

Published in IET Intelligent Transport Systems
 Received on 1st December 2006
 Revised on 3rd January 2008
 doi: 10.1049/iet-its:20060083



Mobile mapping for the automated analysis of road signage and delineation

S. McLoughlin¹ C. Deegan¹ C. Mulvihill¹ C. Fitzgerald²
 C. Markham³

¹School of Informatics and Engineering, Institute of Technology Blanchardstown, Dublin 15, Ireland

²National Roads Authority of Ireland, Kildress House, Pembroke Row, Dublin 2, Ireland

³Department of Computer Science, National University of Ireland, Maynooth, Co. Kildare, Ireland

E-mail: simon.mcloughlin@itb.ie

Abstract: A portable mobile stereo vision system designed for the assessment of road signage and delineation (lines and road studs or 'cat eyes') in low light conditions is presented. This novel system allows both geometric and photometric measurements to be made on objects in a scene. Using the system, it has been shown that retro-reflectors, and in particular road signs, can be identified by nature of their reflective properties. In addition, a novel imaging application has been investigated that facilitates the detection of defective road studs. Any objects examined can also be positioned on a national grid through the fusion of stereo vision with global positioning system technology. Automated feature extraction and analysis routines make the system fully autonomous.

1 Introduction

Accurate terrestrial mobile mapping systems have been in operation for over two decades. Through advances in global positioning system (GPS) technology and photogrammetry, features can be extracted from an image automatically and positioned in a global reference frame to an accuracy of <1 m. The early 1990s, in particular, seen two major developments in mobile mapping through the production GPSVan [1, 2] by the Ohio state University in the USA and the VISAT van [3, 4] developed at the University of Calgary in Canada. Both systems have been modernised through the years and are now commercial exploits. They employ sophisticated GPS sensors, inertial measurement units (IMUs), dead reckoning systems and stereo vision modules. The absolute position and velocity information provided from the GPS data are used to provide updates to the inertial navigation system (INS) to counteract IMU drift errors. The data from these sensors are integrated using a Kalman filter to obtain more accurate position and orientation estimates. The stereo vision component has at least two CCD cameras separated by a

wide baseline for ranging out to 50 m. The epipolar geometry of the stereo rig is also computed to aid correspondence of conjugate pairs. The overall feature positioning accuracy reported by such systems is impressive. GPSVan reports positioning accuracy of 10 cm for features in the stereo field of view, whereas VISAT van reports a positioning accuracy of 30 cm inside of a 50 m range corridor. The values for VISAT van are based on the sub-pixel location of specifically designed targets [5] in the images so performance may not be quite as good when surveying untailored objects where object centres are difficult to extract because of perspective distortions. The stereo systems are calibrated by a bundle adjustment technique [6] using objects whose precise 3D locations and image coordinates are known.

Photobus [7] is a European mobile mapping system that has been used to map road centrelines. The system is distinguished from its predecessors by its ability to geo-reference through a single vertically oriented complementary metal-oxide sensor (CMOS) sensor. Also, the real-time performance is achieved by

transmitting GPS error correction estimates using NTRIP (NTRIP is a wireless Internet protocol for streaming GPS correction data) to the rover GPS receiver so corrections can be done on the fly. More recently at the Ohio State University, a highly autonomous terrestrial mobile mapping system has been developed for the same problem (mapping road centrelines) [8]. The system also operates during periods of moderate GPS signal reception. The system works in near real time and boasts a feature positioning accuracy of less than a decimetre. All feature extraction and image analysis routines have been automated. Less complex and expensive systems have also been developed for surveying purposes. Ellum and El-Sheimy [9] have developed a low-cost minimum complexity backpack mobile mapping system. Using a GPS receiver, a digital compass and a single consumer digital camera, absolute accuracies under 25 cm are obtained at 20 m geometries using a minimum of three images of the features. Applications of the system are urban geographic information system (GIS) data collection, highway inventory, architectural reconstruction and small-scale topographic mapping. The challenges that remain in this area today can be split into two domains. The development of algorithms and techniques to automate feature extraction from the vast amounts of data acquired by such systems and the construction of more accurate, more robust, more portable and less expensive mapping systems for object positioning.

The acquisition system developed for data collection in this work is entirely portable and is capable of obtaining a positioning precision (repeatability measure) of 1.34 m [root mean square (RMS) error] using standard (relatively inexpensive) hardware. The nature of the system construction and the components used also mean the system is versatile and applicable to a number of areas in the ITS domain. Section 2 discusses the acquisition system in more detail.

Currently, the system has been used to identify the position of defective road studs and as an inventory system for road signage. Novel algorithms and techniques to automate the detection of these image features have been developed for imagery acquired in low-light conditions (i.e. dusk and night hours). Defective road studs are detected using classic image analysis routines and geometrical constraints, whereas signage is detected by exploiting the surface properties of retro-reflective materials. Currently, all data processing is done offline. Sections 3.2 and 3.3 discuss these applications.

2 Data acquisition system hardware

The data acquisition system consists of reliable hardware components integrated using multi-threaded application

programming interfaces. Two high dynamic range (10-bit) CMOS cameras comprise the stereo vision system and are used to make both photometric and geometric measurements on objects in the scene. The optical axes of the cameras are close to parallel to a level ground plane, which is not optimal for stud detection. However, this configuration allows for different types of features to be analysed in the same data set. At present, road signs and studs are analysed but this set may also be expanded to consider other objects such as painted line markings and lampposts. A second component of the system is a standard dual frequency GPS receiver (Garmin III) that produces global information about the position of the vehicle. The fusion of post-processed GPS information (Section 3.1.3) along with the local position information provided by the stereo system allows national grid coordinates to be obtained for the features under investigation. GPS data and stereo image data are synchronised using a phase-locked loop, whose oscillation frequency is controlled by the GPS clock. GPS data are captured at a rate of 1 Hz, whereas stereo image pairs are captured at a rate of 2 or 3 frames/s (fps). All hardware components are interfaced through a central notebook PC hosting an Intel Pentium 4 2.8 GHz processor, 512 MB of RAM and an 80 GB hard disk. Most data analysis is currently done offline under relaxed time constraints, so the analysis hardware is machine independent. The complete acquisition system can be controlled from inside the vehicle across a wireless local area network. A sine wave power inverter provides alternating current from a 12 V lead acid battery to the acquisition computer and a FireWire hub, which in turn powers the cameras. The main acquisition system components are depicted in Fig. 1b. The hardware is secured inside a roof box that has two openings enabling the cameras to view the road ahead. In this respect, the system is completely portable and can be mounted on any vehicle possessing standard roof bars, see Fig. 1a. A more detailed overview of the entire acquisition system is available in [10] where the system is used to classify different road signage materials, that is, locally illuminated signs, diffuse reflecting signs and retro-reflecting signs. Recently, the system has also been fitted with a powerful infrared (IR) illuminator that gives higher scene irradiance than the vehicle headlights. The IR wavelength is also invisible so should not cause any problems with other road users. The wide spectral response of the cameras makes this configuration plausible.

3 Software theoretical foundations

In this section, a number of software-based data processing techniques and algorithms will be described for image feature extraction and the subsequent positioning of the extracted features on a national grid. Fig. 2 shows an intermediate level data flow diagram that serves to

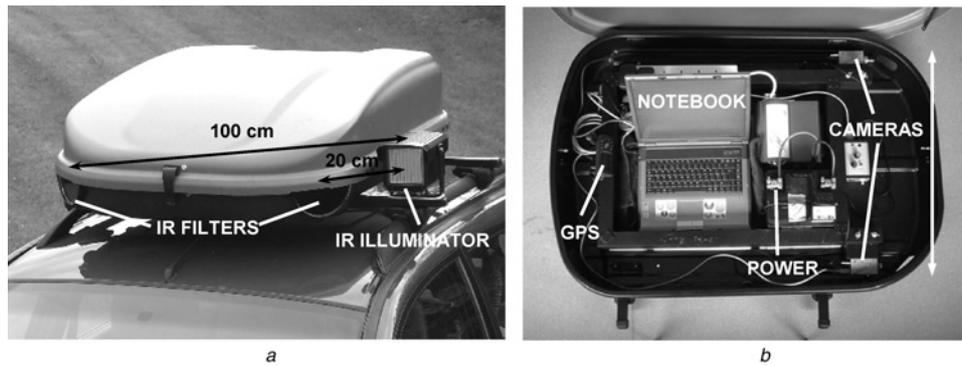


Figure 1 Data acquisition system and acquisition system components

a Data acquisition system mounted on a vehicle

b Acquisition system components

inform the reader on the various processing routines that transforms the raw sensor data into national grid coordinates of ‘interesting features’.

As is evident from Fig. 2, the set of processes or transforms are broken down into two main streams. The leftmost stream deals with the image data and specifically the automated feature extraction and stereo coordinate computation. The rightmost stream deals with the global positioning data and how it can be transformed into a format that is more accurate and compatible with the stereo measurements such that they can be combined.

3.1 Feature position estimation

3.1.1 Local coordinate estimation: The local 3D coordinates for features are calculated using stereo vision. The feature extraction routines (Sections 3.2 and 3.3) return the image pixel coordinates for the feature points in a stereo pair. By knowing these coordinates and the cameras intrinsic/extrinsic parameters through calibration, the projectors (ray from object to image plane) can be intersected to estimate the X , Y and Z coordinates of the feature in a camera-centred Cartesian coordinate system [11]. The camera calibration is based on the work of Zhang [12], which allows the calibration to be carried out by observing a planar pattern undergoing rotations and translation in an image sequence. The motion of the plane need not be known. All that is required is the precise measurements of the features on the plane and the corresponding image coordinates. The planar pattern usually consists of a checker board where the features are the corners of the squares. This calibration technique simplifies the calibration of stereo systems with large baselines.

3.1.2 Global coordinate estimation: The actual position computation of the image features in planar global coordinates (national grid) requires three parameter sets. The local X and Z coordinates of the

feature (estimated from stereo vision), the translation offset between the local and global coordinate systems (provided by GPS), and the rotation offset (vehicle heading) of the local and global coordinate systems. The rotation offset between the coordinate systems is not given directly from the GPS data but can be estimated by differentiating the position with respect to time. Because of the noise in GPS measurements, orientation calculations can introduce large positional errors. The receiver makes global coordinate measurements to an accuracy of 15 m RMS (or 1–5 m using differential global positioning system (DGPS)) at a rate of 1 Hz. Within Ireland, there are a number of GPS post-processing options available to improve accuracy. Ordnance Survey Ireland offers free hourly online downloads of receiver-independent exchange format GPS data for 16 active reference stations throughout Ireland. All 16 stations on the network conform to the EUREF Zero-Order Class B specification (better than 2 cm in X , Y and Z) and so can be used for accurate GPS post-processing. At the time of writing both Leica GeoSystems and Trimble Europe B.V. offered commercial real-time kinematic (RTK) services in Ireland. The Trimble VRS Now correction service is available and is capable of delivering centimetre accuracy to compliant GPS receivers. Wide area augmentation system subscriptions are also available through OmniSTAR who offer a number of services for different accuracy requirements, for example, OmniSTAR virtual base station for sub-metre accuracy and OmniSTAR high performance for decimetre accuracy. The short-time precision in GPS measurements is actually of more interest to the applications discussed here since measured displacements are more critical than absolute accuracies. Despite the sensor limitations, post-processing of the data can significantly improve the estimated trajectories. Section 3.1.3 has a more detailed discussion on this topic.

Once translation and rotation parameters have been established from the post-processed GPS data, the

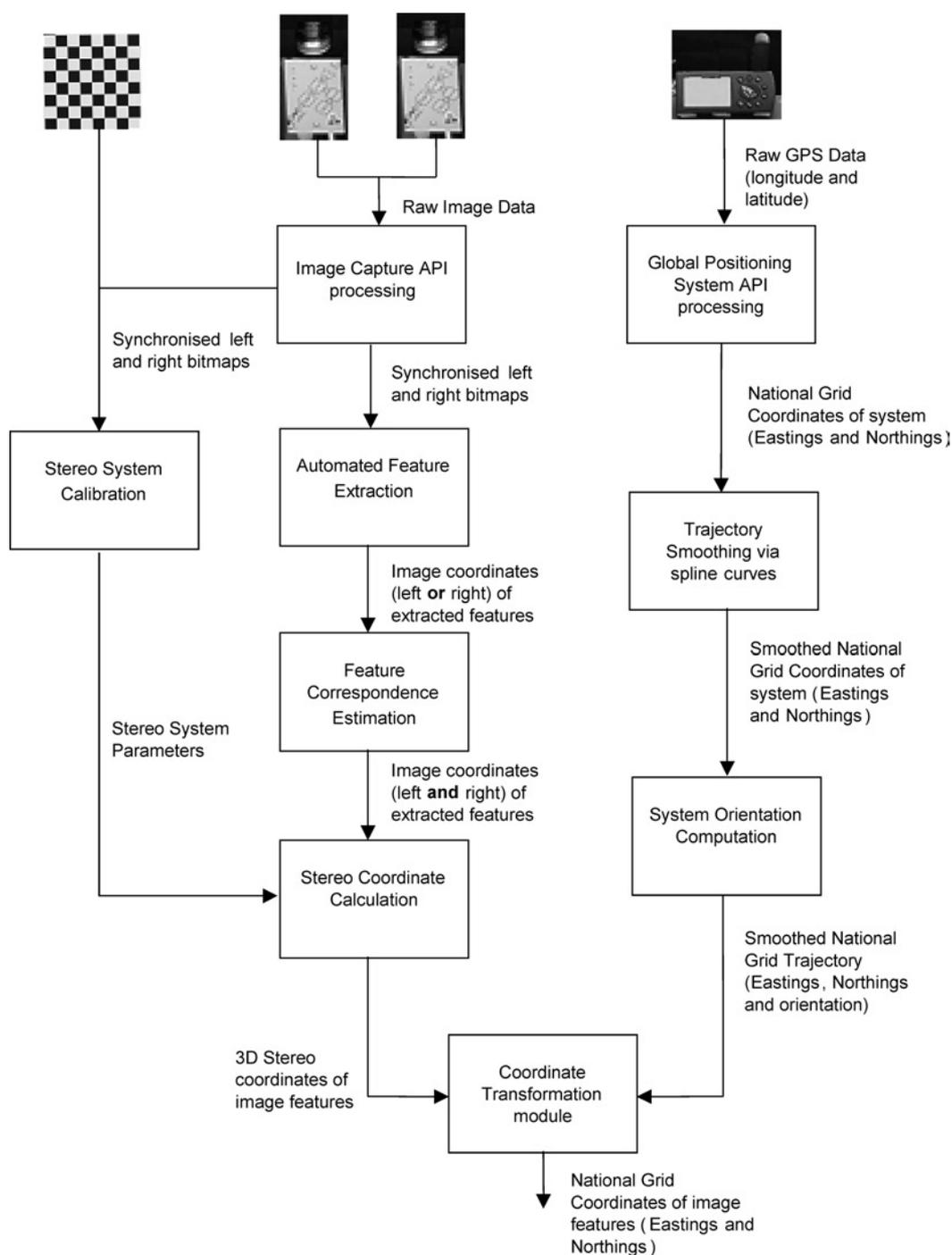


Figure 2 Data flow diagram showing software processing modules

global coordinates (national grid) of the feature points can then be estimated by applying a simple Euclidean transformation (given in homogenous coordinates as)

$$\mathbf{P}_{\text{global}} = (\mathbf{R}(\theta)\mathbf{T})\mathbf{P}_{\text{local}} \quad (1)$$

where $\mathbf{P}_{\text{global}}$ and $\mathbf{P}_{\text{local}}$ are the feature points in the global and local coordinates systems, \mathbf{R} the standard 2D rotation matrix and \mathbf{T} and θ are the translation and rotation offsets between the local and global coordinate systems.

3.1.3 GPS trajectory smoothing: The fact that the receiver is in a terrestrial vehicle that moves along a piecewise smooth road network gives an insight into how position measurements should progress, that is, the resultant GPS curve should be piecewise smooth because of the motion of the vehicle and the nature of the roads it travels on. Using this knowledge, the true trajectory of the vehicle can be better estimated by fitting a smoothing function to the noisy GPS data. Cubic smoothing splines are appealing for this application because they offer high flexibility while maintaining smoothness by enforcing

segments to have the same first and second derivatives at the joining points. The resultant spline curves are continuous which also allows for image position estimation between GPS measurements. The smoothing [13] operates by fitting a curve to the data points while minimising two criteria. The first is that the spline curve should come reasonably close to the data points and is dictated by the uncertainties in the coordinate measurements (GPS uncertainty). To balance this constraint and prevent a direct interpolation, the curvature of the spline is also minimised. A reliable fit occurs when the two terms are balanced properly using weighting factors. Fig. 3 shows how feature positions on a short road segment (~350 m long) are affected by noisy GPS data and how cubic spline smoothing improves estimates. The features shown are road studs on the centreline. The improvement can be seen by noting on the left-hand side (LHS) how the raw heading estimate affects the feature locations. In particular, there is a significant horizontal disparity for the same stud detected in successive frames. This disparity is typically about 3 m but may be as high as 6 m for the worst-affected areas. After post-processing, a disparity of ~1 m is more typical (although this may rise to 2.5 m for the worst-affected areas).

3.2 Feature extraction: road sign detection

The principal cue used in this work for the detection of road signs is different from the conventional techniques that use either colour or shape as the main cues such as in [14–18]. While colour and shape are the obvious and intuitive approaches to use since road sign manufacturers conform to colour and shape restrictions, they are not without their problems. From a digital imaging point of view, there are three

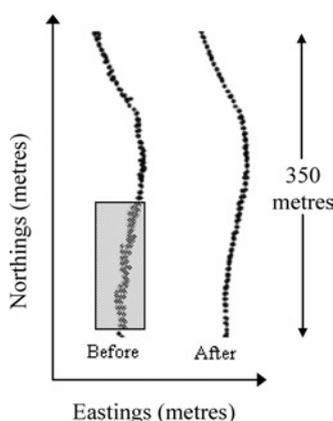


Figure 3 Example of how cubic spline smoothing can improve GPS estimates

The black dots represent detected features (road studs) on the road surface that characterise its shape
Shaded areas are where raw orientation estimates were poor

factors that will influence the pixel red/green/blue (RGB) values acquired from the camera as can be seen as

$$L(r, g, b) = \int_{\text{visible}} G(\lambda)R(\lambda)I(\lambda)d\lambda \quad (2)$$

where L is the image radiance measured in terms of the three primaries, $G(\lambda)$ the spectral response of the imaging device, $R(\lambda)$ the spectral reflectance properties of the surface and $I(\lambda)$ the spectral distribution of the incident light. If a traffic sign, for example, is illuminated with sunlight that is characterised by a spectral distribution with a lot of energy towards the blue wavelengths, then its sensed colour will be different than when illuminated by a vehicle's headlamps where much of the energy is in the red portion of the spectrum [19]. Techniques that rely on shape as the dominant discriminatory factor are also subject to problems caused by occlusion along with the fact that such techniques rarely detect all road sign types because of the significant variation in shape. Many authors who use shape and colour-based techniques report high detection and recognition rates but the number of false positives or false alarms seems to be a problem with many techniques. Gil-Jimenez *et al.* [18] report detection rates of 97% for their fast Fourier transform shape-based technique. False positives are a significant problem although with 175 reported out of 300 images for rectangular signs. Loy and Barnes [16] with the radial symmetry detector (another shape based technique) show similar trends. Detection rates of 95% for octagon-shaped signs, 95% for square-shaped signs and 100% for triangular-shaped signs are reported. The number of false positives again is quite high (19 false positives for 15 square targets) for signs that have a common shape. Miura *et al.* [17] report high detection rates of 100% for speed signs and 97% for warning signs without mentioning the false positive rate. The technique described in the following section is capable of detecting different types (colour and shape) of road sign under the assumption that they are retro-reflective. This also includes signs that are significantly occluded. The results of the technique are presented in Section 4. It is worth remembering that all types of retro-reflective road signs are taken into account when evaluating the performance regardless of the shape or colour.

3.2.1 Exploiting the cone of retro-reflection: By carrying out experimental analysis of the material properties composing signs, interesting insights are noted. Road signs are retro-reflectors meaning they reflect significant amounts of incident light back to the source. They are purposely engineered as imperfect retro-reflectors, meaning they actually reflect light back towards the source in a cone with a non-uniform luminous intensity distribution, see Fig. 4.

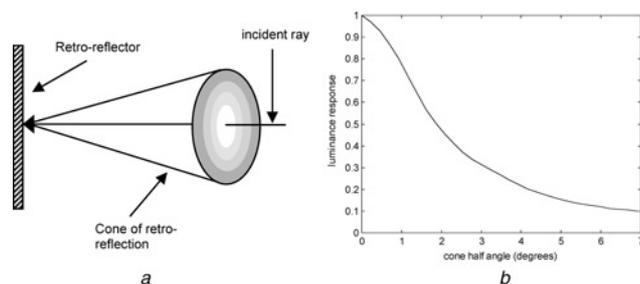


Figure 4 Road signs reflect light back towards the source in a cone with a non-uniform luminous intensity distribution

a Cone of retro-reflection as produced by a road sign
b Reflected light intensity response inside the cone of retro-reflection for a typical road sign

Therefore the luminance levels inside the cone of retro-reflection increase as the observer moves in towards the centre. It is this property that allows retro-reflectors to be identified in a stereo pair of correctly positioned cameras. One camera is placed very close to the cone centre (inner cone camera) where luminance levels are high and the other is placed towards the cone edge (outer cone camera) where luminance levels are lower. For actual data acquisition, one camera is located beside the IR source (20 cm from the source on the LHS of the roof box) and acts as the inner cone camera, where the signal observed from retro-reflectors is significantly greater than the one observed by the outer cone camera, which is located on the opposite side of the roof box (1 m from the source on the right-hand side (RHS) of the roof box) acquisition system. This arrangement can be seen in Fig. 1b. A comparative analysis of the light intensity information in the inner and outer cone cameras is used as the main factor to identify retro-reflective surfaces.

3.2.2 Candidate region extraction: In order to compare the intensities of surfaces in the image pair, candidate surfaces must first be extracted in the left or right image. A single region of interest (ROI) is extracted from each image based on the limits of projection of signs at specific distances and placements [20]. ROI extraction reduces the search space from 1280×1024 to $\sim 1000 \times 500$ pixels. The higher signal in the inner cone (left) image is not invariant to position changes or the quality of retro-reflective surfaces so a lower bound on the image intensity of retro-reflective surfaces is found empirically and used to threshold the image. The resultant binary image contains regions that are sufficiently bright to be a retro-reflective surface. A dynamic soft threshold would be advantageous in this setting but the potentially small area of road signs in the image make thresholding techniques such as grey-level histogram analysis impractical.

There will be a number of other non-road sign objects that survive the thresholding process. At night, this number is quite small and is typically $< 3-4$ objects.

Because of the increased illumination in a dusk setting, this number rises substantially but is typically < 10 objects. Other light sources in the scene, bright or shiny surfaces and the sky area in a dusk setting are all examples of regions that may survive the thresholding process. In fact, the sky area can introduce another problem during segmentation. When the sky is background to a road sign the two become merged in the same connected component so the sign cannot be analysed individually. To reduce the effect of this problem, the exclusive OR of any sky connected components with an edge image is found. Sky connected components are simply considered as those that touch the first row of the ROI. The edge image is formed by hysteresis thresholding [21] on the image gradient so if there is sufficient contrast between the sign and the sky area an edge will form and 'break' the sky connected component releasing the sign. For this technique to work effectively there must be a distinct border and light intensity difference between the sign and the sky background. Other non-sign regions may also be released from the sky connected component which may increase the computational complexity of the algorithm. However, the accuracy benefits outweigh the extra processing time required for this particular application. A real-time application may have different performance requirements. Fig. 5 illustrates the process.

The inclusive OR of the broken sky connected component image with the original binary image (without the sky connected component) is then found so a connected component analysis can be done to reduce the number of candidate regions.

Connected components are rejected from the set based on a number of criteria that are listed below:

1. Components should not be saturated (intensity comparisons cannot be made on the saturated components, e.g. very bright lights including oncoming headlights).
2. No components should be connected to the top row of the ROI (excludes some sky components).
3. Components should have an area within predefined limits. Exceptionally large (greater than 20% of ROI) and small ($< 0.095\%$ of ROI) components are rejected from the set. This criterion will exclude large sky connected components that are not connected to the top row of the ROI as well as small noisy components that form during thresholding.
4. Component bounding box aspect ratio should be within the predefined limits. No aspect ratios should be more than 6:1 (excludes exceptionally long and thin connected components that may sometimes result from specular reflection).

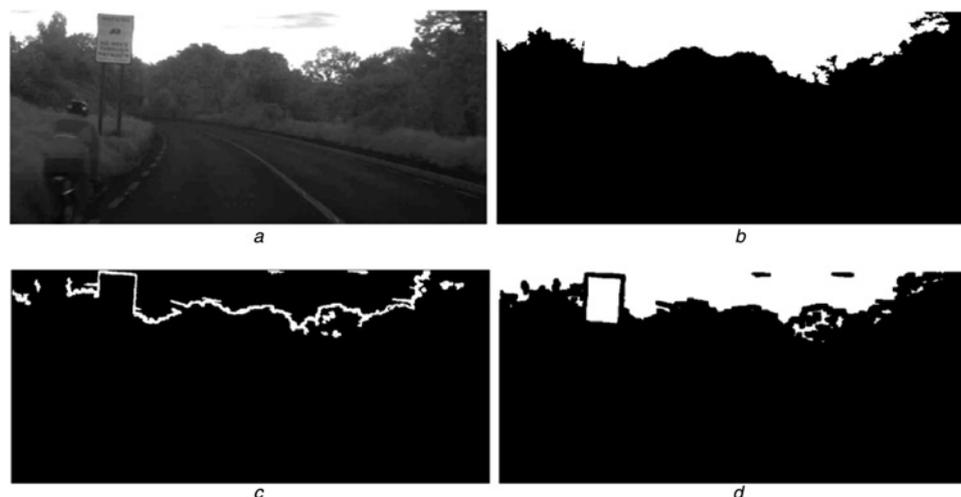


Figure 5 Breaking sky connected components with image edges to release foreground objects

Edges are thickened to improve the process

a Original ROI

b Sky connected component (including sign)

c Edge Image

d Broken Sky Connected Component

Criterion 4 is the only one that considers the shape of connected components in the binary image. More specific and representative shape information (e.g. boundary information) could be used to reject candidates from the set but this may also restrict signs that may be occluded or have irregular shape because of thresholding. Also, a wide variety of sign shapes (i.e. circular, rectangular, diamond, triangular, hexagonal) are of interest to the application, which may lead to a number of valid irregularities being introduced during thresholding. After the constraints above have been imposed on the left binary image, bounding box coordinates are recorded for each connected component and a match is sought in the right image.

3.2.3 Region matching and light intensity comparison: A template matching approach is used to find region correspondences in the greyscale stereo pair. The bounding box coordinates found in the left image are used to create a bounding box in the right image. This is then shifted horizontally across the right image and correlated until the maximum allowable disparity is reached. The maximum correlation value above a fixed threshold (0.7) is used as a match. If there are no matches above the threshold the region is discarded. The correlation function reaches a maximum when the two windows are related by an affine transformation of the intensity values

$$w' = \lambda w + \mu \quad (3)$$

where w and w' are the correlation vectors (bounding box greyscale pixel values) and λ and μ are constants [22]. This is appropriate for this particular application since there will

be an affine transformation between the intensity values for the retro-reflective material in the two images. Note also that the left and right camera reference frames have been transformed so that corresponding image features are located on the same raster scan line in the stereo pair. This procedure is known as epipolar rectification [23] and makes the correlation process somewhat simpler. Epipolar rectification should not be confused with orthorectification, which removes geometric distortions from imagery to reveal a planimetric map.

Because the bounding box will also include pixels from the background when dealing with non-rectangular signs, the binary image is used to form region masks for the bounding boxes (Fig. 6). The masks can then be used to find the greyscale pixel values belonging to the retro-reflective material only. Both the left and right bounding boxes are masked and the resulting non-zero greyscale values inside the bounding box are used to form the correlation vectors.



Figure 6 Road sign bounding boxes and their associated binary masks

The roundabout warning sign above shows how correlation improves when using values that are part of the binary mask only

Using the whole bounding box results in a correlation value of 0.69 where as using only the mask pixels gives a value of 0.87

After all the matching regions are found, the mean region light intensities are analysed. The mean light intensity of the region is also calculated from the masked image pixels. This increases the light intensity differential between left and right-matched regions. If the mean light intensity of the left region (inner cone camera) is sufficiently higher than that of the right region (outer cone camera) ($R_{\mu_l}/L_{\mu_l} < 0.7$, where R_{μ_l} and L_{μ_l} are the right and left mean region intensities), then a retro-reflector is deemed present. Some non-retroreflective regions such as specular reflection and incorrect correlation matches do pass all of the rejection stages. Also non-road sign retro-reflectors such as vehicle licence plates and junction markers also cause problems. Junction markers could be included in the set of correct road sign detections because their purpose is similar and they are of interest to the application. Vehicle licence plates are difficult to reject from the set seen as they have a similar shape and orientation to directional road signs. A combination of local 3D coordinates and region aspect ratio is used to reject these objects from the set. They will typically have a relatively small local Y-coordinate although this can change with the pitch of the vehicle. The addition of an INS to recover the dynamic vehicle pitch would allow true Y-coordinates with respect to a ground plane to be recovered for all features so better position estimates could be obtained. All other non-road sign detections are considered false positives. Classification of the signs is not dealt with in this work seen as all retro-reflective signs are of interest.

3.3 Feature extraction: detection and localisation of defective road studs

Stereo image pairs of retro-reflective road studs are collected at night when the studs are prominent. A number of other objects are also visible in the images

that must be excluded from the set of correctly detected road stud features. In particular, the illumination from overhead lighting can cause a degree of specular reflection from the surrounding area of a road stud. This may have the effect of increasing the image luminance for that area and reducing the contrast between the road stud and the surrounding road surface. A simple image thresholding (using a fixed threshold) solution would be ineffective in this case, see Fig. 7. Thresholding is also made more difficult by the fact that the foreground objects (road studs) constitute a small area of the scene, which makes it difficult to locate them in an image grey-level histogram. An image gradient-based technique for the detection of operational road studs is presented that accounts for the dynamic illumination conditions that are present in a road environment. Once the operational studs have been identified in the images, they are mapped into a global reference frame (according to Section 3.1). Defective road studs can then be identified by analysing the interval between successive operational studs. If this interval exceeds the standard spacing maintained by the relevant authorities, then this area is highlighted as one with absent or defective roads studs that needs maintenance.

3.3.1 Road stud extraction from stereo imagery:

The image analysis for the detection of operational studs proceeds as follows. First, the images are convolved with a 3×3 bidirectional Sobel filter kernel. The highest magnitudes in the image gradient result from pixels on the boundary of retro-reflective objects (e.g. road studs) and other light sources. This makes it possible to find the border of the road studs by thresholding the image gradient. Hysteresis thresholding is applied to the gradient image so that low-intensity borders are not discarded. In ideal images, the road studs are bounded by a closed contour, but this is not always the case and small gaps

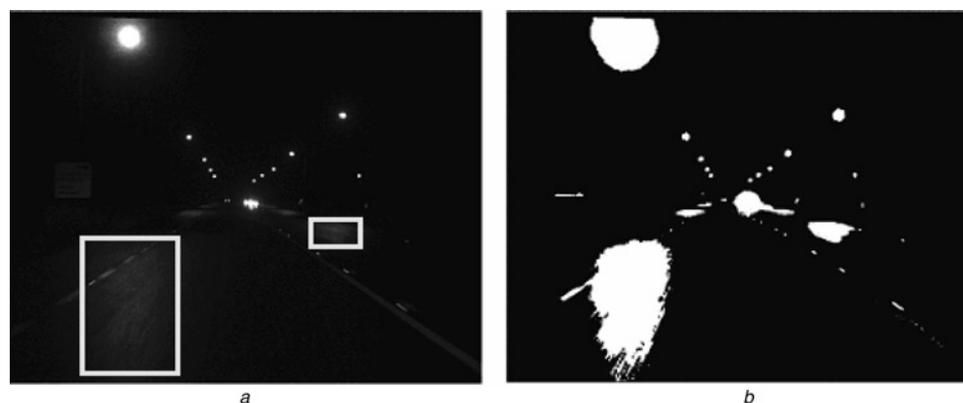


Figure 7 Original image with specular reflection and binary image using a simple global threshold

a Original image with specular reflection

b Binary image using a simple global threshold

Some road studs are lost in the specular reflection regions

may appear in some of the stud's borders. To counteract this problem, the morphological operation 'bridge' is used to complete the object border. The binary object interiors are then filled so a connected component analysis can be carried out. A previous approach to the problem [24] that applied fixed thresholding to the original images suffered from the reflection interference problems mentioned. The improved gradient-based technique is mostly insensitive to the reflections on the surface of the road because they tend to change gradually across the image producing low-gradient magnitudes. The improvement in the gradient-based technique can be seen by comparing the binary images in Fig. 8b with the previous approach in Fig. 7b.

In order to refine the set of candidate road studs even further, some application specific information is used. The areas of all the objects in the binary images are calculated and compared with the maximum allowable area for road studs. Object areas that exceed this are rejected from the set. Another minor refinement uses the fact that in the original image the road studs have a higher contrast with the background that surrounds them. False detections from car head lights and specular reflection do not generally obey this rule. Thus the objects where the mean grey-level intensity of the object pixels is less than the mean grey-level intensity of the neighbouring pixels that surround the object are rejected. Lastly, the object centroids are computed and used as features for stereo correspondence. The features detected are so small that the projective distortion effects are minimal and ignored.

After processing the images using the image processing techniques explained above, the corresponding features are found in the stereo image pair. Most of the time there is only a single possible correspondence for a stud because of image

rectification and the nighttime setting. Where there is more than one possible correspondence, a disparity constrained cross-correlation algorithm is used.

The final rejection factor for road studs is based on an estimation of the ground plane and the knowledge that studs lie on this plane. The road surface within a short-range interval can generally be assumed to be planar. The ground plane is calculated on a frame by frame basis to account for the dynamic pitch of the vehicle because of the road topography and condition. The y -offset from the cameras to a ground plane that is parallel to the optical axis is calculated a priori and gives an initial estimation of where the ground plane should lie. A fixed threshold around this distance is used to determine which feature points are on the ground plane. These feature coordinates are used to update the estimate for the plane so further non-road stud objects can be rejected. Features inside of a 35 m range are considered in order to limit the increasing position error associated with stereo computations beyond this range. The majority of stereo measurements made inside this interval are considered to have sub-metre accuracy. This estimate is obtained by considering a number of factors. A field experiment was conducted where square targets were placed in the field of view of the stereo rig out to 45 m. Their coordinates were then triangulated and compared with the ground truth manual measurements to try and identify the level of systematic calibration error. A linear regression of set against measured values was performed and resulted in a 0.3% deviation in slope and a 0.1 m deviation in intercept from the ideal. The theoretical error in disparity and associated depth is also considered. Given the camera focal length of 16 mm, the stereo baseline of 0.72 m and the pixel dimension of $6.7 \mu\text{m}$, a disparity error of one pixel at 35 m will result in an error of ~ 0.71 m.

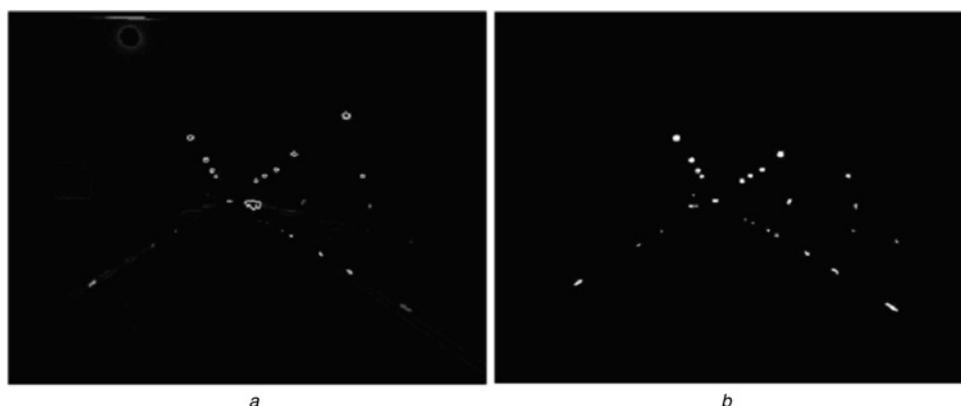


Figure 8 Image gradient using bidirectional Sobel operator and binary image using improved method

- a Image gradient of Fig. 7a using bidirectional Sobel operator
- b Binary image of Fig. 7a using improved method

Incorrect correspondence matches could produce much larger errors but such cases are considered outliers and as a direct result of the correspondence algorithm.

4 Experimental results

This section presents qualitative and quantitative results relating to the problems of road sign detection and defective road stud detection as described in Sections 3.2 and 3.3.

4.1 Experimental results: road sign detection

Results from two road segments are shown in Figs. 9 and 10. In the first segment, the data were collected at nighttime in a rural environment. In the second segment, the data were collected during dusk hours in an urban environment. Both results were obtained under IR illumination to increase scene illuminance and image intensities for retro-reflectors.

Fig. 9 shows the road signs detected at nighttime using the technique described in Section 3.2 and their easting and northing locations on a 5.5 km stretch of road in County Meath, Ireland. The data were acquired in a rural environment.

For the data set presented in Fig. 9, 29 detections of different objects were made by the technique. Of this figure, all 29 objects were road signs. Thirty-one road signs in all are present on this road segment, which implies a detection rate of 94% without any false positives. Road signs that were missed were junction markers that were either too small or too dim in the

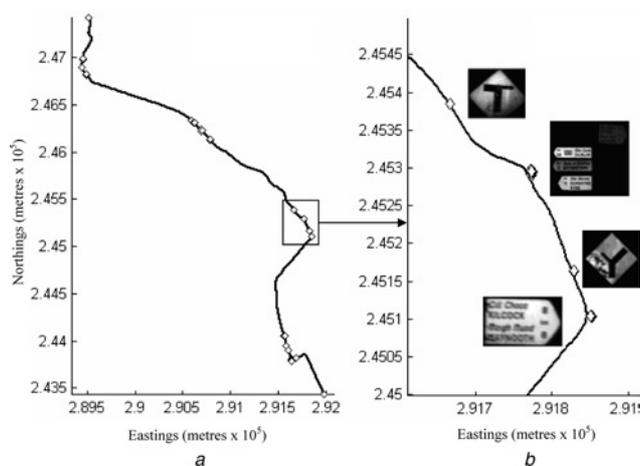


Figure 9 Road signs detected at nighttime and their northing and easting locations

a Locations of the detected features on a 5.5 km GPS trail
b Closer view of part of the road segment showing four features detected and their northing and easting locations

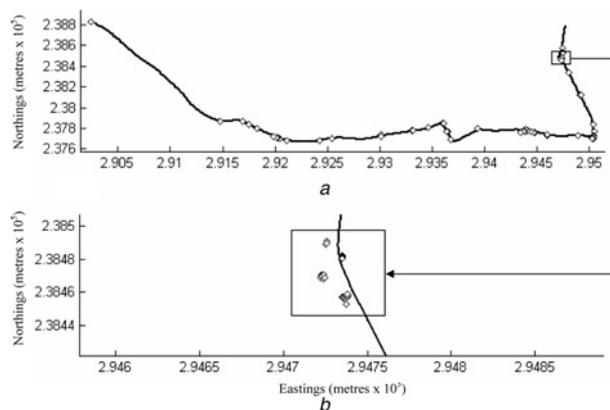


Figure 10 Easting and northing locations of road signs detected on a 6.5 km GPS trail

a Locations of the detected features on a 6.5 km GPS trail heading west
b Closer view of part of the road segment showing five features detected and their northing and easting locations
c Two images showing the detection of the road signs on the GPS trail in b

image. This can happen with a marker that is too small in the image at long ranges and poorly illuminated at short ranges because of its position (e.g. low down on the RHS where the radiant intensity from the IR source is lower).

Fig. 10 shows the easting and northing locations of road signs detected on a 6.5 km GPS trail. The data were acquired in a dusk setting and in an urban environment.

For the data set presented above, 57 detections of different objects were made by the technique. Of this figure, 47 objects were road signs and the remainder were false positives. Fifty-nine road signs in all are present on this road segment, which implies a detection rate of 80% and a false positive rate of 17%. False positives were due to specular reflections and incorrect correspondence matches which both increase in the dusk hours because of increased image information content. Undetected signs included those that were poorly illuminated by the source and did not survive thresholding and those that had a low signal-to-noise ratio in the outer cone image which corrupted the correlation process. The results for road sign detection are summarised in Table 1.

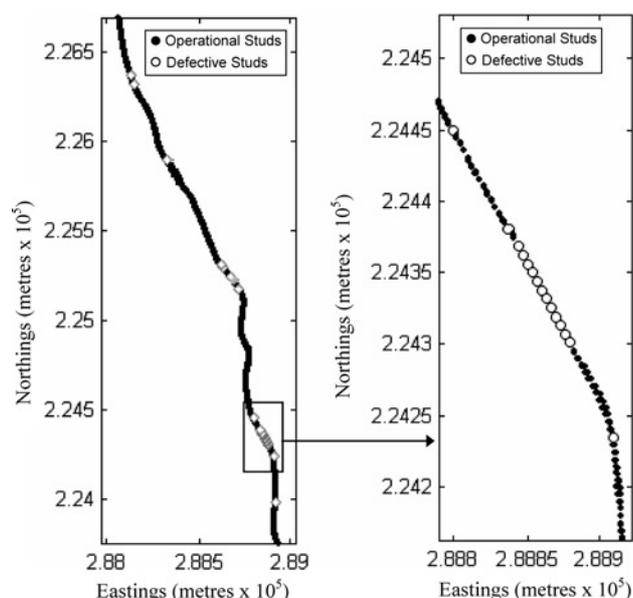
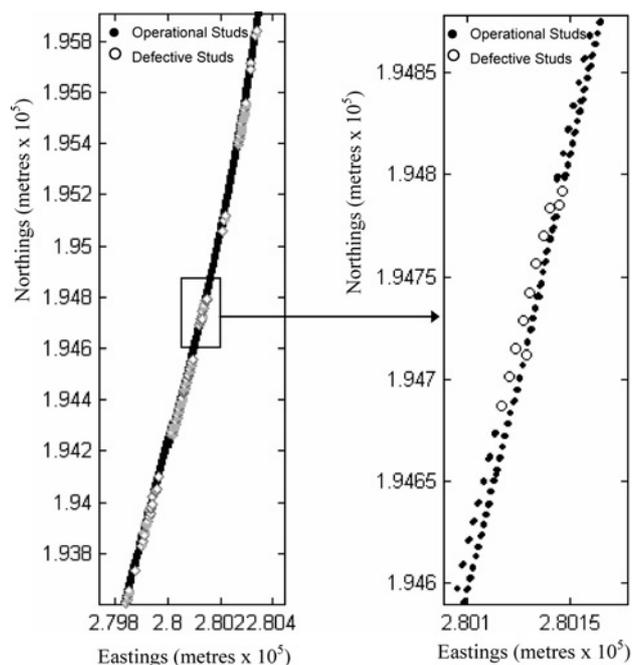
Table 1 Summary of results for road sign detection in dusk and night hours

	Night hours (rural)	Dusk hours (urban)
road signs present	31	59
road signs detected	29 (94%)	47 (80%)
false positives	0 (0%)	10 (17%)

4.2 Experimental results: detection and localisation of defective road studs

The minimum installation spacing for road studs on Irish roads is 12 m although this can be reduced to 6 m on dangerous or busy roads [25]. The technique presented here for detection and localisation of defective road studs assumes the spacing is known a priori. If this is not the case then the technique still detects and locates operational studs so that a manual interpretation can be performed.

Figs. 11 and 12 illustrate how effective the technique is on real-world data sets. The road segment containing studs on the centreline only (Fig. 11) was ~ 3 km in length with a standard stud spacing of 6 m. This implies a total of 500 studs that should be operational. The technique deemed 25 of the studs on this segment to be defective. After a manual inspection of the images, 23 studs were found defective and all of these were present in the set detected automatically.

**Figure 11** National grid locations of defective road studs on the centreline of a 3 km stretch of secondary road in Co. Kildare, Ireland**Figure 12** National grid locations of defective road studs on both sides of a primary road lane stretching for 2.35 km in Co. Kildare, Ireland

This implies a 100% defective stud detection rate with an 8% false positive rate for this data set.

The road segment containing studs on both sides of the lane (Fig. 12) was more demanding because of overhead lighting but high accuracy was maintained [26]. This stretch of primary road was ~ 2.35 km in length. Studs on the centreline should have 6 m spacing but this actually varied between 6 and 12 m which made it difficult to analyse. Six-metre spacing was assumed for centreline studs meaning every other stud was deemed defective when the spacing changed to 12 m. Studs on the edge of the lane had 12 m spacing. The results for the edge of the lane showed that 10 studs were deemed defective from a total of 196. Of this, all were found defective through manual inspection of the images implying a detection rate of 100% without any false positives. For the centreline studs, 65 were deemed defective out of 392. After a manual inspection of the images, 62 studs were found defective and all but one of these was present in the set detected automatically. This implies a detection rate of 98% and a false positive rate of 6%. False positives can generally be attributed to road studs on slip roads that are not illuminated sufficiently and reflections from oncoming traffic and external lighting. The results for defective road stud detection on both road segments are summarised in Table 2.

The results presented are representative of many Irish primary and secondary roads. Longer data sets (e.g.

Table 2 Summary of the defective road stud detection results

	Centreline studs	Left lane edge
actual defective	85	10
detected defective	84 (99%)	10 (100%)
false positives	6 (7%)	0 (0%)

50–100 km) and different operational settings (e.g. motorways) are something that will be acquired and processed post-prototype. However, the real-world segments presented contained many challenges that suggest the system and techniques developed should be applicable to road segments of arbitrary length.

5 Conclusion and future work

In this paper, a mobile and fully portable stereo vision system has been described that has application in many problems in intelligent transportation systems. The system has the capability to make geometric and photometric measurements on objects in a scene and is also a platform to create and test image analysis techniques in the area of ITS. Through the use of stereo vision and GPS technology, the system can position any features analysed on a national grid.

Novel and robust automated feature extraction techniques have also been developed for the detection of defective road studs and road signs. Surface reflectivity properties have been exploited to produce a technique capable of detecting occluded signs and signs of arbitrary shape. Through this synergy of computer vision and robotics, the system is fully autonomous and capable of delivering the said information without any manual inputs.

Although the described system has thus far only been tested on Irish roads, it should have little problems in localisation providing suitable GNSS services are available for the positioning component. There may be an advantage for a GNSS receiver update for some countries to make the most of the available satellite systems (e.g. GLONASS) and augmentation facilities, for example, RTK corrections. The feature extraction routines primarily rely on the retro-reflective properties of road studs and road signs. As retro-reflection is the predominant technology in the developed world for enhancing the visibility of road markings and signage, these algorithms should be easily adaptable for operation in other countries.

Current and future work will see the integration of an INS to the data acquisition system to provide transformation parameters during periods of GPS outage [27]. The fusion of the GPS measurements

with the INS measurements using a Kalman filter approach will also improve positioning accuracy. Daytime operation of the system is also currently being investigated through the use of active illumination and optical filters although current results are confined to dusk and nighttime settings.

Some of the techniques and applications discussed throughout this paper have detected and analysed retro-reflective objects using array-based image sensors. It is also a goal of the system to be able to automatically determine the retro-reflectivity levels (coefficient of retro-reflection, cd/lux/m^2) of these objects from image data acquired from the acquisition system.

Although the systems current purpose is in road feature inventory and analysis, the system is also extendable to other areas such as driver assist systems. An added sign recognition stage could be implemented to identify speed and/or warning signs to alert the driver of any potential dangers. If the global positions of such signs were made available to current satellite navigation systems, then a pre-emptive warning could be issued to the driver revealing information that otherwise may go unnoticed. Obvious system repackaging requirements would exist in such cases.

6 Acknowledgments

The authors would like to thank the Tramore House Regional Design Office (THRDO), the National Roads Authority of Ireland (NRA) and the Council of Directors of the Institutes of Technology for their continued support during the research.

7 References

- [1] BOSSLER J, GOAD C, JOHNSON P, ET AL.: 'GPS and GIS map the nations highways', *GeoInfo Syst. Mag.*, 1991, **1**, (3), pp. 26–37
- [2] BOSSLER J: 'GPS van – input to GIS'. Proc. ION GPS, 1992
- [3] EL-SHEIMY N, SCHWARZ KP: 'Kinematic positioning in three dimensions using CCD technology'. Proc. IEEE Vehicle Navigation and Information Systems Conf. (VNIS), Ottawa, 1993
- [4] SCHWARZ KP, MARTELL HE, EL-SHEIMY N, ET AL.: 'VIASAT – a mobile highway survey system of high accuracy'. Proc. IEEE Vehicle Navigation and Information Systems Conf. (VNIS), Ottawa, 1993
- [5] COSANDIER D, CHAPMAN MA: 'High precision target location for industrial metrology'. Proc. SPIE Videometrics, Calgary, Canada, 1993, pp. 111–122

- [6] GRANSHAW SI: 'Bundle adjustment methods in engineering photogrammetry', *Photogramm. Rec.*, 1980, **10**, (56), pp. 181–207
- [7] GONTRAN H, SKALLOUD J, GILLIERON PY: 'Photobus: towards real-time mobile mapping'. Proc. XXth Int. Society for Photogrammetry and Remote Sensing (ISPRS) Congress, Istanbul, Turkey, July 2004
- [8] TOTH C, GREJNER-BRZEZINSKA D: 'Redefining the paradigm of modern mobile mapping: an automated high-precision road centreline mapping system', *Photogramm. Eng. Remote Sens.*, 2004, **70**, (6), pp. 685–694
- [9] ELLUM CM, EL-SHEIMY N: 'A mobile mapping system for the survey community'. Proc. 3rd Int. Symp. Mobile Mapping Technology, Cairo, Egypt, 3–5 January 2001
- [10] MCLOUGHLIN S, DEEGAN C, FITZGERALD C, ET AL.: 'Classification of road sign type using mobile stereo vision'. Proc. SPIE Opto-Ireland 2005, Dublin, Ireland, 2005
- [11] TRUCCO M, VERRI A: 'Introductory techniques for 3-D computer vision' (Prentice Hall, 1998)
- [12] ZHANG Z: 'A flexible new technique for camera calibration', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2000, **22**, (11), pp. 1330–1334
- [13] REINSCH CH: 'Smoothing by spline functions', *Numer. Math.*, 1967, **10**, pp. 177–183
- [14] HABIB F, UEBBING R, NOVAK K: 'Automatic extraction of road signs from terrestrial color imagery', *Photogramm. Eng. Remote Sens.*, 1999, **65**, (5), pp. 597–601
- [15] DE LA ESCALERA A, ARMINGOL JM, MATA M: 'Traffic Sign recognition and analysis for intelligent vehicles', *Image Vision Comput.*, 2003, **21**, pp. 247–258
- [16] LOY G, BARNES N: 'Fast shape based road sign detection for a driver assistance system'. Proc. IEEE ICIRS (Intelligent Robots and Systems), Sendai, Japan, 28 September–2 October 2004, pp. 70–75
- [17] MIURA J, KANDA T, SHIRAI Y: 'An active vision system for real-time traffic sign recognition'. Proc. IEEE Intelligent Transportation Systems, Detroit, October 2000, pp. 52–57
- [18] GIL-JIMENEZ P, LAFUENTE-ARROYO S, GOMEZ-MORENO H, ET AL.: 'Traffic sign shape classification evaluation. Part II. FFT applied to the signature of blobs'. Proc. IEEE Intelligent Vehicles Symp. 2005, 6–8 June 2005, pp. 607–612
- [19] ARENS JB, SAREMI AR, SIMMONS CJ: 'Color recognition of retroreflective traffic signs under various lighting conditions'. *Public Roads*, vol. 55(1), Federal Highway Administration, 1991, pp. 1–7
- [20] Traffic Signs Manual, Department of the Environment, Irish Government, 1996
- [21] CANNY F: 'A computational approach to edge detection', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1986, **8**, (6), pp. 679–698
- [22] FORSYTH D, PONCE J: 'Computer vision: a modern approach' (Prentice Hall, Upper Saddle River, New Jersey, 2003)
- [23] TRUCCO M, VERRI A: 'Introductory techniques for 3-d computer vision' (Prentice Hall, 1998)
- [24] MULVIHILL C, DEEGAN C, FITZGERALD C, ET AL.: 'Mobile stereo vision for the identification of defective road studs'. Proc. Irish Machine Vision and Image Processing Conf. (IMVIP), Belfast, Ireland, 2005, pp. 135–142
- [25] European Commission Directorate-General for transport: 'COST 331 – requirements for road marking', July 1997
- [26] MULVIHILL C, DEEGAN C, FITZGERALD C, ET AL.: 'Mobile stereo vision for the mapping of defective road studs'. Proc. IEEE Advances in Cybernetic Systems (AICS), Sheffield, England, 2006
- [27] FOY S, MCLOUGHLIN S, DEEGAN C, ET AL.: 'Road sign safety identification through the use of a mobile survey system'. Proc. 5th Int. Symp. Mobile Mapping Technology (MMT '07), Padua, Italy, May 2007