different solutions so that the most effective approach won.

Taking place in the inhospitable Mojave Desert and requiring computing engines that could operate continuously without human intervention, the races forced teams to address reliability and performance in extreme environments. In particular, the races required systems that could perform in high temperatures and dusty conditions. Teams selected reliable sensors, placing them in positions where accidents or rocks were less likely to damage them. The systems also required cleaning devices to keep sensors free of dust, water, mud, and other debris.

VisLab at the Grand Challenge

Claudio Caraffi and Stefano Cattani
VisLab, Università degli Studi di Parma

VisLab, together with Oshkosh Truck Corp. and Rockwell Collins, developed and successfully fielded TerraMax, an autonomous vehicle that reached the finish line of the 2005 DARPA Grand Challenge. TerraMax uses artificial vision, laser scanners, GPS, inertial sensors, and map databases to sense and understand its environment.[1] VisLab developed its vision system for obstacle detection and drivable path sensing.

TerraMax's three-camera system allows precision and efficient computation at a wide range of viewing distances. The cameras sit on a rigid bar over the vehicle hood. By selecting two cameras at a time, the system can get stereo pairs with different baselines—that is, intercamera distances.

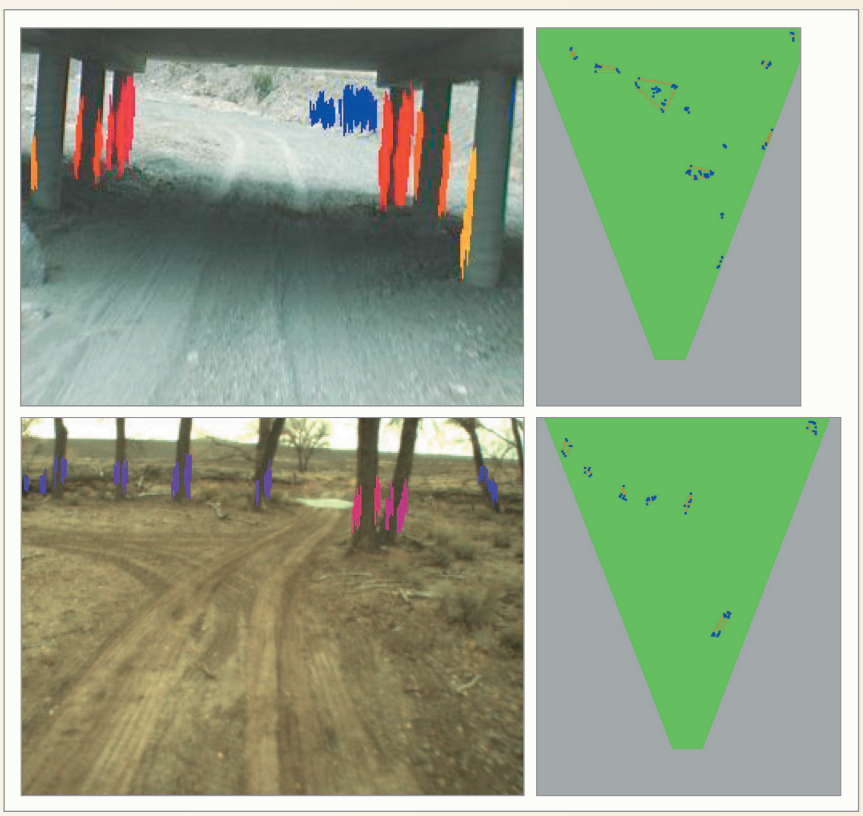

Figure A. Obstacle detection for different scenarios. The color of the obstacles varies with distance. The images on the right map the obstacles in a bird's eye view. The camera's field of view is plotted in green.

During the DARPA Grand Challenge, TerraMax selected the baseline based on vehicle speed. Higher speeds required greater sensing distances and thus wider baselines. We developed image stabilization to overcome vehicle oscillations from off-road environment terrain bumps.

The system relies on a two-step approach for full 3D stereo reconstruction. The first step exploits V-disparity image properties to stabilize images. The V-disparity images are 3D graphical representations of the similarity measures between left and right image rows depending on the shifts used to compare them.[2] V-disparity images contain basic information about the ground's position and shape, taking slope changes into consideration. In this way, the system can estimate the vehicle's pitch at the time of acquisition and stabilize images.

As a second step, using the information about the ground, the system addresses the correspondence problem by computing a disparity search image (DSI). The correspondence problem is the process of finding which pixel in the left image matches which pixel in the right image of a stereo image pair. We compute the DSI using small confrontation windows to allow detection of thin obstacles that other sensors missed. We apply a series of filters to the DSI, highlighting disparity concentrations that are detected as obstacles. We compute the obstacle 3D world coordinates via stereo triangulation. Figure A shows an example of this process.

Code optimizations exploited the processors' MMX and Streaming SIMD Extensions (SSE) instruction set. During the DARPA Grand Challenge, obstacle detection ran with a guaranteed 15-Hz throughput, although it could perform the entire computation within 30 milliseconds on a Pentium IV using a 2.8-GHz processor system, allowing other image-processing routines to run on the same machine.

Along with obstacles, the vision system also provides drivable path

Five prototypes completed the 2005 race. The winning team, Stanford University's "Stanley," mainly used laser-scanner and GPS sensors. H1ghlander and Sandstorm employed a similar approach and technology; in addition to the information coming from onboard sensors, they also exploited a precise premission-planning strategy. The Kat-5 vehicle used lasers as primary sensors as well.

As the "VisLab at the Grand Challenge" sidebar describes, the TerraMax prototype—a Medium Tactical Vehicle Replacement (MTVR) truck that Oshkosh Truck Corp., Rockwell Collins, and VisLab developed—

information. Most of the algorithms used for path detection look for a single homogeneous road surface in front of the vehicle. Since the hypothesis of homogeneity becomes a huge limitation because it bounds the set of detectable roads to the case of medium/well-structured environments, we tried to generalize the problem, considering that roads can also be made of patches of heterogeneous surfaces. To find these potentially heterogeneous surfaces, the algorithm looks for a variable number of small homogeneous terrain portions. They can represent any kind of natural or artificial environment elements, such as gravel or paved roads, grass, water puddles, oil stains, drivable rocks, lane markers, and shadows. Consequently, it's possible to summarize the path-detection algorithm as a two-step process:

1. Divide the image in homogeneous regions made of connected pixels.
2. Decide which combination of the obtained regions could represent the road surface with the highest probability.

Researchers have successfully studied the first step, clustering, using both evolutionary and traditional approaches. However, the Grand Challenge's real-time constraints led to the adoption of a simple—but fast and easily tunable—clustering algorithm as a good tradeoff between performance and computational requirements.

The second step falls into the class of decision problems. The decision process we developed tries to minimize the risk of wrong classifications, taking into account the current vehicle state. The underlying idea is that each cluster belongs to the road with a given probability depending only on its own intrinsic properties: homogeneity, size, shape factors, and covered free-space area. Figure B shows the clusters classified as road and off-road. However, a high probability of being road isn't sufficient, and sometimes not even necessary, to be finally classified as road. In fact, a sensible road classifier's principal goal is to minimize the risk associated with an incorrect classification on the basis of the current vehicle state. The following rule applies to classification decisions: Those with higher risks (requiring sudden changes in vehicle behavior) need higher probabilities of being correct before they can be assigned. Decision networks[3] extend Bayesian networks and provide a general methodology for rational decision making that fits the problem of deciding about the set of clusters that belongs to the road surface.

Path detection provided results in accordance with the output of the laser scanner and extended the perception range beyond the limits of the laser scanner, which is affected by vehicle pitch.

References

1. D. Braid, A. Broggi, and G. Schmiedel, "The TerraMax Autonomous Vehicle," *J. Field Robotics*, vol. 23, no. 9, 2006, pp. 693-708.
2. A. Broggi et al., "The Single-Frame Stereo Vision System for Reliable Obstacle Detection Used During the 2005 DARPA Grand Challenge on TerraMax," *Proc. Intelligent Transportation System Conf.* (ITSC 06), IEEE Press, 2006, pp. 745-752.
3. A. Broggi et al., "A Decision Network-Based Framework for Visual Off-Road Path-Detection Problem," *Proc. Intelligent Transportation System Conf.* (ITSC 06), IEEE Press, 2006, pp. 951-956.

Claudio Caraffi *is a PhD student in the Department of Information Technology at Università degli Studi di Parma. Contact him at caraffi@ce.unipr.it.*

Stefano Cattani *is a PhD student in the Department of Information Technology at Università degli Studi di Parma. Contact him at cattani@ce.unipr.it.*

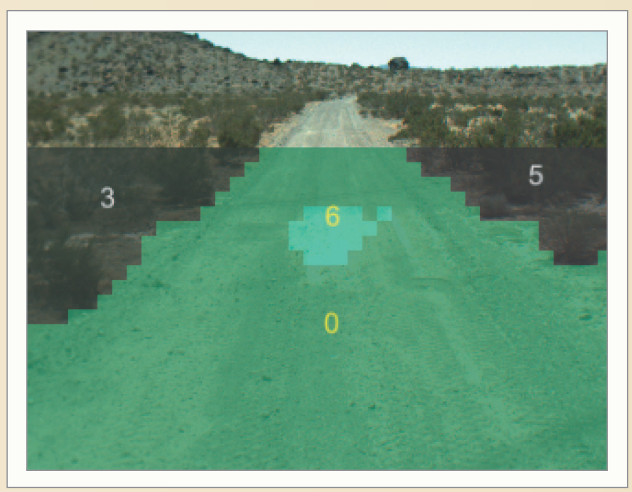

Figure B. White numbers denote off-road clusters, and yellow numbers denote road clusters. The interest area is limited to the region within 50 meters from the vehicle. This information is obtained from the preliminary processing of the obstacle-detection module.