2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Novelty-based visual obstacle detection in agriculture

Patrick Ross, Andrew English, David Ball, *Member, IEEE*, Ben Upcroft,
Gordon Wyeth, *Member, IEEE*, Peter Corke, *Fellow, IEEE*

***Abstract*— This paper describes a novel obstacle detection system for autonomous robots in agricultural field environments that uses a novelty detector to inform stereo matching. Stereo vision alone erroneously detects obstacles in environments with ambiguous appearance and ground plane such as in broad-acre crop fields with harvested crop residue. The novelty detector estimates the probability density in image descriptor space and incorporates image-space positional understanding to identify potential regions for obstacle detection using dense stereo matching. The results demonstrate that the system is able to detect obstacles typical to a farm at day and night. This system was successfully used as the sole means of obstacle detection for an autonomous robot performing a long term two hour coverage task travelling 8.5 km.**

I. Introduction

Robotics and automation have key roles to play in improving farm productivity. A transition from manned vehicles to autonomous robots requires a complete navigation system, including a method to reliably detect obstacles. However, reliable obstacle detection in agricultural fields is challenging due to their complex unstructured nature. In broad-acre fields, stubble (harvested crop residue) and crop are typical features which are visually ambiguous and have 3D structure, but should not be classified as obstacles. State of the art approaches use a suite of sensors, typically lasers rangers and cameras, to overcome each sensor's failure mode [1].

Vision-only solutions have promise in broad-acre fields as they are low cost and provide both appearance information and structural information through stereo matching. Early work in obstacle detection for agricultural robots demonstrated the potential of monocular vision to detect the existence of obstacles using a relatively simple algorithm [2]. Stereo vision provides a method to more accurately determine the obstacle's location, however the highly ambiguous appearance of broad-acre fields presents problems for dense disparity matching.

This paper presents a new vision only system for detecting obstacles that uses a novelty detector to inform stereo matching. The novel contributions of this work are the design of an online novelty detection system capable of updating its model over time for changes in the environment and lighting conditions, and the combination of this novelty detector and stereo matching into a complete obstacle detection system. Performing stereo matching on only the novel image regions overcomes problems with ambiguity in the ground plane and appearance. Computational performance is also significantly improved as only a small percentage of the image is processed with dense stereo matching. The novelty detector continuously updates an appearance model to handle gradual changes in lighting and field appearance and finds novel image regions by estimating the probability density in image descriptor space. In order to improve discrimination, the novelty detector incorporates an image-space positional understanding.

Figure 1. Output of the novelty detector for various obstacles. The obstacles are clearly detected, with a few other small false positives, typically bright sections of grass. The vast majority of stubble is filtered out, making the obstacle detection more robust to stubble.

We demonstrate the performance of the system in detecting several obstacles typical to the farm environment. The system detects these obstacles at day and night, using a strobe light system for night operation. The system is incorporated into a navigation system and is the sole means of detecting obstacles for an autonomous robot performing a two hour coverage task.

The remainder of this paper is laid out as follows. Section 2 details relevant prior art. Section 3 motivates and outlines our strategy for obstacle detection, Section 4 describes experiments and Section 5 presents results.

II. Background

There is a considerable amount of prior work detecting obstacles. We focus here on prior art that is applicable to field robotics, since obstacle detection in field environments has different issues to that of other environments.

Typical solutions to obstacle detection in field environments use expensive sensor suites, including multiple

The authors are with the School of Electrical Engineering and Computer Science at the Queensland University of Technology, QLD, Australia (e-mail: patrick.ross@connect.qut.edu.au)

lasers and cameras [1, 3], and often other sensors such as IR cameras. Combinations of sensors can provide high fidelity environmental information, but is impractical for many applications due to the high cost of the sensors, and high computational cost.

Pure laser has also been considered for obstacle detection and environmental modeling [4]. This approach has been shown to give good results in environments where the ground is not occluded.

Some works have considered the use of pure vision for obstacle detection and traversability estimation. These include the use of pure structure from stereo matching [5, 6], and the use of monocular classification strategies [7-9].

Pure scene structure methods suffer from the inability to determine the traversability of the object, considering only its shape. This incorrectly classifies grass as an obstacle, as discussed in [10]. This is of particular issue in broad-acre fields, as stubble and young plants are a common feature. These are traversable, and in fact it is desirable to do so for many important tasks to be carried out in these fields, such as weed spraying.

In addition, structure from stereo vision requires the use of dense stereo matching. Stereo matching typically performs poorly in fields due to the highly repetitive patterns formed by the crop rows, introducing ambiguity into the problem. Geiger [11] produced a strategy for dense stereo matching which attempts to disambiguate these regions by looking for high quality seed disparities, and interpolating.

Conversely, appearance-based strategies are capable of considering the class of the object, not just its shape. This allows for the understanding that stubble and similar objects are not obstacles [10]. However, typical appearance-based strategies are susceptible to changes in lighting conditions. Ranganathan [12] demonstrated the change in SIFT descriptors given small lighting changes. Valgren [13] explored this effect in more detail, concluding that all common descriptors (such as SIFT, SURF and their variants) were to some degree lighting variable, although it is unclear whether this is due to variation in the descriptors or the detected keypoints. Kim [10] reported issues with their classification algorithm when imagery transitioned from direct sunlight into shaded areas.

Typical structural obstacle detection algorithms utilise ground plane extraction [5, 6], which is almost impossible in fields where the ground is often largely obscured. It is possible to infer the location of the ground as in [1] in this case, however this requires significant modeling of the environment, significant processing effort, and accurate structural information (in this case from a pair of lasers).

Soyer [14] discusses the necessity of a robotic system to be able to focus on areas of interest. This focus enables reduced processing on the image as a whole, while not significantly impacting the quality of the result, subject to the quality of the interest detection. This conclusion is motivated by physiological and psychological studies of biological systems, where this so-called selective attention is shown to be an integral part of biological visual processing [15].

Regions of interest can often be detected using novelty detection [2], where regions of the image which are dissimilar to recent views of the world can be identified. For an in-depth review of novelty detection, we refer readers to the comprehensive two-part review produced by Markou [16, 17]. We focus here on the strategies for novelty detection that have been applied to robotics problems.

Ollis [2] designed a system for visual obstacle detection using novelty detection. Their system provided noisy results since it used the novelty mask directly as the obstacle map. They concluded that further filtering was necessary in order to make the system functional.

Marsland [18] produced a method of novelty detection which emulates the process of habituation, where the human brain learns to ignore repeated stimulus. This was achieved using a Grow When Required (GWR) self-organizing map.

Neto [19] presented a system for novelty detection which utilized the reconstruction error of an autoencoder. When presented with a novel sample, the autoencoder produces a large error, hence error characterises novelty. This strategy is reported to be very sensitive to the initial conditions, and hence doesn't adapt well to a changing environment.

There is a similar field of change detection [20, 21], which determines whether the environment has changed since a prior traversal. Rather than considering the environment as a whole, it compares this scene to prior traversals to determine novelty.

III. System Design

In broad-acre fields the environment's appearance is generally uniform as seen in Figure 1. Aside from a few pathological cases, typical obstacles seen in this environment have a different appearance to the field. Therefore the system incorporates a novelty detector to mask the region to search for obstacles. Figure 3 shows the different obstacles typical to this environment, including people, vehicles, power poles, and other farming equipment. Since the crop is regularly seen the novelty detector learns to filter it, and hence it is rarely marked as an obstacle.

Typical operations in fields require moving along the crop rows. During this type of movement, the appearance of the inter-row gaps should remain very constant in image space subject to small local variations, hence novelty detection can be significantly improved by incorporating a strong image-space positional understanding.

Figure 2 gives an overview of the system. The vehicle's motion over rough ground means that there is considerable frame to frame image change which would limit the use of image-space positional understanding. Visual odometry is used to stabilise imagery prior to novelty detection, by providing the transformation between camera imagery and stabilized imagery. This transformation is filtered before being applied to remove steady state components. Novelty detection is performed on the stabilised imagery, and the resultant mask is then cast back into the camera frame for stereo matching. This point cloud is then filtered to produce an obstacle map. To prevent common obstacles from being incorporated into the model of normal appearance, obstacles detected by the system are transformed into image space and masked from being learnt.

During all trials the complete system operated continuously at 10Hz.

A. Image stabilization

The image stabilization node calculates an image-space transformation from camera imagery to stabilized imagery using a filtered pose estimate. This is used to generate

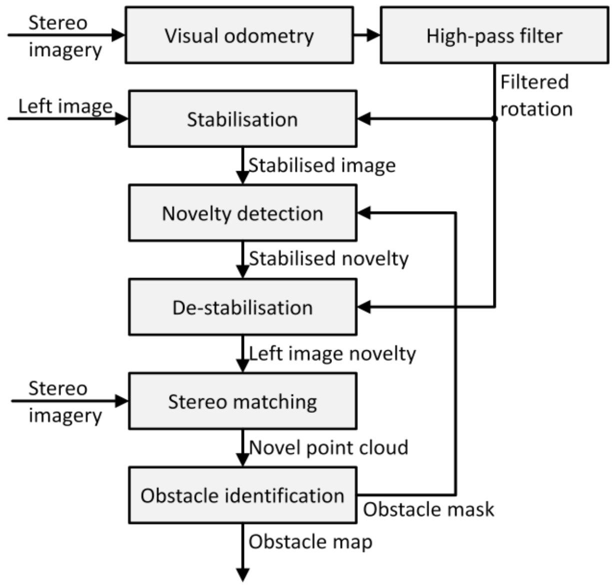

Figure 2. Overview of the system. Visual odometry is utilised to stabilise imagery for novelty detection. This novelty mask is then cast back into the camera's reference frame, and is used to direct stereo matching. The resultant point cloud is filtered to produce an obstacle map, which is fed back into the novelty detection node to prevent obstacles from being added to the model of the normal environmental appearance.

stabilized imagery for novelty detection, and to convert the novelty mask back into the camera frame for stereo matching.

Stereo visual odometry (LIBVISO2[1] [22]) generates a rotational pose estimate. This pose estimate is high-pass filtered so that the rotation has no steady state component. Consider the situation where the vehicle turns around at the end of the row; the unfiltered stabilisation after this maneuver puts the entire image out of view.

At each update step the magnitude of the rotation is scaled down by a factor of $(1-\alpha)$ before adding the relative change from the new frame. The image transformations are done in an approximate manner. The image is rotated using the roll and shifted using the pitch and yaw.

B. Image descriptors

The image is split into a 40x32 grid, each cell of which is 32x32 pixels in size. This node generates a descriptor which comprises the location, the average Lab colour and the standard deviation of Lab colour. Lab colour was chosen since it is perceptually uniform, which makes taking the L2 distance between vectors meaningful.

Each of these variables was initially scaled to be in the range $[0,1]$. The variance was scaled to describe the same range. The variance in this context provides a simple approximation of patch texture.

$$\widehat{\boldsymbol{x}} = \begin{bmatrix} L_\mu & a_\mu & b_\mu & L_\sigma & a_\sigma & b_\sigma & u & v \end{bmatrix}^T \quad (1)$$

where L, a, and b indicate the components of the Lab colour space, the subscripts μ and σ indicate the mean and the standard deviation respectively, and u and v indicate the location of the center of the block in image space.

[1] Available from http://www.cvlibs.net/software/libviso/

C. Novelty detection

The novelty detector uses Parzen windows to determine the probability density of the image descriptors in descriptor space by looking for samples with a probability density below a threshold, ρ. The density is estimated using all prior descriptors in the database.

In standard Parzen windows the probability density function, $p(\boldsymbol{x})$ is approximated by

$$p(\boldsymbol{x}) \approx \frac{1}{N}\sum_i k_\sigma(\boldsymbol{x}-\boldsymbol{x_i}) \quad (2)$$

where $k_\sigma(\boldsymbol{x})$ is the kernel function, $\boldsymbol{x_i}$ are the samples, N is the number of samples. In this application, over a period of hours the definition of normal appearance may change as global lighting levels and weather conditions change. For this reason, the definition of normal, and hence novel, changes over time. To model this, we require the ability to weight samples differently in order to gradually forget old data. We therefore approximate the probability density by

$$p(\boldsymbol{x}) \approx \frac{1}{\sum_i w_i}\sum_i w_i k_\sigma(\boldsymbol{x}-\boldsymbol{x_i}) \quad (3)$$

where w_i is the weight of sample $\boldsymbol{x_i}$. Here we choose the kernel function, $k_\sigma(\boldsymbol{x})$ to be a Gaussian kernel with spherical covariance. This assumption of spherical covariance simplifies calculation of the probability density, however at the expense of some generalisation.

$$k_\sigma(\boldsymbol{x}) = \frac{1}{\sigma^D(2\pi)^{\frac{D}{2}}}\exp\left[-\frac{x^Tx}{2\sigma^2}\right] \quad (4)$$

where D is the dimensionality of the descriptor. In order to mitigate the assumption of spherical covariance the samples are scaled using a scaling vector $\boldsymbol{s}$.

$$\boldsymbol{x} = \boldsymbol{s}^T\widehat{\boldsymbol{x}} \quad (5)$$

Weights are initialised for new samples by:

$$w_i^0 = \begin{cases} 1 & \text{if the first update step} \\ \lambda & \text{else} \end{cases} \quad (6)$$

where λ is the forgetting factor. This means that at startup the model is fully initialised and strongly biased towards the initial data. As time progresses, this bias reduces. Weights are reduced over time using the forgetting factor:

$$w_i^{n+1} = (1-\lambda)w_i^n \quad (7)$$

Given the large volume of data required to construct an accurate model of the environment, this estimation of probability density needs to be extremely efficient, especially for real time operation. To achieve this, FLANN [23] is used to provide a list of the samples within 3 standard deviations of a query. Since FLANN provides the distance to these points, the kernel function evaluation is extremely efficient. Using this form, the above equation for the kernel function can be rewritten as

$$k_\sigma(d) = \frac{1}{\sigma^D(2\pi)^{\frac{D}{2}}}\exp\left[-\frac{d^2}{2\sigma^2}\right] \quad (8)$$

The novelty detector updates the model every 40 frames and executes in a separate thread to not impact the continuous operation of the system. To avoid queuing updates the number of frames between updates is greater than the maximum model update time which has the effect of changing the effective forgetting factor. The effective forgetting factor updates according to

Figure 3. The various obstacles expected to be in the field, including from left to right, vehicles, people, power poles, tanks and drums at both day and night time. Note how the lighting changes significantly affect the appearance of the environment and the obstacles.

$$\lambda_{\text{eff}} = 1 - (1 - \lambda)^{\frac{1}{40}} \tag{9}$$

Hence the choice of the number of frames between updates only impacts the choice of λ, and so isn't directly a system parameter.

The output of the novelty system is put through a rudimentary islanding filter, which keeps only novel regions that are adjacent to other novel regions. To ensure that all of the novel regions are properly marked the result is dilated.

To ensure that obstacles are not ever marked as normal, and hence missed by the novelty system, obstacles detected by the system mask the model updates. The model incorporates only those samples that are at least a minimum distance in image space from an obstacle. Adding a minimum distance in this way allows for the novelty detector to use the mask without needing to pass through stabilisation.

D. *Obstacle detection*

The obstacle detector uses stereo matching to generate a point cloud at regions in the image identified as novel. Our method is independent of the choice of stereo matching algorithm, however here we compare the use of block matching and LIBELAS[2] [11].

Once a point cloud has been generated, the results are filtered to produce an obstacle map. Ground plane detection is typically used to determine which points in the point cloud are obstacles, however in this case there may be very few ground points actually found. In an ideal scenario, no ground points would ever be added to this point cloud, so ground plane detection is unsuitable.

Rather than use a ground plane assumption, the vehicle is assumed to be flat on the ground. Points are cast into the vehicle's reference frame, and any points that are in a given height range are marked as obstacles. This detector generates occupancy grids using these point cloud obstacles which are integrated over multiple frames by the motion planner.

Note that this assumption is similar to the assumption employed when utilising planar laser scanners for obstacle detection, where the plane of the laser is assumed to be parallel to the ground. As with a planar laser this is expected to perform poorly on objects which are low to the ground due to the height filtering.

IV. EXPERIMENTS

The robotic platform used for testing is a John Deere TE Gator modified for autonomous operation. It is Ackermann steered and is approximately 2.6m long and 1.5m wide. It has been fitted with a 200L tank and a 5m spray boom for spray operations. It has two different obstacle detection sensors; a planar SICK LMS151 laser scanner and a stereo pair of forward facing iDS UI-5240CP PoE cameras. The laser was mounted so that the scan line was parallel to the ground, at a height of 0.7m. Note that the majority of the stubble was less than 0.5m. The cameras were mounted at a height of 1.55m, angled towards the ground by 15 degrees, with a 0.75m baseline. The cameras were calibrated using the AMCC MATLAB toolbox[3] [24]. The cameras provided colour imagery at 1280x1024 resolution, operating at 10Hz.

The vehicle has two onboard PCs, one of which was responsible for vehicle control and path planning, the other of which was utilised for perception and obstacle detection. The system utilises the Robot Operating System (ROS) for inter-process and inter-machine communication.

All experiments were carried out in fields on a farm in Emerald, Australia. The bulk of the experiments were performed in a broad-acre sorghum stubble field, where the goal was to perform a weed-spraying coverage operation.

A. *Novelty detection*

The first study investigated the ability of the novelty detection algorithm to adapt to different crops and successfully detect obstacles. Three different datasets were gathered in three different fields: wheat, chickpea and a recently harvested sorghum field. The obstacle in these datasets was a person.

During these experiments the vehicle was driven by a human operator, and was in motion for each of the test cases.

[2] Available from http://www.cvlibs.net/software/libelas/

[3] Available from http://code.google.com/p/amcctoolbox/

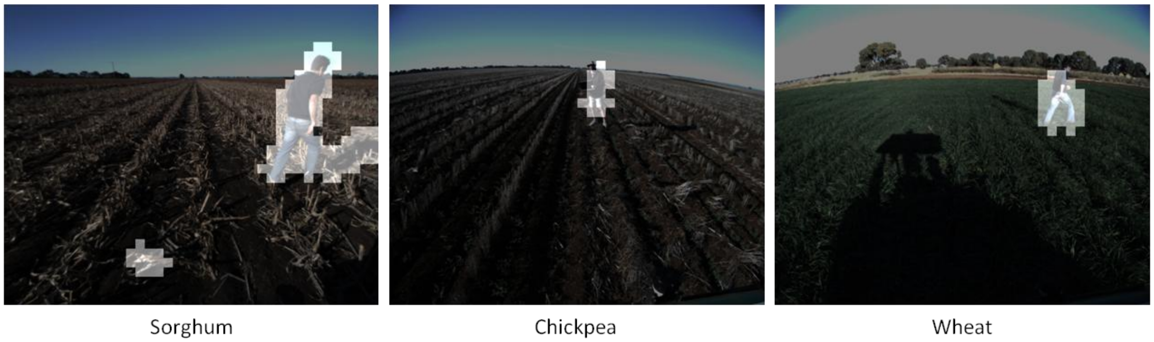

Figure 4. Sample novelty detection when a person walks in front of the cameras. Other small false positives can corrupt the result, however these aren't usually labelled as an obstacle after stereo matching. The novelty detector successfully detects the person in each of the different sample fields.

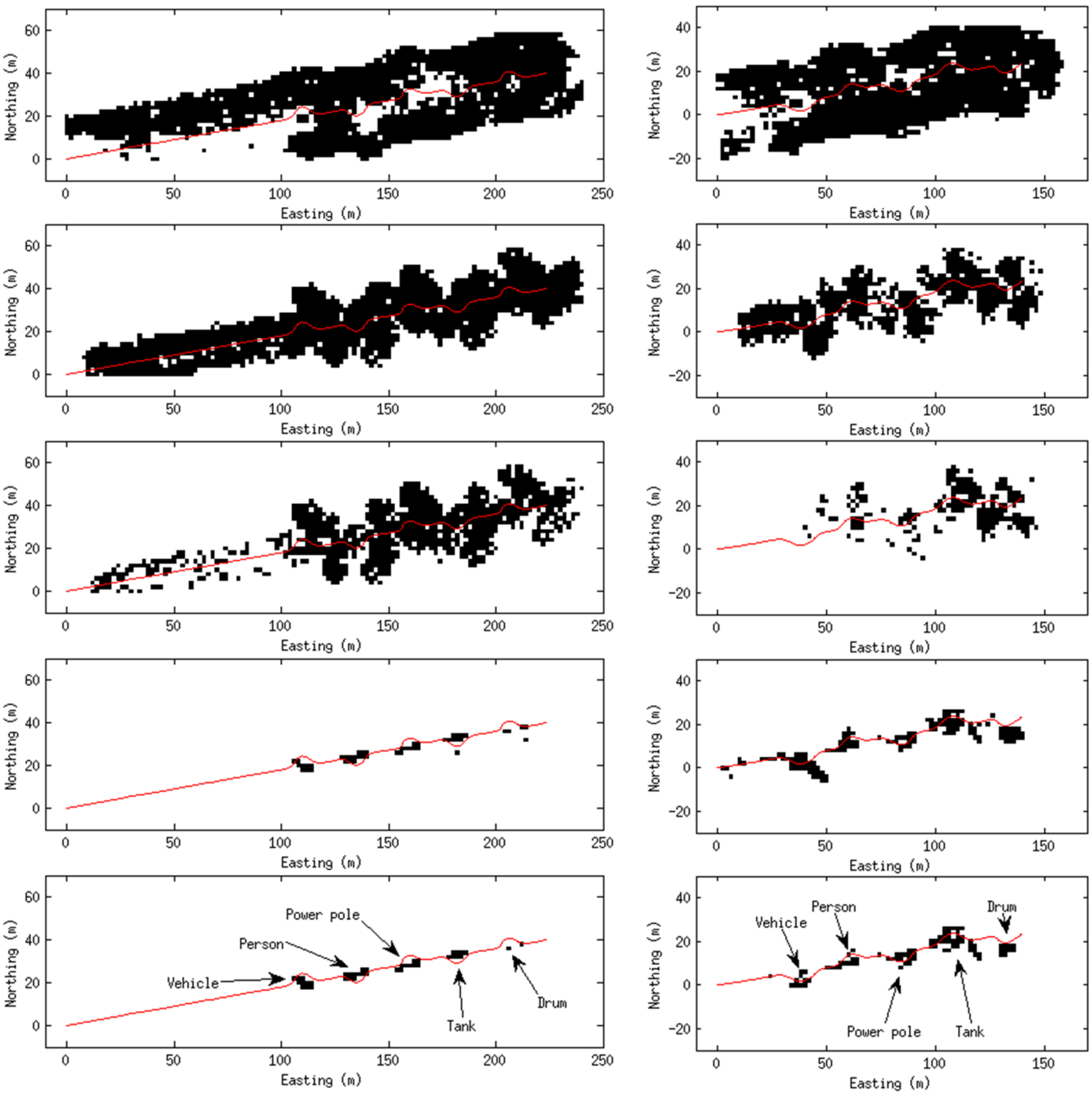

Figure 5. Obstacle detection results. Left column is the day dataset, right column is the night dataset. Both datasets have the same physical obstacle configuration, while the vehicle was started closer to the obstacle course in the night dataset. The results are, top to bottom: laser, block matching, block matching with novelty, LIBELAS, LIBELAS with novelty. The obstacle map is quantised to blocks of 2m here for clarity - in the actual system the block size was 0.1m. The configuration of the planar laser produced obstacle maps which were completely unusable, as did block matching. For all experiments, the novelty detector improved the quality of the obstacle detection while not filtering out any actual obstacles. There was a performance degradation of the system during night time operation due to the reduced quality of stereo matching.

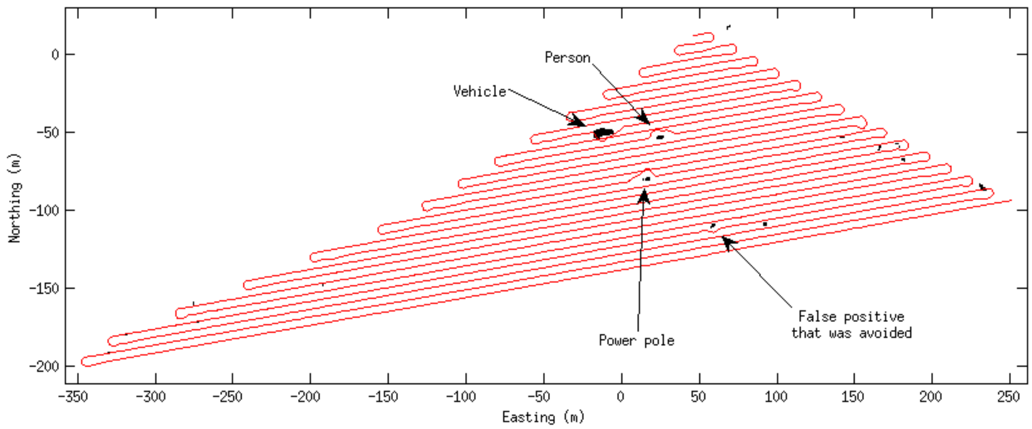

Figure 6. The path taken by the robot for the long term experiment. Each of the obstacles were successfully detected, however there were a number of small false positives. Only a single one of these was avoided, as indicated in the figure. The others were cleared from the obstacle map on the next update step, and so weren't avoided.

B. *Obstacle detection*

This study investigated the accuracy of the obstacle detection algorithm in detecting a person, vehicle, drum, tank and a power pole. Refer to Figure 3 for images of each of these obstacles. Two datasets were collected, one during the middle of the day and at another late at night. For the night dataset, a pulsed light source was used for illumination. The vehicle was driven by a human operator. The location and heading of the vehicle were logged using a high-precision GPS system.

In this study we compare both block matching and LIBELAS for accuracy in this environment to demonstrate that the use of novelty improves results irrespective of the choice of stereo matching algorithm. Additionally, this allows the choice of the best stereo matching algorithm in this field for the extended test. We compare this to the planar laser. The planar laser is expected to perform poorly in this environment, since the vehicle is bouncing around considerably on the uneven ground, and hence the scan-line often intersects the ground.

C. *Long-term operation*

To investigate the long term effectiveness of the system, we conducted a 2 hour experiment where the robot autonomously traversed the field using a lawnmower coverage path. The system described in this paper was the sole method of detecting obstacles. During this test the vehicle travelled approximately 8.5 km. There were three obstacles that required avoidance in the field - a vehicle, a person and a power pole. These were the same obstacles from the previous experiment, however the person and vehicle were in different locations in the field. Refer to Figure 6 for a summary of the field and obstacle locations.

V. RESULTS

A. *Novelty detection*

Figure 4 demonstrates the typical result from the novelty detection in each of these fields. The algorithm readily adapts to each field with no parameter tuning, however there are a few false positives. These false positives typically include the shadows of obstacles, which are typically filtered out by the stereo matching. The algorithm successfully detects the human obstacle in each of these datasets.

B. *Obstacle detection*

Figure 5 demonstrates the accuracy of the system with and without novelty detection using the two different stereo matching algorithms. In both cases the novelty detector improves the robustness of the detection by removing outliers. LIBELAS out-performs block matching for both datasets. Block matching is fast but seems to give poor results on this data, likely due to a combination of ambiguity (repetitive patterns) and the use of the sum of absolute differences (SAD) for matching. LIBELAS is slower, but reduces ambiguity and is capable of interpolating surfaces with little texture. Unfortunately, it has a tendency incorrectly interpolate where there is little texture, such as in the sky, and so can also give erroneous results. The improvements for LIBELAS by adding in the novelty detector were less significant than that for block matching, however these differences became more significant during longer-term testing.

All of the obstacles were successfully detected in the day dataset with our system using LIBELAS for stereo matching, while only the drum was missed on the night dataset. There were other false positives in the night time dataset. However our system still out-performed stereo matching alone and the planar laser. This suggests that the stereo matching algorithm may need more tuning for use at night time.

The stereo matching-based obstacle detection with our novelty filter and LIBELAS was found to be largely independent of the height threshold for these short experiments - values between 0.5m and 1.2m off the ground gave almost identical results. This is due to the robustness of the novelty detector.

The planar laser performed poorly due to a combination of the fact that the vehicle was constantly pitching due to the uneven ground, and the fact that the laser was mounted only just above the typical height of the sorghum stubble. For ideal operation in this environment, the laser should be mounted higher, however this height is dependent on the field. As the crop grows and is harvested, the height of the laser would need to be constantly adjusted to suit the height of the crop.

C. *Long-term operation*

Figure 6 shows the path that the vehicle took, with the obstacle map from the proposed system overlaid. Each of the true obstacles are also indicated. These were all successfully avoided, however there was also a single false positive which the vehicle avoided. This was caused by the vehicle traversing a contour bank. At this time the vehicle was facing towards the ground, violating the assumption that the ground was flat and the vehicle was level with the ground. Despite this, this contour bank was successfully traversed a number of times with only this single false positive, highlighting the robustness of the system.

VI. CONCLUSION

This paper presents a novel method for vision-based obstacle detection in broad-acre fields using a combination of a novelty filter and stereo matching. The system uses a novelty detector to determine areas of interest in the scene to direct stereo obstacle detection. This makes the system more robust in broad-acre fields. Since the novelty detector continuously updates a model of the appearance of the environment, it generalises to different fields and lighting conditions.

The algorithm significantly out-performs planar laser obstacle detection in fields, since laser scanners detect stubble as obstacles, resulting in large amounts of false positives. This makes coverage operations almost impossible using planar laser scanners.

A caveat of our proposed system is that it assumes that the appearance of any obstacles deviate from the typical appearance of the environment. Since the novelty detector operates on colour, when the colour of the obstacle becomes similar to that of the terrain it becomes more difficult to detect. While we consider this a rare case, this is a potential issue with our descriptor design, and future work will investigate more effective descriptors. We expect additional cues from texture and depth information could improve the quality of the result. However, we have shown that this assumption holds for many of the obstacles typical to this environment.

ACKNOWLEDGMENT

This work was supported in part by the Australian Research Council Linkage Project LP110200375 “Robotics for Zero Tillage Agriculture” awarded to the Queensland University of Technology, SwarmFarm Robotics and the Australian Centre for Field Robotics.

The authors would like to thank Kyran Findlater for constructing the strobe light used in the night experiments.

REFERENCES

[1] C. Wellington, A. Courville, and A. Stentz, "Interacting markov random fields for simultaneous terrain modeling and obstacle detection," in *In Proceedings of Robotics: Science and Systems*, 2005.

[2] M. Ollis and A. Stentz, "Vision-based perception for an automated harvester," in *Intelligent Robots and Systems, 1997. IROS '97., Proceedings of the 1997 IEEE/RSJ International Conference on*, 1997, pp. 1838-1844 vol.3.

[3] S. Thrun, M. Montemerlo, H. Dahlkamp, *et al.*, "Stanley: The Robot That Won the DARPA Grand Challenge." vol. 36, M. Buehler, K. Iagnemma, and S. Singh, Eds., ed: Springer Berlin / Heidelberg, 2007, pp. 1-43.

[4] S. Karumanchi, T. Allen, T. Bailey, *et al.*, "Non-parametric Learning to Aid Path Planning over Slopes," *The International Journal of Robotics Research,* vol. 29, pp. 997-1018, July 1, 2010 2010.

[5] R. Hadsell, P. Sermanet, J. Ben, *et al.*, "Learning long-range vision for autonomous off-road driving," *Journal of Field Robotics,* vol. 26, pp. 120-144, 2009.

[6] P. Vernaza, B. Taskar, and D. D. Lee, "Online, self-supervised terrain classification via discriminatively trained submodular Markov random fields," in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, 2008, pp. 2750-2757.

[7] P. Filitchkin and K. Byl, "Feature-Based Terrain Classification For LittleDog," in *IEEE International Conference on Intelligent Robots and Systems*, 2012.

[8] A. Angelova, L. Matthies, D. Helmick, *et al.*, "Learning and prediction of slip from visual information," *Journal of Field Robotics,* vol. 24, pp. 205-231, 2007.

[9] Y. N. Khan, P. Komma, and A. Zell, "High resolution visual terrain classification for outdoor robots," in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, 2011, pp. 1014-1021.

[10] D. Kim, J. Sun, S. M. Oh, *et al.*, "Traversability classification using unsupervised on-line visual learning for outdoor robot navigation," in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, 2006, pp. 518-525.

[11] A. Geiger, M. Roser, and R. Urtasun, "Efficient Large-Scale Stereo Matching." vol. 6492, R. Kimmel, R. Klette, and A. Sugimoto, Eds., ed: Springer Berlin / Heidelberg, 2011, pp. 25-38.

[12] A. Ranganathan, S. Matsumoto, and D. Ilstrup, "Towards illumination invariance for visual localization," in *Robotics and Automation, 2013. Proceedings. ICRA '13. IEEE International Conference on*, 2013.

[13] C. Valgren and A. Lilienthal, "Sift, surf and seasons: Long-term outdoor localization using local features," in *Proceedings of the European conference on mobile robots (ECMR)*, 2007, pp. 253-258.

[14] C. Soyer, H. I. Bozma, and Y. Istefanopulos, "A mobile robot with a biologically motivated vision system," in *Intelligent Robots and Systems '96, IROS 96, Proceedings of the 1996 IEEE/RSJ International Conference on*, 1996, pp. 680-687 vol.2.

[15] J. Duncan, "Selective attention and the organization of visual information," *Journal of Experimental Psychology: General,* vol. 113, pp. 501-517, 1984.

[16] M. Markou and S. Singh, "Novelty detection: a review—part 1: statistical approaches," *Signal Processing,* vol. 83, pp. 2481-2497, 2003.

[17] M. Markou and S. Singh, "Novelty detection: a review—part 2:: neural network based approaches," *Signal Processing,* vol. 83, pp. 2499-2521, 2003.

[18] S. Marsland, U. Nehmzow, and J. Shapiro, "On-line novelty detection for autonomous mobile robots," *Robotics and Autonomous Systems,* vol. 51, pp. 191-206, 2005.

[19] H. Vieira Neto and U. Nehmzow, "Visual novelty detection with automatic scale selection," *Robotics and Autonomous Systems,* vol. 55, pp. 693-701, 2007.

[20] B. Neuman, B. Sofman, A. Stentz, *et al.*, "Segmentation-based online change detection for mobile robots," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, 2011, pp. 5427-5434.

[21] B. Sofman, B. Neuman, A. Stentz, *et al.*, "Anytime online novelty and change detection for mobile robots," *Journal of Field Robotics,* vol. 28, pp. 589-618, 2011.

[22] B. Kitt, A. Geiger, and H. Lategahn, "Visual odometry based on stereo image sequences with RANSAC-based outlier rejection scheme," in *Intelligent Vehicles Symposium (IV), 2010 IEEE*, 2010, pp. 486-492.

[23] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," in *In VISAPP International Conference on Computer Vision Theory and Applications*, 2009.

[24] M. Warren, D. McKinnon, and B. Upcroft, "Online calibration of stereo rigs for long-term autonomy," in *Proceedings of the International Conference on Robotics and Automation*, 2013.