Surveying by motion tracking: modelling 3D subterranean landscape from video imagery

Anna E Mason[1], Dr David M Mountain[2], Prof Jonathan Raper[3]

[1,2,3] City University, Northampton Square, London, EC1V 0HB
Tel. +44 (0)20 7040 5060
Email. a.e.mason@city.ac.uk, dmm@soi.city.ac.uk, raper@soi.city.ac.uk

KEYWORDS: GIS, 2D3 Boujou, Camera Tracking, Video Imagery, Caving

1. Introduction

Computer Vision is concerned with obtaining information from image data, such as the analysis of 2D image sequences to identify and represent real world objects in 3D space. As stated by Hartley and Zisserman (2000) Computer Vision builds upon knowledge from physics, the psychology of perception and the neurosciences, but the strength of achievement rests undoubtedly upon the ability to describe complex geometric concepts and translate them into computer code. The last 15 years of substantial research on motion tracking systems has produced Augmented Reality (AR) applications ranging from robotic navigation, computer assisted surgery, augmented reality, virtual TV production and creation of special effects in the movie making industry (Welch and Foxlin, 2002; Bhanu and Paulidis, 2005). In the last seven years commercial vision-based camera tracking systems, known interchangeably as matchmoving systems, have emerged which utilise the geometry of multiple uncalibrated views to generate digital 3D representations from 2D video or film footage (Hartley and Zisserman, 2000) . A synergy of this technology with GIS could provide a significant step forward in the efficiency of three dimensional data capture and modelling techniques.

One area where this could prove particularly beneficial is surveying and modelling subterranean cave and karst systems, where standard positioning systems such as Global Positioning Systems (GPS) are unavailable and surveying techniques are typically laborious or require costly laser scanner and range finder equipment. Given the challenging nature of caves, underground and confined with no access to infrastructure or view of the sky, these environments will not benefit from the investment going into the next generation of Global Navigation Satellite Systems (GNSS) such as the Galileo (European Commission, 2007) or wireless positioning systems. In this context, survey techniques that could use handheld cameras are worthy of investigation. The emphasis of this research is to present a method for using matchmoving or camera tracking software to perform 3D reconstruction of a scene and to evaluate how this technology can be applied to surveying subterranean environments.

2. Video Data Acquisition in the Field

The experiments were conducted in three UK study areas; one chalk mine and two caves formed by dissolution of limestone. Sections of each mine or cave were surveyed using conventional line survey methods and equipment (eg. measuring tape and compass) as illustrated in Figure 1.

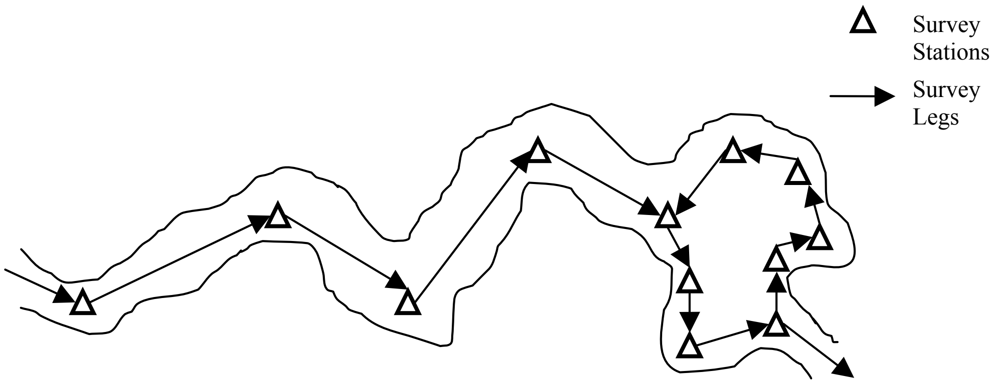


Figure 1. Line Survey. Modified from Day (2002)

The survey centreline of each passage was then filmed using portable video camera and lighting equipment. The lighting was built using a wide angle 50W halogen bulb and torch casing, connected to a lead acid battery. PYSOP markers (2d3a, 2005) were used to create recognisable features for analysis. By identifying key issues relating to technique, equipment, environment and location, each experiment contributed to systematically refine and improve the data acquisition methodology. The process is summarised in Figure 2.

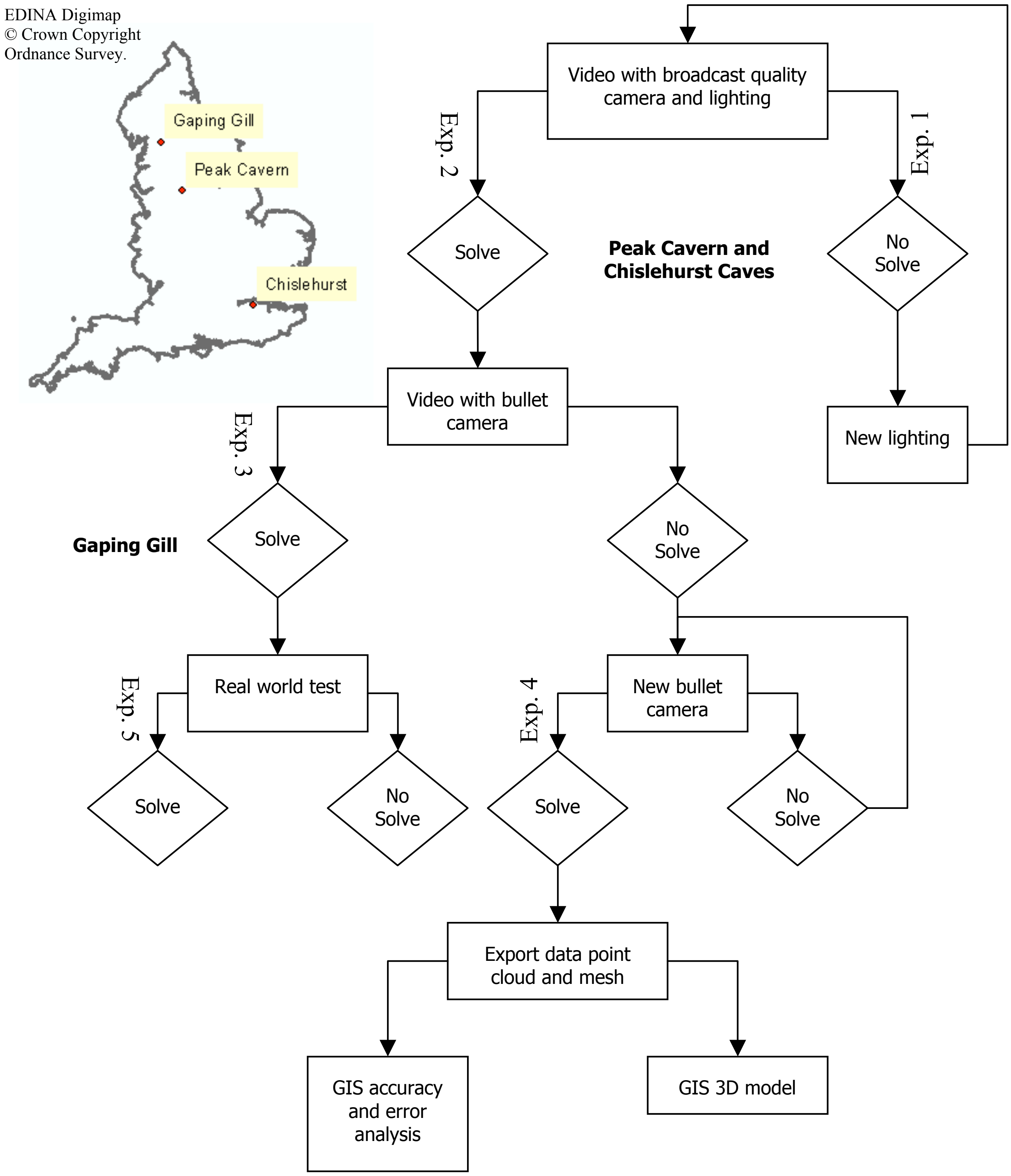
3. Three Dimensional Reconstruction and Modelling

Video image sequences were sectioned to represent survey legs of up to 40m in length. Each image sequence was processed with Boujou4 (2d3, 2007) and evaluated for quality using standard matchmoving guidelines, such as visual cues and reprojection error, as outlined by authors such as Dobbert (2005).

Matchmoving software performs analysis of raw 2D imagery to produce optimal estimates of camera movement and rotation. This is achieved by tracking 2D points of interest as they move in the frame through time (Figure 3a). Selection is based on an algorithm such as corner detection in Boujou4 (2d3, 2005b) or edge detection as in Kemp and Drummond (2004). Corner detection is likely to be a more appropriate algorithm in organic environments as there are few straight lines or edges to match that will not change radically with aspect and illumination. Also, Kemp and Drummond (2004) suggest that high density of similar features (like repeating rock texture) is not beneficial to edge detection.

2D feature points are then matched between frames using a 3D pose estimation algorithm which is often Random Sample Consensus (RANSAC) as in Nister et al. (2006), Mouragnon et al. (2006), Kemp and Drummond (2004), and Zhang et al. (2006). 3D pose estimation is described by Fisher et al. (2005) as determining the transformation of an object in one coordinate frame with respect to the same object in another coordinate frame. The software creates a virtual camera and estimates the camera parameters, movement and rotation in relation to the 2D feature tracks (Dobbert 2005; Fitzgibbon and Zisserman 2003). This process is called 3D calibration. This information is used to calculate the depth of objects in a scene, producing a 3D point coordinate cloud known as a “camera solve” (Figures 3b and 3c). In addition, a surface mesh can be generated from the coordinate point cloud using a form of delaunay triangulation (Figures 4a and 4b).

Figure 2. Summary of the data acquisition and 3D reconstruction workflow



b)

a)

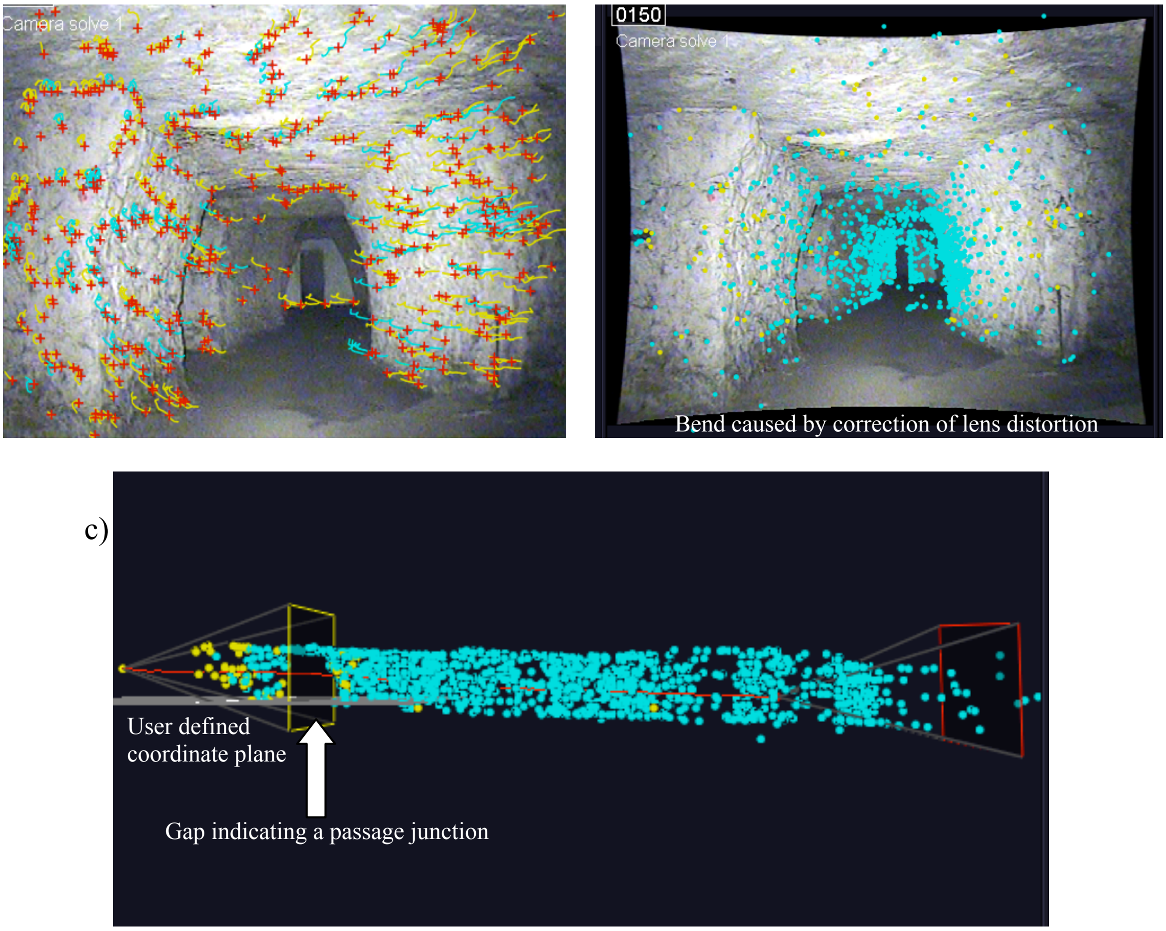
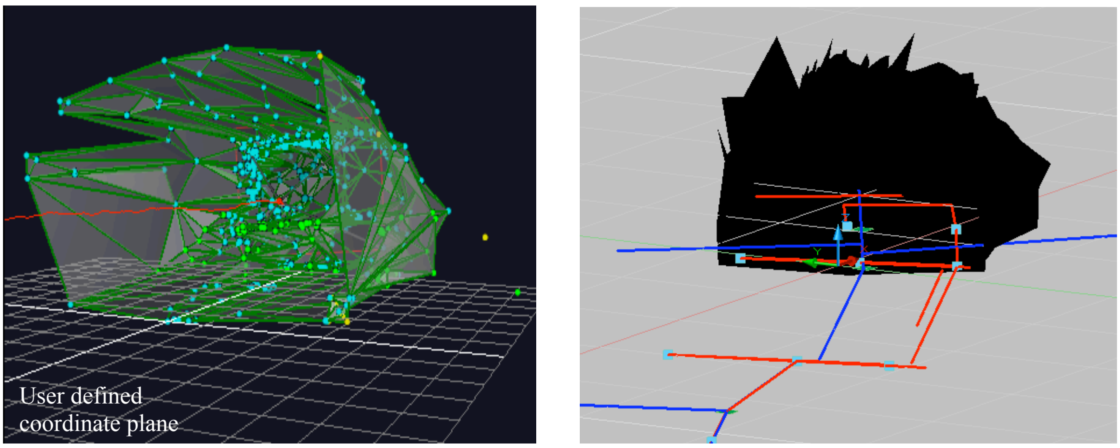


Figure. 3 a) 2D view of 2D feature tracking. The red crosses are points of interest and the yellow vectors are their tracks in previous frames. b) 3D points generated by a successful matchmoving “camera solve”. c) 3D side view of 3D points generated for a survey leg of cave passage. The red line through the centre of the points is the path the camera travelled. The yellow square is the camera’s projected Field of View.

In Figure 4a the cave surface appears angular. CAD, GIS and animation tools were tested available to improve appearance, including use of photorealistic lighting and materials (Figure 4c). Alternative GIS and CAD modelling approaches, such as lofting between vertical sections or creating a rectangular surface grid, were also possible using matchmove data.

a)



b)

c)

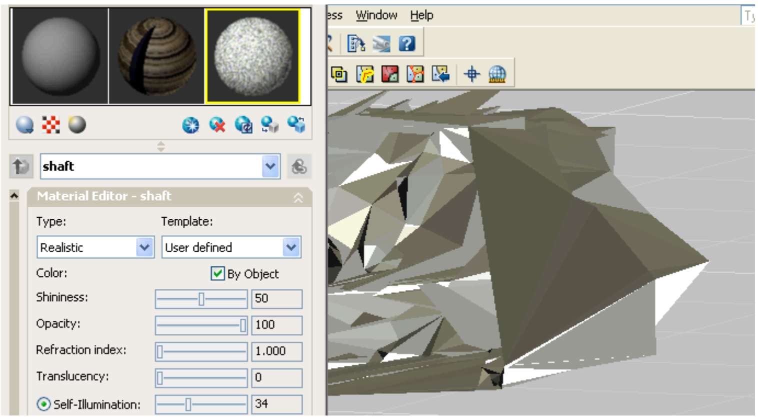


Figure 4 a) 3D view of a mesh in Boujou4. b) The same mesh mapped to the line survey results (red and blue lines) in GIS. c) An example of lighting and colour created from a single frame of video.

4. GIS Analysis and Modelling

GIS/CAD tools were used to assess the results visually and quantitatively. Detailed assessment of positional accuracy was precluded by insufficient control data to adequately establish a 3D model for comparative analysis. Existing LIDAR was incomparable as it maps large chambers (Westerman et al., 2003; Murphy et al., 2005), however potential exists to generate 3D digital representations of entire connected systems by merging LIDAR data of chambers with matchmove data of interconnecting passages.

4.1 Three Dimensional Model

Survey results from Chislehurst were used to create a 3D model connecting surface and subsurface features.

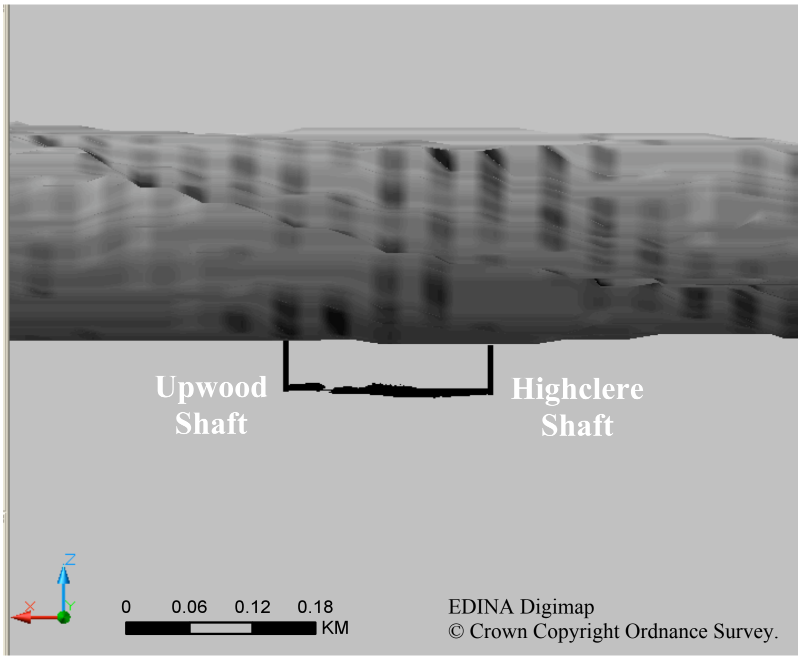


Figure 5. The mine passage connecting Upwood Shaft to Highclere Shaft at Chislehurst, Bromley

4.2 Scale

Relative scale accuracy was tested on a sample coordinate point cloud using a small sample of ground truthed measurements. The scale differences ranged from 0.2cm to approximately 20.3cm. Mean of the four measurements was 13.3cm and RMS deviation from the average was 8.1. When considered in proportion to height and width, these measurements were found to fit well within 10-15% guideline suggested by Day (2002) for sketched cave profiles.

4.3 Shape

Figure 6 shows a sample of 3D polyface meshes sectioned and analysed for accuracy of shape. The interior of each cross section is comparatively similar to the field sketches. Gaping Gill appears to be the least representative in terms of shape. This pattern is explained by the appearance of rocks inside the passage.

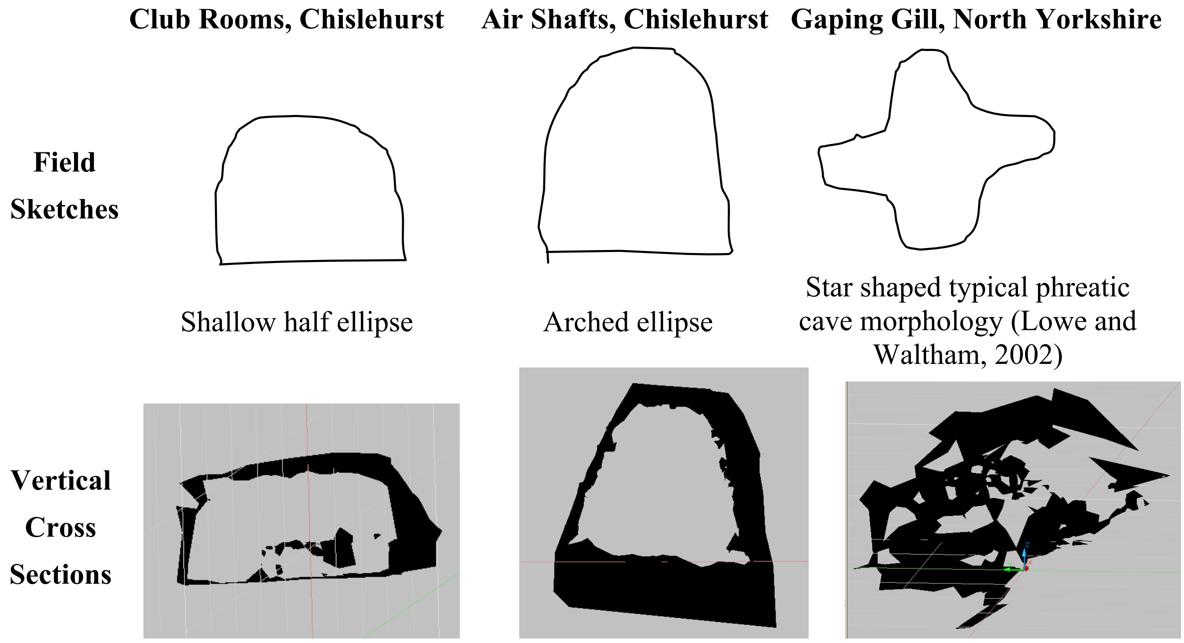


Figure 6. Cave survey sketches compared to 3D polyface meshes generated using Boujou4

5. Discussion and Conclusions

This research showed that matchmoving programs are sufficiently robust to track cave passage scenes using consumer quality portable cameras and lighting. Current research points in the direction that this technology is moving; towards faster, more accurate, real time processing, with less user input required to achieve 3D reconstruction of real scenes. The commonality of motion tracking applications is that they effectively perform digital surveying tasks of real scenes. This research demonstrates how this digital survey data can be used in GIS, thus presenting an example of a new compatible source for 3D data. The data acquisition and processing method appears particularly well suited to challenging environments where other forms of surveying may be less accessible, such as cave systems. Yet if this method was adopted as a mainstream form of surveying, it would be applicable wherever anyone has access to a video camera of reasonable resolution, any place and at any time.

Initial results and observations indicate that where good 3D calibration has been achieved, relative scale and shape are shown to be accurate, although the technique and sample sizes presented here are insufficient to give more than approximate quantitative measures. Even if the optimal pose estimation nature of monocular vision tracking systems limit the process from achieving accuracy comparable to professional surveying equipment, the process still demonstrates potential for many applications where lower spatial sampling and accuracy are required, such as cave surveying, navigation, and search and rescue. Actual comparison of such surveying techniques presents an opportunity for further study.

References and Citations

2d3 Boujou4 (2007) http://www.2d3.com/ (accessed 20/12/07)

2d3 (2005a) What sort of tracking markers does boujou track the best? [Online] http://www.2d3.com/php/support/faqs.php (accessed 20/12/07)

2d3 (2005b) Boujou Three Help - Tutorial 1: Basic Tracking [Offline Electronic Manual]

Bhanu, B. and Pavlidis, I. (2005) *Computer Vision Beyond the Visible Spectrum: Advances in Pattern Recognition*, London, UK: Springer.

Day, A.J., (2002) *Cave Surveying – A guide to the equipment, techniques and methodology of the BCRA system*, Cave Studies Series No.11. Buxton, UK: British Caving Research Association (BCRA).

Dobbert, T. (2005) *Matchmoving: The invisible art of camera tracking*, Sybex Inc., Alameda, CA.

European Commission (2007) EU Galileo Project [Online] http://ec.europa.eu/dgs/energy_transport/galileo/index_en.htm (accessed 20/12/07)

Fisher, R. B., Dawson-Howe, K., Fitzgibbon, A., Robertson, C., and Trucco, E. (2005) *Dictionary of computer vision and image processing*, UK: Wiley

Fitzgibbon, A. and Zisserman, A. (2003) Automatic Camera Tracking. In Shah, M. and Kumar, R. (eds), *Video Registration,* UK: Kluwer, pp18-35

Hartley, R. and Zisserman, A. (2004) *Multiple View Geometry in Computer Vision (Second Edition)*, Cambridge, UK: Cambridge University Press.

Kemp, C. and Drummond, T.W. (2004) Multi-Modal Tracking using Texture Changes, *In Proceedings of the British Machine Vision Conference*, pp73–77

Mouragnon, E., Dekeyser, F., Sayd, P., Lhuillier, M., and Dhome, M. (2006) Real Time Localization and 3D Reconstruction, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Volume 1, pp363 – 370

Murphy P., Parr A., Strange, K., Hunter, G., Allshorn S.L., Halliwell, R.A., Helm, J., Westerman, A.R. (2005) Investigating the nature and origins of Gaping Gill Main Chamber, North Yorkshire, UK, using ground penetrating radar and lidar, *Cave and Karst Science*, Volume 32, pp25-38

Nistér, D., Naroditsky, O., Bergen, J. (2006) Visual odometry for ground vehicle applications, *Journal of Field Robotics,* Volume 23(1), pp3–20

Welch, G. and Foxlin, E. (2002) Motion Tracking: No Silver Bullet, but a Respectable Arsenal, *IEEE Computer Graphics and Applications,* Volume 22(6), pp24-38

Westerman, A.R., Pringle, J.K., Hunter, G. (2003) Preliminary LIDAR survey results from Peak Cavern Vestibule, Derbyshire, UK. *Cave & Karst Science*, Volume 30(3), pp129-130

Zhang, G., An. X; Chen, W; Bao, H., Qin, X. (2006) As-consistent-As-possible compositing of virtual objects and video sequences, *Computer Animation and Virtual Worlds*, 17(3-4), pp305-314

Acknowledgements

Sincere thanks to the British Caving Research Association for funding and the resources to explore the project in such depth. I am equally very grateful to 2d3 for generously providing Boujou4. Thanks to John Farrer of Peak Cavern and Terry Hunt of Chislehurst Caves, for allowing access to the caves. Thank you to Rod LeGear, Kent Underground Research Group, for his help at Chislehurst Caves. Thanks to Shaun Puckering of Hitch n Hike, Hope Village, for his advice on building caving lamps and for lending his cave surveying gear, and to Adam Evans of Outdoor Instruction, for his guidance at Peak Cavern. Thanks to the Craven Potholing Club for making Gaping Gill such an inspiring experience. Thank you to Paul Chambers and Sarah Mason for surveying and AV assistance.

Biography

Anna E Mason is an MSc GIS graduate from City University. Her undergraduate degree is from Canterbury University, New Zealand. Anna currently works as a GIS Developer and is an operational volunteer for MapAction. Research interests include applications for graphics and animation technology in GIS.

Dr David M Mountain is a research fellow on the LOCUS project and lecturer in Geographic Information at City University. Research interests include location-based services, modelling individual accessibility, modelling time and space, virtual and augmented reality environments and mobile computing.

Prof Jonathan Raper is Head of the Information Science Department at City University. Previously he was at Birkbeck College, University of London. Jonathan has led projects and published research on location-based services, GI policy, philosophy of spatial and temporal representation, and the handling and analysis of environmental information.