A REVIEW: DIGITAL STEREO CLOSE RANGE VIDEOGRAMMETRY FOR THREE DIMENSIONAL MEASUREMENTS

Mohd Sharuddin Ibrahim[1], Halim Setan[1] & Albert K. Chong[2]
[1]Medical Imaging Research Group, Fakulti Kejuruteraan Dan Sains Geoinformasi, Universiti Teknologi Malaysia, 81310 Skudai, Johor
[2]School of Surveying, University of Otago, 310 Castle Street, Dunedin, New Zealand
Email: sharuddin@fksg.utm.my, halim@fksg.utm.my & chonga@albers.otago.ac.nz

Key Words: Videogrametry, 3D measurement, Close Range Photogrammetry

ABSTRACT

Usually photogrammetrist used still image for measurement and modeling of object. Videogrametry refers to video images taken using camcorder or movie function on digital still camera. Video movie consists of sequences of images (or frames). If video speed is 25 fps (frame per second) and taken for 1 minute (i.e. 60 seconds), there are 25 frame per second or overall 1500 image. This paper highlights the capabilities of video as a tool for 3D measurement and modeling, as well as still image. Several advantages are discusses in detail. This paper discusses an on-going research to develop a real time video capturing software and procedure for high-accuracy real-time data capture and image analysis. The development also focuses on the optimization, in terms of method, procedure, low-cost and accurate results. The research methodology consists of real time video capturing, image analysis, bundle adjustment, motion detection, motion tracking and 3 D coordinate movement on each frame. With this development, it can be applied to measure moving object such as sport analysis, metrology, inspection, and model the motion of human for medical purposes.

1.0 Introduction

This paper focus on capabilities of video as a tool for 3D measurement and modeling, as well as still image. Several advantages and on-going research to develop a real time video capturing software and procedure for high-accuracy real-time data capture and image analysis are discuses.

Motion capture can also be achieved by image-based methods. They can essentially be split into monocular and multiimage systems. Monocular systems use sequence of image acquired by a single camera. To gain three-dimensional information from 2-D video clips, knowledge of the human motion has to be used. Some systems gain this knowledge by learning from provided sample training data and applying statistical methods to get the 3-D motion (Mahoney 2000, Song et al. 2000, Rosales and Sclaroff 2000). Other systems perform the tracking of defined human body models with constrains by sophisticated filtering processes (Deutscher et al. 2000, Segawa and Totsuka 1999, Cham and Rehg 1999). Multi-image system use sequence of images acquired simultaneously by two or more cameras. Some systems assume a simple 3-D human model which is fitted comparing its projections into the different images to the extracted silhouettes of the moving person (Cheung et al. 2000, Delamarre and Faugeras 1999) or the extracted edges (Gravila et al. 1996). Other systems use image based tracking algorithms to track in 3-D the surface of the

human body (D'Apuzzo et al. 2000) or the different body parts (Ohno and Yamamoto 1999). Mathematical models of the human motion can also be used to track directly in the 3-D data, which can be trajectories of known key points (Iwai et al. 1999) or dense disparity maps (Jojic et al. 1999).

1.1 Close Range Photogrammetry

Still images are commonly used in close range photogrammetry area in order to measure 3-dimensional coordinate. Using advance algorithms such as bundle block adjustment, resection and triangulation, 3D point can be easily determined with high accuracy. Close range photogrammetry technique involves stereo and convergent method. The method is similar to aerial photogrammetry and it uses pair of images that overlaps at 60%. Convergent images consist of several snap shot of images and there should be minimum 4 images and 25 identical points seen in every image. In capturing data via digital device, there are two category i.e. static (still image) and dynamic (video).

1.2 Introduction to Video

To capture and record video, a series of images (sequence) are captured. This task can be easily done by video technology either using only analog or digital mode. Digital video requires advance image processing. Analog video can be used but it must converted to digital form. Digital video can be described in four difference level of detail as shown in figure 1. At the lowest level, it consists of a set of frames (segment of picture). At the next higher level frames are grouped into shots. The term shot refer to a continuous camera recording. Consecutive shorts are aggregated into scenes based on story-telling coherence. All scenes together compose the video. Video comprises both the images and audio tracks. Audio are ignore in this research.

Figure 1 Structural of video component

There are three standard video schemes that are currently in worldwide use - NTSC, PAL, and SECAM. This information is summarized in Table 1 (Tudelft, 2005). This paper used PAL video format to carrier out the task of first step in motion tracking research

Standard	NTSC	PAL	SECAM
Property			
images / second	29.97	25	25
ms / image	33.37	40.0	40.0
lines / image	525	625	625
(horiz./vert.) = aspect ratio	4:3	4:3	4:3
interlace	2:1	2:1	2:1
us / line	63.56	64.00	64.00

Table 1: Standard video parameters

1.3 Basic camcorder capturing video

Generally, camcorders record 25 frames per second (fps), in fact it records 50 individual images per second, intermixing every two consecutive pictures (with half the height) into one frame. In term of video its call fields (not pictures). So two fields (odd and even) are mixed into one frame. This mixing is called interlacing (Figure 2 and 3). The timeline of analog camcorder is usually different. Analog camcorders do not mix the recorded frames. They record frame after frame.

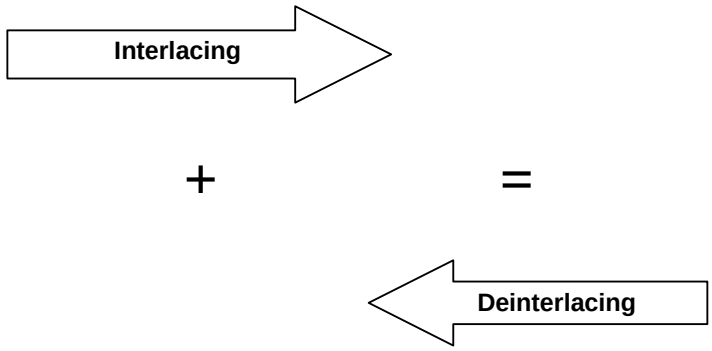

Figure 2 :Record field1 + Record field2 = frame 1

Figure 3 : Basic interlace

1.4 Why interlacing?

Movies with 25 non-interlaced (progressive) frames per second don't look very fluid. Watched a football game with 25 progressive fps would look as if the ball isn't flying fluidly thru the air. With 50 fields per second which are then combined to 25 frames per second this looks much better. Camcorder that record 50 images per second are current technology.

Interlacing is clever way to compress a movie when one cannot use digital compression methods. Interlacing reduces the bandwidth (storage space nowadays) by half, without losing vertical resolution in quiet areas (in motion areas you don't notice very much anyway, because it's moving 50 times per second). So interlacing is a way to display the nonmoving parts with full resolution and the moving parts with half resolution, but fluidly. It's a very clever way to cut bandwidth without sacrificing much quality. But even as technology marches on and camcorders get better, you will have 2 options: To record interlaced (smoother motions) or non-interlaced (higher vertical resolution).

1.5 Why deinterlancing ?

Deinterlacing is opposite of interlacing method, its split frame one into field odd and field even. So the size of vertical is half then the original ones. To keep 25fps for each field has to keep the full quality by duplicated (restored) to original frame. So that, we have a pair of image (frame) can be apply in close range photogrammetry technique.

1.6 Motion tracking in sport science

Sports Motion recognized the need for video motion technology in athletic training (Figure 4) and several years ago began to bring together the various areas of expertise necessary for building effective motion analysis products.

Figure 4 : Sequence of images sport

Motion capture is a vital tool in the study of biomechanics. How people move and react is effected by many different factors as varied as lifestyle, age and environment. Motion capture can aid understanding of human biomechanics; motion analysis can allow researchers to quantify by just how much motion is affected. The used of motion in sport science field can help player to accelerate learning, improve performance and prevent injuries. Digital video cameras and motion analysis software can be used to monitor achievement and standard coaching at all athletic levels. By skillfully capturing an athlete's mechanics at the proper angles several achievement are made:

- Expose mechanical flaws not visible to the naked eye.
- Improve communication between coach and athlete.
- Teach the athlete to properly "internalize" verbal and visual coaching cues.
- Assist the coaching staff in developing appropriate corrective drills.
- Break bad habits that inhibit skill advancement.
- e game.

Sports medicine and performance analyses go hand in hand with analyses aimed at preventing injury. While the protocols are often very similar, the analysis of sports related movements often entails analyzing a variety of highly dynamic movements. The magnitudes for angular velocities and accelerations associated with sports activities are often large and occur while translating through large fields of view.

Motion Analysis provides the tools for the Sports Medicine & Performance professional to perform accurate functional evaluations or analyses for clinical and research oriented purposes.

2.0 Objective and outcome of research :

2.1 Objective:

- To develop gold standard for sports person as a guide to perform and improve skill in sport events via scientific and technological advance video motion tracking technique.
- To improve accuracy of video capturing.
- To develop fast motion tracking by changing target (shape, size, color and lighting) and determine the best algorithms for Improve fast tracking for sport science application.

2.2 Outcome / benefit from research

- Improving our country achievement in sport arena.
- Development also focuses on the optimization, in terms of method, procedure, low-cost and accurate results.
- Can also be applied to measure moving object such as sport analysis, metrology, inspection, and model the motion of human for medical purposes.

3.0 Methodology of research (Flow Chart)

The research involves the developed of a real time video capturing software and procedure for high-accuracy real-time data capture and image analysis. The research methodology consists of real time video capturing, image analysis, bundle adjustment, motion detection, motion tracking. Figure 5 show the flow chart of research approach and figure 6 shown the complete process. Paper focuses on the optimization, in terms of method, procedure, low-cost and accurate results. The research methodology consists of real time video capturing, image analysis, bundle adjustment, motion detection, motion tracking and 3 D coordinate movement on each frame.

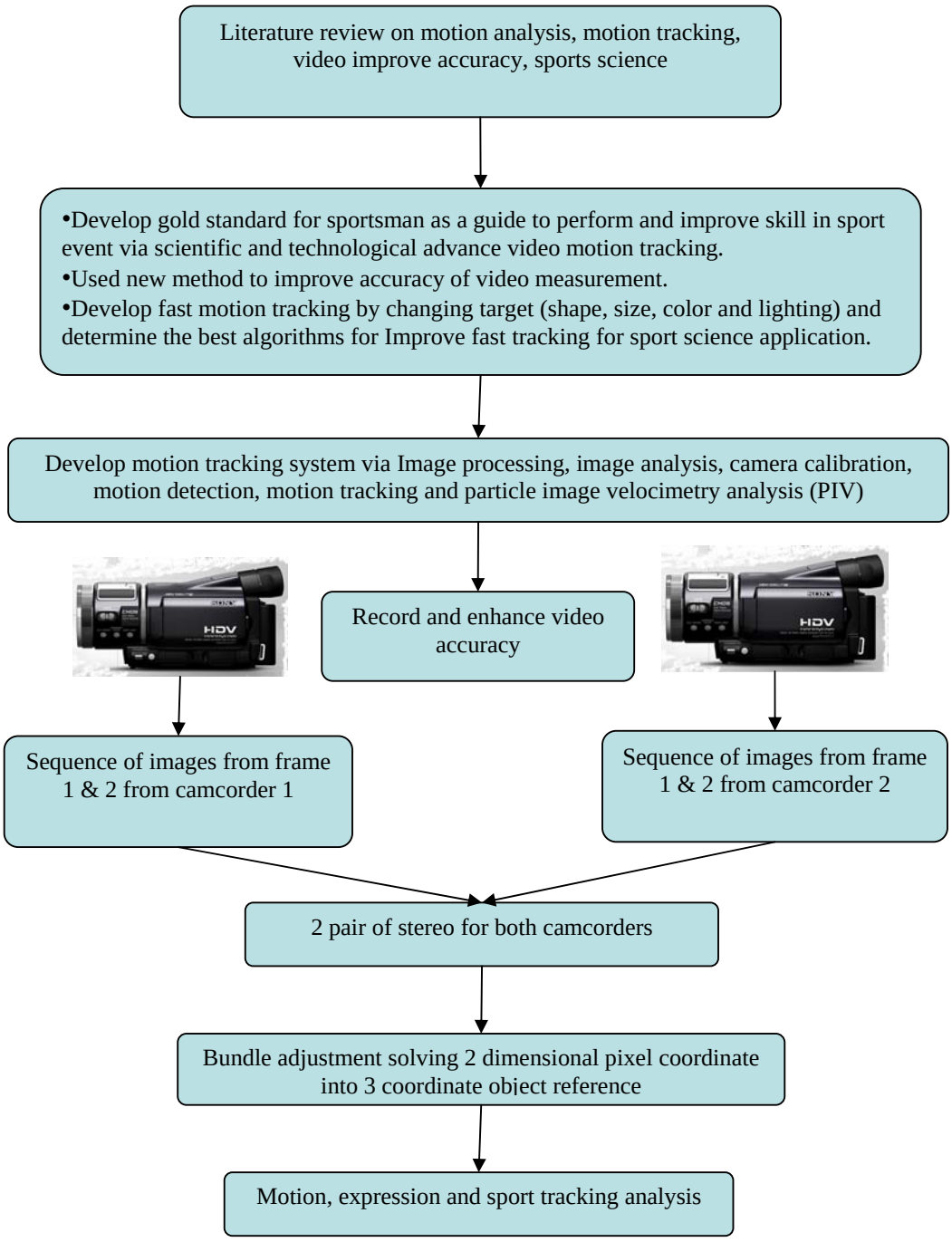

Figure 5 : Flow chart of research approach

Figure 6 : Complete processing

3.1 Instrument involved in this research

With continued research study in close range photogrammetry at the university, we are able to have several imaging tools such as High Definition Video Camcorder from Sony (Figure 7). It is able to record both camcorders in a single snap using lanc port (ACC). In order to develop advance program, several commercial and open source software are used (Table 2).

Table 2 : Hardware and software needed in this research

Hardware	• 2 unit Digital Camcorder with lance controller
Commercial Software	• Visual Studio.Net 2005 Professional Edition • MontiVision Development Kit (SDK)-MontiVision Development Kit • Bundle Software • Australis 6.0 (Close Range Photogrammetry)
Open Source Software	• Intel® Open Source Computer Vision Library (Open Source) • Virtual Dub

3.1.1 Hardware:

Figure 7 : Sony digital High Definition Video Camcorder

3.1.2 Commercial Software :

- **Bundle Software**
  Bundle™ is software package that performs self-calibrating photogrammetric Bundle™ adjustments and terrestrial network adjustments. The primary purpose of a Bundle™ adjustment is to optimally determine the 3-dimensional co-ordinates of points from 2-dimensional image measurements (Figure 9). However, **Bundle™** can also be used to calibrate cameras and adjust small-scale terrestrial networks.

- **Australis 6.0 software**
  Australis software has been developed under supervise of Prof. C.S. Fraser from University of Melbourne, Australia department of Geomatic Engineering. The recent development of the Australis software version 6.0 are using RO (relative orientations). Images are taken using any digital camera. Using Australis software a series of images are than process in order to build up 3 dimensional coordinate (Figure 3). The accuracy is depending on the calibration technique.

- **Interface, image processing & analysis - MontiVision Development Kit (SDK)**
  Hardware Independent Machine Vision Application Design in Industry, Biotechnology, Medicine, etc. Software Development for simultaneous access to Multiple Video Sources, including Capture Cards, USB, FireWire and Network Cameras. Supports Seamless Recording and Playback of Multiple Files, Time Lapse and Motion Detection (Figure 9).

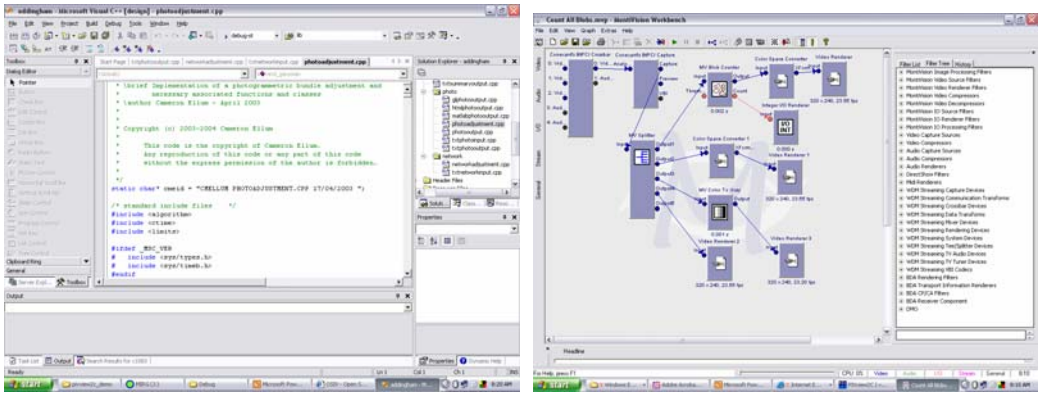

Figure 9 : Bundle™ is software package and Monti Vision SDK

3.2.3 Open Source Software :

There are several open sources

1. Intel® Open Source Computer Vision Library (Open Source).
2. Virtual Dub

4.0 Result and preliminary findings of research

Using low resolution digital camcorder (Panasonic NV-DS30) with 0.8 mega pixel and Australis 6.0 (Figure 10), data captured has been processed and result are shown in Table 3. Calibration to determined camera parameter is done using rotating calibration plate with single camcorder. Video data are split into small frame overall 9 frame (Figure 11) and processing were carry out by deinterlacing the even and old field, then compared with full frame, comparison of the results are shown in Table 4.

Figure 10: Rotating calibration plate

Result show that the differences between full and deinterlace on 25fps (even and old) in term of camera parameter are between 0 to 0.01 mm. We can deinterlace to reduce file size and maintained the accuracy of measurement.

Figure 11 : Panasonic NV-DS30 and processing Australis 6.0

Next, testing involves the incensement of the frame rate of camcorder from 25fps up to 50fps using low cost camcorder. With high definition camcorder Sony HDV we are able to get better result in term of accuracy and motion detection for sport science.

Table 3 : Parameter camcorder

Table 4: Comparison of the camera parameter result between full, even and old field

6.0 Conclusion

Result shown that accuracy of video camcorder and still image are small up to 0.01mm. This happen because of the data of data camcorder are process as a sequence of image. Accuracy of measurement can be increase by using high speed camcorder and high definition camcorder up to mega pixel. Tracking motion with this accuracy will promises the better result on measurement. It also can be applied to measure moving object such as sport analysis, metrology, inspection, and model the motion of human for medical purposes.

7.0 Preferences

1. Cham T.-J. And Rehg J. M., 1999. A Multiple Hypothesis Approach to Figure Tracking. Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado, USA, Vol. 2, pp. 239-245
2. Cheung G. K. M. et al., 2000. A Real Time System for Robust 3D Voxel Reconstruction of Human Motions. Proc. Of IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA, Vol. II, pp.714-720
3. D'Apuzzo N., Plänkers R., Fua P., 2000: Least Squares Matching Tracking Algorithm for Human Body Modeling, International Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B5/1, Amsterdam, The Netherlands, pp.164-171

4. Delamarre Q. and Faugeras O., 1999. 3D Articulated Models and Multi-View Tracking with Silhouettes. Proc. of 7th IEEE International Conference on Computer Vision, Kerkyra, Greece, Vol. I, pp. 716-721
5. Deutscher J. et al., 2000. Articulated Body Motion Capture by Annealed Particle Filtering. Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA, Vol. II, pp. 126-133
6. Gavrila D. M., Davis. L., 1996. 3-D model-based tracking of humans in action: a multi-view approach. Proc. Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, pp. 73-80
7. Iwai Y. et al., 1999. Posture Estimation using Structure and Motion Models. Proc. of 7th IEEE International Conference on Computer Vision, Kerkyra, Greece, Vol. I, pp. 214-219
8. Jojic N. et al., 1999. Tracking Self-Occluding Objects in Dense Disparity Maps. Proc. of 7th IEEE International Conference on Computer Vision, Kerkyra, Greece, Vol. I, pp. 123-130
9. Mahoney D. P., 2000. A New Track for Modeling Human Motion. Computer Graphics World, May 2000, pp. 18-20
10. Ohno H. and Yamamoto M., 1999. Gesture Recognition using Character Recognition Techniques on Two-Dimensional
11. Rosales R. and Sclaroff S., 2000. Inferring Body Pose without Tracking Body Parts. Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA, Vol. II, pp. 721-727
12. Segawa H. and Totsuka T., 1999. Torque-based Recursive Filtering Approach to the Recovery of 3D Articulated Motion from Image Sequences. Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado, USA, Vol. 2, pp. 340-345
13. Song Y. et al., 2000. Towards Detection of Human Motion. Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA, Vol. I, pp.810-817
14. Tudelft 2005, Lecture Note on Image Processing Fundamental.