

# Digital Stereo Close Range Videogrammetry For High Accuracy 3d Measurement Motion Tracking System

Mohd Sharuddin Ibrahim<sup>1</sup>, Halim Setan<sup>1</sup> & Albert K. Chong<sup>2</sup>

<sup>1</sup>Medical Imaging Research Group, Faculty of Geoinformation Science & Engineering, Universiti Teknologi Malaysia (UTM)

<sup>2</sup>School of Surveying, University of Otago, Dunedin, New Zealand

Email: [sharuddin@fksg.utm.my](mailto:sharuddin@fksg.utm.my), [halim@fksg.utm.my](mailto:halim@fksg.utm.my) & [chonga@albers.otago.ac.nz](mailto:chonga@albers.otago.ac.nz)

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

## ABSTRACT

With enhances technology of digital imaging and video resolution up to 1080dpi, research has shown that sequence image from a frame of images can be used to provide high accuracy measurement. Usually photogrammetrist used still image for measurement and modelling of object. Videogrammetry 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. Real time video capturing, image processing, image analysis, motion detection, motion tracking, bundle adjustment, 3D coordinates on each movement and draw a patent of motion are possible. This paper focuses on the on-going study to develop a real time high accuracy motion tracking system. It highlights the capabilities of video as a tool for 3D measurement and modelling, as well as still image. Several advantages are discusses in detail. 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 3D 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

On-going research to develop a real time video capturing software and procedure for high-accuracy real-time data capturing and image analysis comprised of video streaming data, real time data capture, and bundle adjustment and three-dimensional (3D) coordinates. Motion capture can also be achieved by image-based methods. It can essentially be split into monocular and multi-image systems (Mohd Sharuddin et al, 2006).

Monocular systems use sequence of image acquired by a single camera. To gain 3D information from 2D 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 3D motion (Mahoney 2000, Song et al. 2000, Rosales & Sclaroff 2000). Other systems perform the tracking of defined human body models with constraints by sophisticated filtering processes (Deutscher et al. 2000, Segawa & Totsuka 1999, Cham & Rehg 1999).

Multi-image system use sequence of images acquired simultaneously by two or more cameras. Some systems assume a simple 3D human model which is fitted by comparing its projections into the different images to the extracted silhouettes of the moving person (Cheung et al. 2000, Delamarre & Faugeras 1999) or the extracted edges (Gravila et al. 1996). Other systems use image based tracking algorithms to track in 3D the surface of the human body (D'Apuzzo et al. 2000) or the different body parts (Ohno & Yamamoto 1999). Mathematical models of the human motion can also be used to track directly in the 3D 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 vs Videogrammetry

Still images are used in close range photogrammetry either stereo or convergent technique to measure 3D coordinates. Stereo 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 of 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). Dynamic (video) or videogrammetry give more advantages, i.e. tracking discrete point based on time. By using time based measurement, camera image can be synchronized to reduce time error.

### 1.2 Introduction to Video

To record a video, a series of images (sequence) are captured either analog or digital mode. Digital video requires advance image processing. Analog video can be used after being transferred to digital format. 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 shots 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.

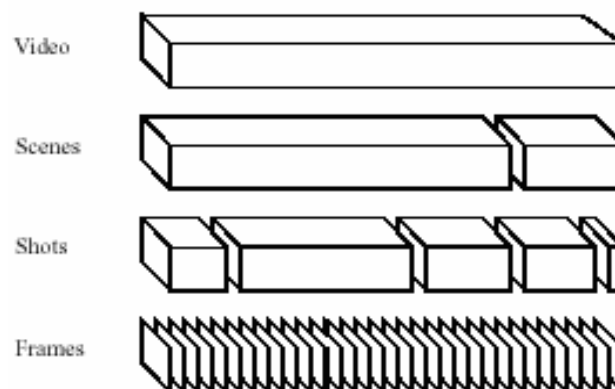


Figure 1: Structural of video component

There are three standard video schemes that are currently used worldwide, i.e. 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.

Property	NTSC	PAL	SECAM
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 terms of video it is called fields (not pictures). Odd and even flied are mixed into one frame called interlacing (Figure 2). The timeline of analog camcorder is usually different. Analog camcorders do not mix the recorded frames. They record frame after frame.

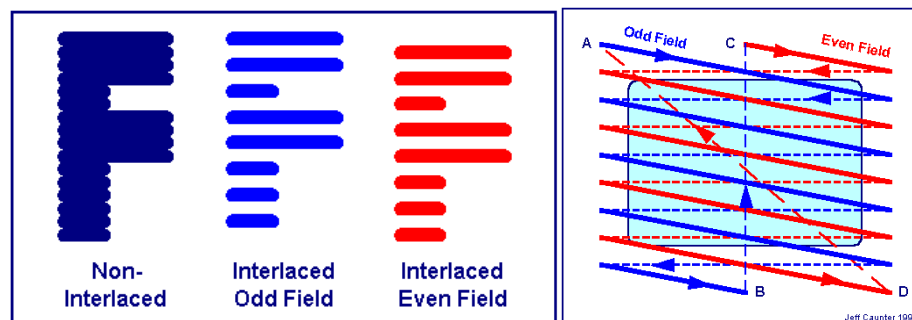


Figure 2: Basic interlace

Movies with 25 non-interlaced (progressive) frames per second do not look very fluid. Watching a football game with 25 progressive fps would look as if the ball is not flying fluidly thru the air. With 50 fields per second which are then combined to 25 frames per second this looks much more better.

Interlacing is a 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 do not notice very much anyway, because it is 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 is a very clever way to cut bandwidth without sacrificing much quality.

Deinterlacing is opposite of interlacing method, i.e. it splits frame one into the odd and even field. So the size of vertical is half than 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) that can be applied in close range photogrammetry technique.

## 1.4 Motion tracking

Motion tracking and motion capture has been used in several area i.e. sport training, metrology inspection, graphic and multimedia. Figure 3 shows examples of biomechanics study. For the easily stage of this on-going research, study are focuses on sport person.

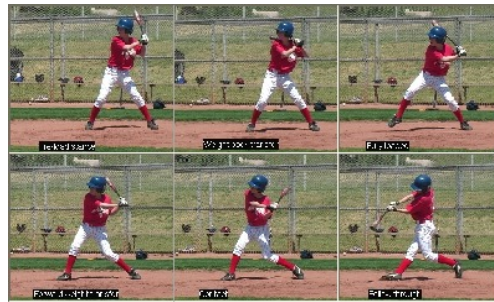


Figure 3: Sequence of images sport

Motion capture is a vital tool in 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.

## 2.0 Objective and outcome of research

### 2.1 Objectives

The objectives of this research are:

- To develop real time or near real time advance video motion tracking technique.
- To improve accuracy of video capturing.
- To develop fast motion tracking for sport science application.

### 2.2 Outcomes from research

The outcomes from this research are the following:

- 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

The research involves the development 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. Flow chart of research approach is shown in Figure 4. The research methodology consists of real time video capturing, image analysis, bundle adjustment, motion detection, motion tracking and 3D coordinate movement on each frame.

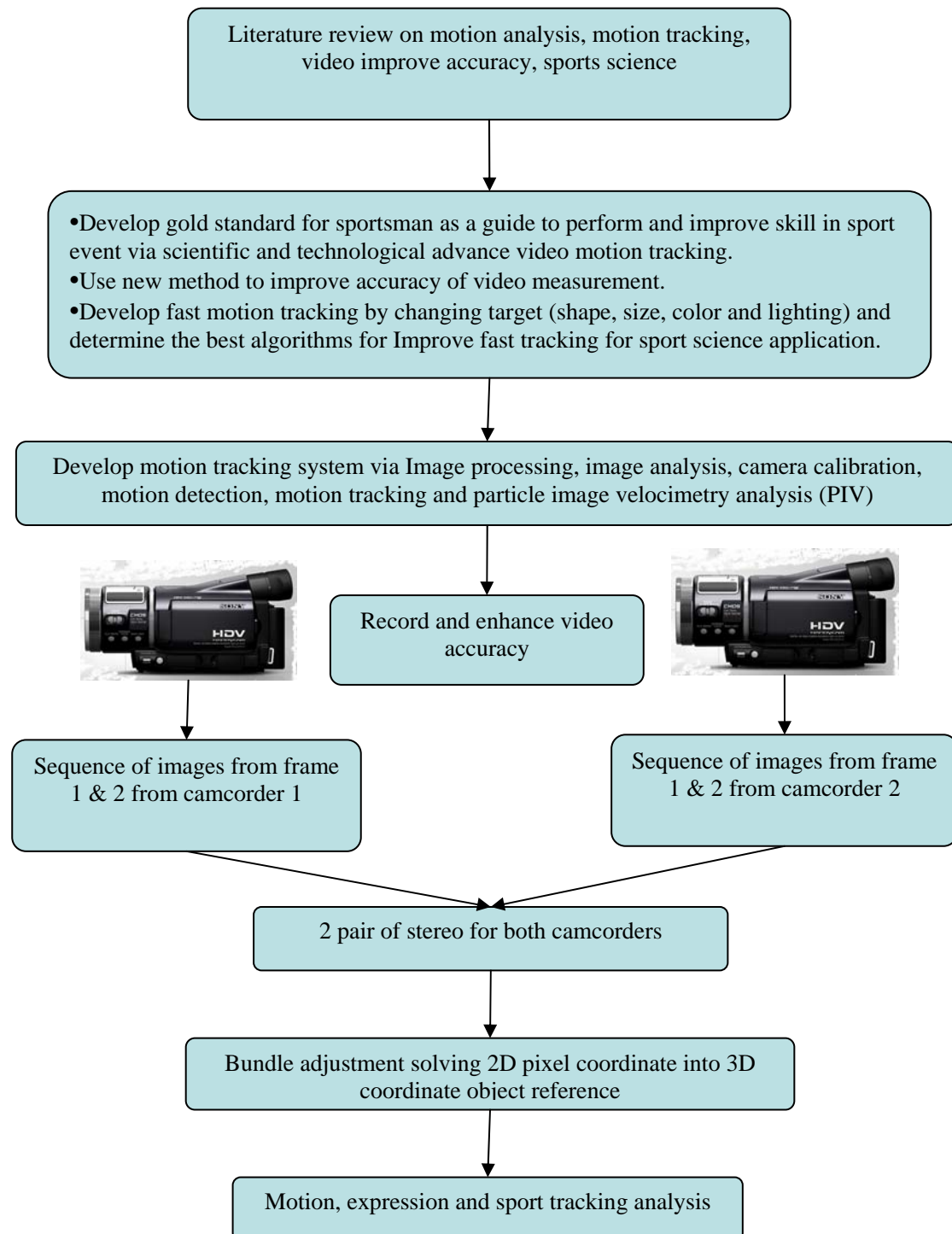


Figure 4: Flow chart of research approach

### 3.1 Research material (hardware and software)

With continued research study in close range photogrammetry at UTM, we are able to have several imaging tools such as High Definition Video Camcorder from Sony (Figure 5). It is able to record both camcorders in a single snap using lance 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	<ul style="list-style-type: none"> <li>• 2 unit Digital Camcorder with lance controller</li> </ul>
Commercial Software	<ul style="list-style-type: none"> <li>• Visual Studio.Net 2005 Professional Edition</li> <li>• MontiVision Development Kit (SDK)-MontiVision Development Kit</li> <li>• Bundle Software</li> <li>• Australis 6.0 (Close Range Photogrammetry)</li> </ul>
Open Source Software	<ul style="list-style-type: none"> <li>• Intel® Open Source Computer Vision Library (Open Source)</li> <li>• Virtual Dub</li> </ul>

#### 3.1.1 Hardware



Figure 5: Sony digital High Definition Video Camcorder

#### 3.1.2 Commercial Software

The following commercial software will be used:

- **Bundle Software**  
**Bundle™** is software package that performs self-calibrating photogrammetric Bundle™ adjustments and terrestrial network adjustments.
- **Australis 6.0 software**  
 Close range photogrammetric software to do bundle adjustment and produce 3D coordinates. The recent development of the Australis software version 6.0 is using RO (relative orientations). Images are taken using any digital camera. The accuracy depends on the calibration technique.
- **Interface, image processing & analysis - MontiVision Development Kit (SDK)**  
 Hardware Independent Machine Vision Application Design in Industry, Biotechnology, Medicine, etc. Support software Development for simultaneous access to Multiple Video Sources, including Capture Cards, USB, FireWire and Network Cameras. Support seamless Recording and Playback of Multiple Files, Time Lapse and Motion Detection.

### 3.1.3 Open Source Software

The following open sources software will be utilized:

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

## 4.0 Result and findings of research

Using low resolution digital camcorder (Panasonic NV-DS30) with 0.8 mega pixel and Australis 6.0 (Figure 7), data captured has been processed and results are shown in Table 3. Calibration to determined camera parameter is done using rotating calibration plate with single camcorder. Video data are splitted into small frame overall 9 frame (Figure 7) and processing were carried out by deinterlacing the even and odd field, and then compared with full frame. The comparison of the results are shown in Table 3.



**Figure 7:** Panasonic NV-DS30, Rotating calibration plate and processing Australis 6.0

Results show that the differences between full and deinterlace on 25fps (even and odd) in term of camera parameter are between 0.00 to 0.01 mm. Deinterlace can be used to reduce file size and maintained the accuracy of measurement (Mohd Sharuddin et al, 2006). Accuracy of image depends on camcorder resolution and fps. 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 shows differences between odd and even field with full frame (normal).

**Table 3:** Comparison of the camera parameter result between full, even and odd field

Camera Variable	Full - Even	Full - Odd
C (mm)	-0.0130	-0.0111
XP (mm)	-0.0030	0.0050
YP (mm)	-0.0026	-0.0059
K1	-0.0010	0.0035
K2	0.0002	-0.0061
K3	0.0000	0.0028
P1	-0.0001	-0.0002
P2	-0.0001	0.0002
B1	0.0002	-0.0009
B2	0.0000	-0.0006

## 5.0 Conclusion

Accuracy of measurement can be increased by using high speed camcorder and high definition camcorder up to mega pixel. Tracking motion with this accuracy will promise better result on measurement. It also can be applied to measure moving objects such as sport analysis, metrology, inspection, and model the motion of human for medical purposes.

## 6.0 Preferences

1. Cham T.-J. & 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., Plänkner 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. & Faugeras O., 1999. 3D Articulated Models and Multi-View Tracking with Silhouettes. Proc. of 7<sup>th</sup> 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. Gavrilu 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. Jovic 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. Mohd. Sharuddin Ibrahim, Halim Setan & Albert K. Chong (2006), "A Review Digital Stereo Close Range Videogrammetry For Three Dimensional Measurement". International Symposium And Exhibition on Geoinformation 2006, 19-21 September 2006, Sheraton Subang Hotel & Towers, Kuala Lumpur.
11. Ohno H. & Yamamoto M., 1999. Gesture Recognition using Character Recognition Techniques on Two-Dimensional.
12. Rosales R. & 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.
13. Segawa H. & 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.

14. 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.
15. Tudelft 2005, Lecture Note on Image Processing Fundamental.