Geoinformation Science Journal, Vol. 9, No. 2, 2009, pp: 21-35

STEREO-IMAGE CAPTURE TOOL FOR FORENSIC MAPPING OF SURVEILLANCE VIDEO FOOTAGE

Mohd Farid Mohd Ariff[1], Zulkepli Majid[1] Halim Setan[1] and Albert K. Chong[2]
[1]UTM-Photogrammetry and Laser Scanning Research Group, Universiti Teknologi Malaysia
[2]Faculty of Engineering and Surveying, University of Southern Queensland
mohdfarid@care2.com, zulkeplimajid@utm.my & halim@utm.my

ABSTRACT

This paper discusses the evaluation of a video stereo-image capture tool for photogrammetric applications. The "NuView Stereo Adapter", a stereo-imaging device and a high-definition video (HDV) were used to capture stereoscopic video footage. The HDV video becomes a complete stereo-imaging system when stereo adapter is mounted onto video lens barrel and could be used to capture stereoscopic video footage for many photogrammetric applications. Relevant photogrammetric applications might include: motion study, 3D animation and aerial/security surveillance. The device permits two distinct views (left and right perspective views) to enter a single lens HDV. A single convergence control in the device allows the user to obtain a stereo view of near and far objects. The system is easy to operate and the stereo-images give a good visual 3D effect. Research was carried out to determine the quality of the stereo-image for 3D measurement. The investigation involved: (1) the calibrations of a customized system for optimum close range photogrammetric application; and (2) a case study involving forensic mapping of surveillance video footage. A semi-automated and manual monodigitising photogrammetric technique was used to determine the accuracy of the system. The results of the research show that the system is capable of obtaining spatial data with 10 mm accuracy using a single video camera. As the resolution of an off-the-shelf video camera is high, the video footage can achieve 30 mm accuracy under 5 m object distance.

Key words: NuView adapter, high-definition video, stereoscopic video, camera calibration, forensic mapping.

1.0 INTRODUCTION

The main objective of this paper is to discuss recent research on the use of stereo-image capture tool, commonly known as a NuView stereo adaptor for close range photogrammetric application. The paper covers: (1) discussion of the principle of the stereo adaptor; (2) a technique for calibrating the customized system; and (3) evaluation of the stereo-image for a case study involving forensic mapping of surveillance video footage.

Nowadays, a number of stereoscopic video capture devices may be purchased ‘‘off-the-shelf’’. These devices are commonly designed based on the following arrangements (Montgomery et al., 2002; Lee et al., 2004):

i. **A single digital video camera (DV) is used to record a scene simply by moving it. A stereoscopic view of the scene can be achieved by selecting sequences which give optimum stereo disparity. The system is not suitable for moving scenes**

ii. **Two DVs are set up at a specific distance apart and both record a scene simultaneously, using, for example, the 3D Video Encoder (3D ImageTek Corp., 1998). To synchronize the DVs a generator locking device is needed and this ensures accurate matching of the recorded (left and right DV) sequences to obtain stereoscopic coverage (Woods, 2007)**

iii. **A biprism adaptor is placed in front of the DV. The equivalent of a stereopair of images is formed as the left and right halves of a single CCD image using a biprism (Lee et al., 1999)**

iv. **A 3D adaptor is placed in front of the DV. Shutters are used to record the left and right views as odd and even fields (the NuView Adapter developed by i-O Display Systems of Sacramento, California)**

Stereoscopic video footage has been used for a range of photogrammetry applications which include aerial surveillance, human and animal motion study (tracking), bank/security surveillance, industrial measurement and underwater work.

Woods (2007) showed that a Mini-3D underwater stereoscopic video camera system designed for use on remotely operated vehicles (ROVs) is suitable for underwater pipe inspection in the oil and gas industry. Woods stresses that a stereoscopic video provides many advantages over a regular 2D video such as:

i. Improved ability to see through suspended matter or turbid water.

ii. Improved ability to see through visual clutter.

iii. Improved ability to judge size and distance.

Chong (2007) reported on the use of the stereo adapter (NuView) for motion study. A customized stereo-imaging system was evaluated. The system is easy to operate and is efficient for stereo-imaging capture and the amount of stereo overlap is always the same. The results of the research show that the system is capable of obtaining spatial data with 13 mm accuracy using a single video camera. As the resolution of an off-the-shelf video camera is coarse, the video footage can only achieve sub-centimeter accuracy less than 10 m object distance. Nevertheless, the stereo-video footage can be used for other application as well, such as virtual athletic training, medical research, surveillance and field performance study.

Chong and Schneider (2001) used a dual video camera technique to study the population age distribution of bottlenose dolphins in New Zealand waters. The stereo system was ideal for the dolphin study because these cetaceans move along the bow of a moving vessel at a convenient speed which was easy to match with the boat used.

Studying human motion, Burmeister *et al.*, (2005) showed that video clips from two synchronized Sony XC-ST50CE monochrome video cameras could be used to detect hand-grasp movements and corresponding biosignals (EMG) at the same time with high temporal and local precision. Similar research was conducted by Chong (2004), and Tzovaras *et al.*, (1997).

According to Chong (2007), NuView is probably the best low-cost device for the planned close range photogrammetric applications because:

i. **It is easy to set up**

ii. **There is no degradation in the speed of recording (number of frames per second)**

iii. **There are very few adverse effects on the quality of the left and right perspective views using this form of synchronization**

iv. **Image processing is uncomplicated**

The drawback of the system is that the row resolution of the captured stereo-image is reduced by half (that is, from 760×560 to 380×560) and the image must be resampled to obtain a full-frame image for high-quality stereodigitising.

2.0 EQUIPMENT & SOFTWARE

2.1 NuView Stereo Adapters

NuView stereo adapter is a video camera add-on (**Figure 1**). It turns almost any ordinary camcorder into a 3D movie machine and works for all current video include HDV. Stereo adapter usually does not interfere with zoom and autofocus functions of the video. It consists of two LCD shutters, a prismatic beamsplitter and an adjustable mirror (Stereo3d, 2007). The mirror/prism system puts the video lens into the center of the light rays of a left and a right eye view (**Figure 2**).

Figure 1. NuView stereo adapter & HDV

The shutters ensure that the camera lens receives only one of the views at a time. The stereo adaptor is connected to the video-out port of the video camcorder. This way, the shutter can synchronize to the recording. Hence, the stereo adaptor provides a 3D field-sequential recording capability, that is, it allows the recording of a field-sequential video onto a DV tape or onto the computer via the firewire connection. The stereo adaptor was originally produced for home movie production and virtual reality animation (for example, for visual fly-through products for computer games).

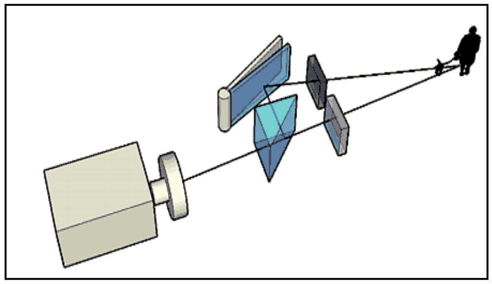

Figure 2. Principle of the stereo adaptor

2.2 High-Definition Video

One unit Sony HC-5E HDV camera was used in the research (**Figure 1**). It has a spatial resolution of 6 megapixels and has a principal distance (PD) ranging from 5.1 to 51 mm. The video format was 1920×1080 pixels and each pixel represents 2.3 μm on the CMOS (Complementary Metal Oxide Semi- conductor). Combining superb picture quality of HDV 1080i, the HC-5E with feature latest-generation ClearVid CMOS Sensor technology provide high-definition pictures with greater resolution and sensitivity. The HDV becomes a complete stereo-imaging system when stereo adapter is mounted onto video lens barrel.

2.3 Software

Three types of software were essential for the case study on the suitability of the system for forensic mapping. They were:

i. **VirtualDub: Field de-interlacing software.**

ii. **Adobe Photoshop: Imaging processing/editing software.**

iii. **Australis: Camera calibration, 3D spatial data editing and mono-image digitizing software.**

3.0 METHODOLOGY

The research entails three stages. The first involves the system calibration. The next comprises the testing of the system evaluation, while the last stage consists of a case study. A case study was conducted on mannequin and human subjects. Initially, mannequins were used to avoid any error due to movement during image capturing.

3.1 System Optic Calibration

Today close-range photogrammetry is mostly based on the use of non-metric solid state sensor cameras. To translate the accuracy potential of the cameras into the object space, they must be modeled and calibrated (Chong and Stratford, 2006; Wolf and Dewitt, 2000). Standard non-

metric camera calibration is well documented (Fryer, 1989; Beyer 1992; Peterson *et al*. 1993; Fraser and Edmundson 1996). This process includes the determination of the principal point of autocollimation (PPA) (x_p and y_p), the principal distance (PD) (c), the radial lens distortion parameters (k_1, k_2 and k_3), and in some instances the dynamic fluctuation.
The calibration of the system was carried out using a portable calibration frame as shown in **Figure 3**. Retro-targets were placed on the frame and on the bracing at three level of depth (0-mm, 75-mm & 150-mm). Retro-targets are highly reflective targets, which are specially made for precise automated digitizing using computer software (Atkinson, 1996). The coordinates of the retro-targets on the portable control frame were obtained photogrammetrically.

Figure 3. A portable calibration frame

Both the HDV lens and NuView were set to focus at infinity. The HDV focal length was set at the maximum wide angle (5.1 mm). The HDV was mounted on a tripod. The position of the HDV mount and the portable calibration frame were adjusted so that the object distance of 3.0 m would be similar to the object distance of the case study. The 3.0 m object distance was found to be suitable for the chosen video.

A short video footage of the portable control frame was taken by tilting the control frame through four different rotations (**Figure 4**). This process was repeated four times. Each piece of the recorded footage was processed by VirtualDub software so that the left and right perspective views could be separated and stored as separate images. As the light-ray paths for left and right views were not exactly the same through the adaptor, the corresponding optics was calibrated separately. Australis camera calibration software was used to compute the individual interior orientation parameters using convergent images

Figure 4. Four different rotations (left image)

3.2 System Evaluation

The task involved testing the equipment in a photogrammetric environment and analyzing the quality of the photogrammetric measurement to determine the accuracy aspect of the computed 3D data. In general, the object distance of the photography was determined by the project specifications of the case study. In this case, the selected distance was 3.0 m between the HDV and the subject. However, a range of distances was considered important in the testing. For an object distance of 2.5 m or shorter, a preliminary test showed that the area of stereo coverage was too small for most photogrammetric applications involving bank/security surveillance. Consequently, the HDV position was set at 4.0 and 5.0 m from the subject.

A mannequin was used in the evaluation process. Retro-targets were placed on the mannequin, which was positioned 0.1 m in front of the control frame (**Figure 5**). The control frame was used to provide the object-space control for the exterior orientation of stereo-images obtained by the stereo adaptor. A short video clip of the mannequin was taken and the process was repeated four times.

The video clips were de-interlaced to extract the individual left and right images of the test range at each object distance. **Figure 6** shows a stereopair of images de-interlaced from the video frame. Each stereopairs were selected from each clip randomly. Therefore, a total of 20 (5 object distance × 4 repeats) stereopairs were analyzed. A semi-automated monodigitising technique using Australis software was used to carry out interior and exterior orientations of these stereopairs. Test points of known coordinates were used to compare with the captured values to determine the 3D accuracy of the computed data.

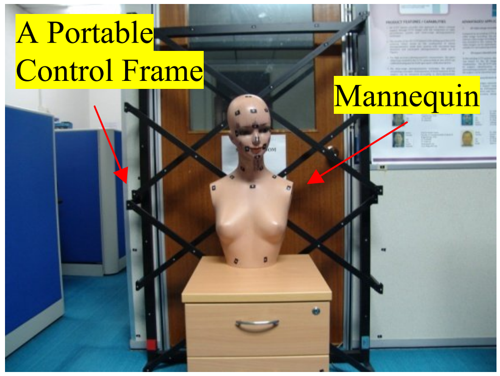

Figure 5. Mannequin, was set up at a short distance in front of a control frame

Figure 6. A stereopair of images de-interlaced from the video frame

3.3 Case Study

The study involved taking video clips of a suspect (**Figure 7**) which was placed at a short distance in front of the photogrammetric control frame. The camera and the adaptor settings were made similar to the values used during the system evaluation exercise. The HDV was set up at 3.0 m from the object.

Figure 7. A suspect was set up at a short distance in front of a control frame

The stereo-images from stereo adapter were analyzed based on manual monodigitising technique using Australis software to carry out interior and exterior orientations of these stereopairs. The computed coordinates of the craniofacial/ anthropometry landmark of the suspect were used to compare with the true coordinates to determine the accuracy of the video for forensic mapping and measurement. Subsequently, the video clips were processed according to the procedure discussed in the system evaluation.

4.0 RESULTS AND DISCUSSION

4.1 HDV Calibration

The results of four sets of camera calibrations are shown in **Table 1**. The results show that the focal length '*c*' of the left view and right view was different by about 0.104 mm. A Student's t-test ($t\alpha(0.05)$) shows that the difference between the means is significant (H_0: $\mu1-\mu2=0$ was rejected). Based on **Figure 8**, the principal point offset (x_p & y_p) and the radial lens distortion (k_1) values vary by a small amount between the two views. Most of the calibrated parameters are slightly higher for the right view. According to Chong (2007), the change in the parametric values could be the result of a longer light path through the right view port.

Table 1. HDV camera parameters

Camera Parameter	Left View		Camera Parameter	Right View	
	Mean (mm)	Std. Dev (mm)		Mean (mm)	Std. Dev (mm)
c	5.20233	0.005786	c	5.30655	0.005647
x_p	0.07950	0.006351	x_p	0.08700	0.008042
y_p	0.05800	0.007394	y_p	0.06875	0.005852
k_1	0.00068	0.000067	k_1	0.00081	0.000024

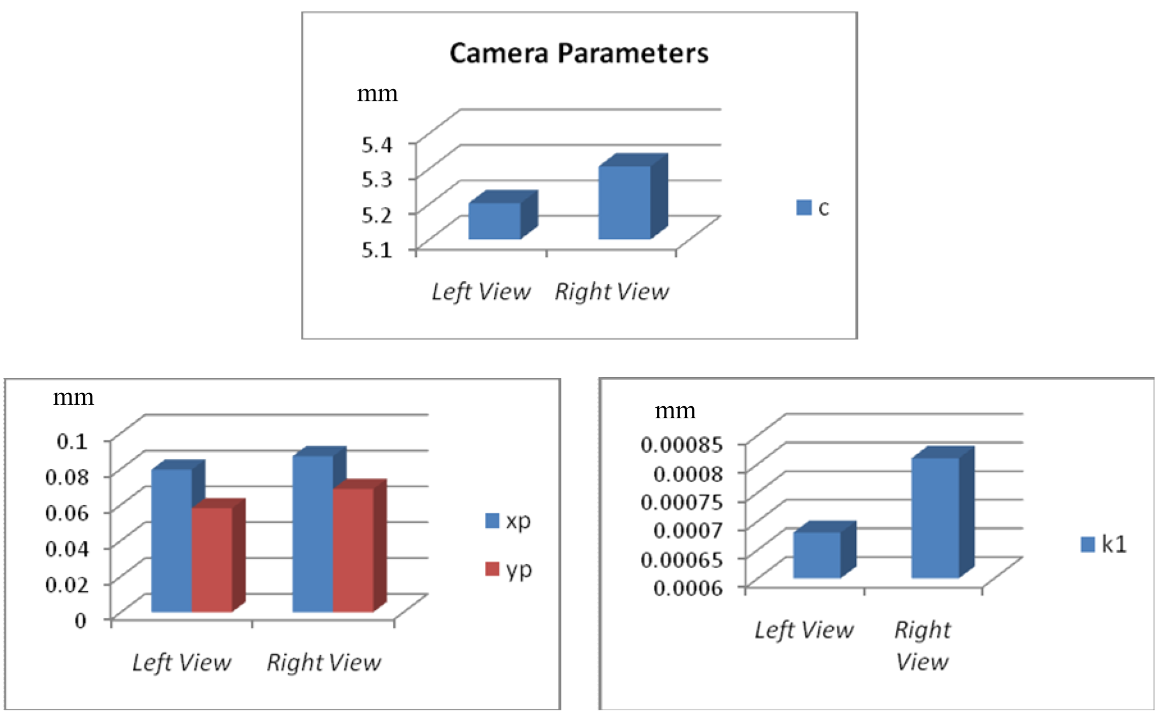

Figure 8. Left view vs. right view

4.2 System Evaluation

At this stage, a set of six test points selected from targets in the test range was used as reference in the experiment (**Figure 9**). The 3D distances between these targets were used as true values in the evaluation of the new technique. The relative positions of these targets were determined by a high-precision bundle technique using multi-convergent photography and the results were checked by Microscribe (mini Coordinates Measuring Machine) (**Figure 10**). The maximum difference of the computed 3D distances between the bundle technique and the direct measurement technique (Microscribe) was less than 2.1 mm.

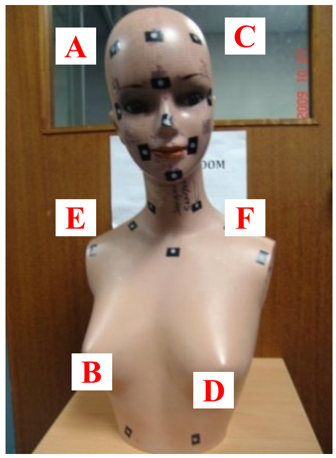

Figure 9. Mannequin Subject

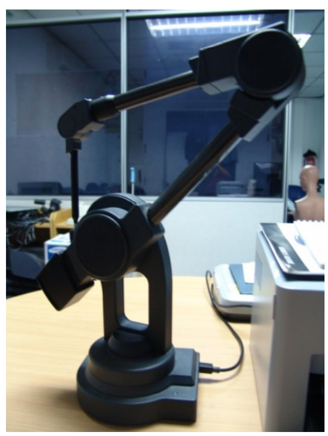

Figure 10. Microscribe

After the video clips were processed, 3D distances between test targets were computed. The average 3D differences (distance discrepancies) and RMS discrepancies between the true values and those computed from the video clips of the five repetitions are listed in **Table 2**. A graphic plot of the data of **Table 2** is shown in **Figure 11**. The 3D distance discrepancy and the object distance were chosen for the y-axis and the x-axis, respectively. Based on **Table 2** and **Figure 11**, the average 3D distance discrepancy increases in size as the object distance increases from 3 to 5 m. The trend is expected to continue because the resolution of the image pixel becomes bigger/lower as the object distance gets longer. Thus, the error of measurement might be eventually become larger.

Table 2. Average discrepancy between the measured and the true 3D distance versus the object distance

Test Points ***A-B***	Object Distance (m)		
	3.0	***4.0***	***5.0***
Difference in 3D distance (mm)	6.9	16.5	21.4
RMS of four sets (mm)	1.5	2.1	3.5
Test Points ***C-D***			
Difference in 3D distance (mm)	9.7	21.4	29.1
RMS of four sets (mm)	2.1	2.4	4.2
Test Points ***E-F***			
Difference in 3D distance (mm)	7.9	17.9	25.4
RMS of four sets (mm)	1.7	2.0	3.7

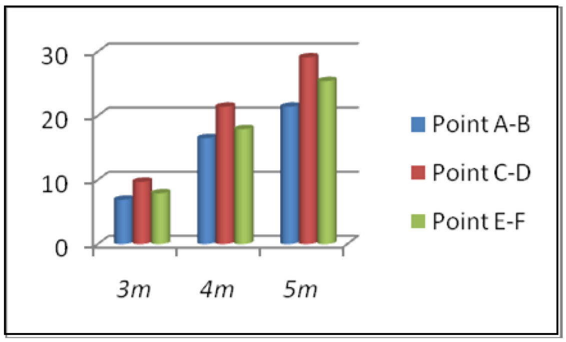

Figure 11. A plot of the average 3D distance discrepancy of Table 2

4.3 Case Study

A set of eight test points (craniofacial/ anthropometry landmark) (**Figure 12**) selected from suspect was used as reference in the experiment. The 3D distances between these points were used as true values in the case study. The relative positions of these points were determined by direct measurement technique using Microscribe coordinates measuring machine. Subsequently, the video clips were processed and analyzed according to the procedure discussed in the system evaluation.

Figure 12. Eight test point based on craniofacial/anthropometry landmark

Table 3 shows the average (mean) 3D differences and standard deviation discrepancies between the true values (measured by microscribe) and those computed from the video clips of the four repetitions. A graphic plot of the data of **Table 3** is shown in **Figure 13**. The 3D distance discrepancy and the test point were chosen for the y-axis and the x-axis, respectively.

Table 3. Average (mean)/std. dev. discrepancy between the measured and true values versus anthropometry landmark

Mean	*n*	*Std. Dev*	
4.31198	4	1.727385	ex-ex
6.55230	4	2.077377	tr-pg
6.92440	4	2.328261	h1-h2
11.14498	4	1.665015	t1-t2
7.23341	16	3.098982	Total

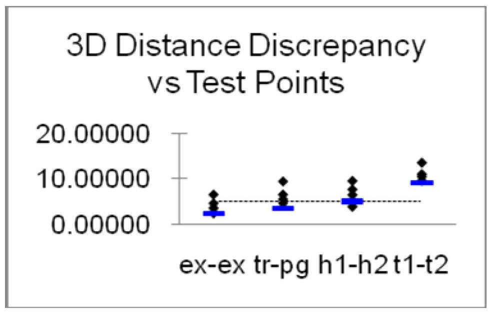

Figure 13: A plot of the average 3D distance discrepancy of Table 3 against the craniofacial/anthropometry landmark

Based on **Figure 13**, most of the mean 3D difference values for all test point are lower than or close to the 7.0 mm, except for 't1-t2' test points provides error more than that value. The errors are caused by a few factors such as; manual digitizing, difficulty of identifying craniofacial landmarks and other factors. The standard deviation discrepancies between the true values and those computed from the video clips of the four repetitions was consistently in a range of 1.6 to 2.3 mm.

5.0 CONCLUSIONS

A customized stereo-imaging system was evaluated for application in human motion research. The system is easy to operate and is efficient for stereo-image capture: the amount of stereo overlap is always the same. With the stereoscopic video footage captured by the customized stereo- imaging system, the measurement of suspect morphology can be carried out based on human body anthropometric concepts. The measurement can provide high accuracy forensic mapping for suspect identification.

The results of the research show that the system is capable of obtaining spatial data with 10 mm accuracy using a single video camera. As the resolution of an off-the-shelf video camera is high, the video footage can achieve 30 mm accuracy under 5 m object distance. This accuracy specification is required for suspect identification from surveillance video footage based on height measurement method (Chong, 2002). Based on the results, the authors found that there could be few important factors which could reduce the accuracy of anthropometric measurement data. Consequently these factors were studied in more detail and the results show that accuracy could be degraded considerably.

Nevertheless, the stereo-video footage can be used for other applications as well, such as security surveillance, forensic mapping, medical research and field performance study. The suitability of the customized system for these applications will be assessed in future research. More research is needed to improve the technique in order to achieve optimum photogrammetric measurement.

ACKNOWLEDGEMENTS

The authors acknowledge financial assistance for this research from Ministry of Science, Technology and Innovation (MOSTI) for Scholarship of Postgraduate Fellowship Scheme (PGD) and e-science fund vote 79353

REFERENCES

3-D ImageTek Corp., 1998. *3D Video Encoder*. http://www.3dimagetek.com/encoder.htm [Accessed: 21th August 2008].

Atkinson, K.B. (Ed), 1996. *Close Range Photogrammetry And Machine Vision*. Whittles Publishing, Caithness, Scotland. 371 pages.

Beyer, H.A., 1992. *Geometric And Radiometric Analysis Of CCD-Camera Based Photogrammetric Close-Range System*. Institute for Geodesy and Photogrammetry, ETH Honggerberg, Zurich. 186 pages.

Burmeister, O. P.,Litza, M. G., Nitschke, M. and Hofmann, U. G., 2005. *Synchronous stereo-Video And Biosignal Recording A Basic Setup For Human Computer Interface Applications*. 2nd International IEEE EMBS Conference on Neural Engineering. 562 pages: 501–505.

Chong, A. K. And Schneider, K., 2001. *Two-Medium Photogrammetry For Bottlenose Dolphin Studies*. Photogrammetric Engineering & Remote Sensing, 67(5): 621–628.

Chong, A.K., 2002. *A Rigorous Technique For Forensic Measurement Of Surveillance Video Footage*. Photogrammetric Engineering & Remote Sensing. Vol. 68, No.7, July 2002.

Chong, A. K., 2004. *Optimising The Accuracy Of A Low-cost Photogrammetric Motion Study Ssystem*. Photogrammetric Record, 19(108): 296–310.

Chong, A.K. and Stratford, P., 2006. *Near Infrared Photography For Craniofacial Anthropometric Landmark Measurement*. The Photogrammetric Records. March 2006, 21(113).

Chong, A.K., 2007. *An Inexpensive Stereo-Image Capture Tool For Motion Study*. The Photogrammetric Records. September 2007, 22(119): 226-237.

Fraser, C.S. and Edmundson, K.L., 1996. *The Metric Impact Of Reduction Optics In Digital Cameras*. Photogrammetric Record. 15(87): 437-446.

Fryer, J.G., 1989. *Camera Calibration In Non-Topographic Photogrammetry. In: Non-Topographic Photogrammetry*. (Ed. H.M. Karara). 2nd Edition. American Society for Photogrammetry. 445 pages: pp. 59-68.

Lee, D. H., Kweon, I. S. and Cipolla, R., 1999. *A Biprism-stereo Camera System*. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 1: 82–87.

Lee, J.-Y., Nam, S.-J., Lee, J.-H., Park, C.-S., Joo, Y.-S., Petrov, D. A., Kim, Y.-G. and Lee, Y.-B., 2004. *Development Of A Reliable And Practical HD Stereoscopic Camera System*. SPIE, 5291: 188–195.

Montgomery, D. J., Jones, C. K., Stewart, J. N. and Smith, A., 2002. *Stereoscopic Camera Design*. 4660: 26–37. SPIE.

Peterson, A.E., Durdle, N.G., Raso, V.J. and Hili, D.L., 1993. *Calibration Of Video Cameras For Scoliosis Mapping*. Geomatica, 47(1): 29-38.

Stereo3d, 2007. *The unofficial Nu-View & Nu-Shades page.* www.**stereo**3d.com/**nuview**.htm [Accessed: 27th August 2006].

Tzovaras, D., Grammalidis, N. and Strintzis, M. G., 1997. *Object-based Coding Of Stereo Image Sequences Using Joint 3-D Motion/disparity Compensation*. IEEE Transactions on Circuits and Systems for Video Technology. 7(2): 312–327.

Wolf, P.R., and B.A. Dewitt., 2000. *Elements Of Photogrammetry With Applications In GIS*. Third Edition, McGraw Hill, New York.

Woods, A., 2007. *Stereoscopic (3D) Imaging*. http://www.3d.curtin.edu.au [Accessed: 16th August 2008].

AUTHORS

Mohd Farid Mohd Ariff is a PhD candidate at the Faculty of Geoinformation Science & Engineering, Universiti Teknologi Malaysia. Currently, he is an active member of Photogrammetry and Laser Scanning Research Group. He holds B.Eng. (Hons.) (Geomatic Engineering) degree and M.Sc (Photogrammetry) degree from Universiti

Teknologi Malaysia, Malaysia. His PhD research lies in the area of Photogrammetry for Forensic Applications.

Dr. Zulkepli Majid is a senior lecturer at the Faculty of Geoinformation Science & Engineering, Universiti Teknologi Malaysia. Currently, he is the head of Photogrammetry and Laser Scanning Research Group. He holds B.Sc (Land Surveying) degree, M.Sc (Photogrammetry) degree and a PhD (Medical Photogrammetry) degree from Universiti Teknologi Malaysia, Malaysia. His research interests lie in the areas of Photogrammetry and Laser Scanning for various applications. His current research projects are the application of photogrammetry in forensic and assets data capture using terrestrial laser scanner.

Dr. Albert K. Chong is a senior lecturer at the Faculty of Engineering and Surveying, University of Southern Queensland. He holds B.Sc in Photogrammetry from California State University, Fresno, M.Sc. (Civil Engineering) from the University of Washington and a PhD (Civil Engineering) from the University of Washington. His research interests lie in the areas of close-range photogrammetry and close-range laser scanning. He has authored over 137 articles in both international and national refereed journals, conference proceedings, and workshop in field of engineering, surveying, photogrammetry, laser scanning and GPS.

Dr. Halim Setan is a professor at the Faculty of Geoinformation Science and Engineering, Universiti Teknologi Malaysia. He holds B.Sc. (Hons.) in Surveying and Mapping Sciences from North East London Polytechnic (England), M.Sc. in Geodetic Science from Ohio State University (USA) and Ph.D from City University, London (England). His current research interests focus on precise 3D measurement, deformation monitoring, least squares estimation and 3D modeling.