Stereo Based Obstacle Detection for an Unmanned Air Vehicle

Jeffrey Byrne, Martin Cosgrove and Raman Mehra
Scientific Systems Company, Inc.
500 West Cummings Park, Suite 3000, Woburn, MA 01801
Email: {jbyrne, martin, rkm}@ssci.com

Abstract— This paper presents the Visual Threat Awareness (VISTA) system for real time collision obstacle detection for an unmanned air vehicle (UAV). Computational stereo performance has progressed such that several commercial or open source implementations are available which operate at frame rate, but suffer from well known correspondence errors. We show that introducing a global segmentation step after commodity stereo can increase robustness and leverage existing stereo software. The global segmentation step is based on a graph structure appropriate for collision detection, human vision inspired foveation, perceptual organization and graph partitioning using the minimum s-t graph cut. This system has been prototyped using the Sarnoff Acadia I vision processor to enable processing of 640x480 resolution imagery at 5-10Hz operation on embedded avionics. We describe system theory, demonstrate segmentation results on scenes of increasing complexity, and show flight experiment results on Georgia Tech's GT-Max autonomous helicopter against real collision obstacles.

I. INTRODUCTION

Unmanned Air Vehicles (UAVs) are envisioned as an integral part of future civilian and military operations, including intelligence, surveillance and reconnaissance (ISR) missions, search and rescue missions and surveillance for homeland security. These missions often require a UAV to fly nap-of-the-earth, risking collision with low level obstacles such as trees and buildings whose position cannot be guaranteed as known before flight. Practical UAVs must include situational awareness based on sensing and perception of the immediate environment to locate collision dangers and plan an appropriate avoidance path.

Sensors considered for collision detection can be characterized as active or passive sensors. Active sensors, such as millimeter-wave and microwave RADAR, are currently under investigation for detect, see and avoid (DSA) capability for large UAVs [1]. These sensors exhibit all weather operation with resolution appropriate for wire detection, however measurements are sparse with a low scan rate, the sensor is non-covert due to active emission of radiation, and the sensor form factor does not currently scale to micro air vehicles. Passive sensors based on visual electro-optical (EO) or forward looking infrared (FLIR) are promising due to low size, weight and power requirements as well as a lack of emitted radiation. However, they require significant image processing to detect obstacles and require a narrow field of view for wire detection which limits sensor coverage. Bhanu et al. [2] argue for a *maximally passive* system that combines narrow field of view

active sensors for wire detection with wide field of view passive stereo sensors for coverage. This paper proposes a passive stereo system for visual obstacle detection suitable for integration into such a maximally passive system.

Visual obstacle detection is a well explored problem in the literature. Techniques include differential invariants of optical flow [3][4][5][6], horopter, multibaseline and omnidirectional stereo [7][8][9][10], structure from motion [11][12], bearing only methods [13] and obstacle classification [14]. A full review is out of the scope of this paper, however the techniques mentioned all require an explicit or implicit solution to the *correspondence problem* to establish matching features in distinct viewpoints to reconstruct scene structure. Local correspondence errors can generate false or missed obstacles, so a practical visual obstacle detection approach must be robust to such errors. In this paper, we use binocular stereo, focusing on a global segmentation technique performed after stereo correspondence to improve robustness to these errors.

Stereo obstacle detection is difficult due to the variability of an unstructured outdoor environment. Computational stereo computes scene structure by establishing correspondence between features in each image, however variability in surface reflectance, lighting and viewpoint may introduce matching errors that manifest as false or missed obstacles. Real time stereo software based on area correlation or block matching is currently commercially available, but this software suffers from these local correspondence errors and performance is sensitive to the choice of a fixed matching confidence threshold. Recent developments in computational stereo show significant improvements over block matching in indoor environments, however the state of the art is not real time and remains sensitive to occlusions [15][16]. Therefore, we propose to combine existing real time stereo with a global segmentation step to increase robustness to correspondence errors, resulting in a practical obstacle detection system suitable for real time operation.

This paper presents the Visual Threat Awareness (VISTA) system for passive, stereo image based obstacle detection for an unmanned air vehicle. The VISTA system combines block matching stereo computed on the Acadia I vision processor designed by Sarnoff Corporation [17] with image segmentation based on a special purpose graph representation appropriate for collision detection, perceptual organization and efficient graph partitioning based on the minimum graph cut. This
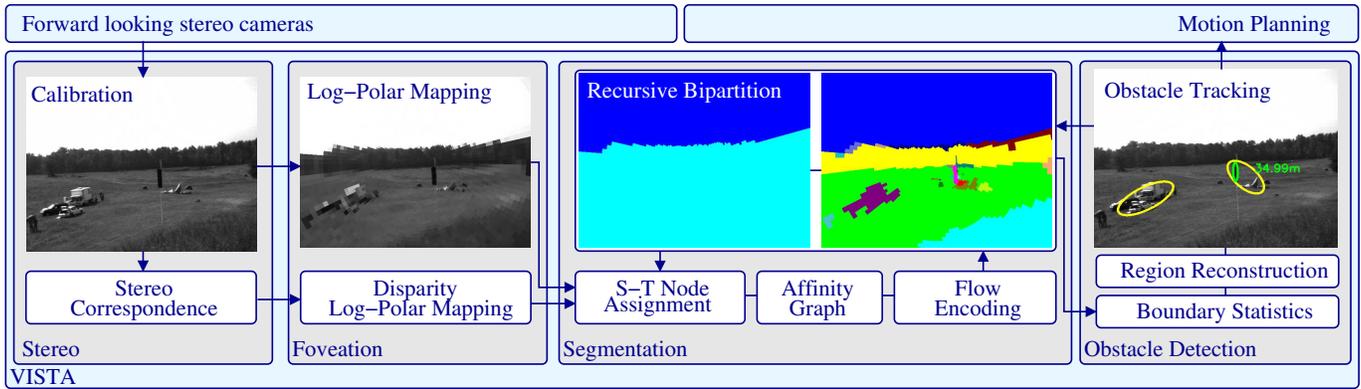
Fig. 1. Visual Threat Awareness (VISTA) system block diagram.

segmentation provides a means to increase robustness to stereo correspondence errors, and provides boundary constraints suitable for motion planning. This paper will describe system theory, demonstrate qualitative segmentation performance on imagery of increasing complexity, and show flight experiment results on Georgia Tech's GT-Max autonomous helicopter [18] in real collision scenarios.

II. SYSTEM ARCHITECTURE

The Visual Threat Awareness (VISTA) system is an approach to collision obstacle detection based on real time stereo, graph partitioning and perceptual organization. A block diagram of the system is shown in figure 1. A stereo pair of cameras is mounted forward looking on the UAV to monitor the region through which the UAV will fly. On each iteration, imagery is captured from a calibrated stereo pair of cameras and passed to the Acadia I vision processor which computes a *stereo disparity map* (section II-A) which is proportional to the 3D scene geometry. The imagery and disparity maps are *foveated* using a log-polar mapping compression (section II-B) and fused into an *affinity graph* representation (section II-C) using perceptual organization techniques. The affinity graph is recursively bipartitioned using a minimum graph cut (section II-C) resulting in an estimate of *k-regions* within the imagery. Those regions which pass an edge consistency test (section II-D) are 3D reconstructed, represented with a bounding ellipse and tracked using a Kalman filter (section II-D) forming an obstacle map suitable for motion planning and avoidance. Each of these blocks will be described in more detail.

A. Computational Stereo

Computational stereo is the process of extracting three-dimensional scene structure from two or more images taken from distinct viewpoints. This computation requires a solution for *stereo correspondence* which is the process of establishing the projection of a scene point in each image. Correspondence techniques attempt to find matching points in the imagery by exploiting constraints such as epipolar geometry, ordering, brightness constancy, edge consistency and uniqueness [15][16]. However, this matching can be ambiguous when features in one image do not have an identical and unique

match in the other image. This may be due to viewpoint changes (foreshortening), multiple feature match (regions of low contrast, periodic features) or no feature match (specular reflections, occlusion, minimum distance violation).

Real time block matching correspondence techniques, such as the commercial products available from Videre Design, Point Grey or Tyzx, include a matching confidence threshold to discard poor matches. However, as shown in figure 2, the accuracy of the correspondence is sensitive to the choice of this threshold, and it is unclear how to choose this threshold in general in an outdoor environment without introducing false alarms or missed detections.

In this system, stereo correspondence is computed on the Acadia I vision processor [17] using a sum of absolute differences (SAD) block matching approach along calibrated epipolar scanlines, with left/right consistency checking and maximum 32 disparity search. The Acadia I vision processor is able to compute 640x480 binocular stereo at 23Hz.

B. Foveation

Foveation refers to a space variant image representation with a high resolution central region or *fovea* surrounded by a lower resolution periphery [19]. This approach is inspired by the structure of the human retina which features a space variant distribution of light sensitive elements.

In the context of collision detection, foveation provides *collision appropriate image compression*. Foveation retains high resolution in the image center, which has a high likelihood of containing a collision obstacle due to alignment of vehicle heading with the optical axis of the sensor. Foveation also reduces the resolution in the periphery which may contain an obstacle, but a low likelihood of containing a collision obstacle. This approach focuses available computational resources on those spatial image regions that are likely to contain collision dangers.

Foveation can be implemented using a log-polar mapping [19], such that the space variant resolution is proportional to the log of the distance from the image center. An example of the log-polar mapping is shown in figure 3.
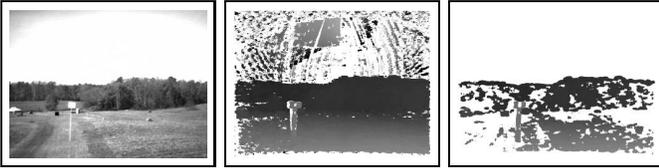
Fig. 2. Tradeoff between stereo matching threshold and correspondence accuracy. (left) Grayscale image (middle) Disparity map with low threshold (right) Disparity map with high threshold. The low threshold introduces severe correspondence errors in the sky (Dark=far,light=near,white=undefined)

Fig. 3. Example of foveation using the log-polar mapping. (left) Original image (middle) Graph representation of foveation (right) Foveated image with high resolution central fovea and low resolution periphery.

C. Segmentation

Segmentation can be defined as the process of labeling features in an image such that features with equal labels are "similar" and "belong together" in some sense, and features with unequal labels do not.

In the context of obstacle detection, segmentation provides a foundation for rejecting stereo correspondence errors. As previously discussed, stereo correspondence errors are generated when image features do not have an identical and unique match. If we introduce a segmentation step such that "similar" is defined as smoothly varying feature measurements, then a segmentation will define k-regions such that within group features are smoothly varying and between group features are discontinuous. Smoothly varying features will generate poor stereo correspondence due to low contrast, so a segmentation can *safely ignore all disparity interior to a region*, and focus only on disparity at the region boundary. Boundary disparity can be checked for the *edge consistency constraint* [20], which states that a true object boundary should exhibit smoothly varying disparity. Those regions passing this check are used for motion planning.

A suitable segmentation algorithm must be stable and scalable. A stable algorithm will generate a consistent segmentation given a slightly different perspective of a scene. Local segmentation approaches such as thresholding, clustering, region growing, split-and-merge and deformable contour techniques are sensitive to initial conditions and often generate inconsistent segmentations in smooth regions [21]. As a result, in this system, we use a global segmentation technique based on recursive minimum cut of an affinity graph. This approach exhibits stability due to the use of global image information, with a polynomial computational complexity for scalability to large images.

The segmentation problem can be posed formally as a Bayesian energy minimization problem [22]. Assume that there exists a finite set of points $P = \{p_1, p_2, \ldots, p_N\}$ that fall within the field of view of the sensor for which measurements can be taken. For each point p_i, a sensor can capture a multidimensional measurement $M(p_i) = \{m_1, m_2, \ldots, m_k\}$, such that the total set of all measurements for all points is $S = \{M(p_1), M(p_2), \ldots, M(p_N)\}$. Each measurement is some descriptive feature of p_i that may include intensity, texture, color, intensity gradient, motion, depth or others. A *labeling* $f(P)$ is a mapping from P to L where L is a finite set of labels. An *energy optimal labeling* f^* minimizes the energy function E [23]
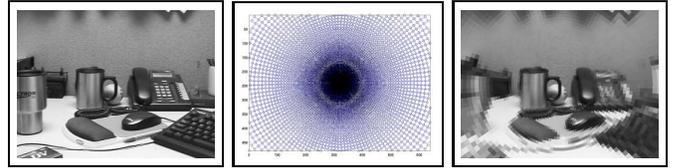
$$E(f) = \sum_{p \in P} D_p(f_p) + \sum_{(p,q) \in N} V_{p,q}(f_p, f_q) \quad (1)$$

$$f^* = \arg \min_f E(f) \quad (2)$$

D_p is a function which encodes the cost of assigning label f_p to p, which represents prior knowledge about the true labeling of p. $V_{p,q}$ is a function which encodes the cost of assigning label f_p to p and a different label f_q to q when (p,q) are neighbors in a given neighborhood set $N \subset P \times P$. This function represents a penalty for violating label smoothness for neighboring (p,q). Solutions f^* to the energy minimization problem are difficult to find in general since (1) can be non-convex in a high dimensional space.

In this application, we approach the energy minimization in (2) as a *recursive maximum network flow* problem. A network flow graph is defined as a directed graph $G = (V, E, W)$ with nodes V, edges E and edge weights W. Edge weights w_{ij} between nodes i and j are interpreted as *capacities*, and certain distinguished nodes s and t in V are interpreted as *terminal nodes*. The maximum network flow problem is that of determining the maximum flow of some commodity between terminal nodes such that the maximum flow on any edge is less than or equal to capacity. Using the Ford-Fulkerson theorem, it can be shown that a solution to the maximum network flow problem is also a solution to the minimum graph cut or mincut problem [24]. The mincut on a network flow graph defines a graph bipartition which is equivalent to a binary labeling, where the binary labeling is an exact solution to the energy minimization in (2) assuming that D_p is equal to the terminal edge capacities, $V_{p,q}$ is equal to the edge capacities, and $V_{p,q}$ is a *regular function* as defined in [23]. Recursive application of the binary labeling generates a k-labeling such that the maximum of the inter-partition flows is minimized among all possible partitions of G into the same number of partitions [24].

This approach requires that imagery be abstracted to a network flow graph representation suitable for obstacle detection. Graph nodes and graph edges are defined using the log polar mapping from section II-B. The log polar mapping is well suited to the minimum cut since the cut value is invariant to certain classes of expansion, such as rotationally symmetric obstacles centered at the fovea. This type of expansion pattern or *looming* may be generated by an approaching collision
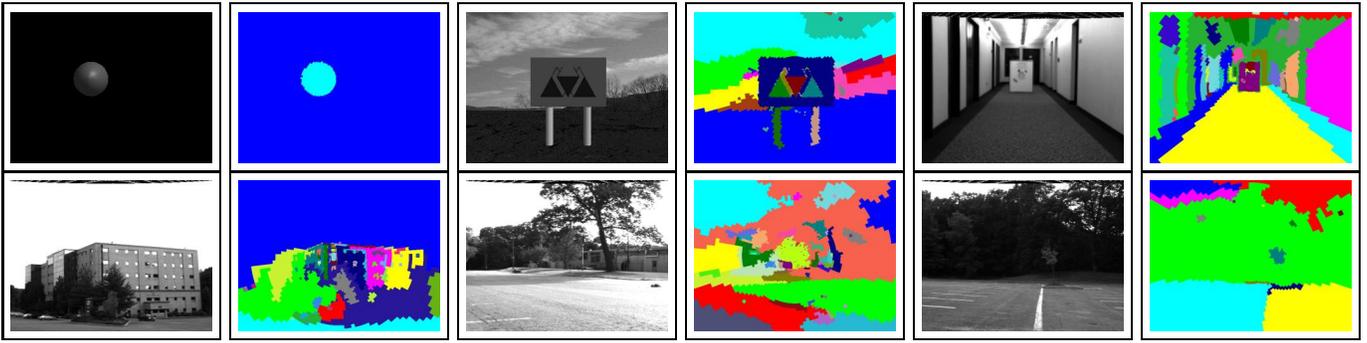
Fig. 4. Segmentation examples. (top row): rendered sphere, rendered billboard, indoor hallway. (bottom row): building, large tree, small tree.

obstacle. Graph edge weights are encoded using the perceptual organization heuristic of similarity [25][26] such that the edge weight w_{ij} or *node affinity* between adjacent nodes i and j can be defined as:

$$w_{ij} = \sum_{d=1}^{D} \alpha_d d_{ij} \qquad (3)$$

$$d_{ij} = \exp\left(\frac{-(\max((m_d(i) - m_d(j)), \mu_d) - \mu_d)^2}{2\sigma_d^2}\right) \qquad (4)$$

Equation 4 is a nonlinear model for feature smoothness, which is similar to the approach by Shi and Malik in [27]. Each node i has D measurements $m(i) = [m_1, m_2, \ldots, m_D]^T$, with the measurement $m_d(i)$ corresponding to the dth measurement for the ith node, such as intensity, color, texture, depth or others. The feature smoothness between nodes i and j is parameterized by μ_d and σ_d which define the mean and standard deviation for smooth feature changes. The resulting node affinity is a linear combination of each feature smoothness with weight α_d. The form of equation 3 is related to *boosting* in machine learning theory [28], which implies that coefficients α may be determined using an offline boosting procedure.

A common drawback of the minimum cut is that solutions tend towards small groups [24][29]. To compensate, we introduce a *flow encoding* $\mathcal{F}(w_{ij})$ which maps affinity w_{ij} to a flow appropriate for the log-polar mapping and the minimum cut.

$$\mathcal{F}(w_{ij}) = C^{\frac{(w_{ij}+1)}{2}} \qquad (5)$$

The choice of the constant C in (5) is chosen experimentally to satisfy three flow constraints as discussed in [30].

In this application, the minimum cut of the network flow graph is computed using a polynomial time augmenting path approach proposed by Boykov and Kolmogorov [31]. Features in equation (4) include foveated intensity and disparity. Additional issues of oversegmentation, perceptual organization heuristics, source-terminal (S-T) node assignment for network flow graphs, recursion stopping criterion and flow encoding are addressed in [30].

Examples of this segmentation approach are shown in figure 4. This figure shows segmentation performance on synthetic, indoor and outdoor imagery. Regions of constant color in the segmentation result correspond to those regions in the intensity imagery with the same label. These images show segmentation results that are qualitatively equivalent to human segmentation, with some acceptable oversegmentation. Examples of the segmentation approach in a flight experiment are shown in figure 5 (row two). Finally, the segmentation sub-block in figure 1 shows the first and last steps of the recursive bipartition.

D. Obstacle Detection and Tracking

Obstacle detection and tracking includes boundary statistics, region reconstruction and obstacle tracking. The k-partition from section II-C defines a set of k-regions such that the kth region is a connected subgraph $G_k = (V_k, E_k)$ with nodes V_k, edges E_k, and *boundary nodes* $B_k = \{i \in V_k | (i, j) \in E_k, j \notin V_k\}$. Boundary nodes have at least one graph edge in a different region. The edge consistency constraint [20] states that a true object boundary should exhibit smoothly varying disparity. If the disparity variance for nodes B_k is above a threshold, then this partition violates edge consistency and is discarded as correspondence noise. Empirical tests show that this boundary statistic threshold is more reliable than the stereo matching threshold.

Those regions which pass the edge consistency check are reconstructed in 3D using the mean boundary disparity and linear triangulation, and are represented using the bounding ellipse of the region. Each obstacle ellipse is independently tracked using a Kalman filter, and the result is a state estimate $\hat{x}_k$ and state estimate covariance P_k for the bounding ellipse of each obstacle in the inertial frame of the vehicle. The set of obstacles that intersect a given collision volume are then passed to the control system for motion planning.

III. FLIGHT EXPERIMENTS

Flight experiments were performed the Georgia Tech GT-Max autonomous helicopter platform [18], outfitted with the VISTA flight computer and stereo cameras. The system parameters were: stereo baseline B=50cm, lens focal length f=12mm, pixel size p=12μm, image resolution=640x480 and maximum stereo disparity=32. Parameters derived from these system specs are: minimum spatial resolution of 0.3"@46.9' (0.7cm@14.3m), and a one disparity distance of 1500ft (457.2m). The VISTA flight computer consisted of the PCI
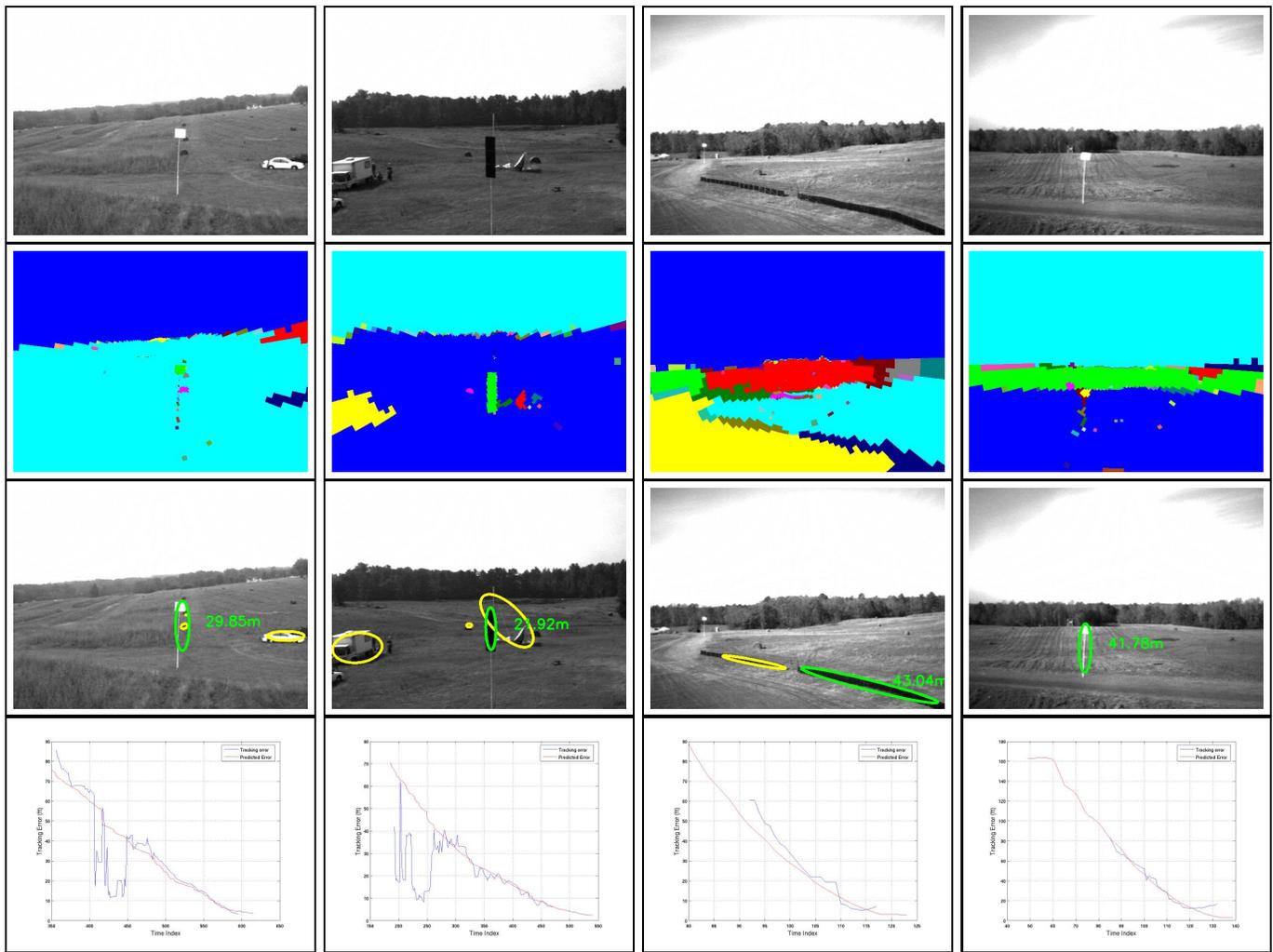
Fig. 5. VISTA flight experiment results for four collision detection scenarios (left-right): flights 1-4 (columns): Calibrated grayscale imagery from left camera, k-partition segmentation (solid color=region), Obstacle detection overlay (yellow=tracked obstacle, green=nearest collision obstacle with collision distance), Quantitative obstacle detection performance evaluation.

based Acadia I vision processor mounted on a battery powered 2.4GHz Pentium 4 motherboard. The system was attached to the RMAX helicopter and all processing was performed on board. Runtime performance of the system for each flight ranged from 5Hz-10Hz due to scene complexity affecting the total number of regions k of the recursive bipartition.

Figure 5 shows imagery and processing results from four of nineteen flight experiments. These flight experiments had the vehicle fly a true collision trajectory approaching either a "sign" obstacle (40"x30" which foamcore) or "pole" obstacle (90"x20" black foamcore) at an altitude of 21' at speeds up to 30ft/s. The obstacle detection results show that both the pole (flight 2) and sign (flights 1,3,4) obstacles are correctly detected at various collision distances, as well as other tracked obstacles shown with yellow ellipses: hay bales and cars (flight 1-2), van (flight 2), gantry (flight 2) and ground tarpaulins (flight 3). In flight 3, notice that the sign obstacle is visible far in the background, but the nearest collision obstacle is

correctly detected as a ground tarpaulin obstacle.

Figure 5 includes graphs of obstacle detection performance. The ground truth position of the collision obstacle was captured after each flight, and the obstacle estimation error was computed by comparing the tracking estimate of the green ellipse centroid to the ground truth obstacle centroid. The blue plot shows the Euclidean distance between the estimated position of the obstacle $\hat{P}$ and the ground truth position P, such that the error at time index i is $|P - \hat{P}|$. The red plot shows the *predicted* estimation error given the ground truth distance to the obstacle from the UAV navigation solution and the known range resolution of stereo. Stereo range resolution is nonlinear in range due to pixel quantization error in stereo triangulation, where the range uncertainty Δ_z for a single disparity at distance r is $\Delta_z(r) = r^2 p / Bf$. A correct detection result should exhibit error $E(r) \leq 2\Delta_z(r)$, where E is the range uncertainty due to a two pixel disparity error. The error E is also a function of the vertical and horizontal position

error, but these errors is dominated by range uncertainty for the distances considered in UAV flight, and are negligible in practice [32]. The plots show that the position estimation error does follow the expected error bound E, occasionally improving the error due to subpixel disparity estimates and disparity averaging in the filter, or worsening the error due to subpixel correspondence errors. Finally, the flight experiments were carried out under various lighting conditions, but with a fixed stereo threshold. The plots show that there are no false alarm obstacle detections due to stereo correspondence errors.

IV. CONCLUSIONS

This paper has described the Visual Threat Awareness (VISTA) system for passive stereo based obstacle detection. The VISTA system augmented the foundation of 640x480@23Hz stereo with a 5-10Hz global segmentation step that improved performance over stereo alone. Proof of concept was demonstrated in four flight experiments against real obstacles on embedded hardware with no false alarms and position estimation accuracies of $\pm6ft@50ft$. Future work includes extending the testing to include urban environments with multiple obstacles, and to integrate an online motion planner for avoidance.

V. ACKNOWLEDGEMENTS

REFERENCES

[1] R. B. et al., "Detect, see and avoid compliance using 35ghz radar for unmanned vehicles operating in national airspace," in *Proceedings of Unmanned Vehicle Systems (UVS) International*, June 2004.

[2] B. Bhanu, B. Roberts, D. Duncan, and S. Das, "A system for obstacle detection during rotorcraft low-altitude flight," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 3, pp. 785–897, July 1996.

[3] R. Cipolla and A. Blake, "Image divergence and deformation from closed curves," *International Journal of Robotics Research*, vol. 16, no. 1, pp. 77–96, 1997.

[4] T. Camus, D. Coombs, M. Herman, and T. Hong, "Real-time single-workstation obstacle avoidance using only wide-field flow divergence," *Videre: A Journal of Computer Vision Research*, vol. 1, no. 3, 1999.

[5] T. Netter and N. Franceschini, "A robotic aircraft that follows terrain using a neuromorphic eye," in *Proc. 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2002)*, 2002.

[6] S. E. Hrabar, P. I. Corke, G. S. Sukhatme, K. Usher, and J. M. Roberts, "Combined optic-flow and stereo-based navigation of urban canyons for a uav," in *Submitted to IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005.

[7] Y. Zheng, D. Jones, S. Billings, J. Mayhew, and J. Frisby, "Switcher: A stereo algorithm for ground plane obstacle detection," *IVC*, vol. 8, pp. 57–62, 1990.

[8] H. Moravec, "Robot spatial perception by stereoscopic vision and 3d evidence grids," CMU Robotics Institute, Tech. Rep. CMU-RI-TR-96-34, September 1996, http://www.seegrid.com.

[9] J. M. H. Koyasu and Y. Shirai, "Realtime omnidirectional stereo for obstacle detection and tracking in dynamic environments," in *Proc. 2001 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Maui, Hawaii, November 2001, pp. 31–36.

[10] S. E. Hrabar and G. Sukhatme, "Omnidirectional vision for an autonomous helicopter," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2003.

[11] O. Shakernia, R. Vidal, C. S. Sharpy, Y. Ma, and S. Sastry, "Multiple view motion estimation and control for landing an unmanned aerial vehicle," in *IEEE Conference on Robotics and Automation*, 2002.

[12] Y. C. A.E. Johnson and L. Matthies, "Machine vision for autonomous small body navigation," in *IEEE Aerospace conference*, 2000.

[13] T. Gandhi, M.-T. Yang, R. Kasturi, O. Camps, L. Coraor, and J. McCandless, "Detection of obstacles in the flight path of an aircraft," *IEEE Transactions on Aerospace Electronics*, vol. 39, no. 1, pp. 176–191, January 2003.

[14] D. Gavrila and V. Philomin, "Real-time object detection for smart vehicles," in *Proceedings of International Conference on Computer Vision*, 1999, pp. 87–93.

[15] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1/2/3, pp. 7–42, April-June 2002.

[16] M. Brown, Z. Burschka, and G. Hager, "Advances in computational stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 993–1008, 2003.

[17] G. van der Wal, M. Hansen, and M. Piacentino, "The acadia vision processor," in *IEEE proceedings of International Workshop on Computer Architecture for Machine Perception (CAMPS 2000)*, September 2000.

[18] E. Johnson and D. Schrage, "The georgia tech unmanned aerial research vehicle: Gtmax," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, 2003.

[19] R. Wallace, P.-W. Ong, B. Bederson, and E. Schwartz, "Space variant image processing," *International Journal of Computer Vision*, vol. 13, no. 1, pp. 71–90, 1994.

[20] Y. Ohta and T. Kanade, "Stereo by intra- and inter-scanline search using dynamic programming," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 7, no. 1, pp. 139–154, 1985.

[21] N. Pal and S. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, no. 9, pp. 1277–1294, 1993.

[22] B. P. D. Greig and A. Seheult, "Exact maximum a posteriori estimation for binary images," *Journal of the Royal Statistical Society*, vol. 51, no. 2, pp. 271–279, 1989.

[23] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 2, pp. 147–159, February 2004.

[24] Z. Wu and R. Leahy, "An optimal graph theoretic approach to data clustering: Theory and application to image segmentation," *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1101–1113, November 1993.

[25] K. L. Boyer and S. Sarkar, "Perceptual organization in computer vision: Status, challenges, and potential," *Computer Vision and Image Understanding*, vol. 76, no. 1, pp. 1–5, October 1999.

[26] J. Byrne, A. Gandhe, R. Prasanth, B. Ravichandran, M. Huff, and R. Mehra, "A k-partition, graph theoretic approach to perceptual organization," in *IEEE conference on integration of knowledge intensive multi agent systems (KIMAS '03)*, October 2003, pp. 336–342.

[27] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, August 2000.

[28] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting," *Annals of Statistics*, vol. 28, no. 2, pp. 237–407, 2000.

[29] P. Soundararajan and S. Sarkar, "Investigations of measures for grouping by graph partitioning," in *International Conference on Computer Vision and Pattern Recognition*, December 2001.

[30] J. Byrne, M. Cosgrove, and R. K. Mehra, "Vista: Visual threat awareness for an unmanned air vehicle," Scientific Systems Company Inc, Tech. Rep. SSC-05-12912-11, DARPA Final Report DAAH01-00-C-R187, April 2005.

[31] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in computer vision," in *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, ser. LNCS, vol. 2134. Springer-Verlag, September 2001, pp. 359–374.

[32] S. Blostein and T. Huang, "Error analysis in stereo determination of 3d point positions," *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 9, no. 6, pp. 752–765, November 1987.