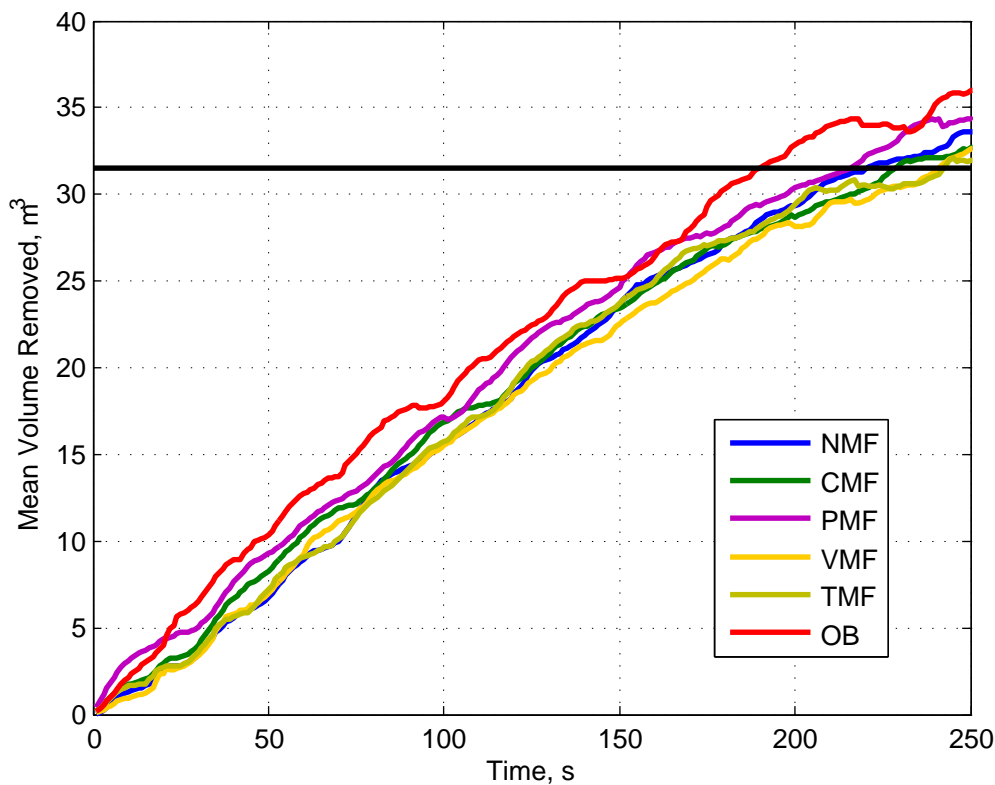


Figure 3.10: Motion cues - mean volume removed from within slot over time.

The progress trajectories followed by the no motion (NMF) and combined motion (CMF)

configurations show little difference. The rate of volume removal appears consistent for both configurations. This is consistent with the minimal difference observed in task completion time for these two configurations.

The vibration motion configuration (VMF) and the translational acceleration / deceleration cue configuration (TMF) display a similar trajectory to the default configuration (NMF) up until approximately 120s. At this point, the vibration motion configuration slows in volume removal terms while the translational acceleration / deceleration cue configuration (TMF) continues to closely match the static configuration (NMF). The pitch only configuration (PMF), however, leads all other configurations in volume removal terms throughout the duration of the time series shown. This appears to indicate on average higher productivity as represented by rate of volume removal.

The progress trajectory observed in on board operation (OB) leads all teleoperation configurations throughout the time series as might be expected. The isolated pitch motion feedback (PMF) configuration provides a trajectory that falls roughly between the conventional non-line-of-sight teleoperation configuration (NMF) and on board operation (OB).

An alternative representation of Figure 3.10 is provided by Figure 3.11. The slot bulldozing task is broken down into three phases based on volume: 0-10m³, 10-20m³ and 20-30m³. Figure 3.11 shows the average time taken to complete each volume phase.

Figure 3.11a shows that the combined motion (CMF) and the isolated pitch motion (PMF) configuration complete the first phase marginally faster than the conventional non-line-of-sight teleoperation configuration (NMF). There is limited distinction observable in the latter two phases of the task.

The above results appear to indicate some distinction in performance between the early stage and later stage of the task. Figure 3.12a presents the mean volume per push over pushes 1 to 5 in the tests. Figure 3.12b presents the mean volume per push for the remainder of the test.

In Figure 3.12a, the mean volume per push is distinctly elevated in the pitch motion configuration (PMF) compared with the conventional non-line-of-sight configuration (NMF). The average volume per push in the first five pushes is 19.6% higher in the pitch motion configuration (PMF) than in the conventional configuration. This compares with 14.6%

3.5 Results

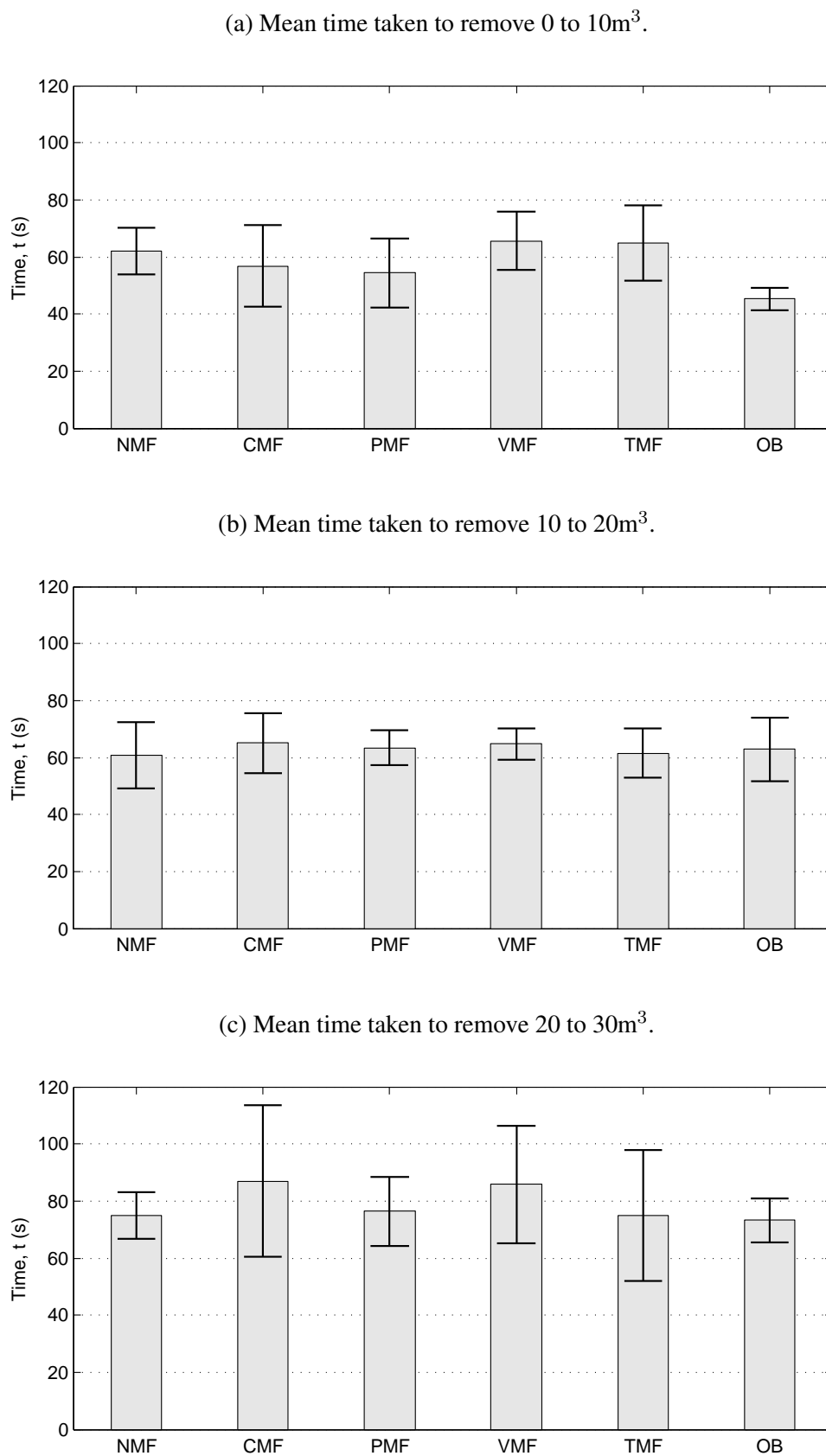
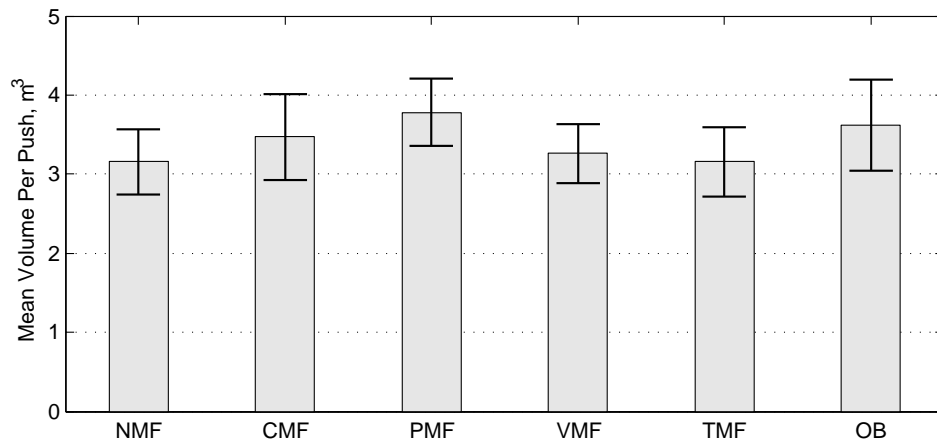


Figure 3.11: Motion cues - mean time taken to remove volume segment. Error bars of one standard deviation.

(a) Mean volume per push for pushes 1 to 5.



(b) Mean volume per push for pushes 6 to end of test.

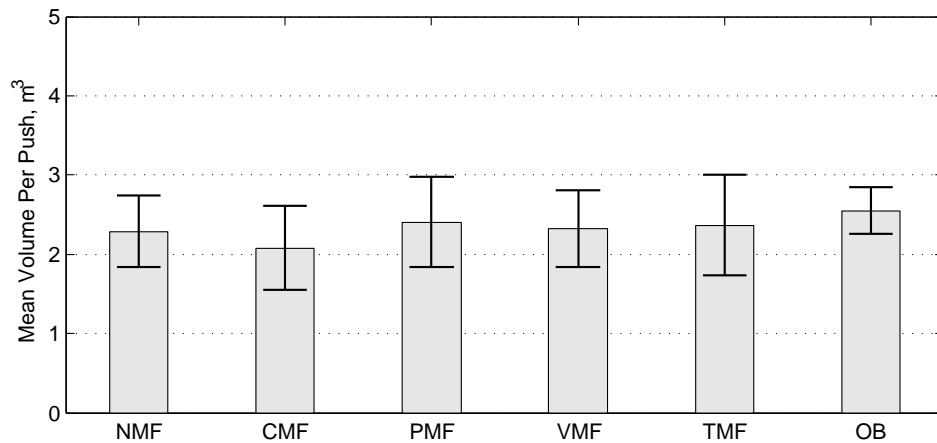


Figure 3.12: Motion cues - mean volume per push. Error bars of one standard deviation.

higher volume pushes in on board operation.

The later stage of the task reveals less distinction between configurations. The mean volume per push in the different configurations across the sixth to last push is between approximately 15 and 35% less than in the initial 5 pushes. This result can be interpreted in the context of the test by considering the different phases of the task. Initially, the task is focused on rapid volume removal but later transitions to a focus on correcting geometry defects and smoothing terrain.

3.5 Results

Table 3.12: Operator A - total workload score.

Configurations	Total Workload Score	
	<i>mean</i>	<i>std</i>
No motion feedback, NMF	36.8	5.4
Combined motion feedback, CMF	45.2	11.1
Pitch only motion feedback, PMF	37.9	9.1
Vibration only motion feedback, VMF	44.1	6.1
Translation only motion feedback, TMF	56.2	4.1

Influence of motion cues on perceived workload

A summary of the total workload results for Operator A is presented in Figure 3.13 and Table 3.12. The survey was not conducted for the first five of each of the combined motion (CMF) and no motion (NMF) configurations nor for the on board (OB) tests.

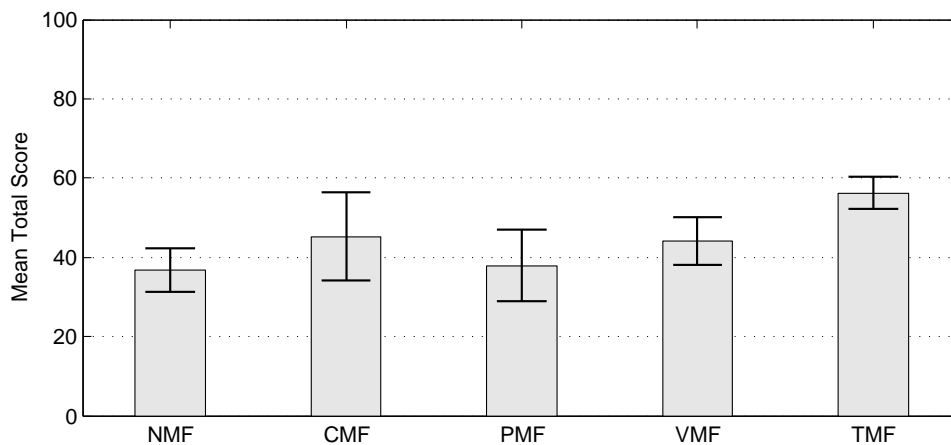


Figure 3.13: Operator A - mean workload total score. Error bars indicate one standard deviation.

The lowest average workload was observed in the default configuration (NMF = 36.8), marginally lower than the pitch only configuration (PMF = 37.1). The highest average perceived workload was encountered in the translation only configuration (TMF). Of all the configurations, only the pitch only configuration (PMF) had an average total workload score which was not statistically different from the the baseline configuration (NMF).

These results indicate that the introduction of motion feedback is generally associated with an increase in perceived workload for Operator A. However, the degree of workload

increase varies across the feedback configurations. Pitch only motion feedback (PMF) goes against this general trend with a perceived workload only marginally different from the baseline configuration. The total workload score can be divided into its individual workload elements as shown in Figure 3.14.

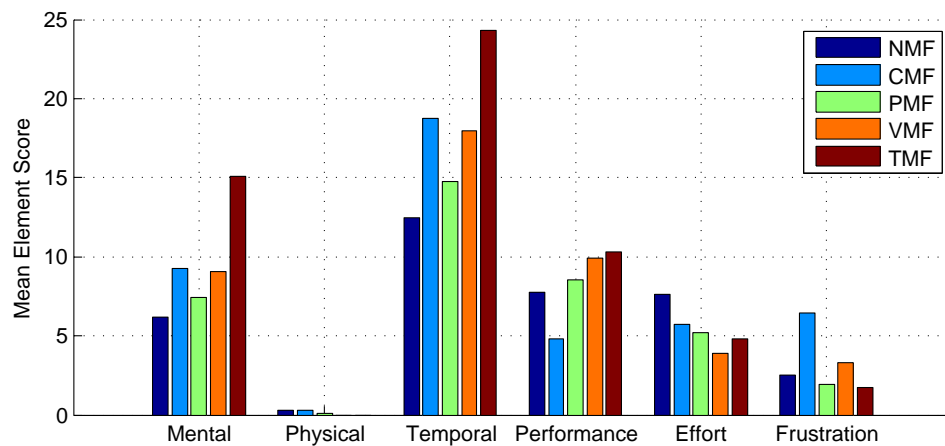


Figure 3.14: Operator A - mean workload element scores.

Mental Demand was observed to be distinctly higher when only translational motion feedback (TMF) was provided. It is difficult to interpret why the provision of only translational motion feedback (TMF) should yield such high Mental and Temporal Demand.

The major contribution to the elevated workload scores against the baseline (NMF) for the vibration motion (VMF) and combined motion (CMF) configurations was also in Mental Demand and Temporal Demand. Frustration was also elevated in the combined motion (CMF) configuration when compared to the baseline (NMF).

There were minimal differences between the pitch motion (PMF) configuration and the baseline (NMF) across the six workload elements. It would appear that the provision of pitch does not significantly impose on any of the aspects of workload while supporting improved performance as discussed earlier.

3.5.3 Depth of view dominates visual quality and audio

The results in this section focus on visual quality and audio feedback and correspond to stage two of the data collection. The influence of the visual configuration augmentations was examined through tests conducted with enhanced visual quality and the 3D vision im-

3.5 Results

plementation. The influence of audio was examined by replacing the high quality stereo headphones with earmuffs. The baseline and three alternate test configurations are described in Table 3.13. The specific augmentations applied in each configuration are listed as per Table 3.3.

Table 3.13: Audio-visual configurations.

Designation	Test Configurations	Augmentations
2DSR	Remote station in standard audio-visual configuration	AUG9
2DHR	Remote station with high resolution 2D video and stereo audio	AUG1, AUG3, AUG9
2DHR-A	Remote station with high resolution 2D video and no stereo audio	AUG1, AUG9
3DHR	Remote station with high resolution 3D video and stereo audio	AUG2, AUG3, AUG9

Figure 3.15 shows the average completion time in each configuration. Descriptive statistics are presented in Table 3.14.

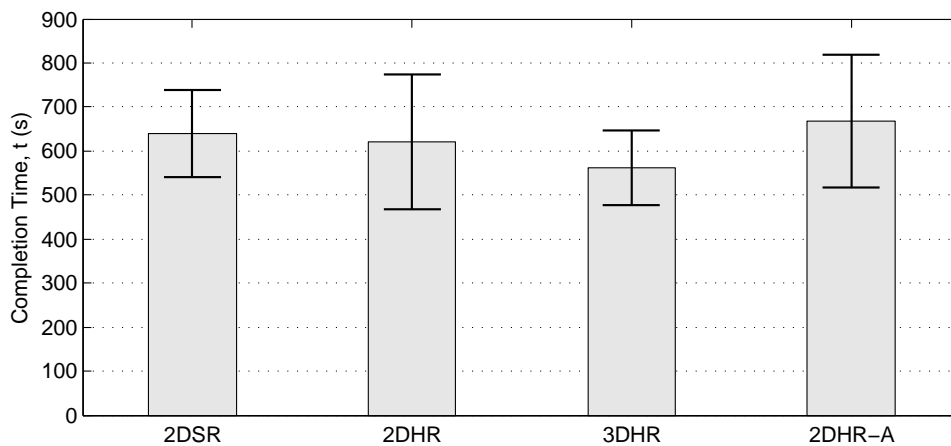


Figure 3.15: Completion under different visual and audio quality configurations. Error bars show standard deviation.

The fastest configuration was 3D high resolution video (3DHR), 12.1% faster than the mean completion time in the conventional non-line-of-sight teleoperation configuration (designated in this section as 2DSR). The 2D high resolution video configuration (2DHR) was on average 7.1% faster than the baseline but with comparatively high standard deviation.

Table 3.14: Descriptive statistics for visual and audio quality test configurations.

Configurations	Completion Time (s)		Sample Size, n
	<i>mean</i>	<i>std</i>	
Standard AV arrangement, 2DSR	639.3	99.2	20
2D high resolution AV, 2DHR	621.0	153.4	20
2D high resolution V no audio, 2DHR-A	668.4	151.1	15
3D high resolution AV, 3DHR	561.8	85.6	15

A two-way analysis of variance on visual quality and depth is performed with factors as shown in Table 3.15. This shows significance for depth ($p = 0.0459$) at the 5% level but a non-significant result for visual quality.

Table 3.15: Vision factors and levels.

Configurations	Vision quality	Depth
Standard AV arrangement, 2DSR	0	0
2D high resolution AV, 2DHR	1	0
2D high resolution V no audio, 2DHR-A	1	0
3D high resolution AV, 3DHR	1	1

The influence of audio quality is examined at 3 levels as shown in Table 3.16. The 2D high resolution configuration without audio (2DHR-A) is 4.6% slower than the standard AV configuration (2DSR). While the averages show an improvement in performance at each level of audio quality, the effect of audio quality is found to be not significant according to analysis of variance.

Table 3.16: Audio factors and levels.

Configurations	Audio quality
2D high resolution V no audio, 2DHR-A	0
Standard AV arrangement, 2DSR	1
2D high resolution AV, 2DHR	2

The combined effect of enhanced vision and motion

The combined effect of vision and motion is examined through two consolidated configurations. These two configurations are shown in Table 3.17. The specific augmentations applied in each configuration are listed as per Table 3.3.

3.5 Results

Table 3.17: Combined vision and motion configurations.

Designation	Test Configurations	Augmentations
3DHR+PMF	Remote station with high resolution 3D video, stereo audio and pitch motion cues	AUG1, AUG3, AUG5, AUG9
3DHR+FMF	Remote station with high resolution 3D video, stereo audio and full motion feedback	AUG1, AUG3, AUG4, AUG5, AUG6, AUG7, AUG8, AUG9

The mean completion time in these two configurations is plotted in Figure 3.16. The baseline configuration (2DSR) and the 3D high resolution video without motion configuration (3DHR) are also included for comparison in Figure 3.16. Descriptive statistics are presented in Table 3.18.

The 3D high resolution AV with full motion configuration (3DHR+FMF) provided the fastest mean completion time among all configurations, 12.8% faster than the mean completion time for the baseline configuration (2DSR). The 3D high resolution AV with pitch motion configuration (3DHR+PMF) provided a slower average completion time within 1% of the baseline configuration (2DSR).

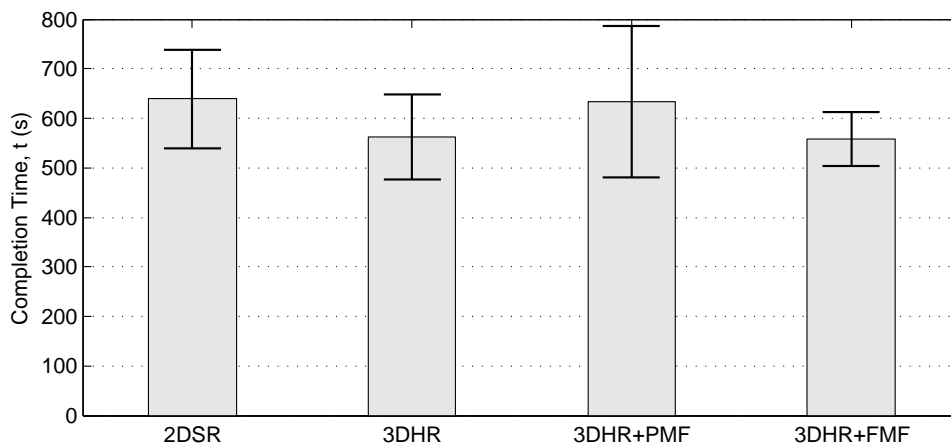


Figure 3.16: Completion with audio, visual and motion enhancements. Error bars show standard deviation.

The results are analysed based on the motion factors shown in Table 3.19. No significant effect is identified for either factor.

Table 3.18: Descriptive statistics for configurations with audio visual and motion enhancements.

Configurations	Completion Time (s)		Sample Size, n
	<i>mean</i>	<i>std</i>	
Standard AV arrangement, 2DSR	639.3	99.2	20
3D high resolution AV, 3DHR	561.8	85.6	15
3D with pitch motion, 3DHR+PMF	633.6	153.0	16
3D with full motion, 3DHR+FMF	557.3	54.7	9

Table 3.19: Vision and motion factors and levels.

Configurations	Pitch	Full
3D high resolution AV, 3DHR	0	0
3D with pitch motion, 3DHR+PMF	1	0
3D with full motion, 3DHR+FMF	1	1

The difference in completion time between 3D video with and without full motion feedback is marginal. On the basis of completion time, it would appear reasonable to suggest that combining visual and motion feedback can improve performance but that the specific effect of the additional motion cues appears insignificant compared with the enhanced visual cues.

Performance over task duration

Figure 3.17 plots the mean volume removed from the slot over time for each of the vision, audio and combined motion configurations of stage two of data collection. The conventional teleoperation configuration (2DSR), the high resolution 2D (2DHR) and high resolution 3D (3DHR) configurations show very similar trends up to approximately 150s. At this time, the two enhanced video configurations (2DHR and 3DHR) diverge from the baseline configuration. The divergence appears to indicate that the enhanced video configurations maintain a more consistent rate of volume removal for longer while the baseline configuration provides lower productivity. Considered within the context of the task performed, this is the point at which the task transitions to a focus on finer levelling of the ground at the target depth.

The high resolution 2D video configuration without audio (2DHR-A) follows a very similar trend to the baseline configuration (2DSR). It also appears to lag behind the other configurations at times up to approximately 150s. It does not exhibit the divergence seen

3.5 Results

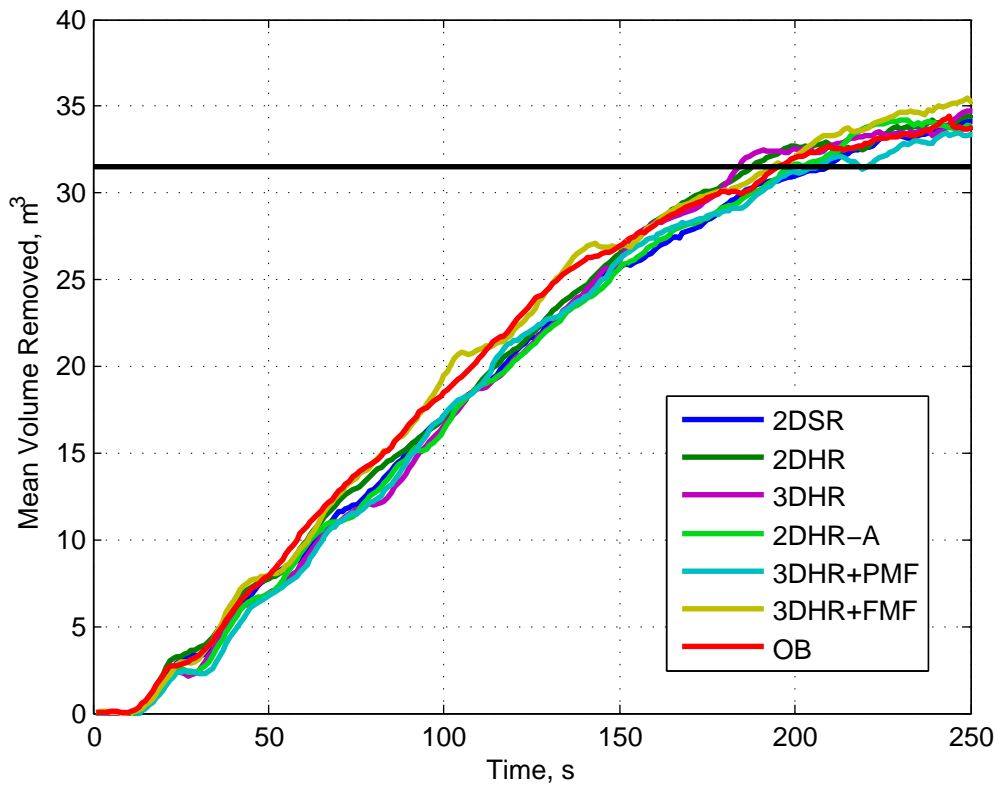


Figure 3.17: Combined vision, audio and motion cues - mean volume removed from within slot over time.

for the two high resolution video configurations with audio (2DHR and 3DHR) at 150 s.

The 3D video with full motion configuration (3DHR+FMF) leads all other configurations for the majority of the time series shown. This can be compared against the 3D video configuration without motion (3DHR) which only diverges from the baseline configuration (2DSR) at approximately 150s. By contrast, the configuration with 3D video and pitch motion only (3DHR+PMF) tracks fairly closely to the baseline configuration (2DSR).

These results suggest that motion feedback and in particular, full motion feedback, provides the most significant impact during the early to middle phase of the task which is focussed on removing material as quickly as possible.

It is interesting to note that the progress line for the 3D video configuration with full motion (3DHR+FMF) follows a very similar trend to on board operation (OB).

As introduced in the previous section, an alternative representation of the task progress plot is provided by the separation of the task into three volume segments as shown in Figure 3.18. Figure 3.18a and 3.18b show that the 3D video with full motion configuration (3DHR+FMF) marginally leads the other teleoperation configurations in phases 0-10m³ and 10-20m³. Figure 3.18c shows that in phase three, the conventional non-line-of-sight configuration is instead lead by the two enhanced video configurations (2DHR and 3DHR).

The apparent differences observed in volume removal over time are again examined by splitting the task and analysing mean volume per push. Figure 3.19a plots the mean volume per push over pushes one to five. Figure 3.19b plots the mean volume per push for the remainder of pushes required to complete the task.

In Figure 3.19a, the mean volume per push is distinctly elevated in the 3D high resolution video with full motion feedback configuration (3DHR+FMF). The average volume per push was 21.9% higher during the first five pushes in the 3D high resolution video with full motion feedback configuration compared with the conventional configuration (2DSR). This difference in mean performance for this combined configuration is significant at the 5% level based on a one-sided t-test comparing with the baseline. The mean volume per push is also elevated in on board operation and significant at the 5% level. No other differences with the baseline are statistically significant.

The later stage of the task again reveals less distinction between configurations. However, the mean push volume in the 3D visual configuration (3DHR) is elevated against the baseline (2DSR). This difference in mean performance for this enhanced video configuration is significant at the 5% level based on a one-sided t-test comparing with the baseline.

The results presented suggest that the presence of full motion feedback supports higher volume pushes during the early phase of the task. This benefit is not apparent as the task transitions to a different focus during the later stage of the task. At this time, the visual quality become a significant contributor to push volume.

Joystick input frequency - a proxy for operator workload

Joystick inputs of the operator were also recorded during execution of the tests described in this section. No data was recorded in the on board configuration. The principal control signals were recorded on the left and right joystick. In the arrangement used for testing,

3.5 Results

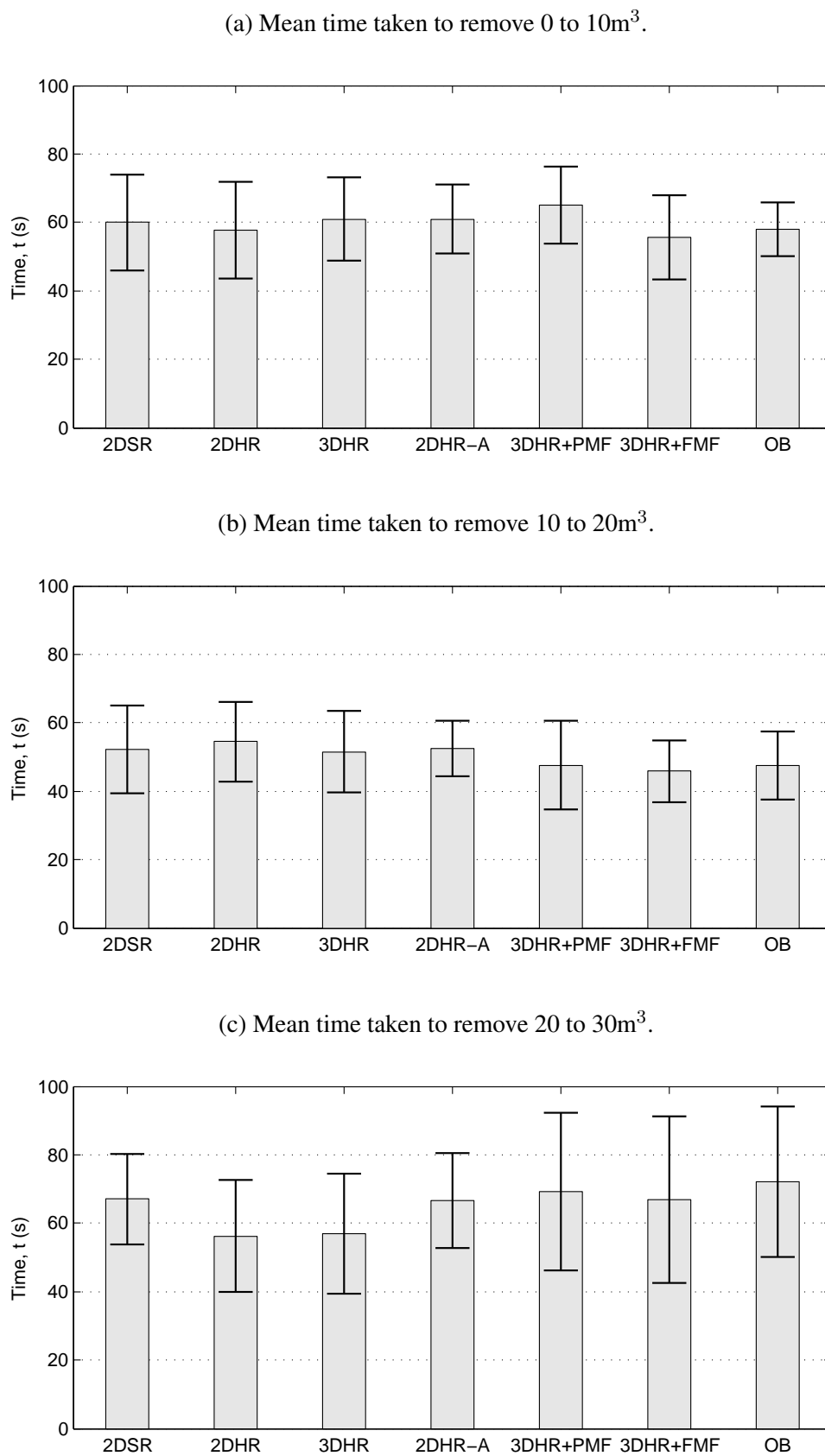
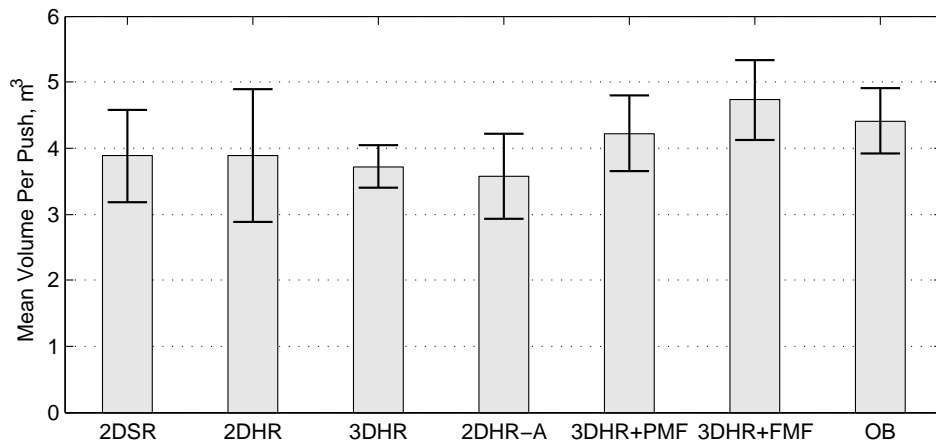


Figure 3.18: Combined vision, audio and motion cues - mean time taken to remove volume segment. Error bars of one standard deviation.

(a) Mean volume per push for pushes 1 to 5.



(b) Mean volume per push for pushes 6 to end of test.

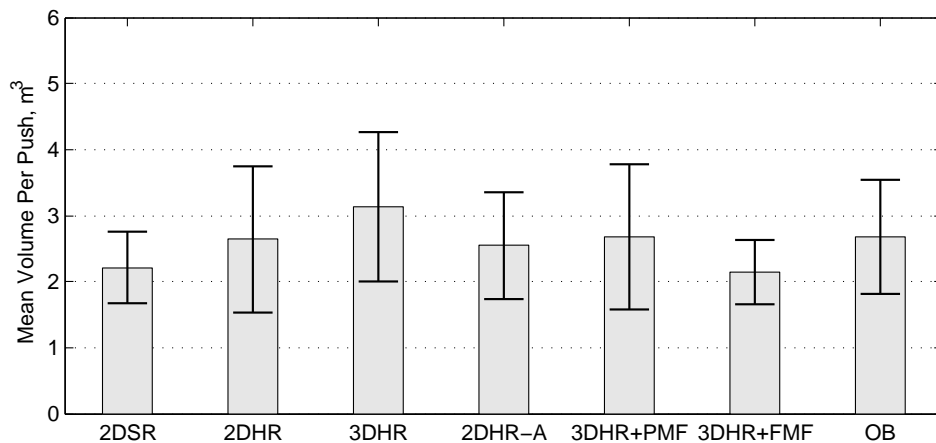


Figure 3.19: Combined vision, audio and motion cues - mean volume per push. Error bars of one standard deviation.

the left joystick controlled steering with the principal axes allocated to forward, reverse, left and right. The right joystick controlled blade orientation and position with the principal axes allocated to raise, lower, tilt-left and tilt-right. A thumb switch on the right joystick also provided control of blade pitch. The control signal recorded was binary (either on or off).

An example of the left and right joystick signals during execution of a test are shown in Figure 3.20 and Figure 3.21 respectively. To aid interpretation, the x-position of the

3.5 Results

bulldozer within the slot over time is shown in Figure 3.22. Dashed red lines indicate slot x-direction boundaries. The solid green line indicates completion criteria met.

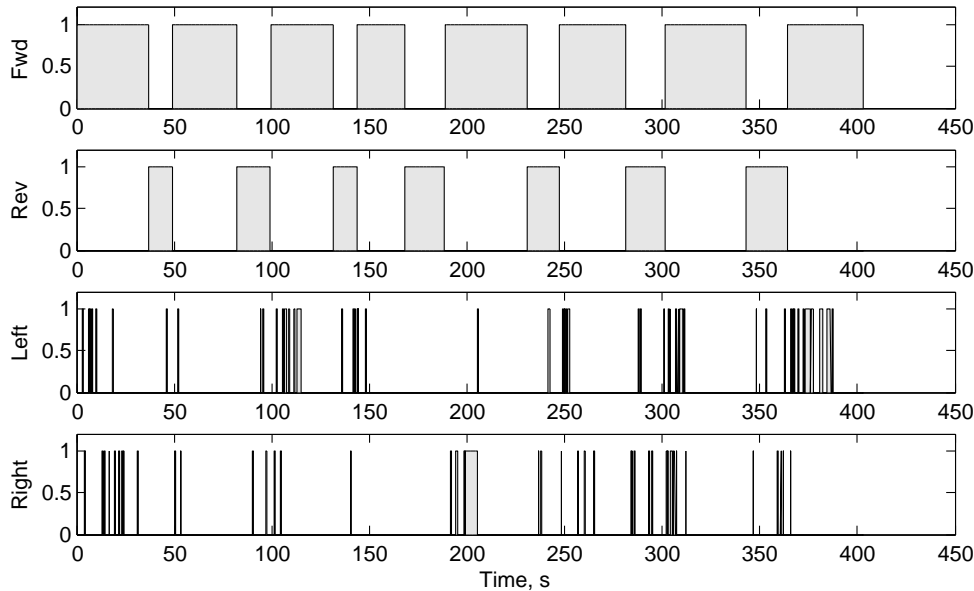


Figure 3.20: Left joystick control inputs over test duration.

It is useful to briefly examine the joystick signals and what they represent prior to evaluating them collectively. The forward and reverse commands shown in Figure 3.20 are easily associated with the push and reverse cycles shown in Figure 3.22. The left and right signals shown in Figure 3.20 illustrate the fine corrections applied by the operator to maintain the lateral position of the bulldozer. The dominant control activity on the right joystick as shown in Figure 3.21 is the control of the blade height (raise and lower commands). From Figure 3.21 it can be seen that the lower command leads the raise command as would be expected as part of the dig-push-dump cycle. Minor adjustments are made in the remaining signals to achieve desired blade orientation in different circumstances.

It is proposed that a suitable metric for quantitatively analysing control input behaviour would be counting the number of control adjustments made by the operator over the duration of the test. Specifically, counting how many times the signal goes from ‘off’ to ‘on’ in all the different axes over the execution of a test. Table 3.20 contains the count of input corrections for data displayed in Figure 3.20 and Figure 3.21. It is suggested that a low number of corrections may represent a more relaxed control style, and hence lower workload. It is useful to normalise this count by dividing by the test duration so that tests

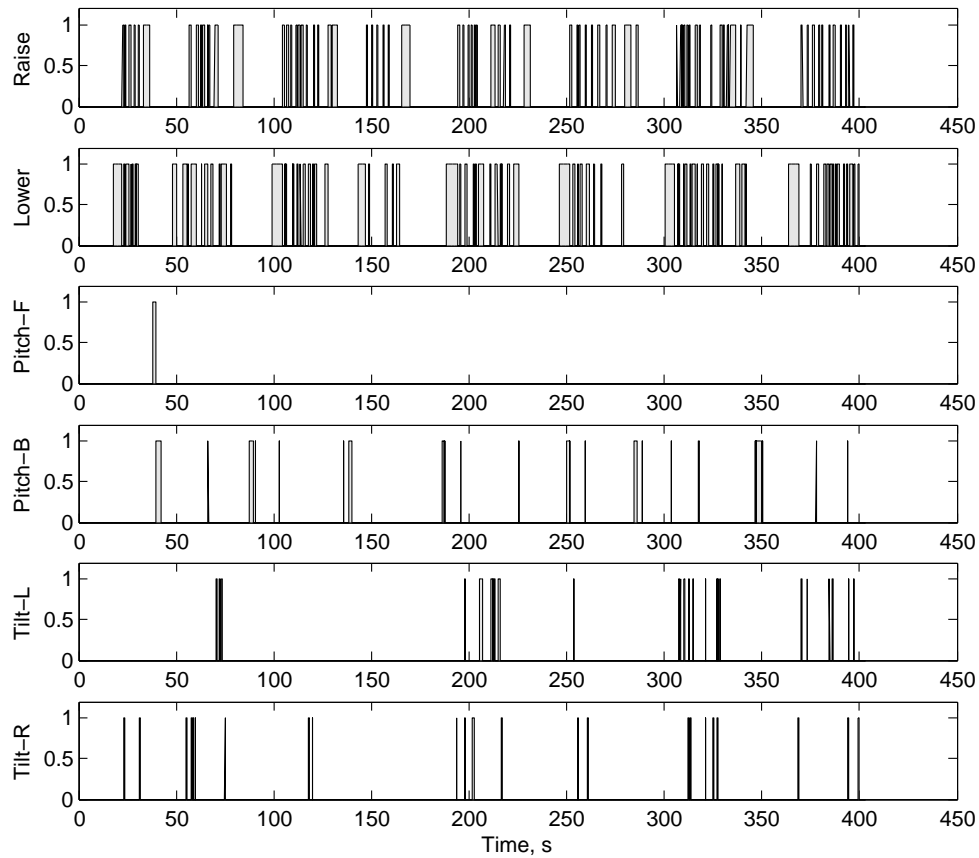


Figure 3.21: Right joystick control inputs over test duration.

and configurations can be compared regardless of task completion time. The frequency metric then becomes joystick inputs per minute.

Figure 3.23 plots the combined mean joystick correction frequency under the six different teleoperation configurations. Distinctly reduced joystick input frequency is observed in the three configurations with 3D vision. Among these, the mean input frequency for the 3D vision configuration with (3DHR+FMF) and without (3DHR) full motion feedback are found to be significant at the 5% level in a one-sided t-test comparison with the baseline (2DSR). This reduced frequency may indicate a more relaxed control style, and possibly lower actual workload. It is possible that the enhanced and additional cues enabled earlier perception of required adjustments which could then be implemented with less urgency and repetition.

3.5 Results

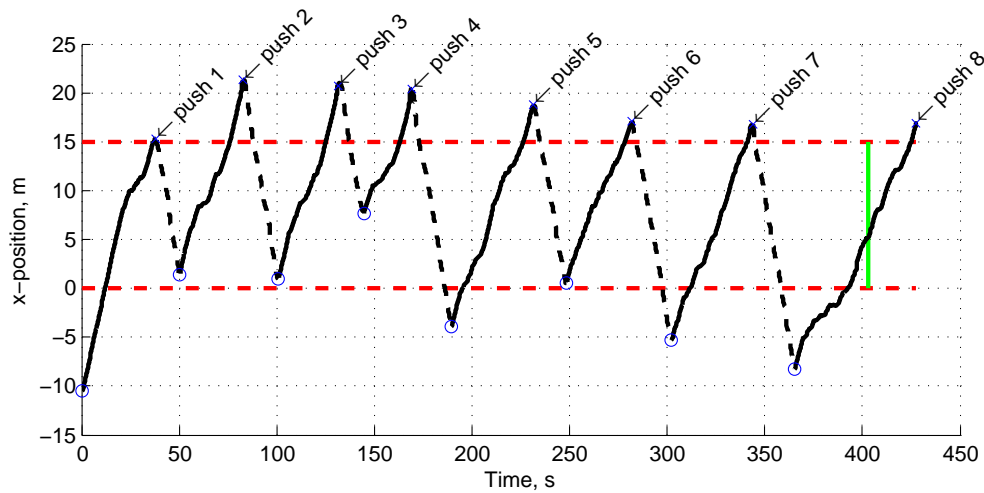


Figure 3.22: x-position of bulldozer within the slot over time.

Table 3.20: Joystick correction count for typical test.

Joystick Signal	Input Corrections
Raise	79
Lower	85
Pitch Forward	1
Pitch Back	25
Tilt Left	25
Tilt Right	23
Forward	7
Reverse	7
Turn Left	50
Turn Right	42

In addition to examining mean input frequency, it also is interesting to consider the distribution of joystick corrections over the different control axes. The nature of this distribution may represent the operator's distribution of control effort. Figure 3.24 shows the average joystick inputs per minute for each control axis, and indicates that control effort is largely concentrated among blade raise/lower commands, pitching backwards and left/right steering. This distribution suggests that the operator appears to be chiefly engaged in adjusting blade height and pitch to achieve desired cut depth and in rejecting lateral disturbances to maintain a straight heading. The identified concentration of control effort may inform where operator assist technologies could be of value as a means of reducing operator workload and allowing greater focus on higher-level control tasks.

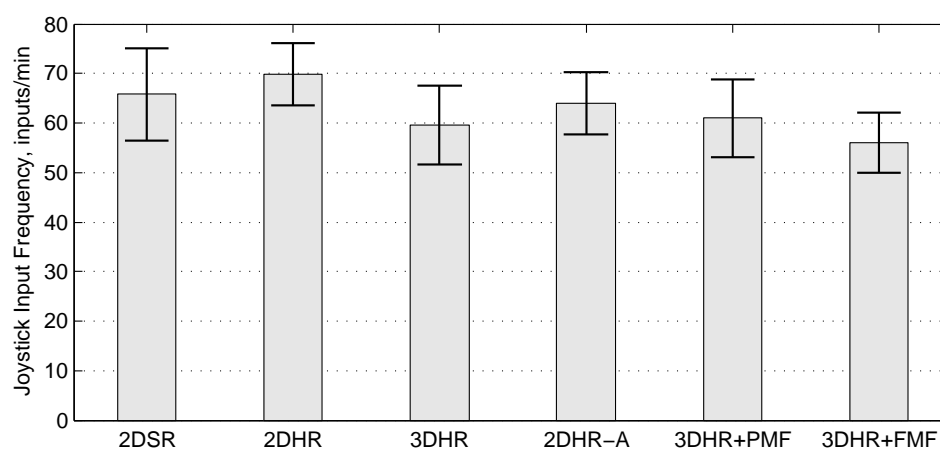


Figure 3.23: Mean number of joystick corrections per minute. Error bars of one standard deviation.

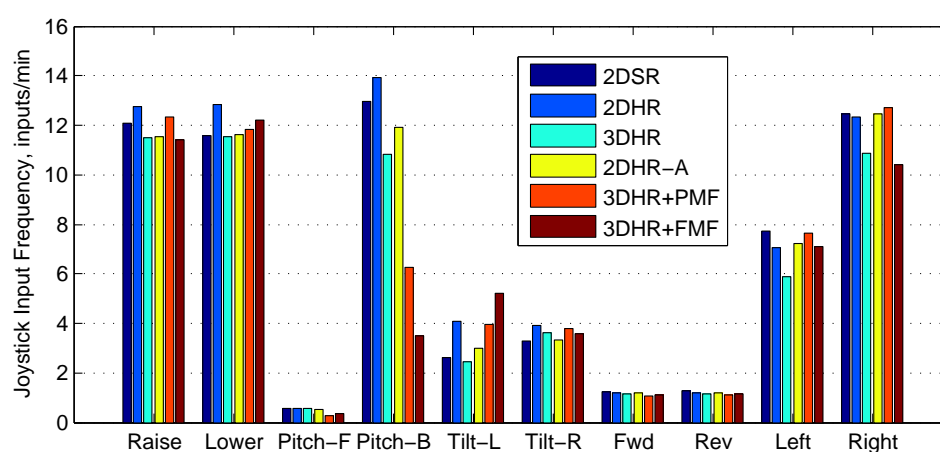


Figure 3.24: Mean joystick inputs per minute on each control axis.

Perceived workload

A summary of the total workload results for Operator B is presented in Figure 3.25 and Table 3.21.

The lowest average total workload score was observed for the 3D video configuration (3DHR). The reduction in workload observed in the 3D video configuration (3DHR) against the baseline (2DSR) was found to be statistically significant. This result complements the typically higher performance across other metrics observed for the 3D video

3.5 Results

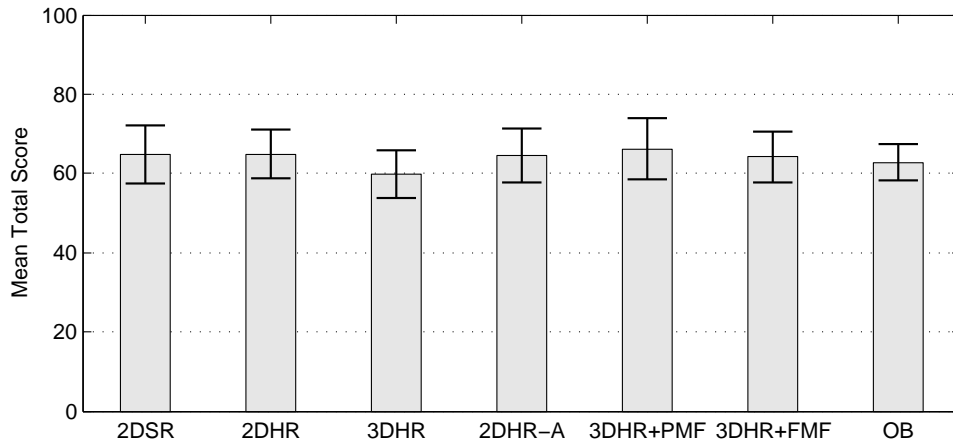


Figure 3.25: Operator B - mean workload total score. Error bars indicate standard deviation.

Table 3.21: Operator B - total workload score.

Configurations	Total Workload Score	
	<i>mean</i>	<i>std</i>
Standard AV arrangement, 2DSR	64.7	7.3
2D high resolution AV, 2DHR	64.9	6.1
3D high resolution AV, 3DHR	59.8	6.1
2D high resolution V no audio, 2DHR-A	64.5	6.8
3D high resolution AV with pitch motion, 3DHR+PMF	66.2	7.7
3D high resolution AV with full motion, 3DHR+FMF	64.1	6.5
On board operation, OB	62.8	4.7

configuration (3DHR).

Figure 3.26 presents the workload score broken into the six individual elements. Among the teleoperation configurations, the 3D video configuration with pitch motion (3DHR+PMF) and with full motion (3DHR+FMF) were found to be the most physically demanding with distinct differences from the baseline (2DSR) observed. It would appear reasonable to suggest that this is due to the need for the operator to maintain control in the presence of motion feedback. It is interesting to note that although the Physical Demand score was elevated for the teleoperation configurations with motion (3DHR+FMF and 3DHR+PMF), it is still considerably lower than the score observed for on board operation.

The results also show heightened Temporal Demand in the high resolution 2D video con-

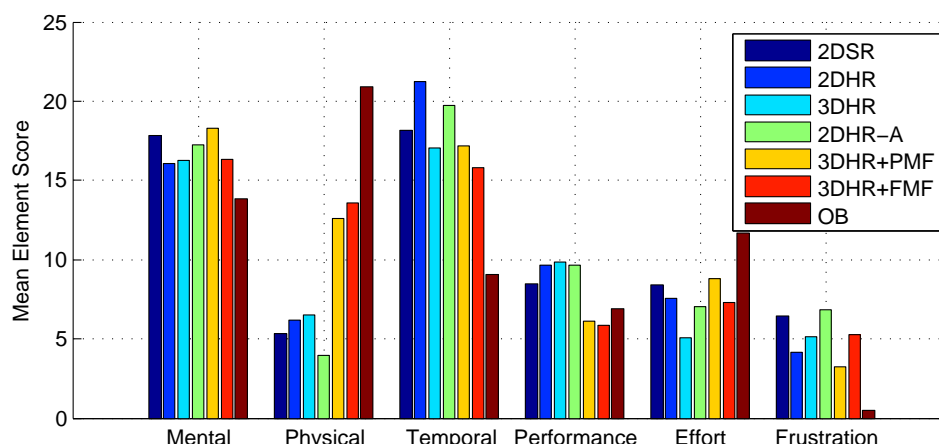


Figure 3.26: Operator B - mean workload element scores.

figurations with (2DHR) and without audio (2DHR-A). This observation is difficult to interpret but may be a consequence of the additional features that the high resolution vision allows the operator to focus on. The absence of this increase in Temporal Demand in the 3D vision configurations may then be a consequence of the fact that the vision provides a sense of realism that is comparable to being on board and thus makes the task appear less frantic.

The 3D video configuration (3DHR) required the least amount of Effort and this result was the key contributor to the overall low score for this configuration. Significant differences against the baseline (2DSR) were observed in Frustration for the 3D AV configuration with pitch motion (3DHR+PMF) and the high resolution 2D AV configuration (2DHR). The increase in Frustration observed for the 3D AV configuration with full motion feedback (3DHR+FMF) seems at odds with the decrease observed for the configuration with only pitch motion feedback (3DHR+PMF), particularly when viewed in the context of the comparatively fast completion times with full motion feedback. Notably, on board operation (OB) was by far the least frustrating for Operator B.

From Figure 3.26, it can be seen that overall, Mental Demand and Temporal Demand were the dominant contributors to the loading in the teleoperation configurations.

Some interesting observations can be made in comparison with other results obtained. The lower frequency of joystick input corrections for the motion configurations shows some correlation with the NASA TLX results. The addition of motion feedback to the

3.5 Results

3D video configuration was observed to increase perceived physical loading but reduce temporal loading according to the NASA TLX surveys. The analysis of joystick corrections shows a reduced frequency of inputs under the motion configurations, especially for blade commands, which, it appears, the operator does indeed appreciate and reflect in their NASA TLX responses.

3.5.4 Task visualisation is critical to task completion

The role of task visualisation in operator perception was examined by removing the task visualiser from the conventional and combined motion configurations in stage one of data collection. The test protocol without task visualisation (the principal means of feedback to the operator on whether tolerance at grade had been achieved) involved instructing the operator to continue until they believed they had met the completion requirements. The operator would thus stop the test when they believed the task was complete. To reiterate, the completion requirement was at least 90% of the slot surface (discretised into a 0.25m by 0.25m grid) within $\pm 100\text{mm}$.

With the task visualisation removed, it was observed that the operator was incapable of accurately judging depth and grade preventing completion requirements from being met (in all but one test run). Figure 3.27 shows the average peak at-grade percentage achieved for no motion feedback (NMF-TV) and combined motion feedback (CMF-TV) without task visualisation.

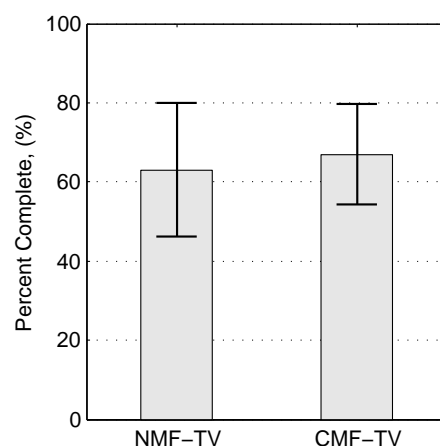


Figure 3.27: Without task visualiser - percentage complete at termination of test. Error bars show one standard deviation.

This failure to achieve completion makes it difficult to assess performance based on task completion time. Instead, it is useful to examine the volume removal history over the duration of the task as shown in Figure 3.28. The reduced rate of volume removal is clearly discernible in the two configurations without task visualisation (NMF-TV and CMF-TV). The presence of combined motion feedback appears to make a negligible contribution in supplementing the loss of the task visualiser.

The impact of removing the visualiser is detrimental to the operator's productivity from an early stage. The significantly lower gradient in the initial phase of the task reveals that average push volumes are low.

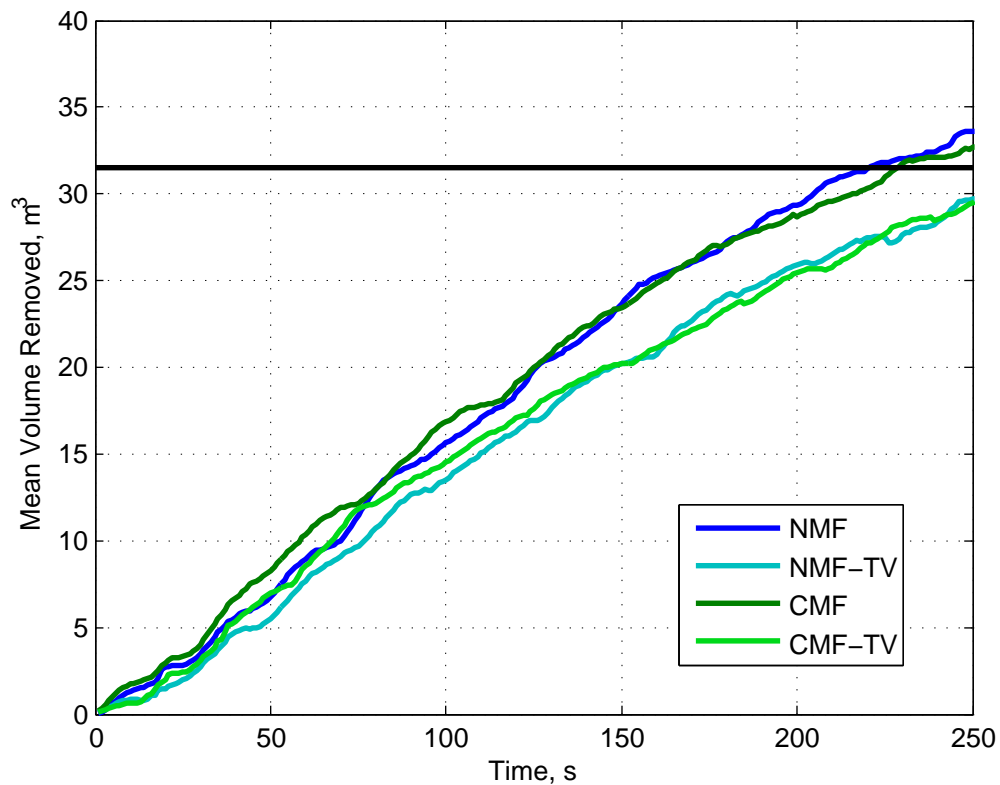


Figure 3.28: Without task visualiser - mean volume removed from within slot over time.

Influence on perceived workload

Table 3.22 presents the total workload scores for the two configurations without task visualisation. Total workload scores are also plotted in Figure 3.29.

3.6 Discussion of key results

Table 3.22: Without task visualiser - total workload score.

Configurations	Total Workload Score	
	<i>mean</i>	<i>std</i>
No motion feedback, NMF	36.8	5.4
No motion and no task visualiser, NMF-TV	73.1	6.7
Combined motion feedback, CMF	45.2	11.1
Combined motion and no task visualiser, CMF-TV	62.7	12.6

Figure 3.29 shows that removal of the visualiser distinctly elevated Operator A's perceived workload. In all tests performed by Operator A, the highest and second highest average total workload scores were given to the conventional configuration without visualisation (NMF-TV) and combined motion without visualisation (CMF-TV) respectively.

The workload score is broken down into its individual elements in Figure 3.30. Mental Demand was observed to be distinctly higher in the absence of the task visualiser (NMF-TV and CMF-TV). A plausible hypothesis for this result is that the task visualisation represents complex information in a simple manner thus freeing the operator to focus on other aspects of the task. When the operator must not only perform the task but also maintain a mental model of the terrain geometry, their concentration and sense of urgency increases.

It is interesting to note that the default configuration without the task visualiser (NMF-TV) also exhibited the largest average level of Frustration, significantly higher than that of the baseline (NMF) as well as the combined motion configuration without task visualisation (CMF-TV). This result may suggest that although the presence of combined motion cues did not substantially influence performance based on the metrics evaluated, the presence of these cues made the task less frustrating to perform.

3.6 Discussion of key results

The experimental results obtained appear broadly consistent with the afferent filter concept introduced in Chapter 1. It is non trivial to interpret these results given the complexities of examining the role of the human in the control loop. As Weir and McRuer (1968) note, 'the driver may change his dynamic characteristics or may alter the system structure (close other loops) to obtain the required increase in control fidelity'. Nevertheless, we see a detectable degradation associated with the transition from on board operation

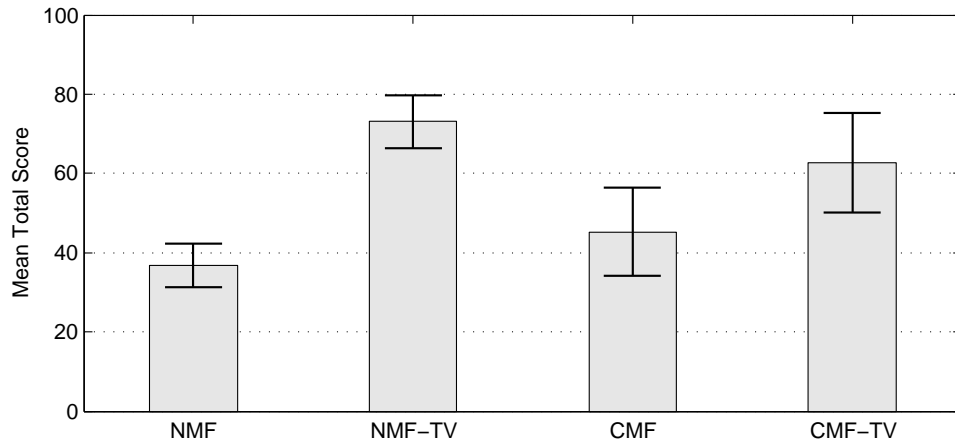


Figure 3.29: Without task visualiser - mean workload total score. Error bars indicate one standard deviation.

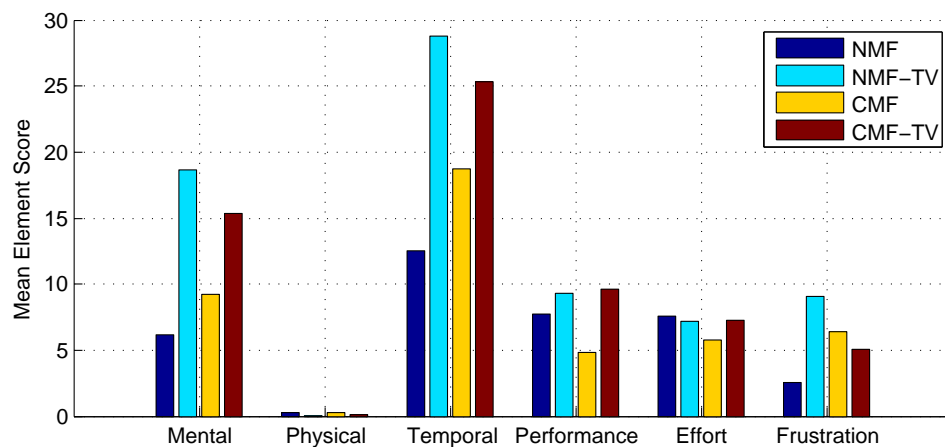


Figure 3.30: Without task visualiser - mean workload element scores.

to the conventional non-line-of-sight teleoperator and partial improvement through the addition/enhancement of various feedback channels.

Degradation from on board operation

The performance degradation from on board operation was most pronounced for Operator A, with the task completed on average 29.8% slower under the conventional teleoperation system. By comparison, Operator B was on average 13.8% slower. Assuming that the perception component of the teleoperator is at least partly responsible for this degrada-

3.6 Discussion of key results

tion, these results would appear to confirm that certain feedback channels necessary for effective operation are being ‘filtered’.

The influence of motion cues

In examining potential methods to bridge this performance gap, the introduction of isolated and combined motion cues to the teleoperator was observed to have an insignificant influence on task completion time. However, analysis of operator productivity over the duration of the test indicated elevated productivity during the early phase of the task when pitch motion only was provided. The average volume per push in the first five pushes was 19.6% higher when pitch motion was added to the conventional teleoperator configuration. This result suggests that pitch motion may be of assistance during the initial phase of the task which is largely focused on slot setup and bulk volume movement. Also notable is the fact that all motion configurations apart from pitch motion only elevated Operator A’s perceived workload compared with levels indicated in the conventional teleoperator configuration.

The limited influence observed for the combined motion configuration provided to Operator A, which also contains pitch motion, may indicate that the other motion cues were poorly constructed and/or distracting.

The influence of visual and audio cues

Examination of the potential for enhanced vision and audio to address teleoperator performance degradation indicated that depth perception was a significant cue. 3D vision (3DHR) provided a statistically significant reduction in task completion time, providing on average 12.1% faster completion times. The configuration also reduced Operator B’s perceived workload compared with the conventional configuration. Enhanced visual quality (2DHR) but without the depth perception provided by the 3D vision implementation also improved completion times but not to the same extent. The impact on completion time of removing audio feedback (2DHR-A) was found to be relatively minor.

The combined influence of enhanced visual and motion cues was also tested and it was observed that the addition of motion did not further reduce completion times. The mean task completion time when 3D video was coupled with full motion feedback (3DHR+FMF) was approximately the same as without motion (within 1%). The addition of pitch motion only to the 3D video configuration (3DHR+PMF) was observed to deteriorate perfor-

mance in terms of task completion time.

However, the coupled 3D video and full motion feedback configuration was found to improve push productivity during the initial high-volume stage of the task. The average push volume during the first five pushes was 21.9% higher than in the conventional teleoperator configuration.

The combination of the 3D vision with full motion feedback (3DHR+FMF) did not show reduced perceived workload for Operator B. However, the frequency of joystick inputs was found to be significantly reduced under the 3D visual configuration with (3DHR+FMF) and without (3DHR) full motion feedback. The frequency of joystick inputs is hypothesized to reflect operating style with reduced frequency indicating a more relaxed operating style and reduced actual workload.

The influence of task visualisation

The role of visualisation in enabling execution of the structured task was found to be critical. When the visualisation was removed from the set of cues provided to the operator, they were unable to effectively determine completion criteria. In addition, removal of task visualisation reduced push productivity in the initial phase of the task and elevated perceived workload. This result suggests that the task visualisation makes a significant contribution in supporting cut profile planning and therefore blade fill. This may indicate that the visualiser provides useful feedback in positioning and orienting the bulldozer for high volume pushes during the early phase of the task.

In summary, the results obtained from experiments indicate that perception characteristics influence the multifaceted aspects of performance in many and complex ways. This makes it difficult to draw truly definitive conclusions on the impact of individual cues but does allow for some broad statements to be made. Among the cues involving remote perception replication, depth perception would appear to be the dominant factor influencing performance with the benefits of 3D found to be particularly distinct. The provision of motion cues were observed to improve certain aspects of performance, i.e. early stage push productivity. Task visualisation also plays a critical role in enabling effective planning and task execution.

Task analysis for bulldozer operation

4.1 Aim

This chapter examines the operation of a bulldozer from the perspective of determining the correspondence between feedback pathways and control actions. The aim of this analysis is to contextualise the effects observed in Chapter 3 within an appreciation of the feedback pathways constructed by the operator.

The results obtained in Chapter 3, as well as task observation, targeted surveys and qualitative operator feedback are used to formulate a perception-control model for the experimental task performed.

4.2 Linking perception to task behaviour

Identifying the links between perception and control actions first requires a method of capturing, describing and decomposing the bulldozer operation task. Interface designers perform task analysis to better understand the control activities of the user. A broad set of techniques have emerged to assist in drawing out and describing tasks.

Kirwan and Ainsworth (1992) split the challenge of task and interface design into two stages: (i) determining what information must be supplied to permit understanding of the current status and requirements of the system; and (ii) what actions and controls must be applied by the operator to control the system. This provides the foundation for determining how this information and control capability are best provided.

A commonly used task analysis technique is Cognitive Task Analysis (CTA). Hoffman and Militello (2008) define CTA as, ‘the determination of the cognitive skills, strategies and the knowledge required to perform tasks’. Cognitive Task Analysis can be achieved through operator interviews with questions targeted at three key areas: the concepts important to the domain and how they are related; the procedures that are followed and how various reasoning ‘rules’ govern these procedures; and the exception cases in which special procedures are required (Hoffman and Militello, 2008).

Once such task analysis has been performed, there are a variety of techniques available to decompose and represent this information. A popular method of describing work organisation is hierarchical task analysis (HTA). HTA captures the operations that are performed as part of a task as well as the plans that represent the conditions under which operations are performed. The HTA applies a hierarchy that permits specification of operations with different levels of detail. The traditional purpose of such analysis is to ‘identify actual or possible sources of performance failure and to propose suitable remedies’ (Annett, 2010). While this express purpose is not exactly the intent in the case under consideration, the formal process is still relevant in applying decompositional thinking to the task in order to identify its elemental components.

4.3 Modelling the human

As described in Section 1.4.5, there has been significant effort applied to creating human models of control and behaviour. Pilot or driver models have been developed to both better understand the role of the human in a control system and to support vehicle design through modelling closed loop performance. These pilot models attempt to explain how humans take various inputs and convert them into appropriate control actions. Such models range from relatively simple disturbance rejection control of path errors in driving a vehicle (Weir and McRuer, 1968) up to highly detailed models that seek to capture each stage in the sensory process (McRuer and Krendel, 1965). Driver/pilot models continue to be used in assessment of vehicle dynamics and fidelity assessment of simulators.

Another area of focus for researchers of perception-to-action modelling is the concept of Situation Awareness. Endsley (1995) defines Situation Awareness as, ‘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’. Situation Awareness (SA) is considered a key aspect of task behaviour and determinant of performance.

4.4 Analysis of bulldozer operation and teleoperation

According to Endsley (1995), consideration of SA represents a step beyond the simplistic view of the human as a processor of inputs to yield appropriate control outputs. As Endsley (1995) notes, ‘operators must do more than simply perceive the state of their environment. They must understand the integrated meaning of what they are perceiving in light of their goals.’ It is important to recognise the separation that the concept of SA seeks to enforce between comprehension and decision: perfect comprehension can still lead to an incorrect decision and imperfect comprehension may lead to a decision that is theoretically correct given that comprehension but not appropriate for the ‘true’ situation.

Endsley (1995) describes a three level model of SA as illustrated in Figure 4.1. Level 1 is perception of the current situation. Level 2 is comprehension of the current situation. Level 3 is projection to determine future status. To explain these levels within the context of bulldozer operation, let us examine the situation in which the operator perceives sudden and excessive pitching of the bulldozer (level 1). The operator comprehends that the blade has cut too deep (level 2). Appreciating this, the operator projects that in the next reverse phase there may be a need to avoid impact due to see-sawing over the edge of the beginning of the steep gradient created (level 3). The separation into these levels suggests that training and experience become more important in making each level transition. It has been observed (Klein, 1993) that in time critical tasks, conscious deliberation of alternative solutions is rare. Instead people recognise the most appropriate action based on established experience in similar situations. However, there is a tendency back towards the analytical model when there is disputed information.

4.4 Analysis of bulldozer operation and teleoperation

The literature provides a variety of useful frameworks and theoretical models for examining and capturing the perception-to-control-action behaviour of the bulldozer operator. The following sections present the results of an investigation of the operators’ qualitative assessment of their own perception and control behaviour. These results as well as the findings of Chapter 3 are used to develop a task model and ultimately a perception-to-control model for the bulldozer operator.

4.4.1 Task and cue surveys

After completion of each test in the experiments described in Chapter 3, the operator was given a one page survey to complete. The survey presented a series of statements to which

4.4 Analysis of bulldozer operation and teleoperation

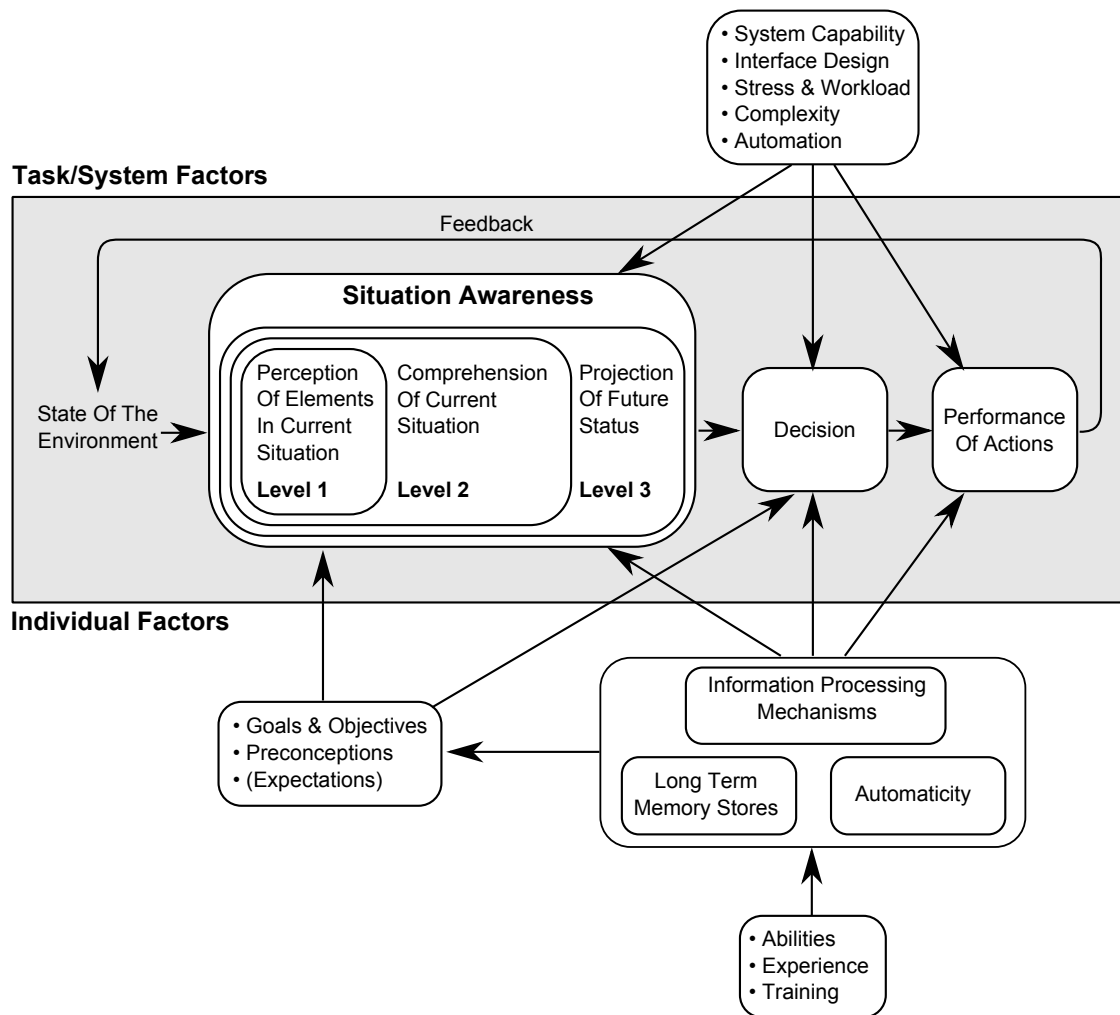


Figure 4.1: Situation awareness model (figure recreated based on Endsley (1995)).

the operator indicated their level of agreement or disagreement. The level of agreement with the statement was scored on a scale between strongly disagree (1) and strongly agree (10). The survey was split into two sections: questions related to performing the task and questions related to perceiving relevant operational cues. The task focus section targets the ability of the operator to complete the assigned task while the cue focus section is concerned with the different kinds of adjustments the operator is able to make based on the cues provided to them.

The exact questions used in the survey differed slightly between stage one and two of data collection based on the perception factors in focus. A section for general feedback was

4.4 Analysis of bulldozer operation and teleoperation

also provided.

Motion focus

The different statements presented in the survey for stage one of data collection are listed in Table 4.1 and Table 4.2. The cue focus section of the survey for stage one targets the cues made available through motion feedback. Note that there are no responses in the cue focus section for the stationary configurations (NMF and NMF-TV). Presentation of results and discussion is split into the task and cue focus sections.

Table 4.1: Motion focus - task focus statements.

Statement
Q1 <i>I was able to fill the blade at desired speed.</i>
Q2 <i>I was able to fill the blade consistently.</i>
Q3 <i>I was able to recover in instances where I detected a deviation from desired fill.</i>
Q4 <i>I felt my productivity was equivalent to being on board.</i>
Q5 <i>I would feel comfortable performing this task for a full shift</i>

Table 4.2: Motion focus - cue focus statements.

Statement
Q6 <i>I made steering adjustments based on motion cues in certain instances.</i>
Q7 <i>I made speed adjustments based on motion cues in certain instances.</i>
Q8 <i>I made implement adjustments based on motion cues in certain instances.</i>

Task focus

The mean response and standard deviation for each statement in the task focus section are summarised in Table 4.3. The operator provided similar responses to statements 1 to 3 (targeting perceived ability to maintain fill speed, consistency and reject disturbances) across all configurations excluding those where task visualisation was absent (NMF-TV and CMF-TV). The absence of the task visualisation tool appears to have had a significant detrimental effect on the operator's perception of their ability to control blade fill.

4.4 Analysis of bulldozer operation and teleoperation

Interestingly, this perceived degradation was less extreme when the operator was provided combined motion feedback (CMF-TV).

Table 4.3: Motion focus - task focus statement responses. Mean and (std).

	NMF	NMF-TV	CMF	CMF-TV	PMF	VMF	TMF
Q1	8.0 (0.8)	3.4 (0.8)	8.3 (1.3)	6.5 (1.4)	8.4 (0.5)	8.0 (0.5)	8.0 (0.5)
Q2	7.6 (1.1)	3.6 (0.8)	8.0 (1.8)	6.9 (1.0)	8.3 (0.9)	8.0 (1.1)	8.1 (0.7)
Q3	7.3 (1.0)	2.0 (0.9)	8.1 (2.1)	5.1 (1.6)	7.5 (1.0)	7.2 (1.2)	7.5 (1.5)
Q4	4.5 (2.1)	1.0 (0.0)	6.3 (1.9)	3.8 (1.5)	6.7 (0.8)	5.7 (1.7)	6.6 (1.3)
Q5	4.3 (2.0)	1.0 (0.0)	6.7 (2.1)	3.6 (1.6)	7.0 (1.6)	5.6 (1.9)	6.7 (1.1)

The addition of motion feedback, with the exception of isolated vibration motion feedback (VMF), significantly enhanced the operator's perceived sense of productivity. The removal of the task visualisation degraded the operator's sense of productivity but the effect was less when combined motion feedback (CMF-TV) was still available.

The same configurations in which heightened productivity was perceived, were also viewed favourably in terms of imagined comfort with full shift operation. The mean response in combined motion (CMF), pitch motion (PMF) and translation motion (TMF) were all better than neutral (score > 5). This suggests that motion feedback can play a role in supporting operator engagement. The stationary configuration without task visualisation (NMF-TV) presented the most significant level of discomfort.

Cue focus

The survey responses in the cue focus section are summarised in Table 4.4. Responses indicate that the operator felt combined motion feedback (CMF) provided information that was of use in informing decisions around steering. Generally, the operator indicated that the other motion feedback configurations did not support steering adjustments. This may be a consequence of the inclusion of roll and yaw in the combined motion feedback (CMF) which are absent from the other configurations.

The operator's responses appear to indicate that combined motion (CMF) and vibration motion (VBF) feedback provide cues that are relevant to speed modulation. This suggests

4.4 Analysis of bulldozer operation and teleoperation

Table 4.4: Motion focus - cue focus statement responses. Mean and (std).

	CMF	CMF-TV	PMF	VMF	TMF
Q6	7.6 (1.1)	3.9 (1.7)	3.3 (2.1)	3.1 (1.5)	2.1 (0.3)
Q7	8.7 (1.2)	6.8 (1.0)	5.7 (2.7)	8.0 (1.8)	5.1 (1.0)
Q8	9.6 (1.1)	8.4 (0.5)	9.0 (0.9)	4.6 (1.6)	6.5 (1.0)

that vibration cues provide information that is useful in determining appropriate tracking speed.

The combined motion feedback (CMF) and pitch motion feedback (PMF) configurations appear to have provided the most useful information relating to implement control. This suggests that pitch motion may provide the bulk of the cues concerning implement adjustments and be of particular use when high degree of implement control is required.

In summary, the broad spectrum of cues available in the combined feedback configuration (CMF) would appear to be used and favoured by the operator. This is degraded by the removal of the task visualisation. Pitch motion feedback (PMF) and the isolated pitch cues it provides would appear to be the next most useful as perceived by the operator.

Vision, audio and motion focus

The statements presented in the survey during stage two of data collection are summarised in Tables 4.5 and 4.6. The cue focus section targets the operator's sense of awareness of the state of the bulldozer as well as the surroundings relevant to the accomplishment of the task. The task and cue focus sections are examined separately below.

Task focus

The task focus section results are presented in Table 4.7. The responses indicate limited distinction between configurations in terms of the operator's perceived ability to control blade fill (Q1 to Q3). All enhanced perception configurations are elevated against the baseline (2DSR), but only marginally in some instances.

The operator's perceived sense of productivity (Q4) was distinctly improved in all configurations with high resolution 3D video (3DHR, 3DHR+PMF, 3DHR+FMF). The added

4.4 Analysis of bulldozer operation and teleoperation

Table 4.5: Vision, audio and motion focus - task focus statements.

Statement
Q1 <i>I was able to fill the blade at desired speed.</i>
Q2 <i>I was able to fill the blade consistently.</i>
Q3 <i>I was able to recover in instances where I detected a deviation from desired fill.</i>
Q4 <i>I felt my productivity was equivalent to being on board.</i>
Q5 <i>I would feel comfortable performing this task for a full shift</i>

Table 4.6: Vision, audio and motion focus - cue focus statements.

Statement
Q6 <i>I had a good sense for the location and relative depth of the blade.</i>
Q7 <i>I was able to get an impression of the blade fill by observing material roll over.</i>
Q8 <i>I was able to perceive contours and features in the slot that helped me.</i>
Q9 <i>I had a good sense for the location and orientation of the dozer within the slot.</i>
Q10 <i>I was able to perceive machine loading.</i>

depth perception from 3D vision possibly increased the level of confidence the operator had in the visual information, allowing for judgements and decisions closer to those that would have been made on board the bulldozer compared to all other configurations. The 3D visual configuration with full motion feedback (3DHR+FMF) in particular significantly elevated perceived productivity. This was the only configuration in which the operator allocated a score that was close to neutral and at times in agreement with the statement.

The operator appeared most comfortable (Q5) with the motion configurations (3DHR+PMF and 3DHR+FMF), indicating distinctly higher potential for operation over a full shift compared with the baseline (2DSR). Motion may have provided a sense of familiarity with the on board experience thus making the operator more comfortable. Interestingly, the operator found the high resolution 2D configuration (2DHR) more comfortable on average than the 3D configuration (3DHR) when motion was not included. The removal

4.4 Analysis of bulldozer operation and teleoperation

Table 4.7: Vision, audio and motion focus - task focus statement responses. Mean and (std).

	2DSR	2DHR	3DHR	2DHR-A	3DHR+PMF	3DHR+FMF
Q1	7.7 (0.5)	7.8 (0.4)	8.0 (0.4)	7.9 (0.3)	8.1 (0.4)	8.1 (0.3)
Q2	7.5 (0.7)	7.6 (0.5)	7.9 (0.5)	7.4 (0.6)	8.0 (0.5)	7.8 (0.7)
Q3	7.0 (0.6)	7.3 (0.6)	7.1 (0.6)	7.1 (0.6)	7.3 (0.6)	7.6 (0.5)
Q4	3.9 (0.3)	4.0 (0.4)	4.3 (0.5)	3.7 (0.9)	4.3 (0.6)	5.0 (0.5)
Q5	7.0 (0.5)	7.5 (0.8)	7.3 (0.5)	6.9 (0.8)	8.1 (0.4)	8.1 (0.3)

Table 4.8: Vision, audio and motion focus - cue focus statement responses. Mean and (std).

	2DSR	2DHR	3DHR	2DHR-A	3DHR+PMF	3DHR+FMF
Q6	7.3 (0.6)	7.8 (0.6)	7.9 (0.3)	7.7 (0.5)	7.9 (0.3)	8.0 (0.5)
Q7	6.9 (0.6)	7.8 (0.5)	8.0 (0.0)	7.7 (0.5)	7.9 (0.3)	8.0 (0.5)
Q8	6.4 (0.6)	7.1 (0.6)	7.2 (0.6)	6.9 (0.4)	7.3 (0.6)	7.4 (0.7)
Q9	7.3 (0.5)	8.0 (0.3)	7.9 (0.3)	7.9 (0.5)	7.9 (0.3)	7.8 (0.4)
Q10	5.1 (0.3)	5.3 (0.6)	5.1 (0.3)	5.3 (0.6)	5.0 (0.0)	5.0 (0.0)

of audio (2DHR-A) degraded the operator's comfort to a level indistinct from the baseline (2DSR), suggesting that audio provides cues that support engagement.

Cue focus

The cue focus survey results are presented in Table 4.8. The operator responses suggest that enhanced visual quality distinctly improved sense of awareness around blade location and relative depth (Q6) as well as blade fill (Q7). The addition of motion (3DHR+PMF and 3DHR+FMF) to the 3D visual configuration (3DHR) does not appear to have registered with the operator as having further enhanced this awareness.

4.4 Analysis of bulldozer operation and teleoperation

The improved visual quality appears to have also elevated the operator's perceived ability to identify contours and features within the slot (Q8). The 3D configurations performed marginally better than the 2D configurations. The configurations with motion feedback received the highest mean scoring.

Similarly, enhanced visual quality appears to have elevated the operator's sense of the location and orientation of the bulldozer (Q9). This sense appears slightly degraded with the addition of full motion feedback (3DHR+FMF), possibly suggesting that the diversity of the motion cues can at times be confusing and difficult to resolve.

No distinct differences were observed between configurations with respect to the operator's perceived ability to discern machine loading (Q10). The slight elevation in the mean response for the high resolution 2D video without audio (2DHR-A) configuration is confounding when considering the assumption that qualitative feedback from operators suggest that audio provides a useful cue in detecting machine loading.

4.4.2 Operator feedback

Operator feedback was collected throughout the execution of all tests in post block interview sessions. Questions focussed on which cues were considered important and how different cues were used and transformed into command decisions.

Operator A

The following points represent a consolidation of feedback provided in interviews with Operator A. The points are a summary of the operator's opinion.

- Pitch is the most important and beneficial motion feedback cue. It supports blade lift decisions particularly around slot entry and exit. It highlights areas where there are significant deviations from a flat floor. It provides information that is useful in establishing an optimal spoil pile.
- Roll is useful in certain circumstances but distracting in others. On the whole, roll is not particularly beneficial for the task being executed. Motion feedback is not disorienting without roll.
- Yaw is useful in certain circumstances but distracting in others. Yaw is most useful for detecting when the edge of the blade has caught the side of the slot causing the

4.4 Analysis of bulldozer operation and teleoperation

machine to swing about. On the whole, yaw is not particularly beneficial for the task being executed. Motion feedback is not disorienting without yaw.

- Vibration is a useful cue but less important than pitch. The information it provides around track slip is used to perform blade lift and/or throttle commands. The vibration cues can be processed and incorporated without significant mental effort and therefore provides useful information without being distracting.
- Translation is a useful cue but is less important than pitch. Translation feedback provides some information about bogging down and highlights directional changes. As with vibration, the information provided by translation feedback can be processed and incorporated without significant mental effort.

Operator B

The following points represent a consolidation of feedback provided in interviews with Operator B. The points are a summary of the operator's opinion.

- The 3D high resolution video is the most favoured of all visual configurations. The 3D visual implementation provides access to the majority of the visual cues that would be seen in the bulldozer itself.
- The high resolution video feeds provide superior vision of the amount of material being carried by the blade. This information supports better decisions around raising or lowering the blade, or modulating speed to adjust how much load is being carried.
- Depth perception provided by the 3D video feed allows better appreciation of proximity to slot walls and position within the slot. This supported more accurate corrections and appropriate timing.
- No dizziness, headaches or eye pain were experienced while wearing the 3D glasses.
- The 3D video feed had the highest mental loading due to the amount of visual cues available.
- The stereo audio provided cues indicating track slip that are more difficult to detect only through video. The pitch of the drive whine can also be utilised to discern

engine and blade loading. Without audio, concentration levels were elevated, requiring greater focus on vision.

- Pitch provided cues relating to cut depth and slot geometry. The pitch response when cutting can be used to quantify the relative depth achieved. The pitch cue also gives good indication of slot entry and exit geometry.
- The roll cue is not regarded as being as important as pitch, but its presence was preferred. It was considered useful in discerning lateral grade.
- Vibration was the most important motion feedback cue. The same cues are obtained from the replicated vibration but without the aggressive motion of being on board the bulldozer.
- Translational acceleration motion is helpful in providing an accurate feeling of the state of the bulldozer. This cue provided an accurate feeling of motion and useful feedback when overcutting.
- It was easier to see past shadows using the enhanced vision, particularly in the 3D configuration.

4.5 Modelling bulldozer control behaviour

The execution of a slot bulldozing operation involves multiple coordinated activities. The control behaviour of the operator must adjust according to the specific requirements at a given stage of the operation. The operator switches between six operational states to execute the overall task goal. These are:

1. **Aligning** - Lining up with the planned slot boundaries
2. **Entering** - Driving into the slot up to the point of beginning to cut
3. **Cutting** - Executing a desired cut trajectory
4. **Carrying** - Pushing cut material through the slot to the dump point
5. **Dumping** - Unloading material from the blade at the dump point
6. **Reversing** - Returning to the start point for the next push

Figure 4.2 presents a state transition model of this behaviour. This state model provides a starting point for the Hierarchical Task Analysis (HTA). Each state corresponds to a

4.5 Modelling bulldozer control behaviour

sub-goal within the broader task. The HTA further dissects these sub-goals into further subordinate goals. The purpose of this HTA is to highlight where the operator's control behaviour varies in nature and how it is concentrated in different stages. The HTA for the slot bulldozing operation is illustrated in Figure 4.3 and described in the following sections.

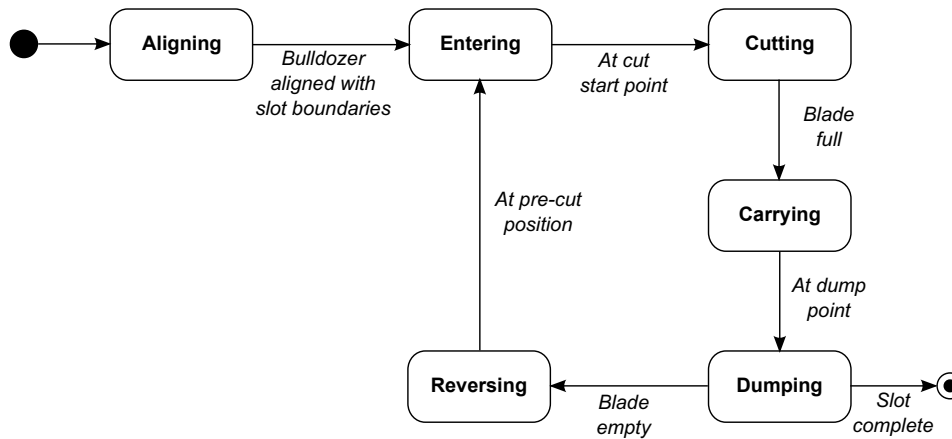


Figure 4.2: State transition model.

Sub-goals 1, 2 and 6

Sub-goals 1, 2 and 6 relate to comparatively simple operations around maintaining a desired position and orientation within the slot boundary. The operator must make steering and velocity adjustments to control the bulldozer's pose.

Sub-goal 1 represents the alignment performed prior to entry within the slot boundary. By contrast, sub-goals 2 and 6 demand slightly elevated focus in that the slot boundary may be physically defined by the slot sides. There is typically no or limited clearance between the blade and the competent material in the sides of the slot. The operator must therefore achieve correspondingly fine control over the orientation of the bulldozer so as to prevent unnecessary lateral cutting into the slot walls while executing sub-goal 2 and to avoid tracking up or against the slot walls while reversing in sub-goal 6.

Furthermore, the reverse operation executed as part of sub-goal 6 may necessitate additional velocity adjustments to avoid impacts caused by abrupt see-sawing type motion of the bulldozer as a result of excessive undulations in the slot surface.

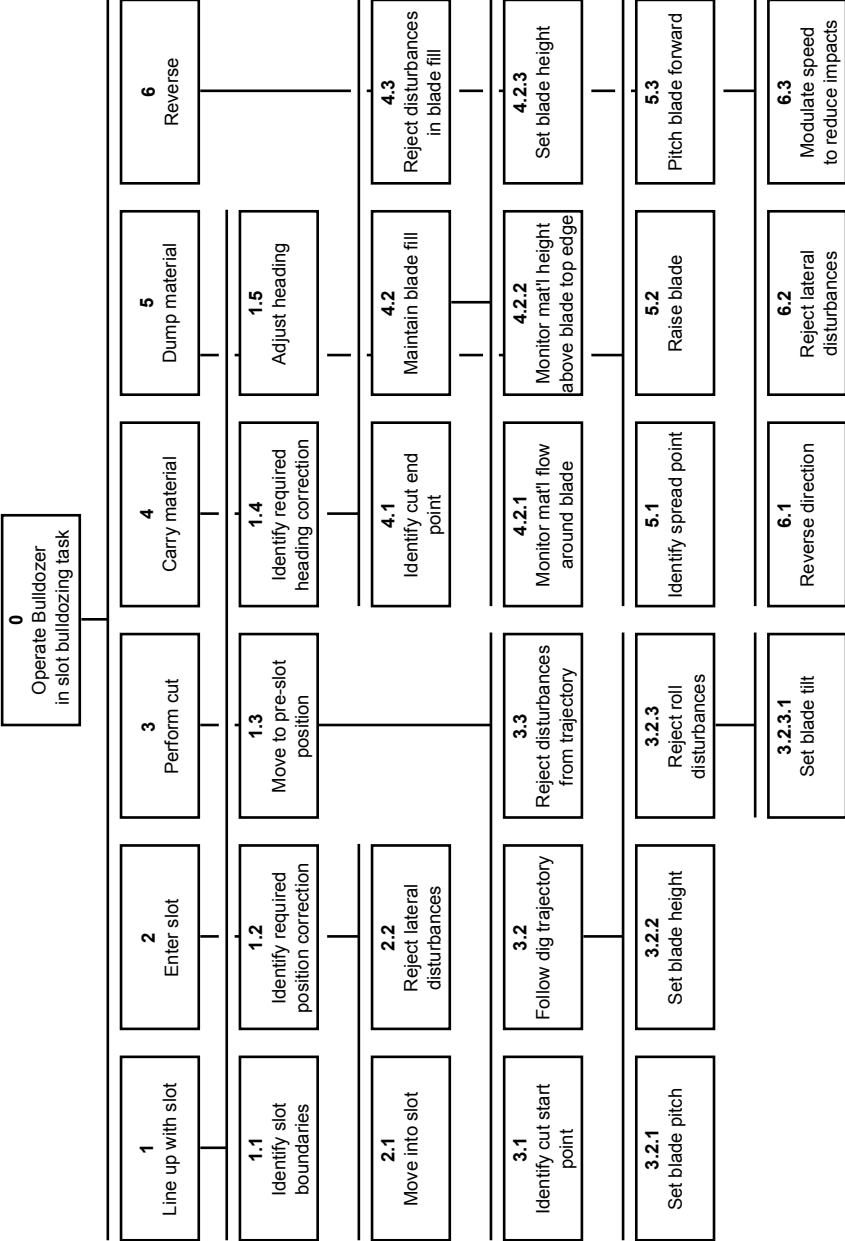


Figure 4.3: Hierarchical Task Analysis for slot bulldozing.

4.5 Modelling bulldozer control behaviour

Sub-goal 3

Sub-goal 3, *Perform cut*, is arguably the most involved task in executing the slot. The operation requires selection and following of the required dig trajectory. Following the required dig trajectory involves modulating blade height and pitch. The competent material typically prevents precise tracking and random disturbances must be rejected.

Loading on the bulldozer must also be considered. Excessive cutting may lead to excessive loading and cause a loss of traction. The operator must monitor loading to reduce instances of traction loss and correct for this when it occurs.

In addition to rejecting disturbances from the required dig trajectory, the operator must also reject any roll and lateral disturbances.

Sub-goal 4

Once competent material has been cut, it must be carried to the end of the slot. The loading of the carried material on the blade alters the dynamic response of the bulldozer. The control response is subsequently affected and this must be accommodated by the operator.

A key sub-goal during the carry stage is maintaining blade fill. The operator must make blade adjustments to prevent losing material and also to prevent further cutting. This activity requires fine control of blade height.

While carrying material, the operator must also attempt to reject any roll and lateral disturbance.

Sub-goal 5

After the material has been carried through the slot, it must be spread or dumped into a pile. The operator must determine the desired dump point for the current push, recognising that subsequent pushes may follow. Successive pushes will typically help build up a pile that can also be used as a ramp for subsequent dumping into clear space. The efficient dumping of material is closely related to the geometry of this outramp. A ramp that is too steep is inefficient in that material must be lifted higher against gravity. A ramp that is not steep enough is inefficient in that material must be pushed over a longer distance.

Once the dump point has been reached, the operator must release material by raising and

pitching the blade forward. After all material has been dumped, the operator transitions to reverse.

4.6 Mapping perception to control

The hierarchical task analysis (HTA) highlights the various control demands on the operator. It provides a basis for the proposed control model and exploring the relationship between perception and control actions.

The operator, the bulldozer and the environment are the three parts that together form the closed-loop vehicle element. The commanded input to the operator is the desired slot geometry. In addition, any environmental features or obstacles represent an external input to the operator.

A multiloop model is proposed in which there is a main outerloop focused on path / trajectory control and three inner loops managing subsidiary aspects of the bulldozer control task. These four feedback loops are:

1. Path / trajectory motions
2. Angular motions
3. Machine loading
4. Blade loading

The model presented attempts to describe the on board operation of the bulldozer, permitting a theoretical assessment of introducing a teleoperator. This model is illustrated in Figure 4.4. This model has been developed with strong influences from Weir and McRuer (1968).

The feedback pathways illustrated in Figure 4.4 are examined in more detail below. The impact on these feedback pathways of transitioning from on board operation to teleoperation is also discussed.

4.6 Mapping perception to control

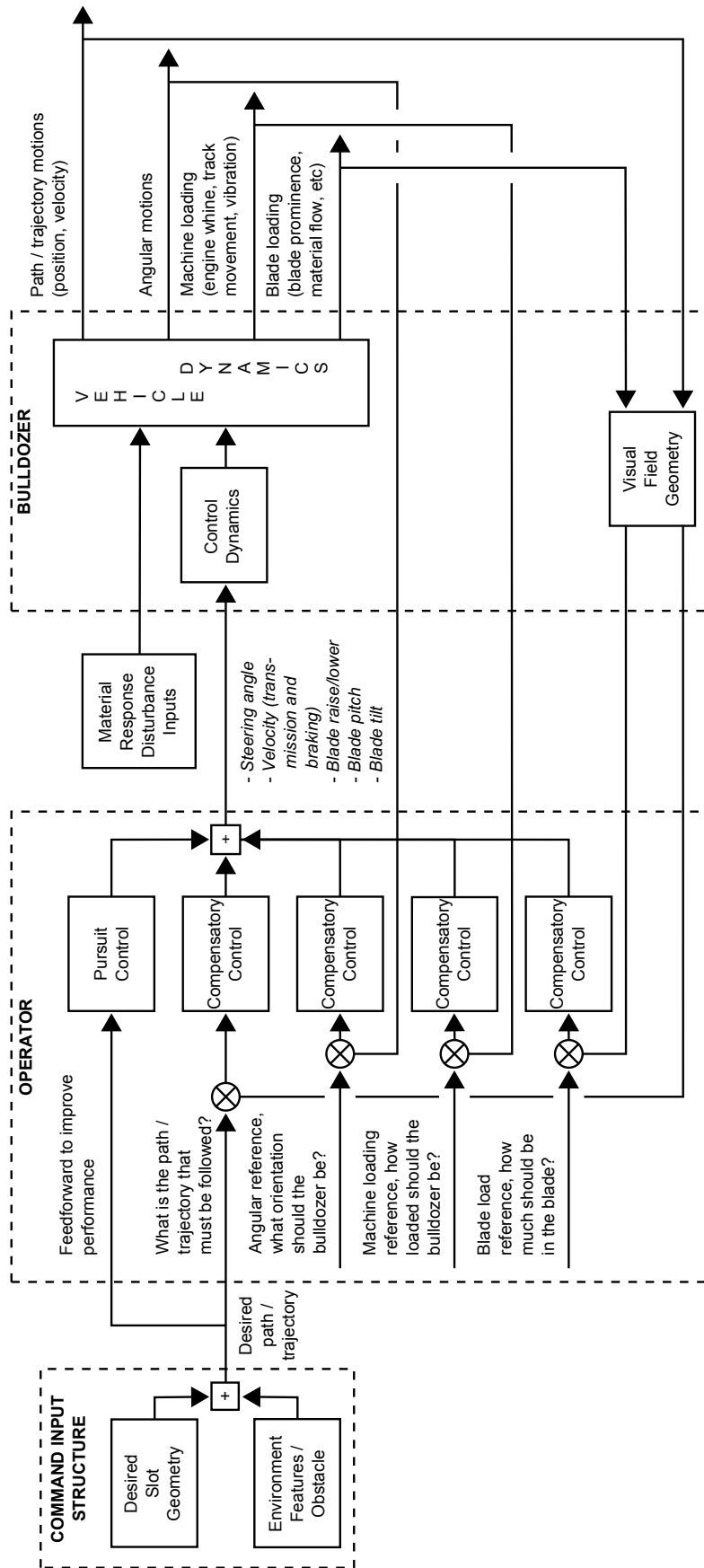


Figure 4.4: Model of bulldozer operation.

1. Path / trajectory motions

At all times, the operator must seek to maintain a desired path and cut trajectory within the slot. This control task would appear to largely depend on vision and situational awareness. Vision allows the operator to perceive key terrain features that enable localisation for high level path control. The position of the blade relative to the terrain surface and slot boundaries similarly provides important visual feedback for controlling the cut trajectory. On board, the operator can make use of a wide field of view, peripheral vision and prior awareness of site geometry. Under teleoperation, the operator's field of view is reduced and there is no access to useful peripheral vision. The operator may also have limited prior awareness of site geometry given that the machine was not physically accessed and boarded.

The execution of an individual slot task involves primarily linear motion. The control of heading relates principally to rejection of disturbances. Heading disturbances may be observed visually through relative motion of the terrain. Such cues may be reduced under teleoperation due to the lack of peripheral vision. On board, abrupt heading disturbances may be perceived through jerks in yaw motion. However, such sudden jerks in yaw typically occur when the blade catches the side of the slot or when the operator over-corrects to prevent this. In both instances, the cue provided by motion would arguably occur only after lesser disturbances have failed to be rejected.

2. Angular motions

The inner compensatory loop to control the angular orientation and motions of the bulldozer is performed with reference to a learnt sense of what orientation the bulldozer should be in a given circumstance. The cutting operation involves tracking a desired cut trajectory that must be matched to a given vehicle orientation. In addition, the operator will seek to establish a desired gradient at slot entry and exit. Pitch motion would appear to be a key cue in informing these particular control tasks.

On board, the operator may perceive a change in the gravity vector and/or dynamic roll motion. However, due to the offset between the end-effector (blade) and the centre of gravity of the bulldozer, by the time the operator perceives roll it may be too late to make any necessary correction. In such instances, the operator must store and recall this information when conducting the subsequent pass.

4.6 Mapping perception to control

In a static teleoperation system, the operator must rely on vision to interpret the vehicle's orientation. However, the lack of peripheral vision removes one key pathway through which relative motion might be perceived visually.

3. Machine Loading

The loading on the bulldozer must be carefully monitored both during the cutting and carrying operation. The feedback loop may not be used during reversing and moving to position.

The operator seeks to track a level of loading that permits maximum blade fill while maintaining forward motion. The key perception pathways related to this tracking task are relative blade position, pitch, vibration, engine and machine acoustic response and translational acceleration/deceleration.

Variation in loading on the bulldozer may be sensed through pitch motion and linear acceleration/deceleration. For example, a lightening of load may cause an acceleration and forward pitching. The bulldozer pitch will also influence the level of loading which the bulldozer can sustain.

The acoustic response of the engine can be interpreted in much the same way as the driver of a manual car may listen to the engine for cues to change gears. The power of the bulldozer may be managed by controlling the machine loading in this fashion.

When traction is lost, the bulldozer tracks will continue to rotate under a condition termed track slip. Track slip will cause a unique vibration and aural response. On board, track slip is easily perceived through this vibration and aural response as well through direct view of the freely rotating tracks against the stationary background. These cues inform the operator that machine loading must be reduced. Poor visual quality in teleoperation can make it difficult to identify track slip visually. The audio and motion feedback pathways may also be unavailable to the teleoperator.

4. Blade loading

The innermost loop is maintenance of a set blade fill during the carry operation. The blade fill can only be observed visually by means of the presence of material above the top edge of the blade or flowing around the side of the blade. The perception of these cues relies on perception of depth, texture, contrast and independent flow. These cues are very subtle

compared to other cues such as gross motion and relative motion of the blade. Poor visual quality in teleoperation can significantly reduce the operator's ability to perceive these cues.

4.7 Overview of assembled model

The bulldozer perception-control model illustrated in Figure 4.4 highlights the complexities of bulldozer operation. The control of path and cut trajectory motions is suggested to be the highest level activity performed by the operator. Within this loop, the operator seeks to reject disturbances related to orientation changes of the bulldozer that would prevent them from achieving the desired cut trajectory. During cutting and pushing the operator must then also maintain control over machine and blade loading. Complicating a simple representation of the perception and control activities performed by the operator is the fact that the importance of the different feedback loops varies depending on the operational state the bulldozer is currently in.

The transition from on board operation to teleoperation degrades the feedback pathways illustrated in Figure 4.4. The reduction in field of view, depth perception and visual quality would appear to have the most significant impact on the path / trajectory and blade fill control loops. The elimination of motion cues would appear to be most detrimental to effective control of vehicle orientation. Degraded vision, audio and elimination of motion cues would appear to all combine to limit effective control of machine loading during teleoperation.

The perception-control model provides context for consolidating the results presented in Chapter 3 and considering how operator perception might be best enhanced to support effective bulldozer teleoperation. This is pursued in the following chapter.

Consolidating perception requirements

5.1 Aim

This thesis has examined the role of perception in supporting high levels of performance in bulldozer teleoperation. The experiments conducted and analysis applied provide insight into the specific influence of a diverse set of perception cues. The aim of this chapter is to consolidate these findings into a cohesive summary of perception for bulldozer teleoperation and its influence on performance and user acceptance.

5.2 Perception groupings

The major findings across the four perception groupings (vision, audio, motion and task visualisation) and their implications on requirements for effective bulldozer teleoperation are summarised in the following sections. Where relevant, comments around practicality of achieving adequate closure of these feedback pathways are also provided.

5.2.1 Vision

The perception-control model presented in Section 4.6 suggests that vision plays a key role across most control activities performed by the operator. The experimental results obtained indicate that, among the aspects of vision, depth perception provides the most distinct influence on performance for non-line-of-sight bulldozer teleoperation in the experimental task conducted. The 3D high resolution visual configuration evaluated in this study resulted in a reduction in task completion time by approximately 12%. The operator

for which this result was obtained was able to achieve a performance level in teleoperation equivalent to their on board operation performance.

Results from evaluation of volume removal progress and push productivity suggest that vision plays a more important role during the finishing phase of the task. This phase of the task requires establishing flatness at the target depth and has similar aspects to flat level grading. Finer control is required during this phase and visual quality appears to play a role in enabling this capability.

The performance difference between the 3D and 2D configurations may be attributed to the additional depth cues provided. The operator commented that there were cues that could be obtained in the 3D configuration that were not available in the 2D configuration. In particular, the additional depth information enabled the operator to better distinguish between terrain features with similar colouration and texture but at different depths in the scene.

It must be noted that the visual quality improvements evaluated in this study are coupled with a reduction in capture-to-display latency of approximately 100ms. The latency reduction represents approximately a 30% decrease in latency from the standard visual configuration employed as the baseline. However, the closed loop latency, defined as the time taken from when the operator instructs a correction to when the correction is observed in the video, is reduced by a figure closer to 10 to 15% as a result of this 100ms decrease.

This study has sought to explore the raw potential for visual quality to improve performance and the experimental apparatus was developed to support this capability without consideration of commercial and technical practicality. There is significant additional technical complexity and bandwidth utilised in achieving the low latency, high resolution visual feed. The 3D arrangement introduces further complexities due to the need for additional hardware. When bandwidth is constrained, there is a balance that must be struck between latency and visual quality.

The closed loop latency is the latency that the operator perceives. This study found that performance comparable with on board operation could be achieved for at least one operator despite this closed loop latency in the teleoperation system. Cunningham et al. (2001) also found that, with practice, humans can develop ways to accommodate delay and still operate effectively. However, variability in this delay hinders accommodation

5.2 Perception groupings

and is highly detrimental to performance.

It would seem more advisable to pursue improved visual quality at the cost of latency when the full system latency is high. Bandwidth allocation may be optimised by prioritising visual quality improvements in views that are most relevant to the visual cues required and the task being performed. Appropriate placement and use of cameras with wide field of view may allow for redundant views or unused views to be removed. Alternate means may be able to help minimise issues associated with latency through intelligent task visualisation and/or applying operator assist technologies.

5.2.2 Audio

The bulldozer operator is suspected of primarily using audio as an assisting cue in identifying machine loading. The advantage of hearing is that it is highly sensitive to temporal structure, i.e. humans can pick up patterns and fine changes in structure. However, as Kelley (1964) notes, ‘hearing is not a spatial sense’ and this limits the extent of information it can provide. The experimental results obtained do not show the provision of audio to be a key determinant of performance.

Operator feedback does indicate that audio is a helpful cue and allows information obtained through other senses to be reconciled without significant mental effort. This is consistent with the observation of Macadam (2003) that, ‘auditory information is generally seen to be most beneficial when acting as a supplementary cue within a multi-channel environment’.

High quality in cab audio is also suspected of supporting operator engagement. The benefits related to operator acceptance and engagement, and the limited complexity involved suggest that audio should be provided to the operator at the teleoperation station.

5.2.3 Motion

Isolated testing of the influence of individual motion cues did not reveal any significant influence on task completion time. However, through examining the progress of the experimental task according to how much volume is removed over time there appears to be a benefit of motion during the initial stage of the task. The pitch only motion configuration provided to Operator A resulted in an average volume per push for the first five pushes that was 19.6% higher than the average achieved in the conventional (static) configura-

tion. Similarly, the coupled 3D vision and full motion feedback provided to Operator B resulted in an average volume per push for the first five pushes that was 21.9% higher than the average achieved in the conventional (static) configuration. It is during the first five pushes of the task in which the control activity is primarily focused on deep cut trajectories and high volume pushes. These results suggest that motion cues may be useful to the operator at this time in supporting command decisions related to appropriately controlling the bulldozer orientation for productive cutting and pushing.

When interaction with motion feedback was examined for the enhanced visual configuration as part of the second stage of testing, it was observed that the addition of motion feedback provided no distinct added benefit in task completion time. The provision of isolated pitch motion feedback in the enhanced visual configuration in fact appeared to be detrimental to performance indicating a possible interaction potentially caused by mismatch between observed visual cues and motion cues when visual quality is high. As noted by McRuer and Krendel (1965), ‘Motion effects which conflict with the visual modality can cause illusions which distort the pilot’s perception of the state of affairs. These can be so severe as to affect the pilot’s control capability.’

Both operators commented and indicated in surveys that they were able to obtain useful information from the motion feedback cues and generally found that motion improved their engagement and comfort while at times increasing their perceived workload. For Operator B, the combination of full motion feedback with the high quality 3D vision implementation resulted in a significant reduction in the frequency of joystick corrections suggesting lower actual workload.

If it is not feasible to provide motion feedback as part of the teleoperation system, it may be possible to close the same feedback loops through alternate means to a reduced degree. Orientation of the bulldozer can be accurately represented in visualisation or through overlay in the visual feed. The use of motion by the operator to build up a mental terrain model may similarly be assumed by the visual representation of terrain. However, operators commented that where features in the terrain visualisation were suspected of being incorrect, motion feedback provided secondary evidence against which their suspicions could be tested.

5.2 Perception groupings

5.2.4 Task visualisation

Tests conducted where the task visualisation was not provided to the teleoperator revealed its critical role in supporting effective completion of the task and reducing perceived workload. The task visualisation is able to encapsulate relatively complex information and present it to the operator in a format that is readily interpretable. The task visualiser enhances the operator's preview capability. As Macadam (2003) notes in the context of examining human driver models, 'the use of preview allows the human driver to not only provide anticipatory control responses for upcoming or developing driving scenarios, but to also conduct certain planning activities in response to developing situations.'

The provision of motion and enhanced vision at the teleoperation station may support a sensation consistent with on board operation but the lag between perception and the opportunity to make a correction remains. By contrast, a task visualisation which shows the slot surface contour may in fact allow the operator to determine the appropriate control to correct for deviations from desired slot geometry. Furthermore, the task visualisation provides a secondary and corroborating reference for other observed and or sensed cues.

The observations regarding the task visualisation suggest that virtual replications of the scene are of significant use in the teleoperation of bulldozers. Operator feedback suggests that simplicity is an important characteristic. As highlighted earlier in this section, there may be potential for addressing latency issues by enhancing the operator's ability to foresee required actions and plan ahead through intelligent visualisation design. It is possible that task visualisation may be further enhanced by introducing prediction of vehicle and environment states. Predictive displays have been shown to be highly effective in assisting manual control (Kelley, 1964). Kelly et al. (2011) have demonstrated the effectiveness of using a completely virtualized view provided to the teleoperator with predictive positioning of the remote vehicle to compensate for latency. Endsley's model of Situational Awareness also suggests that it may be desirable to provide information on the future state and that this would be of particular use to inexperienced operators. In general, the key objective should be to present information in terms that are closely aligned to the major goals of the operator.

5.3 Ranking perception requirements

The broad effects noted for each perception grouping permit a superficial ranking that is specific to the task conducted. Attempts have been made to rank perception requirements in other control activities.

Macadam (2003) provides a top-to-bottom ranking of sensory channels used in driving:

1. Vision
2. Vestibular and Kinesthetic
3. Tactile
4. Auditory

Kelley (1964) lists the senses most relevant to manual control. Listed in order from inner loop to outer loop:

1. Tactual-kinesthetic senses
2. Balance
3. Hearing
4. Vision

In the context of bulldozer teleoperation, it would appear that vision is the most critical component in the feedback provided to the operator during most operation tasks. Vision provides concise and readily interpretable cues that, it is hypothesised, allow the operator to improve their feedforward control of the bulldozer.

Task visualisation provides a very effective way of bridging inadequacies in the visual feedback implementation. The task visualisation would appear to be used in a similar feed forward fashion but at a higher strategic level.

The replication of motion does appear to be less significant than vision but still useful to the remote operator. However, not all aspects of motion are equally useful. It would appear that motion is used in relatively tight feedback loops to control the bulldozer's

5.3 Ranking perception requirements

orientation and loading. Based on the dynamics of the task, pitch is a key cue for controlling the dig trajectory, levelling and dumping operations. Vibration provides a cue that is useful both in providing passive feedback on terrain roughness, and thereby informing models of material behaviour, as well as highlighting the occurrence of track slip; a timely identification of which can enable a rapid correction. Translation provides cues that may highlight velocity changes that are not immediately apparent visually. However, high visual quality may degrade the relevance of this. Roll feedback informs the operator's ability to reject lateral level disturbances. Yaw feedback is similar in that the task conducted does not require yaw control but rather just disturbance rejection.

The results suggest that audio feedback provides no direct influence on performance. However, it does serve as a confirmatory cue on many other important inputs. This facilitates resolution of ambiguous cues, improving the efficiency of operator command selection.

In summary, the following ranking of perception requirements is proposed for enhancing teleoperation performance of bulldozers:

1. Vision
2. Goal oriented task visualisation
3. Motion
4. Audio

Conclusions and recommendations for further investigation

6.1 Conclusion

This thesis has sought to answer the key research question: *what feedback attributes are critical to maximising bulldozer teleoperation performance and user acceptance?* This question has been investigated experimentally through application of an enhanced perception cell specifically developed to augment and provide control over the perception feedback available to the teleoperator.

The enhanced perception cell provided the means to explore the influence on the bulldozer teleoperator of individual perception cues as well as their combined effect. The key perception groupings examined were visual quality, audio quality, motion feedback and task visualisation. A structured slot bulldozing task was executed by the operator under varying perception configurations. Task completion time, volume removal progress, joystick activity and perceived workload scores were measured. Task observation and operator feedback was also collected to inform the development of a hierarchical task analysis and control behaviour model.

The experimental results obtained and analysis performed suggest the following ranking in order of degree of influence on operator performance and acceptance:

1. Vision

2. Goal oriented task visualisation
3. Motion
4. Audio

The high resolution 3D visual configuration was found to provide the most significant improvement in performance as measured based on task completion time. Operator B achieved a reduction in average completion time of approximately 12%. The influence of a simple visualisation closely aligned to the task under execution was found to be critical to successful completion of the task. In addition, the task visualisation significantly reduced operator perceived workload. Motion feedback was found to provide some benefits around operator engagement and performance when executing heavy cutting. Operator A achieved an average volume per push for the first five pushes in the pitch only motion configuration that was 19.6% higher than the average achieved in the conventional (static) configuration. Operator B similarly achieved an average volume per push for the first five pushes in the 3D vision and full motion feedback configuration that was 21.9% higher than the average achieved in the conventional (static) configuration. However, no additional benefit was observed in terms of task completion time when motion was combined with the enhanced visual configuration. The results indicate that audio feedback is beneficial in terms of engagement and supporting the operator's ability to reconcile and corroborate other cues but has no observable impact on task performance.

A number of limitations of the current study are acknowledged. Due to availability constraints, the experiments were conducted with only two operators with limited overlap in the configurations investigated. This limits any understanding regarding how perception requirements might vary subtly or substantially from operator to operator. In addition, the findings are specific to the structured slot bulldozing task described in Section 3.4.2. While the learning obtained can be generalised to some extent given the different dimensions involved in the structured task, there may be unique requirements that appear due to the specific characteristics of a particular task. An attempt is made to support generalisation of results through the task analysis and behaviour modeling described in Chapter 4.

Recognising these limitations, the study was able to establish a correlation between perception capability and the various dimensions of performance in teleoperation. For one operator in the study, the 3D visual configuration with and without motion enabled per-

6.2 Original contributions of the thesis

formance at a level equivalent to on board operation. This finding is useful not only in the information it provides around what perception characteristics are required to achieve high levels of performance but also in demonstrating that under certain conditions, teleoperation can match on board performance.

6.2 Original contributions of the thesis

This thesis presents an experimental investigation of the influence of key feedback pathways on bulldozer teleoperation performance and user acceptance. It is believed that this experimentation is original in that no prior studies have examined a control element of this type. The study is also unique in the examination of a teleoperation platform for which operators have extensive on board operation experience. This is in contrast to similar teleoperation studies where the controlled element is a form of unmanned vehicle for which on board operational experience is not obtainable.

The experiments conducted have identified the relative contributions of distinct perception cues on a range of performance metrics. The thesis has also presented a theoretical model in Section 4.6 that attempts to explain the link between feedback pathways and task control behaviour/decisions.

These two avenues of investigation have been brought together to deliver a consolidated set of recommendations on perception requirements for the bulldozer teleoperator.

6.3 Recommendations for future work

The results presented and recommendations made around perception requirements suggest both further confirmatory and exploratory work.

The following priority areas for future work are suggested:

1. Isolated testing of individual cues on operator performance when executing a repeatable virtual task, i.e. error rejection or trajectory tracking.
2. Testing with multiple levels of video resolution to determine appropriate point of balance between teleoperator performance/acceptance and transmission loading.
3. Testing of predictive virtual displays as a means to address latency issues in the visual feedback transmission path.

4. Testing of the influence of control latency on teleoperator performance/acceptance and transmission loading.
5. Evaluation of the system models put forward in Chapter 4 using a bulldozer dynamic model from a control theory perspective along the lines of Weir and McRuer (1968).

Several limitations of the current study are acknowledged and should be remedied in any further work. Specifically, future studies should consider:

1. Execution of experiments with multiple operators to examine operator-to-operator variability.
2. Execution of experiments involving the performance of various other bulldozer tasks.

6.4 Concluding remarks

The investigation of the influence of perception on operator performance in bulldozer teleoperation was motivated by the potential opportunity to remove operators from significant operating hazards. The underlying technology to achieve bulldozer teleoperation exists and is viable. However, the degradation in performance that is often associated with teleoperation presents an obstacle to its broad uptake in the mining industry.

The teleoperator must be provided with some replication of the cues that are necessary to constructing the feedback loops that allow for effective control of the bulldozer. The experiments conducted in this thesis have attempted to reveal where interfaces should prioritise effort in cue replication. Results from these experiments have also informed the development of a perception-control model that, it is hoped, will be useful to those considering the role of perception in bulldozer operation.

It is hoped that the information revealed by this study supports the maturation of bulldozer teleoperation technology and broader utilisation within the industry. The benefit is in the potential to free operators from the risks of the operating environment while still allow them to work at their full potential.

Bibliography

- James L. Adams. An investigation of the effects of the time lag due to long transmission distance upon remote control: Phase ii - vehicle experiments, phase iii - conclusions. Technical Report D-1351, NASA, 1962.
- John Annett. *Hierarchical Task Analysis*. CRC Press, 2010.
- K. Beykirch, F. M. Nieuwenhuizen, H. J. Teufel, H.-G. Nusseck, J. S. Butler, and H. H. Blthoff. Control of a lateral helicopter side-step maneuver on an anthropomorphic robot. In *AIAA Modeling and Simulation Technologies Conference and Exhibit*, pages 1–8, 2007.
- Shane A. Bowen, Brian P. Oakley, and John S. Barnett. Effects of motion on skill acquisition in future simulators. Technical report, United States Army Research Institute for the Behavioural and Social Sciences, 2006.
- Caterpillar. *Operation and Maintenance Manual: D8T Track-Type Tractor*, 2009.
- J.Y.C. Chen, E.C. Haas, and M.J. Barnes. Human performance issues and user interface design for teleoperated robots. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 37(6):1231–1245, 2007.
- Douglas W. Cunningham, Vincent A. Billock, and Brian H. Tsou. Sensorimotor adaptation to violations of temporal contiguity. *Psychological Science*, 12(6):532–535, 2001.
- Luis de Florez. True blind flight. *Journal of the Aeronautical Sciences*, 3(5):168–170, 1936.
- D Drascic, Paul Milgram, and J Grodski. Learning effects in telemanipulation with monoscopic versus stereoscopic remote viewing. In *IEEE International Conference on Systems, Man and Cybernetics, 1989*, pages 1244–1249, 1989.

- Richard Edmondson, J. Larry Pezzaniti, Justin Vaden, Brian Hyatt, James Morris, David Chenault, Andrew Bodenhamer, Bradley Pettijohn, Joe Tchon, Tracy Barnidge, Seth Kaufman, David Kingston, and Scott Newell. 3D display for enhanced tele-operation and other applications. *Proceedings of the Society of Photo-Optical Instrumentation: Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV*, 7690:76901D–1–76901D–10, 2010.
- Mica R. Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1):32–64, 1995.
- Manuel Ferre, Rafael Aracil, and Manuel Navas. Stereoscopic video images for telerobotic applications. *Journal of Robotic Systems*, 22(3):131–146, 2005.
- P Robuffo Giordano, H Deusch, J Lächele, and H H Bühlhoff. Visual-vestibular feedback for enhanced situational awareness in teleoperation of UAVs. In *66th American Helicopter Society International Annual Forum 2010*, pages 2809–2818, 2010.
- A. Halme, J. Suomela, and M. Savela. Applying telepresence and augmented reality to teleoperate field robots. *Robotics and Autonomous Systems*, 26:117–125, 1999.
- Sandra G. Hart and Lowell E. Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati, editors, *Human Mental Workload*, volume 52, pages 139–183. North Holland Press, Amsterdam, 1988.
- Robert R Hoffman and Laura G Militello. *Perspectives on cognitive task analysis: historical origins and modern communities of practice*. CRC Press, 2008.
- R. Hosman and H. Stassen. Pilot’s perception in the control of aircraft motions. *Control Engineering Practice*, 7(11):1421–8, 1999.
- R. Hosman and J.C. van der Vaart. Thresholds of motion perception. In *Twelfth Annual Conference on Manual Control*, pages 956–984, 1976.
- R. Hosman and J.C. van der Vaart. Effects of vestibular and visual motion perception on task performance. *Acta Psychologica*, 48:271–287, 1981.
- Davood Karimi and Danny Mann. Role of motion cues in straight-line driving of an agricultural vehicle. *Biosystems Engineering*, 101(1-3):283–292, 2008.

BIBLIOGRAPHY

- Charles R. Kelley. Manual control. Technical Report AD449586, Engineering Psychology Branch Office of Naval Research, 1964.
- Alonzo Kelly, Nicholas Chan, Herman Herman, Daniel Huber, Robert Meyers, Pete Rander, Randy Warner, Jason Ziglar, and Erin Capstick. Real-time photorealistic virtualized reality interface for remote mobile robot control. *The International Journal of Robotics Research*, 30(3):384–404, 2011.
- Barry Kirwan and Les K Ainsworth. *A guide to task analysis: the task analysis working group*. Taylor and Francis, 1992.
- Gary A Klein. A recognition-primed decision (RPD) model of rapid decision making. In Gary A Klein, Judith Orasanu, Roberta Calderwood, and Caroline Zsombok, editors, *Decision making in action: Models and methods*. Ablex Publishing, 1993.
- Corinna Lathan and Michael Tracey. The effects of operator spatial perception and sensory feedback on human-robot teleoperation performance. *Presence*, 11(4):368–377, 2002.
- Sangyoon Lee and Gerard Jounghyun Kim. Effects of haptic feedback, stereoscopy, and image resolution on performance and presence in remote navigation. *International Journal of Human-Computer Studies*, 66:701–717, 2008.
- William H. Levison. Use of motion cues in steady-state tracking. In *Twelfth Annual Conference on Manual Control*, pages 895–917, 1976.
- Rong Liu and Max Meng. Acoustic display for navigation in internet-based teleoperation. In *IEEE International Conference on Intelligent Robots and Systems*, pages 4161–4165, 2005.
- Rong Liu and Yong-Xuan Wang. Auditory feedback and sensory substitution during teleoperated navigation. In *IEEE Transactions on Mechatronics*, pages 680–686, 2012.
- Salvatore Livatino, Giovanni Muscato, S. Sessa, C. Koffel, C. Arena, A. Pennisi, D. Di Mauro, and E. Malkondu. Mobile robotic teleguide based on video images. *Robotics Automation Magazine, IEEE*, 15(4):58–67, 2008.
- Charles C Macadam. Understanding and modeling the human driver. *Vehicle System Dynamics*, 40(1-3):101–134, 2003.

- Michael J Massimino. *Sensory substitution for force feedback in space teleoperation*. PhD thesis, Massachusetts Institute of Technology, 1992.
- John P. McIntire, Paul R. Havig, and Eric E. Geiselman. What is 3D good for? A review of human performance on stereoscopic 3D displays. *Proceedings of the Society of Photo-Optical Instrumentation: Head- and Helmet-Mounted Displays XVII; and Display Technologies and Applications for Defense, Security, and Avionics VI*, 8383: 83830X–1 – 83830X–13, 2012.
- Duane T. McRuer and Ezra S. Krendel. Mathematical models of human pilot behavior. Technical Report AGARD-AG-188, NATO Advisory Group for Aerospace Research and Development, 1965.
- Daniel M. Merfeld, Lionel Zupan, and Robert J. Peterka. Humans use internal models to estimate gravity and linear acceleration. *Nature*, 398:615–618, 1999.
- J. R. Murphy. Application of panospheric imaging to a teleoperated lunar rover. In *IEEE International Conference on Systems, Man and Cybernetics, 1995. Intelligent Systems for the 21st Century*, pages 3117–3121, 1995.
- Yasufumi Nagai, Shigeru Tsuchiya, Takashi Iida, and Shinichi Kimura. Audio feedback system for teleoperation experiments on Engineering Test Satellite VII: System design and assessment using eye mark recorder for capturing task. *IEEE Transactions on Systems, Man and Cybernetics*, 32(2):237–247, 2002.
- F. M. Nieuwenhuizen, K. Beykirch, M. Mulder, and H. H. Blthoff. Identification of pilot control behavior in a roll-lateral helicopter hover task. In *AIAA Modeling and Simulation Technologies Conference and Exhibit*, pages 1–11, 2007.
- Jesús Ortiz, Cecilia Tapia, Lorenzo Rossi, Jean-Guy Fontaine, and Mario Mazza. Description and tests of a multisensorial driving interface for vehicle teleoperation. In *Proceedings of the IEEE 11th International Conference on Intelligent Transportation Systems*, pages 616–621, 2008.
- Oxford English Dictionary. OED online, September 2013. URL <<http://www.oed.com/>>.
- Jens Rasmussen. Skills, rules, and knowledge: signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics*, 13(3):257–266, 1983.

BIBLIOGRAPHY

- R. F. Ringland and R. L. Stapleford. Experimental measurements of motion cue effects on stol approach tasks. Technical Report CR-114458, NASA, 1972.
- R. F. Ringland, R. L. Stapleford, and R. E. Magdaleno. Motion effects on an IFR hovering task: Analytical predictions and experimental results. Technical Report CR-1933, NASA, 1971.
- Bill Ross, John Bares, David Stager, Larry Jackel, and Mike Perschbacher. An advanced teleoperation testbed. *Field and Service Robotics*, 42:297–304, 2008.
- J. A. Schroeder. *Helicopter flight simulation motion platform requirements*. PhD thesis, Stanford University, 1998.
- David R. Scribner and James W. Gombash. The effect of stereoscopic and wide field of view conditions on teleoperator performance. Technical Report ARL-TR-1598, Army Research Laboratory, 1998.
- Rober Sekuler, Allison B. Sekuler, and Renee Lau. Sound alters visual motion perception. *Nature*, page 308, 1997.
- T. B. Sheridan and W. R. Ferrell. Remote manipulative control with transmission delay. *IEEE Transactions On Human Factors In Electronics*, pages 25–29, 1963.
- Thomas B. Sheridan. *Telerobotics, Automation, and Human Supervisory Control*. MIT Press, Cambridge, Massachusetts, 1992.
- Richard S. Shirley and Laurence R Young. Motion cues in man-vehicle control: Effects of roll-motion cues on human operator’s behaviour in compensatory systems with disturbance inputs. *Transactions on Man-Machine System*, 9(4):121–128, 1968.
- R. L. Stapleford, R. A. Peters, and F. Alex. Motion effects on an IFR hovering task: Analytical predictions and experimental results. Technical Report CR-1325, NASA, 1969.
- Jarno Uusisalo and Kalevi Huhtala. Analysis of effects of remote control on usability of hydraulic mobile machines. In *8th JFPS International Symposium on Fluid Power*, pages 688–695, 2011.
- Jarno Uusisalo, Kalevi Huhtala, and Matti Vilenius. Effects of remote control on usability of hydraulic excavator. In *ASME 2009 Dynamic Systems and Control Conference*, pages 321–328, 2009.

- Jacqueline Vivéash, Joanna White, Jenny Boughton, Stuart King, and Martin Kaye. Remote control of vehicles. In *RTO HFM Symposium on Spatial Disorientation in Military Vehicles: Causes, Consequences and Cures*, 2002.
- D. Weir and D. McRuer. Models for steering control of motor vehicles. In *Proceedings of the Fourth Annual NASA-University Conference on Manual Control*, pages 135–169, 1968.
- R Welch and D Warren. Intersensory interactions. In K Boff, L Kaufman, and J Thomas, editors, *Handbook of Perception and Human Performance*, chapter 25. John Wiley and Sons, 1986.

APPENDIX A

Experimental Record

The following tables contain the completion times recorded in each experimental run.

Table A.1: Completion time results - stage 1

(a) No motion feedback		(b) Combined motion feedback		(c) Pitch only motion		(d) Vibration only motion	
NMF	Completion Time, s	CMF	Completion Time, s	PMF	Completion Time, s	VMF	Completion Time, s
NMF-1	688	CMF-1	693	PMF-1	558	VMF-1	806
NMF-2	566	CMF-2	674	PMF-2	521	VMF-2	662
NMF-3	583	CMF-3	683	PMF-3	532	VMF-3	672
NMF-4	773	CMF-4	497	PMF-4	532	VMF-4	621
NMF-5	681	CMF-5	622	PMF-5	819	VMF-5	671
NMF-6	612	CMF-6	781	PMF-6	985	VMF-6	573
NMF-7	603	CMF-7	985	PMF-7	607	VMF-7	717
NMF-8	690	CMF-8	662	PMF-8	588	VMF-8	631
NMF-9	725	CMF-9	585	PMF-9	699	VMF-9	662
NMF-10	797	CMF-10	662	PMF-10	630		
NMF-11	669	CMF-11	577				
NMF-12	711	CMF-12	673				
NMF-13	685	CMF-13	981				
NMF-14	660	CMF-14	754				
NMF-15	641	CMF-15	761				

Table A.2: Completion time results - stage 1 continued

(a) Translation only motion		(b) On board operation	
TMF	Completion Time, s	OB	Completion Time, s
TMF-1	814	OB-1	526
TMF-2	794	OB-2	469
TMF-3	722	OB-3	533
TMF-4	739	OB-4	544
TMF-5	765		
TMF-6	663		
TMF-7	678		
TMF-8	564		
TMF-9	543		
TMF-10	491		

Table A.3: Completion time results - stage 2

(a) Standard resolution video			(b) High resolution video			(c) 3D high resolution video			(d) 2DHR without audio		
2DSR	Completion Time, s		2DHR	Completion Time, s		3DHR	Completion Time, s		2DHR-A	Completion Time, s	
2DSR-1	738		2DHR-1	458		3DHR-1	438		2DHR-A-1	531	
2DSR-2	726		2DHR-2	828		3DHR-2	704		2DHR-A-2	919	
2DSR-3	677		2DHR-3	774		3DHR-3	546		2DHR-A-3	557	
2DSR-4	658		2DHR-4	560		3DHR-4	524		2DHR-A-4	640	
2DSR-5	495		2DHR-5	661		3DHR-5	669		2DHR-A-5	1018	
2DSR-6	590		2DHR-6	688		3DHR-6	579		2DHR-A-6	597	
2DSR-7	740		2DHR-7	440		3DHR-7	631		2DHR-A-7	630	
2DSR-8	620		2DHR-8	905		3DHR-8	590		2DHR-A-8	635	
2DSR-9	780		2DHR-9	683		3DHR-9	552		2DHR-A-9	674	
2DSR-10	760		2DHR-10	595		3DHR-10	656		2DHR-A-10	862	
2DSR-11	759		2DHR-11	659		3DHR-11	403		2DHR-A-11	559	
2DSR-12	685		2DHR-12	425		3DHR-12	560		2DHR-A-12	682	
2DSR-13	443		2DHR-13	565		3DHR-13	445		2DHR-A-13	471	
2DSR-14	727		2DHR-14	695		3DHR-14	573		2DHR-A-14	591	
2DSR-15	604		2DHR-15	458		3DHR-15	557		2DHR-A-15	660	
2DSR-16	532		2DHR-16	475							
2DSR-17	534		2DHR-17	699							
2DSR-18	579		2DHR-18	427							
2DSR-19	562		2DHR-19	526							
2DSR-20	577		2DHR-20	899							

Table A.4: Completion time results - stage 2 continued

(a) 3D video with pitch motion		(b) 3D video with full motion		(c) On board operation	
3DHR+PMF	Completion Time, s	3DHR+FMF	Completion Time, s	OB	Completion Time, s
3DHR+PMF-1	592	3DHR+FMF-1	553	OB-1	513
3DHR+PMF-2	854	3DHR+FMF-2	479	OB-2	611
3DHR+PMF-3	679	3DHR+FMF-3	539	OB-3	654
3DHR+PMF-4	544	3DHR+FMF-4	533	OB-4	663
3DHR+PMF-5	455	3DHR+FMF-5	484	OB-5	662
3DHR+PMF-6	608	3DHR+FMF-6	615	OB-6	522
3DHR+PMF-7	504	3DHR+FMF-7	629	OB-7	479
3DHR+PMF-8	588	3DHR+FMF-8	575	OB-8	486
3DHR+PMF-9	654	3DHR+FMF-9	609	OB-9	575
3DHR+PMF-10	733			OB-10	651
3DHR+PMF-11	668			OB-11	598
3DHR+PMF-12	531			OB-12	478
3DHR+PMF-13	593			OB-13	499
3DHR+PMF-14	484			OB-14	616
3DHR+PMF-15	580			OB-15	556
3DHR+PMF-16	1070			OB-16	563
				OB-17	572
				OB-18	573
				OB-19	471
				OB-20	491