Towards Discrimination of Challenging Conditions for UGVs with Visual and Infrared Sensors

Christopher Brunner, Thierry Peynot and James Underwood
ARC Centre of Excellence for Autonomous Systems
Australian Centre for Field Robotics
The University of Sydney, NSW 2006, Australia
c.brunner/tpeynot/j.underwood@acfr.usyd.edu.au

Abstract

This work aims to contribute to reliability and integrity in perceptual systems of autonomous ground vehicles. Information theoretic based metrics to evaluate the quality of sensor data are proposed and applied to visual and infrared camera images. The contribution of the proposed metrics to the discrimination of challenging conditions is discussed and illustrated with the presence of airborne dust and smoke.

1 Introduction

This work presents a step towards understanding and developing integrity in perceptual systems for Unmanned Ground Vehicles (UGVs) in unknown and unstructured environments. UGVs are generally equipped with one or a combination of exteroceptive sensors, such as laser and radar range sensors and colour and infrared cameras, which are used to perceive the environment. The perception systems of UGVs use various algorithms with the sensor data to provide a meaningful interpretation of the environment, by identifying features such as obstacles, slopes and moving objects, which are then used for safe navigation. To achieve long duration of autonomous operation (the order of days), high integrity perception systems are a requirement. In addition, the scope of perception failures is unbounded, with perceptual failures causing catastrophic consequences.

Traditional fault detection focuses on simple hardware failures where sensors provide corrupted (or no) data. Simple, low-level sensor data errors such as noise are well known effects and are properly modeled and handled in data fusion methods such as SLAM [Durrant-Whyte and Bailey, 2006]. A more significant and less systematic category of error occurs when sensed data are misinterpreted in the perception system, generating an inaccurate representation of the environment. For example, a simplistic perception system utilising a laser range scanner with the task of detecting solid objects will interpret dust and smoke as equivalent to solid objects. This is not a sensor fault according to the specifications of the sensor, rather, an inability and failure of the perception system design to accurately model the environment from the perception data. Therefore, perception systems require a more general, higher-level view of failure where errors in the interpretation of data by algorithmic models need to be detected. Indeed, unless all possible environmental conditions are explicitly modeled and interpreted, failure can still occur if a sensor is not discerning the environment as anticipated by the perception module.

The importance of reliable perception in outdoor environments is exemplified in the DARPA PerceptOR program [Kelly *et al.*, 2006] where environmental effects had significant affect on the performance of a state-of-the-art UGV. In PerceptOR, the vast majority of poor performance conditions could be described as sensor interpretation errors due to unmodeled environmental conditions and reliable perception was identified as the performance bottleneck for the whole system. Dust, smoke, rain and fog are recognised as conditions that can cause UGV perception systems to misinterpret the environment [Narasimhan and Nayar, 2003a]. A few partial solutions have been studied in the literature. In the case of cameras, software options include removing or reducing the effects of specific weather conditions from visual images [Narasimhan and Nayar, 2003b], [Nayar and Narasimhan, 1999], [Tan, 2008]. A hardware option for lasers was found with the Ibeo LUX Laserscanner [Ibeo, 2009] which uses multiple returns to reduce the impact of incorrect range readings caused by reflections from rain or snow, although it will still return readings corresponding to thick dust as if they corresponded to hard obstacles. However, these examples target specific instances of environmental misinterpretation for particular applications and general perception problems are still largely unsolved for autonomous perception systems [Urmson *et al.*, 2008] [Leonard *et al.*, 2008].

When interpreting complex outdoor environments,

there is no single physical sensor that can be regarded as perfectly perceiving the environment at all times and in every situation. Therefore, a robust perception system must combine the strengths of a range of sensors, thus mitigating the individual sensor weaknesses. However, if data from multiple sensors are simply fused at a low level, failure will occur when the sensors data are inconsistent. In order to achieve long-term autonomy with a high level of integrity, the response of the sensors to all environmental conditions must be managed as part of the interpretation of the data. Common (and different) characteristics between sensed data must be fused in an intelligent way to discriminate and handle challenging environmental conditions and identify failure states in the perception system.

A simplified general perception block is shown in Figure 1. The figure shows the perception block taking measurements of the environment ε with two sensors and then fusing and interpreting that data to produce an internal representation of the environment α that is suitable for the application. A real world example is the use of two range sensors, like a laser and a radar. In that case, the purpose of the fusion is to make an array of 3D data points using both sets of sensor data. This 3D point cloud is then interpreted, typically to identify obstacles for the purpose of path planning. Another example, as considered in this paper, is the association of a colour camera and an infrared camera.

A failure will occur when data with inconsistencies are fused, causing an inaccurate representation of the environment α. It may be possible to detect such failures at point α, although discriminating the cause of the failure is likely to be difficult due to the multiple sources of input. The earlier that it is possible to detect a failure, the more solutions that become available to adapt or recover from the failure. Metrics may be developed to detect sources of failure by checking the consistency of data between sensors A and B prior to fusion of the data. Alternatively, there may be characteristics of data at A or B that are likely to generate a failure of interpretation.

The reliability and integrity of the perception model may be improved by exploiting information theory to identify general features of known challenging conditions that may cause failure situations. This work proposes to test some information theoretic metrics as applied to perception data from visual and infrared cameras in the presence of smoke and dust to help discriminate challenging conditions for perception. Information theory is chosen due to its potential to work on arbitrary combinations of sensors. This paper evaluates a number of techniques that may be used on a whole class of sensors and discusses metrics that could be used to determine the quality of perception data.

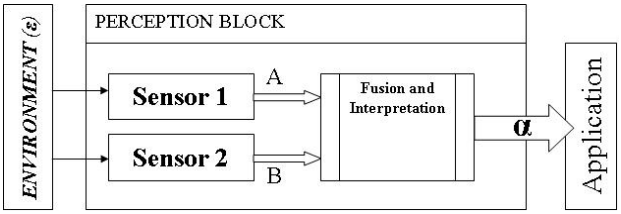


Figure 1: Simplified Perception Block. Sensors 1 and 2 provide data from the environment, A and B respectively. The sensed data is fused and interpreted to provide an internal representation of the external environment at α, which is suitable for the required application.

This article is structured as follows. Section 2 presents the multi-sensor UGV that was used for this work and describes the specific data sets that were chosen to be representative of challenging perception conditions for consideration in this paper. Section 3 introduces redundancy and failure detection as in the context of perception systems. Section 4 discusses information theory and how it may be used in discriminating challenging perception conditions and presents results of information metrics for representative data sets. Finally, Section 5 provides a conclusion and suggests different directions for future work.

2 Experimental Setup

The experimental study proposed in this paper is based on public data sets presented in more details in [Peynot *et al.*, 2009a]. These data were gathered by a UGV equipped with a wide variety of sensors, including laser range scanners, a radar, a colour camera and an infrared (IR) camera, in controlled and variable environmental conditions. These included challenging conditions such as the presence of airborne dust, smoke or rain. Previous work using those data sets has shown that airborne dust and smoke have a significant effect on perception using laser scanners [Peynot *et al.*, 2009b]. The study in this paper focuses on the effect of these environmental elements on visible and IR camera images. Details on those cameras are the following:

- a Prosilica 1360×1024 resolution mono-CCD colour camera, acquiring images at a nominal frame rate of 15 images per second;
- a Raytheon thermal Infrared camera, with a spectral response range of $7-14\mu m$ (average frame rate: 12.5 images per second).

Although the conclusions of this study were drawn from the analysis of various data sets, the illustrations in this paper will focus on three particular representative sequences of images, which are referred to as data sets 1: *Clear* (clear conditions), 5: *Dust* (presence of airborne

dust) and 7: *Smoke* (presence of smoke) in [Peynot *et al.*, 2009a]. The corresponding data were collected from a static vehicle, observing a 'reference' scene (Figure 2) where the positions and sizes of objects are known (see [Peynot *et al.*, 2009a]).

Figure 2: The UGV sensing the trial area

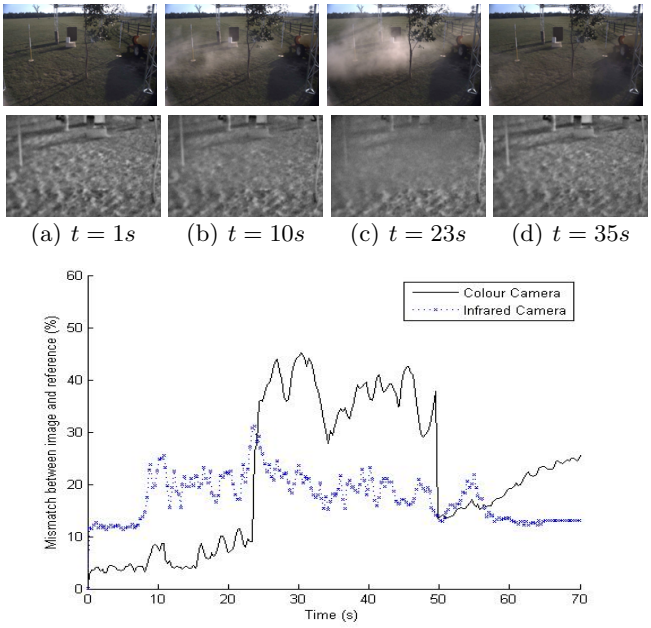
Therefore, for the purpose of this paper, a crude metric evaluating the quality of the data can be computed by making a direct comparison of the R,G,B pixel values at any time step with an average of these values over four early reference images when there are no known environmental impacts. In practice, a pixel is considered as matching the reference average if the distance between the corresponding points in the 3D (R,G,B) space is lower than a pre-defined threshold, chosen to account for noise in the data.

Figures 3 and 4 show the percentage of pixels per scan that do not match the reference image over the course of the data captured for the *Dust* and *Smoke* data sets. Each figure has four representative images from both the colour and infrared cameras (that will be used throughout this paper) to demonstrate characteristic changes in the environment over the course of data collection. The introduction of dust or smoke causes an important increase in the number of pixels that are significantly not matching the reference, which can be visually confirmed in the images.

While dust causes change in both the visual and infrared images, smoke impacts the visual image but has no observable effect on the infrared image due to the particle size of smoke not affecting infrared wavelengths. This extends to the results of Section 4, where smoke has no noticeable effect on any infrared analysis.

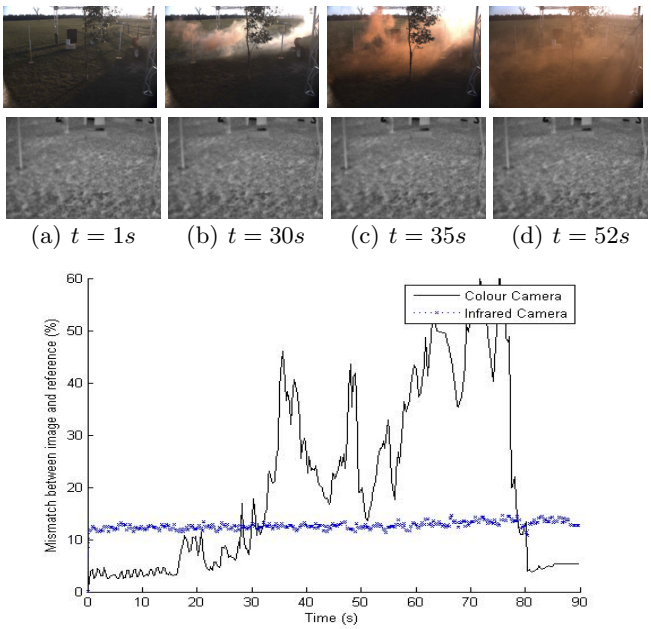
3 Multi-Modal Redundancy for Failure Detection in Perception Systems

Physical redundant sensor systems are used to diagnose hardware failures in many applications [Gertler, 1998].



(a) $t = 1s$ (b) $t = 10s$ (c) $t = 23s$ (d) $t = 35s$

Figure 3: Representative colour (top) and infrared (bottom) images of Data Set 5 (*Dust*) and the percentage of the image that matches a ground truth reference image over time. a) Reference image at $t = 1s$; b) Light dust cloud at $t = 10s$; c) Thick dust cloud at $t = 23s$; d) Very light dust covering most of image at $t = 35s$.



(a) $t = 1s$ (b) $t = 30s$ (c) $t = 35s$ (d) $t = 52s$

Figure 4: Representative colour (top) and infrared (bottom) images of Data Set 7 (*Smoke*) and the percentage of the image that matches a ground truth reference image over time. a) Reference image at $t = 1s$; b) Light smoke at $t = 30s$; c) Thick Smoke at $t = 35s$; d) Thick smoke covering most of scene at $t = 52s$.

In such examples, the observations from multiple sensors measuring the same property are directly compared to each other. For example, two temperature sensors can be directly compared to each other because they are measuring the same property in the same way. If sensors do not agree, then this is an indication that a fault has occurred. However, onboard exteroceptive UGV sensors are limited in physical redundant systems in two significant ways. Firstly, data gathered by differing exteroceptive sensors can never be considered directly redundant as data from a sensor is dependent on the spatial position of the sensor in relation to the environment. Physical limitations of the sensors mean that they can never be mounted in exactly the same position on a UGV. Two identical sensors will always perceive the environment from different vantage points and although the differences can be small with good alignment, significant difference in raw data is possible due to occlusions and relative spatial geometry. Therefore, a mismatch in a direct comparison between two identical sensors may not necessarily be indicative of failure. Secondly, if redundancy is achieved with multiple identical sensors, any impact due to the environment will have the same effect on all of the sensors, thus circumventing the redundancy as incorrect perception data are verified by multiple sensors. As mentioned in an earlier example, multiple laser range sensors will be similarly affected by dust.

There is benefit in a UGV equipped with a range of multi-modal perception sensors that measure the same underlying quantities using vastly different physical processes. Exploiting multi-modal redundancy will be more robust than simple hardware redundancy because the physical processes used by the sensors are less liable to fail simultaneously due to unforseen environmental phenomena. A low level example is in measuring the range to obstacles using laser and radar scanners as in [Peynot *et al.*, 2009b], where aligned laser and radar range sensors were directly compared as a method of removing laser data that had been affected by dust. A higher-level interpretation may involve comparing edges detected with an infrared camera to those found with a colour camera.

Multi-modal redundancy requires that different strategies be incorporated into the perception model to find metrics that can be compared between different sensors types. Finding commonality between multi-modal sensed data will provide redundancy to discriminate between situations that may cause perception failures. Information theory is expected to be a useful starting point in finding general features that are common in perception sensors in challenging conditions without relying on direct comparison of sensor data. A pairing of visual and infrared cameras is considered a good candidate for information theoretic redundancy metrics as both provide an image of the environment with visual similarities in geometry but they are intrinsically different in the properties of the environment they measure so cannot be directly compared.

4 Sensor Data Analysis Using Information Metrics

Information theory was established in the context of communication over noisy channels. The amount of information can be interpreted as "the size of a file that encodes the outcomes of a random experiment" [Mackay, 2007] but the concept can be generalised to deal with any data that adds to the knowledge of a situation. Specifically, information theory deals with the quantification of information. Generally, the greater the amount of information in a data set, the more knowledge that can be gained from the data.

In a UGV perception system, techniques may monitor the level of information provided by a sensor and evaluate the utility of raw sensor data for general perception systems with some independence from the algorithm that uses the data.

An analysis of the characteristics of perception data for the purpose of integrity will require various metrics that compare correlating information across a range of multi-modal sensors. Recognising common perception traits from multiple sensors and multiple metrics should be exploited to identify and discriminate possible challenging perception scenarios. Information is common regardless of the sensor and so it is proposed that a general framework based on information could be used to evaluate the quality of perception data.

In the following sections, this work proposes some information metrics for visual and infrared sensors and evaluates the potential to use these as part of a UGV perception system. The evolution of information theoretic measures such as Shannon Information, Spatial Perceptual Information (SI) and Temporal Perceptual Information (TI) of the raw sensor data in recognised challenging conditions is analysed. It is demonstrated that each method has some strength of discrimination that could be used in combination with other metrics to help improve the integrity of general UGV perception systems.

4.1 Shannon Information Measurement

The Shannon information content of an outcome x is defined to be [Mackay, 2007]:

$$h(x) = log_2 \frac{1}{p(x)} \tag{1}$$

It is measured in bits and is considered a natural measure of the amount of information content provided by a particular observation $x = a_i$ where a_i belongs to a

discrete set of possible observations ($a_i \in A_X$). The less probable an outcome is, the more information it provides when it is observed.

The average Shannon information content or entropy H of an ensemble, X, from a discrete set of possible observations ($x \in A_X$), is given by:

$$H(X) = \sum_{x \in A_X} P(x) log_2 \frac{1}{P(x)} \quad (2)$$

and is expressed in average bits of information per observation.

The more variety in a set of observations, the more information content contained within that set, that is, the most random set of observations contains the most information as it can be interpreted in any way and cannot be compressed. A uniform distribution of all possible observations will provide the maximum entropy that is possible for that set of observations.

For this work, the colour images were converted to three 8-bit images in hue, saturation and luminosity scales using the Matlab *rgb2hsv* function. The infrared images were received as 8-bit images in the luminosity scale. Thus, each resulting image yielded the discrete set of pixel values $A_X = [0, 255]$ and a maximum information content of 8 bits per pixel. The probabilities were calculated using a histogram. Given a single image, for each possible integer value $x \in A_X, P(x)$ was calculated by counting the number of pixels in that image with the value x, and dividing by the total number of pixels. The average Shannon information was then calculated using Equation (2).

When applied to single images, the entropy is a measure of the amount of information in an entire image, which relates to the variety in values of pixels. In video images, an increase in entropy over time means that something has changed in the image that is unique or unusual and has increased the distribution of pixel values. Conversely, if the entropy drops over time, this means pixel values that were rare are no longer in the image or have become less rare as more have appeared. Novel objects, because of the rareness of their pixel value compared to the rest of the image, are considered to provide relatively high levels of information so long as the pixel values remain rare. If those same objects become more common in the image, they are considered to provide less information as they are no longer novel.

The evolution of entropy calculations for the *Dust* data set are shown in Figures 5 and 6 and the *Smoke* evolution is shown in Fig. 7 and 8. The corresponding representative images show the relative localised pixel-by-pixel contribution to the entropy calculation; darker pixels are considered more common and thus provide less information content.

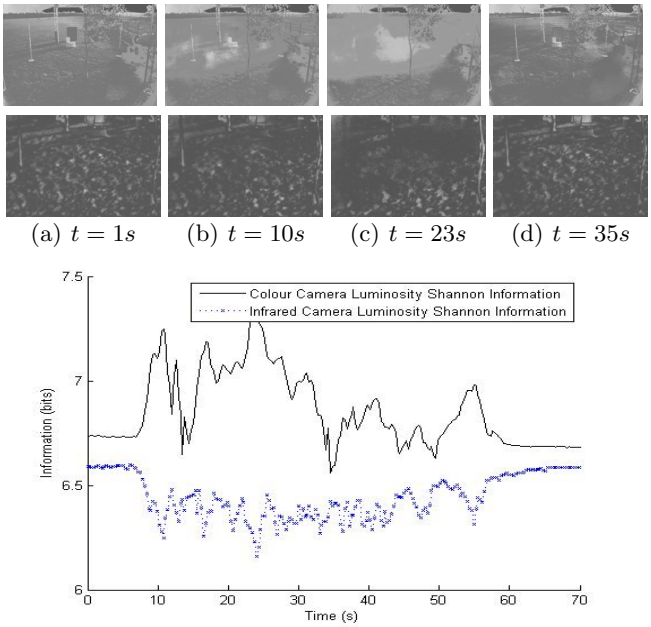


Figure 5: Average Shannon Information from the luminance images are plotted over time for *Dust*. The images show the corresponding local contribution to Shannon Information from the visible (top) and infrared (bottom) images.

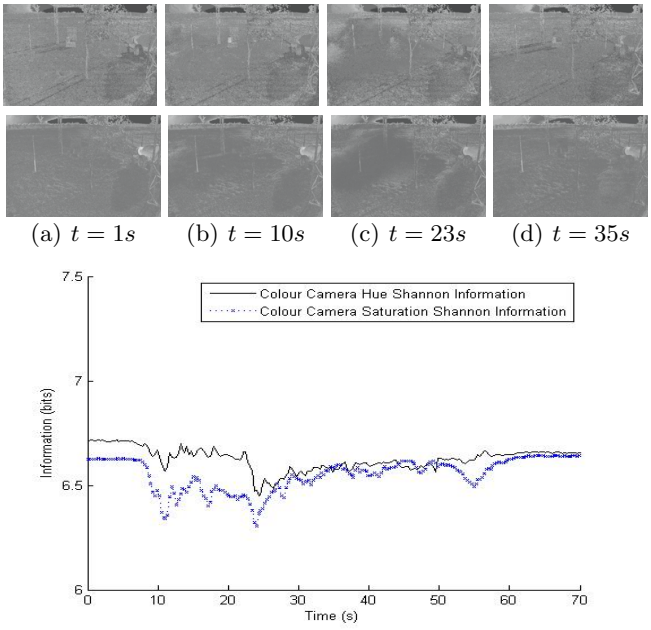
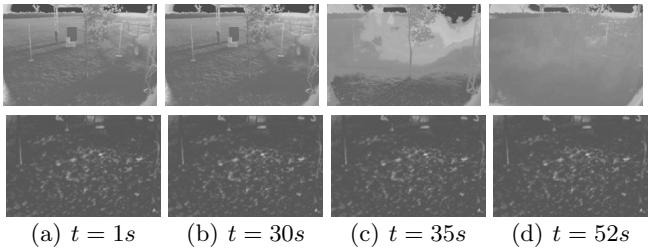


Figure 6: Average Shannon Information from the hue and saturation images are plotted over time for *Dust*. The images show the corresponding local contribution to Shannon Information from the hue (top) and saturation (bottom) images.



(a) $t = 1s$ (b) $t = 30s$ (c) $t = 35s$ (d) $t = 52s$

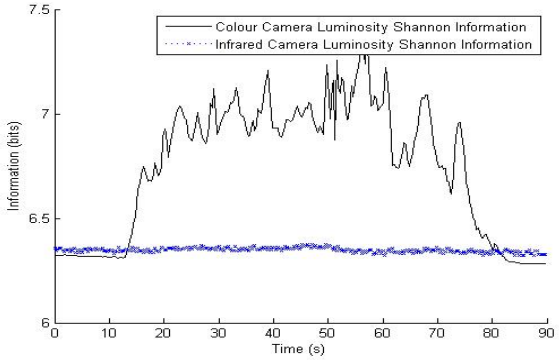
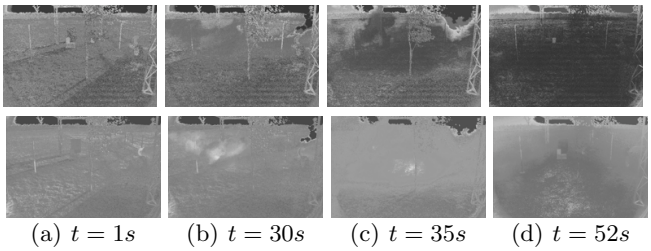


Figure 7: Average Shannon Information from the luminance images are plotted over time for *Smoke*. The images show the corresponding local contribution to Shannon Information from the visible (top) and infrared (bottom) images.



(a) $t = 1s$ (b) $t = 30s$ (c) $t = 35s$ (d) $t = 52s$

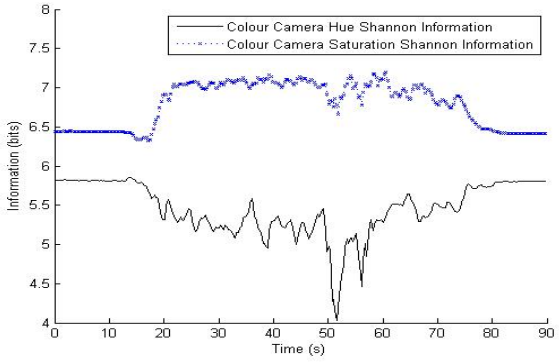


Figure 8: Average Shannon Information from the hue and saturation images are plotted over time for *Smoke*. The images show the corresponding local contribution to Shannon Information from the hue (top) and saturation (bottom) images.

Figures 5 and 7 show that in static clear environments (the first and last seconds of each data set), the Shannon information content remains relatively steady. In the visual images, the appearance of dust and smoke causes an increase in entropy. Although background areas are concealed by dust and smoke in the colour camera, in general dust and smoke clouds are considered to add Shannon information content in the visible spectrum because the mean value of intensity of luminosity is different from the general background scene and so the entropy of the total image increases.

However, in infrared images for *Dust*, the Shannon information consistently drops. This is the case because dust clouds are approximately the same average temperature (and thereby same average intensity in the infrared image) as the background that they occlude. As the relatively homogeneous (low entropy) dust cloud obscure the background, some features that are considered higher information are blocked by the cloud, thereby reducing the overall information content of the image.

Comparing Figures 6 and 8 shows that in *Dust* the Shannon information content of hue and saturation are slightly reduced whereas *Smoke* causes significant increases in saturation information content and significant reduction in hue information content. This was to be expected as the dust is in general similar to much of the background image whereas the smoke used in this data set is bright orange.

From the analysis of *Dust* and *Smoke* data sets, the entropy measurement appears to be a promising way to capture the evolution in environmental conditions and is likely to assist in discriminating some recognised challenging conditions. In these data sets, dust and smoke clouds consistently increase the level of entropy in the visual spectrum, while dust reduces entropy in infrared images and smoke has no effect in the infrared spectrum. The Shannon information may be used to evaluate the utility of sensor data for use in perception systems. However, without context, this metric is not robust enough as an individual measure, it should be used in combination with other information metrics.

4.2 Spatial Perceptual Information Measurement (SI)

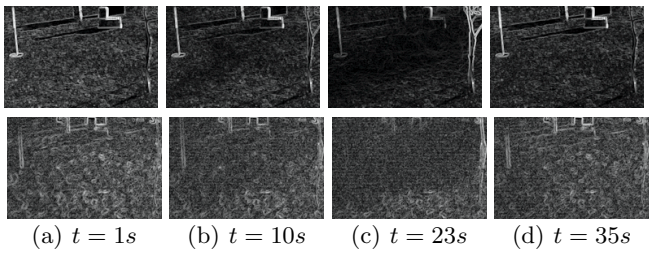
The spatial perceptual information measurement (SI) is used as a test method in the telecommunication industry to determine the level of impairment that is suffered when a scene is transmitted over a digital service channel [ITU-T, 1999]. Each video frame (luminance plane) at time n, (F_n), is filtered with the Sobel filter for edge detection and the standard deviation (std) over the pixel values in each Sobel-filtered frame is then computed.

$$SI_n = std[Sobel(F_n)] \tag{3}$$

This metric provides a measure of the spatial detail in an image by representing the distribution of texture in the image assuming a Gaussian spread and thus is similar to the entropy measurement of pixel intensity by taking higher values for the more spatially complex (or informative) scenes [Correia and Pereira, 2003].

A metric that measures the SI of an image could provide a measurement for the quality of data in an image. Streams of raw sensor data that contain very low spatial information may not be utilised effectively for some perception algorithms that rely on feature identification and may indicate perception or sensor failure. The method can be adapted to monitor the evolution in time of the spatial information by calculating the SI for each image and plotting the result as a time series.

For this work, the colour image was converted to an 8-bit greyscale image using the Matlab *rgb2hsv()* function and taking the luminosity component. The infrared image was received as an 8-bit image in the luminosity scale. The average spatial information was calculated using Equation (3) and with the Matlab *Sobel* edge detector function.



(a) $t = 1s$ (b) $t = 10s$ (c) $t = 23s$ (d) $t = 35s$

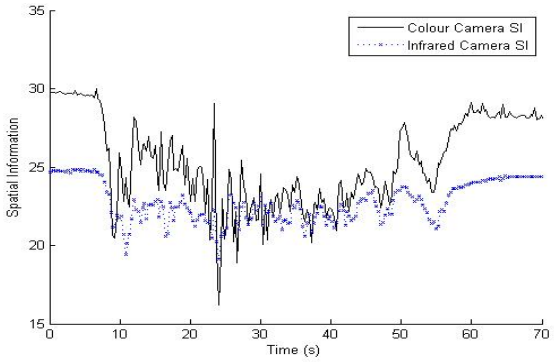
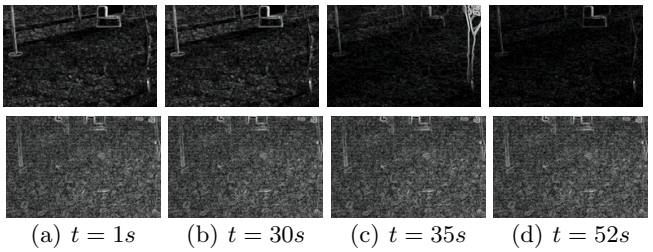


Figure 9: Two synchronised sets of visible (top) and thermal infrared (bottom) camera images are shown, after coarse alignment and application of the Sobel filter used to compute SI. The SI values from both the visible and infrared cameras are shown for *Dust*.

Figure 9 and 10 show the evolution in time of the spatial information in *Dust* and *Smoke* respectively. It should be noted that due to different outdoor lighting conditions at the time of data gathering, the initial value for spatial information is significantly less for *Dust* than



(a) $t = 1s$ (b) $t = 30s$ (c) $t = 35s$ (d) $t = 52s$

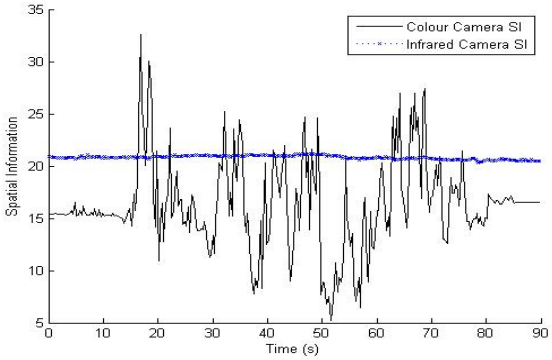


Figure 10: Two synchronised sets of visible (top) and thermal infrared (bottom) camera images are shown, after coarse alignment and application of the Sobel filter. The SI values from both the visible and infrared cameras are shown for *Smoke*.

for *Smoke*. Lower lighting conditions mean a lower contrast in the image and subsequently, minor edges, such as from the texture of the ground, are lost or diminished in filters such as the Sobel filter. In stationary environments (the first few seconds of each data set), the spatial information remains relatively constant

In *Dust*, there is a corresponding decrease in the level of spatial information with the introduction of airborne dust in both the infrared and the visual images. The individual images show that the background features in the environment are obscured (compare image a to b and c in Fig. 9) and in particular, minor edges that are produced by the textures in the environment can be lost, for example, the edges found on the ground are obscured completely by the dust. Although the edges of the tree in the foreground are enhanced with thick background dust contrasting with the edges of the tree (image c), the overall effect on the image is a drop in SI.

In *Smoke*, the general background features in the visual camera are similarly obscured as in *Dust* but the effect is not as pronounced for the total information measurement as many of the background edges that are being covered are already dulled in the lower light conditions. However, there is an additional effect from the orange colour of the smoke providing a strong contrast to the general background that means many edges are actually enhanced and the Sobel filter consequently identifies very

intense edges. Most notably, the edges of the tree on the right side of the image become very strong features and this is a major contributor to increased SI. At the point where smoke fills almost all of the scene ($t = 52s$, in Fig. 10 d), all objects are thickly shrouded in smoke and the SI drops significantly.

SI appears to be a promising way to monitor conditions where the level of detail in the image has changed. Images that are very low in SI may be considered unhelpful for feature or object identification. Sudden variations in SI could indicate that features are being obscured. Indeed, this study has shown that known challenging perception conditions such as dust and smoke tend to obscure edges in the environment and reduce SI. However, the level of SI can also vary depending on the environment and the lighting condition as edges are more likely to be detected when a scene has good contrast. SI could be a powerful discriminatory metric but will require further context.

4.3 Temporal Perceptual Information Measurement (TI)

The temporal perceptual information measurement (TI) [ITU-T, 1999] is based on the motion difference feature, $M_n(i, j)$, which is the difference between the pixel values (of the luminance plane) at the same location in space but at successive times or frames. This metric provides a measure of the amount of temporal changes in a video sequence, thus taking higher values for high motion sequences [Correia and Pereira, 2003]. It can be expressed as:

$$M_n(i, j) = F_n(i, j) - F_{n-1}(i, j) \tag{4}$$

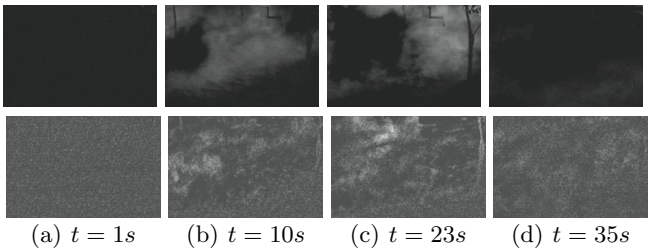
where $F_n(i, j)$ is the pixel at the i^{th} row and the j^{th} column of the n^{th} frame in time.

The measurement of Temporal information (TI) is computed as the standard deviation of the pixel values in $M_n(i, j)$ over all values of i and j.

$$TI = std[M_n(i, j)] \tag{5}$$

If frames are similar then TI will be close to zero. The more motion seen in adjacent frames results in a higher value of TI.

The temporal perceptual information measurement is a method of formalising change in time between images. Figures 11 and 12 show that in situations where the environment is stable, the TI value is very low, whereas with the introduction of dust and smoke the TI increases markedly. The TI measure provides an indication of whether there is a dynamic element to the scene. In this form, temporal information may not directly provide significant information when applied to mobile robotics as simple motion of the robot will always result in changes



(a) $t = 1s$ (b) $t = 10s$ (c) $t = 23s$ (d) $t = 35s$

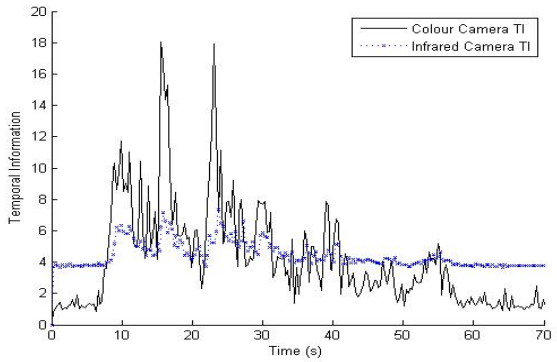
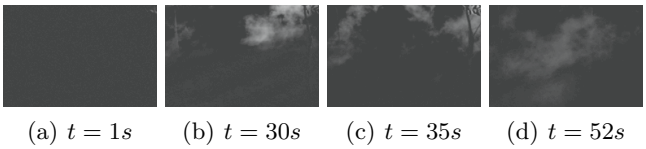


Figure 11: Two synchronised sets of visible (top) and thermal infrared (bottom) camera images are shown after coarse alignment and pixel-by-pixel computation of TI. The TI values from both the colour and infrared cameras are shown for *Dust*.



(a) $t = 1s$ (b) $t = 30s$ (c) $t = 35s$ (d) $t = 52s$

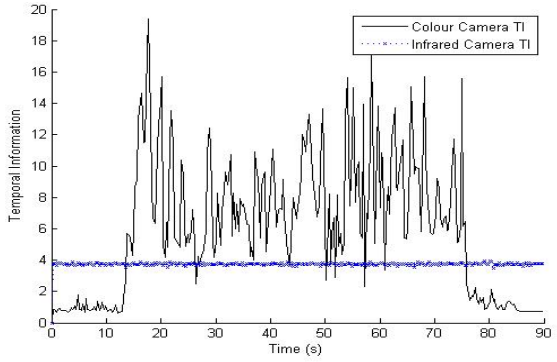


Figure 12: A set of visible images are shown after coarse alignment and pixel-by-pixel computation of TI. The TI values from both the colour and infrared cameras are shown for *Smoke*.

Table 1: Response of Metrics to Dust and Smoke

Metrics	Dust		Smoke	
	Visual	IR	Visual	IR
Shannon Info	$\nearrow^{(*)}$	$\searrow$	$\nearrow$	$\longrightarrow$
Spatial Info	$\searrow$	$\searrow$	$\nearrow\searrow$	$\longrightarrow$
Temporal Info	$\nearrow$	$\nearrow$	$\nearrow$	$\longrightarrow$

$^{(*)}$ The Shannon Information consistently increased when dust appeared in these examples. However, as discussed previously, it has been observed to be dependent on the background.

from frame to frame. However, if smaller sections of two images are presumed to be looking at the same scene (e.g. from a transformation using onboard navigational equipment), these could be compared for temporal changes.

5 Conclusion and Future Work

The purpose of this work is to propose and illustrate metrics that can contribute to the improvement of the perception model by evaluating the quality of sensor data. In particular, the focus of this paper was to show that it is possible to obtain a better discrimination of recognised outdoor challenging conditions (such as airborne dust or smoke) using those metrics computed from visual and infrared data.

Table 1 summarises the response of the metrics in general terms without discussing the absolute value or relative change. The symbol "$\nearrow$" indicates that the value of the metric was observed to consistently increase when the condition described appears in the scene. Similarly, "$\searrow$" and "$\longrightarrow$" indicate that the metric consistently decreased or no effect was observed, respectively. "$\nearrow\searrow$" illustrates that although the condition had an effect on the metric output, the response was variable, thus no straight-forward systematic conclusion could be driven in that case (see Section 4.2).

Relying on information theory, the proposed metrics should have the potential to be generalised to various types of sensors. This work is a preliminary study that calls for several areas of future developments, which are discussed below.

Extension to the dynamic case

The experimental study presented in this paper has been based on sensor data gathered in a static environment, with a motionless vehicle. Naturally, the movement of the vehicle is likely to generate changes in the background of the images which will make the interpretation of the evolution of the considered metrics more difficult. However, it is believed that if the overlap between successive images is sufficient and is estimated, a similar exploitation of these metrics would still be possible.

Local information analysis

In this paper the analysis of the metrics was made on full images, with full resolution. Nevertheless, local analyses are likely to be more efficient and possibly more accurate in the future, as smaller areas of study may allow for the localisation of failures. This will become more relevant especially when dealing with multiple sensors that have overlapping regions of perception space which does not include the whole scan or image.

Combination of Information

An important part of future work will be dedicated to the development of a framework for combining information provided by various metrics and also various sensors. Evaluating the correlation between data from different sensors, for example through *mutual information* [Mackay, 2007, §8.1], offers promising perspectives [Underwood, 2009, §5.2]. Indeed, studying the agreement, and more importantly the *disagreement* between sensors should lead to an increasing power of discrimination of challenging situations. A compelling example is in the presence of smoke, which does not affect the IR camera, contrary to the visual camera (as illustrated in Fig. 7).

Acknowledgements

This work was supported in part by the Asian Office of Aerospace Research and Development (AOARD), the Centre for Intelligent Mobile Systems (CIMS), funded by BAE Systems as part of an ongoing partnership with the University of Sydney, and the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government.

References

[Correia and Pereira, 2003] P. L. Correia and F. Pereira. Objective evaluation of video segmentation quality. *IEEE Transactions on Image Processing*, 12(2):186–200, 2003.

[Durrant-Whyte and Bailey, 2006] Hugh Durrant-Whyte and Tim Bailey. Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms. *Robotics and Automation Magazine*, June 2006.

[Gertler, 1998] Janos Gertler. *Fault detection and diagnosis in engineering systems*. Marcel Dekker, 1998.

[Ibeo, 2009] Ibeo. *Ibeo LUX Laser Scanner*. http://www.ibeo-as.com/english/products_ibeolux.asp, 2009.

[ITU-T, 1999] ITU-T. Subjective video quality assessment methods for multimedia applications, 1999.

[Kelly *et al.*, 2006] Alonzo Kelly, Omead Amidi, Mike Bode, Mike Happold, Herman Herman, Tom Pilarski,

Pete Rander, Anthony Stentz, Nick Vallidis, and Randy Warner. Toward reliable off road autonomous vehicles operating in challenging environments. In Springer Berlin / Heidelberg, editor, *Experimental Robotics IX*, volume 21, pages 599–608. 2006.

[Leonard *et al.*, 2008] John Leonard, J. How, and S. Teller et al. A perception-driven autonomous urban vehicle. *Journal of Field Robotics*, 25(10):727–774, September 2008.

[Mackay, 2007] David Mackay. *Information Theory, Inference & Learning Algorithms*. Cambridge University Press, 2007.

[Narasimhan and Nayar, 2003a] S. G. Narasimhan and S. K. Nayar. Contrast restoration of weather degraded images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(6):713–724, 2003.

[Narasimhan and Nayar, 2003b] S. G. Narasimhan and S. K. Nayar. Shedding light on the weather. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 665–672, 2003.

[Nayar and Narasimhan, 1999] S. K. Nayar and S. G. Narasimhan. Vision in bad weather. In *Proceedings of the Seventh International Conference on Computer Vision*, volume 2, pages 820–827, 1999.

[Peynot *et al.*, 2009a] Thierry Peynot, Sami Terho, and Steve Scheding. Sensor data integrity: Multi-sensor perception for unmanned ground vehicles. Technical report, Australian Centre for Field Robotics (ACFR), The University of Sydney, http://www.acfr.usyd.edu.au/techreports/thierry.shtml, 2009.

[Peynot *et al.*, 2009b] Thierry Peynot, James Underwood, and Steven Scheding. Towards reliable perception for unmanned ground vehicles in challenging conditions. In *International Conference on Intelligent Robots and Systems*, 2009.

[Tan, 2008] R. T. Tan. Visibility in bad weather from a single image. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[Underwood, 2009] James Underwood. *Reliable and Safe Autonomy for Ground Vehicles in Unstructured Environments*. PhD thesis, 2009.

[Urmson *et al.*, 2008] Chris Urmson, J. Anhalt, and D. Bagnell et al. Autonomous driving in urban environments: Boss and the Urban Challenge. *Journal of Field Robotics*, 25(8):425–466, 2008.