

## Computer Vision in Support of Autonomous Capabilities

R J Evans, C G Harris  
Roke Manor Research Limited  
Roke Manor, Romsey, Hampshire SO51 0ZN, UK

### Abstract

*This 3D Vision project aims to demonstrate that a small-scale platform can journey through man-made and natural environments, surviving, navigating, mapping and planning, largely on the basis of passive computer vision. To this end, in our first year of work, we have set up a demonstration platform placing 3D computer vision algorithms in a robotic infrastructure so that we can challenge the computer vision algorithms by using them to support autonomous vehicle capabilities, such as route following and obstacle avoidance. The main elements of this system are now in place. This paper describes the scope of work to date, discusses some related computer vision issues and outlines future plans.*

*Keywords: 3D computer vision, stereovision, autonomous vehicles, robotics.*

### Introduction

This paper provides a view of the EMRS DTC's 3D Computer Vision project, which is being carried out by Roke Manor Research (Project reference EMRS/DTC/1/12). This work was originally proposed as a three-year project with the aim of demonstrating that a small scale platform can journey through man-made and natural environments, surviving, navigating, mapping and planning, largely on the basis of passive computer vision. We have just completed the first year's work, which has seen the setting up of an Initial Development Platform and the implementation of early versions of the main system modules, including navigation, obstacle avoidance and route-following as well as the underlying computer vision system. Work proposed for early in the second year, should see the completion of the link between software generated steering control commands and electrical control signals, and the creation of our first autonomous system.

In the limited space available here we will try to provide a summary of the scope of

the work so far, but a goal is also to put it into context. Consequently this paper continues with sections on a review of relevant work elsewhere (a major element of which is previous work by Roke Manor on structure-from-motion computer vision algorithms), project aims and motivation, a summary of the work done in this project to date, and discussion of technical issues and future plans.

### Previous Work

Computer vision in robotic applications was the subject of much academic research, in the 1980s and early 1990s. Some successful systems were developed but generally solutions were insufficiently robust outside of a few well-controlled environments. DeSouza and Kak (1) provide a survey of computer vision for mobile robot navigation. They identify three classes of system for indoor environments, outdoor structured environments and outdoor unstructured environments. In indoor environments they conclude that substantial progress has been made and this reflects the fact that the indoor environment is more simple and

controllable than the outdoor environment and a greater degree of prior knowledge can be exploited. Feature extraction is potentially easier, and maps and models more readily available and applicable. Success in outdoor structured environments generally refers to road-following systems. Here the environment is also well known and controllable to some extent, but illumination, shadows, and weather conspire against vision systems, though we would assert that the road environment is predisposed towards the passage of vehicles - manned or autonomous - in a way which the indoor environments are not. In 1979 Tsugawa et al (2) reported a vision-based vehicle which could travel at 30km/hr in highly constrained circumstances. About 10 years later Dickmanns & Zapp (3) developed algorithms for "Autobahn" environments.

The unstructured outdoor environment has been more challenging. The prior knowledge that can be applied is minimal and little is available besides the perceived 3D structure. Nevertheless, aided by active sensors, and supplementary positioning sensors, some progress has been made. A notable example is the Mars Pathfinder rover, which uses rudimentary hazard detection and avoidance algorithms to explore the planet surface at a maximum speed of 10m/day. See Matthies et al (4).

DROID is a 3D vision system initially developed by Roke Manor Research in the Alvey initiatives of the late 1980s. DROID is the algorithmic starting point for the current work. Harris (5) provides a detailed description of the original algorithms. DROID is a "structure from motion" algorithm using "point features". Briefly, point features are extracted from each processed frame of a video sequence and tracked. From the apparent motion of these features DROID calculates both the scene structure, represented by the estimated 3D position of the tracked features, and the

path taken by the sensor platform, the ego motion.

Two aspects of DROID contribute to its success: it uses a particularly robust feature extraction algorithm and the calculation of ego-motion is decoupled from the calculation of the 3D feature positions. This decoupling makes the computation tractable in software in real-time. DROID requires a geometrically calibrated camera system and, if operating with a single camera, an estimate of the initial frame-to-frame ego-motion; thereafter DROID is self-supporting. Feature tracking is performed directly in 3D with Kalman filters refining feature locations. Stereo operation obviates the requirement for the initial motion estimate and the fixed stereo base line stabilises the speed-scale ambiguity and long terms drift which occur in mono operation.

### **Aims and Motivation**

The aim of the current DTC project on 3D vision is to "demonstrate that small-scale sensor platforms ... are capable of operating in directed autonomous mode within complex environments, both containing natural (e.g. vegetative) entities and human artefact (e.g. buildings). We aim to show that such platforms can journey through such environments, surviving, navigating, mapping and planning, largely on the basis of passive computer vision".

Aims can be seen on three levels of military interest. The first is straightforward - a vehicle, equipped with a vision system able to support autonomous capabilities, would be useful in its own right. Example autonomous applications include ferrying equipment and personnel from A to B, communications relay, and vehicle control while searching for mines. On another level 3D vision can provide capabilities like mapping and identification of locations for

other sensors. At the sensor technology level, embedding 3D capabilities in an imaging sensor provides a multipurpose device, a subsystem whose workings should be transparent to the user. In the military context we require robust operation and a minimum of operational constraints and support, and covert operation.

While our prime interest in this project is 3D computer vision algorithms, we have put them in a practical robotics context. Firstly this is because we believe, as with other image processing tasks, that processing real data is the only reliable way of testing algorithms. Secondly we wish to demonstrate that computer vision techniques *can* support autonomous capabilities such as obstacle detection, obstacle avoidance, route following and rendezvous etc. Thus our demonstrator vehicle will act both as a platform for exercising computer vision algorithms and demonstrating that the results are adequate for the proposed purposes.

### EMRS DTC Programme Activities

In this section we will summarise the activities to date, in terms of the main project work.

**Concepts and Architectures:** We decided to adapt a convenient motorised platform for use as an Initial Development Platform (IDP), and move to a more capable military-style platform in a later phase. We also identified a control architecture, based on work by Stentz & Hebert (6) who use an arbiter module to balance the requirements of sensor-based activity (e.g. obstacle avoidance) against model-based activity (e.g. following path plans).

**Demonstrator Platform:** The IDP is shown in figure 1. This has an equipment cabinet attached to a commercial motorised base. The main onboard elements of this system are a stereo camera rig, GPS

receiver, vehicle-mounted laptop PC, secure wireless LAN (IEEE802.11g), and a watchdog safety system. A major feature of this system is the ability to record time-stamped data for repeatable off-line analysis, as well as providing on-line work.

**3D Vision Software (& obstacle detection):** We installed the existing DROID algorithms on the IDP and in exercising it we quickly identified a number of pressing algorithm issues. These concerned the calculation of platform motion from visual data to accommodate the bumpy movement associated with military and off-road vehicles. We also addressed various issues relating to use of computer vision outdoors, namely the spatial distribution of extracted image features, stereo image balance and feature matching/tracking issues. Some of these topics will need to be revisited however. We also developed new algorithms to interpret the visually perceived terrain in terms of accessible or “drivable” paths.



Figure 1: The Initial Development Platform

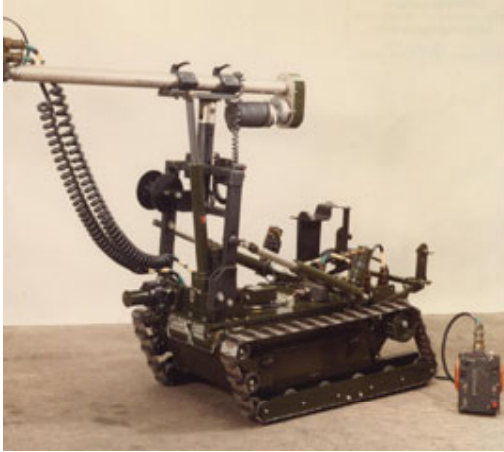


Figure 2: A robust military format platform (courtesy Remotec UK)

**Navigation & Supplementary Sensors:**

We compared visual and GPS data, and developed algorithms to integrate data from the two sources. We found that DROID navigational data is highly accurate - more so than the GPS data - though it only provides relative positioning. Figure 3 shows a plot of the path taken by the IDP on a trial mission. Blue indicates the GPS position and magenta the visual data plot, displaced for clarity. The real-time integrated track is shown in red.

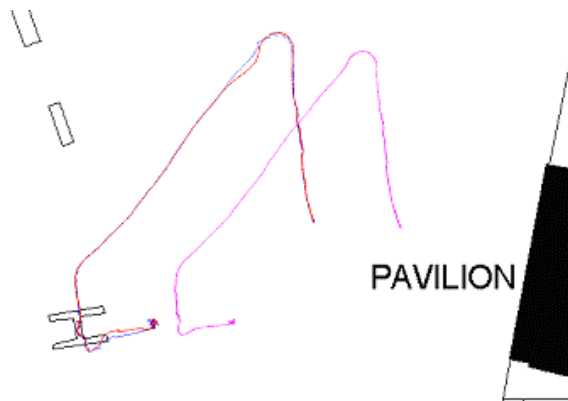


Figure 3: Example navigation data - GPS, Visual data and integrated traces.

**Path Planning & Control:** Here we have implemented a basic system enabling the vehicle to follow a global path plan, which has been manually defined and input as a series of waypoints. Following the approach of Stenz & Hebert, a Global

Navigator module is responsible for monitoring progress and selecting a single target waypoint. In parallel a Local Navigator assesses the viability of the available steering options, considering the shape of the terrain immediately in front of the vehicle, as perceived by the vision system. A Steering Arbiter selects from the viable options the best path to meet the target waypoint. At the time of writing we have not fully closed the control loop and the vehicle remains manually controlled, but we expect to close the loop shortly. Exercising these algorithms with pre-recorded data and hypothetical global path plans, we have found that very simple algorithms can be effective in dealing with more isolated obstructions in open terrain.

**Active 3D Sensors:** The aim of this initial study was to identify the relevant issues and scope the next steps in the integration of active sensors on the demonstrator. Our main findings were that active sensors might be required for a variety of reasons, but no requirement has been identified as the single overwhelming need at this stage, and there is no clear single solution in terms of sensor hardware. There are also issues in the management of the active sensors and the data integration algorithms themselves. A non-trivial issue is the calibration and registration of the two sensing systems; if the sensors are poorly registered, data integration may be counter-productive.

**3D Edge Features:** The need for this work on edge features arose from practical experience with point-based computer vision software and the observation that the quality of 3D surface representation provided by points could be much improved in some situations by additional information from edges. We found that a very promising algorithm could be devised by taking a simple approach - that of processing sampled edges and processing the sample edge elements independently. The simplicity is gained at the cost of

losing information about the connectivity of edge elements, but the trade-off appears to be beneficial with a fair degree of accuracy and processing speed being achieved in our experimental implementation.

### **Computer Vision Issues & Discussion**

Having set up the IDP we are in a position to do a greater level of trialling and assessment than before. Some topics have already emerged, however, from our more subjective assessments.

Given satisfactory feature matching and tracking, the calculation of platform motion and 3D feature positions seems a relatively understood process. The underlying matching and tracking is still problematic, as the appearance of features may vary between image frames. A further issue is that not all features, of course, are stationary in the scene and some are subjective; these do not fit our 3D model and consequently are a source of error and a processing overhead, as they clutter the tracking process. Image noise *per se* does not seem to be a problem.

At a higher level there are questions of integration, interpretation and interpolation of the perceived 3D structure. Our current tool for interpretation and interpolation is based on Delaunay triangulation. This is an unintelligent device and one rogue point can damage an area of the interpreted surface. In the longer term we envisage more intelligent, scene-understanding techniques might be applied to aid the interpretation of structure.

The structure integration issue is related to the question of mapping. The representation of terrain in terms of 3D point features may be a suitable representation for detecting obstacles in the current field of view, but we will require a different representation for long-term use as variations in lighting (weather etc.) and

viewpoint change the visibility of features. In the shorter term, given the susceptibility of triangulation to rogue data, there is the question of how to combine two surface representations that do not share a wholly corresponding set of vertices.

Returning to the issue of feature matching, there are several known reasons why features may look different in different image frames. Inevitably the changes in perspective, on which the structure-from-motion method is based, will change the appearance of individual features. We are obliged to use wide-angle camera lenses for a large field of view, but these have significant distortions and these can affect the appearance of features. Lens distortions are taken into account in our 3D calculations, but not yet in our feature matching & tracking. We are also obliged to use cameras with AGC, because of the variations in lighting encountered outdoors. As a result of AGC there can be substantial brightness differences between left and right stereo pairs, and, as standard AGC systems are designed to respond quickly, there can be smaller but still detrimental changes between successive frames.

While we will be working to improve robustness we should emphasise we have demonstrated basic capabilities in extracting 3D structure and detecting and avoiding obstacles. Figure 4 shows a sample DROID display with an overlay of acceptable steering options.

### **Future Plans**

Our proposal for the second year's work places a priority on trials activities and we believe we will reap the reward of setting up a practical demonstrator.

We plan work in three main areas of computer vision, implementing, on-line, the 3D edge processing algorithms, which we have developed off-line, and addressing the

issue of moving and subjective features. In the course of trials and adjustments we also plan to improve feature matching and tracking.

Continuing our emphasis on practical issues and testing 3D vision results, we plan work on mapping and the use of vision-based mapped data. We also plan an initial study with a view to transferring equipment from the Initial Development Platform to a more robust, military format vehicle, such as shown in figure 2, at a later stage.

### Concluding Remarks

The first year of this project has been largely concerned with groundwork and infrastructure. We have successfully brought together a number of system elements into an integrated system. Overall this bodes well for the later stages of the project when we expect more dramatic results will be achieved.

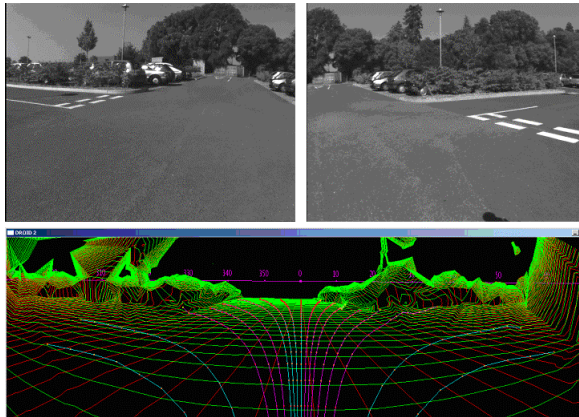


Figure 4: Example DROID display - camera views and reconstructed panorama with assessed steering options

### Acknowledgements

We wish to acknowledge the contributions of numerous colleagues in this project: Carl Stennett, Chris Spence, Esin Turkbeyler, Chris Day, Ryan Orgill, Richard Winterbottom and Chris Gillham.

### References

1. DeSouza and Kak, 2002, "Vision for Mobile Robot Navigation: A Survey", IEE Trans Pattern Analysis and Machine Intelligence 24, pp237-267
2. Tsugawa et al, 1979, "An Automobile with Artificial Intelligence", Proc. Sixth Int'l Joint Conf. Artificial Intelligence, pp893-895
3. Dickmans and Zapp, 1987, "Autonomous High Speed Road Vehicle Guidance by Computer Vision", Proc 10<sup>th</sup> World Congress on Automatic Control, vol. 4.
4. Matthies *et al*, 1995, "Mars Microrover Navigation: Performance Evaluation and Enhancements", Proc IEEE Int'l Conf Intelligent Robots and Systems, vol. 1, pp433-440.
5. Harris, 1993, "Geometry from Visual Motion" in "Active Vision", Eds Blake & Yuille, MIT Press,
6. Stentz and Hebert, 1995, "A Complete Navigation System for Goal Acquisition in Unknown Environments", Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems.

© Roke Manor Research Limited 2004