DAYTIME WATER DETECTION AND LOCALIZATION FOR UNMANNED GROUND VEHICLE AUTONOMOUS NAVIGATION

A. L. Rankin* and L. H. Matthies
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA, USA 91109

ABSTRACT

Detecting water hazards is a significant challenge to unmanned ground vehicle autonomous off-road navigation. This paper focuses on detecting and localizing water bodies during the daytime using a stereo pair of color cameras. A multi-cue approach is taken. Evidence of the presence of water is generated from color, texture, and the detection of terrain reflections in stereo data. A ground detection algorithm is used to estimate the elevation of detected water bodies and locate them within instantaneous and world terrain maps. Temporal filtering in a world map suppresses false detections and relocates detected water as water body elevation estimates improve. This software has been implemented into a run-time passive perception system on an unmanned ground vehicle and tested at Ft. Indiantown Gap, PA.

1. INTRODUCTION

Perception systems for unmanned ground vehicles (UGVs) identify and locate terrain that is hazardous to traverse (Bellutta et al., 2000; Rankin et al., 2005a). Hazardous terrain can be identified as binary obstacles (Rankin et al., 2005b) or can be assigned a traversability cost (Lacaze et al., 2002). Water bodies are challenging terrain hazards for several reasons. Traversing through deep water bodies could cause costly damage to the electronics of UGVs. Additionally, a UGV that is either broken down due to water damage or stuck in a water body during an in-theater autonomous mission may require rescue, potentially drawing critical resources away from the primary mission and soldiers into harms way. Thus, robust water detection is a critical perception requirement for UGV autonomous navigation.

In (Matthies et al., 2003), we cataloged the environmental variables affecting the properties and conditions of surface water, and discussed the sensors applicable to detecting it under each condition. The appearance of water can greatly vary, depending upon the color of the sky, the level of turbidity, the time of day, and the presence of wind, terrain reflections, underwater objects visible from the surface, surface vegetation, and shadows. The large number of possible scenarios and appearances of water makes water detection particularly challenging using a single cue. Although laser sensors, commonly used for UGV autonomous navigation, often get no return value on free-standing water (Hong et al., 1998), fusing laser cues for water with color cues for water can increase water detectability (Hong et al., 2001).

Because there are military operations when it may be desirable for UGVs to operate without emitting strong, detectable electromagnetic signals, a passive perception solution to water detection is desirable. In (Rankin and Matthies, 2004), we described a strategy for extracting and fusing multiple cues for water from a stereo pair of color cameras. To be useful to autonomous navigation, however, water detection results must also be accurately transferred to a world map where vehicle level path planning decisions are made. In this paper, we focus on multi-cue water detection and the accurate placement of detected water into instantaneous and world terrain maps. Temporal filtering in the world map suppresses false detections and relocates water as water body elevation estimates improve.

Figure 1 illustrates the class of UGV this work has been tested on and a severe water hazard from a test range at Ft. Indiantown Gap, PA. In the following section, we summarize our approach to passive perception based multi-cue water detection.

Figure 1. An experimental unmanned vehicle (XUV) navigating on a flooded portion of a road on a test range at Ft. Indiantown Gap, PA.

2. WATER DETECTION APPROACH

Because of the multiple appearances of water, a multi-cue approach to water detection is desirable. In (Rankin and Matthies, 2004), we described techniques applicable to detecting water hazards during the daytime using passive sensors, and a strategy for fusing multiple water detection cues into a terrain map, from which a UGV can perform autonomous navigation. Here, we summarize the approach and describe some improvements to the algorithm.

The multi-cue water detector uses a rule base to combine water cues from color, texture, stereo range reflections, and zero stereo disparity. Hue, saturation, and brightness levels are thresholded to generate the water cue from color. These thresholds are tuned to detect sky reflections in water. Local image intensity variance is thresholded to generate the water cue from texture. Stereo range data is analyzed to detect patterns consistent with range reflections (for example, stereo range data on reflections of trees and other terrain extends below the ground surface). Zero disparity pixels can also provide evidence of a reflection. Zero disparity occurs when the stereo correlator matches the same column in rectified left and right images. When zero disparity pixels occur in the lower half of the disparity image, it is likely caused by reflections of objects that are far away (such as clouds or tree lines). Thus, zero disparity pixels can be a reflection-based water cue.

Figure 2 illustrates multi-cue water detection. The water body in the scene in the upper left image contains reflections of the sky and clouds, and reflections of trees and other ground cover. In addition, the leading edge of the water body is not visible in the scene. The blue regions show the water cue from color (upper right), texture (middle left), stereo range reflections (middle right), and zero stereo disparity (lower left). The lower right image shows the fused water detection image. A union of the four water cues is performed, eliminating regions that are either above the ground or horizon, small, or contain an invalid combination or proportion of water cues, according to the rule-base described in (Rankin and Matthies, 2004).

An advantage of using a multi-cue approach is that each cue can be designed to target a specific water attribute. Perfect detection of an entire water body is thus not expected by any single cue. The fusion of water cues enhances detection of water bodies with multiple attributes. The rules for fusing the water cues are designed to maximize water body detection while minimizing false detection. Since (Rankin and Matthies, 2004), the water detection algorithm has been improved in two primary ways: a new approach to detecting the ground surface has been implemented and close detection regions from a single cue are connected prior to multi-cue fusion.

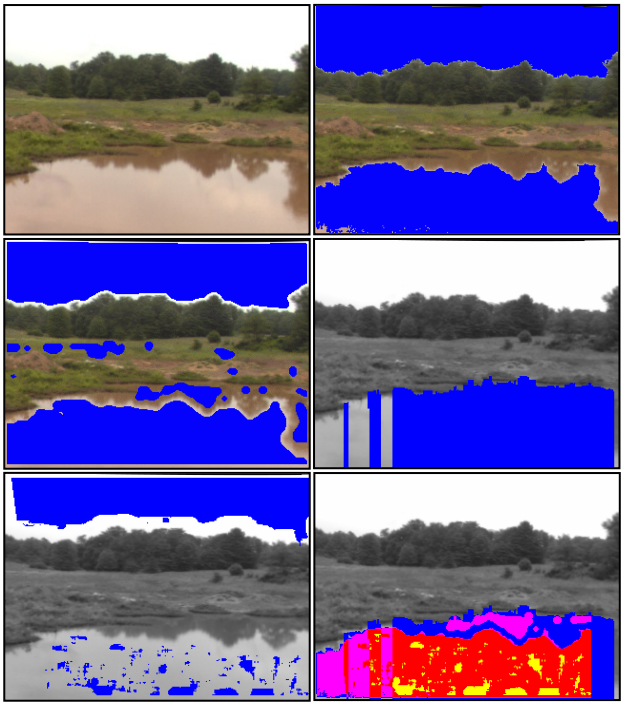

Figure 2. A multi-cue water detection approach is implemented. A rule-base is used to merge cues from color (upper right), texture (mid left), stereo reflections (mid right) and zero stereo disparity (lower left). In the fused water detection image (lower right) blue = single cue, magenta = two cues, red = 3 cues, yellow = 4 cues.

In the past, we've used a simple and fast way to determine the ground height near the leading edge of a candidate water body. For each column in a stereo range image, we stepped up the column looking for the first point that contains a significant change in range and height. This simple technique has worked well when the ground is fairly smooth. But when the ground contains vegetation that is not short, this technique is susceptible to failure. To address this problem, a new ground detector has been implemented that applies two criteria for finding the ground: 1) we expect the elevation variance to be low on the ground, and 2) we expect to find ground near the local minimum elevation.

To apply the first criteria, we create a 50m grid map and bin the range data into course voxels. Then we find the mode elevation for each grid cell and threshold elevation variance for the range data close to the mode elevation. Where there are vertical structures, such as trees, we expect the elevation variance to exceed the selected threshold. The mode elevation for grid cells containing vertical structure may very well be above the ground. Thus, mode elevation alone is not sufficient to locate the ground. We apply the second criteria to

determine if the mode elevation is close to the minimum elevation. To apply the second criteria, we generate a radial minimum elevation map that contains circular swaths centered on the vehicle's current position. Each radial swath is 2 meters wide. The minimum elevation map is filtered for spikes and interpolation is used to fill gaps. Figure 3 illustrates the pixels detected as part of the ground for a scene containing a large puddle. Accurate ground detection helps to isolate and eliminate false water detection (as water cannot occur above the ground surface).

Figure 3. Ground detection around the boundary of a water body is used to estimate the elevation of the water body.

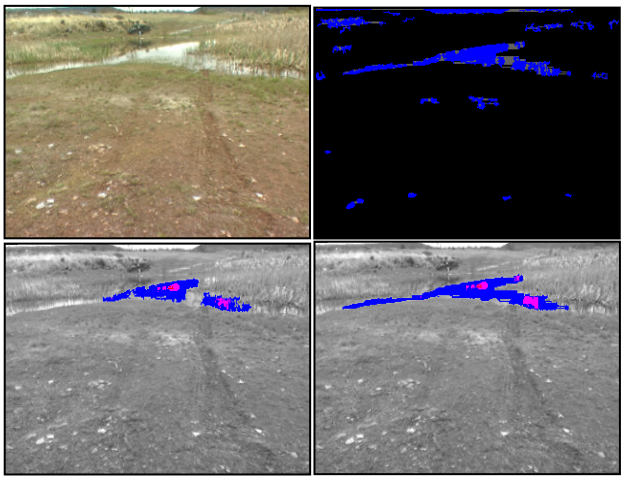

Figure 4. To improve the detection of thin water regions in image space (upper left), close cues are connected prior to region size filtering (upper right). The lower left and right images show water detection without and with this extra step, respectively.

A second modification to the algorithm has improved the detection of water bodies that are narrow in image space. Figure 4 shows water detection results for a particularly challenging water body that is narrow in image space and contains tall vegetation growing through the surface. Here, the predominant water cue is from color. Because detection regions are fragmented, the lower left fused water detection result is fragmented. But by joining close color cue regions prior to fusion (upper right image), the thin water regions in image space are preserved during fusion (lower right). In the following section, we discuss localizing detected water in terrain maps that can be used to plan safe paths during autonomous navigation.

3. WATER LOCALIZATION IN TERRAIN MAPS

Stereo reconstruction of water surfaces is not trivial due to water's specular reflectance and refractive nature. Thus, the stereo range data for water detection pixels cannot be directly used to determine the elevation of water. We estimate the elevation of water bodies by averaging the elevation of the detected ground surface around the perimeter of the fused water detection regions. Figure 5 illustrates the localization of a detected water body. Stereo range data (upper right) is used to locate the ground surface. The ground surface is then used to give each water detection pixel an elevation in an instantaneous terrain map. (An instantaneous terrain map is a snapshot of the terrain, generated from a single stereo pair of images.)

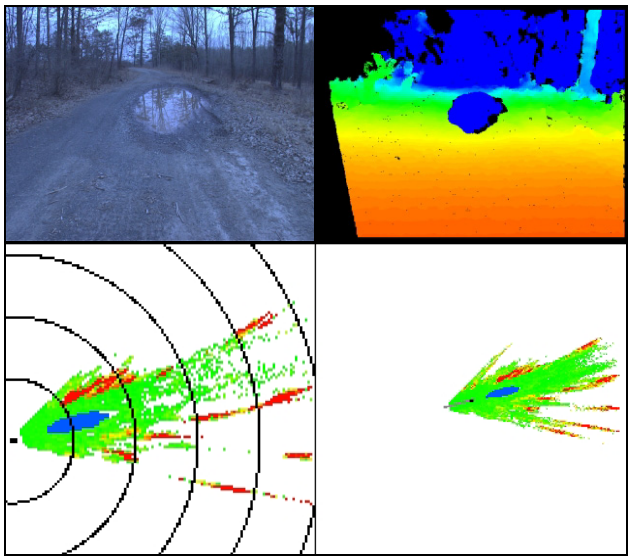

Figure 5. Water detection results transferred to a 50m instantaneous traversability cost map (lower left) and a 100m world traversability cost map (lower right) for the scene in the upper left image, and the 512x384 stereo range image (upper right).

The lower left image in Figure 5 shows a 50m x 50m birds-eye view of an instantaneous traversability cost map for the scene in the upper left image. A traversability cost for each cell was generated by using the stereo range data to determine the local slope and height of the terrain. Green terrain has a low traversability cost, red terrain is non-traversable, intermediate colors between green and red have an intermediate traversability cost, and blue represents water. The rings are in 10m increments. Instantaneous terrain maps are fused into a single 100m x 100m vehicle centered world map. This map retains a history of the terrain the vehicle has recently "seen". Information beyond 50m of the vehicle's current position,

however, falls off the world map to make room for new data. The lower right image in Figure 5 shows a birds-eye view of the world traversability cost map.

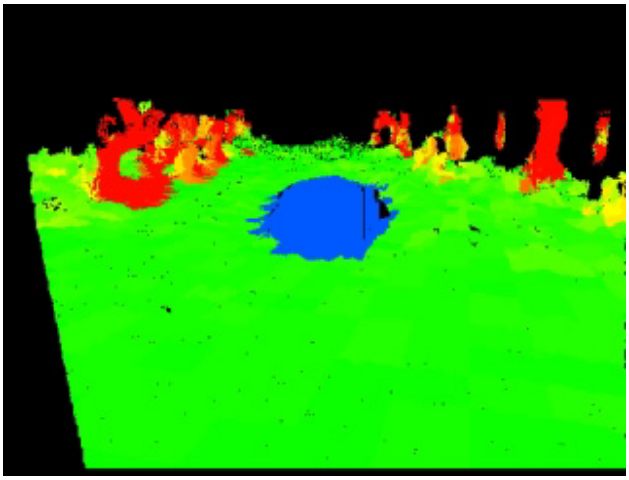

Figure 6. Traversability cost (green-red) and water detection (blue) overlaid on the upper left intensity image in Figure 5. Green has low traversability cost and red has high traversability cost.

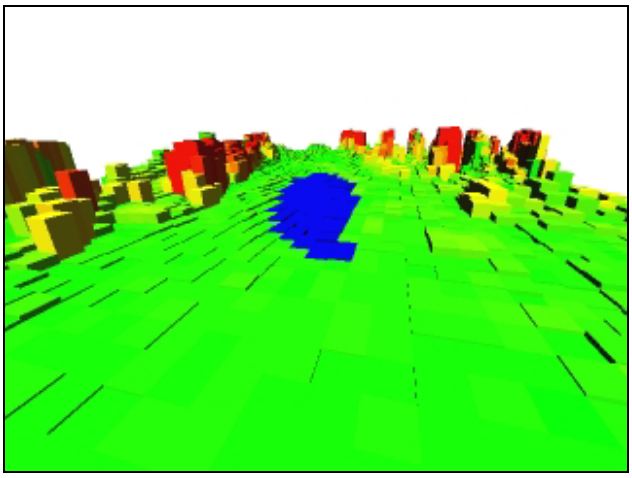

Figure 7. 3D rendering of the world map for the scene in Figure 5.

Figure 6 shows a representation of the instantaneous traversability cost map in Figure 5 projected into image space. Figure 7 shows a 3D representation of the world traversability cost map that corresponds to the scene in Figure 5. To suppress false water detection and noisy traversability cost, temporal filtering is performed in the world traversability cost map in a couple of ways. Although a history of water detection information from recent instantaneous traverasbility cost maps are maintained in the world traversability cost map at a low level, detected water is not output by the world traversability cost map until it has been "seen" at least twice in the instantaneous traversability cost maps. In addition, the last N=3 traversability costs for each cell in the world traversability cost map is averaged.

In the water detection and localization example in Figures 5-7, the UGV approached the water body normal to it. To autonomously navigate around a water body that is in a UGV's path, however, water bodies must also be detected while traveling parallel to them. Figure 8 shows a large pond at Ft. Indiantown Gap from a distance, and up close when traveling parallel to it. Figure 9 shows example water detection and localization results from traveling parallel to the water's edge. The upper left image contains multi-cue water detection results, the low left image contains a 320x240 stereo range image, the upper right image contains a 25m instantaneous terrain map, and the lower right image contains a 50m world map. In the maps, the blue regions indicate the location of the detected water. Note that there is a lapse in water detection in the world map. This occurred when the UGV momentarily steered away from the pond (causing the pond to leave the field of view). Despite the lapse, there is sufficient information in the world map to keep the UGV from trying to plan a path through the water body.

Figure 8. In order to autonomously navigate around a water body in the path of a UGV, it needs to be detected while moving perpendicular (left) and parallel (right) to it.

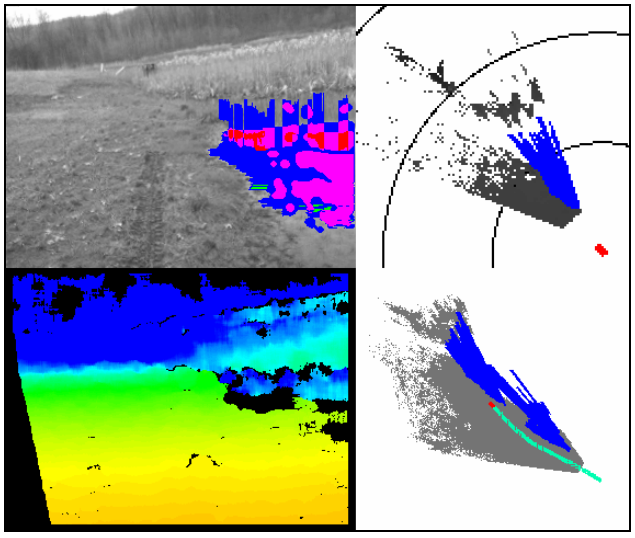

Figure 9. Example water detection and localization result while moving parallel to a water body. Also shown is a 320x240 stereo range image (lower left), a 25m instantaneous map (upper right), and a 50m world map (lower right).

4. RE-LOCALIZING WATER IN A WORLD MODEL

We rely on locating the ground surrounding a water body to estimate the elevation of a water body. When a water body's elevation is incorrect, the location of a water body in the world map will be incorrect. This is illustrated in Figure 10. While overestimating the elevation of a water body (case A) leads to locating the water body closer to the UGV, underestimating the elevation leads to locating the water body further from the UGV (case C). Underestimating the water body elevation also leads to oversizing the water body.

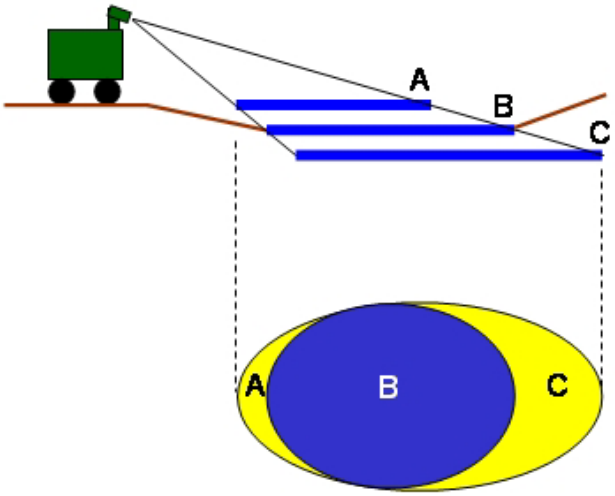

Figure 10. When the elevation of a water body is overestimated (A), the leading edge of the water body appears closer than it really is. When the elevation of a water body is underestimated (C), the trailing edge of the water body appears further than it really is.

Multi-cue water detection can occur at distances beyond which stereo range data is available. When this occurs, the closest ground elevation is used to estimate the elevation of detected water, resulting in poorer water localization at far range with improving water localization as a water body is approached. To avoid oversizing water bodies in the world map as a result of poor water localization at far range, the elevation and location of a water body is updated in the world map each cycle in real time.

Figure 11 illustrates the benefits of re-localizing water during a long drive up to a water pond's leading edge. The upper right image shows the multi-cue water detection for the scene in the upper left image. This image is the final frame in the drive up to the water body. In the world map in the bottom image, yellow represents all the map cells detected as water during the drive, and blue represents the map cells containing water after updating the water's location. Without re-localizing the detected water in the world map, the world map would indicate the UGV is in the water at the end of the drive, when clearly it is not.

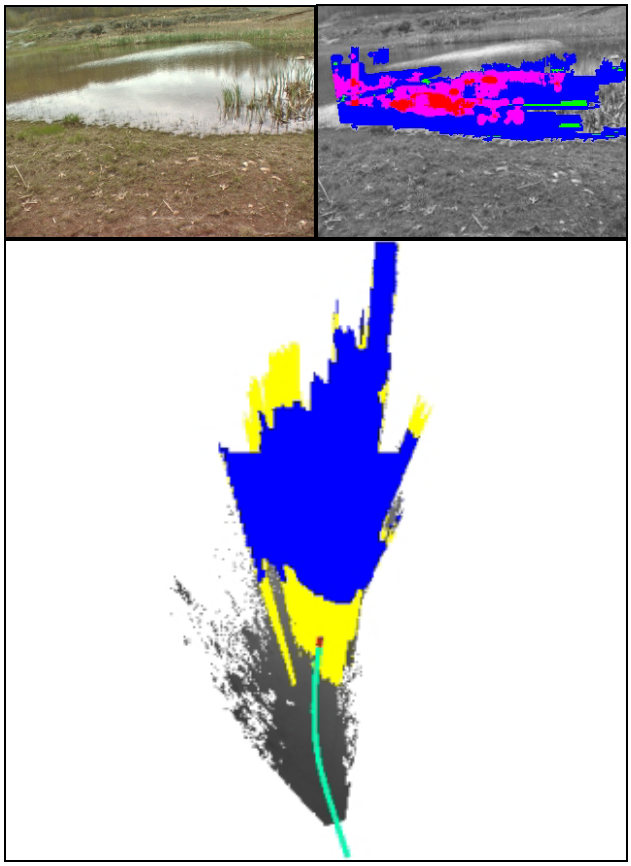

Figure 11. An example of relocating detected water in the world map over time as the water's elevation estimate improves. In the world map in the second row, yellow indicates cumulative water placement during a head on approach. Blue indicates water placement after relocalization. Note that without water relocalization, the world map indicates the UGV is in the water in the last frame (first row), when clearly it is not.

CONCLUSIONS

Robust water detection is a critical perception requirement for UGV autonomous navigation. In this paper, we have summarized an approach to detecting water hazards during the daytime using passive sensors and localizing them in instantaneous and world maps from which a UGV can perform autonomous navigation. Our multi-cue water detector uses a rule base to combine water cues from color, texture, stereo range reflections, and zero stereo disparity. It is robust to narrow water regions in image space. In addition, it rejects small regions, regions above the ground or horizon, and regions

that contain an invalid combination or proportion of water cues.

Because of water's specular reflectance and refractive nature, stereo yields unreliable range data on the surface of water for water localization purposes. We estimate the elevation of a water body by averaging the elevation of the detected ground surface around the perimeter of the fused water detection region. Using this elevation, water is transferred to instantaneous and world maps. Temporal filtering in the world map is used to suppress false detections and relocates detected water as water body elevation estimates improve. This software has been implemented into a run-time passive perception system on a UGV and tested on a 6.9km robotic vehicle test course (called the *Forever Loop*) at Ft. Indiantown Gap, PA.

The *Forever Loop* consists of dirt and gravel trails that weave through cross country terrain containing forested regions, flat and hilly regions, tall and short vegetation, and several ponds (one of which is shown in Figure 8). In March 2006, 12,265 stereo image pairs were collected during XUV teleoperation of the entire course. During this time of year, the ponds were all dry, but there were three puddles on the course. All three of the puddles were consistently detected. Two small traversable puddles were detected in 12 consecutive world maps, starting at a distance of 7 meters. A larger traversable puddle (shown in Figure 5) was detected in 59 consecutive world maps, starting at a range of 13 meters. False positive water detection occurred in 26 of the 12,265 world maps (0.2%). In 24 of these maps, the false positive water detection occured on nontraversable log barriers lining the sides of a trail. In the other 2 world maps, the false positive water detection occurred in tall vegetation lining the side of a trail. On a Linux Dell Precision computer containing an Intel Xeon 3.4GHz processor, the water detection and localization software runs at approximately 1Hz when stereo images are processed at a resolution of 512x384 pixels.

FUTURE WORK

In fiscal year 2007, we have proposed to extend the range of water detection by performing reflection detection in the visible imagery, and to address the problem of detecting mud hazards with UGV sensors. In addition, we will use the shape and size of a detected water body to estimate the level of hazard it poses to a UGV.

ACKNOWLEDGEMENTS

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, and was sponsored by the Army Research Laboratory Collaborative Technology Alliances (CTA) Program through an agreement with the National Aeronautics and Space Administration. Reference herein to any specific commercial product, process, or service by trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

REFERENCES

Bellutta, P., Manduchi, R., Matthies, L., Owens, K., and Rankin, A., 2000: Terrain Perception for Demo III, *Proceedings of the 2000 Intelligent Vehicles Conference*, Dearborn, MI, 326-331.

Hong, T., Legowik, S., and Nashman, M., 1998: Obstacle Detection and Mapping System, *National Institute of Standards and Technology (NIST) Technical Report NISTIR 6213*, 1-22.

Hong, T., Rasmussen, C., Chang, T., Shneier, M., 2001: Fusing Ladar and Color Image Information for Mobile Robot Feature Detection and Tracking, *7th International Conference on Intelligent Autonomous Systems*, Marina Del Ray, CA.

Lacaze, A., Murphy, K., and Del Giorno, M., 2002: Autonomous Mobility for Demo III Experimental Unmanned Vehicles, *Proceedings of the AUVSI Symposium on Unmanned Systems,* Orlando.

Matthies, L., Bellutta, P., and McHenry, M., 2003: Detecting Water Hazards for Autonomous Off-Road Navigation, *Proceedings of SPIE Conference 5083: Unmanned Ground Vehicle Technology V*, Orlando, 231-242.

Rankin, A., Bergh, C., Goldberg, S., Bellutta, P., Huertas, A., and Matthies, L., 2005a: "Passive perception system for day/night autonomous off-road navigation", *Proceedings of the SPIE Defense and Security Symposium: Unmanned Ground Vehicle Technology VI Conference*, Orlando, 343-358.

Rankin, A., Huertas, A., and Matthies, L., 2005b: "Evaluation of Stereo Vision Obstacle Detection Algorithms for Off-Road Autonomous Navigation", *Proceedings of the 32nd AUVSI Symposium on Unmanned Systems*, Baltimore.

Rankin, A., and Matthies, L., 2004: "Daytime Water Detection by Fusing Multiple Cues for Autonomous Off-Road Navigation", *Proceedings of the 24th Army Science Conference*, Orlando.