Application Challenges from a Bird’s-Eye View

Davide Scaramuzza

University of Zurich, Robotics and Perception Group

6.1 Introduction to Micro Aerial Vehicles (MAVs)

An Unmanned Aerial Vehicle (UAV), commonly known as a drone, is an aircraft without a human pilot aboard. The International Civil Aviation Organization (ICAO) of the United Nations classifies UAVs into two types: (i) autonomous aircrafts, (ii) remotely piloted aircrafts. UAVs were initially conceived for military applications but in the recent years we have witnessed also a growing number of civil applications, such as law enforcement and firefighting, security and surveillance, agriculture, aerial photography, inspection, and search and rescue.

Micro Aerial Vehicles (MAVs)

With the term Micro Aerial Vehicle (MAV) is meant a miniature UAV that is less than 1 meter in size and below 2kg in weight. Some MAVs can even be as small as a few centimeters and weigh only a few grams (c.f. Ma et al. (2013); Troiani et al. (2013)).

MAVs can be seen as the logical extension of ground mobile robots. Their ability to fly allows them to easily avoid obstacles on the ground and to have an excellent bird-eye view. MAVs can be classified into rotorcrafts (or rotary wing), fixed or flapping wing, or hybrid (c.f. Figure 6.1).

Rotorcraft MAVs

Small rotorcrafts have several advantages compared to those based on fixed-wings: they are able to take off and land vertically, hover on a spot, and even dock to a surface (c.f. Kumar and Michael (2012)). This capability allows them to navigate easily in unstructured, indoor environments (Shen et al. (2012)), pass through windows (Achtelik et al. (2009)), traverse narrow corridors (Zingg et al. (2010b)), climb stairs (Bills et al. (2011b)), and navigate through or over damaged buildings for rescue or inspection operations (Faessler
et al. (2015b); Michael et al. (2012b)). Thus, they are the ideal platform for exploration, mapping, and monitoring tasks in search-and-rescue and remote-inspection scenarios.

Multicopter MAVs come usually in the form of quadrotors (also known as quadcopters), hexacopters, or octocopters and have matched sets of rotors turning in opposite directions. The smaller the number of rotors, the better the efficiency of the vehicle. On the other hand, the achievable dynamics and, therefore, the maneuverability of the vehicle can be enhanced by a larger number of propellers and a smaller ratio between rotor surface and total weight (Achtelik et al. (2012)). Additionally, hexacopters and octocopters offer redundancy against single-rotor failure. However, quadrotors have become very successful nowadays because of their relatively-simple design.

6.2 GPS-denied Navigation

To date, most autonomous MAVs rely on GPS to navigate outdoors. However, GPS may not be reliable in case of low satellite coverage or multipath, two phenomena that are very frequent in urban settings when flying at low altitudes and close to buildings. Furthermore, GPS is completely unavailable indoors, thus limiting the use of drones in search-and-rescue or remote-inspection operations. At the current state, most MAVs used in search and rescue-and-remote inspection-scenarios are teleoperated under direct line of sight with the operator (c.f. Murphy (2014)). If wireless communication with the MAV can be maintained, there is the possibility to teleoperate the MAV by transmitting video streams from onboard cameras to the operator. However, teleoperation from video streams is extremely challenging in indoor
environments. Furthermore, wireless communication cannot be guaranteed after a certain range. For these reasons, there is a large need of flying robots that can navigate autonomously, without any user intervention.

The key problem in MAV navigation is attitude and position control. Today’s systems handle well the attitude control using proprioceptive sensors, such as Inertial Measurement Units (IMU). However, without position control, they are prone to drift over time. In GPS-denied environments, this can be solved using offboard sensors (such as motion-capture systems) or onboard sensors (such as cameras and laser rangefinders). Motion-capture systems (e.g., Vicon or OptiTrack) consist of a set of external cameras mounted on the ceiling, which track the position of the robots with submillimeter accuracy and at high frame rates (more than 350 Hz). They are very appropriate for testing and evaluation purposes (c.f. Lupashin et al. (2014); Michael et al. (2010b)), such as prototyping control strategies or fast maneuvers, and serve as a ground-truth reference for other localization approaches. However, for truly autonomous navigation in unknown, yet-unexplored environments, sensors should be installed onboard.

A journal special issue on MAV onboard perception and control was published by Michael et al. (2012a). The literature can be divided into approaches using range sensors (e.g., lidars or RGB-D sensors) and camera sensors.

6.2.1 Autonomous Navigation with Range Sensors

Lidar has been largely explored for ground mobile robots (c.f. Thrun et al. (2007)) and similar strategies have been extended to MAVs (c.f. Achtelik et al. (2009); Bachrach (2009)). Using an RGB-D camera and a 2D laser, multi-floor mapping results have recently been demonstrated using an autonomous quadrotor (c.f. Shen et al. (2012), Figure 6.2). Although lidars and RGB-D sensors are very accurate and robust, they are still too heavy and consume too much power for lightweight MAVs. Therefore, cameras are the only viable sensors in the medium-long term; however, they require external illumination to see, and a certain computing power to extract meaningful information for navigation.

6.2.2 Autonomous Navigation with Vision Sensors

Reactive Navigation

Most works on vision-based reactive navigation of MAVs have relied on biologically-inspired vision algorithms, such as optical flow (c.f. Floreano et al. (2009); Hrabar and Sukhatme (2009); Ruffier and Franceschini (2004); Zufferey (2009)). Optical flow has been applied to MAVs for tasks such as on-spot hovering, take-off, landing, and, more generally, reactive navigation (e.g., for obstacle avoidance or to keep the MAV in the center of a canyon by balancing the optical flow on both sides of the robot field of view). While optical flow is crucial for reactive navigation, it cannot be used for precise maneuvers, such as trajectory-following. Furthermore, optical flow only measures the relative velocity, leading the MAV to inevitably drift over time. Nevertheless, due to the limited computational power required by optical flow, this approach has been successfully integrated in several commercial drones, such as the Parrot AR.Drone and the SenseFly products, for autonomous hovering and landing.
Map-based Navigation

The alternative to reactive navigation is map-based navigation, which proved very successful for ground mobile robots equipped with laser rangefinders (c.f. Thrun et al. (2007)). Breakthrough work on vision-controlled map-based navigation of MAVs was done within the European project SFLY (Scaramuzza et al. (2014)), where visual-SLAM (Simultaneous Localization And Mapping) pipelines (e.g., Chiuso et al. (2002); Davison et al. (2007); Forster et al. (2014b); Klein and Murray (2007)) were used in combination with inertial sensors to enable autonomous basic maneuvers, such as take-off and landing, trajectory following, and surveillance coverage. Building upon that work, several vision-based systems have been proposed using both monocular (c.f., Achtelik et al. (2011); Brockers et al. (2014); Forster et al. (2014b); Weiss et al. (2013)) and stereo-camera configurations (c.f. Achtelik et al. (2009); Fraundorfer et al. (2012); Meier et al. (2012); Schmid et al. (2014); Shen et al. (2013b)).

6.2.3 SFLY: Swarm of Micro Flying Robots

The Swarm of Micro Flying Robots (SFLY) project\footnote{Project website: www.sfly.org} (Scaramuzza et al. (2014)) was a European-Union-funded project with the goal of creating a swarm of vision-controlled MAVs capable of autonomous navigation, 3D mapping, and optimal surveillance coverage in GPS-denied environments. The SFLY MAVs did not rely on remote control, radio beacons, or motion-capture systems but could fly all by themselves using only a single onboard camera and an IMU.

The first contribution of the SFLY was the development of a new hexacopter equipped with enough processing power for on board computer vision. The hexacopter was designed and manufactured by Ascending Technology and later sold under the name of Firefly, which

---

*Figure 6.2* (Left) Autonomous MAV exploration of an unknown, indoor environment using RGB-D sensor (image courtesy of Shen et al. (2012)). (Right) Autonomous MAV exploration of an unknown, indoor environment using a single onboard camera (image courtesy of Faessler et al. (2015b)).
has become very popular. The second contribution of the SFLY was the development of a local-navigation module based on the Parallel Tracking and Mapping (PTAM) framework by Klein and Murray (2007) that run in real-time on board the MAV (an Intel Core 2 Duo). The output of PTAM was fused with inertial measurements (c.f. Weiss et al. (2012)) and was used to stabilize and control the MAV locally without any link to a ground station. The third contribution was an offline dense-mapping process that merges the individual maps of each MAV into a single, global map that serves as input to the global navigation module (c.f. Forster et al. (2013)). Finally, the fourth contribution was a Cognitive, Adaptive Optimization (CAO) algorithm to compute the positions of the MAVs, which allowed the optimal surveillance coverage of the explored area (c.f. Doitsidis et al. (2012)). Experimental results demonstrating three MAVs navigating autonomously in an unknown GPS-denied environment and performing 3D mapping and optimal surveillance coverage were presented. A detailed description of the SFLY can be found in (Scaramuzza et al. (2014)). Open-source code is publicly available to the robotics community.

6.2.4 SVO, a visual-odometry algorithm for MAVs

A visual-odometry and mapping algorithm, named SVO, specifically designed for MAV navigation with computationally-limited computers, such as Odroid, was recently proposed by (Forster et al. 2014b). Contrary to state-of-the-art visual-odometry and SLAM algorithms relying on costly feature extraction and matching pipelines (c.f. (Davison et al. 2007; Klein and Murray 2007)), SVO (Semi-direct Visual Odometry) uses a combination of features and direct methods (from which the nickname “semi-direct”) to achieve unprecedented real-time performance (up to 70 fps on Odroid boards and more than 400 fps on an i7 laptop) and high-precision visual odometry (less than 1% drift). The semi-direct approach eliminates the need for costly feature extraction and robust matching techniques for motion estimation. The algorithm operates directly on pixel intensities, which results in subpixel precision at high frame-rates. Precise and high frame-rate motion estimation brings increased robustness in scenes characterized by little, repetitive, and high-frequency textures.

SVO uses a probabilistic mapping method that explicitly models outlier measurements to estimate 3D points; this results in fewer outliers and more reliable points (c.f. Figure 6.3). Image points are triangulated from multiple views using recursive Bayesian estimation. This probabilistic depth estimation allows using every image for incremental depth estimation and provides a depth uncertainty that can be directly used for path planning.

Figure 6.3  Probabilistic depth estimate in SVO. Very little motion is required by the MAV (marked in black at the top) for the uncertainty of the depth-filters (shown as magenta lines) to converge. Image courtesy of (Faessler et al. 2015b).
SVO has been so far used for MAV state-estimation in GPS-denied environments in combination with inertial sensors and runs on the onboard embedded computer. The integration of SVO onboard on an MAV, its fusion with the IMU, and use for closed-loop control and navigation are detailed in (Faessler et al. 2015b). Open-source code is publicly available to the robotics community4. Instructions on how to integrate the SVO position measurements into the popular PX4 autopilot are provided on the PX4 webpage.5

6.3 Applications and Challenges

6.3.1 Applications

Drones have several applications in search-and-rescue, remote inspection, law enforcement, video surveillance, agriculture, aerial photography, photogrammetry, mapping, entertainment, and parcel delivery. However, localization and position tracking is not the sole use of vision sensors. In agriculture, for instance, drones with high-resolution spectral imaging devices are used to gather insight of crops, thus allowing for targeted fertilizing and better use of water and labor. This information can then be used to reduce the need of common fertilizers, which typically pollute local waterways. The main drone-based observation technique is called Normalized Difference Vegetation Index, a measure that assesses the crop productivity, which is calculated on the basis of visible and infrared radiation. When crops are viewed from a standard camera, crops normally look like an indistinct green and brown mass; however, when viewed with an infrared camera many colors suddenly pop out, such as yellow, orange, red, and green; software then stitches together hundreds of images to form a complete picture. In architecture, archeology, geography, and nature conservation, drones are used as mapping tools to get high-resolution 3D models of a construction, building, or terrain. The drones are usually set to take pictures at regular time intervals and a trajectory is planned through GPS. The images must be then downloaded to a laptop PC and powerful photogrammetry software, such as Pix4D or Agisoft, uses state-of-the-art structure-from-motion tools to build dense, photorealistic 3D models with centimeter accuracy. This mapping technology is also used for disaster management to get an overview picture after a flood or an earthquake. Finally, drones are also used as a remote camera in video surveillance and inspection. A live video stream is sent wirelessly from the drone to a tablet screen or video glasses, which are utilized as a feedback to the operator.

In the applications listed above, drones use GPS to navigate autonomously or are remotely operated by an expert pilot. In order to authorize the operation of autonomous drones in different countries in the near future, several challenges need to be overcome in terms of safety and robustness. Furthermore, additional sensors should be used than just cameras and GPS, such as lidars, radars, sonars, thermal cameras, and so on. Redundancy allows coping with sensor failures and operation in harsh conditions, such as night, low-light, smoke, and so on. Since the focus of this book is on computer vision, we will review works dealing with safety and robustness of MAVs using mainly vision sensors.

4https://github.com/uzh-rpg/rpg_svo
5https://pixhawk.org/dev/ros/visual_estimation
6.3.2 Safety and Robustness

If a quadrotor's vision pipeline fails, there is typically a small set of options left: (i) a pilot must take over; (ii) the quadrotor must land immediately; (iii) the quadrotor must use simple fail-backs for stabilization in order to continue its mission. In the following two sections, the state-of-the-art research on failure recovery and emergency landing is reviewed.

Failure Recovery

In (Shen 2014), a linear sliding window formulation for monocular visual-inertial systems was presented to make a vision-based quadrotor capable of failure recovery and on-the-fly initialization. The approach assumed that visual features could be extracted and correctly tracked right from the beginning of the recovery procedure.

Along with possible failures of their state-estimation pipeline, monocular-vision–based quadrotors present the drawback that they typically require an initialization phase before they can fly autonomously. This initialization phase is usually performed by moving the quadrotor by hand or via remote control. Since this is time consuming and not easy to perform, attempts have been made to perform the initialization automatically. For instance, in ((Brockers et al. 2014; Weiss et al. 2015)), the authors presented a system that allows the user to toss a quadrotor in the air, where it then initializes a visual-odometry pipeline. Nevertheless, that system still required several seconds for the state estimate to converge before the toss and several more seconds until the visual-odometry pipeline was initialized. A closed-form solution for state estimation with a visual-inertial system that does not require initialization was presented in (Martinelli 2012). However, at the current state of the art this approach is not yet suitable for systems that rely on noisy sensor data.

A system enabling a monocular-vision–based quadrotor to autonomously recover from any initial attitude and quickly re-initialize its visual-inertial system was recently proposed by (Faessler et al. 2015a) and demonstrated in a scenario where a quadrotor is thrown in the air (c.f. Figure 6.4). In contrast to (Shen 2014), their system did not require the observation of visual features at the beginning of the recovery procedure but only once its attitude is stabilized, which simplifies feature tracking greatly and reduces computational complexity. In contrast to (Brockers et al. 2014) and (Weiss et al. 2015), no preparation time before launching the quadrotor was required and the entire recovery was performed more quickly.

Emergency Landing

Early works on vision-based autonomous landing for Unmanned Aerial Vehicles (UAV) were based on detecting known planar shapes (e.g., helipads with "H" markings) in images (c.f. (Saripalli et al. 2002)) or on the analysis of textures in single images (c.f. (Garcia-Pardo et al. 2002)). Later works (e.g., (Bosch et al. 2006; Desaraju et al. 2014; Johnson et al. 2005)) assessed the risk of a landing spot by evaluating the roughness and inclination of the surface using 3D terrain reconstruction from images.

One of the first demonstrations of vision based autonomous landing in unknown and hazardous terrain is described in ((Johnson et al. 2005)). Structure-from-motion was used to estimate the relative pose of two monocular images and, subsequently, a dense elevation map was computed by matching and triangulating regularly sampled features. The evaluation of the roughness and slope of the computed terrain map resulted in a binary classification of
Figure 6.4  Autonomous recovery after throwing the quadrotor by hand: (a) the quadrotor detects free fall and (b) starts to control its attitude to be horizontal. Once it is horizontal, (c) it first controls its vertical velocity and then, (d) its vertical position. The quadrotor uses its horizontal motion to initialize its visual-inertial state estimation and uses it (e) to first break its horizontal velocity and then (f) lock to the current position. Image courtesy of (Faessler et al. 2015a).
safe and hazardous landing areas. This approach detected the landing spot solely based on two selected images rather than continuously making depth measurements and fusing them in a local elevation map.

In (Bosch et al. 2006), homography estimation was used to compute the motion of the camera as well as to recover planar surfaces in the scene. A probabilistic two-dimensional grid was used as map representation. The grid stored the probability of the cells being flat.

While previously mentioned works were passive in the sense that the exploration flight was pre-programmed by the user, recent work by (Desaraju et al. 2014) was done on how to actively choose the best trajectory autonomously to explore and verify a safe landing spot. However, due to computational complexity, the full system could not run entirely on-board in real-time. Thus, outdoor experiments were processed on datasets. Additionally, only two frames were used to compute dense motion stereo; hence a criterion, based on the visibility of features and the inter-frame baseline, was needed to select two proper images.

A real-time approach running fully onboard an MAV was recently proposed by (Forster et al. 2015) (c.f. Figure 6.5). The authors proposed to generate a 2D elevation map that is probabilistic, of fixed size, and robot-centric, thus, always covering the area immediately underneath the robot. The elevation map is continuously updated at a rate of 1 Hz with depth maps that are triangulated from multiple views using recursive Bayesian estimation. This probabilistic depth estimation not only allows using every image for incremental depth estimation but also provides a depth uncertainty that can be directly used for planning trajectories minimizing the depth uncertainty as fast as possible, as proposed by (Forster et al. 2014a).

6.4 Conclusions

This chapter gave a description of the challenges of GPS-denied autonomous navigation of drones. Laser-based SLAM can outperform by several orders of magnitude the precision of GPS; however, laser rangefinders consume too much power and are too heavy for lightweight micro drones. The chapter then presented alternative techniques, based on visual odometry and SLAM technologies as a viable replacement of laser-based navigation. However, they require external illumination and sufficient texture in order to work reliably. The optimal sensor suit of a drone should be a combination of GPS, laser, ultrasound, and vision sensors (both standard and infrared) to provide sufficient redundancy and success in different environment conditions. However, robustness to changes in the environment and how to handle system failures still remains an open challenge for both engineers and researchers.
Figure 6.5  (a) A quadrotor is flying over a destroyed building. (b) The reconstructed elevation map. (c) A quadrotor flying in an indoor environment. (d) The quadrotor executing autonomous landing. The detected landing spot is marked with a green cube. The blue line is the trajectory that the MAV flies to approach the landing spot. Note that the elevation map is local and of fixed size; its center lies always below the quadrotor’s current position. Image courtesy of (Forster et al. 2015).
References


Adnirulka M, Roth S and Schiele B 2008 People-tracking-by-detection and people-detection-by-tracking Proceedings of the conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA.


Bachrach A 2009 Autonomous flight in unstructured and unknown indoor environments Master’s thesis Massachusetts Institute of Technology.


Badino H 2004 A robust approach for ego-motion estimation using a mobile stereo platform Proceedings of the IWC, Workshop on Complex Motion, Ginzburg, Germany.

Badino H, Franke U and Pfeiffer D 2009 The stixel world - a compact medium level representation of the 3D-world Proceedings of the German Conference on Pattern Recognition, Jena, Germany.


Broggi A, Bertozi M, Fasoli A and Sechi M 2000 Shape-based pedestrian detection Proceedings of the IEEE Intelligent Vehicles Symposium, Dearborn, MI, USA.


Cohen-Steiner D and Da F 2004 A greedy Delaunay-based surface reconstruction algorithm. The Visual Computer 20, 4–16.


Davis J 1998 Mosaics of scenes with moving objects *Proceedings of the conference on Computer Vision and Pattern Recognition*, Santa Barbara, CA, USA.


Diaz J and Torres R 2006 Classification of underwater color images with applications in the study of deep coral reefs *Proceedings of the International Midwest Symposium on Circuits and Systems*.


Dickmanns E 2007 *Dynamic Vision for Perception and Control of Motion* SPRINGER.


Geiger A, Lenz P and Urtasun R 2012 Are we ready for autonomous driving? the KITTI vision benchmark suite
Proceedings of the conference on Computer Vision and Pattern Recognition, Providence, RI, USA.

Geng H, Chien J, Nicolescu R and Klette R 2015 Egomotion estimation and reconstruction with Kalman filters and
gps integration Proceedings of the conference on Computer Analysis of Images and Patterns.

Gerónimo D and López A 2014 Vision-based Pedestrian Protection Systems for Intelligent Vehicles. SpringerBriefs
in Computer Science.

Gerónimo D, López A, Sappa A and Graf T 2010 Survey of pedestrian detection for advanced driver assistance


Girshick R, Donahue J, Darrell T and Malik J 2016 Region-based convolutional networks for accurate object
detection and segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 72, 4861.

M and Kramer P 2007a Documenting hurricane impacts on coral reefs using two-dimensional video-mosaic

Gleason A, Reid R and Voss K 2007b Automated classification of underwater multispectral imagery for coral reef
monitoring Proceedings of the IEEE/MTS OCEANS Conference.

Glover A, Maddern W, Milford M and Wyeth G 2011 FAB-MAP + RatSLAM: Appearance-based slam for multiple
times of day Proceedings of the IEEE International Conference on Robotics and Automation.

Goedeme T, Tuytelaars T and Geel LV 2006 Visual topological map building in self-similar environments
International Conference on Informatics in Control, Automation and Robotics.

Sonomologie 14, 187–199.

Técnico, Universidade Técnica de Lisboa Lisbon, Portugal.

Gracias N and Santos-Victor J 2000 Underwater video mosaics as visual navigation maps. Computer Vision and
Image Understanding 79, 66–91.

Gracias N, Costeira J and Santos-Victor J 2004 Linear global mosaics for underwater surveying Proceedings of the
IFAC/EURON Symposium on Autonomous Vehicles, Lisbon, Portugal.

Gracias N, Negahdaripour S, Neumann L, Prados R and Garcia R 2008 A motion compensated filtering approach
to remove sunlight flicker in shallow water images Proceedings of the IEEE/MTS OCEANS Conference, Quebec
City, Canada.

Gracias N, Zwaan S, Bernardino A and Santos-Victor J 2003 Mosaic based navigation for autonomous underwater

Grimm G, Zelinsky A, Nilsson L and Rilbe M 2004 3D vision sensing for improved pedestrian safety Proceedings
of the IEEE Intelligent Vehicles Symposium, Parma, Italy.

Grzonka S, Grisetti G and Burgard W 2012 A fully autonomous indoor quadrotor. IEEE Transactions on Robotics
28, 90–100.


Haesler R and Klette R 2010 Benchmarking stereo data (not the matching algorithms Proceedings of the German
Conference on Pattern Recognition.

Haesler R and Klette R 2012 Analysis of KITTI data for stereo analysis with stereo confidence measures
Proceedings of the European Conference on Computer Vision, Workshops and Demonstrations.

Haralick R, Shamumagam K and Dinstein I 1973 Textural features for image classification. IEEE Transactions on

Haribaran B, Arbelaez P, Girshick R and Malik J 2015 Hypercolumns for object segmentation and fine-grained


Herisse B, Russotto FX, Hamel T and Mahony R 2008 Hovering flight and vertical landing control of a vtol
unmanned aerial vehicle using optical flow Proceedings of the IEEE/RSJ International Conference on Intelligent
Robots and Systems.

Hermann S and Klette R 2009 The naked truth about cost functions for stereo matching. Technical Report MItech-
TR-33, Multimedia Imaging, Auckland.


Kazhdan M, Bolitho M and Hoppe H 2006 Poisson surface reconstruction Eurographics/ACM SIGGRAPH symposium on Geometry processing, Cagliari, Sardinia, Italy.


Khan W, Suaste V, Caulidellio D and Klette R 2013 Belief propagation stereo matching compared to iSGM on binocular or trincular video data Proceedings of the IEEE Intelligent Vehicles Symposium.


Kratz L and Nishino K 2009 Factorizing scene albedo and depth from a single foggy image ICCV09, pp. 1071 – 1078.


Negahdaripour S 2012a On 3D scene interpretation from FS sonar imagery *Proceedings of the IEEE/MTS OCEANS Conference*.


Negahdaripour S and Madjidi H 2003a Robust optical flow estimation using underwater color images *Proceedings of the IEEE/MTS OCEANS Conference*.


Negahdaripour S, Aylkin MD and Sinnarajah S 2011 Dynamic scene analysis and mosaicing of benthic habitats by fs sonar imaging - issues and complexities *Proceedings of the IEEE/MTS OCEANS Conference*.


Nister D and Stewenius H 2006 Scalable recognition with a vocabulary tree *Proceedings of the conference on Computer Vision and Pattern Recognition*.

Noh H, Hong S and Han B 2015 Learning deconvolution network for semantic segmentation *Proceedings of the International Conference on Computer Vision*.

NVIDIA at CES 2016 Nvidia drive px2 autonomous car processing engine demo - ces 2016.


Ohtake Y, Belvaev A and Seidel HP 2004 3D scattered data approximation with adaptive compactly supported radial basis functions *Proceedings of the Shape Modeling International*.


Pfeiffer D, Gehrig S and Schneider N 2013 Exploring the power of stereo confidences *Proceedings of the conference on Computer Vision and Pattern Recognition*.


Pizarro O 2004 Large scale structure from motion for autonomous underwater vehicle surveys PhD thesis Massachusetts Institute of Technology.


Pizarro O, Eustice R and Singh H 2004 Large area 3D reconstructions from underwater surveys *Proceedings of the IEEE/MTS OCEANS Conference*.


Pomerleau D 1995 RALPH: rapidly adapting lateral position handler *Proceedings of the IEEE Intelligent Vehicles Symposium*, Detroit, USA.

Porter T and Duff T 1984 Compositing digital images *Proceedings of the SIGGRAPH*.


Rezaei M and Klette R 2011 3D cascade of classifiers for open and closed eye detection in driver distraction monitoring *Proceedings of the conference on Computer Analysis of Images and Patterns*.


Richmond K and Rock SM 2006 An operational real-time large-scale visual mosaicking and navigation system *Proceedings of the IEEE/MTS OCEANS Conference*.

Roman C and Singh H 2005 Improved vehicle based multibeam bathymetry using sub-maps and SLAM *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Edmonton, Canada.


Roman C, Inglis G and Rutter J 2010 Application of structured light imaging for high resolution mapping of underwater archaeological sites *Proceedings of the IEEE/MTS OCEANS Conference*.


Rublee E, Rabaud V, Konolige K and Bradski G 2011 ORB: An efficient alternative to SIFT or SURF. *Proceedings of the International Conference on Computer Vision*.


Shum HY and Szeliski R 1998 Construction and refinement of panoramic mosaics with global and local alignment Proceedings of the International Conference on Computer Vision, Washington, DC, USA.


Szeliski R and Shum HY 1997 Creating full view panoramic image mosaics and environment maps *Proceedings of the SIGGRAPH*.


Tan RT 2009 Visibility in bad weather from a single image *Proceedings of the conference on Computer Vision and Pattern Recognition*.


Tao S, Feng H, Xu Z and Li Q 2012 Image degradation and recovery based on multiple scattering in remote sensing and bad weather condition. *Optics Express* 20, 16584 – 16595.

Tarel JP and Hautiere N 2009 Fast visibility restoration from a single color or gray level image *Proceedings of the International Conference on Computer Vision*.

Tena I, Reed S, Petilito Y, Bell J and Lane DM 2003 Concurrent mapping and localisation using side-scan sonar for autonomous navigation *Proceedings of the International Symposium on Unmanned Untethered Submersible Technology*.


No Hands Across America Webpage


vCharge Project


VisLab PROUD-Car Test

VisLab PROUD-Car Test 2013 [vislab.it/proud/](http://vislab.it/proud/).


Tri 2015. Retrieved February 26, 2015, from [www.tritech.co.uk/product/gemini-720i-300m-multibeam-imaging-sonar](http://www.tritech.co.uk/product/gemini-720i-300m-multibeam-imaging-sonar).


Viola P and Jones M 2001a Rapid object detection using a boosted cascade of simple features *Proceedings of the conference on Computer Vision and Pattern Recognition*, Kauai, HI, USA.


Vedron M and Erwin C 1973 The heart rate variability correlates of spontaneous drowsiness onset *Proceedings of the International Automotive Engineering Congress*, Detroit, USA.

Wahlgren C and Duckett T 2005 Topological mapping for mobile robots using omnidirectional vision *Swedish Workshop on Autonomous Robotics*.


Wang Q, Gaidon A and Vig E 2016 Virtual worlds as proxy for multi-object tracking analysis *Proceedings of the conference on Computer Vision and Pattern Recognition*.

Wei J, Snider J, Kim J, Dolan J, Rajkumar R and Litkouhi B 2013 Towards a viable autonomous driving research platform *Proceedings of the IEEE Intelligent Vehicles Symposium*.


Xie Y, Liu LF, Li CH, and Qu YY 2009 Unifying visual saliency with hog feature learning for traffic sign detection *Proceedings of the IEEE Intelligent Vehicles Symposium*, Xi’an, China.

Xiong Y and Pulli K 2009 Color correction based image blending for creating high resolution panoramic images on mobile devices *Proceedings of the SIGGRAPH Asia*, Yokohama, Japan.


Zhang H and Negahdaripour S 2003 On reconstruction of 3D volumetric models of reefs and benthic structures from image sequences of a stereo rig Proceedings of the IEEE/MTS OCEANS Conference, San Diego, USA.


Ziegler et al. J 2014b Video based localization for bertha Proceedings of the IEEE Intelligent Vehicles Symposium, Dearborn, MI.


