THE EFFECT OF MULTIMODAL DATA SOURCES IN PARALLEL ON THE ACCURACY AND RELIABILITY OF OPTICAL MOTION CAPTURE OF HUMAN SUBJECTS

by

NICHOLAS ROACH

A Thesis Submitted to the
University of Birmingham
for the degree of

DOCTOR OF PHILOSOPHY

School of Electronic, Electrical and Computer Engineering
College of Engineering and Physical Sciences
University of Birmingham
June 2011

ABSTRACT

Optical motion capture is a powerful tool for human motion analysis. However, an intrinsic limitation of this method is its dependence upon line of sight, which when occluded, prevents measurement of position, resulting in data loss and failure of trajectory tracking.

The thesis aims to investigate whether sources of inertial measurement can be used in tandem with passive optical motion capture data, to obtain improvement beyond conventional measures in tracking and occlusion recovery of human motion data.

A novel method is developed, where a conventional optical motion capture system is augmented with miniature inertial sensors. These two measurement modalities are integrated via a forward kinematic and sensor model, which facilitates the reconstruction and tracking of occluded marker trajectories.

The method is validated using “gold standard” trajectory data obtained experimentally from a human subject. By synthetically degrading this data using a model of line of sight, the reconstructor is tested over a range of conditions conducive to occlusion.

Following testing of trajectory data with up to 30% loss to occlusion, the method is shown to be capable of tracking marker position with 100% reliability. Additionally, trajectory data is recovered with useful accuracy (average ~8mm), demonstrating an improvement by a factor of ~10 over conventional interpolation methods. A restriction of the method is its reliance upon future trajectory data for reconstruction of occlusion, which prohibits real-time applications.

To my darlings Vicky and Marina.

For all the love, support and tea.

Without you this would have never been possible.

I love you both with all my heart.

ACKNOWLEDGEMENTS

The author would like to acknowledge the contribution of the following people and parties to this thesis.

Prof. Martin Russell

Prof. Alan Wing

Dr. Satoshi Endo

The School of Psychology at the University of Birmingham

The Kai Ming Association for Tai Chi Chuan, West Midlands, UK.

Equipment used in this project was provided via funding from the BBSRC.

CONTENTS

INDEX OF FIGURES

INDEX OF TABLES

ABREVIATIONS

The following abbreviations are used in this thesis:

3R	–	3-Axis Rotation
AIM	–	Automatic Identification of Markers
ASCII	–	American Standard Code for Information Interchange
ASIC	–	Application Specific Integrated Circuit
CGI	–	Computer Generated Imagery
CoR	–	Centre of Rotation
DLT	–	Direct Linear Transform
DoF	–	Degrees of Freedom
FoV	–	Field of View
GPIO	–	General Purpose Input Output
GPS	–	Global Positioning System
GUI	–	Graphical User Interface
IDE	–	Integrated Development Environment
IMU	–	Inertial Measurement Unit
LAN	–	Local Area Network
LED	–	Light Emitting Diode
LoS	–	Line of Sight
MCU	–	Micro(C)ontroller Unit
MEMS	–	Micro Electro-Mechanical Systems
MSSP	–	Master Slave Serial Port
OD	–	Outer Diameter
PC	–	Personal Computer
PCA	–	Principle Component Analysis
PCB	–	Printed Circuit Board
PDF	–	Probability Density Function
PID	–	Proportional-Integral-Derivative
QTM	–	Qualisys Track Manager
RAM	–	Random Access Memory
RISC	–	Reduced Instruction Set
RMS	–	Root Mean Square
RoM	–	Range of Motion
RS-232	–	Recommended Standard 232 (A telecommunications standard)
SD	–	Standard Deviation
SPI	–	Serial Peripheral Interface
TCC	–	Tai Chi Chuan
TCP/IP	–	Transmission Control Protocol/Internet Protocol
TTL	–	Transistor-Transistor Logic
VCL	–	Visual Component Library

CONVENTIONS

Within the body of the text the following conventions are used:

- Cross references between chapters and subsections are signified by brackets ().

- References are signified by square brackets [].

- Figures and tables are numbered within chapters in the form:

  Chapter number.Figure number.

- Matrix, vector and scalar quantities are typically referred to by two letter identifiers. These are distinguished from quantities with subscripted indices by the placement of letters above the baseline. For example, $\mathbf{a_b}$ is a quantity, however, $\mathbf{a}_b$ is the b^{th} index of **a**.

1. INTRODUCTION

1.1 Optical motion capture

The technique of marker based optical motion capture represents a powerful tool for human motion analysis. Typically, it provides a means by which the 3D positions of a collection of identifying points (markers) attached to the human body can be measured with high spatial and temporal precision. Through analysis of the trajectories of these markers it is possible to make inferences regarding the posture of the body [1], the configuration of the musculoskeletal system [2] and the forces produced during movement [3]. The versatility of this technique has led to its popularity over a diverse range of disciplines, for example, the assessment of pathologies of gait [4], the ergonomic design of sports equipment [5] and computer animation [6]. This has given rise to a considerable market for motion capture technology and the development of a plethora of commercial systems [7, 8, 9]. Although a variety of alternative methods of motion measurement exist, marker based optical motion capture dominates the field of human motion capture [10, 11].

A typical optical motion capture system consists of a set of cameras whose pose (position and orientation) about a fixed measurement volume are known. Through the automated detection and triangulation of points observed by these cameras, it is possible to estimate the 3D position of markers within the measurement volume [12]. These positions are associated with locations upon a subject's body and tracked over sequential frames of camera data. By analysing the relationships between the resulting trajectories it is possible to derive the pose of the body in space, perform anthropometric measurements and estimate dynamic properties such as acceleration and velocity.

1.2 Occlusion

A fundamental limitation of optical motion capture is its dependence upon line of sight (LoS) between each camera and marker. Under practical measurement conditions, LoS can be occluded though obstruction by another object in the measurement volume or by the marker leaving the field of view of the cameras. This results in the marker's position being unavailable to the measurement system and a gap in the trajectory data for the duration of the occlusion event. In addition to loss of data, the interruption of trajectory continuity has a disruptive effect on the process of 'tracking', where measured trajectories are associated with locations of markers upon a subject's body. Specifically, it can be difficult to associate a marker with its previous trajectory upon reacquisition of LoS. This results in the division of marker trajectories into multiple fragments whose association with the measurement subject is undefined. This trajectory fragmentation is typically corrected by manual re-association of trajectories by a human operator or by basic trajectory prediction algorithms [7, 8]. The first of these techniques is intensely laborious and results in inefficient processing of marker data. The second is strictly limited to occlusions of short duration, and if applied incorrectly can result in misassociation of trajectories [10]. In cases where trajectory continuity can be restored, a gap for the duration of the occlusion will still be present. The conventional method by which this is addressed is to synthesise a visually acceptable continuation of the trajectory by interpolation. However, the resultant data often bares little correspondence to actual marker position and is therefore considered unacceptable in many applications [13].

In the majority of cases, occurrence of occlusion can be reduced by the addition of cameras to a measurement system or the repositioning of existing cameras to improve coverage of the measurement subject [14]. However, the addition of cameras is often financially and technically prohibitive and the possibilities of repositioning can be limited by the measurement environment. Furthermore, in many cases the body of the measurement subject

is itself the cause of occlusions. Under these circumstances, certain postures will result in occlusion of markers regardless of the specification of the camera system (for example, when markers are completely covered). Therefore, even under ideal measurement conditions some degree of occlusion will occur, making data loss and tracking failure unavoidable.

Due to the ubiquity of this problem in applications of optical motion capture and the problems with conventional solutions, a reliable means of addressing occlusion is of interest both commercially and academically. Approaches which have been successfully applied to maintain trajectory continuity and estimate marker position during occlusion events have involved Principle Component Analysis (PCA) [13] and kinematic modelling [10, 15]. Although these methods have been found to considerably improve quality of tracking when compared to standard methods, none offer reliable trajectory recovery over occlusion events of extended duration. In all cases this is due to marker position being based upon parameters inferred from previous observations of marker movement, geometric constraints or statistical relationships. As these models only accurately predict the marker position in proximity to the point of occlusion, or under conditions of appropriate constraint, their ability to predict occluded marker location is limited both spatially and temporally. Therefore, although these methods can greatly improve occlusion resistance of a system, they cannot be considered to offer complete immunity or reliable accuracy.

1.3 Inertial motion capture

An alternative method by which body movements can be measured uses a network of Inertial Measurement Units (IMUs) distributed over the body [16, 17, 18]. Each IMU typically integrates an accelerometer, gyroscope and compass whose outputs are combined by an appropriate recursive filter to give an estimate of absolute orientation [19, 20]. By using this orientation data to drive a hierarchical model of the body it is possible to estimate the pose of each body segment by forward kinematics. This method provides several advantages over optical motion capture, such as an effectively unlimited measurement volume, no data loss by occlusion and substantially reduced equipment costs [19, 21]. However, this technology presents three major limitations. Firstly, for accurate representation of the body, the kinematic model must be tailored to the anatomical details of each measurement subject. As this is usually achieved by manual measurements of limb lengths the accuracy is typically poor [18]. This problem is compounded by the recursive process of forward kinematic pose estimation by which errors in IMU orientation estimation and anatomical measurements accumulate throughout the hierarchy of the model. The second limitation is the need for an initial 'root' position on which forward kinematic pose estimation is based. As this position cannot be obtained directly from IMU measurements it is typically provided by estimation of contact points of the body with the floor. This method gives poor estimation of translational movements and requires that the body be in constant contact with the ground [18]. Finally, the accuracy of the IMU orientation estimate depends upon the long term stability of magnetic north and the measured acceleration due to gravity. Therefore, strong perturbations of these quantities (due to ferromagnetic metals or prolonged accelerations) can introduce large errors in orientation [19]. Due to these limitations inertial motion capture presents an inherently inaccurate means of human motion measurement with limited usefulness in practical applications [21].

1.4 Aims of the thesis

1.4.1 Research proposal

The thesis aims to investigate whether sources of inertial measurement can be used in tandem with passive optical motion capture data to obtain improvement beyond conventional measures in tracking and occlusion recovery of human motion data. It was considered that the problem of occlusion in optical motion capture could be addressed by the development of a method using multiple sensing modalities in combination. It is proposed that a mutually supporting system could be realised by augmenting passive marker based optical motion capture with inertial sensing. By exploiting the strengths of each measurement modality it may be possible to counter the weaknesses of the other. To test this hypothesis, a functional kinematic body model is proposed as a means of relating optical position measurements and the inertial orientation estimates. Through this relationship, orientation estimates could be used to extrapolate the trajectories of occluded markers to their points of reacquisition by the optical system. It is expected that these trajectories will bear close resemblance to the missing optical data, presenting a means by which trajectory recovery may be conducted and continuity of tracking maintained. Furthermore, it is intended that the partial trajectory data may provide a reference frame for the IMU orientation estimates and correct for sensor drift. Finally, through estimation of joint centres from marker trajectories it is possible that the kinematic model could be calibrated to individual subjects. This would minimise error in forward kinematic estimation of pose and improve accuracy of trajectory reconstruction.

Due to the irreversible effects of occlusion on trajectory data the performance of the proposed occlusion reconstruction technique would not be verifiable if applied to real occluded data. It is therefore intended that an occlusion simulation method be devised, by which gold standard trajectory data may be artificially degraded in a manner resembling the effects of occlusion.

This would provide a means by which the positional accuracy of reconstructed trajectory data could be evaluated.

It is assumed that full body inertial measurement data can be provided via an inertial sensor model applied to the optical trajectory data. This model would be used to provide simulated gyroscopic and accelerometric data for actuation of the kinematic body model. For the purpose of validation of this model, a single physical IMU will be included in the experimental setup against which this simulation will be compared.

1.4.2 Research aims

From the proposal above, the following research aims were derived. These are divided in to the primary aim, which represents the overall goal of the project, and the secondary aims, which are prerequisites for the development of the primary aim.

Primary Aim

- The development and demonstration of a novel method by which optical and inertial measurements can be used in tandem to give an improvement over conventional methods of tracking performance and occlusion recovery.

Secondary Aims

- The development of a kinematic body model and associated calibration techniques to achieve pose reconstruction via orientation estimates obtained from inertial measurements.

- The development of techniques for simulated degradation of gold standard trajectory data to validate the occlusion reconstruction method.

- The development of an inertial sensor model to simulate accelerometer and gyroscope measurements from gold standard trajectory data.

2. OPTICAL AND INERTIAL MOTION CAPTURE

2.1 Optical Motion Capture

2.1.1 3D reconstruction of from camera data

Optical motion capture facilitates the representation of movement via the 3D trajectories of physical markers attached to the surface of a human subject's body [1]. Marker position in 3D space is typically reconstructed from multiple 2D observations made by a calibrated array of cameras.

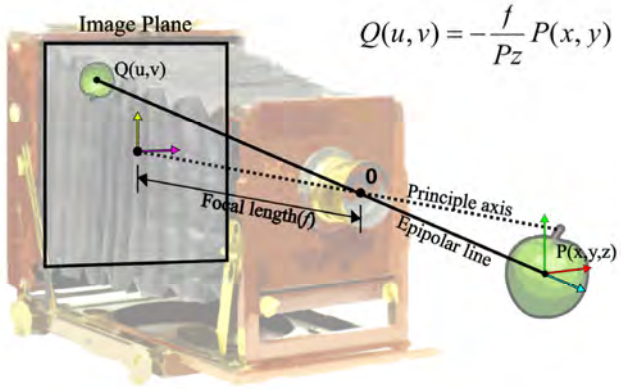


Figure 2.1 – Pinhole Camera Model

The reconstruction of 3D position is analogous to the resolution of depth by the human eyes, whereby the angular disparity between the images of an object on each retina is inversely proportional to its distance from the viewer. In the case of optical motion capture, the disparity between imaginary "epipolar" lines resolves the distance of a marker from an observing camera.

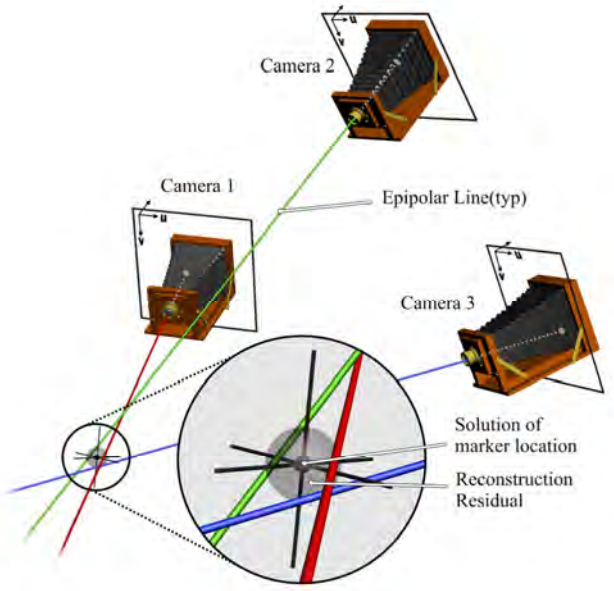


Figure 2.2 – Epipolar reconstruction with multiple cameras

From the example of the "pinhole" camera model shown in figure 2.1, an epipolar line is constructed by projecting

a line between a camera's principle focus and the point corresponding to a marker in its image plane. Extending this line into the three dimensional space in front of the camera, it will eventually intersect with the observed marker. By combining multiple cameras with overlapping fields of view it is possible to construct a 'measurement volume' where multiple epipolar lines are constructed to a common marker location (Figure 2.2). Theoretically, by solving for the point in space at which these lines intersect, the 3D position of a marker can be reconstructed. However, due to the presence of measurement errors perfect intersection is not possible, and instead the lines will pass each other in proximity to the marker. Therefore, the marker location lies at the point where the distance between these lines is at its minimum. This point is typically obtained through the minimisation of a mean-square error cost function derived from the equation of each epipolar line [22]. An estimate of the convergence of the epipolar lines can be obtained from the mean square error at this minimum. This value is typically termed the 'reconstruction residual' and infers the spatial accuracy of the reconstruction.

2.1.2 Camera calibration

For the reconstruction of marker position it is essential that the orientation and position of each camera's imaging plane and principle focus are known. In practice, this pose information is provided in situ through a process of camera calibration. This utilises a

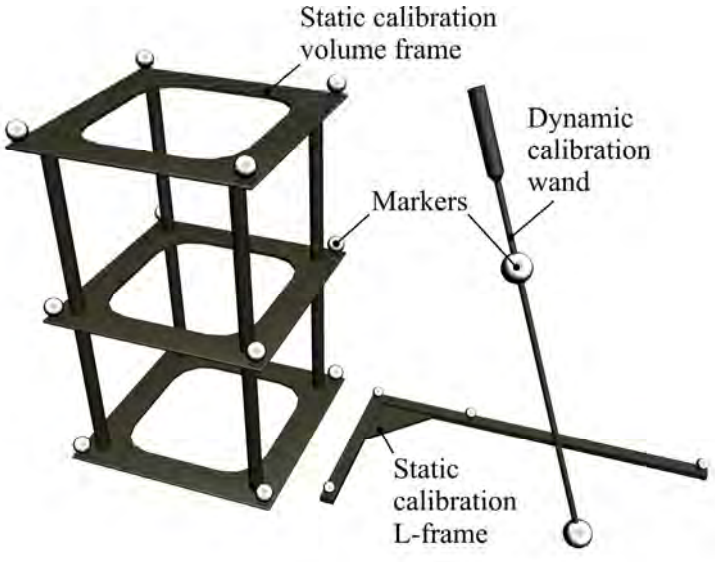


Figure 2.3 – Assorted Calibration objects

"calibration object", which consists of a number of target points arranged in a rigid configuration (Figure 2.3). The pose of a camera can be estimated via the Direct Linear Transform (DLT) [23]. This method provides the transform, which maps the 3D coordinates of the calibration object to their corresponding 2D points in the image plane of an observing camera. For accurate estimation of a camera's pose the DLT requires many calibration points distributed over the entire measurement volume. Older motion capture systems typically involved large calibration frames outfitted with many markers. However, more recent implementations have dispensed with this requirement, relying instead upon the movement of a dynamic calibration object around the capture volume to provide the required calibration points. When used in combination with a static reference object (commonly an 'L-frame'), this allows for convenient calibration of very large measurement volumes [24]. By extension of the DLT it is possible to account for deviations from the pinhole approximation due to the optical distortions introduced by lens geometry. This correction is typically referred to as 'linearisation' and is obtained in a similar manner to camera calibration [25].

2.1.3 Tracking

To measure marker motion over time, sequential frames of measurement data are reconstructed at a fixed frequency (or frame rate). Each of these frames can be considered as a discrete set of points with no explicit association to the points in adjacent frames. Therefore, before this data can be further analysed it is necessary to associate these points with continuous trajectories. The process of establishing these associations is commonly referred to as 'tracking' [26].

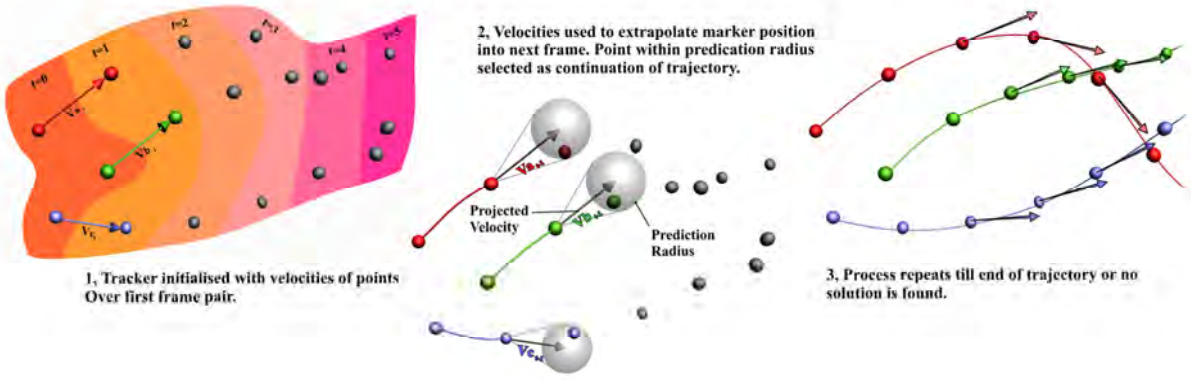


Figure 2.4 – Tracking by dead reckoning

Tracking is conventionally performed by dead-reckoning [27], which predicts the most likely progression of each point from one frame to the next. A typical example of tracking by this method is shown in figure 2.4.

This method operates by extrapolating a point's position via its instantaneous velocity to provide its expected position in future frames. It relies on the assumption that the sample rate of the measurement system is much higher than the bandwidth of any marker movement observed. Therefore, any change in point velocity between one frame and the next will be minimal. Although this method is capable of producing acceptable results [28], its success relies upon markers being continuously visible. This makes the method susceptible to failure when trajectory data is interrupted, for example, by occlusion. Under these circumstances, manual intervention will be required to reinitialise the tracking process on marker reacquisition.

The above example illustrates an extremely simple tracking algorithm, which offers little or no resistance to perturbation by occlusion or positional error. However, in the majority of

state-of-the-art motion capture systems, more sophisticated tracking algorithms are applied. Examples of these include those based upon Hidden Markov models [29] or Kalman filtering [26]. Although such techniques are usually resistant to trajectory interruption, extended gaps in visibility will still result in tracking failure.

2.1.4 Marker systems

For reconstruction from the intersection of epipolar lines it is essential that the image data corresponding to markers is distinguished, or 'segmented' from their background. This is achieved by artificially increasing the contrast of these points so that their brightness exceeds that of other pixels. Following this, the images are segmented by thresholding to identify groups of pixels of maximum brightness. By averaging the pixel coordinates for each of these groups the locations of the marker centres are established (Figure 2.5) [30].

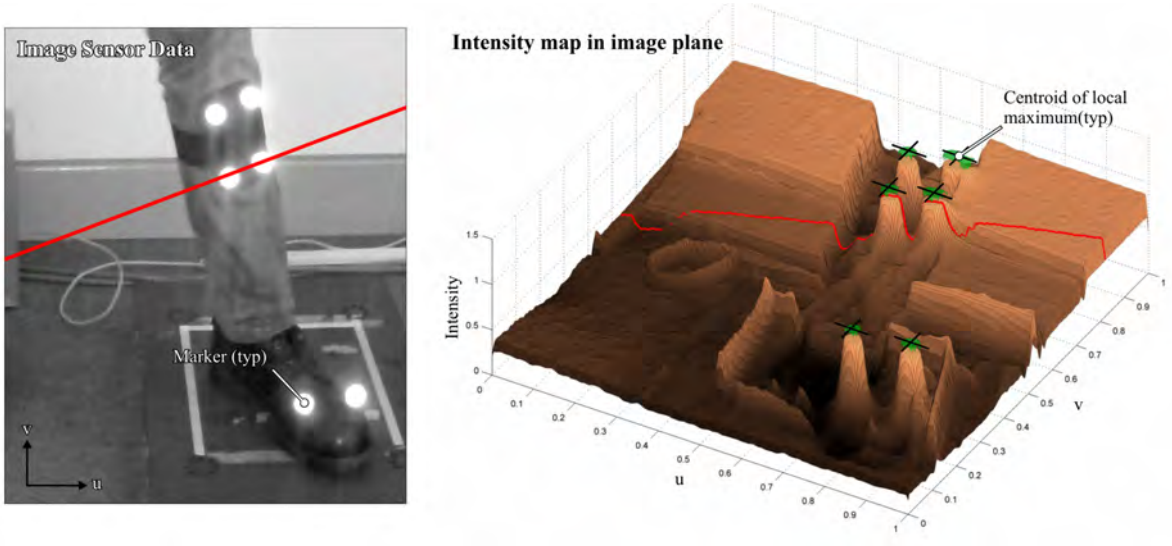


Figure 2.5 – Thresholding of image data

To maximise this difference in contrast, markers in optical motion capture either actively emit light or passively reflect light from a remote source. To reduce susceptibility to ambient

lighting conditions this light is usually monochromatic or infra-red and the cameras are selectively sensitised to its wavelength via optical filters.

In the case of actively emissive markers, high-efficiency light emitting diodes (LEDs) are attached to points of interest upon a subject and pulsed in synchrony with the camera's shutter to provide the desired peak in image brightness. Conversely, for passive reflective types, plastic spheres are covered in highly reflective tape (such as "3M Scotch-lite"), which retro-reflect light from an LED strobe aligned with each camera's optical axis. These two methods, illustrated in figure 2.6, divide optical motion capture systems into those based upon "passive marker" [7, 8] or "active marker" [9, 31] technologies.

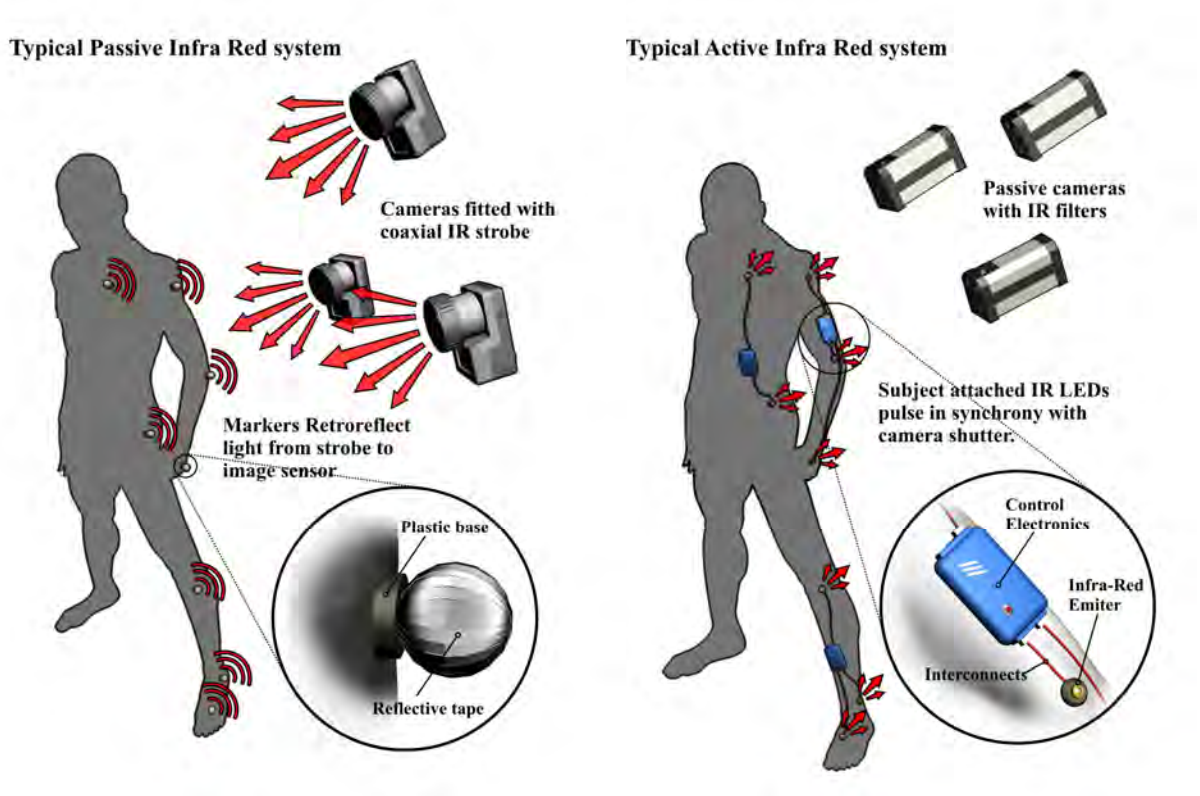


Figure 2.6 – Active and passive marker systems

Although the two technologies operate by the same basic principle, there are several tradeoffs in performance and reliability. For example, passive systems use simpler, more robust markers, which can be discretely installed upon the body and easily replaced. In contrast, active marker systems require control electronics, a power source and cabling to the emitters. Although these emitters are usually smaller than passive markers, the requirement for peripheral hardware discounts any advantage this conveys. Through the coding of strobe frequencies, active markers can be explicitly identified during measurement, something which is not possible with passive markers. Although this effectively eliminates the need for marker tracking, it is associated with a substantial reduction in frame rate [32].

As the two marker technologies operate by identical sensing and reconstruction principles, typical values of positional accuracy are similar. For both passive and active marker based systems the centre of a marker can be estimated with accuracies of less than 1*mm* over a range of several meters [7,8,9]. Active markers are capable of producing considerably higher luminous intensities than their passive equivalents, resulting in greater detection range of markers and resistance to ambient lighting [33]. Absence of a remote strobe system eliminates the influence of artefact reflections from non-marker objects. Advantages of passive systems with respect to reliability include a much wider angle of visibility around a marker due to their spherical construction, as well as independence from restrictions imposed by battery endurance.

2.2 Occlusion in optical motion capture

2.2.1 Causes of occlusion

The technology of optical motion capture is fundamentally dependent upon line of sight (LoS). Should the LoS be interrupted between a marker and camera then an epipolar line between the two cannot be projected. In the event of less than two cameras having LoS to a marker, a solution for epipolar intersection, and consequentially 3D reconstruction will be impossible. This leads to the 'occlusion' of the marker and the loss of trajectory data.

In a practical motion capture setting there are two primary causes of occlusion. Firstly, markers leaving the field of view of the cameras and secondly, LoS being interrupted by objects in the measurement volume. Occurrence of occlusion due to the first case is relatively simple to predict and can be eliminated by ensuring adequate camera coverage for a desired range of motion. However, in the second case the situation is considerably more complex as LoS is typically obstructed by the measurement subject's own body. In these circumstances, occlusion is dependent upon both the pose of the measurement subject's body and the positions of cameras (Figure 2.7). This "self occlusion" is difficult to predict due to the enormous range of poses the human body can assume.

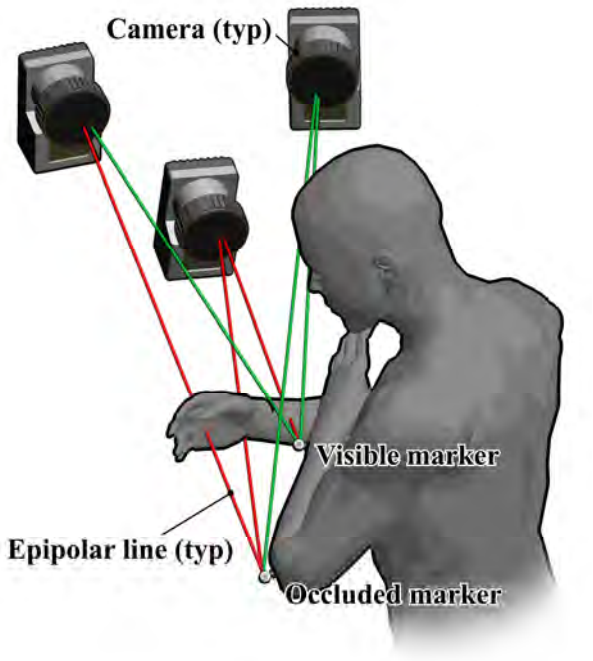


Figure 2.7– Self occlusion of markers

A conventional means of addressing the problem of self occlusion is the addition of cameras to a measurement system [14]. This aims to increase the likelihood of two cameras having a clear LoS ‘around’ an occluding body segment allowing for epipolar reconstruction. However, given the cost of purchasing additional hardware and the effort associated with its installation, this method of occlusion reduction is often impractical. Furthermore, each camera added to a system gives a progressively smaller increase in occlusion resistance. This can be seen in figure 2.8, where the number of occlusion events and relative duration of marker visibility is plotted against the number of observing cameras. The diminished improvement in performance beyond the optimal number of cameras (~7 in this setting) usually makes the addition of further cameras financially uneconomical.

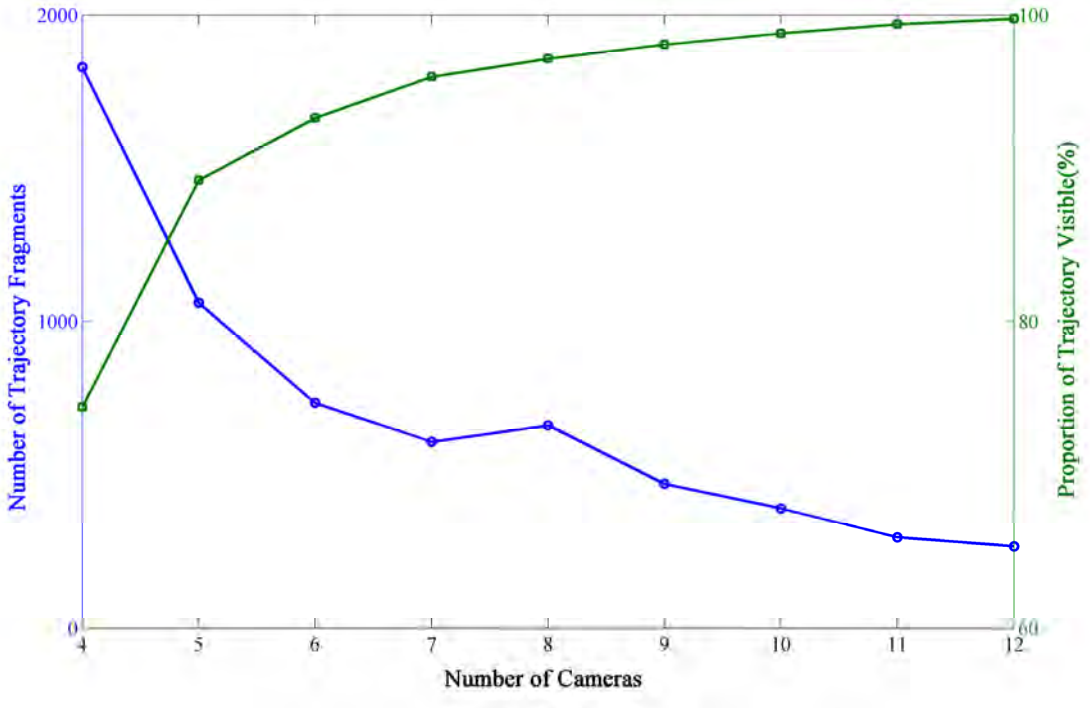


Figure 2.8- Occlusion falloff with increasing cameras

An alternative means of reducing occlusion is to rearrange available cameras to provide increased coverage of regions where occlusion is frequent. However, this method is only of use where body movements are constrained to a specific pose or task, and the resulting improvement in performance is limited. Additionally, as there is no established protocol for performing such optimisations, these adjustments are typically accomplished by trial and error. Although attempts have been made to develop algorithms by which optimal camera configurations can be derived [14], these have provided little benefit beyond confirming the common guidelines for camera placement (for example, that cameras should be distributed evenly around a subject).

2.2.2 Effects of occlusion

In addtion to the loss of measurement data, the effects of occlusion can complicate the tracking of marker trajectories. As was discussed in (2.1.3), extended occlusion events will frequently result in tracking failure, where the tracking algorithm fails to predict the point at which a marker emerges from occlusion. Under these circumstances, this point of emergence will be considered the start of a new unidentified trajectory ‘fragment’, distinct from the original trajectory. A second more severe mode of tracking failure occurs when an erroneous point is chosen for trajectory continuation following an occlusion. This typically occurs when two or more markers in close proximity suddenly change velocity during an occlusion event.

As tracking commonly relies upon the assumption of minimum acceleration for prediction of future marker position [27, 26], the tracking solution provided under these conditions can be incorrect. An example of this failure can be seen in figure 2.9, where two markers exchange identities during an occlusion event. These ‘marker swaps’ typically occur more frequently

when the tracker's sensitivity to velocity discontinuities is lowered to reduce trajectory fragmentation. This presents a trade-off between trajectory fragmentation and marker misidentification.

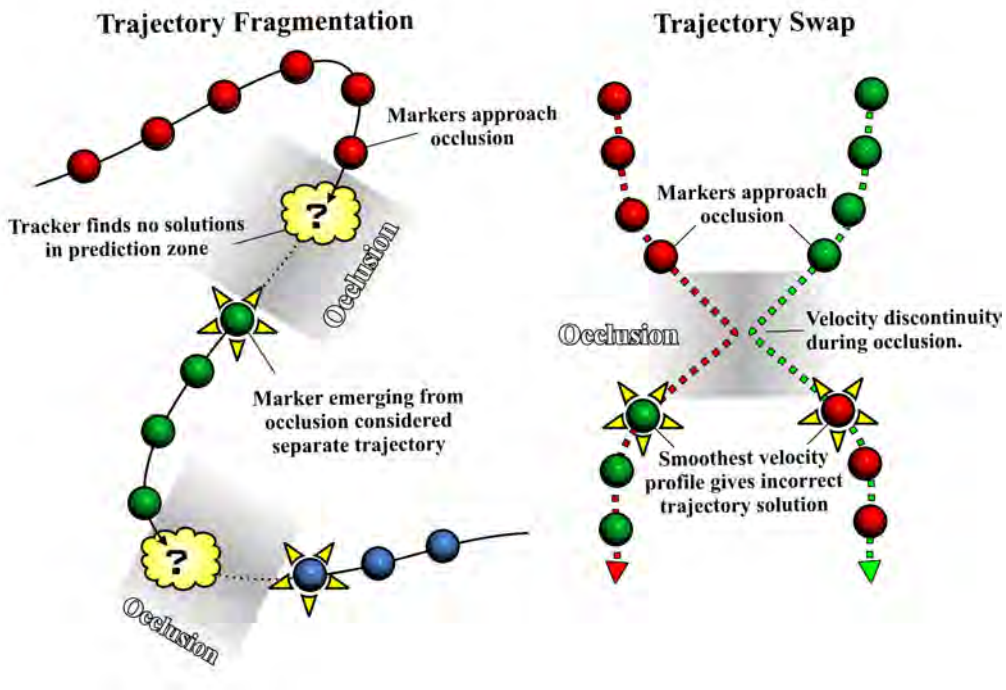


Figure 2.9 - Examples of Tracking Failure

2.2.3 Methods of recovering from occlusion

Recovery from tracking failure due to occlusion is traditionally performed manually during post processing of data [10]. In this process, associations between trajectory fragments are made using their positions on the measurement subject as a cue to their identities. In the case of marker misassignment, it is necessary to locate the point of trajectory confusion, deliberately fragment and then reassign the trajectories. In cases where the confusion is severe, this process can become problematic due to co-dependence of multiple trajectories. Following this process of marker identification, occluded sections of trajectories are typically recovered by spline interpolation [7,8].

For datasets with few markers and minimal occlusion events, manual identification methods can give satisfactory results. However, when applied to larger data sets, such techniques quickly become impractical due to the labour involved. Additionally, reconstruction of occluded trajectory data by interpolation can often result in large errors, which are unacceptable in the majority of applications [13]. In an attempt to address these problems, numerous automated means of trajectory recovery have been developed, relying upon a variety of methods to improve identification of markers and reconstruct occluded trajectories.

The simplest trajectory identification schemes operate by comparing the Euclidean distances between unidentified trajectory fragments with the inter-marker distances of a predefined marker set. By locating trajectory fragments with similar inter-marker distances to those in the predefined set it is possible to infer marker identity [30, 34]. This method is popular in a number of commercial motion capture implementations such as Qualisys's "Automatic Identification of Markers (AIM)" technology [7] and the "Kinematic assisted labelling" used by the 'Vicon' motion capture system [8]. In this capacity, these methods reduce the need for manual intervention during post-processing. However, they provide no capability for the recovery of the trajectory data lost to occlusion.

Although occlusion effectively prevents any inference of the position of an isolated marker, the attachment of markers to a human subject imposes constraints upon the position of a marker with respect to its neighbours. Through knowledge of these constraints and the trajectories of other markers, it is possible to infer the position of a marker during occlusion. This concept has been utilised in a variety of studies, such as the Principal Component

Analysis (PCA) method of Liu et al. [13], which was found to successfully recover trajectory data, though with positional errors of several hundred millimetres.

More practical methods have utilised a kinematic representation of the human body, such as the solution proposed in Herda et al. [10]. This method combined the kinematic constraints provided by estimated joint locations with a dead-reckoning prediction algorithm to estimate the positions of occluded markers and their points of emergence. Although the method was shown to give ~99% recovery of marker trajectories, it relies upon a sufficient number of markers being visible to constrain the model. Also, as the technique was applied to a comparatively small sample of test data and no figures were presented for positional accuracy, it is difficult to assess its true effectiveness.

An alternative method of kinematic trajectory recovery was presented in Todorov [15], where the kinematic interpolation of marker positions was computed via an extended Kalman filter. This augmented the kinematic approximation with a model of the dynamics of body motion, facilitating prediction of occluded marker positions in the absence of constraining trajectory data. Although this method was shown to recover trajectories with up to 50% occlusion loss with high accuracy (~10*mm*), no analysis of the effects of increasing duration of occlusion upon this accuracy was provided. Additionally, the method made no attempt to address the problems of trajectory fragmentation due to tracking failure.

Although these studies present several means by which occlusion recovery can theoretically be addressed, none have comprehensively demonstrated and tested a practically viable solution to all the problems presented by occlusion. Additionally, all these methods rely upon

redundant constraints provided by non-occluded markers or statistical approximations of marker behaviour. Therefore, in conditions where occlusion events of extended duration occur frequently, high levels of positional error and tracking failure are highly likely.

2.3 Inertial motion capture

2.3.1 MEMS Devices

Inertial sensors, such as accelerometers and gyroscopes, are devices capable of measuring relative changes in velocity or orientation with reference to a local coordinate system. Although these sensors are conventionally large mechanical devices, through adaptation of technology originally developed for the construction of integrated circuits, their scale and complexity can be minimised (Figure 2.10). This fabrication process known as MEMS (Micro-Electro-Mechanical-Systems), allows for the integration of ultra miniature mechanical sensing systems alongside appropriate control circuitry. In addition to reduced size, such systems generally exhibit excellent mechanical resilience, low power consumption and can be produced at a fraction of the cost of those built using traditional technologies [35]. By packaging several MEMS accelerometers and gyroscopes in combination with an appropriate ASIC (Application Specific Integrated Circuit), a low cost, monolithic Inertial Measurement Unit (IMU) is obtained. Typically such devices provide measurement over 3-axes of acceleration and 3-axes of angular velocity and have size of a few centimetres and weight of less than 20*g* [36].

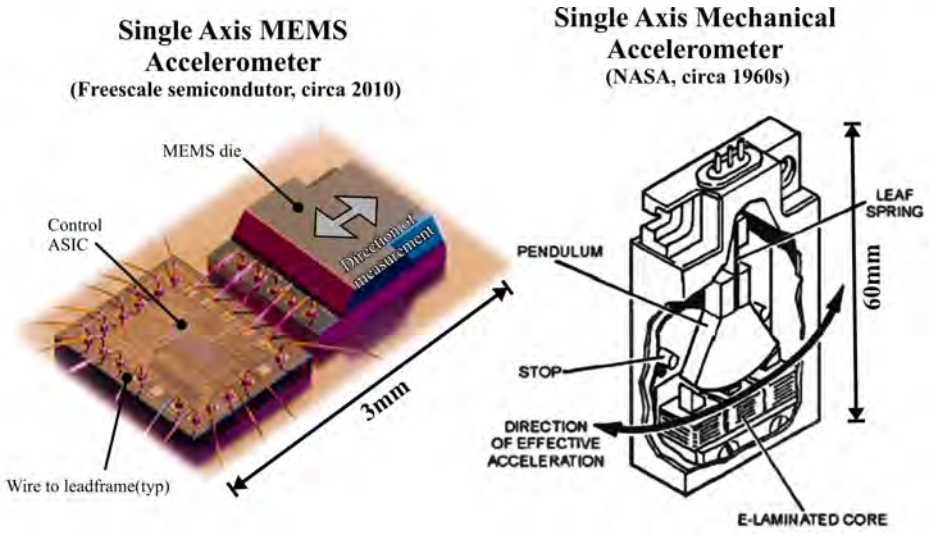


Figure 2.10 – Comparison of a MEMS and mechanical accelerometer

2.3.2 Estimation of orientation

An inertial motion capture system consists of a network of multiple IMUs, 3-axis magnetometers (compasses) and associated data acquisition hardware (Figure 2.11). This network is attached to the human body with one IMU per body segment, allowing for the measurement of the acceleration and angular velocity of individual limb segments as well as their headings with reference to magnetic north. By combining these sensor readings using a data fusion method such as the complimentary Kalman filter [16], an estimate of absolute orientation of body segments against a pre-calibrated frame of reference is obtained. With precision alignment and calibration of sensor axes, accuracies within 0.5^{o} of actual orientation can be achieved. However, due to the dependence of measurement stability upon the resolution of local gravity and magnetic north, considerable drift can occur when a sensor is subjected to strong accelerations or operated in proximity to magnetic materials.

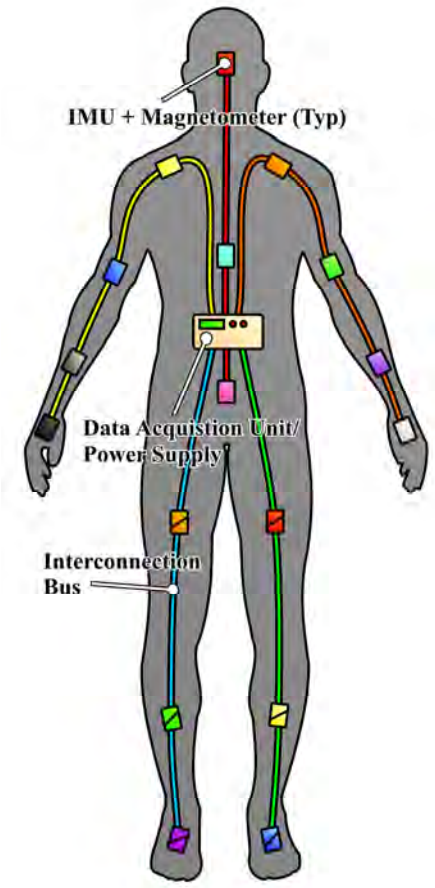


Figure 2.11 – Inertial motion capture setup

2.3.3 Estimation of body pose

In isolation, the orientation data provided by such devices gives little useful information regarding the global position of the body. However, in conjunction with a model of the kinematic constraints of the subject's skeleton, the calculation of angles about joints and hence a rudimentary estimate of pose can be obtained [17]. Practically, this kinematic model

considers the human body as a number of rigid segments (one per IMU), whose orientations are related by fixed transforms to the frame of reference of their attached IMUs. These segments form a hierarchy with rotational constraints at the joints between segments, and fixed translational constraints between these joints (Figure 2.12). By forward kinematics the segments are ‘driven’ about these rotational joints using segment aligned IMU data to define their orientations. The complete pose of the body is reconstructed by following the sequence of transforms down the hierarchy from an origin point to its extremities.

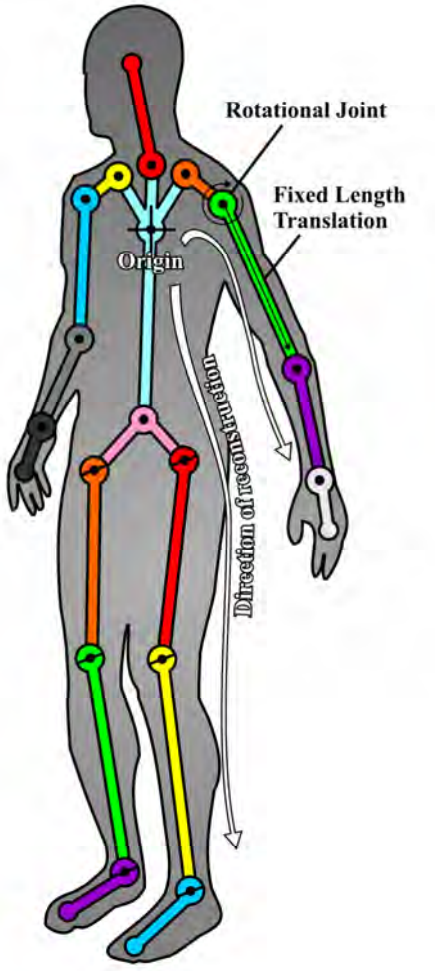


Figure 2.12 – Kinematic Hierarchy

The accuracy of this kinematic reconstruction process is highly dependent on the accuracy with which the defining parameters of the kinematic model are known. However, in the majority of implementations, these parameters are defined using manual methods, which exhibit poor repeatability and precision. For example, the distances between joint centres are typically provided via measurement of the limb length by callipers. Additionally, the alignment between IMUs and body segments is achieved by the subject standing in a predefined calibration pose [18]. Significant improvement in accuracy can be obtained by extracting these initial parameters by the use of a secondary measurement system such as optical motion

capture. However, such means tend to complicate setup procedures and impact upon the portability and flexibility of the method.

The second hurdle to successful inertial pose estimation is the definition of the origin of the kinematic model, which needs to be ‘anchored’ to a point in space for pose to be established in the global environment. This origin is typically obtained from the point at which the feet make contact with the ground. By moving this point from one foot to the other throughout the cycle of a subject’s gait, it is possible to maintain a stable origin point [17] while providing limited positional accuracy in the global environment (Figure 2.13). Although this method can produce visually convincing motion, accuracy from step to step tends to decrease due to accumulation of errors during traversal of the kinematic model from one foot to the other. Additionally, movements where both feet leave the ground will result in a rapid deterioration of measurement accuracy.

An alternative method of generating a stable origin involves combining data from the Global Positioning System (GPS) in tandem with available IMU data [37]. This method allows for improved long range measurement of global translation. However, dependence of this solution on the availability of GPS signals restricts indoor use.

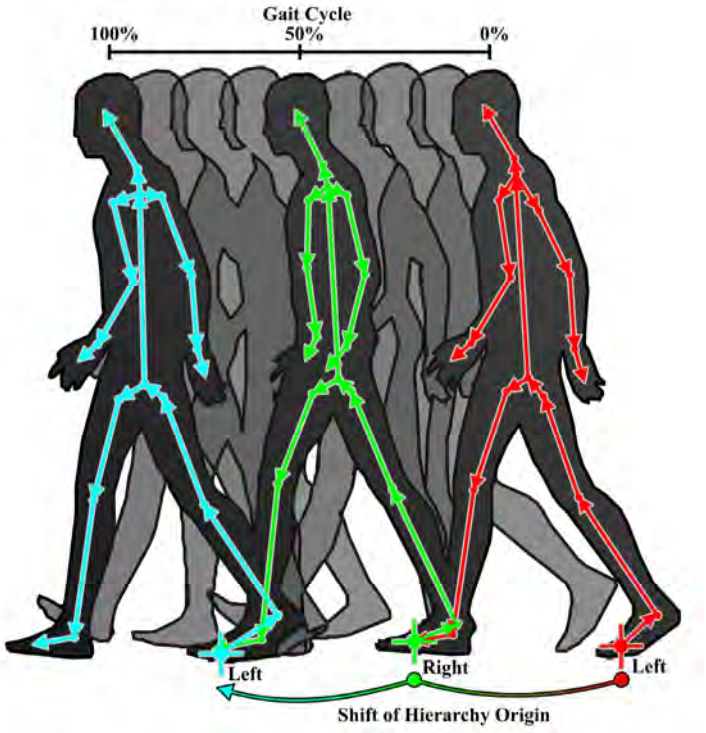


Figure 2.13 – Deriving origin by foot contact point

In terms of practicality, the method of inertial motion tracking presents a number of advantages over more conventional motion capture techniques such as its reduced cost, the simplicity of the equipment involved and its independence from occlusion. However, due to the multitude of approximations on which the method relies, accuracy of pose estimation falls below the requirements of most applications [16, 21].

3. INERTIALLY ASSISTED OCCLUSION RECONSTRUCTION AND TRACKING

3.1 Concept for a trajectory reconstructor

To address the primary aim of this thesis, it is necessary to derive a relationship between the positions of markers upon a subject's body and orientation estimates provided by inertial sensors. This would allow for the prediction of marker positions from orientation estimates independently of the optical measurement system, and facilitate the recovery of their position during occlusion. This relationship will be modelled as a kinematic hierarchy approximating the degrees of freedom of the human body. In this model the body is broken into several segments rotating about estimated centres of joints and the positions of markers are represented as vectors with reference to the pose of these segments. A simplified example of this model is shown in figure 3.1, for two segments corresponding to the upper and lower arm. These segments rotate about centres corresponding to the shoulder and elbow joints respectively. The relationship between these two centres can be described by the vector $\mathbf{v}_c$, referenced to the upper arm segment. Optical markers ($\mathbf{p}_{1\text{-}5}$) are attached to the body segments and are related to their local joint centres by the vectors $\mathbf{v}_{\mathbf{m}1\text{-}5}$. Additionally, an inertial sensor provides a measure of change in the local orientation of the segment to which it is attached. By integrating this change in orientation (ΔRA_{IMU} and ΔRB_{IMU}) against a reference orientation (a process typical described as strapdown integration [38]), an inertially derived estimate of limb orientation is obtained (RA_{SEG} and RB_{SEG}).

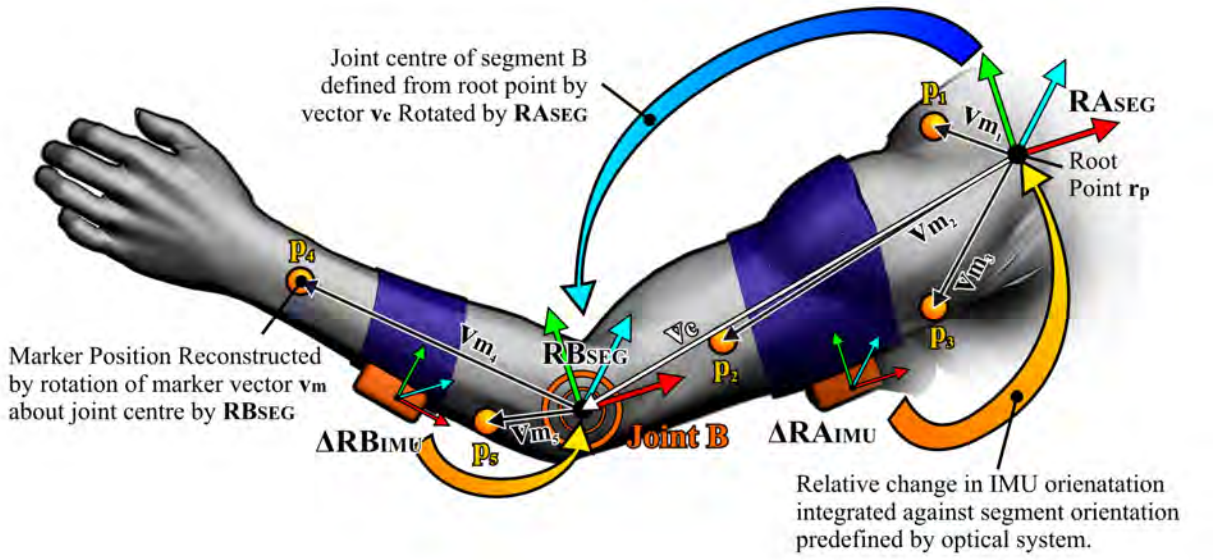


Figure 3.1 – Marker reconstruction by inertial measurement

Given a root point (**$\mathbf{r_p}$**) which describes the position of the local origin of the upper arm segment, it is possible to estimate the position of each marker upon the arm by forward kinematics. By this method, each marker is modelled as the sum of the local vectors between itself and the root rotated by their corresponding segment's orientation estimate. For example, the wrist marker $\mathbf{p}_4$ can be described by the joint centre vector ($\mathbf{v_c}$) rotated by the upper arm orientation, plus its defining vector $\mathbf{v_{m4}}$ rotated by the lower arm orientation. Therefore, using a matrix representation of rotation the five points in the diagram can be estimated as:

$$\mathbf{p}_i = \begin{cases} \mathbf{r_p} + \mathbf{RA_{SEG}}\mathbf{v_{m_i}}^T & i = 1,2,3 \\ \mathbf{r_p} + \mathbf{RA_{SEG}}\mathbf{v_c}^T + \mathbf{RB_{SEG}}\mathbf{v_{m_i}}^T & i = 4,5 \end{cases}$$

As both the orientation estimates and vectors of this model can be obtained independently of marker visibility, this method provides a theoretically sound means by which occluded marker trajectories could be recovered. Additionally, by providing an estimate of position and

velocity of a marker at its point of emergence from occlusion, re-identification of trajectory fragments can be performed.

The performance of inertial sensing based techniques of motion measurement is dependent upon the accuracy of both the vectors defining the body model and the estimates of orientation. In addition, forward kinematic methods are only capable of describing the relative position between rigidly connected segments. Therefore, pose estimation is dependent upon an external reference for the model's root point ($\mathbf{r_p}$) in the global frame. As conventional inertial motion capture methods typically rely upon manual measurements or gross assumptions to define these parameters [18], they give poor repeatability and positional accuracy. Furthermore, inertial estimates of orientation suffer from drift due to the integration of noisy angular velocity data, which results in gradually deteriorating accuracy [16]. Therefore, a kinematic model calibrated and driven by conventional means would struggle to produce position estimates of sufficient reliability for either marker trajectory recovery or tracking.

To address this limitation optical marker data obtained during periods of marker visibility could be used to refine the kinematic model, compensate for orientation and position errors as well as define the root for pose estimation. Calibration of the kinematic model will be performed by application of techniques widely used in biomechanics for the extraction of kinematic parameters from human motion data. These include the estimation of functional joint centres [39] and rigid body pose estimation [40]. To minimise error due to kinematic distance, the root position will be provided dynamically based upon the location of the nearest visible marker to an occluded segment in the kinematic hierarchy. To reduce accumulation of

error in estimates of orientation, periodic realignment will be performed using optical pose estimates. Additionally, the drift in sensor readings will be compensated for by modelling the trend in measurement bias against changes in orientation obtained from optical pose estimates. By these correction mechanisms a mutually supporting system will be realised, in which the kinematic model maintains trajectory continuity and the optical measurements provide calibration and correction for the model.

3.2 Implementation of the trajectory reconstructor

The trajectory reconstructor can be divided into functions based upon either inertial or optical measurements. This allows its operation to be represented in terms of two interacting systems: A kinematic trajectory estimator and optical parameter estimator. Figure 3.2 shows a plan for the proposed trajectory reconstructor detailing the sub components of each system and the estimates and parameters shared between them.

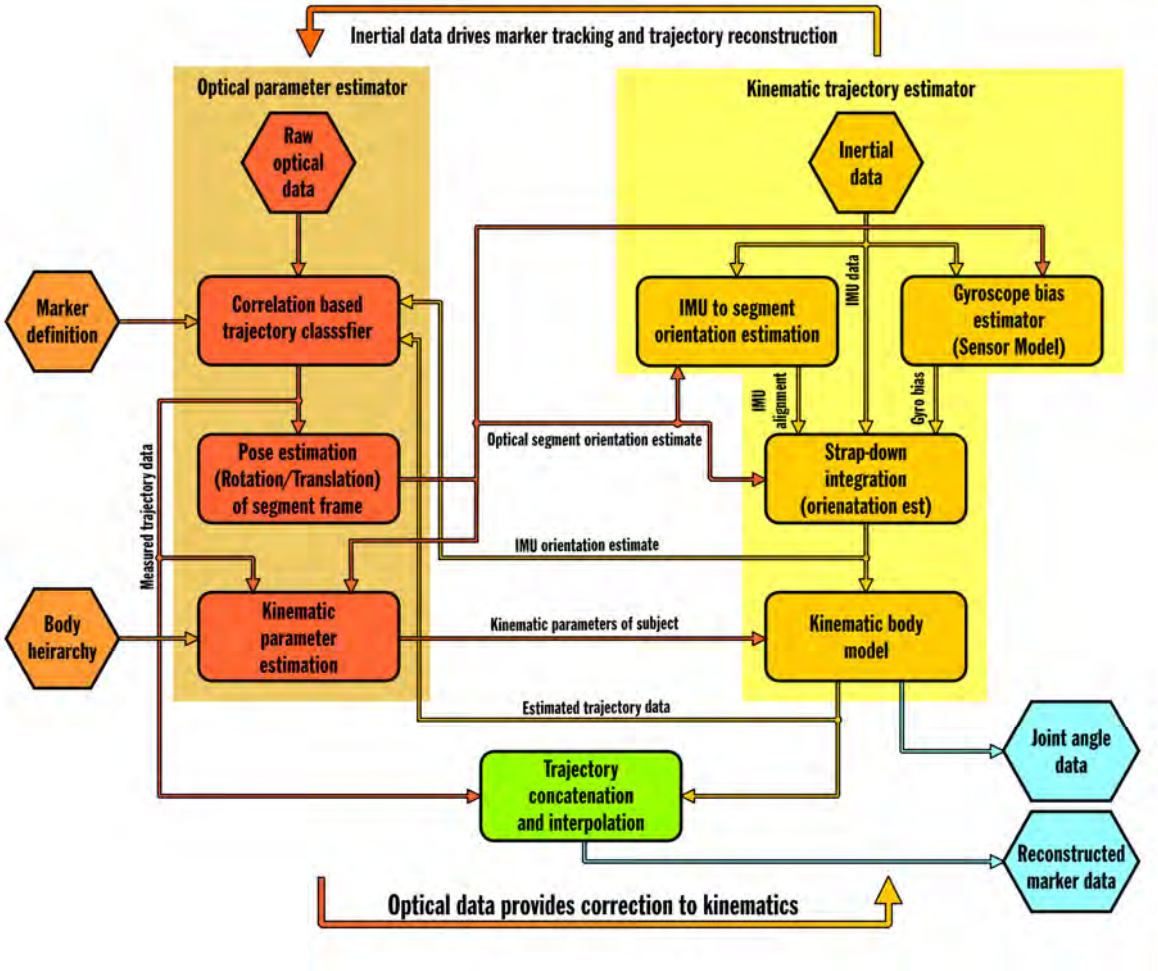


Figure 3.2 – Concept for reconstructor algorithm

It is intended that the optical parameter estimator will process raw coordinate data provided by optical motion capture to extract both the joint parameters for the kinematic model and the pose of body segments. However, before these operations can be performed it will be necessary to identify and associate the raw trajectories with the body segment to which they are attached. It is proposed that this will be conducted by comparing all the visible marker fragments, grouped into clusters, to the estimates of marker trajectories and segment pose provided by the kinematic estimator. By finding clusters whose trajectories produce strong correlations and whose geometry gives high correspondence to predefined cluster configurations [30], marker identity may be inferred. Following identification, it is envisaged that by pose estimation a local origin and rotational frame of reference will be provided for each body segment [40]. This would allow for marker trajectories of neighbouring segments to be expressed in terms of these local frames of reference. By exploiting the relationships between markers in this local frame, distances, which are invariant over the course of a trajectory could be used to infer the locations of the joint centres between segments [39]. These relationships would facilitate the calculation of the vectors relating each joint and marker, which represent the calibration parameters of the kinematic model.

The kinematic trajectory estimator will implement the forward kinematic reconstruction of marker trajectories and estimation of orientation from inertial measurements. It is intended that estimation of orientation will be performed via strapdown integration of gyroscope data acquired from inertial sensors [38]. To ensure correspondence between this estimate and the optically measured segment orientation, the integration will be initialised using the most recently available pose estimate provided by the optical parameter estimator. The errors introduced due to misalignment between segment and sensor axes will be compensated for by

exploiting the relationship between static measurements of acceleration due to gravity and the global reference frame [20].

To obtain further reduction of drift error due to bias in inertial measurements, a sensor model will be used to estimate the expected angular velocity from marker trajectory data. By comparing the long term average of the differences between this estimate and the measured angular velocity, the error in gyroscope offset will be approximated. Finally, by driving the kinematic body model using the corrected orientation data, the occluded marker trajectory data and joint angle parameters will be estimated. To ensure accuracy this kinematic model is calibrated and maintained by the optical parameter estimator.

Despite the error compensation provided by the trajectory reconstructor, it is expected that there will be some disparity between the optical measurements and inertial estimates of marker trajectory. Therefore, to minimise discontinuities between concatenated trajectory fragments, the handoff between trajectories will be performed via an appropriate pose interpolation techniques [41]. The profile of this interpolation over the duration of a trajectory fragment will be controlled by a model of errors accumulated in the kinematic model. It is intended that this scheme will minimise deviation from actual marker trajectories.

3.3 The role of simulation in validation of reconstructed trajectories

In practical measurement scenarios, an occluded marker is completely unknown to the measurement system and can only truly be recovered if additional measurement data is available. However, for validation of the proposed occlusion reconstructor it is necessary that any recovered trajectory data be verified against actual trajectory measurements to provide a benchmark of reconstruction accuracy. Therefore, occluded trajectories obtained by actual measurements will be unsuitable for the purposes of this project as no such benchmark would be available.

To overcome this difficulty it is proposed that the effects of occlusion on measurement data are simulated, providing a means by which ‘gold standard’ trajectory data can be degraded in a controlled manner. Following application of the trajectory reconstructor to this artificially degraded data, accuracy of reconstruction will be assessed by comparison to the original gold standard data. Furthermore, such a simulation would allow for the testing of a wider variety of conditions (i.e. camera configurations) due to simplified experimental requirements. Consistency across experimental conditions can also be guaranteed by virtue of the identical source data.

It is intended that the model of occlusion used in this thesis will be based upon the intersection of epipolar lines (representing the line of sight between cameras and markers) with a geometric model of the human body. By solution of these intersections it will be established whether the pose of a subject results in occlusion of specific markers for a given spatial configuration of cameras [14]. It is envisaged that this body model will be specified through structural parameters extracted from the kinematic model (such as joint positions and

limb lengths) and anthropometric measurements obtained from marker data. The final implementation of the body model will be performed using techniques common to 3D modelling to provide a polygonal mesh with close correspondence to the anatomy of the measurement subject.

The gold standard data set will ideally cover as much of the entire range of human body motions as is possible. This is considered necessary to ensure prevalence of occlusion for efficient testing of the reconstructor and appropriate ranges of joint motion for kinematic calibration. Appropriate ranges of motion may be found in dance and other athletic activities due to their regimented structure and dynamic qualities [10]. The suitability of these activities to the experimental requirements of the thesis will be explored further.

With regard to the acquisition of inertial measurements, it is considered that instrumentation of each segment of the human body with an inertial measurement device would be practically unfeasible. This is due to the expected time and effort required for the implementation of such an apparatus lying beyond the resources available to this project. Additionally, it is assumed that a wired system of inertial sensors would considerably encumber the measurement subject, disturbing marker visibility and restricting subject movement. It is therefore proposed that a model of sensor behaviour driven by the gold standard trajectory data be used to simulate these measurements. By inclusion of a single inertial sensor in the experimental apparatus it will be possible to validate the results of the model against actual measurements.

3.4 Requirements of accuracy and flexibility

The trajectory reconstructor will preserve details of movement, which would otherwise be lost by typical trajectory interpolation methods. However, the positional accuracy it is expected to produce is typically lower than that of optical measurements. Therefore, it is intended that the reconstructor will be suitable for non-medical applications of motion capture, such as the creation of Computer Generated Imagery (CGI) applications, sporting analysis and video games. In these applications, the preservation of nuance of motions is often more important than the precision of measurement. Although the technique may be suitable for some gait analysis applications, it is not intended to be applied to precise biomechanical modelling due to the uncertainties it would introduce.

The error contributed by the reconstructor should be comparable to those intrinsic to optical motion capture. The accuracy of optical motion capture for marker position estimates against the pose of the skeleton is approximately 10-30*mm* [3, 42]. Therefore, this range of error is considered a target for performance of the trajectory reconstructor.

With regard to occlusion, the reconstructor should be capable of tracking and reconstructing gaps in marker trajectories produced by a minimal optical motion capture setup. Although a minimum of six cameras is suggested by manufacturers of optical motion capture equipment for full body measurement [7,8], this project would use four cameras to test the reconstructor under adverse occlusion conditions. Ideally, reconstruction accuracy should remain within the specified limits for the maximum duration of occlusion events produced by this camera system.

As the aim of this project is to develop and validate a novel method of occlusion recovery, rather than produce an integrated software solution, computational efficiency of the reconstructor is not a priority. For this reason the high level language MATLAB (Version 7 R14) will be used for implementation of the reconstructor. This language is considered particularly appropriate for this task, given its optimised handling of matrix operations and integrated analytical functionality. However, the efficiency and portability of the code produced in this language is generally poor.

3.5 Safety and ethical considerations

For the experimental sessions of this project only low impact physical exercises will be required, posing minimal risk of injury. To reduce the risk of allergic reaction, all tape and adhesives used in proximity to the subject's skin will be confirmed to be hypoallergenic.

As the experimental sessions will involve the attachment of active electronic devices to the measurement subject, care will be taken to eliminate the possibility of electrocution. Preventative measures will include double insulation of the mains supply, battery power of equipment where possible, and appropriate isolation measures at interfaces between low and high voltage apparatus.

This investigation will involve a single human participant, therefore, it is necessary to comply with the ethical guidelines required for such studies. Therefore, the participant will be informed of the following:

- The purpose of the experiment and its aims
- Their data will remain anonymous.
- Their consent can be withdrawn at any point during the experiment.
- Their withdrawal from the experiment would incur no penalty whatsoever.

This information will also be provided in writing and the subject will be asked to sign confirming their agreement to participate in the experiment.

4. DERIVATION OF THE BODY MODEL AND MARKER SET

This chapter describes the marker set used in the pilot experiment of the thesis and the simplified body model from which it was derived. Firstly, the concept of the functional body model is introduced and the assumptions required for its implementation are discussed. Secondly, the choices of functional joint locations are explained with reference to the prediction of marker trajectories. Finally, a full body marker set is presented which attempts to minimise measurement errors due to soft tissue movement.

4.1 Assumptions about human motion and the functional model

The body model is based upon the assumption that the motion of the human body can be described completely by the system of constraints imposed by the bones and joints of the skeleton. These constraints are commonly modelled as a kinematic hierarchy, composed of multiple rigid bodies connected by rotating and translating articulations [10, 17]. Through measurement of points on the surface of the body, such a model can be referenced to the anatomy of a subject [43, 44, 45], inferring the pose of the joints, muscles and ligaments of the body. Simplified kinematic models find applications in fields such as computer animation [6], sports science [37] and human computer interaction [46].

For the reconstruction of marker trajectories it was necessary that the body model was capable of accurately reproducing the positions of superficial points on the body. As an anatomical model represented an inefficient means of achieving this aim, a simplified ‘functional’ model was favoured.

4.1.1 Functional body model

By abstracting the underlying anatomy of the human body and considering its motion in terms of fundamental degrees of freedom, a much simplified 'functional' body model was implemented [47]. This model represented the human body in terms of rigid representations of its gross segments connected by the ideal centres of rotation of the major joints. This is in contrast to an anatomical model, which represents the pose of each bone of the skeleton with reference to the physical joints between them. The functional model gives a good approximation of the pose of the body in proximity to specific points of interest (such as sites of marker placement), at the expense of accuracy in other locations. Figure 4.1 shows an example of two models of the forearm, one modelling the functional behaviour of the wrist, and the other accounting for the movement of the ulna and radius. In the functional model, as rotation of the forearm occurs about the elbow, the model only produces accurate pose around the wrist. However, in the case of the anatomical equivalent, the rotation due to pronosupination occurs over the whole length of the forearm, as for the actual skeleton.

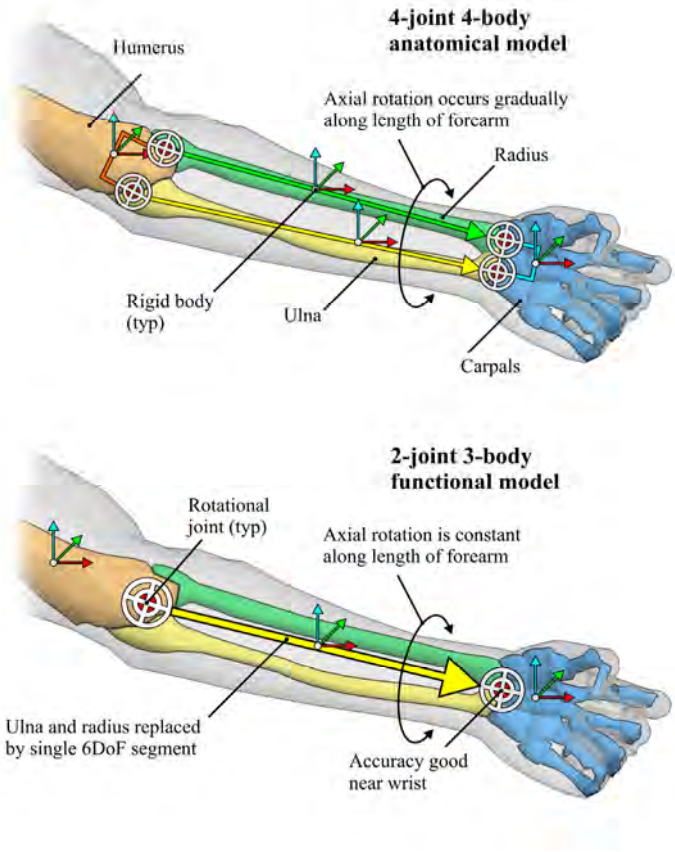


Figure 4.1 – Equivalent functional and anatomical models of the lower arm

To provide the simplest possible representation of body posture, the model used in this thesis was restricted to two components: Rigid bodies representing the pose (translation and orientation) of marker clusters and rotational joints connecting them. All joints were implemented as fixed 3 DoF (3R) joints. As shown by the ball linkage in figure 4.2, this joint model allows the complete pose of any rigid segment to be expressed by forward kinematics as a series of summed vector rotations.

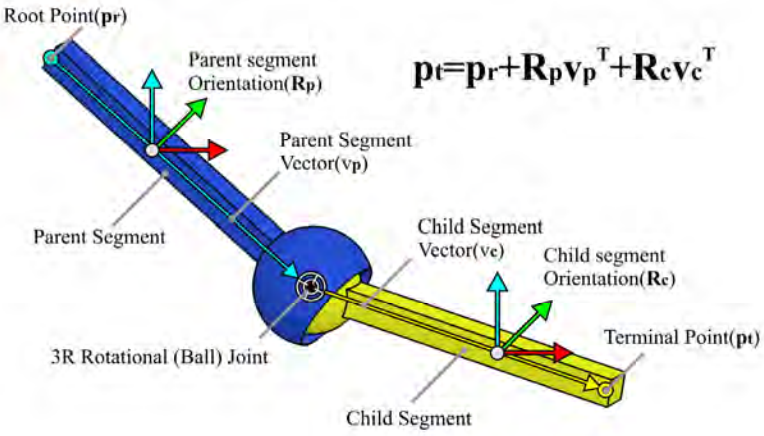


Figure 4.2 – Typical 3R Constraint and reference frames

4.1.2 Anatomical deviation from functional approximation

Although the use of 3R articulations for the modelling of functional joints is common practice in both human movement analysis [48] and animation [6], it has several drawbacks which compromise the accuracy of representation. The first of these is that most anatomical joints exhibit some degree of translation of their centre of rotation when loaded. However, it has been found in laboratory studies that even under extreme loading of structurally compromised joints such movement does not exceed 10*mm* [49].

The second, related problem is due to the convention of defining rigid bodies based on the functional segmentation of the body rather than its underlying skeleton (Figure 64). Under these conditions, the assumption of rigid body motion can be questionable as a functional segment may be composed of several articulated bones. The severity of error introduced by

this simplification varies depending on the complexity of the underlying anatomy and the additional degrees of freedom it introduces. For example, the simplification of the lower arm to a single rigid body gives excellent accuracy with respect to the pose of the hand, whereas at points closer to the forearm the accuracy decreases (4.1.1).

4.1.3 Accumulation of kinematic error

The third limitation of the functional model is the tendency of error to increase rapidly with successive traversal of joints when segment pose is estimated by forward kinematics. This problem affects any orientation based method of pose estimation such as the occlusion reconstruction techniques developed in chapter 10. Its effect is dependent on the level of error in segment orientation estimates and the distances between joints in the model. Figure 4.3 shows the relationship between the worst case angular (orientation) error (ε_n), segment length (l_n) and maximum positional error radius (d_n) for the example of an N segment kinematic 'chain'.

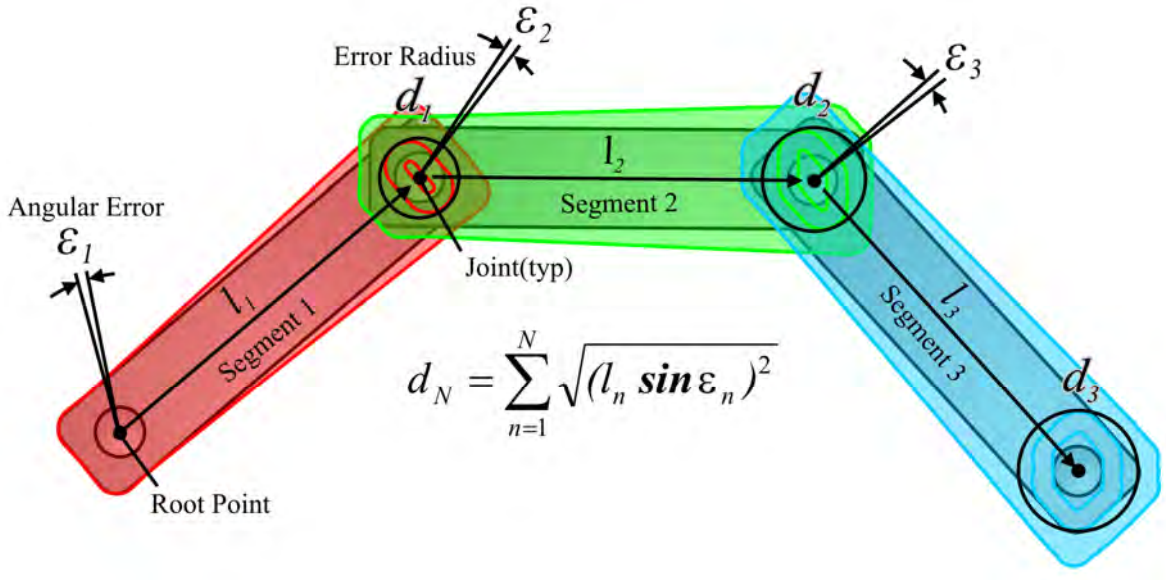


Figure 4.3 – Trigonometric model of the accumulation of positional errors across a kinematic chain due to angular error at its joints

This error occurs due to the length of each segment acting as a lever, magnifying small changes in the angles of joints. This results in large differences in the position of the end of a segment.

In a large kinematic model, even with relatively low levels of orientation measurement error, increasing kinematic chain length will eventually lead to unacceptable levels of error in pose estimation [17]. Therefore, it is essential that both kinematic chain length and errors in orientation estimates are minimised.

4.1.4 Soft tissue movement

In deriving and actuating a functional kinematic model using marker based motion measurements, the largest source of error is typically due to soft tissue movement [50, 51, 52, 53]. Generally, these errors stem from the combined movement of all compliant materials (for example, fat, muscle, skin and clothing) between a marker location and the skeleton during body motion. Due to the complex dynamics and mixed physical characteristics of such materials, compensating numerically for the effect of these errors upon marker position is not trivial. Examples of methods, which have attempted to address this problem are the landmark calibration method given in Cappello et al. [52] and the surface cluster method of Alexander et al. [53]. However, such methods typically only produce marginal reductions in soft tissue artefacts and depend upon extensive anatomical calibration or dense arrays of markers. Therefore, they are considered too cumbersome for integration with the body model and experimental methods of this thesis. Alternatively, passive means of reducing the influence of soft tissue movement upon measurements were utilised, such as plate mounted clusters and anatomical landmarks (discussed further in 4.3.1). Beyond these contingencies, this artefact

was considered to be an inherently unpredictable component of measurement noise and no further attempt was made to compensate for its effects.

The scale of positional error introduced by soft tissue motion widely varies depending on factors such as the position of markers, acceleration and the physical build of a measurement subject. However, from studies of the relative motion between surface mounted markers and bone embedded fixators, a mean error of ~10*mm* (peak ~30*mm*) for displacements [3] and ~3^{o} (peak ~8^{o}) for rotations [42] has been observed. This value is considered to approximate the scale of error which would be expected between measured marker positions and those predicted by a perfect model of the skeleton (functional or anatomical). Therefore, this range of errors is considered a guideline for the expected error of kinematically predicted marker position and hence the optimum accuracy of trajectory reconstruction.

4.2 Design of body model

The functional body model was based on those presented in [10,45,47]. It consisted of 18 rigid segments connected by 16 3R rotational joints. These segments are shown in figure 4.4 along with the naming convention by which they were identified. The model describes the associations of segments with their functional joints (or functional centres of rotation) and local coordinate frames. The actual positions of the functional centres of joints were estimated via a model calibration algorithm, and the local origins of segments were defined via the centroids of associated marker clusters (described in 6.1.2).

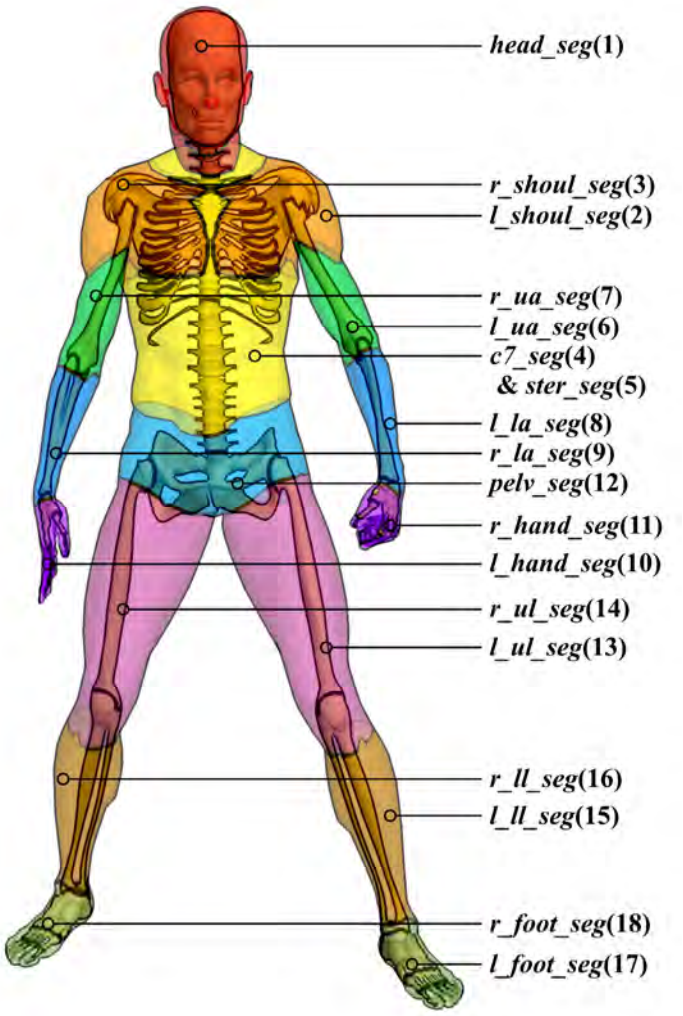


Figure 4.4 – Segments of functional body model and naming conventions

4.2.1 Head and shoulder segments

Figure 4.5 shows the local coordinate systems and approximate functional centres of the head and shoulder segments. The distal *head_seg* segment pivots around the *neck* functional centre (distal to *c7_seg*) simplifying all the cervical vertebrae to a single joint. As this approximation gives poor anatomical correspondence, some translational ability of the head segment is lost.

For the purposes of this project, this was considered a reasonable compromise when compared with the reported difficulties of measuring and modelling the pose of the cervical vertebrae [10, 54]. The convention for the local axis system of the head is for the *z*-axis to point toward the top of the head and *x*-axis toward the forehead.

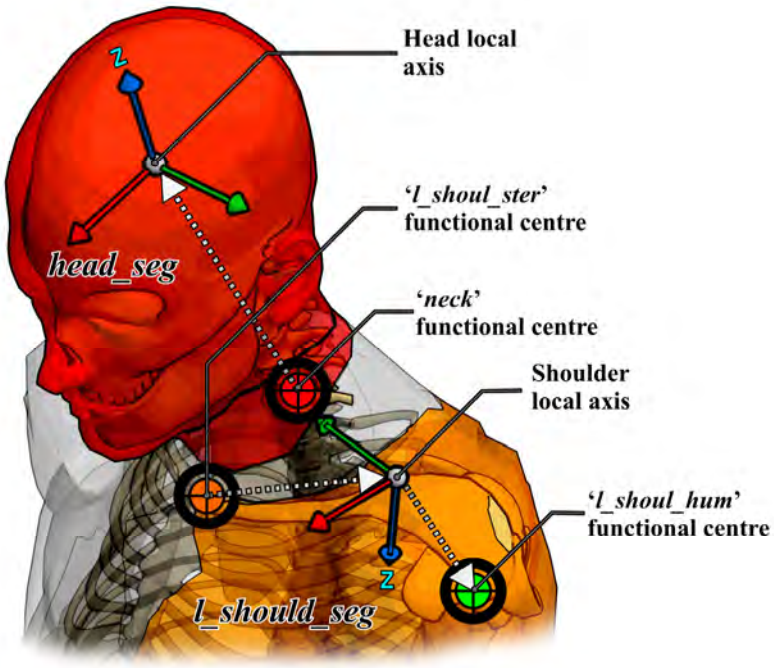


Figure 4.5 – Head and shoulder segments and their functional centres

The *l_should_seg* pivots about its proximal functional centre, *l_shoul_ster*, which loosely approximates the anatomical location of the sternoclavicular joint (distal to *ster_seg*). Its distal end defines the *l_shoul_hum* functional centre which approximately corresponds to the glenohumeral anatomical joint. This allowed the pose of the clavicle to be represented using a single rigid body and two 3R joints. Although this representation was sound in theory, problems were encountered in calibrating these functional centres (6.4.2). The right shoulder segment *r_should_seg* is effectively identical to the left reflected in the midsagittal plane. The

local *z*-axis of the shoulder points toward the ground when standing normally, while the *y*-axis points medially.

4.2.2 Arm and hand segments

The functional model for the left arm (Figure 4.6) was a simple kinematic chain running from its proximal origin at the *l_shoul_hum* functional centre to its distal terminus at the *l_hand_seg*. Two 3R joints, *l_elb* and *l_wrist* approximated the centre of the elbow and the radioulnar joint respectively. The simplification of radius and ulna to a single rigid body limited accuracy of supination/pronation to the distal end of the segment. Therefore, the local origin of the *l_la_seg* was placed close to *l_wrist*.

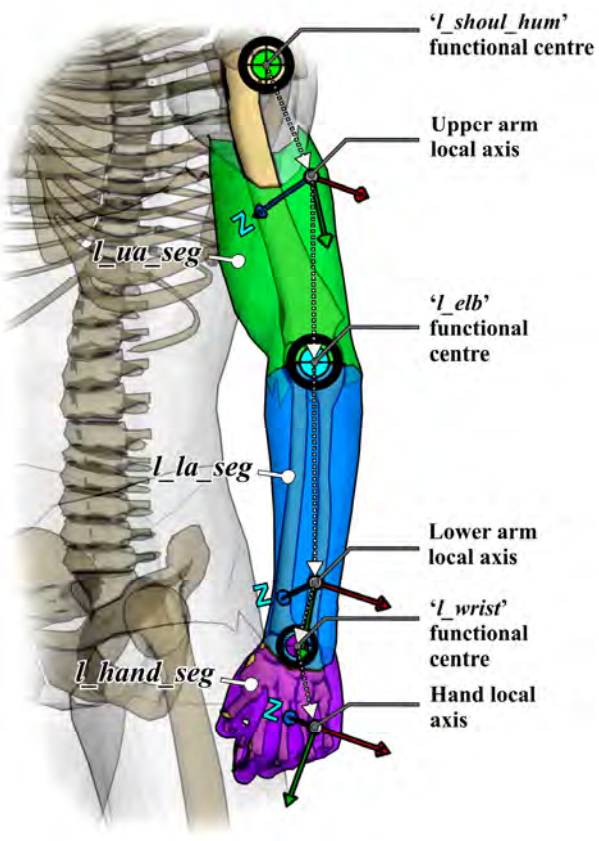


Figure 4.6 – Left arm and hand segments and their functional centres

The entire hand was represented as a single rigid body, *l_hand_seg*, with no consideration of the digits. This greatly simplified the body model and avoided the difficulties inherent in measuring the pose of the hand [55].

The local *z*-axes of all segments of the arm point medially with the arm relaxed. The *y*-axes are aligned with the length of the limb in the distal direction. The right arm model is identical to the left save for reflection in the midsagittal plane.

4.2.3 Torso and pelvis segments

The two segments of the torso and pelvis are shown in figure 4.7. The torso formed the root of the functional model and was composed of two sub-segments, *c7_seg* and *ster_seg.* These two segments approximated the pose of the torso at both the sternoclavicular joint and the 7th cervical vertebra (C7). The purpose of this redundant representation was twofold, firstly to provide two points by which the root of the model was defined and secondarily to facilitate estimation of the stability of the two *shoul_ster* joints with reference to C7. In the torso axis system the *z*-axis protrudes posteriorly from the back and the y-axis points down the spine.

The proximal articulation of the pelvis segment (*pelv_seg*) was about the *pelv* functional centre. As in the case of the *neck*, this centre approximated all the thoracic vertebrae as a single rigid body with the consequence of eliminating translational movement. The two distal joints of the *pelv_seg* were *l_hip* and *r_hip*, which approximately corresponded to the acetabulofemoral anatomical joints of the hip. The z-axis of the pelvis's local co-ordinate system points toward the head with the x-axis aligned anteriorly.

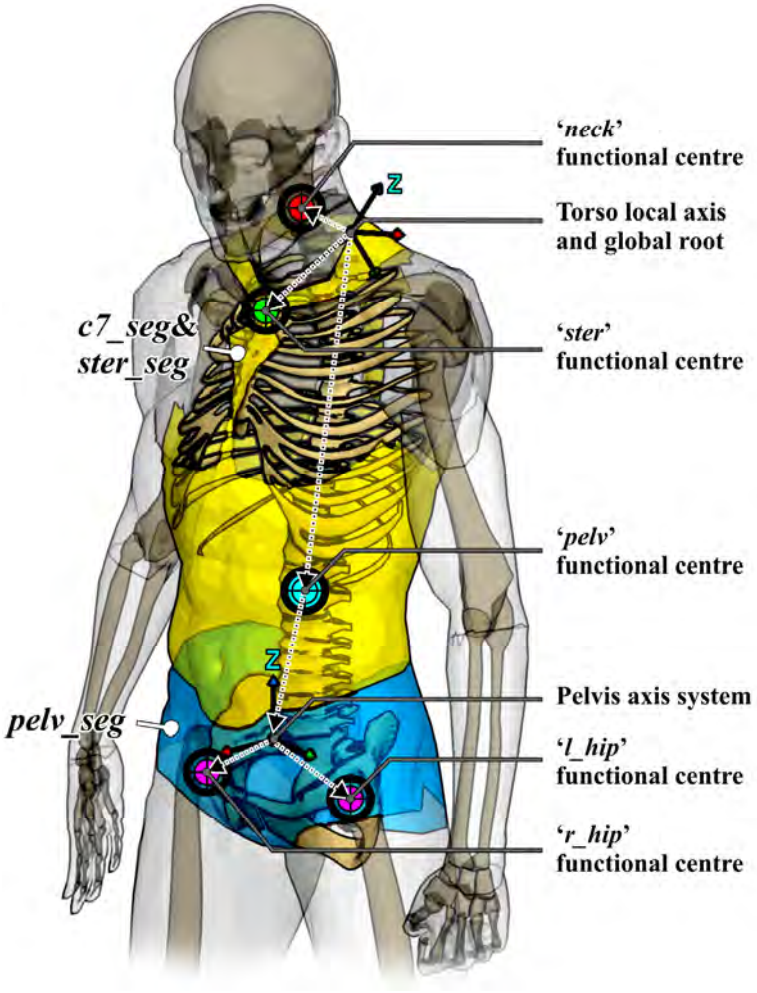


Figure 4.7 – Torso and pelvis segments and their functional centres

4.2.4 Leg and foot

The model of the left leg and foot (Figure 4.8) was analogous to that of the arm and hands. It consisted of a three jointed kinematic chain originating at the *l_hip* centre and terminating at *l_foot*. The fibula and tibia were simplified to a single rigid body, *l_ll_seg*, which gave accurate estimates of the pose around the *l_ankle* joint. The *l_knee* centre roughly corresponded to the tibiofemoral joint. The local *x*-axis of the leg segments points medially when standing normally. The *y*-axis points distally along the length of each segment.

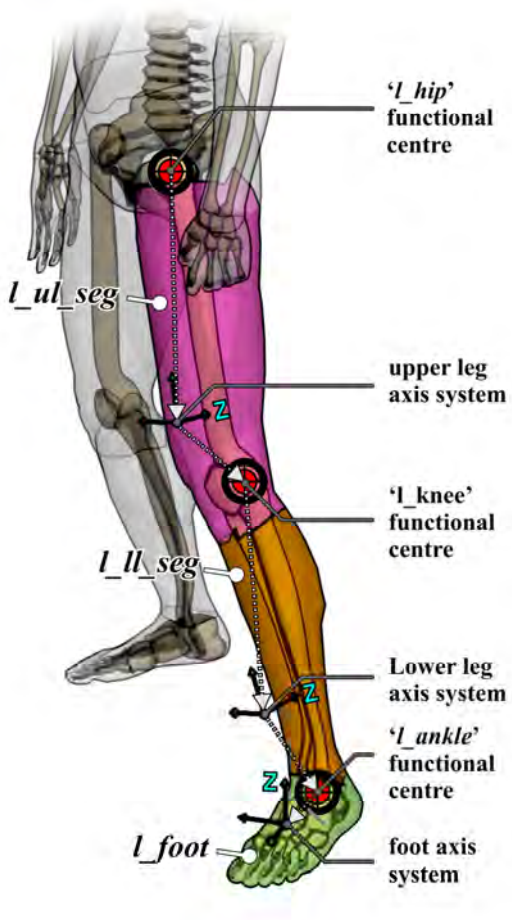


Figure 4.8 – Leg and foot segments and their functional centres

As was the case with the hand, the foot was simplified to a single rigid body pivoting around the l_ankle centre, which approximated the location of the talocrural joint. The right leg model is a sagittally reflected version of the left. The local *x*-axis of the foot points medially and the *y*-axis toward the toes.

4.3 Derivation of full body marker set

This section derives a full body marker set for the purposes of calibrating and actuating the body model via human motion measurements. It discuses the requirements for the placement of markers and several passive methods by which measurement error can be minimised. Finally, the preliminary marker set is presented as used in the pilot experiment.

4.3.1 Considerations for marker placement

In isolation, the pose of a rigid body segment is described as having six degrees of freedom (6DoF) due to its ability to both rotate and translate in the three axes of the Cartesian coordinate system. To completely constrain this pose the positions of at least three non-collinear points upon the surface of the rigid body must be known [48, 40]. Therefore, each segment of the body model must be represented by at least three non-collinear markers, an arrangement referred to as a ‘cluster’.

For the pose obtained from a marker cluster to have good correspondence with the pose of a body segment, it is essential that the site of marker placement exhibits minimal relative movement with respect to the underlying bone. This implies that markers within a cluster should also exhibit a similarly stable relationship with each other if accuracy is to be maintained. As discussed in (4.1.4), the dominant source of motion between marker and bone, and therefore error in pose estimation is due to soft tissue motion. Although the possibility of eliminating or modelling this error has been ruled out, various conventions were adopted in the design of the marker set to reduce its influence. These are listed as follows:

- *Marker placement away from joints and mobile areas:*

  Skin movement artefacts are most prominent in areas with large quantities of fat or muscle. Additionally, large surface displacements may be observed on or around the joints of the skeleton. Therefore, marker sites should be chosen to avoid these locations [3].

- *Placement of markers near anatomical landmarks/bony ridges:*

  The human body presents several locations where the bones of the skeleton come close to the surface of the skin to form palpable landmarks or bony ridges [42]. These landmarks are commonly used as natural reference points to assist in marker application and are often associated with areas where soft tissue artefact is minimal [56]. However, it is not always desirable for markers to be placed directly over bony protuberances, as the slippage of skin over these points leads to increased error.

- *Cluster geometry:*

  To minimise error in orientation estimation Cappozzo et al. [57] states that the average radius of a marker cluster about its centroid should be at least 10 times the standard deviation of the measurement system's resolving error. The resolving error of the system used in this project was approximately 2*mm* [7] resulting in a minimum cluster radius of 20*mm* (though it is later shown that 40mm is a more realistic minimum radius). To ensure even sensitivity to rotation in all axes, the ratio of the least to greatest dimension across a cluster should be greater than half. In the clusters where this aspect ratio is less than one, sensitivity of orientation to marker perturbation will be highest around the long axis of the cluster. Therefore, it is prudent to align this axis in a direction where the error due to soft tissue movement is the largest.

- *Use of plate mounted clusters*

  The area around an appropriately stable point for marker attachment can be insufficient to satisfy the required dimensions of a cluster. To avoid placing markers in unstable locations the cluster is assembled upon a rigid plate, which is taped or strapped to the body. This technique has been found to significantly reduce surface to bone displacement in comparison to skin mounted marker schemes [58]. Additionally, this arrangement virtually eliminates intra-cluster marker motion improving rigid body correspondence. Finally, the use of plates greatly simplifies the process of applying markers to a subject and provides a stable mounting location for inertial sensors (8.2.2).

- *Redundant markers*

  Although only three markers are required for 6DoF representation, there are several advantages of using more. Firstly, reliability of measurement is improved as markers which become occluded will not prevent pose estimation of the cluster. Secondly, accuracy of pose estimation is increased as the additional position data over-constrains orientation of the segment [53]. This allows the influence of perturbations of single markers to be reduced. Finally, the addition of extra markers can assist in the identification of clusters [46] and the rejection of erroneous orientation solutions (12.2.1). From the experiments conducted in Cappozzo et al. [57] it was found that four markers per cluster gave a good compromise between redundancy, accuracy and practicality.

- *Practical and visibility concerns*

  It is essential that marker locations should be practical for the subject and visible to the camera system. For example, markers placed on the inside surfaces of limbs will be prone to both occlusion and collisions during motion.

4.4 Description of full body marker set

By applying the conventions for marker placement to the requirements of the body model, a 68 marker set for the pilot experiment was devised. The remainder of this chapter presents this marker set and discusses its design. A nomenclature by which each marker can be identified is introduced and conventions adopted from standard marker schemes are noted.

4.4.1 Head and shoulder markers

The *head_seg* cluster consisted of 4 markers arranged around the circumference of the head, level with the forehead at ~90^{o} degree intervals (Figure 4.9). This resulted in an average cluster radius of ~100*mm*. This configuration is common to several established marker sets, for example: the Vicon “plug in gait” and “golem” schemes [59]. The markers were attached to the subject via an elasticated headband.

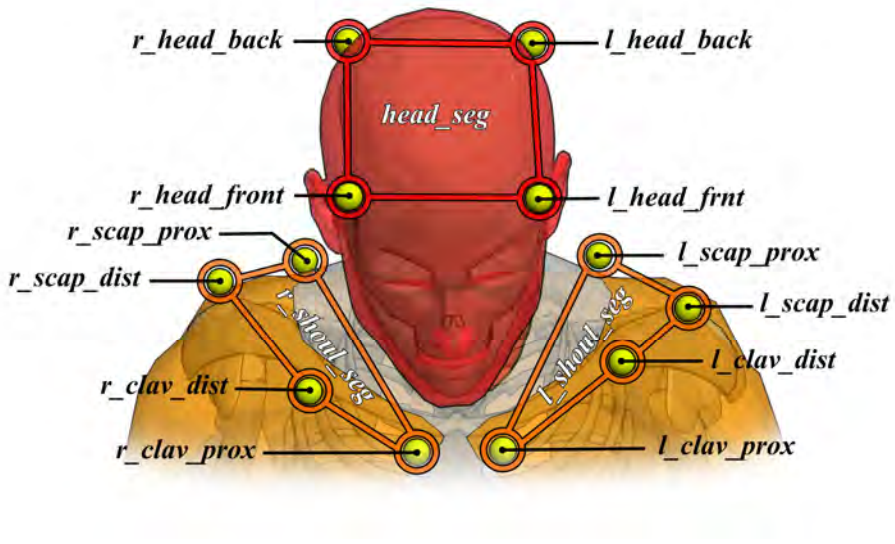


Figure 4.9 – Head and Shoulder Markers

In the case of the shoulders, a four marker cluster was placed over the approximate plane formed by the upper landmarks of the shoulder bones giving an average radius of ~70*mm*. The cluster was composed of surface mounted markers as the use of a plate marker was prohibited by the mobility of the anatomy. The landmarks chosen were the sternal head, coracoid process, tip of the acromion process and the superior angle of the scapula corresponding to the *r_clav_prox*, *r_clav_dist*, *r_scap_dist* and *r_scap_prox* markers respectively. This choice of marker locations was taken from the International Society of Biomechanics (ISB) compliant configuration presented in Jones et al. [60].

4.4.2 Arm and hand markers

In the marker configuration of the arms and hands, plate mounted markers are used extensively (Figure 4.10). The upper arm plates (*r_ua_seg* and *l_ua_seg*) were placed medially/laterally to the bicep and the lower arm plates (*r_la_seg* and *l_la_seg*) approximately corresponded to the distal end of the ulna and radius. Four markers were applied per plate in a rectangular configuration with an average inter marker distance of ~60*mm*. Some asymmetry was intentionally introduced in each marker cluster to reduce the possibility of multiple orientation solutions and facilitate automatic recognition of clusters (12.2.1). The marker clusters of the hands were composed of three surface-mounted markers placed approximately over the carpals. The limited space upon the anatomy restricted average cluster radius to ~30*mm*, which was approaching the lower limit for optimal orientation resolution.

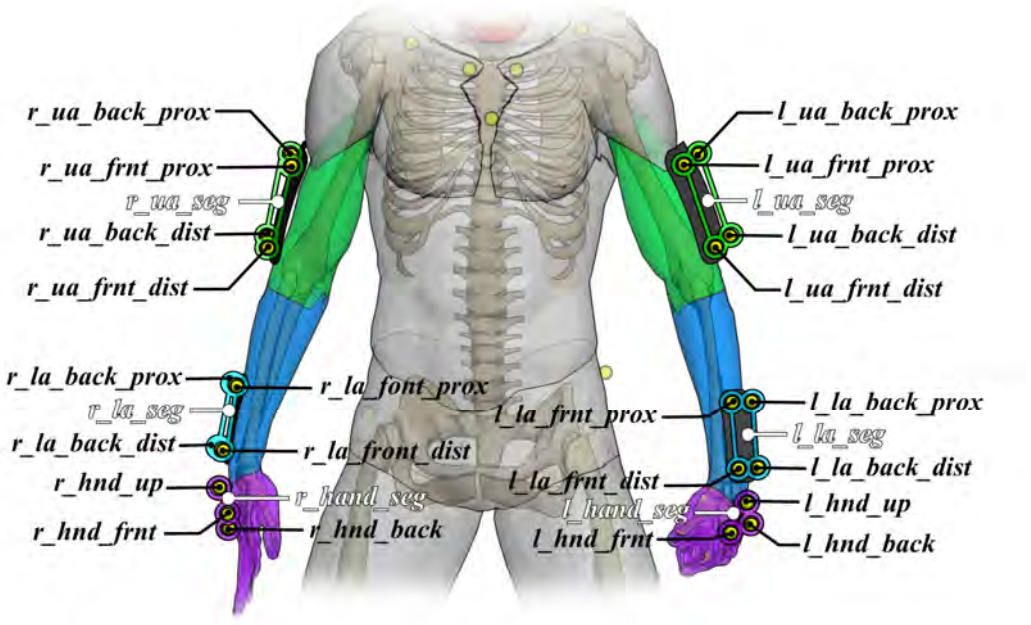


Figure 4.10 – Arm and hand markers

4.4.3 Torso/pelvis markers

To provide the additional redundancy required in the torso segment (4.2.3) six surface mounted markers were used to define the segment (Figure 4.11). Two markers in the sternum cluster (*ster_seg*), *r_clav_prox* and *l_clav_prox*, were shared with the shoulder clusters. The c7 (*c7_seg*) cluster was arranged around the prominence of the spinuous process of the 7th cervical vertebra. Average radii of the clusters were ~90*mm* and ~60*mm* for the sternum and c7 clusters respectively.

The hip cluster configuration (*pelv_seg*) was a variation of the configuration seen in the 'plug-in-gait' model [59] and the 'Helen-Hayes' maker set [4]. The marker locations about the horizontal circumference of the hips were obtained by two carbon fibre rods (or pylons) attached to plates located upon the left and right-hand iliac crest. These pylons projected the markers away from the body providing increased visibility to the camera system. The use of

sacral markers was avoided as they were prone to large perturbations when applied over clothing. Average radius of the hip cluster was ~170*mm*.

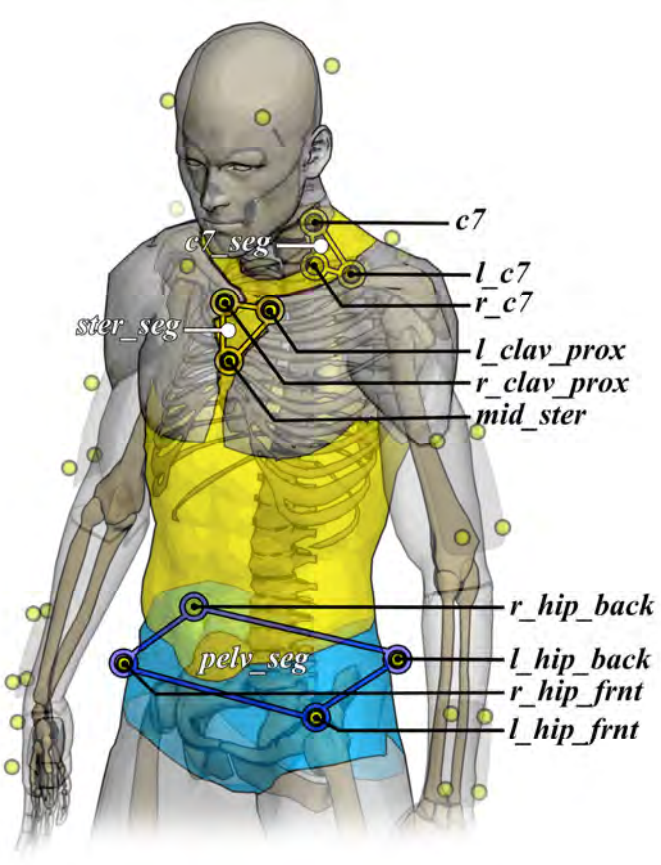


Figure 4.11 – Torso and pelvis markers

4.4.4 Leg and foot markers

Like the arm, the marker set of the leg (Figure 4.12) is predominantly composed of plate mounted, rotationally asymmetrical, semi-rectangular clusters. The positions of the clusters for the thigh and shank segments (*ul_seg* and *ll_seg*) were derived from the set used in Collins et al. [61]. The thigh clusters were moved distally to reduce soft tissue errors. Average radius was ~70*mm* for all clusters.

The foot marker clusters (*l_foot_seg* and *r_foot_seg*) consisted of four surface mounted markers. Three of these markers formed a triangle of average radius ~120*mm* for the estimation of pose. The third marker (*l_foot_up* and *r_foot_up*) provided a reference by which flexion/extension of the toes was assessed. Markers were attached to the shoes of the subject with the aim of constraining movement of the foot to a single rigid body.

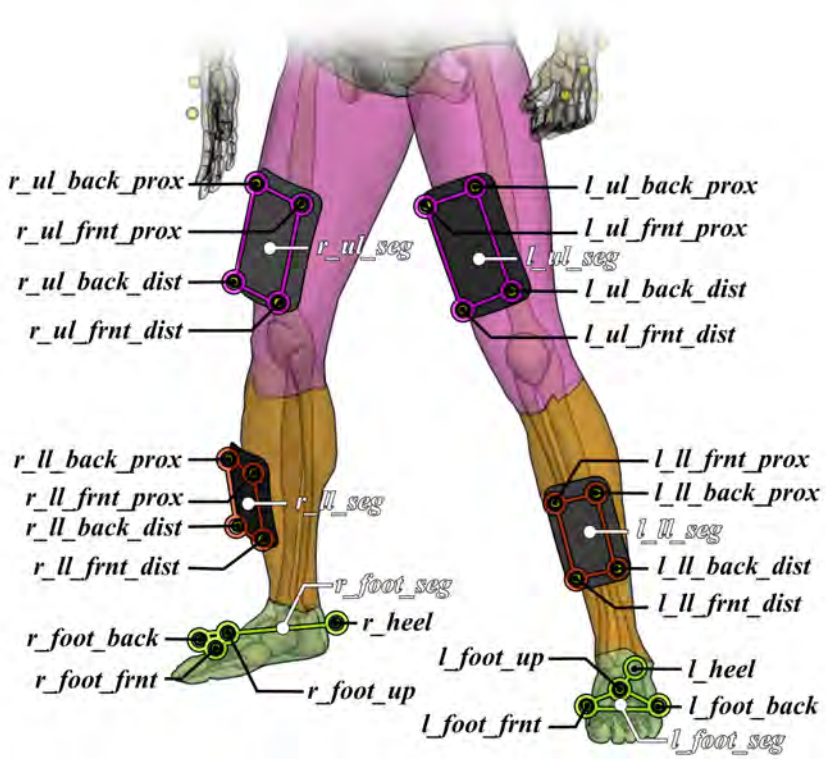


Figure 4.12 – Leg and foot markers

4.5 Summary

This chapter has described the development of an 18 segment body model with 16 functional joints and its corresponding marker set. The model was composed of purely rotational 3R joints and fixed translations resulting in a total of 48 DoF. The marker set was optimised to provide 6DoF representation of each body segment and minimise the influence of soft tissue movement and measurement errors upon the estimation of segment pose. It consisted of 66 markers arranged in clusters of 3 or 4 markers, with 10 plate mounted clusters and 9 clusters assembled directly upon the subject's skin.

5. PILOT EXPERIMENT

This chapter describes the pilot experiment which provided the motion data for the initial development of the body and occlusion models. Firstly, the requirements of the experimental measurement system are specified. Secondly, an appropriate set of reference motions are selected for testing and calibration of the body model. Finally, the protocol of the experiment is discussed and a summary of its execution is presented.

5.1 Specification of primary measurement system

The primary measurement system was a 12 camera Qualisys Oqus 3-series system [7]. This system was chosen preferentially over two older and well established measurement systems (a 12 camera Qualisys Proreflex [28] and 6 camera Vicon 512 system [62]). Although the use of a new and relatively untested system presented possible reliability problems, these were outweighed by the performance advantages offered by the Oqus system. In particular, the improved resolution of 1.3 million pixels (vs. 0.3 million pixels for both systems) was considered a key requirement given the large number of markers in the full body marker set. In addition to the motion capture data, a conventional video camera provided a visual record of the experiment. As the inertial measurement apparatus was unavailable for the pilot experiment, direct measurements of limb acceleration and rate of turn were not performed. This was due to the apparatus being implemented later in the project from a specification based on the results of this experiment.

5.1.1 Camera installation

The measurement system was installed within the 8 x 8 x 5m^3 space of the Posture and Balance Laboratory in the School of Psychology, University of Birmingham. Positioning of the cameras was achieved via photographic clamps (Manfrotto super clamp) and tripod heads (Manfrotto 804RC2) attached to the existing mounts of a previously installed 6 camera system (Figure 5.1). Additional camera positions were provided by standard photographic tripods. The resulting configuration was approximately circular with cameras distributed at regular intervals around the perimeter of the room with an average height from the ground of approximately 1.8*m*. This provided a capture volume radius of approximately 1.5*m*. Acquisition and processing of captured data was performed using a standard PC (Intel Core2 2.8*GHz*, 2Gb RAM) running Qualisys Track Manager (QTM) version 2.0.324. The origin of

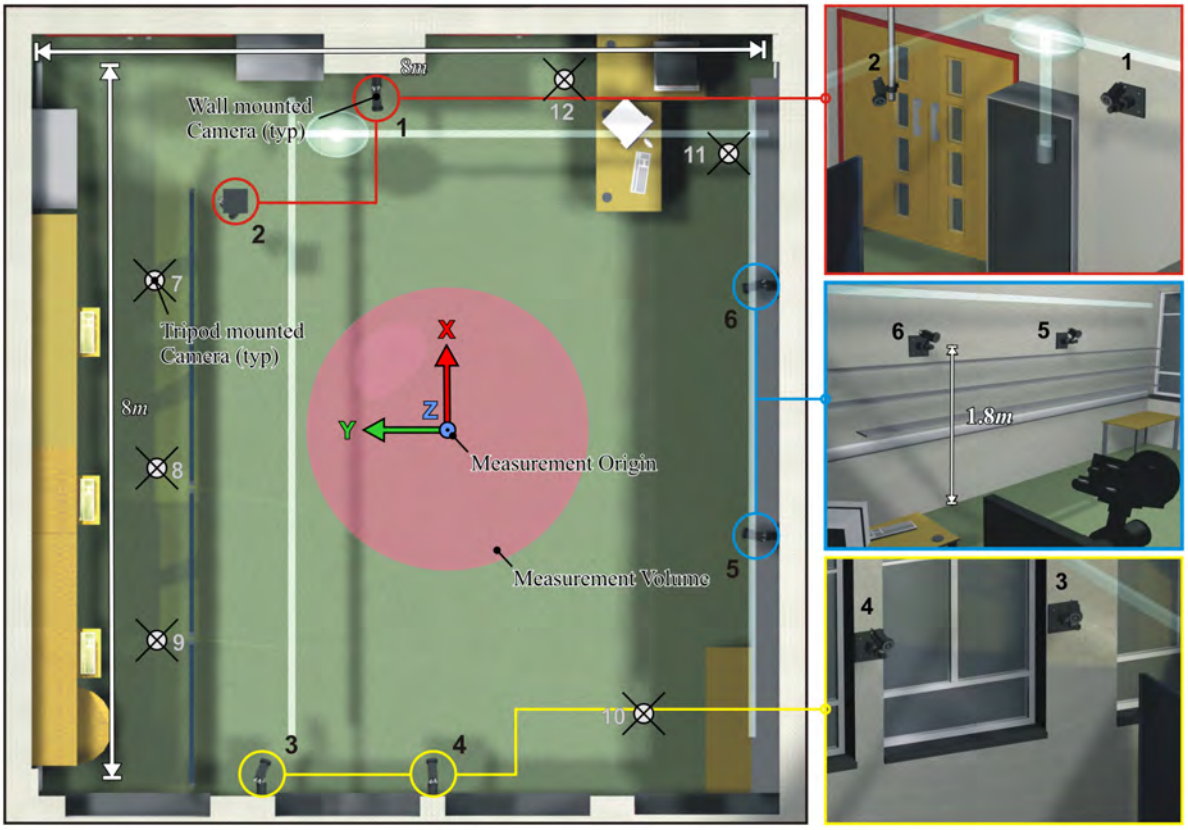


Figure 5.1 – Measurement configuration for pilot in the posture and balance lab

the co-ordinate frame was defined using the standard 750*mm* L-frame at the centre of the lab's floor space. Average calibration residual for all cameras was found to be approximately 0.5*mm*. Measurement was conducted at a frame rate of 120*Hz*, providing a Niquist frequency of 60*Hz*, which is more than adequate given the ~30*Hz* bandwidth of human movements [63].

5.1.2 Implementation of marker set

The full body marker set was implemented using 18*mm* diameter passive markers. This marker size was established experimentally as being visible at the maximum distance from each camera required to establish the desired measurement volume. For skin mounted markers, hypoallergenic double sided tape was used to affix the marker bases directly to the subject's skin. For plate mounted marker clusters, markers were attached to compliant carbon fibre plates supported by elasticated Velcro bands about the circumference of each limb. For the hip markers two smaller plates were used, curved to conform to the ridge of the pelvis. A carbon fibre pylon was attached across each of these plates projecting two 18*mm* markers away from the body. Finally, the head cluster was composed of elastic belt worn around the forehead.

5.2 Experimental Protocol

5.2.1 Source of appropriate test data

To ensure adequate calibration of the body model it was required that any motion data should exercise a good range of motion about the majority of body joints (6.3.2). Additionally, these motions should be regimented and repeatable, to allow for conditions to be maintained across multiple experimental sessions. A review of previous studies of accuracy and occlusion in optical motion capture revealed that typical reference movement sequences were strongly typed athletic movements such as karate [10], repetitive "daily routine" movements [16] or general range of motion exercises [45]. Although these activities offered some of the characteristics required by this project they did not satisfy the range and variability of motion required.

Figure 5.2 - Poses of Tai Chi Chuan
[Adapted from http://www.chinashaolins.com]

After consideration of various alternatives, the practice of Tai Chi Chuan (TCC) was chosen as the movement reference for this project. TCC originated in China in the 13th century as a "soft" martial art and is generally practiced with the aim of improving health and longevity. Although more pugilistic forms of the practice exist, TCC is most commonly practiced without an opponent or partner [64].

From a technical standpoint TCC consists of a relatively small set of body and limb postures (approx 30, depending upon style), which are connected via flowing movements to form complex patterns (Figure 5.2). With reference to the requirements of this project, TCC presented a number of appealing characteristics as a motion reference. These are discussed below:

- Slow to moderately paced movement: TCC is performed at a moderate pace with few impacts or uncontrolled accelerations. This limited the required frame rate for optical measurement and ensured that the dynamic ranges of inertial sensors were respected.

- Range of movement: TCC movements exercise a good range of motion for each joint of the body. An exception to this is the motion of the spine, which is kept straight during practice. However, this is advantageous considering the simplified model of the neck and torso (4.2.1).

- Fluid movement: Movements in TCC typically follow smooth trajectories. This allows for optimal performance of spline interpolation, which is used as a benchmark for assessment of reconstruction techniques (13.3.5).

- Repeatability: TCC practitioners strive to perfect their practice against an ideal. This combined with its slow and controlled pace make TCC motions extremely repeatable when performed by skilled practitioners.

- Segmentation: TCC consists of a sequence of well defined poses connected by smooth motions. Therefore, it is possible to segment each sequence by identifying points of minimum velocity, which is advantageous for analysis.

5.2.2 Recruitment of subject and discussion of patterns

Following research into the popularity of TCC in the local community several associations and societies were identified. Among these was the “Kai Ming Association” [65], whose literature indicated previous experience collaborating with the scientific and medical community through studies of TCCs role in fall prevention. An appointment was made to participate in one of the local workshops and further evaluate the suitability of TCC to the project. During this session the class instructor gave a comprehensive demonstration of the general repertoire of movements encompassed by TCC. The aims and experimental requirements of the project were explained alongside a simple demonstration of motion capture technology. Response to the session was positive and a date was set for a preliminary meeting and pilot experiment.

During the preliminary meeting the requirements of the motion sequences were discussed with the participant and the protocol for the pilot experiment was established. It was decided that the experiment should consist of two repeated blocks of four sequences: The first sequence was a set of general range of motion exercises of each joint of the body. This served as a subject calibration and fallback, had the TCC patterns failed to provide adequate range of motion for testing. For the remaining three sequences the following requirements were presented to the participant:

- Patterns should have durations between 30 and 60 seconds.
- Each motion should be moderately paced and extended stationary periods should be avoided.
- The sequences should exercise a large range of motion over as many joints as possible.
- The patterns should demonstrate varying levels of complexity. It was suggested that one should be performed on the spot, another with transverse footwork motion and the last with complete freedom of motion.

Based upon these requirements the following patterns were agreed upon for the pilot experiment:

Sequence 1: "Preparation to single whip"

A slow sequence performed on the spot. Small steps are followed by several upper body rotations and sweeps of the arms.

Sequence 2: "Fist under elbow to toe kicks"

Fore and back transverse footwork, while circular reaching motions are performed with alternate arms. Several short kicking motions complete the pattern.

Sequence 3: "Four corners to lotus kick"

The feet follow a circular path about four points. Circular hand movements follow into low sweeps of arms. The pattern finished in a slow kicking motion.

5.3 Pilot experimental session

5.3.1 Execution of experiment

Before data collection, the requirements of the experiment were reiterated to the participant and informed consent was taken. It was made clear that all data would be used anonymously and that the participant was free to withdraw from the experiment at any time. The subject was then allowed to conduct his warm up routine while the measurement equipment system was calibrated. This calibration was conducted using the approximate centre of the laboratory as the global measurement origin, with x and y axes parallel with the floor and the axis pointing upwards vertically.

Over the course of the session all eight trials were successfully captured. Due to software failure and experimenter error it was necessary to repeat two trials, however, this had little consequence for the resultant dataset. A cursory reconstruction and examination performed during the experiment showed that all trials produced usable data. Considerable occlusion was noted for the markers located upon the hands, feet and sternum. However, these problems were temporarily rectified for the pilot dataset through spline interpolation.

5.3.2 Post-processing of raw data

Reconstruction and labelling of marker data was performed using the proprietary QTM software, shown in figure 5.3. Following epipolar 3D reconstruction of the raw 2D camera data the first sequence of motion data was manually labelled. Following this, QTM's Automatic Identification of Markers (AIM) system was used to partially label the remaining sequences based on this data. Despite this feature and the good measurement conditions of the lab, all sequences required some manual labelling before the results were considered

satisfactory for further use. Finally, transient and ghost markers were deleted from the sequences and spline interpolation was applied to fill gaps of up to 60 frames (0.5 seconds). The resultant trajectory data was exported as several tab delimited ASCII text files to facilitate simplified import into MATLAB.

Although 0.5 seconds represents an unusually large interpolation duration, given the limited flexibility in placement of cameras, it presented the only available means of recovering the occluded trajectory data. Additionally, as the pilot dataset would be predominantly used for development of the kinematic model, the quality of the trajectory data at this stage of the project was of reduced importance. For the primary experiment adaptations were made to the camera configuration to reduce the extent of occlusion and eliminate the need for spline interpolation entirely.

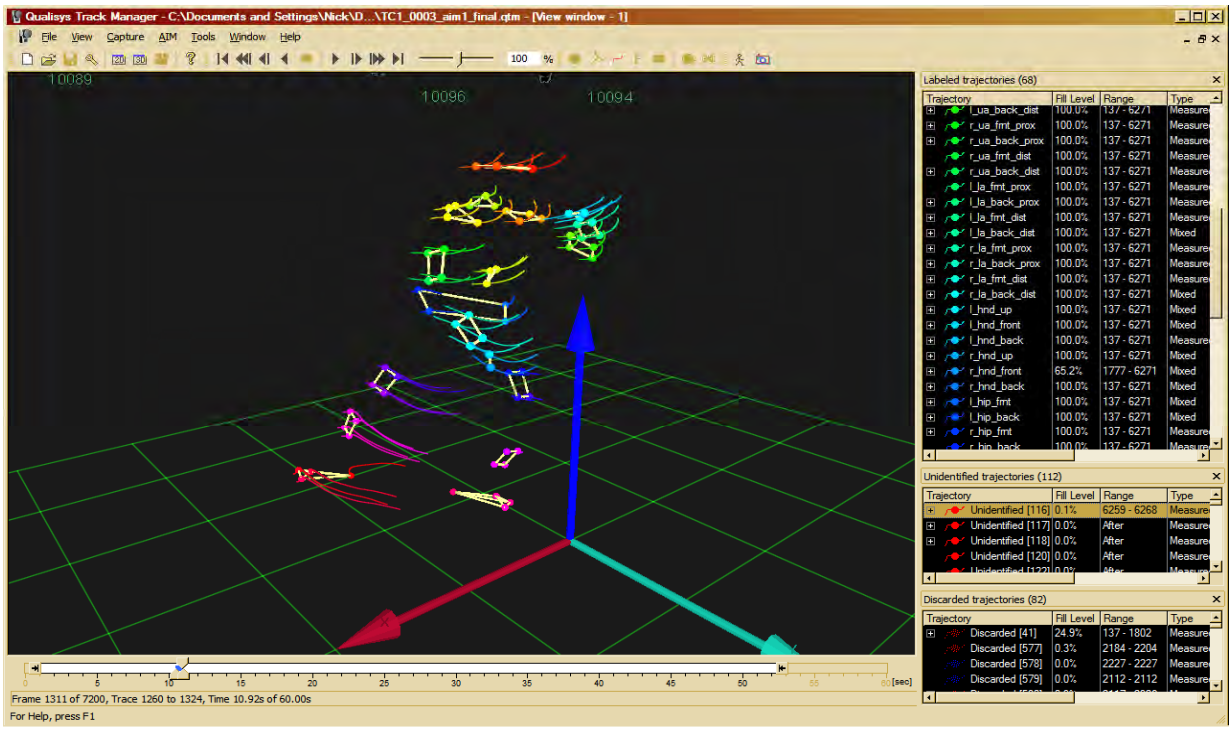


Figure 5.3 - QTM user interface

6. CALIBRATION OF THE KINEMATIC BODY MODEL

The body model used in this project represents a limited approximation of the actual anatomical structure of the human body. Therefore, regardless of the accuracy with which its defining parameters are estimated, deviations will be observed from the actual pose of the subject. However, as the primary aim of this thesis is to estimate the trajectories of body mounted markers from orientation estimates, its behaviour and performance will be characterised solely in this capacity. This chapter describes the implementation of the body model from marker trajectory data, tests its performance and discusses the suitability of the model with regard to the project's aims.

6.1 Cluster pose estimation

From the raw trajectory data collected in the pilot experiment, it was necessary to establish the pose of each marker cluster against the global coordinate system defined by the measurement system. As well as providing the basis by which the body model was calibrated and tested, this pose data was used extensively throughout the remainder of this thesis.

6.1.1 Initial data segmentation

Trajectory data obtained in the pilot experiment was imported into Matlab as tab delimited ASCII data and stored as a structure type. In addition to a 3D array containing the raw co-ordinate data organised in *dimension* x *marker* x *sample number* format, this structure contained the labels assigned to each marker, sample rate and acquisition duration.

Following data import, trajectory data was grouped according to each marker's parent cluster. This association was established via an external configuration file containing the name of

each cluster and its constituent markers as defined in (4.4). The format of this file was simple, with cluster definitions preceded by an underscore and followed by a list of marker names. An example of the cluster definition for the head segment is shown below:

```
_head_seg — Segment Name
      l_head_frnt
      r_head_frnt      } Constituent Marker Names
      r_head_back
      l_head_back
```

The raw trajectory data structure was sorted via this configuration file to produce a 4-D array containing all trajectories in terms of their parent cluster. By assigning each parent cluster a numerical index (1 to 18) and each constituent marker a sub index (1 to 4) this array was organised as *dimension* x *cluster index* x *sub index* x *sample number*. This array and the cluster-marker associations defined in the configuration file were stored in a new structure designated the 'locked data set', which provided storage for all further body model parameters (discussed in 6.4.2) and data. Through this structure, individual marker trajectories can be indexed hierarchically in terms of their cluster membership and associated with the body segment to which they are attached (Figure 6.1).

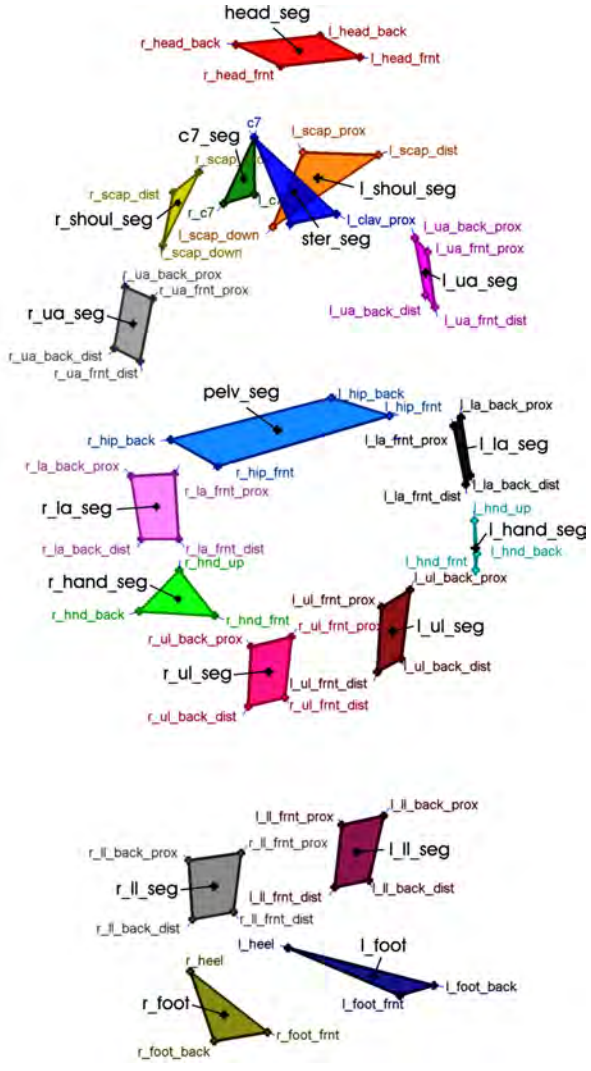


Figure 6.1 – Segmentation of Marker Set (Primary experiment)

6.1.2 Pose extraction and testing

The pose for each marker cluster is represented by the transform between the axes of the global co-ordinate system at the measurement origin and a local co-ordinate system whose origin lies at the centroid of that cluster. The *x* and *y* axes of this local system are aligned with the plane formed by the segment's constituent markers, and the *z*-axis is perpendicular to this plane with a clockwise convention.

This transform is defined by finding the translation and rotation, which relate the positions of a measured cluster's ($\mathbf{p}_{r_{1..4}}$) markers and a fixed 'reference' cluster ($\mathbf{p}_{l_{1..4}}$). The orientation of this reference with respect to the global axes specifies the alignment of the local axis system of the cluster. Assuming that both these sets of coordinates are geometrically identical, the translation between them is the vector $\boldsymbol{v_t}$ relating their centroids ($\boldsymbol{c_r}$ and $\boldsymbol{c_l}$). Additionally, by translating each set by subtraction of the vectors $\boldsymbol{v_l}$ and $\boldsymbol{v_r}$, it is possible to reference each set to the global origin, providing a 'pivot' point about which orientation can be defined (Figure 6.2).

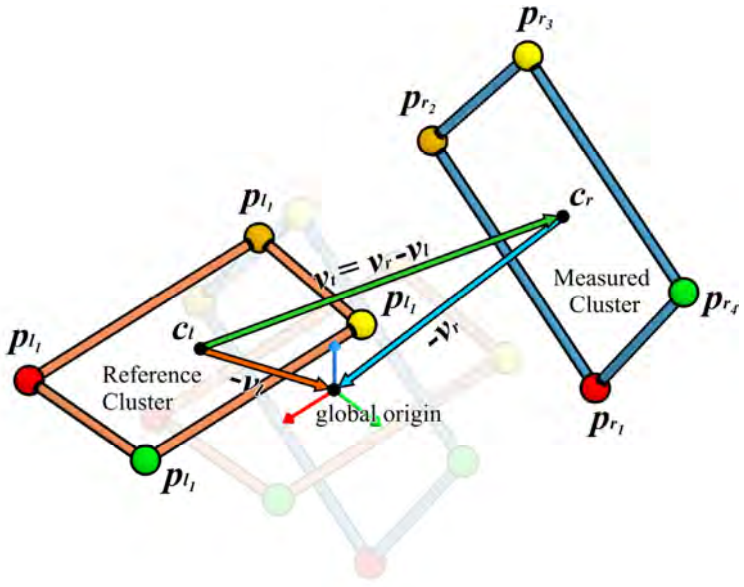


Figure 6.2 – Translation between centroids of two clusters

As shown in figure 6.3, the centroid represents the exact geometric centre of a group of points. Therefore, its use as the origin for each cluster ensures the even influence of each marker position over the total translation of the set. This is particularly important in the case where

coordinate data is inaccurate or prone to noise (as is the case for measurement data). Perturbations of individual points within the set will tend to average out over the group, leading to a reduction in overall positional error. A similar advantage is obtained in estimation of orientation, due to each point being approximately equidistant from the pivot and therefore possessing an equal moment [40].

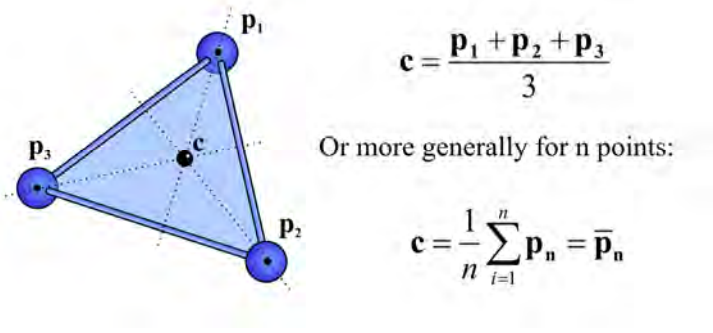


Figure 6.3 – Calculation of a cluster's centroid

6.1.3 Generation of reference cluster

To define the reference cluster it is necessary to generate a set of coordinates around the global origin, which are geometrically identical to each measured cluster. As the position of these points relative to the global axis system describes the alignment of the local axis system of a cluster, it is essential that the orientation of the reference cluster can be precisely controlled.

Although the majority of the clusters in the marker set share a similar rectangular layout, variation in geometry due to manual attachment of markers makes each cluster unique. Therefore, each reference must be defined for each cluster individually. To allow for good correspondence and precise control of orientation, the reference is established using the angular relationships between markers in each measured cluster. This process is depicted in figure 6.4, where simple trigonometry is used to estimate the angle of each marker about, and its distance from the centroid. These parameters are used as polar co-ordinates to generate a plane aligned and origin centred reference cluster using a clockwise convention about the z–

axis. These reference axes can be aligned with the segment axis conventions defined in (4.2) by arbitrary rotation.

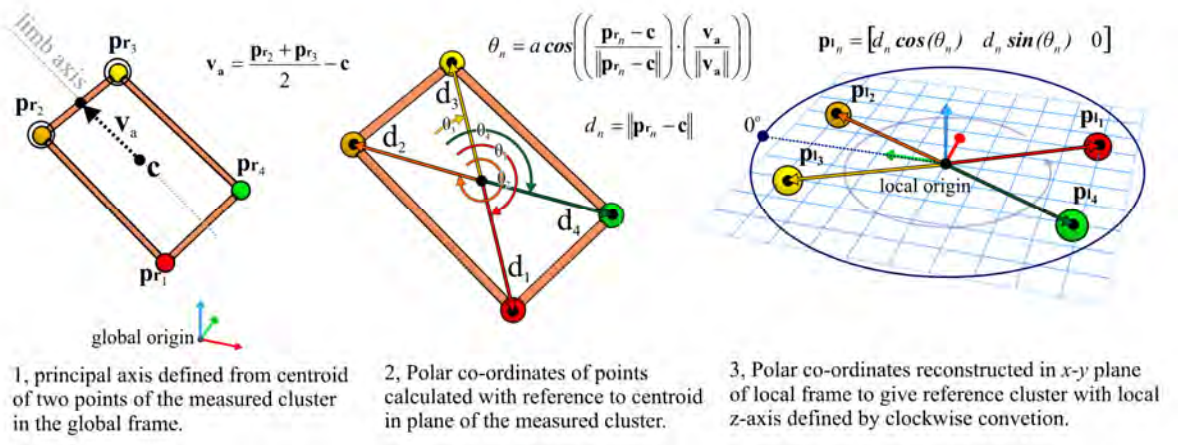


Figure 6.4 – Generation of a planar reference.

The enforced planarity of this reference makes this method unsuitable for use with extremely non-planar cluster configurations. However, as all the clusters used in this project have a planar or near planar configurations, this was not found to result in reduced accuracy of pose estimation.

6.1.4 Orientation estimation

The rotational transform relating the reference and measured clusters can be estimated by the minimisation of the cost function of the error between the reference and measured marker sets for a rotation ($\mathbf{R}_{lr}$). Examples of typical cost functions are given below for the co-ordinate sets $\mathbf{p}_{l_n}$ and $\mathbf{p}_{r_n}$, (where n is the marker index within the cluster) representing the cluster co-ordinates after subtraction of the centroid. The first is expressed via the norm in terms of the sum of Euclidean distance between pairs of markers (e_d) and the second using the dot product to give the average angle between corresponding vectors to markers from the cluster origin (e_θ):

$$e_d = \sum_{i=1}^{n} \left\| \mathbf{p}_{\mathbf{r}_n} - \mathbf{R}_{\mathbf{lr}} \mathbf{p}_{\mathbf{l}_n}^{\ T} \right\| \quad e_\theta = \sum_{i=1}^{n} \left(\mathbf{R}_{\mathbf{lr}} \mathbf{p}_{\mathbf{l}_n}^{\ T} \right) \cdot \mathbf{p}_{\mathbf{r}_n}$$

Although it is possible that both these functions could be minimised by standard least squares techniques, a constraint upon any valid solution is that $\mathbf{R}_{\mathbf{lr}}$ must represent a valid rotation and must therefore be orthonormal (Assuming rotation is represented in matrix form). This complicates the minimisation of either cost function and therefore the estimation of orientation.

The ubiquity of the problem of estimation of absolute orientation over three dimensional photogrammetric applications has led to a large number of proposed solutions to the minimisation problem given above. A review of the four of the most popular methods is given in Lorusso et al. [66] where a direct comparison is made in terms of accuracy and computability in a variety of applications. Based on the findings of this paper, the method first described in Horn [40] was selected as the means of orientation estimation to be used in this thesis. In addition to giving superior performance in terms of accuracy and computability, an attractive aspect of this method is its use of the unit quaternion representation of rotation (Appendix 10). This representation is noted to be inherently more stable than the conventional orthonormal matrix due to the simplicity, by which unit magnitude of a quaternion can be guaranteed, in contrast to the difficulty in ensuring orthogonality of a matrix. Furthermore, efficiency of storage and computation is improved due to the 4 vector format of a quaternion versus the 3x3 equivalent rotation matrix. Finally, the unit quaternion can easily be transformed into matrix, axis-angle or Euler representations using a reduced number of

operations than required by alternative representation of rotation. This is a desirable feature as these conversions are used frequently throughout this thesis.

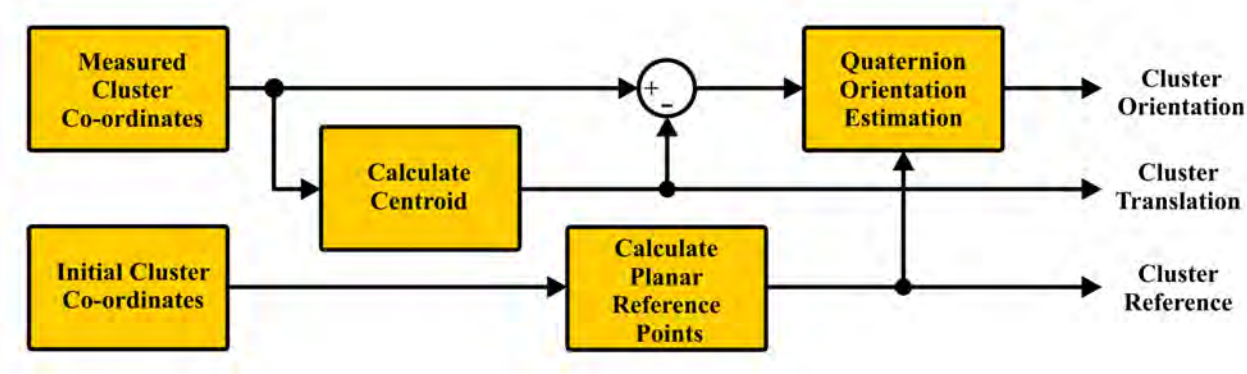


Figure 6.5 – Schematic of Pose Estimation Function

Combining the quaternion orientation estimation, centroid subtraction and reference generation results in the final pose estimation algorithm. Its basic structure is given in the block diagram of figure 6.5, depicting the flow from cluster position measurements to pose estimates for a single frame. The algorithm was implemented in Matlab based around the working given in [40]. Some deviation from the original implementation was necessitated by the observations of Lorusso [66], that greater accuracy is obtained through the use of an iterative solution to the final eigen-parameters. The full code listing of the algorithm as implemented for the body model is given in (Appendix 1).

6.1.5 Full body orientation estimation

The pose of each body segment in the marker set was calculated by sequentially applying the pose orientation estimation algorithm to each frame, cluster-wise of the segmented co-ordinate data of the locked data set. This is shown in figure 6.6 for a single frame from the pilot data set. The unit quaternion orientation estimates were converted to rotation matrices and superimposed upon the clusters centroids to give an impression of the local cluster axes. As the alignment of local axis systems provided by the method of (6.1.3) did not correspond to those defined in the body model (4.2), the relative rotational transforms between these two systems were provided by an external ASCII configuration file.

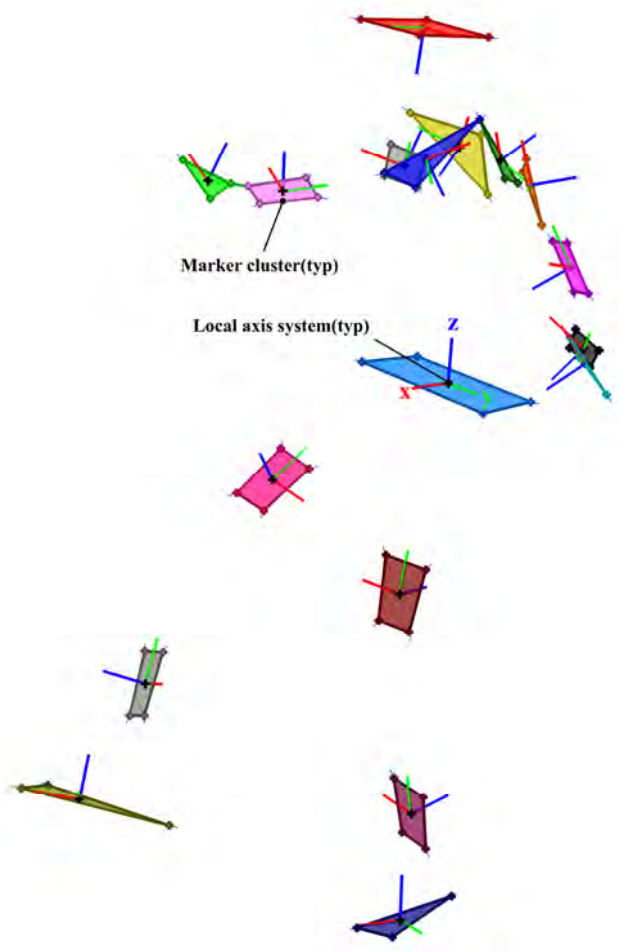


Figure 6.6 – Momentary Cluster Pose Estimates

6.2 Testing of the pose estimation algorithm

Baring invasive techniques such as bone pins [67] or medical imaging [68] there is no method by which the underlying pose of human skeleton can be directly measured. Therefore, it was necessary to test the pose estimation algorithm via a combination of simulation and other indirect techniques. These are described in this section and the results of their application discussed.

6.2.1 Deviation of orientation with marker position error

To assess the effects of marker measurement noise upon estimation of orientation, a 4 point cluster was simulated with dimensions approximating the lower size limit of the full body marker set (60*mm* x 80*mm*). This cluster was then transformed by 200 uniformly distributed randomly generated rotations and pose estimation was performed. Sensitivity to positional error was assessed by repeating this process while a randomly generated positional offset was applied to each transformed co-ordinate. Scaling this offset between 0*mm* and 15*mm* over 200 steps generated the deviations in orientation estimation shown in figure 6.7. Average error in orientation is shown by the solid black line and the range of one standard deviation by the surrounding pink area. Orientation error was calculated as the average angle component of the axis-angle transform (Appendix 9) relating each of the noise to no-noise conditions over the 200 orientations.

An approximately linear increase in orientation error is observed over the range of positional errors expected to occur due to residual noise and soft tissue artefact (4.1.4). This suggests the method will provide a graceful falloff in performance with increasing marker measurement noise when applied to the estimation of body segment pose.

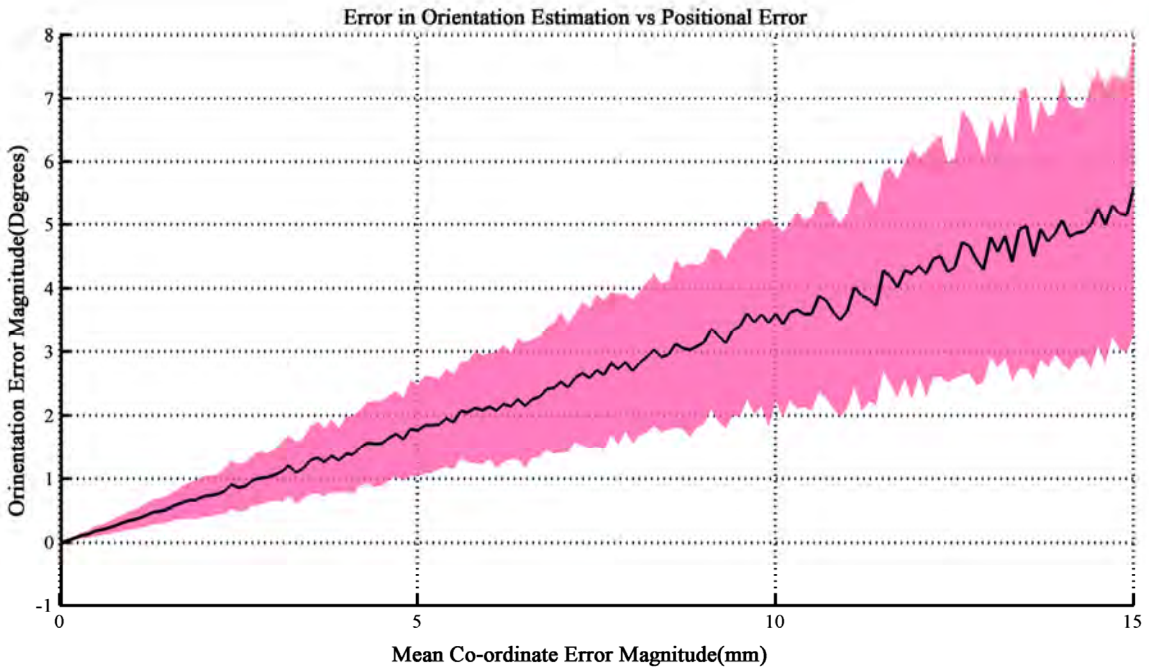


Figure 6.7 – Orientation Error vs. Positional Noise

A second parameter, which influences the accuracy of orientation estimation is cluster size (4.3.1). The effect of this variable is examined by repeating the previous simulation and varying the average cluster radius over five values of positional error. The resultant relationship is shown in figure 6.8, with each solid line representing the change in orientation error for increased cluster radius.

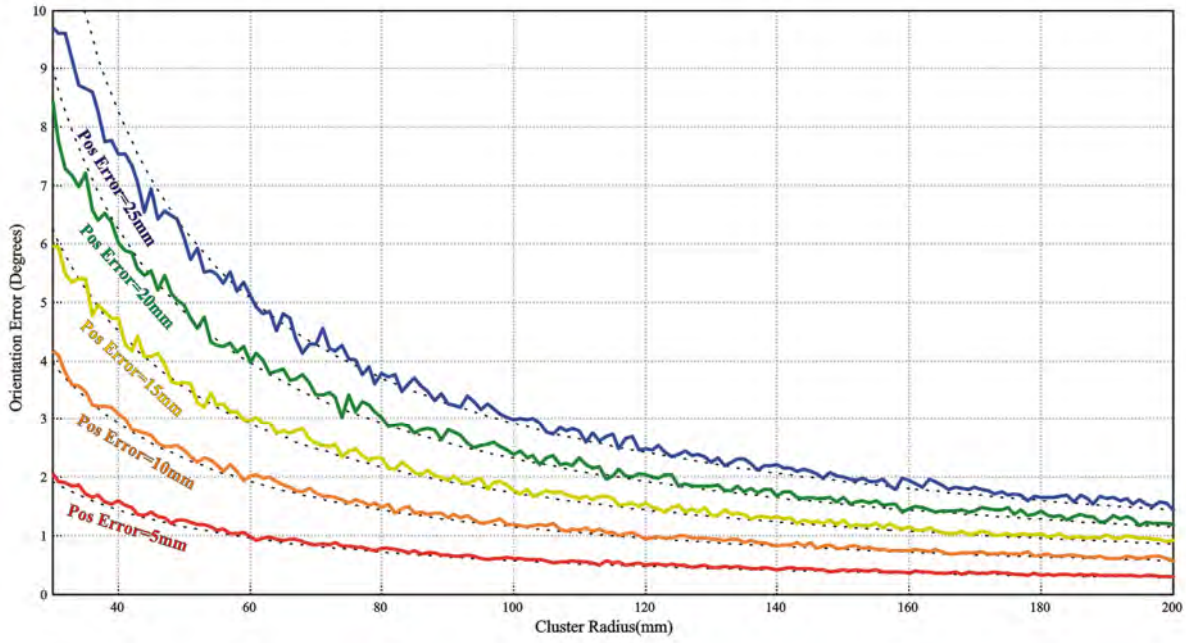


Figure 6.8 – Orientation error vs. cluster radius

Unlike positional error, the relationship between cluster radius and orientation error is clearly not a linear function. However, through consideration of the orientation estimation operation in terms of simple trigonometry, the relationship between cluster size, measurement noise and orientation can be modelled by the 'moment' of each marker about a pivot formed by the centroid. Simplifying the situation to two dimensions, the perturbation of marker position above or below the plane of the cluster represents the opposite side of a right angle triangle and the cluster radius the adjacent. From figure 6.9, it is shown that the angle of this triangle is given as the arc tangent of the ratio between the two sides and can be considered to represent the error in orientation.

$$\theta_{error} = \tan^{-1}\left(\frac{P_{error}}{Radius}\right)$$

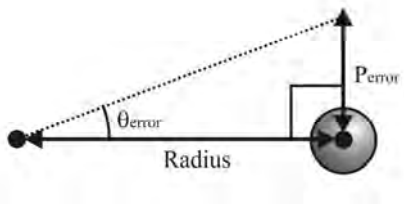


Figure 6.9 – Tangent error

Referring back to figure 6.8, the broken black lines represent this function fitted to the simulation data. The fit was adjusted by scaling the arc tangent by a factor of ~0.2 chosen to minimise mean square distance from the observations. This factor is required to approximately compensate for positional errors, which are not directly perpendicular to the plane of the cluster and are therefore not accounted for by the simple trigonometric model.

From this model it is possible to infer that, given the average positional error discussed in (4.1.4) of ~10*mm*, if an angular accuracy of ~3^o is to be maintained then clusters must have an average radius of at least 40*mm*. This criterion is met by the majority of markers in the full body marker set apart from those upon the hands.

6.2.2 Stability of orientation estimation for human movements

To assess the performance of the pose estimation algorithm when applied to real data, an alternative method of inferring orientation error was devised. This makes use of the previously defined error model, cluster size and the deviation of a cluster from rigid body motion to estimate orientation error. A block diagram of the operation of this method is shown in figure 6.10.

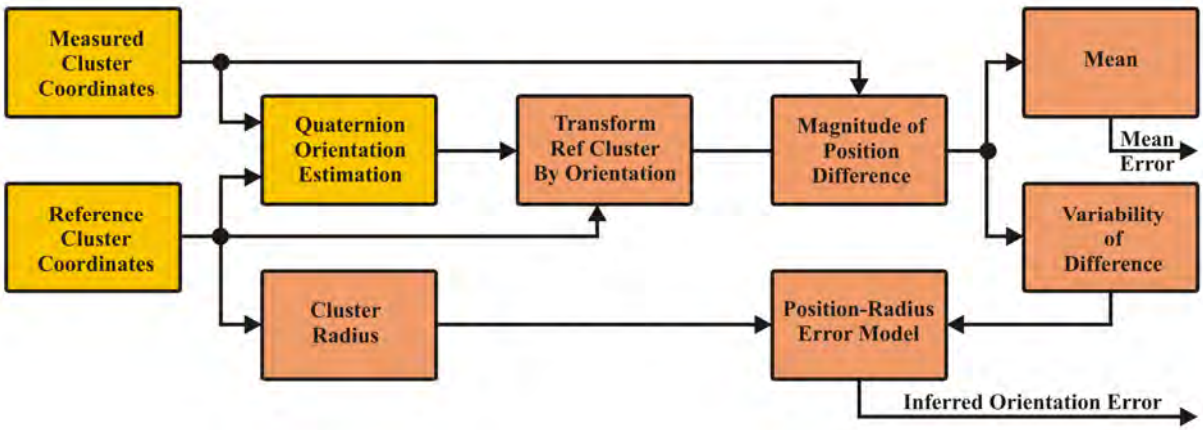


Figure 6.10 – Inferred orientation error model

The basis of this error metric is that any measured cluster co-ordinates can be transformed to the frame of the reference cluster using the inverse of their orientation estimate. Assuming perfect rigidity and therefore correspondence between the geometries of both measurement and reference clusters, the distance between corresponding markers in these clusters should be zero. However, due to the errors in both orientation estimation and measurement error, some deviation will be observed between the two. As the error introduced by the orientation estimation is minimal in comparison with that due to measurement, by subtracting one cluster from the other the measurement error per marker is approximated. By taking the variance of this parameter over a short sample window, the contribution of any constant difference between the two sets of points is eliminated, leaving only the transient components of measurement error. Applying the trigonometric orientation error model with the cluster radius provided via the reference cluster, total orientation error can be inferred over the course of measurement. Figure 6.11 shows the orientation and position error obtained by application of this method to the pilot ‘range of motion’ trial.

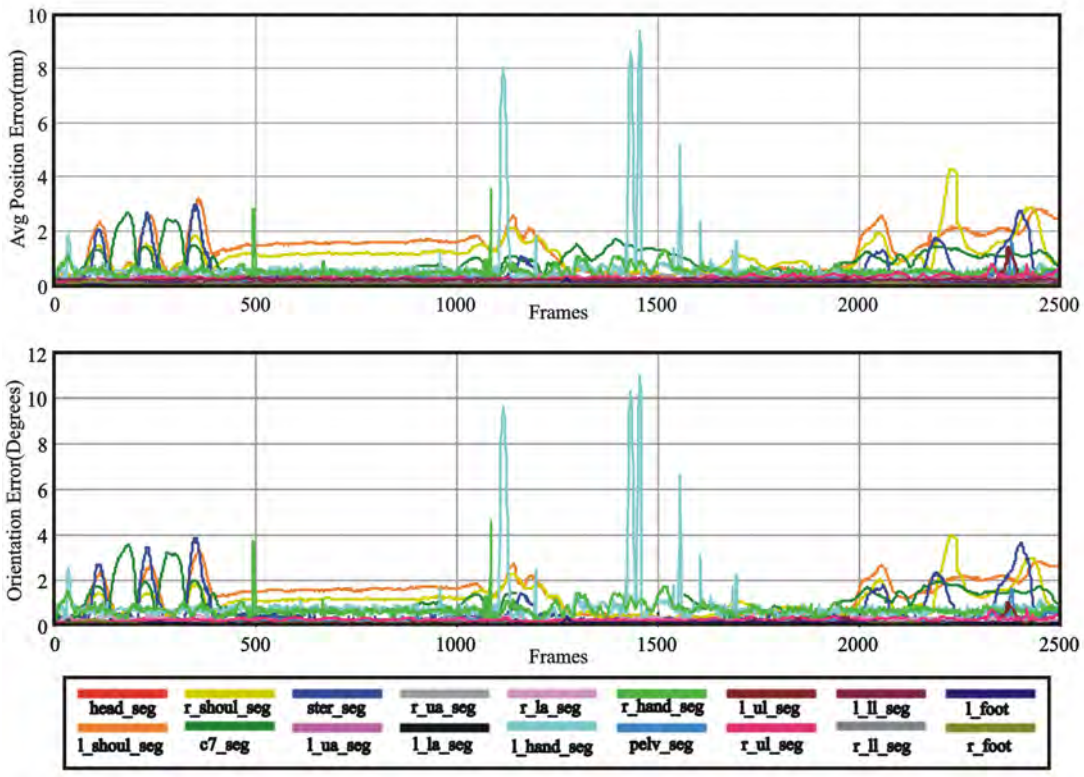


Figure 6.11 – Graph of inferred orientation error

This figure indicates that both deviation of marker position and inferred error are relatively well behaved, presenting positional deviations averaging ~1*mm* and equating to an orientation error of less than one degree. However, in the case of a few clusters sporadic errors of considerably larger magnitude are noted. This is better represented by the data in figure 6.12, which shows the orientation error magnitude in terms of the mean, maximum and standard deviation.

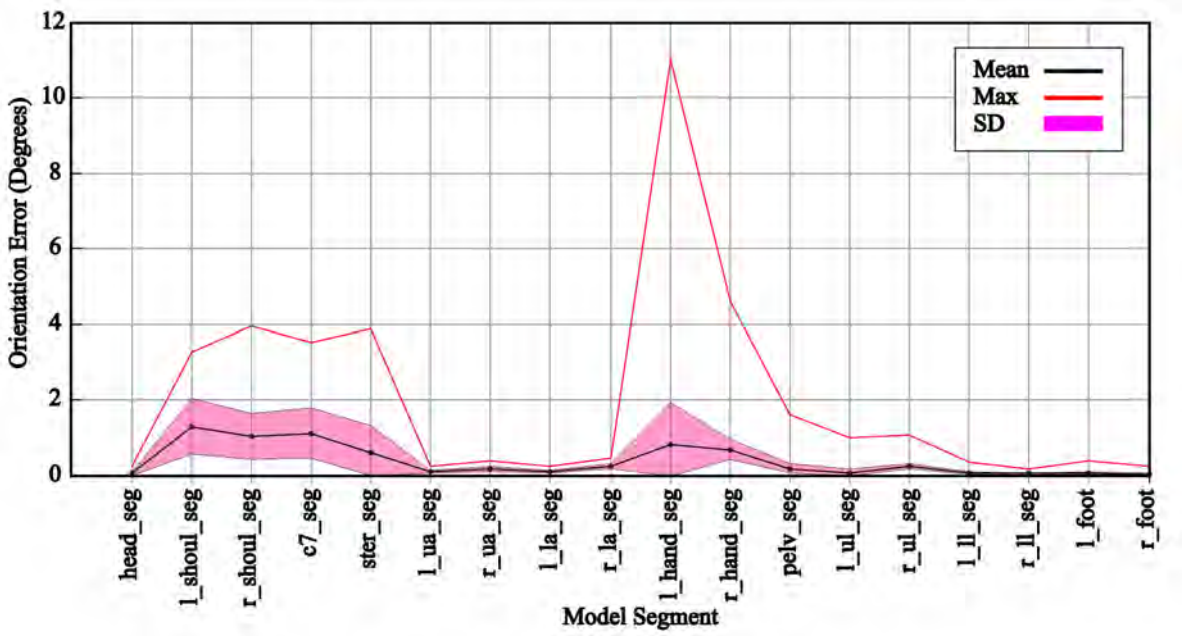


Figure 6.12 – Average orientation error for full body

From this data it is noted that clusters implemented as rigid plates, or attached to more rigid body segments such as the head arms and legs, produce errors around the level of optical reconstruction residual (<1*mm*). However, for clusters composed of skin mounted markers, such as the shoulders, torso and hands, error levels are considerably higher. For the shoulders and torso this level of error is expected, as the separate bones of the segment allow for large relative motion. However, in the case of the hands, the average cluster radius is considerably lower than the established minimum of 40*mm* and leads to exaggerated levels of error. Refinements to address these limitations for the primary experiment are discussed in (10.1.2).

6.3 Extraction of centres of rotation and testing

6.3.1 Isolation of joint centres in local frame

To estimate the functional centres of rotation in the local co-ordinate frame of each body segment, the inverse of the estimated cluster pose is used to selectively transform a ‘target’ cluster and its neighbour to the local co-ordinate frame of its reference cluster. This effectively isolates the target in space, referencing the trajectories of the neighbouring cluster to the local co-ordinate system of the target cluster. These trajectories are examined with the aim of identifying common orbits occurring around a fixed point in the local space. The existence of such orbits over the range of motion of a cluster is considered to infer the presence of a spherical rotational (or 3R) constraint between the segments. Therefore, by resolving the centre of this constraint, a single functional joint in the body model is defined. To obtain this functional centre in terms of the global measurement frame, and hence its context within the body model, transformation by the uninverted pose of the target cluster is performed.

Figure 6.13 illustrates this local transformation applied to the upper arm cluster for a short range of measurement. In the case of the globally referenced measurement data, little relation is observed between the trajectories of the upper and lower clusters. However, the same data transformed to the local reference frame of the upper arm segment clearly indicates the presence of a spherical orbit about a fixed point. This point defines the functional centre of rotation (CoR) of the elbow.

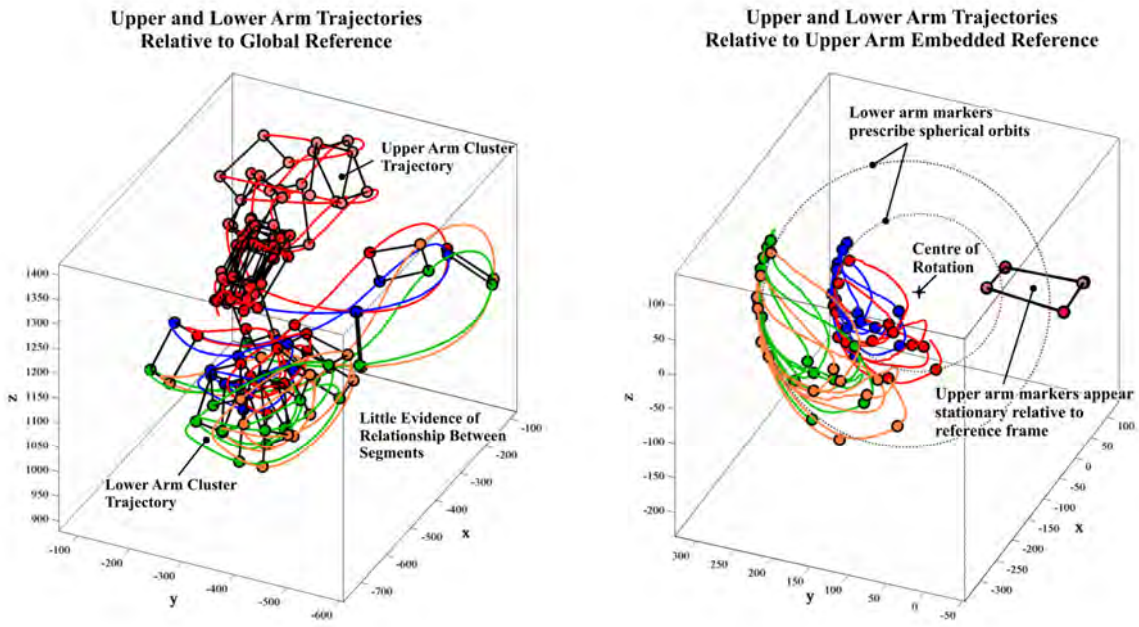


Figure 6.13 – Global and locally aligned trajectory data

6.3.2 Estimation of functional joint centres

Following transformation into the local frame, a method is required by which the fixed orbit relationships between marker trajectories can be identified and their centre estimated. This estimation is achieved through consideration of the mechanical behaviour of the 3R joint.

The idealised form of the 3R joint is depicted in figure 6.14. It consists of two arbitrary rigid bodies: A fixed parent segment and mobile child segment, connected by a mobile ball within a socket. By exercising the range of motion (RoM) of this joint and observing the trajectory of any

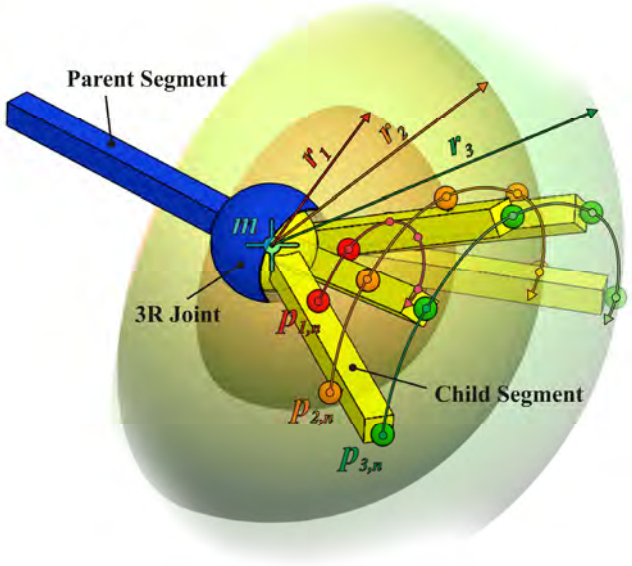


Figure 6.14 – Relationships in 3R joint

point upon the surface of the child segment a spherical orbit with a fixed radius will be observed. The joint produces a set of trajectories over *N* successive samples as it exercises its range of motion with *K* points attached to the surface of the child segment. If $\boldsymbol{p}_{k,1}$,.. , $\boldsymbol{p}_{k,N}$ represents the spherical trajectory of the k^{th} marker (where k = 1 to K) with radius r_k from the CoR ***m***, for a set of trajectories to satisfy the requirements of a perfect 3R joint, the following expression must hold (where *n* is the sample index and *k* is the trajectory index):

$$\sum_{k=1}^{K}\sum_{n=1}^{N}\left(\sqrt{\left|\mathbf{p}_{k,n}-\mathbf{m}\right|^2}-r_k\right)=0$$

However, due to the influences of measurement noise, marker movement and imperfections in the construction of any real joint, this expression will not hold true in practice. Therefore, by squaring to ensure positive error values a cost function (f_m) for the error between the marker trajectories and the sphere radius is obtained.

$$f_m=\sum_{k=1}^{K}\sum_{n=1}^{N}\left(\sqrt{\left|\mathbf{p}_{k,n}-\mathbf{m}\right|^2}-r_k\right)^2$$

By minimising f_m, for **m** and orbit radii r_k the best fit CoR for the selected set of trajectory points $\mathbf{p}_{k,n}$ will be found. Depending on the actual value of f_m at which this minimum occurs, the existence of a purely rotational joint may be inferred.

Due to the loss of the sign for the subtraction $\mathbf{p}_{k,n}-\mathbf{m}$ within the radicand of the square root, minimisation of this cost function is not trivial. A brute force approach to its solution is given in Piazza et al. [69] where a gradient following iterative search is applied to the minimisation.

However, the inherent inefficiency and tendency to solve for local minima of such techniques makes this method suboptimal. A definitive review of alternative methods by which this problem has been approached is given in Chang et al. [39]. From this work a method was selected for the solution of functional joints in the body model. This technique, adapted from the two dimensional algebraic surface fitting technique of Pratt [70] was shown to provide both superior accuracy and excellent immunity to measurement error. Most importantly, it is shown to produce accurate solutions for even small ranges of motion about joints (RoM). A full code listing of the Matlab implementation of this algorithm can be found in Appendix 2.

6.3.3 Testing of joint centre estimation

Although the performance of the functional joint solution algorithm is well established in the literature [39], testing was considered necessary to verify the method in the context of the body model used in this project.

For initial testing several simulated sets of points were generated approximating trajectories of a cluster about an ideal 3R joint. These points were randomly distributed upon the surface of several spheres lying at different radii from a fixed central point. The solid angle of the spheres arc was varied to control for joint range of motion. An example of such a test data set for a 90^o RoM over four marker orbits is shown in figure 6.15. The predefined centre of orbit is marked with a fine black cross and test coordinate data are shown by the red, orange, yellow and green points. The solution provided by the CoR algorithm, marked in red, exactly corresponds with the centre of orbit.

Unlike real measurement data, this scheme allows for the parameters of RoM and measurement noise to be precisely controlled and their effects on the accuracy of CoR estimation observed. In figure 6.16 the average Euclidean difference between the CoR estimation and orbit centre is plotted as a function of RoM angle between 5^o-180^o. The red plateau of the graph represents errors in excess of 100*mm*. Measurement noise is simulated as a uniformly distributed random perturbation to each coordinate with a peak value ranging between 0*mm* and 10*mm*. In line with typical values in the real marker set, average cluster radius was set to 60*mm*, radius of orbit radius was 100*mm* and the number of markers was set to four. Each trajectory set tested consisted of 400 data points distributed over the 4 marker orbits and 100 fits. These levels of repetition were chosen as they were found to give stable averages over multiple runs of the simulation.

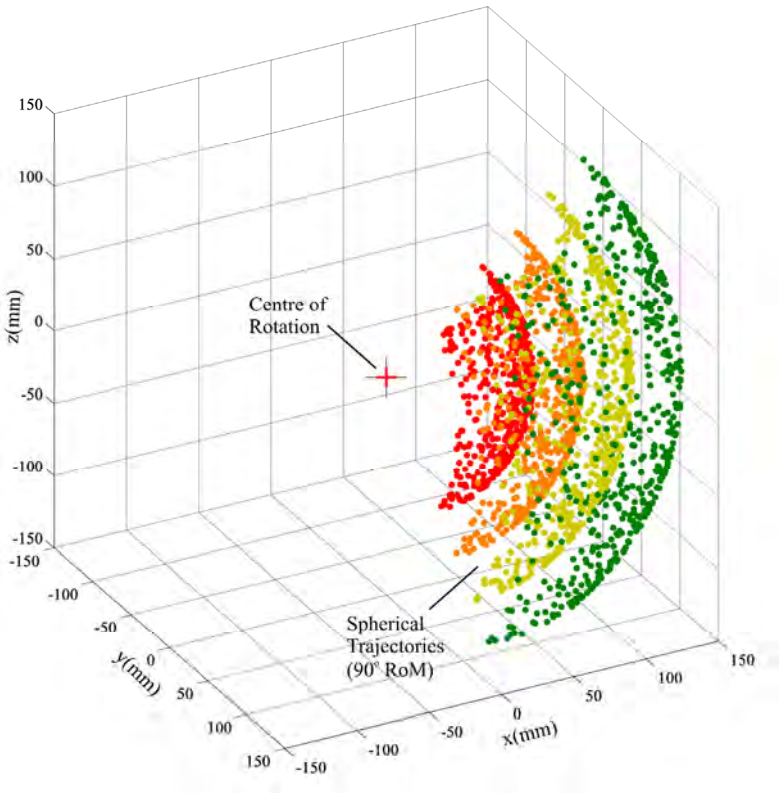


Figure 6.15 – Simulated spherical orbit

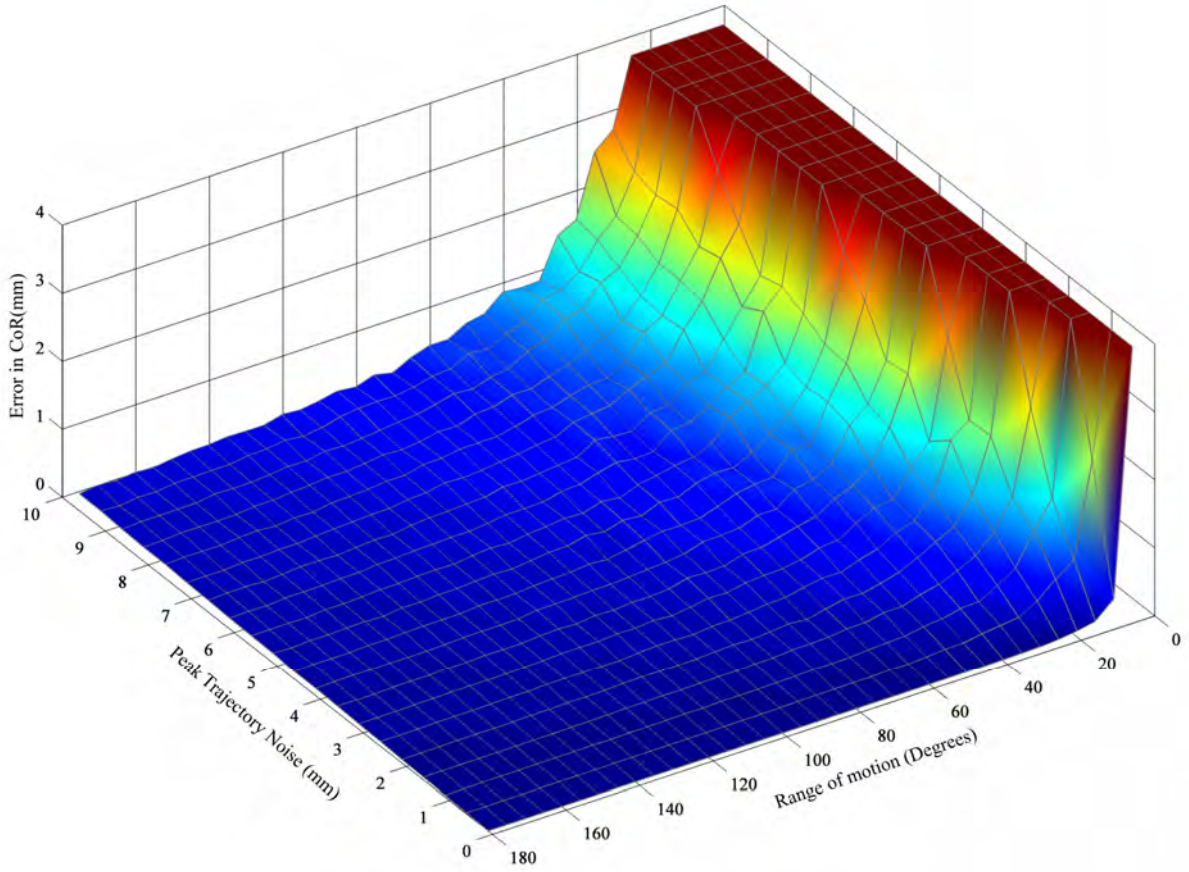


Figure 6.16 – Error in estimation of CoR with varying RoM and positional noise

Under virtually all conditions, the algorithm was seen to produce an extremely accurate fit with less than 1*mm* between actual and estimated CoR Error. However, beyond these stable values, the quality of the fit deteriorates rapidly. It was observed that reduced RoM has the most profound effect upon fit quality, with the range of 20-40^{o} resulting in a sharp increase in error and failure becoming increasingly probable beyond this point.

Therefore, in practical measurements where errors below 10*mm* cannot be guaranteed, a 40^{o} minimum range joint motion should be maintained to ensure accurate estimation of functional CoRs. However, as the movement sequence used to establish these CoRs in the body model

calibration (the ‘range of motion’ dataset of 5.2.2) exhibits deliberately exaggerated RoM about the majority of joints, this limitation does not present a problem.

Two other factors affecting accuracy of CoR estimation are the radius of orbit from the CoR and the number of trajectories used in the estimation. The effect of these parameters was assessed by repeating the previous simulation with values of RoM and position noise fixed at the ‘worst case’ values of 45^{o} and $10mm$ respectively. Orbit radius was varied between $10mm$ and $300mm$ over sets of trajectory data with between 1 and 6 individual orbits. The results of this simulation are shown in figure 6.17, averaged over 1000 fits per sample to accentuate the very slight variation observed.

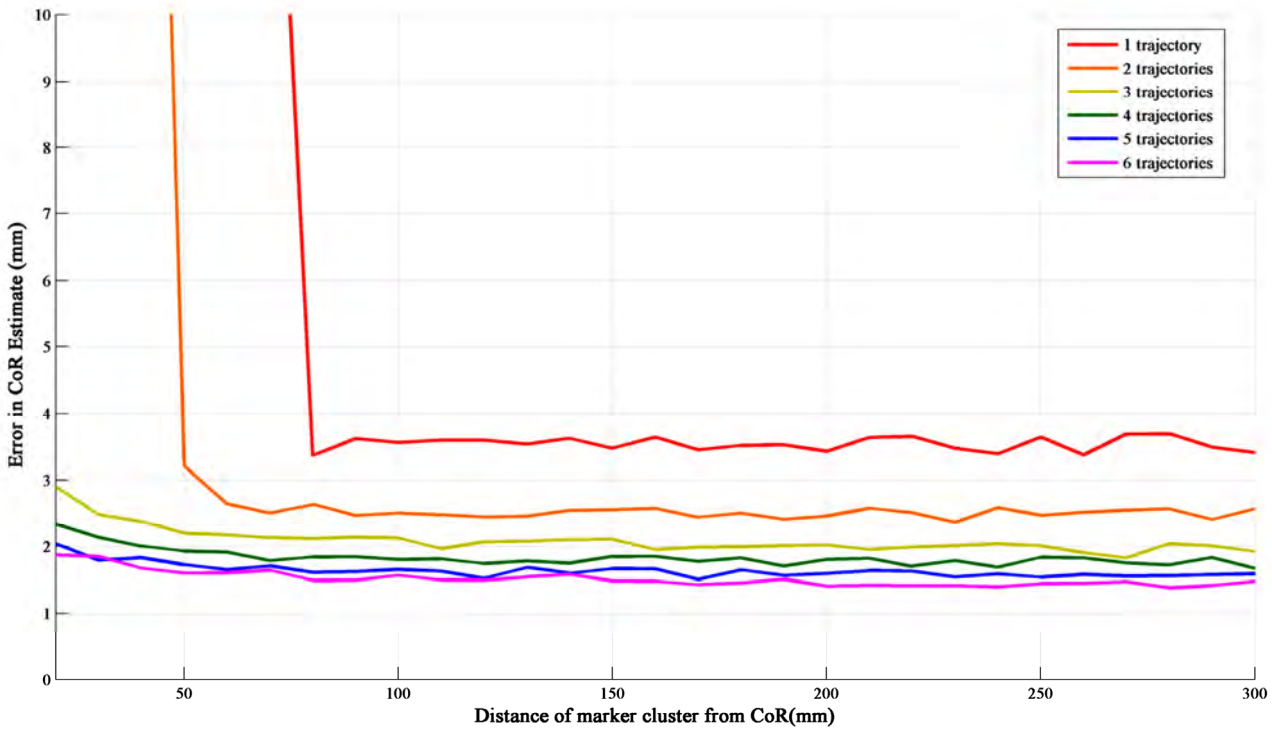


Figure 6.17 – CoR error vs. distance from Joint

From this graph, the dependence of accuracy of CoR estimation upon orbit radius is minimal and is only noticeable for radii of 100*mm* or less. However, the effect of number of orbits is considerably stronger, with one or two point trajectories giving consistently worse results than higher numbers of points. Additionally, fitting is seen to fail for low orbit radii in both one and two trajectory cases. However, in the case of the three and four marker clusters, stable results are observed across the entire range. This combined with the relative lack of improvement in accuracy given by five and six marker clusters, provides further evidence for the suitability of the 3-4 marker clusters chosen for the full body marker set.

6.4 Implementation of final body model.

6.4.1 Implementation of two segment kinematic model

Returning to the example of the upper and lower arm clusters (Figure 6.13), the CoR estimation is now applied to the estimation of the functional centre of the elbow with reference to the local frame of the upper arm cluster. As each marker in the lower arm cluster will orbit this centre at an approximately constant radius, with knowledge of these radii and the cluster's orientation it is possible to estimate the position of these constituent markers in the local frame of the upper arm cluster by forward kinematics. Furthermore, by transforming these local positions by the pose estimate of the upper arm, the trajectories can be brought into the global measurement frame. This simple model allows for the trajectory of any cluster to be described in terms of only its orientation and the pose of a neighbouring 'root' segment. Through this ability it provides the basis of the body model, occlusion model and trajectory reconstruction algorithm.

Figure 6.18 details the working of this reconstruction, allowing the estimated positions of the markers of the lower arm cluster $\tilde{\mathbf{p}}_l$ in the global frame to be obatined. The 'root' pose of the upper arm cluster is described by the quaternion $\dot{\mathbf{q}}_{ru}$ and vector $\mathbf{t}_{ru}$ which are obtained via the pose estimation algorithm. The vector $\dot{\mathbf{v}}_m$ represents the difference between the estimated centre of rotation and origin in the local frame of the upper arm cluster. The vector $\dot{\mathbf{v}}_l$ represents the difference between elbow CoR and centroid of the lower arm marker cluster $\mathbf{p}_l$, obtained through transformation of the global position of elbow CoR into the local frame of the lower arm cluster (6.3.1). Finally, $\mathbf{r}_l$ represents the reference cluster for the lower arm segment (6.1.3) and the orientation of the lower arm segment is described by the

quaternion $\dot{\mathbf{q}}_{rl}$. The diacritic dot for the vectors $\dot{\mathbf{v}}_m$ and $\dot{\mathbf{v}}_l$ indicates unit quaternion representation (Appendix 10).

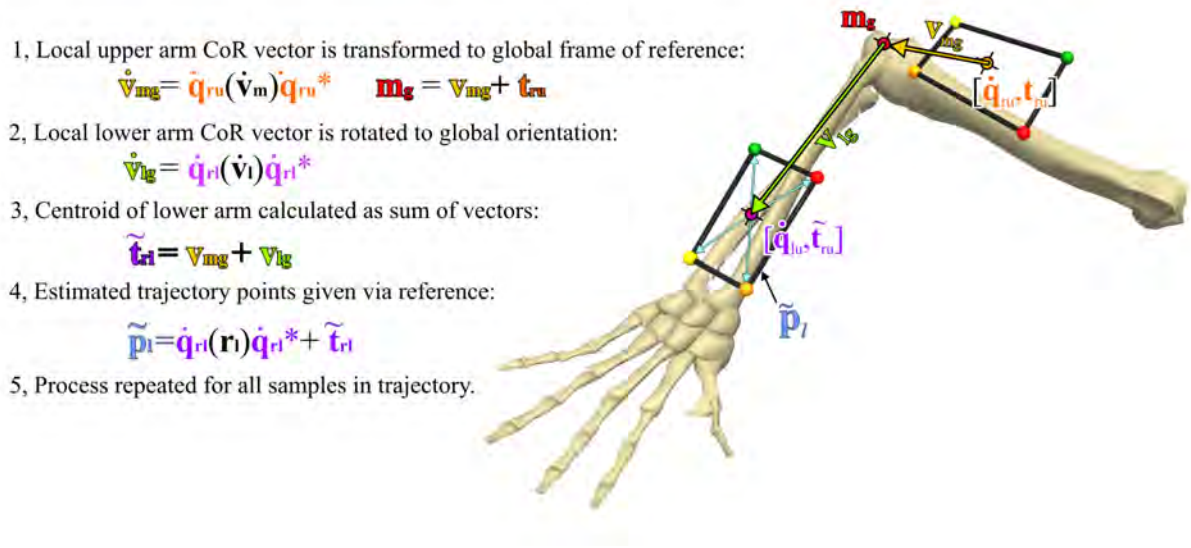
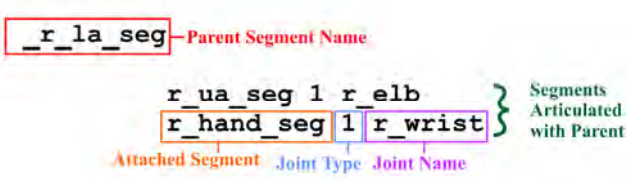


Figure 6.18 – Forward kinematic reconstruction of the pose of the arm

6.4.2 Implementation of full body kinematic model

Estimation of the joint centres for the full body was achieved by applying the single joint model to each pair of distal and proximal segments of the body model in turn. The hierarchy of the body model was defined manually via an external ASCII configuration file. This file has a format similar to the original cluster configuration (6.1.1), where the definition of each parent segment is preceded by an underscore. The definition is followed by a list of child segments with which the parent is articulated, and the global identifiers for each joint.



The joint configuration file is parsed to generate two lookup tables, referred to as the 'kinematic lookup'. The first containing the associations given in the configuration and the second the reverse information, describing the segments associated with each joint. By using these tables in conjunction with the pose data of the body segments, it is possible to retrieve the joint associations of each segment in turn and apply the CoR estimation algorithm. The resultant vectors defining the joint model ($\dot{\mathbf{v}}_m$ and $\dot{\mathbf{v}}_l$) are added to the 'locked data' structure, indexed according to the kinematic lookup. This process tailors the body model to the subject's physical dimensions and will be referred to as the 'body model calibration'.

Following calibration of the body model using the range of motion data set, a preliminary test was performed to verify approximately correct operation. Each local joint centre obtained by the calibration was transformed into the global measurement frame (step 1 of figure 6.18). By plotting the resultant global joint centres superimposed upon the original marker data, correspondence with approximate locations of anatomical joints was verified visually. To improve legibility of results, associations between clusters and joints were also marked, indicating the 'skeleton' of the body model. An example of this plot for three frames of the range of motion dataset is shown in figure 6.19. Although joint centres are calculated in both proximal and distal directions, only distally derived centres are plotted.

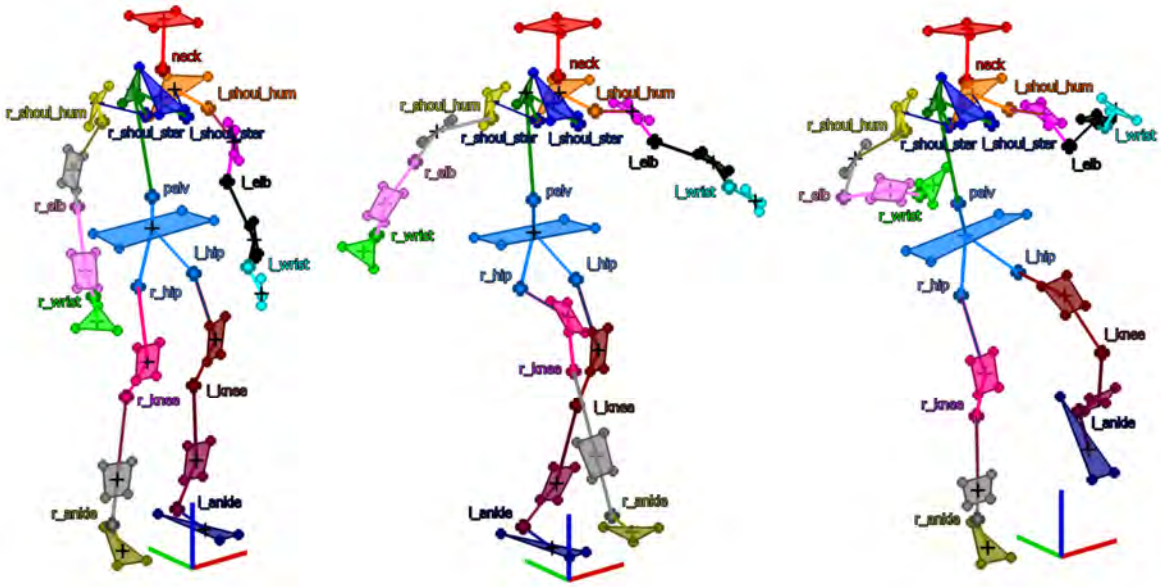


Figure 6.19 – Estimated functional joints with reference to marker locations

From this assessment it was seen that the majority of joints exhibited stable correspondence to their anatomical equivalents over the course of the tested motion sequence. However, inconsistent results were observed in the joints in the vicinity of the shoulder complex, which failed to produce acceptable solutions for either clavicle. This is indicted by the unexpectedly low and asymmetrical placement of the shoulder-humeral functional joint and shoulder-sternum joint, shown in figure 6.20.

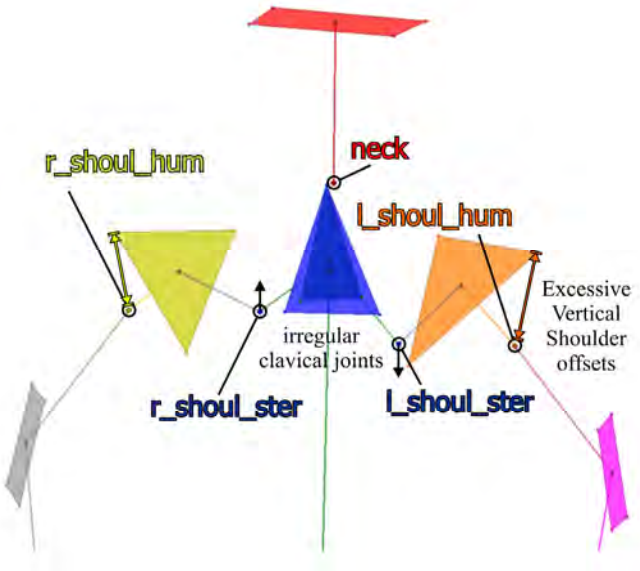


Figure 6.20 – Offsets in shoulder joints

It is likely this is due to poor RoM exhibited by the markers of its cluster as well as the over simplification of the shoulder complex in the body model. However, to conduct any further analysis of the performance of the body model a more objective means of assessment than visual inspection is required.

6.5 Validation of body model

Due to the idealised nature of functional joints and the inability to characterise their anatomical counterparts, it was not possible to assess the accuracy of the body model against any externally verifiable benchmark. Therefore, it was necessary to utilise several indirect methods to test the resemblance of measured joint motion to the assumptions of the model and the extent to which marker motion satisfies the requirements of CoR estimation.

6.5.1 Suitability of spherical joint model

If the 3R spherical approximation is valid then all neighbouring trajectories to a joint will be seen to lie on the surface of one or more spheres about the CoR. Therefore, any deviation from this relationship observed in measurements can either be attributed to measurement noise or deviation from spherical motion. As the expected range of measurement noise is known, values found above this range suggest inadequacy of the 3R model. The scale of this deviation is used to provide a rudimentary indication of the appropriateness of the 3R assumption for each joint.

Returning to the minimisation of the spherical cost function given in (6.3.2), for a group of trajectories $\mathbf{p}_{k,n}$, (where k denotes trajectory index for k=1, .., K and n is the point along the trajectory for n=1, .., N) the set of perfect concentric spherical orbits with centre $\mathbf{m}$ and radii r_k, which best fit the trajectory data will give the minimum average value of f_m expression for all k and n.

$$f_m = \frac{1}{KN}\sum_{k=1}^{K}\sum_{n=1}^{N}\left(\sqrt{\left|\mathbf{p}_{k,n}-\mathbf{m}\right|^2}-r_k\right)^2$$

However, the effects of measurement noise and deviations from a perfect spherical orbit will cause this value to deviate from its mean value over its RoM. Therefore, by modifying this expression to give the absolute error in trajectory radius from **m** against the fitted value r_m, an error metric of deviation from the spherical fit is obtained. By averaging across all radii and indexing by n the instantaneous value for this error e_n, is given as:

$$e_n = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(r_k - \left|\mathbf{p}_{k,n} - \mathbf{m}\right|\right)^2}$$

Following application of this metric to the joint centre solutions obtained in (6.4.2), the mean, maximum and standard deviation of radial deviation about each joint for all frames of the range of motion data set is shown in figure 6.21.

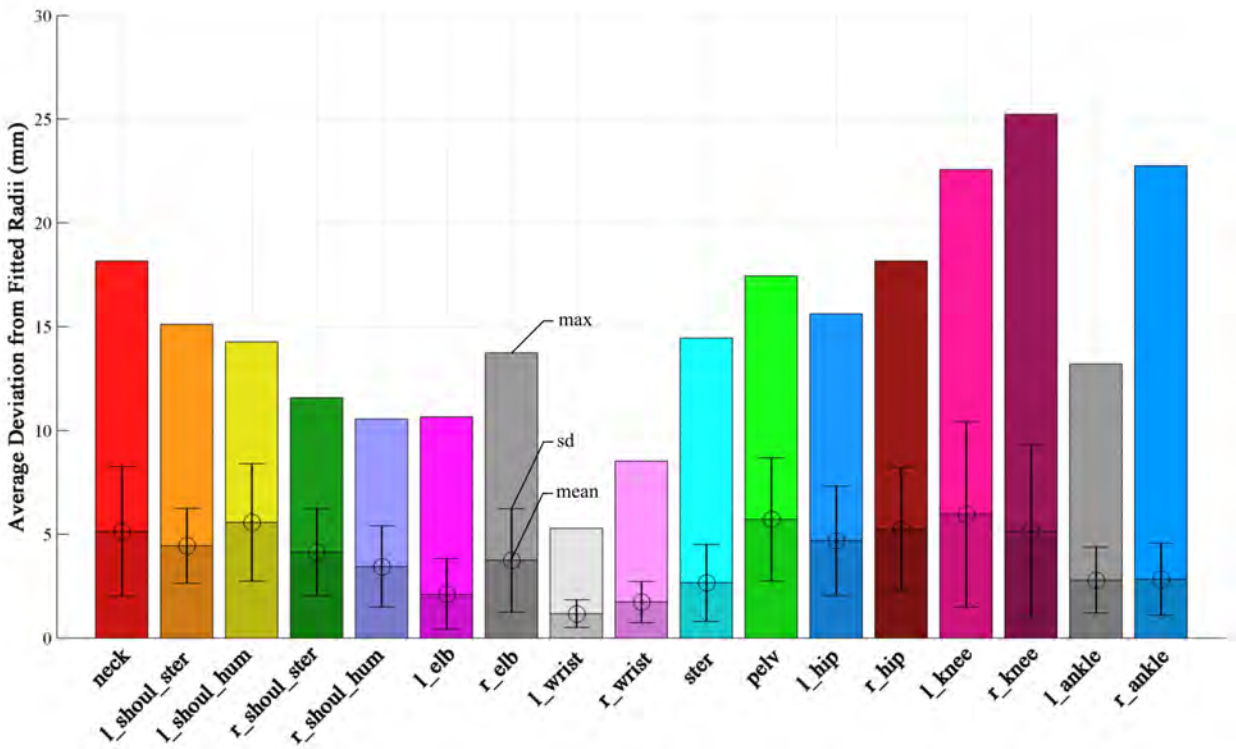


Figure 6.21 – Deviation of trajectories from spherical orbit

From this plot it appears that the resemblance of the spherical model to the trajectory data is good, with an overall maximum radial deviation of 25*mm*, mean of 4*mm* and standard deviation of 2.5*mm*. This places the majority of errors within the 10*mm* upper limit for measurement error, meaning they could have arisen solely from the measurement process, rather than any inadequacy of the model. However, in cases of low RoM a low variability of radial error does not necessarily preclude the possibility of a poor CoR fit. Specifically, this occurrence is due to low RoM leading to trajectory profiles, which are near planar rather than the desired spherical sector. Under these circumstances, a large fit radius relative to marker displacement will be required to minimise the cost function. Therefore, the small deviation of the trajectory from the near planar fit over the measurement RoM will lead to artificially improved error performance. Additionally, the slight changes in apparent curvature, which occur due to measurement noise can lead to major perturbations of the CoR estimate under these conditions. This accounts for the failure of CoR estimation shown in (6.3.3).

6.5.2 Approximation of joint range of motion

To test which segments are likely to produce problematic CoR solutions due to diminished RoM, the angular displacement of each joint was estimated. As the measurement subject was encouraged to exercise their full range of joint motion in the test data, any problematic physical limits were easily identified. To simplify the analysis of range of motion only the maximum angular displacement of each limb about its joint was considered. Direction and orientation of components of motion were disregarded to avoid the complexity inherent in their representation [48]. Each joint RoM was approximated as the maximum difference in angle between the initial vector defined by the joint centre and the distal segment centroid

$\mathbf{v_{md_1}}$, and the same vector for all remaining samples $\mathbf{v_{md_n}}$. This angle was given by the dot product of the normalised vectors as shown in figure 6.22.

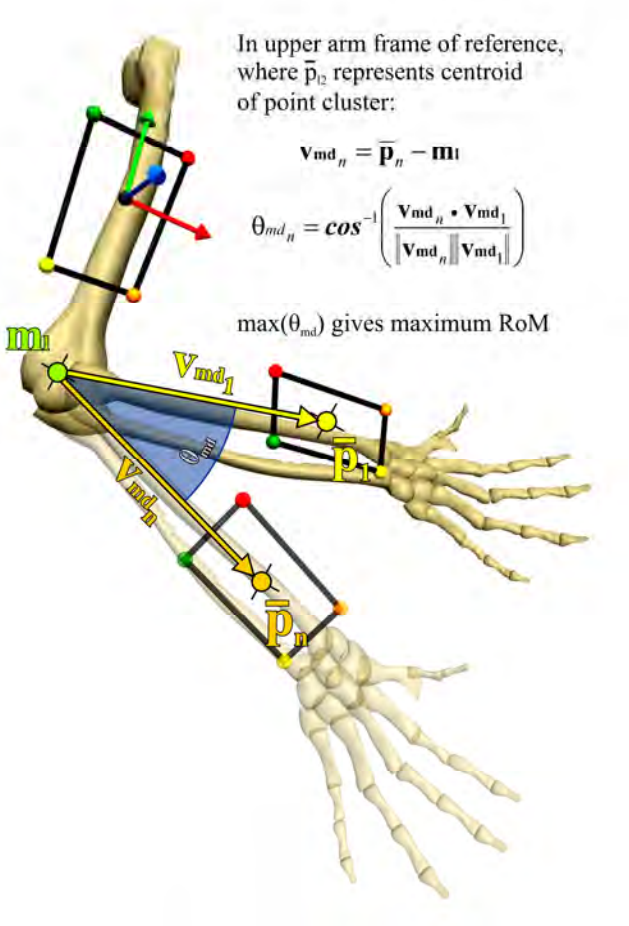


Figure 6.22 – Simplified range of motion

As this method gives the angular displacement from a fixed initial vector from which rotation in either direction can occur, the actual range of motion can be up to twice that returned by the dot product method for *n=1, .., N*. However, this limitation is taken into account in the analysis of results. The maximum angular displacement for each joint of the body model for the range of motion dataset is shown in figure 6.23.

With reference to the criteria for CoR stability established in (6.3.3), the majority of joints are seen to exhibit a RoM of 40^o or greater and therefore present no concern. However, from the results of (6.4.2), both shoulder-sternum functional joints (r_shoul_ster and l_shoul_ster) exhibit RoMs below the acceptable margin for accurate CoR placement, having a maximum possible range of ~20^o. The particularly low RoM of the sternum ('ster') joint is of little concern as this joint is effectively immobile and is omitted from the model from the primary experiment onward.

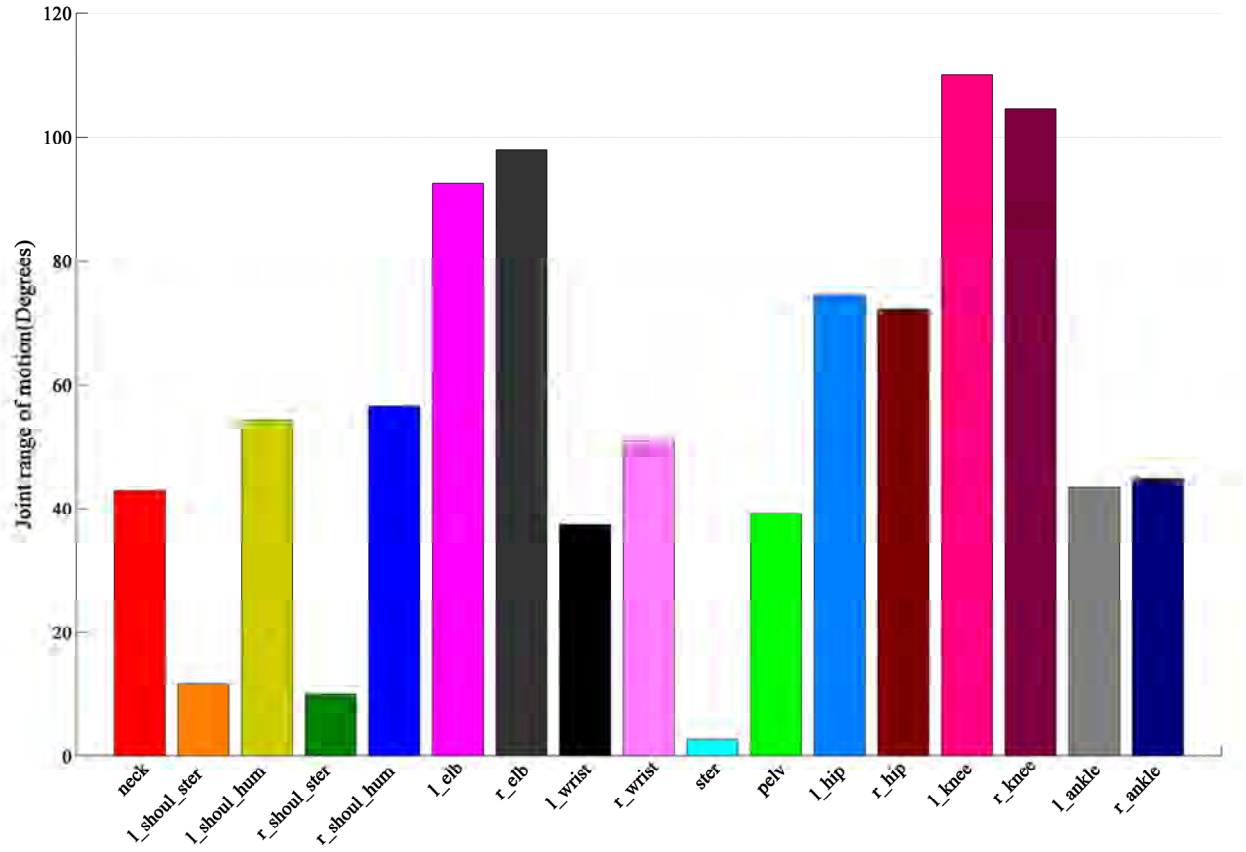


Figure 6.23 – Approximate joint range of motion

Regarding the shoulder-sternum joints, the observed marker motion about the CoR is primarily due to the motion of the bones of the clavicle and scapula about the sternoclavicular joint. As the scapula can freely translate (across the back of the ribcage), it is capable of moving relative to the clavicle and disrupting the rigid body correspondence of markers. This limitation in combination with the low RoM indicates that the simplified 3R model and CoR estimation are insufficient for proper characterisation of the shoulder complex.

6.5.3 Forward reconstruction of marker position

The performance of the body model was further assessed by driving it in a forward kinematic manner using the cluster orientation data obtained from measurements. Through the method given in (6.4.1) the trajectories of markers in any cluster were reconstructed in terms of the cluster's orientation and the complete pose of a neighbouring cluster. By calculating the average Euclidean error between marker trajectories reconstructed in this manner and those measured directly by the optical system, a metric for the resemblance of the body model to actual measurements is obtained. As this method is the basis for the occlusion recovery algorithms developed later in this thesis, the results of this analysis represent the baseline of error for the inertially assisted reconstruction.

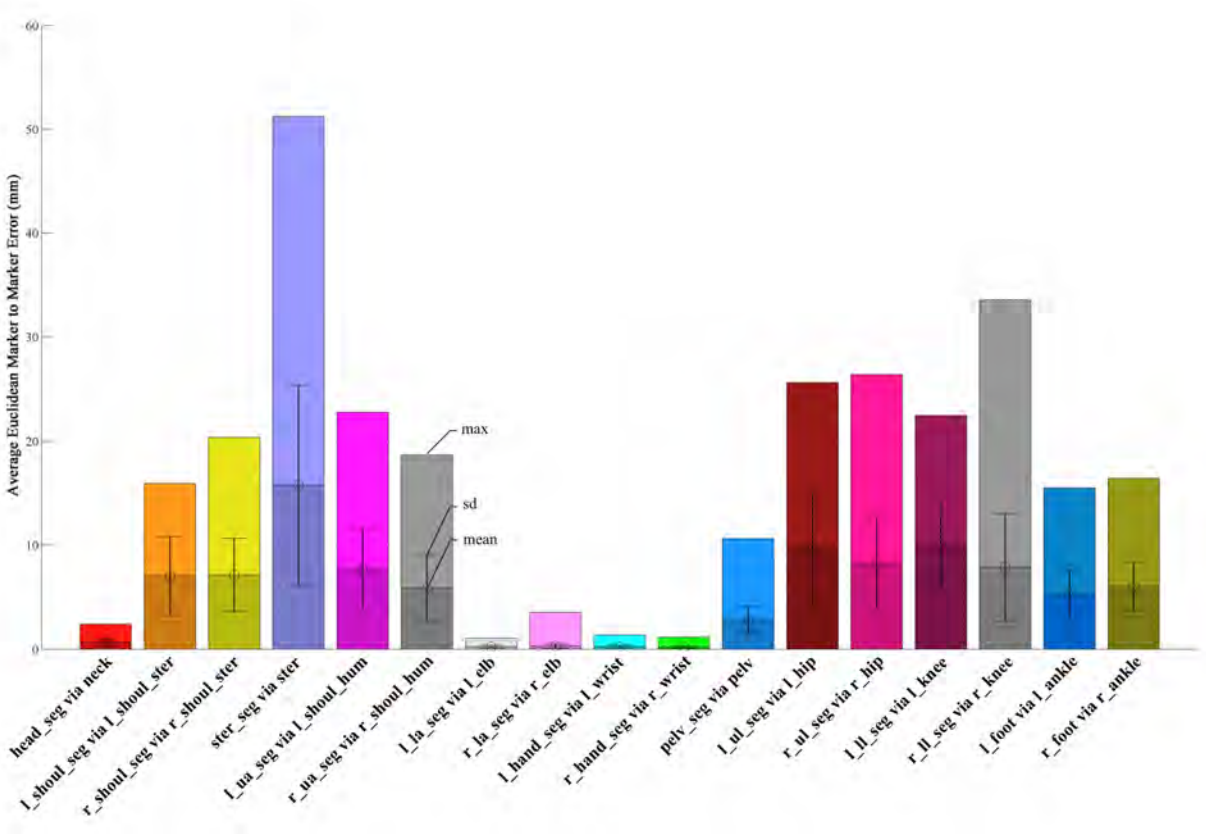


Figure 6.24 – Kinematic reconstruction error

Figure 6.24 shows the mean, maximum and standard deviation of average Euclidean error following application of the forward reconstruction. Trajectory comparisons are made against the range of motion dataset and each result is described in terms of the cluster reconstructed and parent joint name. Therefore, a result labelled ‘x via y’ signifies that the cluster ‘x’ was reconstructed about the joint ‘y’.

The kinematic model performs adequately, with the average positional error for all joints, barring the sternum (‘ster’), lying below the 10*mm* threshold of measurement error (4.1.4). In the case of the upper body joints, this average lies at least one standard deviation below this threshold. However, the joints of the lower limbs are all noted to produce slightly higher levels of error. This is attributed to the distances between lower limb cluster centroids and their associated joints being considerably greater than in the upper limbs. Due to the increased moment at this distance, effects of positional perturbation are magnified per unit of orientation error (Figure 6.9). The unexpectedly large peak error in the case of the right knee (r_knee) is due to spline interpolation causing a momentary deviation from rigid body correspondence, which is not reflected in the body model.

6.5.4 Conclusions on body model performance

This chapter has described the derivation of the kinematic body model from marker data and tested its performance. For the joints over the arms and legs, acceptable performance has been noted for marker reconstruction error against expected limits of measurement error. Furthermore, the required ranges of motion for calibration of the model are well within the ranges provided by the test data and the physical limits of these limbs. This presents a good

performance baseline for the inertially assisted reconstruction and suggests useful accuracy is possible, assuming that the error contribution from inertial measurements can be minimised.

With regard to the orientation estimation, the chosen cluster size and marker population gives desired performance in all cases apart from the hands, shoulders and c7 cluster. For the hands, this is easily addressed by increasing the cluster radius beyond the threshold defined in (6.2.1). The c7 is noted to give poor rigid body correspondence as two of its constituent markers are placed in proximity to the scapula, making them prone to relative movements. A measure to correct this was to combine the c7 and inner clavicle markers, discussed further in (10.1.2).

In the case of the shoulder clusters, in addition to problems with the stability of orientation estimates, the range of motion calculated for the sternoclavicular joint is below the threshold for accurate CoR estimation. Although the analysis conducted in (6.5.1) suggested that reconstruction via the estimated CoRs produces acceptable results, this was disconfirmed by the visual inspection of (6.4.2). It is therefore likely, that these unexpectedly positive results are due to the reduced range of motion leading to reduced deviation from the spherical approximation (6.5.1). Although an alternative shoulder model could be used in the place of the 3R method [71, 72], it is desirable that homogeneity of the body model be maintained to provide a straightforward foundation on which further stages of this project are based. Therefore, more sophisticated attempts to accurately model the articulation of the shoulder complex will be left for future work. Due to the potential for this limitation to introduce instability in further methods, this joint will be excluded from some of the later phases of analysis.

Due to their capability for horizontal translation, the simplification of the neck and pelvis articulations to 3R joints was a source of concern. However, the low levels of error obtained in the reconstruction, combined with adequate values of RoM suggest adequate performance. It is believed that this disparity can be accounted for by the absence of this translational motion in the sample data set (verified by inspection). It is possible that these additional degrees of freedom could be accounted for by additional joint types or the use of a predictive model of spinal movement [73]. However, both of these options would require revision of the marker set and departure from the exclusively 3R design of the body model.

7. SIMULATING OCCLUSION

Due to the inherently destructive nature of real occlusion, it was necessary to provide a means of simulating the effects of occlusion upon measurement data. This was necessary to preserve original trajectory data as a benchmark for assessment of reconstruction accuracy. Therefore, two simulations of the effects of occlusion were devised allowing for gold standard measurement data to be degraded in a predictable and non-destructive manner. The first was based upon a geometric model of Line of Sight (LoS) and the second utilised a more general statistical approximation. This chapter describes the derivation of these models, their testing against real occluded data, and compares their performance.

7.1 Volume based occlusion model

The volume model of occlusion simulated the interaction of the LoS between each camera and marker (epipolar line), with a geometric model of the subject's body. By driving this model using the pose data acquired from the primary experiment, the approximate occlusion performance of any arbitrary camera configuration could be simulated.

7.1.1 Camera model

The camera representation used in the volume occlusion model was based upon adaptation of the basic pinhole camera approximation introduced in (2.1.1). By transforming this model with reference to the measurement origin, it is possible to represent the pose of each camera in the measurement setup. As shown in figure 7.1, this pose is described by the translation ($\mathbf{t}_c$) and rotation ($\mathbf{R}_c$) in the measurement frame of reference. Additionally, the horizontal and vertical field of view about the camera's principle axis can be described by the angles (θ_u) and (θ_v) respectively.

The first two column vectors of the rotation matrix ($\mathbf{R}_c$) describe the alignment of the imaging plane ($\mathbf{v}_u$ and $\mathbf{v}_v$) and the third the principle axis ($\mathbf{v}_z$). A fifth parameter ($\mathbf{r}_c$) is also included to account for the maximum distance from $\mathbf{t}_c$ at which a marker can be resolved.

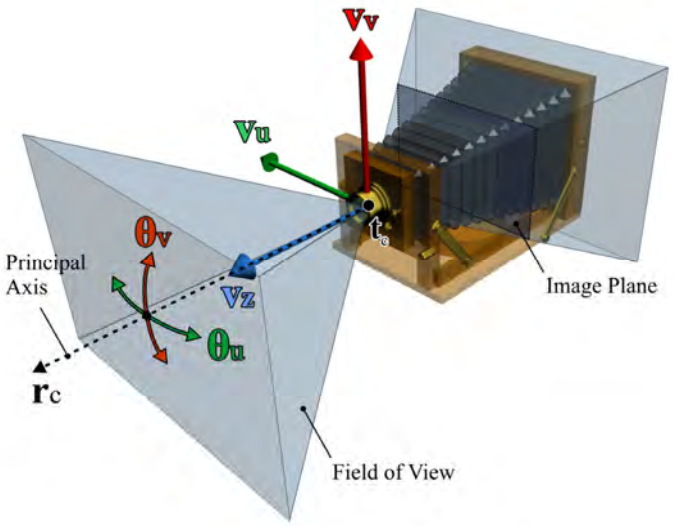


Figure 7.1 – Detail of camera model

7.1.2 Capture volume model

The basic condition for visibility of a marker is that it lies within the maximum resolving range and field of view (FoV) of a camera. Considering a marker to be represented as a point ($\mathbf{p}_m$) in measurement volume, its distance (d_c) from the principle focus of a camera and hence observation distance is expressed as:

$$d_c = \left\| \mathbf{p}_m - \mathbf{t}_c \right\|$$

Under the conditions of the pilot experiment, the maximum range (r_c) at which a camera could reliably detect an 18*mm* marker was approximately 4 meters. By thresholding d_c against this value for each camera in the configuration, markers exceeding the range of visibility for each camera were detected.

To test for a marker's presence within the camera FoV, a conical approximation rather than the actual pyramidal frustum was utilised. Although technically incorrect, the conical frustum

greatly simplifies the testing of camera FoV. Furthermore, due to the elevated placement of cameras in the measurement setup, occlusions occurring at the limit of the vertical FoV were far more common than those at the wider horizontal FoV. Therefore, the simulated conical FoV was set to equal the vertical FoV of the measurement system.

The angle of a marker within the conical FoV was calculated as the angle (θ_d) between the epipolar line of a marker and the principle axis of the camera:

$$\theta_d = cos^{-1} \frac{\mathbf{v}_z \cdot (\mathbf{p}_m - \mathbf{t}_c)}{\|\mathbf{v}_z\| \|\mathbf{p}_m - \mathbf{t}_c\|}$$

Occlusion due to FoV was modelled by thresholding this value against half the measured vertical FoV of the cameras. This was found to be approximately 22.5^o ($45^o/2$).

7.1.3 Volume model for occlusion simulation

Interruption of LoS by the subject was modelled as the intersection between epipolar lines and a geometric representation of the subject's body. For this purpose, a polyhedral technique was utilised [74], where each body segment was represented as a closed mesh composed of multiple polygons. This mesh was represented numerically via two matrices, one containing the coordinate data of each vertex of the mesh, and the second the order in which these vertices were associated to define each polygon of the mesh. Figure 7.2 shows a simple triangular pyramidal volume composed of a vertices matrix ($\mathbf{V}_m$) and faces matrix ($\mathbf{F}_m$).

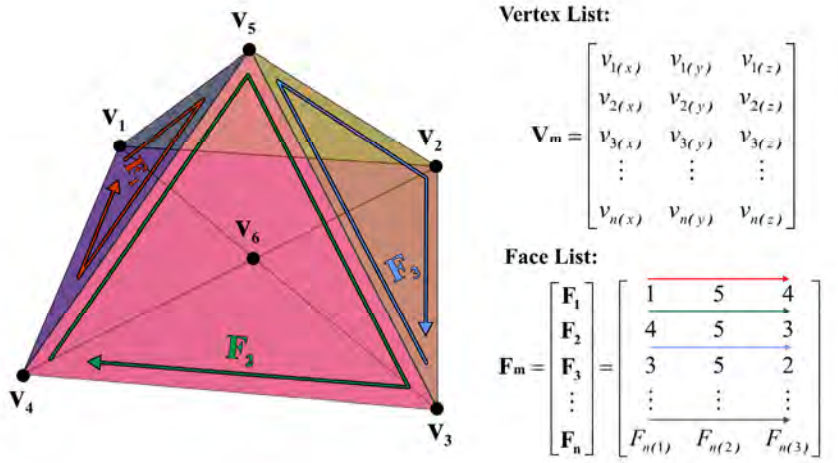


Figure 7.2 – Polygon pyramidal mesh

An advantage of this method is that the pose of a mesh can be adjusted by transforming the coordinates of the vertex matrix by a combination of translations and rotations. In a similar manner, the centroid ($\mathbf{c_m}$) of the mesh is calculated as the mean of each vertex in $\mathbf{V_m}$:

$$\mathbf{c_m} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{V_{m}}_i \qquad \text{where } M \text{ is the number of points in } \mathbf{V_m}$$

As with the marker cluster described in (6.1.2) this provides a means of establishing the local origin of the mesh. Using this point as the origin for rotation, the geometry can be transformed to a pose described by the Rotation matrix $\mathbf{R_m}$ and translation vector $\mathbf{t_m}$, resulting in the transformed vertices $\mathbf{V'_m}$:

$$\mathbf{V'}_m = \mathbf{R}_m(\mathbf{V}_m - \mathbf{c}_m) + \mathbf{t}_m$$

Deriving the value of $\mathbf{t}_m$ from the centroid of a measured marker cluster and $\mathbf{R}_m$ from its estimated orientation, it is possible to register any arbitrary mesh to the pose of the cluster. In this manner, the geometry can be approximately aligned with the volume of the body segments of a measurement subject. This is shown in figure 7.3, where two square frusta are registered to the pose of right arm giving a crude representation of its volume.

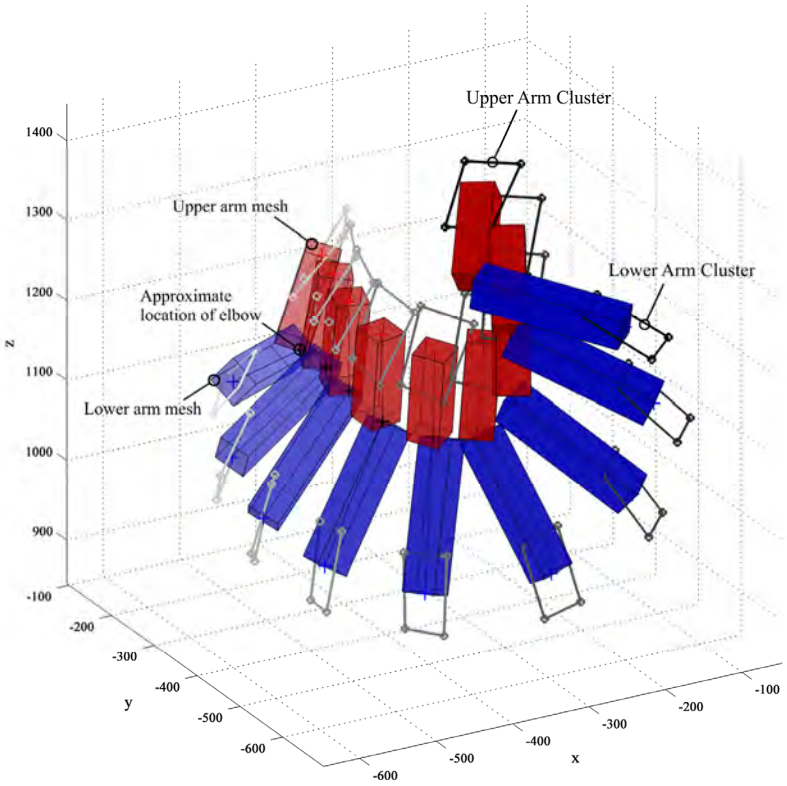


Figure 7.3 – Simple volume model of the arm

7.1.4 Ray-polyhedron occlusion model

Occlusion by body geometry was assessed by solving for the intersection of an epipolar line with each polygon of a mesh in turn. The basis for this solution was the ray-triangle intersection algorithm given in Moller et al. [75]. This allowed for the point of intersection ($\mathbf{P}(u,v)$) to be found between a line segment defined by a start and end point ($\mathbf{p}_s$ and $\mathbf{p}_e$), and a triangle defined by three vertices ($\mathbf{v}_1$,$\mathbf{v}_2$ and $\mathbf{v}_3$) ,as shown in figure 7.4. This method utilised

the transform, which mapped the vertices to an origin aligned unit triangle in the global x-y plane. By applying this transform to the line segment and solving for its intersection with the x-y plane, the point of intersection within the local (*u, v*) plane of the triangle was established. From the working of this method (Appendix 3) the solution in terms of the local co-ordinates and distance (*t*) along the line segment is given by:

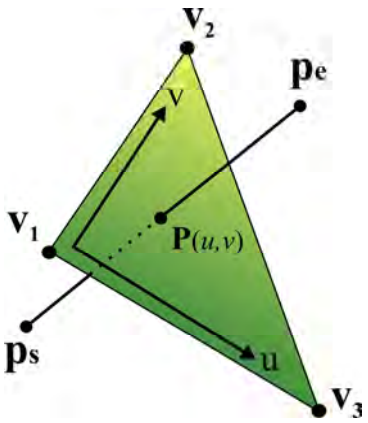


Figure 7.4 –Line segment triangle intersection

$$\begin{bmatrix} t \\ u \\ v \end{bmatrix} = \frac{1}{\mathbf{e}_1 \cdot (\mathbf{p}_n \times \mathbf{e}_2)} \begin{bmatrix} \mathbf{e}_2 \cdot (\mathbf{f} \times \mathbf{e}_1) \\ \mathbf{f} \cdot (\mathbf{p}_n \times \mathbf{e}_2) \\ \mathbf{d} \cdot (\mathbf{f} \times \mathbf{e}_1) \end{bmatrix} \quad \text{where} \quad \begin{aligned} \mathbf{e}_1 &= \mathbf{v}_2 - \mathbf{v}_1 \\ \mathbf{e}_2 &= \mathbf{v}_3 - \mathbf{v}_1 \\ \mathbf{f} &= \mathbf{p}_s - \mathbf{v}_1 \end{aligned}$$

Intersection between the line and triangle was confirmed by testing whether the *u* and *v* co-ordinates lie within the area of the unit triangle and that the following inequality is met:

$$0 \le u \le 1 \quad \text{and} \quad 0 \le (u + v) \le 1$$

To calculate the point of intersection in the original three dimensional coordinate frame, the equation for the plane of the triangle or line segment was used with the parameters given by the intersection solution. Alternatively the equation of the line segment can produce an identical result.

$$\mathbf{P}(u,v) = (1 - u - v)\mathbf{v}_1 + u\mathbf{v}_2 + v\mathbf{v}_3 = \mathbf{p}_s + t\mathbf{p}_n$$

The algorithm was implemented as a C++ Microsoft windows DLL via Matlab's MEX headers. This provided a substantial increase in processing speed, necessitated by the large number of line polygon intersections generated by the full body model. To extend the method's usefulness beyond triangular meshes, a polygon subdivision function was also devised allowing for processing of polygons with higher numbers of vertices.

Combining the intersection and subdivision functions provided the basis of the volume occlusion algorithm. This allowed any geometry expressed in the vertex-face list convention to be tested for intersection with any arbitrary line segment. Figure 7.5 shows the result of this algorithm applied to a simple test mesh, which is intersected by multiple radial line segments. The red lines indicate intersections and the black crosses mark points of intersection.

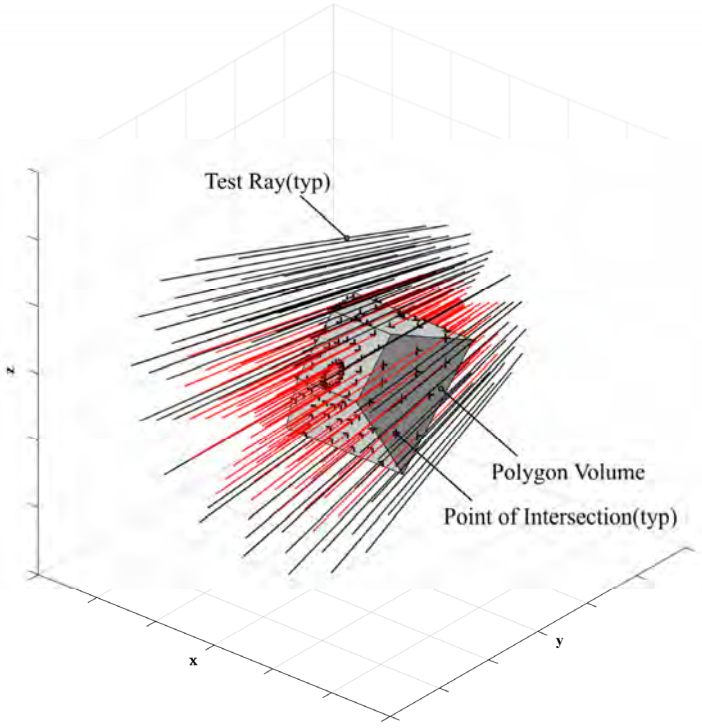


Figure 7.5 – Mesh intersection with radial line segments

7.1.5 Body segment pose

For alignment of the mesh representation of a body segment and the subject's anatomy, it was necessary to establish a coordinate system local to each segment, in addition to that of the marker cluster. This was achieved via the reconstruction of a local system of axes about the joint centres provided by the kinematic body model. This ensured alignment between abutting ends of mesh segments as well as anatomical correspondence.

Figure 7.6 represents the derivation of this local axis system for the lower right forearm in the form of a globally referenced rotation matrix ($\mathbf{R}_s$) and an origin defined by the proximal joint centre ($\mathbf{c}_p$). By taking the difference between the globally referenced proximal and distal joint centre coordinates, $\mathbf{c}_p$ and $\mathbf{c}_d$, the vector ($\mathbf{v}_s$) was obtained defining the principal axis of the segment. By normalising this vector, the *y-axis* of the local axis system and second column of the rotation matrix were defined. The point $\mathbf{p}_i$ was obtained by finding the line, which passed through the centroid ($\mathbf{p}_c$) and was perpendicular to $\mathbf{v}_s$. Taking the difference between this point and the centroid gave the vector $\mathbf{v}_i$, which when normalised defined the local *z-axis* and the third column of $\mathbf{R}_s$. Finally, by exploiting the redundancy in the matrix representation of rotation, the *x-axis* and third column were given by the cross product of $\mathbf{v}_s$ and $\mathbf{v}_i$.

Calculation of the co-ordinates of the intersection point $\mathbf{p}_i$ and hence vector $\mathbf{v}_i$ was achieved by the solution of the minimum distance between the point $\mathbf{p}_c$ and line described by $\mathbf{v}_s$ projected from $\mathbf{c}_p$. The actual coordinates of this point were found by the generic method described in [76]. This intersection operation will be represented by the notation $\mathbf{a} \perp \mathbf{b} \rightarrow \mathbf{c}$. This gives the point on the line segment $\mathbf{b} \rightarrow \mathbf{c}$, which is closest to point $\mathbf{a}$.

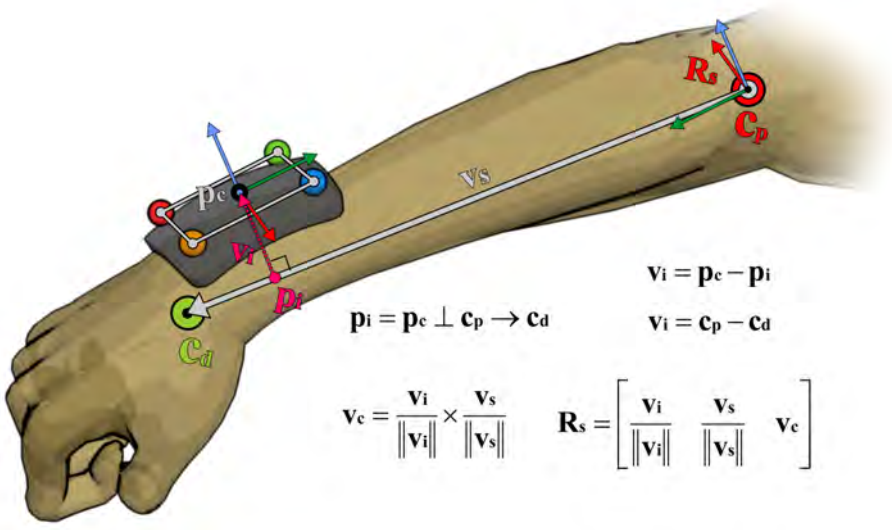


Figure 7.6 – Segment relative axis system

This technique was applied to all body segments with single proximal and distal joint centres, such as the upper and lower arms and legs. However, in the case of segments possessing more than two joint centres (such as the torso and pelvis), $\mathbf{c}_p$ and $\mathbf{c}_d$ were defined as the centroids of collected proximal and distal centres respectively. In the case of terminal segments (such as the head, hands and feet), where only a single joint centre is available, the segment frame is obtained directly from cluster orientation. Any misalignment between cluster and anatomy is corrected for by an additional arbitrary rotation. This rotation is defined manually from commonsense observations about segment posture, for example, that the feet lie flat upon the floor when standing normally.

7.1.6 Estimation of body mesh

To achieve an improved geometric representation from the example given in figure 7.3, key dimensions of each body segment were obtained from marker trajectory data. These measurements were used to scale geometric primitives to resemble each limb segment before alignment to the segments pose. Figure 7.7 shows a method by which the vertices of a unit cube ($\mathbf{v_f}$) were scaled to give a rectangular frustum ($\mathbf{v_c}$) approximating the volume of the lower right arm. The scaling was described in terms of the cross section of the frustum at its two bases, given by (w_1, h_1) and (w_2, h_2) and its length (l). Transforming $\mathbf{v_f}$ by the pose estimate of a body segment ($\mathbf{R_s}$) gave the segment registered vertices ($\mathbf{v_c}'$) approximating the volume in the global frame.

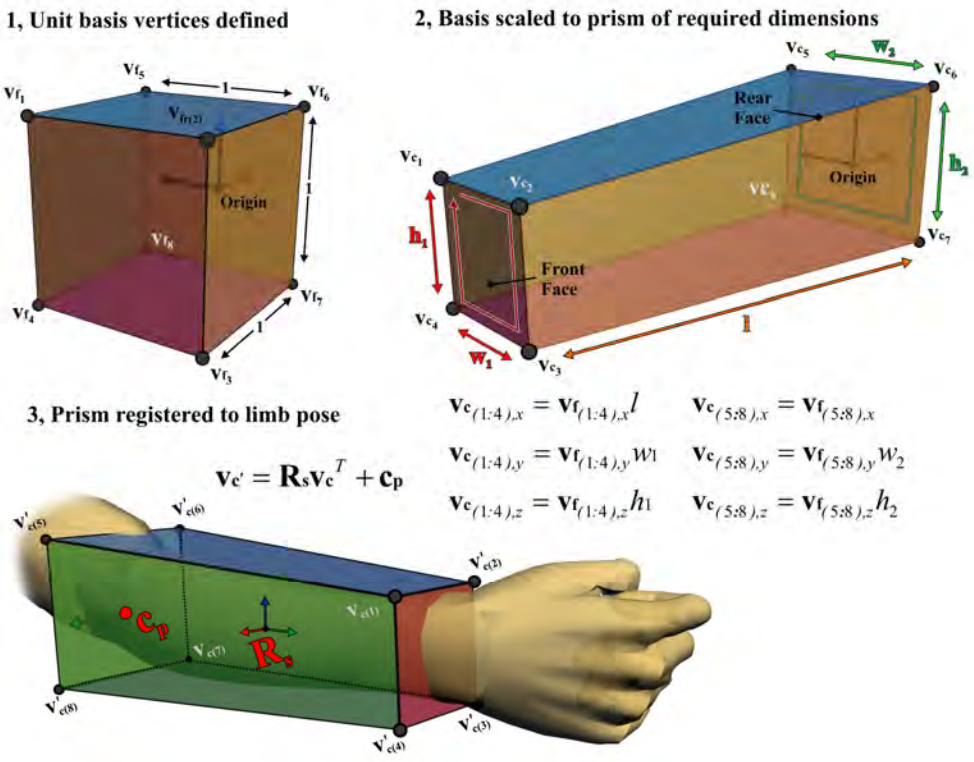


Figure 7.7 – Square frustum approximating volume of the forearm

For body segments with two functional joints, the dimensions of the frustum were obtained via estimation of the taper of each body segment along its length. This taper was estimated as the gradient (Δh) along the local *y-axis* of the segment between its principle axis $\mathbf{v}_s$ and the plane defined by the marker cluster (assumed to run parallel with the limb surface). By extrapolating a line with this gradient from the cluster centroid along the length of the segment, the diameter of the segment was approximated. Figure 7.8 shows this operation applied to the lower arm, defining the dimensions h_1 and h_2 at each joint. Coordinates of the marker cluster are defined as $\mathbf{m}_n$ (where n is the index of the marker), $\mathbf{p}_p$ represents the centroid of the two proximal markers of the cluster and $\mathbf{p}_k$ the closest point on $\mathbf{v}_s$ to $\mathbf{p}_p$.

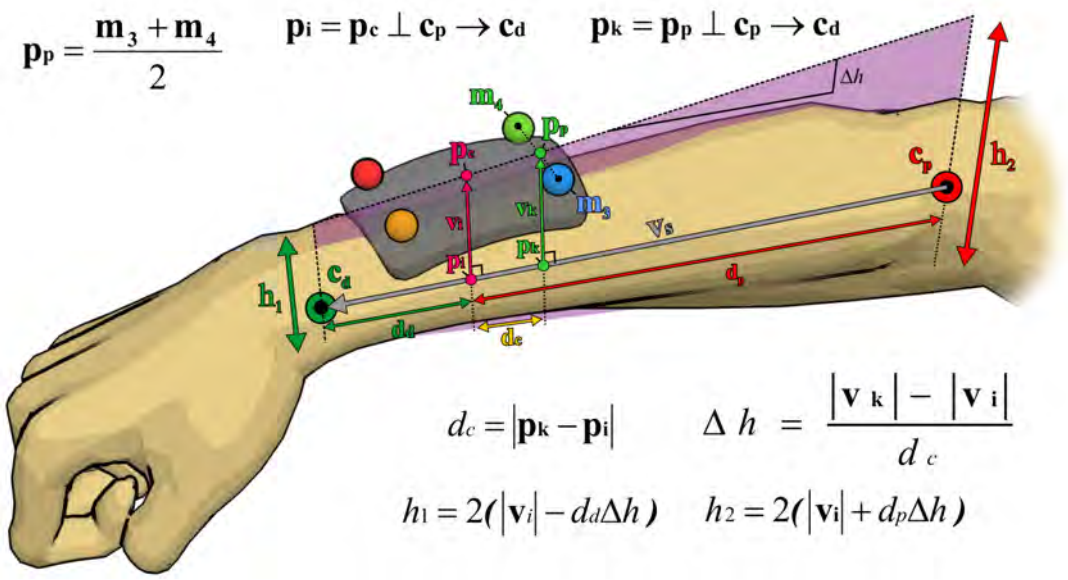


Figure 7.8 – Estimating the taper of the forearm

From rough anatomical measurements of the subject, it was found that the mean average of the ratio of lower arm breadth to height was approximately 1.5 at the wrist and 1.2 at the elbow. The h_1 and h_2 parameters were scaled by this value yielding w_1 and w_2. Similar aspect

ratios were calculated for other limb segments. Finally, the segment length l was obtained directly from the magnitude of the proximal-distal vector $\mathbf{v_s}$.

The meshes of segments having greater or less than two joint centres were estimated by alternate means. Although similar techniques were applied in the calculation of segment taper, the sequence of operations and the relative locations or markers were different. The dimensions and ratios involved are given in figure 7.9.

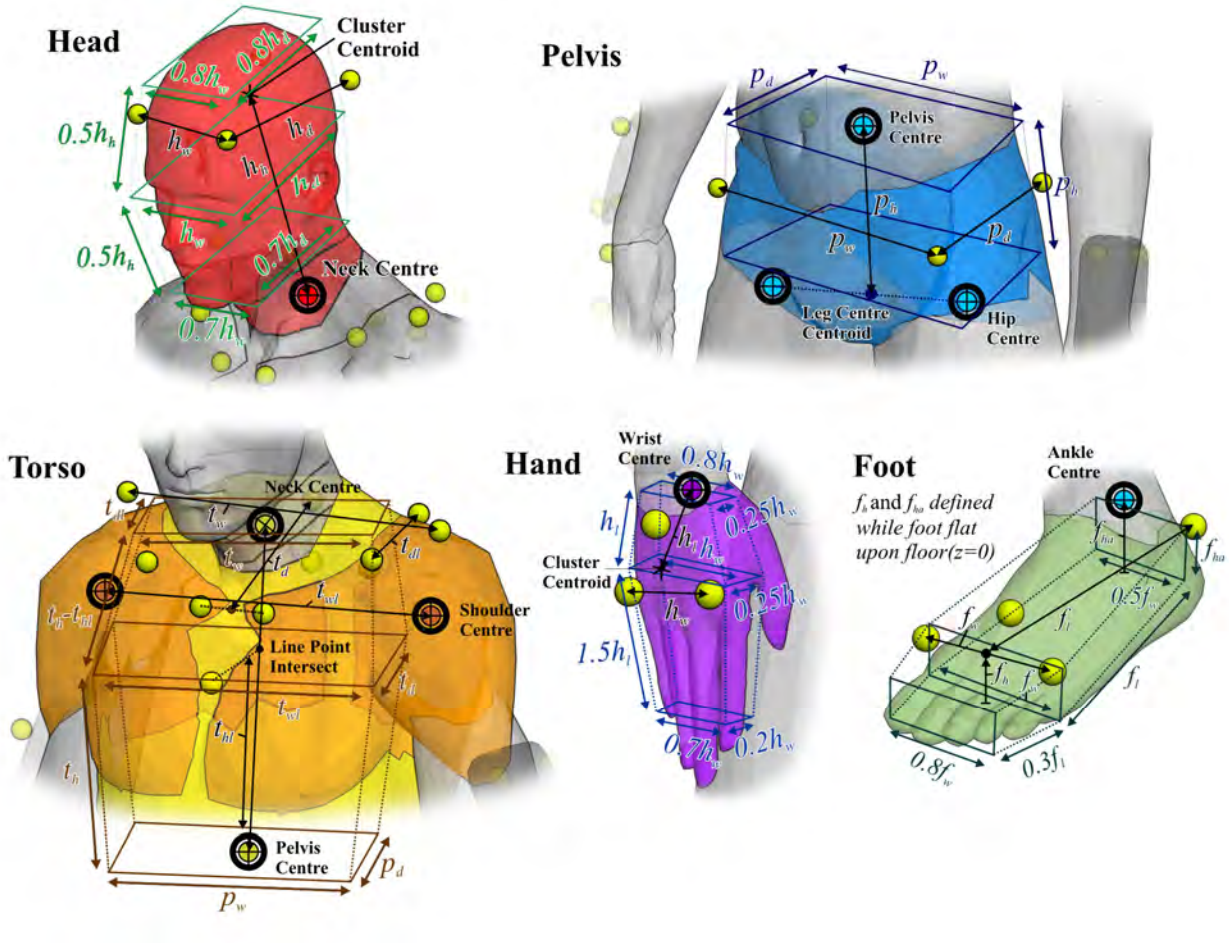


Figure 7.9 – Dimensions of non-limb segments

The linear approximation of taper with limb length did not give good agreement of dimensions between abutting segments. Therefore, to ensure matching cross sections at joints the *h* and *w* parameters of opposing faces were averaged. A further improvement was the addition of approximately spherical ‘gap filling’ meshes about the model’s joint centres. These meshes prevented open notches from appearing between abutting prisms at extreme joint angles. Finally, the cuboid primitives were replaced with hexagonal prisms to obtain a better approximation of the roughly cylindrical anatomy of the human body. The final implementation of the volume body model was conducted in Matlab and is presented in Appendix 3.

7.1.7 Conclusions on volume body model

The volume body model was applied to the pilot data set and the resulting geometries were combined to give a single mesh for each frame of the measurement data. This data is plotted for selected frames of the ‘preparation to single whip’ motion sequence in figure 7.10. The original measured marker coordinates and function joint centres are superimposed for reference.

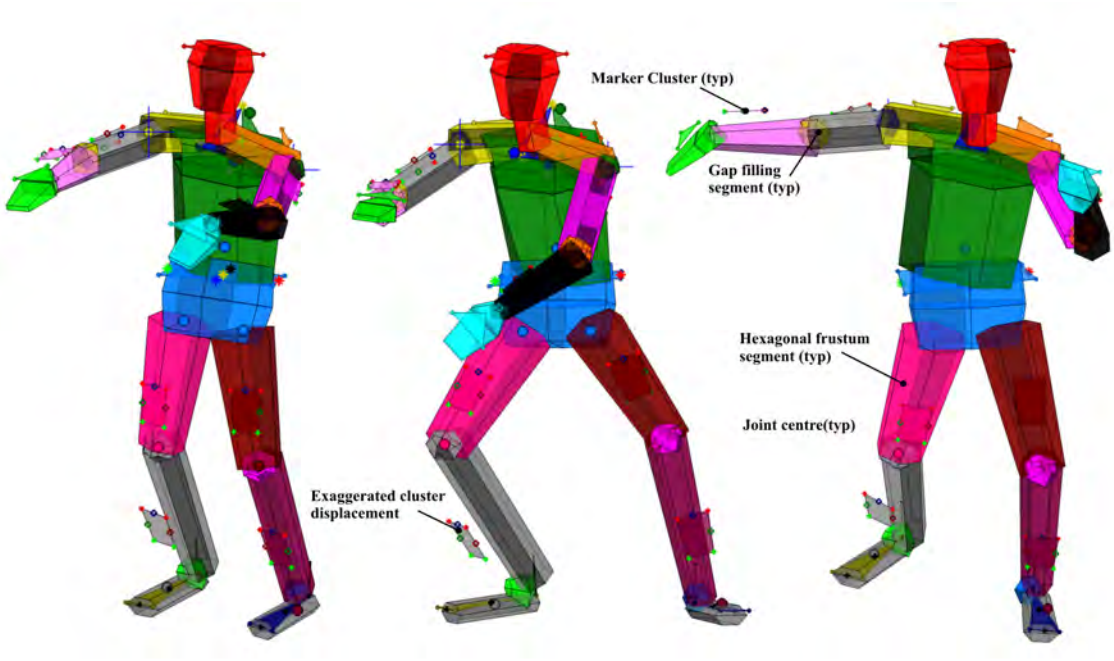


Figure 7.10 – Output of volume model

From examination of this data several shortcomings of the volume model are apparent. Despite the addition of hexagonal frusta, the model presents a crude representation of the human body, accommodating none of its complex curves or changes in dimension over its envelope of movement. This is compounded by the limited range of anthropometric measurements, which can be obtained from the marker data. Additionally, the cross section estimation relies upon the assumption that each segment is symmetrical with volume evenly distributed around its principle axis. However, this is rarely the case for real anatomical joints and their positions relative to the surface of the skin. To accommodate for this asymmetry an additional transformation was applied to move the mesh away from the axis. These transforms were assigned manually to give acceptable appearance as opposed to being based upon actual subject measurements. An effect of this was the exaggerated distance between cluster and mesh seen in figure 7.10. As these clusters were flush with the surface of the skin during measurement, this evidently indicates a deviation of the model from actual segment pose.

7.1.8 Testing of the volume occlusion model

Integration of the camera model, volume body model and mesh intersection algorithm yielded the final volume occlusion model. Using this model the points of intersection of each epipolar line with the body mesh for each frame of experimental data were approximated. By identifying markers visible to less than two cameras, the points of marker occlusion were established. A visualisation of the output from the occlusion simulation for a six camera configuration can be seen in figure 7.11. Epipolar lines, which intersect with the geometry are coloured black and those with LoS to markers match the colour of the camera from which they originate.

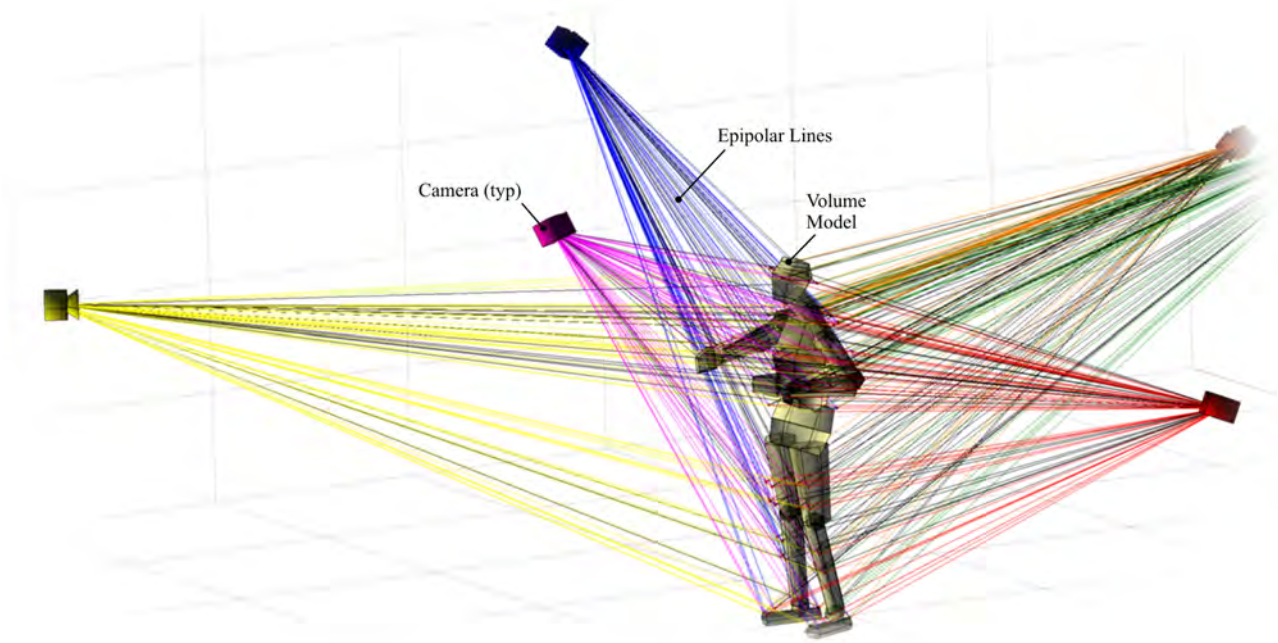


Figure 7.11 – Epipolar lines in a typical six camera ray-volume intersection test

Due to the combinatory nature of the occlusion model, the number of line-polygon intersections required to define the occlusions in a single motion sequence was a factor of the number of cameras, markers, polygons and frames. This presented significant computational requirements. For the measurement configuration used in the pilot experiment there were 12

cameras and 68 markers over 12k frames. With regard to the volume model, the 30 solid segments in its structure consist of 584 triangles, resulting in approximately 476k ray-triangle intersections per frame. Using the optimised geometry intersection algorithm (7.1.4) it was found that processing of a single frame of 12 camera occlusions took around 3.6 seconds on the test hardware (3.5*GHz* Pentium 4, 2Gb RAM). This gave a total processing time of about 12 hours for a complete 12k frame trial.

For testing of the occlusion simulation, a small range of trajectory data degraded by actual occlusion was produced via the 'QTM' software of the motion capture system [7]. Two sub-sets of four cameras were selected from the twelve camera measurement configuration, designated 'even' and 'odd'. Configurations of four cameras were considered appropriate for test purposes as they produced a high degree of occlusion (~30%) while minimising the occurrence of unresolvable fragments (i.e markers whose identities could not be verified manually). By repeating the 3D reconstruction of experimental data using these camera sets two fragmented datasets were produced. Identities of these trajectory fragments were re-established manually. To provide the simulated data set for comparison, the parameters of these two camera configurations were extracted from QTM and used in combination with the volume occlusion model to synthetically degrade the original experimental data.

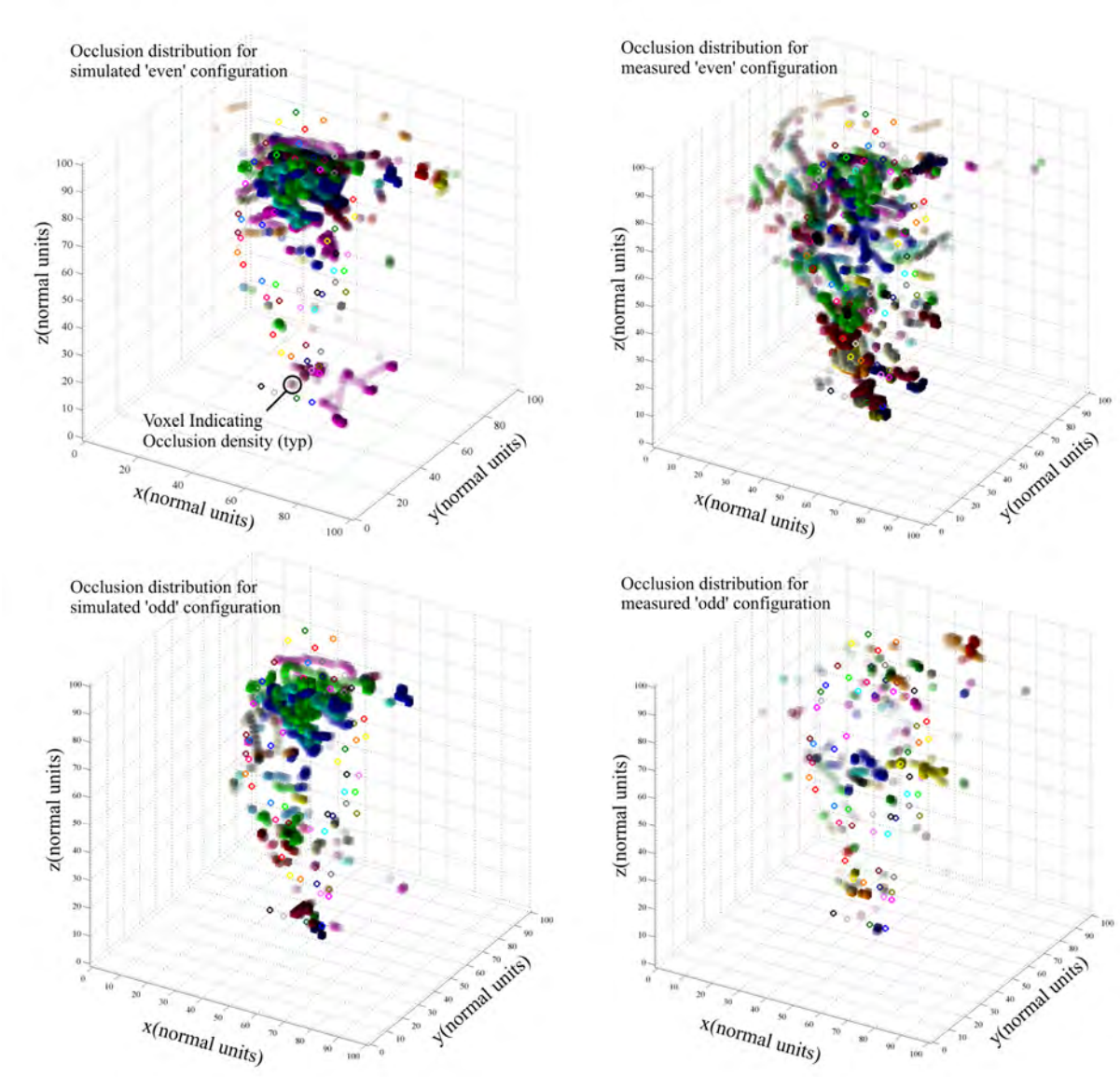


Figure 7.12 – Voxel based occlusion comparison

Comparison of occlusion performance was achieved by the subdivision of the measurement volume into cubic cells (or voxels). Using a 100^3 unit array scaled to accommodate the full range of marker motion, the number of frames each marker spent in occlusion per voxel was calculated. Normalising these values for each marker over the entire motion sequence resulted in a volumetric map of relative occlusion density. This was used for the comparison of occlusion performance between camera configurations. Figure 7.12 shows the results of this

method applied to measured and simulated trajectory data for the odd and even camera configurations. Increasing occlusion density is proportional to voxel opacity and voxel colour indicates marker identity.

By comparing the distribution of voxels between simulated and measured datasets, little resemblance was noted between them. Although the overall spatial distribution of voxels is similar due to each set being derived from the same data, the locations where particular markers are occluded are quite different. This was verified visually by comparing the colour and position of groups of voxels between simulated and measured datasets. Given this poor resemblance, little improvement was expected from further tests.

In addition to the deviation of the body model from the subject's anatomy (7.1.6), it is suggested that the varying reflectivity of markers, the occurrence of partial occlusions and soft tissue movement account for this disparity. Although the realism of the body model's geometry could have been improved, this would have done little to accommodate the stochastic behaviour of these factors. Furthermore, during the development of the occlusion reconstruction algorithms it was noted that the frequency of marker occlusion varied substantially across the marker set, with occlusion occurring more frequently for specific markers and body poses (13.3.1). To allow for performance of the reconstructor to be assessed over the widest range of movement conditions, it was desirable that all marker trajectories exhibited extensive fragmentation. However, this bias in occlusion only allowed for the reconstructor to be tested over a small range of motions and markers. The effects of this bias could have been reduced by conducting tests over a very large number of camera configurations. However, the low computational efficiency of the volume simulation of

occlusion made this impractical. Therefore, an alternative model capable of producing even distributions of occlusion was required.

7.2 Statistical method of occlusion simulation

In an attempt to address the shortcomings of the volume model, a statistically based model of occlusion was derived based upon Chen et al. [14]. This method described occlusion solely in terms of period and duration of events. This eliminated dependence upon body pose and produced an even distribution of occlusion events appropriate for testing of the occlusion reconstruction described later in the thesis.

7.2.1 Generation of occluded data for characterisation of the statistical model

To characterise the distribution of occlusion, samples of real occlusion for multiple camera configurations were required. To avoid extensive manual re-labelling of experimental data as described in (7.1.8), trajectory fragments and the labelled trajectories of the gold standard data were compared. By finding the gold standard trajectory with minimum distance to each trajectory fragment, marker identity was inferred. By applying this method to data reconstructed with varying numbers of cameras, multiple sets of labelled occluded trajectory data were obtained.

7.2.2 Distribution of occlusion duration and period

In isolation from the body model, only two parameters are required to describe an occlusion event: its period and duration. As illustrated in figure 7.13, period is the time between the start of one occlusion and the next, and duration is the time over which a marker is occluded.

The distribution of period and duration was obtained for several sub-sets of the experimental camera configuration for the "preparation to single whip" motion sequence. To produce an even degradation of occlusion performance, cameras were eliminated from the experimental configuration with the aim of maintaining a symmetrical distribution around the capture volume. Figure 7.14 and 7.15 show these distributions as overlaid histograms for occlusion period and duration respectively. Note the absence of the 8 camera data which was lost due to operator error.

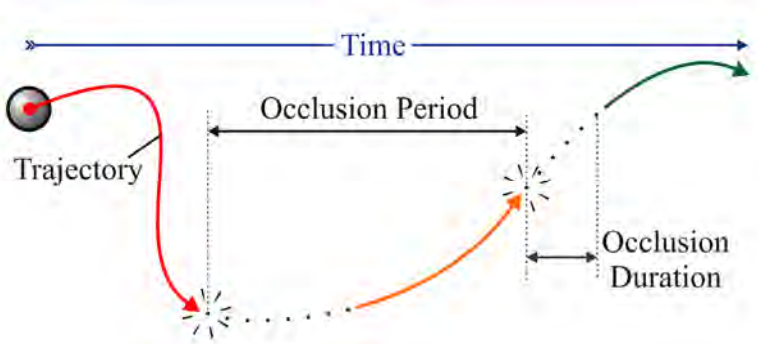


Figure 7.13 – Occlusion duration and period

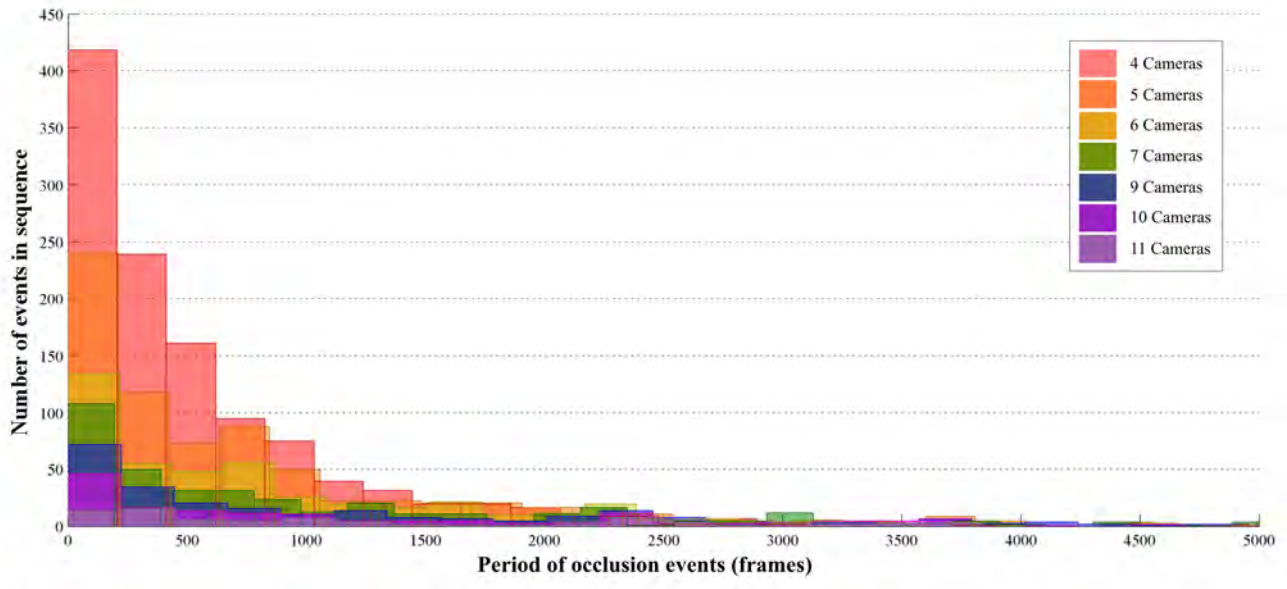


Figure 7.14 – Distribution of occlusion period

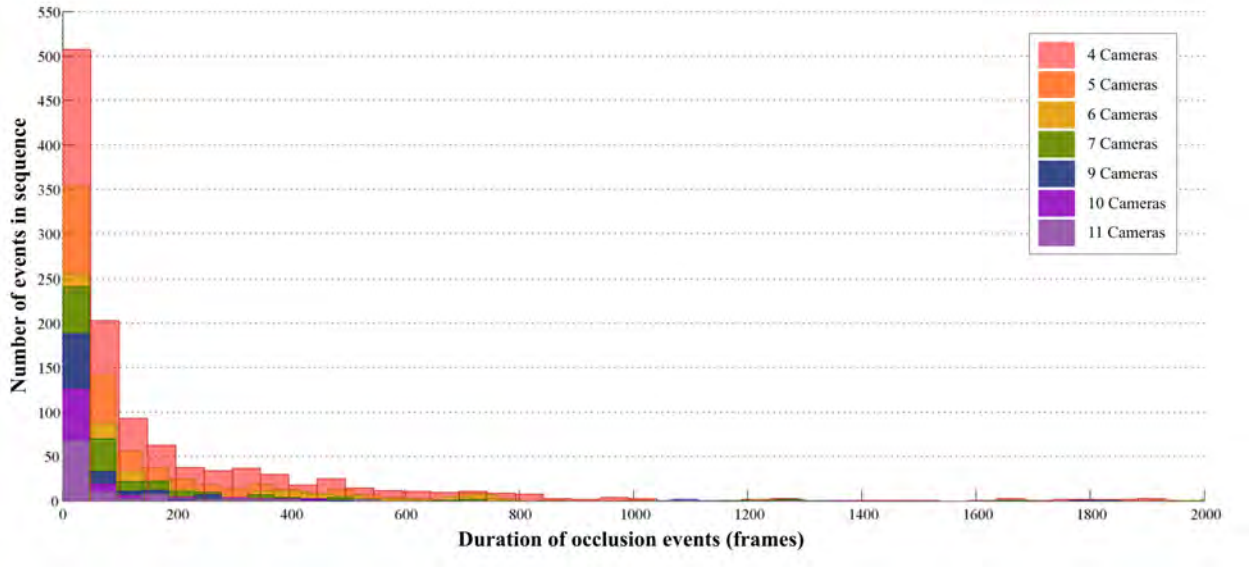


Figure 7.15 – Distribution of occlusion duration

By increasing the number of cameras the period and duration of occlusion events is continuously reduced. More interestingly, for each condition both distributions were seen to have a close to inverse exponential distribution. This suggested that the relationship could be modelled as a Poisson process, where the time between independent events of a continuous random process follows such a distribution. Although occlusion is technically not an independent event as marker visibility depends upon a large number of variables, which interact in a non linear fashion, its behaviour was sufficiently complex to be approximated as such [77].

By fitting the exponential probability density function (PDF) to the histogram data, the probability of occlusion event period and duration was modelled as a function of the number of cameras in the measurement configuration. This is achieved by normalising the volume under each histogram and fitting the rate values (μ_p and μ_d) for the exponential PDFs of

occlusion period and duration respectively. Figures 7.16 and 7.17 show the resulting distributions for event period and duration.

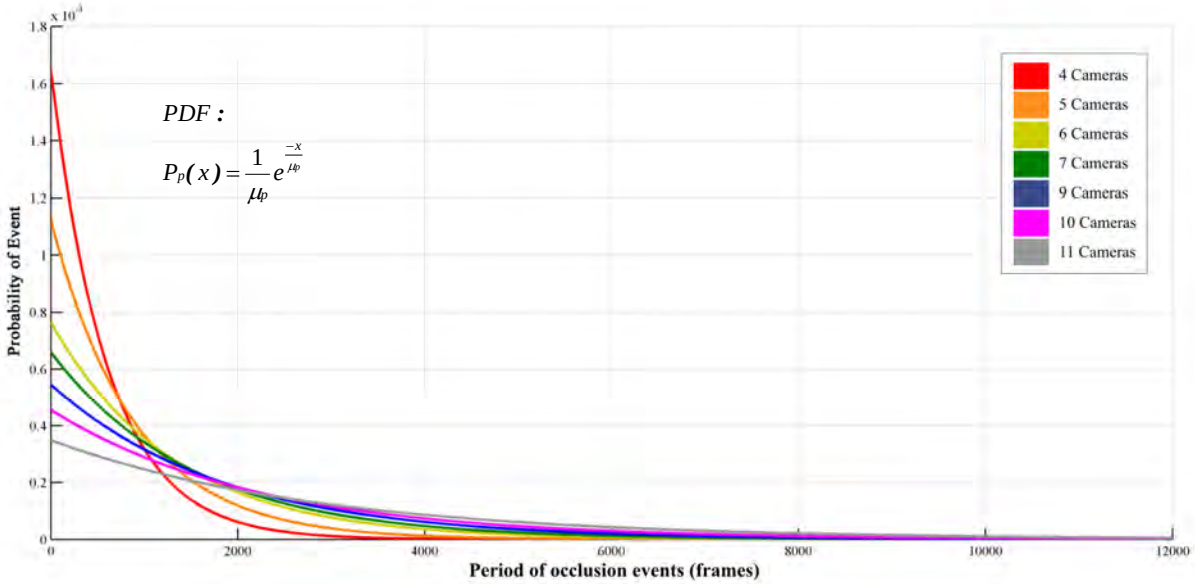


Figure 7.16 – Fitted probability of occlusion period

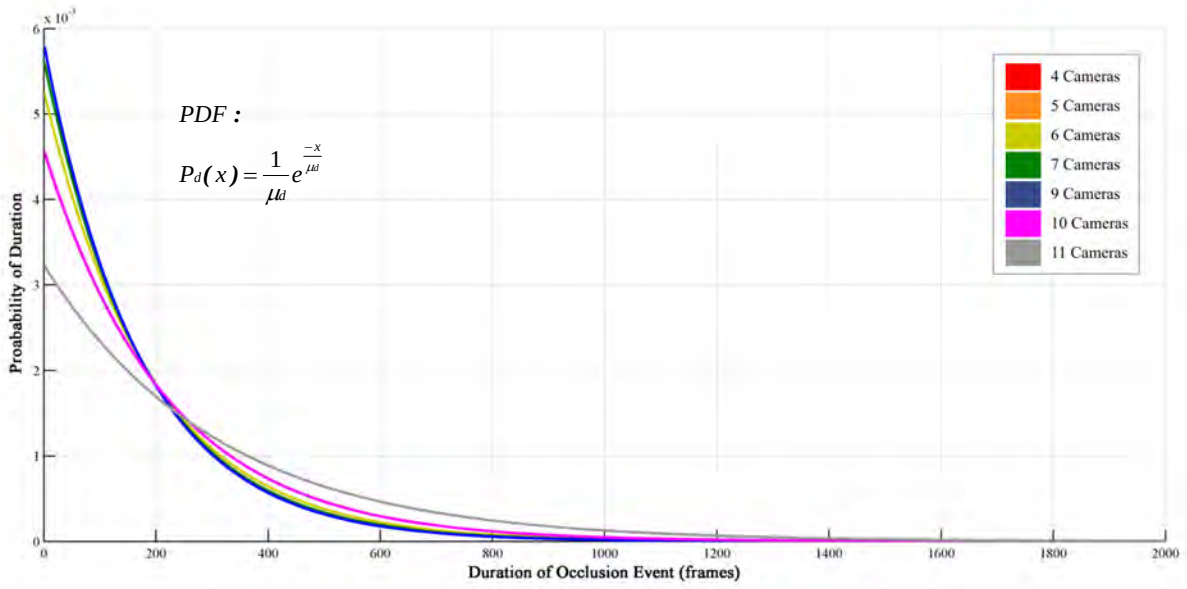


Figure 7.17 – Fitted probability of occlusion duration

Increasing the number of cameras has a 'flattening' effect on the period distribution of occlusion events. In the case of the duration distribution, after the difference in numbers of events have been corrected for by normalisation, the actual duration of occlusion events varies only slightly with the number of cameras. Additionally, in all cases, occlusion events of over 2000 frames in length were found to be extremely improbable, occurring in approximately 1/100kth of occlusion events. This value provided the upper limit of occlusion duration for the testing of the occlusion reconstruction algorithm.

As shown in figure 7.18, the variation of the rate parameter μ_P for occlusion period can be approximated as a linear function of the number of cameras in the measurement configuration. Therefore, by finding the least squares fit to this series and substituting into the exponential PDF, the probability distribution of occlusion period (P_P) is approximated. For the experimental data, where c is the number of cameras, this was found to be:

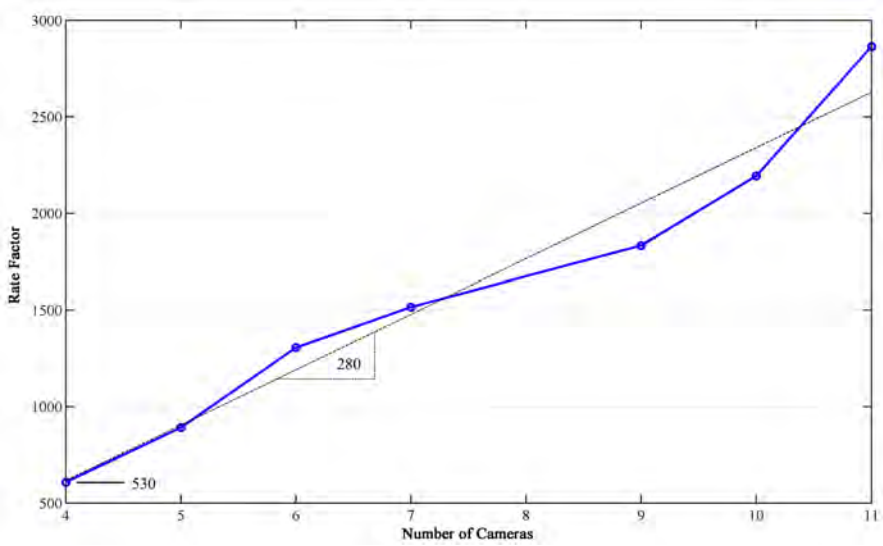


Figure 7.18 – Rate as a function of number of cameras

$$P_P(x,c) = \frac{1}{280c + 530} e^{\frac{-x}{280c+530}}$$

Correspondingly, for occlusion event duration (*Pd)*, the average distribution across all camera sets was:

$$P_d(x) = \frac{1}{210} e^{\frac{-x}{210}}$$

For the simulation of occlusion, each occlusion event was determined by picking a point from each of these distributions via a pseudo random number generator. This generates a table of start and end points for occlusion of each marker over the duration of a motion sequence. This data was then used in combination with the trajectory anonymisation algorithm of (12.1.2) to artificially fragment the gold standard data.

7.2.3 Behaviour of statistical model

As this model did not account for marker position, all markers in the set had equal probability of occlusion. Although the model could not account for variation due to pose of camera location, it would have been possible to provide a different exponential fit for each marker in the set. However, as noted in (13.3.1) uneven occlusion performance across the marker set was undesirable. Additionally, this model was specific to the experimental conditions from which it was defined and was therefore less flexible than the volume model. However, as the conditions over the experimental sessions were consistent this method was well suited to this project. Finally, the statistical method was considerably more computationally efficient than the volume model.

7.3 Comparison of occlusion models

The two occlusion models presented in this chapter addressed the problem of simulating occlusion by two different approaches. The volume model used a geometric method to estimate the interruption of LoS as a function of the position of the cameras, markers and measurement subject. It was intended that this method would provide close correspondence between simulated and actual occlusion events. However, difficulties in specifying the geometry of the subject's body and accommodating the stochastic processes inherent to occlusion severely compromised its accuracy.

In contrast to the volume model, the statistical model estimated the occlusion of each marker as a continuous random process. Although this method preserved the distribution of occlusion events over time, they were independent of both camera position and subject pose and therefore, bore little spatial resemblance to measured occlusion. Additionally, for any change in experimental conditions it was necessary to redefine the model from the distribution of occlusion obtained from measured trajectory data.

Due to the poor resemblance of the volume model to the actual occluded data and the necessity for even distribution of occlusion (13.3.1), the data produced by the statistical model was better suited for testing of the occlusion reconstructor. Furthermore, its improved computational efficiency facilitated the generation of a much larger occluded data set, allowing for thorough testing to be conducted. For these reasons the statistical method was chosen as the preferred means of occlusion simulation for the remainder of this thesis.

8. SPECIFICATION OF A PRACTICAL IMU SYSTEM

This chapter presents the specification and implementation of an inertial measurement system which fulfils the requirements of this project. Firstly, the requirements of an appropriate sensor are established and such a device selected. Secondly, the electrical and software requirements for interface and data collection are discussed. Finally, a number of basic tests are conducted to verify the functionality and performance of the apparatus.

8.1 Selection of an Inertial Measurement Unit (IMU)

8.1.1 The ADIS16350

Regarding selection of an appropriate IMU for the project, the ADIS family of sensors produced by Analog Devices Inc. was identified. Of particular interest was the ADIS16350 IMU module [36], which at the time of specification represented the state of the art in MEMS IMU design. The module integrates both a tri-axial accelerometer and tri-axial rate-gyroscope, as shown in figure 8.1. These sensors are capable of resolving the instantaneous linear acceleration and angular velocity about a central "point of percussion" in three orthogonal axis local to the devices body. Each of these axes are guaranteed calibrated and aligned (to within $<0.05^{o}$ of orthogonality between

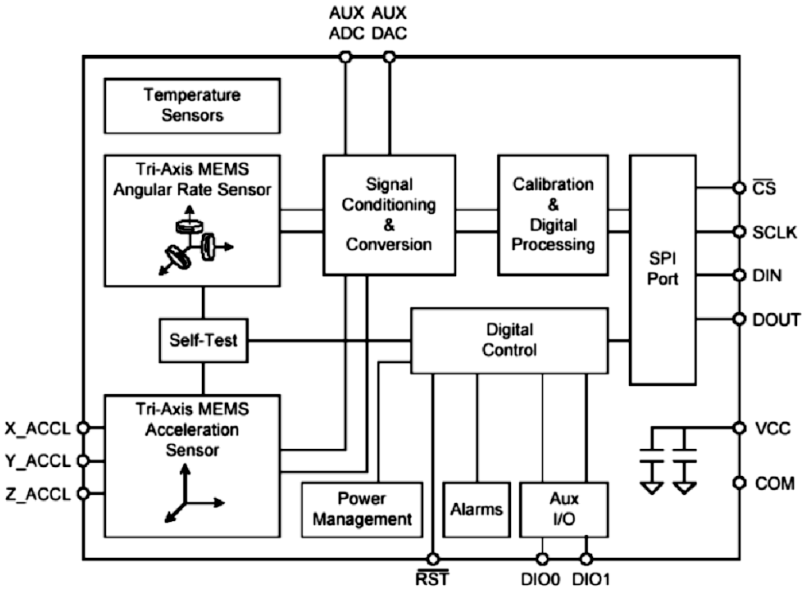


Figure 8.1 – IMU device structure (from [36])

x, *y* and *z*) at manufacture, eliminating the need for user calibration. Other features of the device include a self testing/calibration system, internal digital filtering and several general purpose auxiliary digital and analogue interface lines. The device has volume of approximately 8 cm^3 and weight of 16*g*, making it appropriate for attachment to a human subject without encumbrance. A maximum over-range of 2000g before the device suffers physical damage is also an advantage, given the rough treatment expected of the device during motion measurements.

The host interface to the device is via a high speed synchronous Serial Peripheral Interface (SPI) port. The provision of a digital rather than analogue interface to the device reduces noise pickup over long cable runs, and the relatively few wires required by the bus (4 + power) reduces the complexity of connections to the device and measurement subject.

8.1.2 Performance requirements

For this application it is important that inertial sensor performance is adequate to capture the range of accelerations and rates of turn generated by human body segments. The chosen ADIS IMU provides a maximum analogue-to-digital conversion rate of 820*Hz*, and its maximum mechanical sensor bandwidth is 350*Hz* for both accelerometer and gyroscope. This sample rate provides excess beyond the Nyquist limit for the maximum measurement frequency of the sensor. An integrated low-pass digital filter performs anti-aliasing automatically with the selected sample rate.

To examine whether sensor bandwidth was sufficient for the requirements of this project, linear accelerations and rotational rates about joints were calculated for the extremities over

the range of motion sequence acquired in the pilot experiment. This sequence was chosen as it presents rapid and discontinuous motions, which better characterise peak segment acceleration than the tai-chi sequences. Using marker data from the forearm, lower leg and foot clusters, instantaneous translational acceleration was estimated by the 100*Hz* low-pass filtered second derivative of centroid trajectory. Rate of turn was found via the first derivative of twice the arc cosine of the real part of the quaternion cluster orientation estimate (6.1.4). The power spectrum of each of these signals was taken and is shown against frequency in

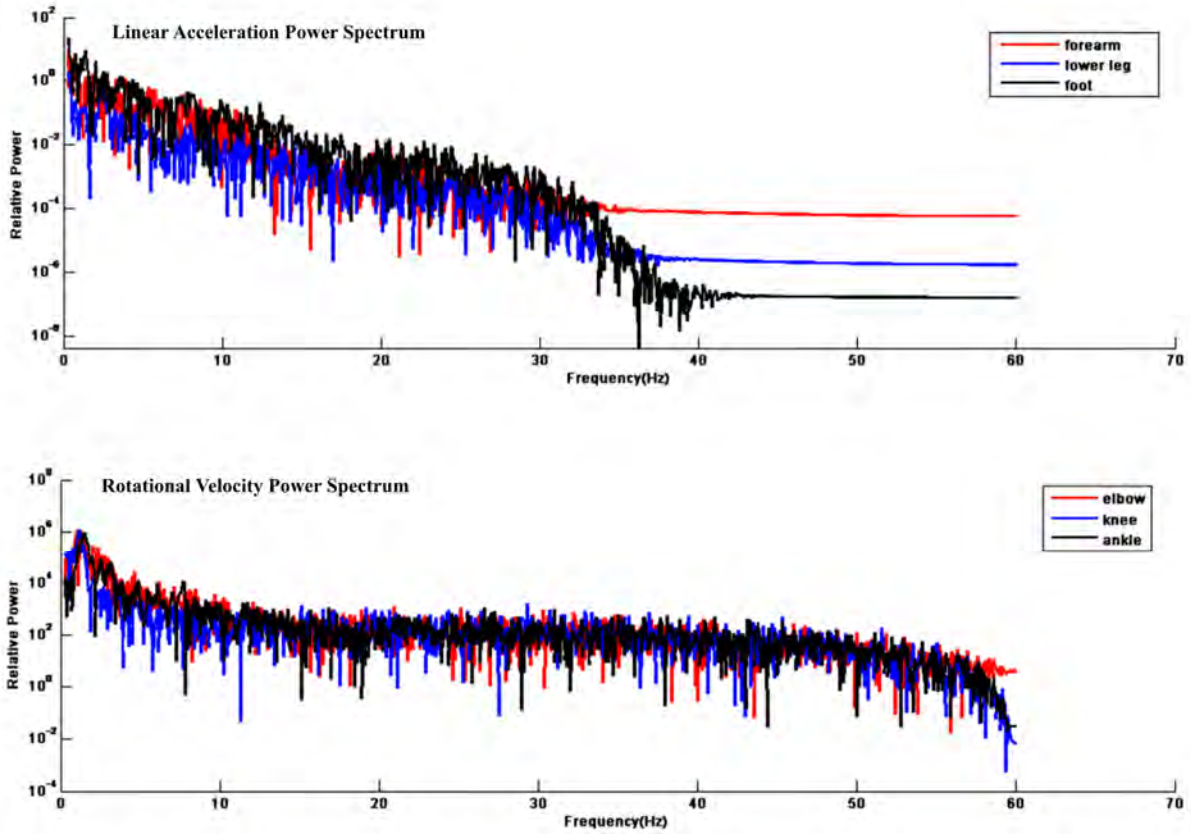


figure 8.2.

From this power distribution, the entire signal bandwidth for linear acceleration is seen to be in the sub 20*Hz* band. In the case of angular velocity, the majority of signal bandwidth is

below 10*Hz* with additional noise components occurring due to position quantisation. These results indicate that the 350*Hz* bandwidth of the ADIS16350 is more than adequate for purpose. Given this excess, the internal sampling rate of the IMU is reduced to ~550*Hz* (Actually 546.1347*Hz* due to hardware limitations) to reduce data transmission and storage requirements. The internal anti-aliasing filter automatically compensates for this change.

As well as the need to provide the necessary measurement bandwidth, the IMU must also provide adequate dynamic range for normal accelerations/angular rates of subject's limbs. Using the data from the frequency analysis, the absolute accelerations and angular rates over the course of a trial are plotted in figure 8.3.

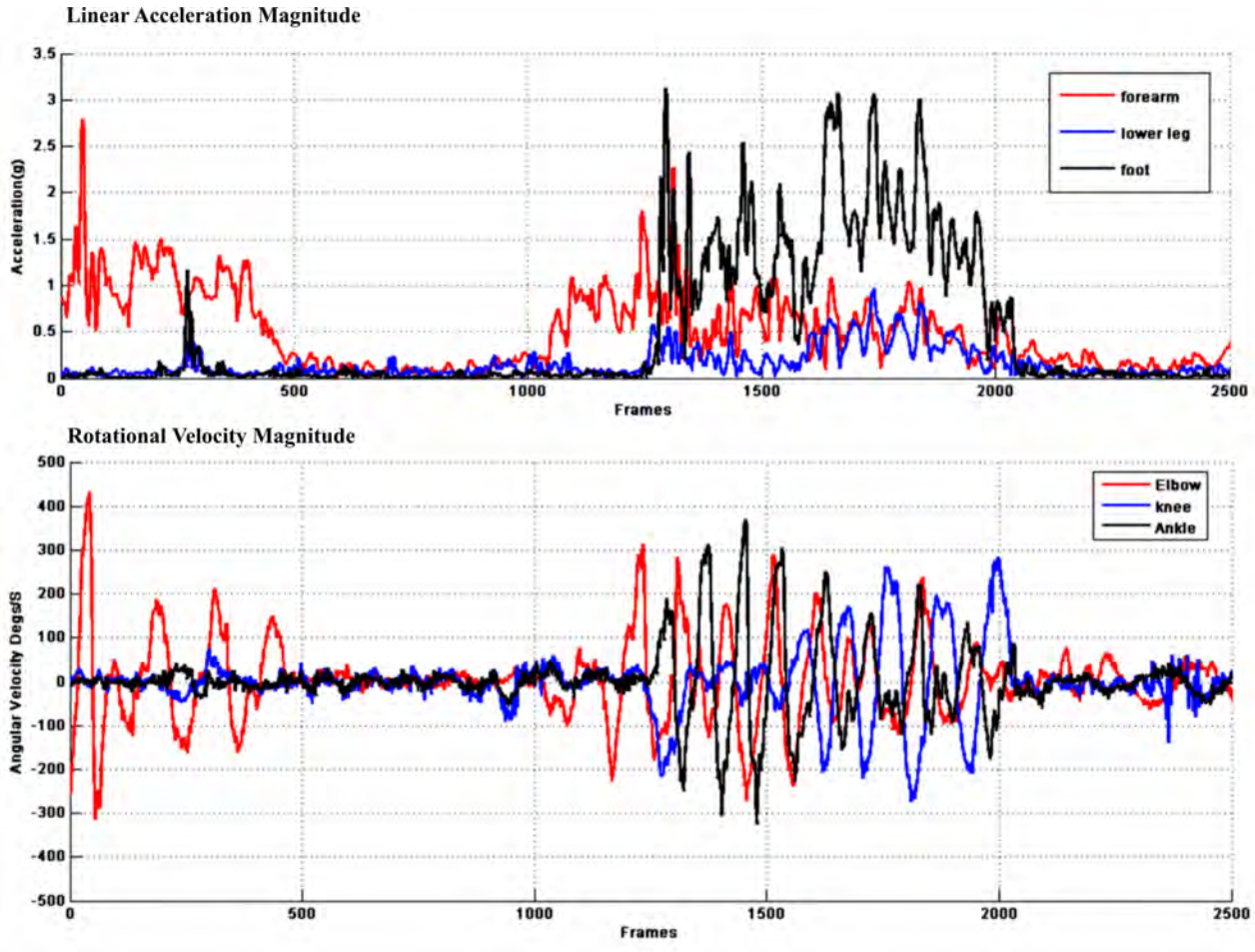


Figure 8.3 – Range of acceleration and rate of turn

The ADIS device features several gyroscope sensitivities which are selectable in software, with the widest of these giving +/- $300^{o}/s$ measurement range. The accelerometer, however, has a fixed range of +/- 10g. With reference to the plot, peak acceleration of body segments is around 3g, therefore the 10g range is adequate. However, with regard to angular velocity, both the elbow and ankle joints are noted to generate peak rates of up to $400^{o}/s$, in excess of the maximum range. Although this presents a problem for the use of the device, these rates occur infrequently in normal motion data, particularly in the Tai-Chi sequences of the experimental session. Furthermore, as no higher dynamic range IMU with similar specification was available, there was little choice but to accept this limitation. Any errors introduced by saturation will be mitigated by either software error correction or simulation, as discussed later in this thesis.

As the measurements from the IMU are used for the calculation of body pose via estimation of sensor orientation, the impact of sensor resolution upon results is difficult to predict. Therefore, until further evidence is available it will be assumed that the devices internal 14-bit converter will be sufficient for purpose (2.5mg and $0.07^{o}/sec$ at $300^{o}/sec$ range)

In addition to the parameters discussed, many other factors such as offset drift, sensor noise and nonlinearity will also influence the suitability of an IMU to a particular application. However, like resolution, the extent to which these parameters will affect performance will need to be established empirically.

8.2 Implementation of a single body mountable IMU unit

The following section discusses the integration of the IMU with the optical motion capture system. Mechanical construction of an instrumented marker cluster is discussed, as well as the electrical implementation of an appropriate embedded data acquisition system.

8.2.1 Embedded microcontroller system

All communication and control functions for the IMU are implemented via a dedicated microcontroller system, whose function is to bridge the low level serial communication protocol of the IMU to a host PC for the purposes of mass data storage and simplified user interaction.

The acquisition system was based around a third party development board. The Modtronix SCB65EC [78] was found to satisfy the requirements of the project and is shown in its stock condition in figure 8.4. The device is based around an 8-bit PIC18F6621 Microcontroller Unit (MCU) [79] and features a maximum instruction cycle frequency of 40*MHz*, a 64kB program memory, 4kB data memory as well as many integrated peripherals. The most important of these are the dual Master

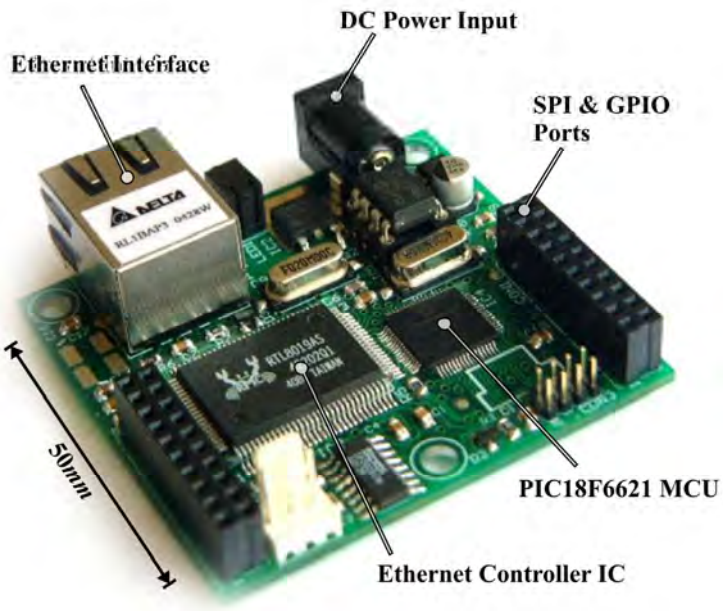


Figure 8.4 – The SCB65EC board [77]

Synchronous Serial Port controllers (MSSPs) which can be configured for SPI operation up to10*MHz*.

For the required IMU data rate of ~550*Hz*, reading all six (acceleration plus gyroscope) 16-bit measurement registers (14-bits measurement + flags) of the IMU results in a total bandwidth of 52.8*kbits/s*. In relation to the 40*MHz* clock speed of the MCU this allows for the execution of ~760 instruction cycles per full sample frame. This is sufficient for the necessary control and TCP stack operations for data transmission. Additionally, the 64kB program space is more than adequate for firmware given the RISC (Reduced Instruction Set) architecture of the MCU (actual firmware totalled 42kB). Finally, as little measurement data is stored upon the micro controller, the 4kB data memory provides ample buffer space for intermediate measurement data.

A feature which is not native to the MCU but provided by the SBC65EC module, is a 10Mbit Ethernet chipset (RTL8019AS) and connector. When accompanied by a firmware implementation of the TCP/IP protocol stack this allows the MCU to communicate via any Ethernet based Local Area Network (LAN). This provides a standardised interface by which measurement and control data may be exchanged with a host PC. Other useful characteristics of the SBC module include a small physical footprint (approx 50x50mm^2), integrated voltage regulation and an on-board RS-232 level converter for basic serial communications. Programming of the microcontroller firmware is performed over Ethernet, using an integrated boot-loader.

8.2.2 Instrumented cluster

The marker clusters used in the full body marker set are composed of rigid, epoxy impregnated carbon fibre plates which reduce inter-marker motion and soft tissue movement artefacts [58]. Additionally, they present a convenient and stable mounting for the IMU (Figure 8.5). To accommodate the curved surface of the plates and the flat base of the IMU, the interface between the two bodies was achieved via a 0.5*mm* steel plate upon which the IMU was attached. A 4*mm* layer of firm rubber between the two parts provided the compliance necessary to ensure reliable contact. The steel plate also served as a mounting location for a 9-pin female D-sub socket and adapter Printed Circuit Board (PCB), which provided a robust electrical interface between the IMU and tethering cable. Power supply decoupling was also included on the PCB.

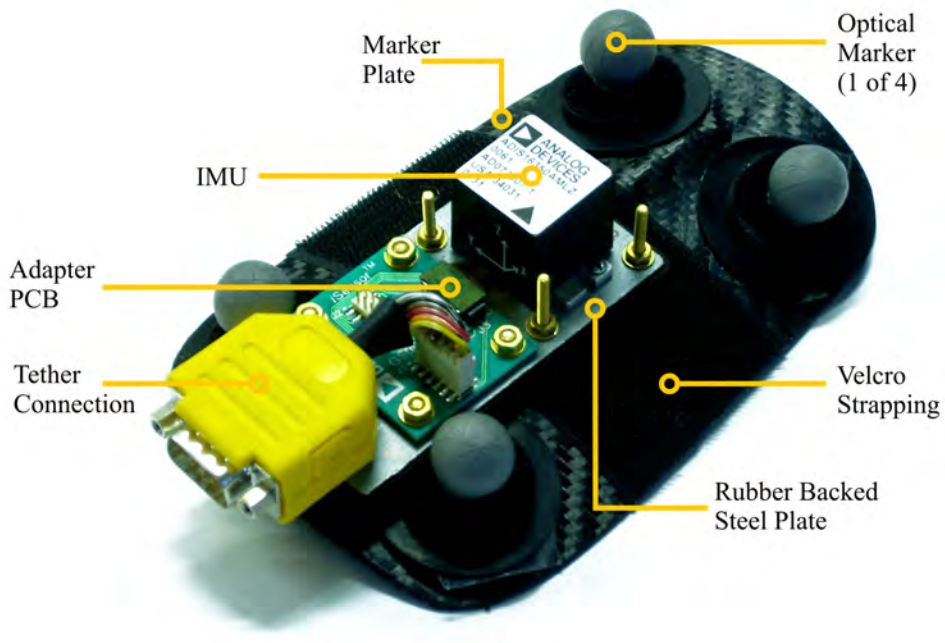


Figure 8.5 – Instrumented cluster

The additional weight of the IMU and mounting hardware (~80*g*) will increase soft tissue movement artefacts somewhat, due to its inertia. To compensate for this additional strapping was applied to the cluster. To reduce the effects of the tethering cable pulling upon the cluster during movement, the cable was looped at the cluster and the excess was run through several extra straps away from the cluster. This is detailed further in (10.1.1).

8.2.3 Electrical specification

Electrical connection between the IMU and microcontroller was achieved via a 5 meter 5 way tether cable linking the instrumented cluster and SCB. Power for the IMU (5*V* DC) was derived directly from the SBCs internal regulator. The four lines of the IMU SPI interface were tied directly to the three microcontroller GPIO (General Purpose Input/Output) pins, which were mapped to the first MSSP controller (Port RC3-5). A forth line from this port (RC6) was configured as a generic digital output for control of the device's Chip-select line (CS). As line RC3 was not routed to a header on the SCB, a wire link was added to the PCB to accommodate this. Due to impedance and timing issues encountered during firmware design a NAND gate (74HCT00) was used to buffer for the CS signal between MCU and IMU.

To provide testing and status indication during operation and debugging, the pins of the MCU's second port (RB) were connected via limiting resistors (330Ω) to a 8 segment LED bar display. For synchronisation with the frame pulse of the motion capture system, a buffered TTL input was added to line RC7. Finally, IMU, Ethernet and synchronisation signals were broken out via appropriate connectors and the power to the system was provided by a

medically approved 12*V* DC Power supply. A simplified block diagram of the completed system is shown in figure 8.6.

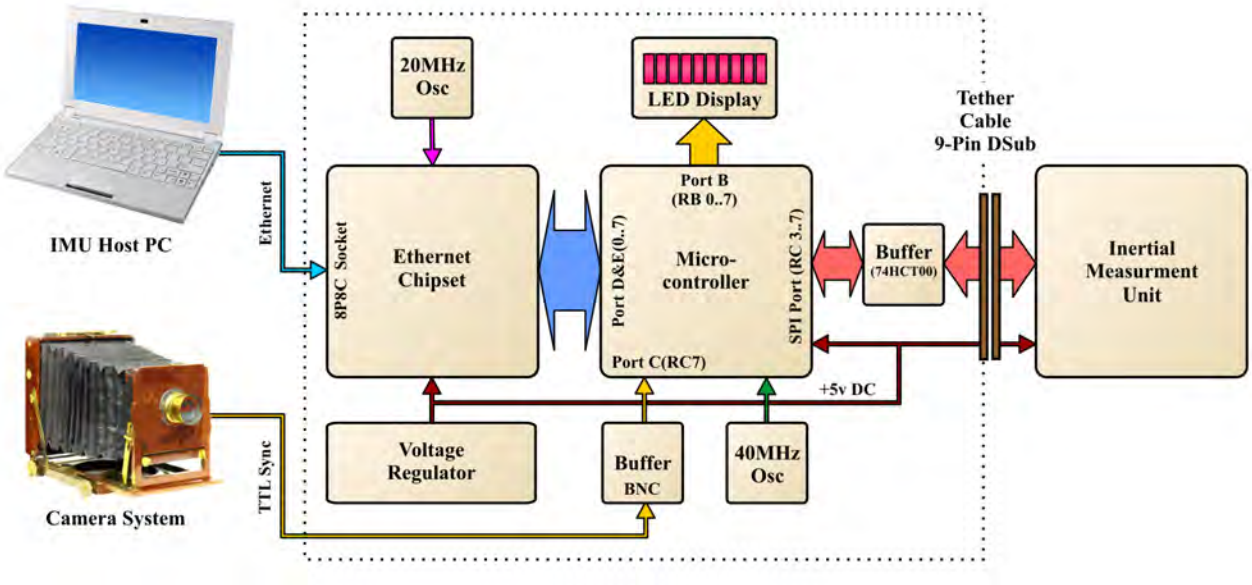


Figure 8.6 – Acquisition device block diagram

8.3 Software development

This section describes the development of the firmware for low level control of the IMU and the software application for data storage on the host PC. This effort is divided into three tasks: Interfacing the IMU with MCU via the SPI port, establishing a bidirectional TCP connection between host PC and MCU and implementation of a control system and Graphical User Interface (GUI) for the apparatus on the host PC.

8.3.1 IMU sampling routines

All firmware development was conducted using the standard PIC18 MCU development suite of Microchip MPLAB, C18 libraries and compiler for C language support. Sampling of the IMU was achieved via a polling mechanism, where the main program loop controlled the initiation of SPI transactions. In each transaction the MCU acts as the master, initiating data transfer by sending a read/write instruction corresponding to an internal address of the slave IMU [36]. For each instruction to be received properly by the IMU, 16-bits must be transmitted over sixteen rising transitions of a 2*MHz* Clock signal. As the SPI bus operates in a symmetrical manner, while instruction bytes are being transmitted to the IMU, data bytes are simultaneously being received. This complicates parsing of measurement data as the response to a previously transmitted command is only received as the next command is transmitted. Additional problems were caused by the 16-bit frame length used by IMU, as the MCU is limited to 8-bit operation. To overcome this, two successive 8-bit transmissions were combined and timing adjusted by software delay to meet bus requirements. Details of a typical transaction for the data input and output (SI and SO), clock (SCLK) and Chip select (CS) lines of the bus are shown in figure 8.7. Key timings maintained by the firmware are: CS to clock delay (T_{CS}), inter byte delay (T_{HF}) and inter frame delay (T_F).

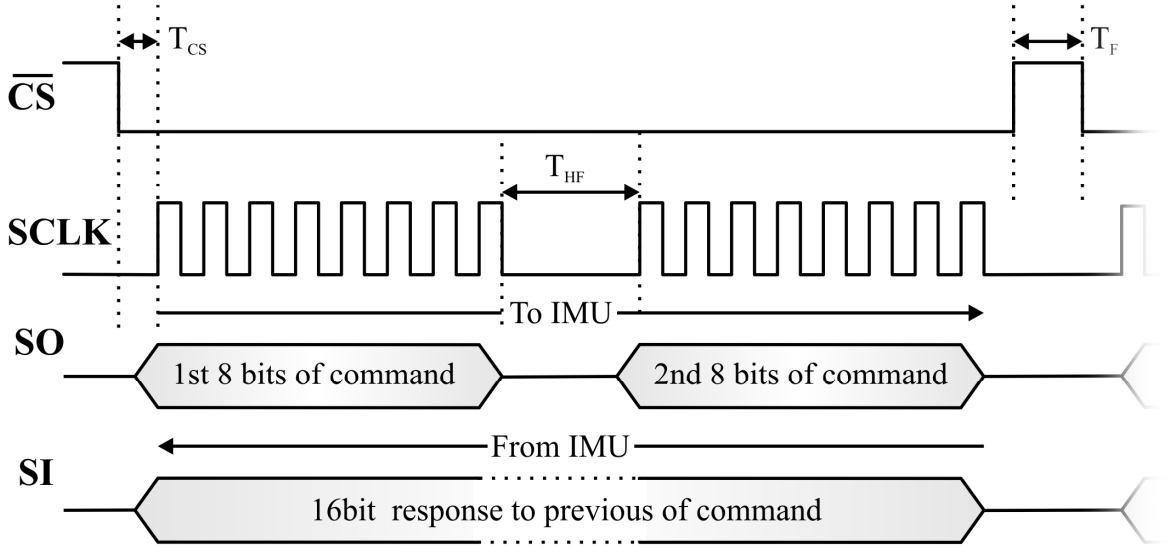


Figure 8.7 – IMU SPI timing diagram

In response to an IMU read instruction specifying the address of an measurement register internal to the IMU, a 16 bit value is returned, consisting of a padded 14-bit 2's complement sensor reading for a single axis. The upper two padding bits serve as flags indicating new data or sensor error. Via polling the 'new data' bit, the MCU interrogates the IMU until an updated sample is available. On reception of an updated sample the MCU increments the read address over successive instructions, retrieving data for all sensor axes. By repeatedly executing this sequence of operations the MCU collects all new samples from the sensor. Although an interrupt based system would have produced more efficient results, clashes occurred when operated alongside the TCP stack. Additionally, an extra signal line would have been required between MCU and IMU.

The completed IMU read/write operation was implemented in the firmware as a single C function call, simplifying further firmware development. For testing of the IMU read

operations, the data read from the IMU was transmitted to the development PC via RS232 link at a reduced IMU sample rate.

8.3.2 Ethernet - TCP connection

Implementation of the TCP stack on the microcontroller was achieved using the Modtronix TCP/IP libraries supplied with the SCB [78]. These libraries allowed for basic network functionality while avoiding direct interaction with the Ethernet chipset.

To test suitability to the application, firmware was written allowing the MCU to present itself as a server capable of receiving a simple TCP connection on a fixed IP address and port. When a connection to this server was established continuous TCP data packets were transmitted consisting of the repeated ASCII string “HelloWorld” and a 16-bit integer indicating the number of packets transmitted since connection (12 bytes). Rate of transmission was set to approximately 50*kbps*, to approximate the bit rate of a single IMU.

For the client side of the connection a simple 32-bit Microsoft windows application was built using the Borland C Builder 5 Integrated Development Environment (IDE). This IDE was chosen for its superior networking support via the integrated TCP/IP Visual Component Library (VCL) component. On initialisation this application attempts to open a TCP/IP connection to the MCU server on a specified IP address and waits for transmitted data. On reception of a packet it is compared with the expected test string and packet count, with any mismatch being noted in a simple GUI. Running this application in tandem with the MCU for several hours produced no detectable errors, suggesting adequacy of the link for transmission of IMU data.

8.3.3 Final acquisition system

For the implementation of the acquisition system, the IMU read/write function was implemented within the TCP server test firmware. For each call of the read function which returned new data, the six (*x*, *y*, *z*) accelerometer and gyroscope readings plus a 16-bit sample index and checksum were appended to the header of the TCP packet before transmission. Therefore, a packet was generated for each sample event, giving a rate of 550 packets per second. This data rate lay well within the 10*Mbps* bandwidth provided by the Ethernet link.

To facilitate reception of IMU data via TCP the client side application was modified to parse received data packets into discrete measurements, decode two's complement sensor readings and apply factory calibration factors. In addition to the primary measurement and storage functionally, supplementary features were added to the software and firmware. These included triggering of data collection via the MCU synchronisation input; reading of gyro temperature and supply voltage; as well support for the transmission of various sensor command codes. These codes allow for configuration of the IMU sample rate, sensor bias, anti-aliasing filter, as well as initiation of self diagnostic functions.

To improve usability of the system, all functions were made accessible via a simple GUI. For the purposes of user monitoring of signal quality, a real-time plotting function was included to display the most recent sensor data. This user interface is shown during measurement in figure 8.8.

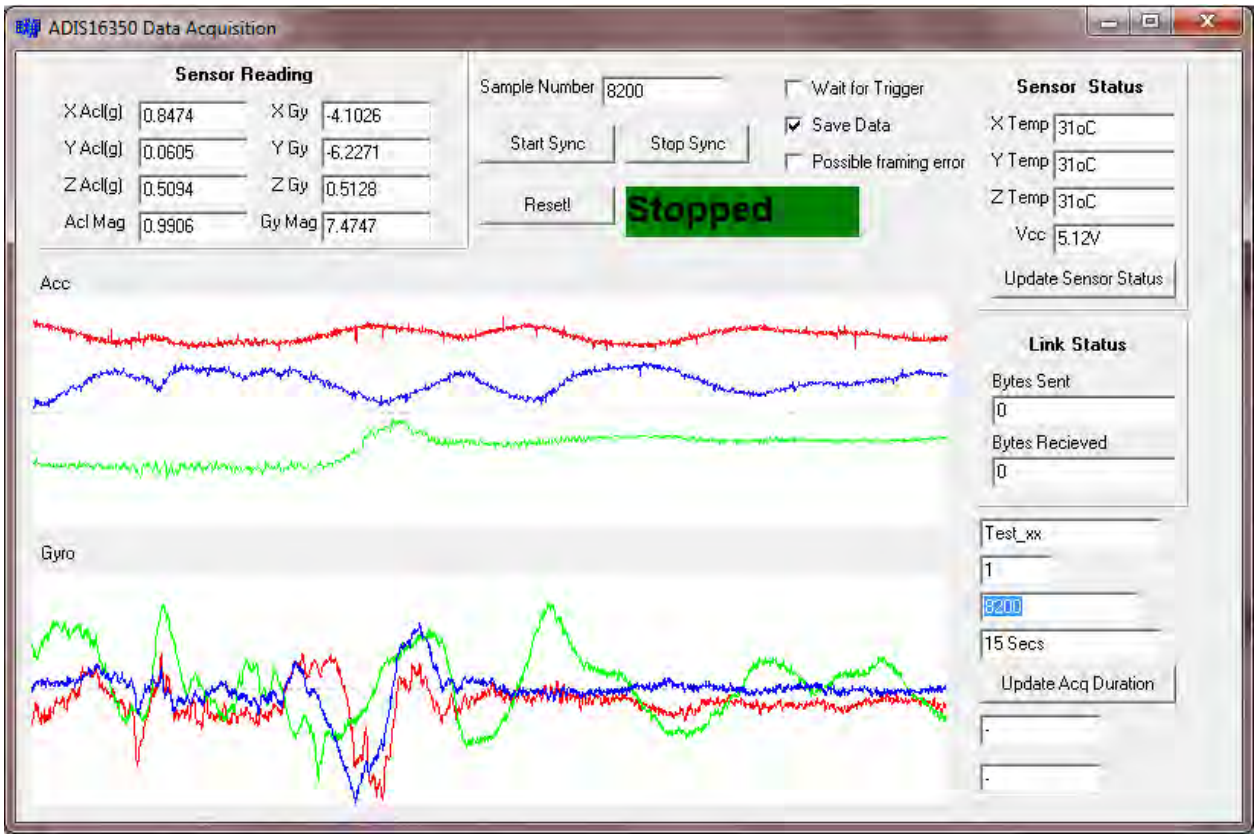


Figure 8.8 – IMU acquisition GUI

8.4 Testing and calibration of measurement system

Before experimental measurements were performed basic testing of the apparatus was performed. In this section the sensors self test functions are used to verify factory calibration and a limited comparison is made between known references and sensor measurements.

8.4.1 Self test

The integrated self-test functions of the IMU permit two levels of testing to be performed: The first, which reports the status of the IMU's internal logic and sensors; and the second, where a constant mechanical bias is applied to the accelerometer and gyroscope, verifying integrity of the entire signal path. Following construction of the apparatus both tests were executed and found to yield results in agreement with the sensor specification [36].

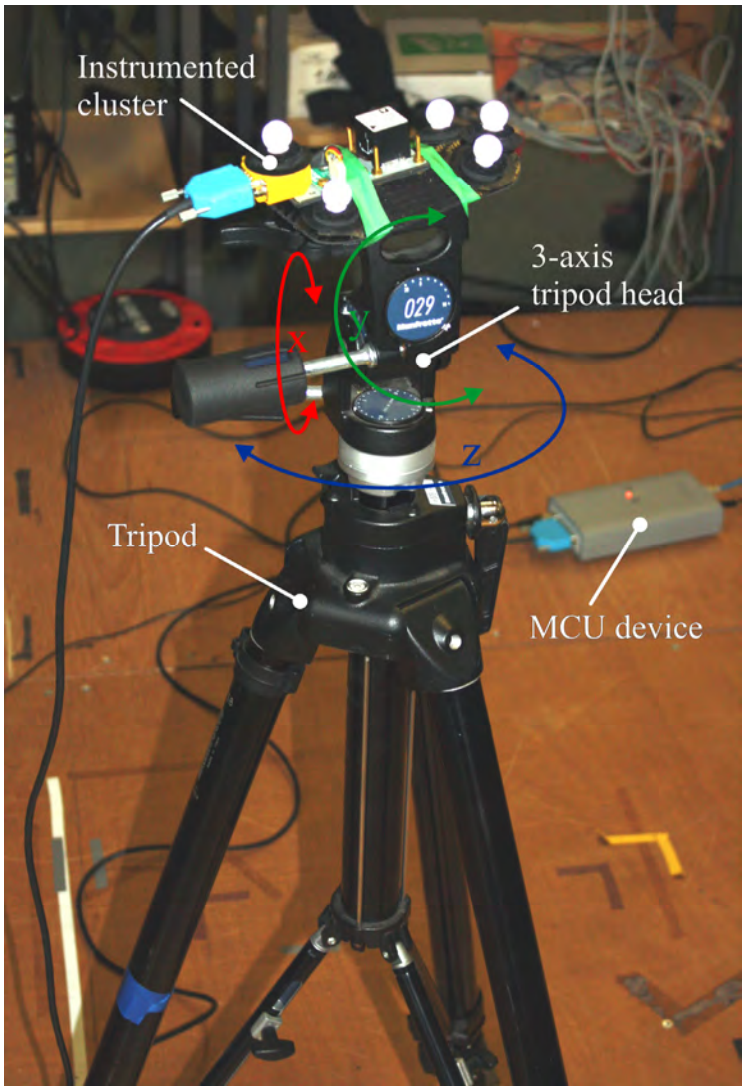


Figure 8.9 – Tripod test setup

8.4.2 Field test

For experimental testing of the apparatus, the instrumented cluster was mounted atop a standard 3DoF photographic tripod to allow for control of its attitude. Alignment of the

articulated axes of the tripod with the sensing axes of the IMU was approximated manually. Reference data was provided by the optical motion capture system and the passive makers on the cluster. Synchronisation between the two systems was achieved manually due to limitations of the MCU firmware (Though in later experiments synchronisation was performed automatically via the optical system's TTL frame pulse). This setup is shown in figure 8.9 with articulated axes labelled (note the housed MCU device is also shown in the right of this image).

During the data collection optical and inertial measurement systems were triggered simultaneously for a duration of 10 seconds with an optical frame rate of 200*Hz*. For the first sequence the tripod head was rotated sequentially by +/- 90^o over each of its articulated axes and for the second a static measurement was taken. Approximately one hour later a second static acquisition was performed. The optical data for all three trials were then reconstructed and labelled via QTM. Both data sets were imported into Matlab and the IMU data downsampled via nearest neighbour to the frame rate of the optical data.

8.4.3 Accelerometer Test

During correct operation, if the sensor is not subjected to translational motion the accelerometer will only record a constant vertical acceleration of 1g due to gravity. Therefore, regardless of sensor orientation, the total magnitude of measured acceleration across all sensing axes should also equal 1g (assuming zero bias). Figure 8.10 shows the 10*Hz* low-pass filtered x, y and z acceleration components recorded by IMU during the rotational movement of the tripod head. The total magnitude of acceleration, shown by the black line, approximates 1g throughout the trial suggesting correct operation of the device. Although some transient

deviations are observed during rotation, these are attributed the IMU being mounted somewhat off the axis of rotation of the tripod.

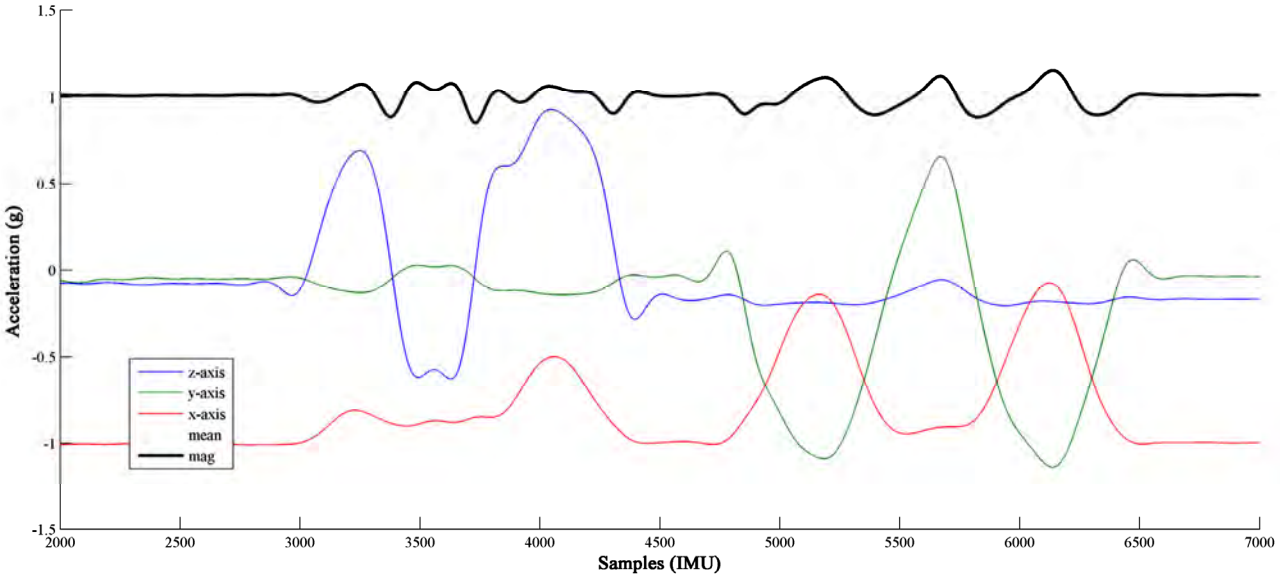


Figure 8.10 – Acceleration vector components and magnitude over yaw, pitch and roll motion

8.4.4 Gyroscope Test

Firstly, an assessment of 'in-run' bias stability (The stability of bias within a single continuous measurement) for the gyroscope was conducted. This was obtained from the difference of average gyroscope readings for each axis over the two static acquisitions. Over the hour long interval between measurements these differences were found to be $-0.0081^{o}/sec$, $0.0048^{o}/sec$ and $-0.0215^{o}/sec$ for the x, y and z axes respectively. This lies within the specification of $0.015^{o}/sec$ in run stability given in the IMU specification.

Preliminary testing of the dynamic response of the gyroscope was achieved through comparison of IMU and optical measurements over the 3-axis tripod movement. Instantaneous magnitude of angular velocity was obtained as the first derivative of the axis angle orientation estimate (as in 8.1.2). This value was low-pass, zero lag filtered at $10Hz$ to

reduce differentiation noise. In the case of the IMU angular rate, the magnitude for all 3 gyroscope axes was calculated and filtered in an identical manner. The two rate magnitude estimates are shown in figure 8.11 and exhibit excellent correspondence to each other. Peak error is approximately $25^{o}/sec$, occurring at ~1050 samples where angular rate exceeded the sensors dynamic range.

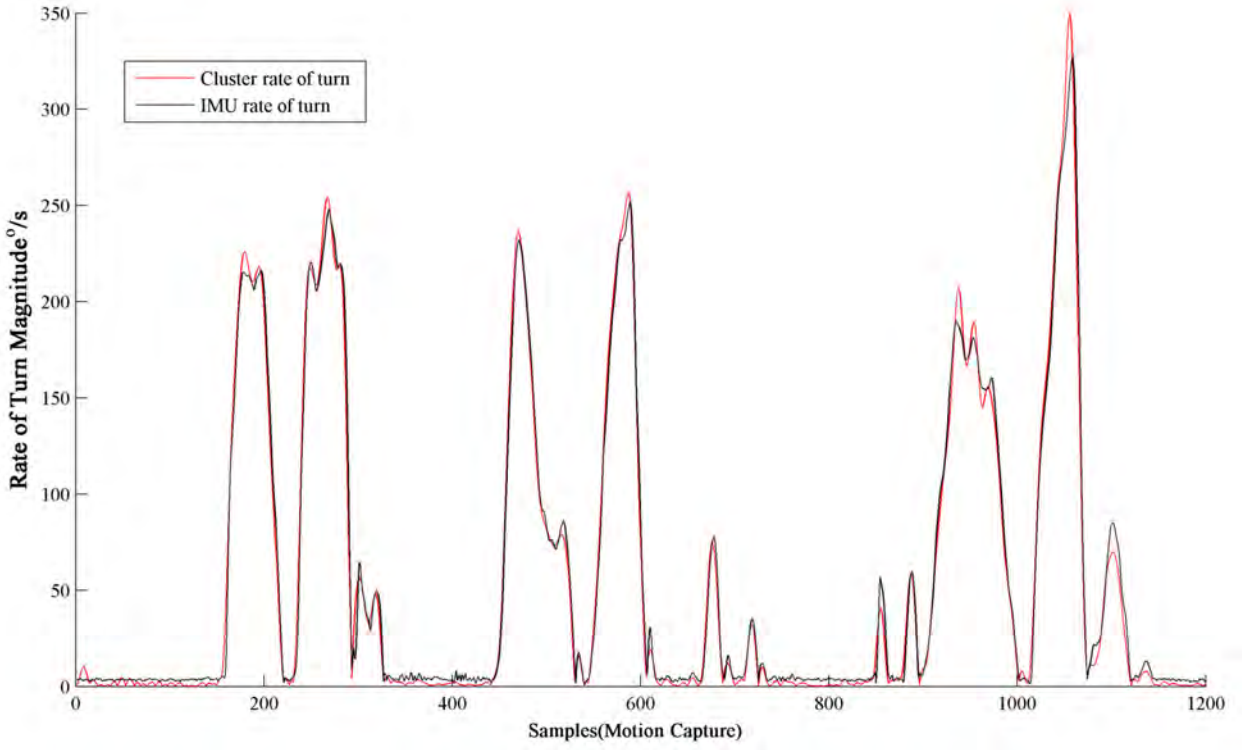


Figure 8.11 – Measured and simulated rate of turn magnitude over yaw, pitch and roll motion

Despite good magnitude correspondence the arbitrary alignment between the sensing axes of the IMU and thecluster coordinate frame leads to poor agreement in direction of acceleration and rate estimates. As successful application of IMU readings depend upon agreement of these two frames it is essential that this alignment be quantified. This forms the basis for the sensor model and is discussed in the next chapter.

9. MODELLING OF FULL BODY INERTIAL MEASUREMENTS

This chapter describes a sensor model, which allows simulation of inertial measurements from optical motion capture data. In the first two sections the requirements of the model are addressed: The alignment of the sensor and reference frames, estimation of gyroscopic and accelerometric measurements and modelling of noise and bias characteristics. In the third section the model is tested against real measurement data and applied to full body data.

9.1 Alignment of IMU frame of reference with cluster

To make useful comparisons between the IMU and optical measurements it is necessary to establish the rigid body transform between the axis system of the sensor and the orientation estimation of the marker cluster.

9.1.1 Alignment by gravitational acceleration

The estimation of rotational alignment between optical and IMU axes was obtained by exploiting the constant influence of gravity upon the measurement of the accelerometer. By calibrating the optical system to ensure alignment between its –*z* axis and acceleration due to gravity, a common vector by which the two frames were related was defined. Therefore, the alignment between the two frames was described by the rotation which mapped the gravitational vector in the sensor frame to the -*z* axis of the optical frame over several distinct orientations of the IMU. The relationships between the frames of the sensor, cluster and global axes are shown in figure 9.1. The difference in alignment between sensor and cluster is given by the rotation matrix ($\mathbf{R}_{al}$), the estimated cluster orientation ($\mathbf{R}_{cl}$) and the IMU orientation estimate ($\mathbf{R}_{imu}$). Estimates of the global -*z* axis in optical and IMU reference frames are given by ($\mathbf{z}_{glob}$) and ($\mathbf{z}_{imu}$) respectively.

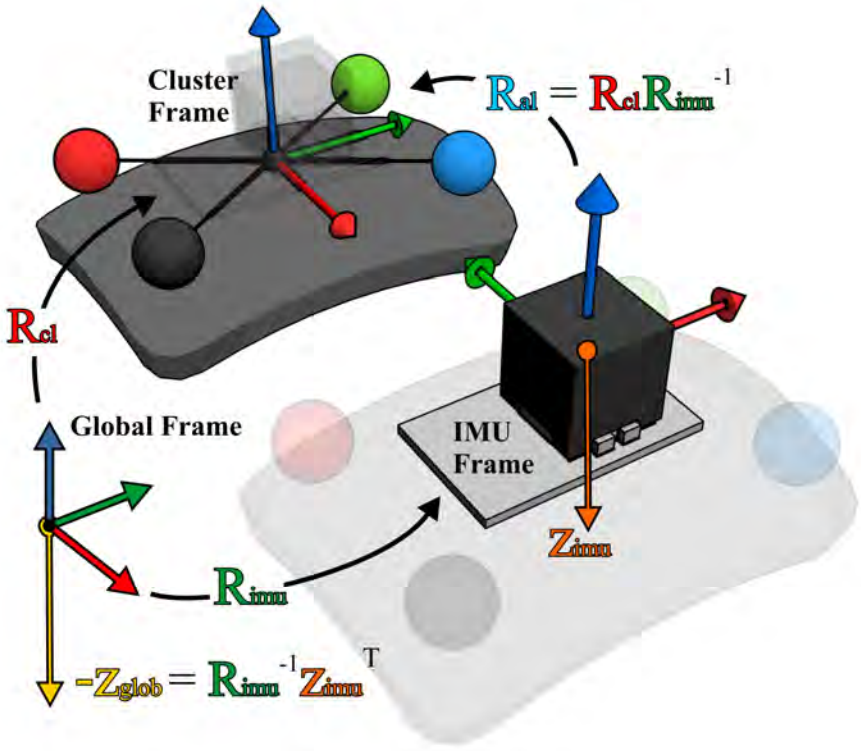


Figure 9.1 – Basic alignment transform

9.1.2 Detection of stationary points

Translational movement of the IMU during alignment calibration resulted in accelerations which perturbed the estimate of the gravitational vector. This made estimation of alignment unreliable in all but stationary settings.

To minimise the influence of this translation upon estimation of cluster-IMU alignment from dynamic data, such as a subject calibration trial, it was necessary to detect points where translational acceleration was at a minimum. This was achieved by taking the standard deviation over a sliding window of the sample acceleration data. In the conditions where this value was found to be at a local minimum and below a fixed threshold, each component of acceleration was averaged over the window to provide the reference gravitational vector. At

the same point in the sequence the -*z* axis in the clusters local frame of reference is also recorded. This process was repeated for the whole sequence, giving multiple pairs of vectors for the -*z* axis. Figure 9.2 shows the results of applying the method to a dynamic test sequence obtained via the tripod setup of (8.4.2) and a window length of 100 samples.

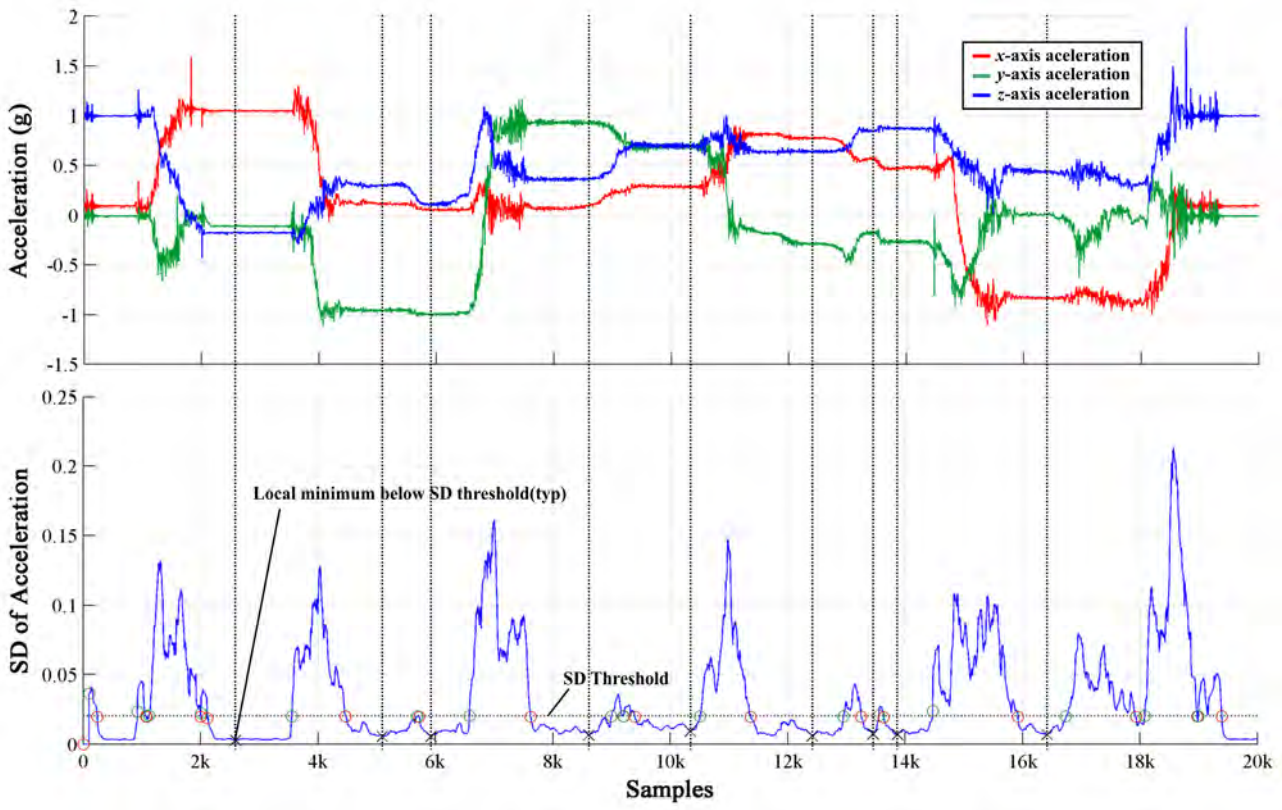


Figure 9.2 – Gated variance of acceleration

9.1.3 Cumulative estimation of rotational alignment

As multiple rotational transforms exist between two vectors, a single measure of gravitational acceleration will only partially define the relationship between the two frames of reference. This is due to freedom of rotation about the axis defined by the direction of the gravitational vector. However, by combining several sets of gravitational readings obtained over a range of different orientations it is possible to constrain the alignment estimate to a single rotation.

This was achieved by adapting the quaternion orientation estimation method of (6.1.4) to minimise the angle between vectors rather than points about a centroid. By normalising the sets of vectors obtained via the stationary point detection and applying this orientation estimation, an approximation of the alignment rotation was obtained. Furthermore, as this method relies upon the average angular error between vectors, it will tend to reject any perturbation from the actual alignment. This provides resistance to perturbation by translational acceleration.

The application of this method to the averaged pairs of vectors obtained in (9.1.2) is shown in figure 9.3 with each pair of vectors plotted before and after application of the estimated alignment rotation. Colours represent each corresponding set of vectors, with circles and crosses indicating optical and IMU derived vectors respectively.

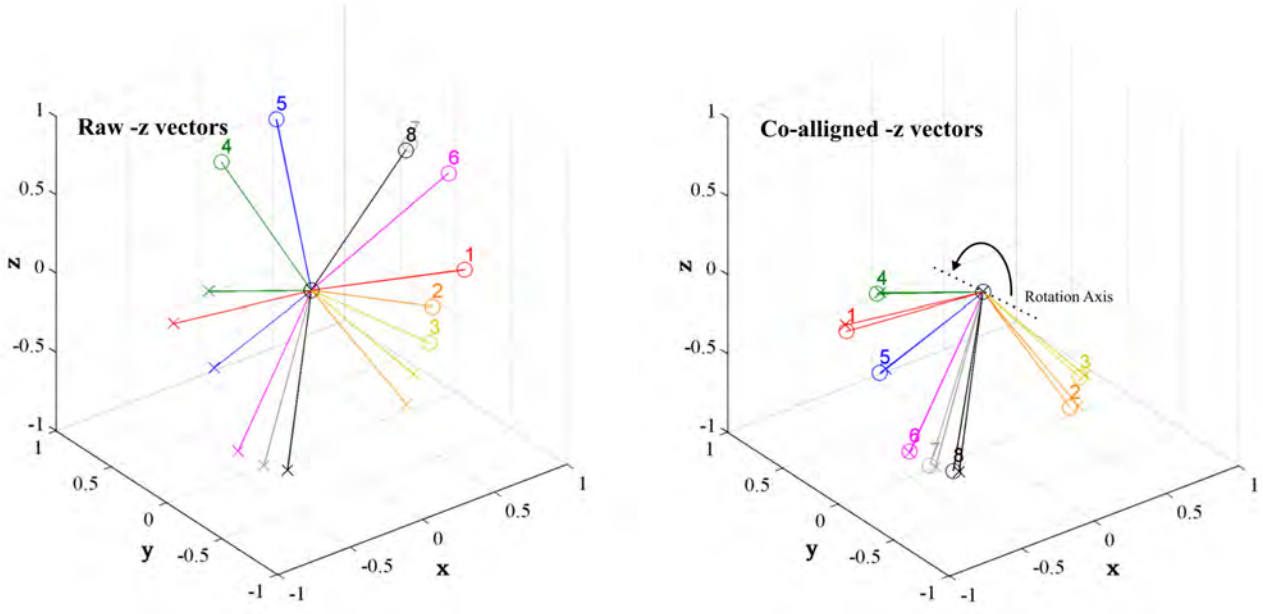


Figure 9.3 –Alignment of local gravity vectors by quaternion orientation method

Through testing of this alignment against simulated data it is found that for a set of approximately 10 pairs of vectors an error of 0.5-1^{o} between estimated and actual alignment is achieved.

As translational alignment ($\mathbf{t_{al}}$) cannot be established directly by inertial measurement, this parameter is obtained though manual measurements taken via a calibrated digitising pointer and the optical motion capture system. The alignment vector is approximated as the difference between the centroid of the marker cluster and the corner points of the IMU housing.

9.2 Simulation of IMU readings from optical data

The IMU model is based upon the alignment transform defined previously, which allows estimates of the translational acceleration and rate of turn measured in the global optical frame to be expressed in terms of the local reference frame of the IMU. By augmenting these estimates with models of the effects of gravity, sensor drift and noise, a complete simulation of IMU behaviour is obtained.

9.2.1 Modelling of IMU acceleration

Assuming the instantaneous orientation and translation of the instrumented cluster at sample time *t* is described by the rotation matrix ($\mathbf{R}_{cl_t}$) and vector ($\mathbf{t}_{cl_t}$); and the fixed orientation and translation between cluster and IMU by $\mathbf{R}_{al}^{-1}$ and $-\mathbf{t}_{al}$, any vector in the optical global frame ($\mathbf{v}_g$) can be described in the IMU local frame as ($\mathbf{v}_l$), by the combined transform:

$$\mathbf{v}_l = \mathbf{R}_{al}\mathbf{R}_{cl}{}^{-1}\mathbf{v}_g{}^{T}$$

Substituting $\mathbf{v}_g$ with the 2nd derivative of $\mathbf{t}_{cl_t}$, the estimated instantaneous linear acceleration of the cluster centroid in the local frame of the IMU is obtained ($\tilde{\mathbf{a}}_{lc}$):

$$\tilde{\mathbf{a}}_{lc_t} = \mathbf{R}_{al}\mathbf{R}_{cl_t}{}^{-1}\left(\frac{d^2\mathbf{t}_{cl_t}}{dt^2}\right)^{T}$$

As there is a translational offset between cluster centroid to point of percussion (the point to which measurement of acceleration and rates of turn are referenced), it is necessary to account for acceleration due to the moment about this point. In the global reference frame the

instantaneous vector from the centroid to point of percussion $\mathbf{v}_{\mathbf{ig}_t}$, is given as $-\mathbf{t}_{\mathbf{al}}$ rotated by the orientation of the cluster $\mathbf{R}_{\mathbf{cl}_t}$, or:

$$\mathbf{v}_{\mathbf{ig}_t} = \mathbf{R}_{\mathbf{cl}_t}(-\mathbf{t}_{\mathbf{al}})^T$$

The global instantaneous acceleration due to the moment about $\mathbf{t}_{\mathbf{cl}}$ is obtained by differentiating this vector. Adding this to the transformed component of the expression for $\tilde{\mathbf{a}}_{\mathbf{lc}_t}$, an expression for the estimated motion derived linear acceleration is obtained:

$$\tilde{\mathbf{a}}_{\mathbf{lm}_t} = \mathbf{R}_{\mathbf{al}}\mathbf{R}_{\mathbf{cl}_t}^{-1}\left(\frac{d^2\mathbf{t}_{\mathbf{cl}_t}}{dt^2} + \frac{d^2\mathbf{v}_{\mathbf{ig}_t}}{dt^2}\right)^T$$

In addition to acceleration derived from translational motion, the accelerometer registers acceleration due to mechanical shock and gravity. As shock typically exhibits frequency components of hundreds of hertz and displacements less than a millimetre, it is beyond the measurement range of the optical motion capture system. Furthermore, given measurement data was low-pass filtered to reduce the influence of marker quantisation error, influence of these frequency ranges was considered to be negligible. Therefore, the effects of shock and impact were omitted from the sensor model.

Gravity is considered to impart a constant acceleration vector along the -*z* axis of the global frame of reference. Assigning this vector the identifier ($\mathbf{g}_{\mathbf{g}}$) and adding it to the global component of $\tilde{\mathbf{a}}_{\mathbf{lm}_t}$ the expression for total acceleration is:

$$\tilde{\mathbf{a}}_{\mathbf{l}_t} = \mathbf{R}_{\mathbf{al}}\mathbf{R}_{\mathbf{cl}_t}^{-1}\left(\frac{d^2\mathbf{t}_{\mathbf{cl}_t}}{dt^2} + \frac{d^2\mathbf{v}_{\mathbf{ig}_t}}{dt^2} + \mathbf{g}_{\mathbf{g}}\right)^T$$

The final acceleration estimate is divided by the constant of standard gravity g_0 (9806 mm/s^2), converting to the measurement units of the IMU (g).

Figure 9.4 illustrates each of the vectors and rotations which comprise the acceleration model in terms of the relationships between the cluster and IMU reference frames.

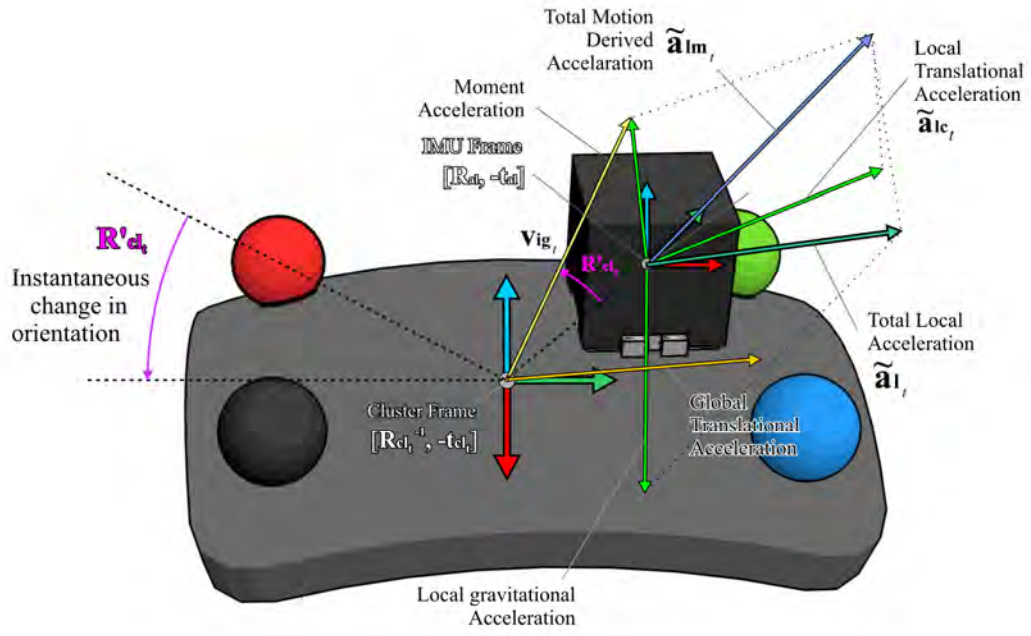


Figure 9.4 – Transform and Vectors in Acceleration Model

9.2.2 Modelling of IMU rate of turn

The model of the IMU gyroscope utilises the Euler decomposition of the IMU orientation to provide three scalar rotations about the cardinal axes of the sensor. To avoid the problem of 'gimbal lock' inherent in this representation (Appendix 8), the estimate is based on the relative change in orientation per sample, rather than the absolute orientation.

In the terms of the instantaneous cluster orientation estimate $\mathbf{R}_{\mathbf{cl}_t}$, the relative rotational transform from one instant to the next is given as:

$$\mathbf{R}'_{\mathbf{cl}_t} = \mathbf{R}_{\mathbf{cl}_{t-1}}{}^{-1}\mathbf{R}_{\mathbf{cl}_t}$$

Transforming this into the frame of reference of the IMU via $\mathbf{R}_{\mathbf{al}}$ gives:

$$\mathbf{R}'_{\mathbf{imu}_t} = (\mathbf{R}_{\mathbf{al}}{}^{-1}\mathbf{R}_{\mathbf{cl}_{t-1}})^{-1}(\mathbf{R}_{\mathbf{al}}{}^{-1}\mathbf{R}_{\mathbf{cl}_t})$$

This rotation matrix is then decomposed into an *x-y-z* Euler sequence, by application of the arc-sine and 4-quadrant arc tangent. By indexing the rotation matrix as $\mathbf{R}'_{\mathbf{imu}_{t(i,j)}}$, where i is the row and j is the column offset, the axial angular velocity vector per time step is given by:

$$\tilde{\boldsymbol{\varphi}}_{\mathbf{s}_t} = \left[sin^{-1}(\mathbf{R}'_{\mathbf{imu}_{t(3,2)}}) \quad tan^{-1}(-\mathbf{R}'_{\mathbf{imu}_{t(2,3)}}, \mathbf{R}'_{\mathbf{imu}_{t(2,2)}}) \quad tan^{-1}(-\mathbf{R}'_{\mathbf{imu}_{t(3,1)}}, \mathbf{R}'_{\mathbf{imu}_{t(1,1)}}) \right]$$

To ensure compatibility with real IMU measurements the components of this vector are multiplied by 180/π and the sample frequency of the acquisition system to convert from radians per sample to degrees per second.

By taking the difference in rotation before conversion to a Euler sequence the problems of gimbal lock associated with the Euler representation are avoided. Due to the relatively small changes in orientation per sample step ($<2^{o}/sample$ at the maximum expected rate of turn of $400^{o}/sec$), the estimate will always remain within the range of orientations permitted by the *x-y-z* Euler sequence. Therefore, the stability of the method is guaranteed for the ranges of angular velocities produced by body segments. Additionally, these small rotation angles will minimise the introduction of coupling errors between rotation axes introduced by non-commutativity of the Euler rotations.

9.2.3 Bias and measurement noise model

Taking the derivative of marker displacement for estimation of acceleration can result in high levels of noise due to marker quantisation error. To reduce the effect of this upon the sensor model it is necessary to filter the trajectory data before differentiation. To this end, a zero delay Butterworth low-pass filter with cutoff of $15Hz$ is used [38]. It is noted that although this filter has the potential to attenuate useful components of subject motion, typical bandwidths of motion for the body segments of interest are typically below the filter cutoff [63].

To establish the noise and offset characteristics of the accelerometer and gyro components of the IMU apparatus, three static trials of 100 second duration were collected via the tripod setup described in (8.4.2). After subtraction of the mean, the RMS value was taken for each axis of the sensor. The values obtained are shown in the following table:

	x-axis	y-axis	z-axis
Gyro Noise ($^o/s$) RMS	0.364	0.419	0.417
Accel Noise (mg) RMS	4.17	10.07	8.00

Table 9.1 – Measured gyroscope and accelerometer noise

In the case of the gyro noise level, the measured value was found to be in agreement with the IMU documentation of 0.60°/s RMS [36]. However, the value given for the accelerometer in the same documentation is 35mg RMS, which is considerably higher. This disparity was due to the internal digital filter being enabled for the lab tests, whereas it was disabled for the figures given in the documentation. With the filter disabled a value of ~40mg was recorded.

The second key parameter in quantifying the gyro performance is the rate at which the zero bias of the gyro will change over time (11.1.2). This was established experimentally using a 10 second sliding window mean over the length of each acquisition. The resultant series was then filtered with a cut-off of 0.1*Hz* and differentiated to find the instantaneous rate of bias drift. By taking the RMS value of this drift, the values of $0.0288^o/h$, $0.0338\,^o/h$ and $0.0340\,^o/h$ were obtained for the x, y and z axes respectively. These values were found to be considerably lower than the manufacturer's specified value of $4^o/h$. It is suspected that this disparity was due the relatively short period over which the bias drift was averaged. However, the later value is chosen for use in the sensor model to ensure bias drift is not underestimated.

To integrate these sources of error into the sensor model, the accelerometer and gyro sensor noise is modelled as white noise, first order low pass filtered with cutoff of 50*Hz*. This was found to yield a power spectrum approximating that found experimentally. The resultant data set is then normalised to give the desired RMS amplitude. The gyroscope bias drift is modelled as a random walk, randomly integrating either the positive or negative rate of bias

drift per sample with a probability of 50%. Both noise estimates are superposed upon the simulated measurements of the sensor model.

A number of characteristics of the IMU are neglected by the model. Most noticeably these are non-linearity, temperature effects, non-orthogonality of axes and coupling between accelerometer and gyro readings. These were omitted from the model as their contribution to total sensor error was considered minimal in comparison to errors introduced by IMU-optical misalignment, filtering of trajectory data and the Euler decomposition of rotation.

9.2.4 Simulation Validation

Validation of the IMU model was achieved by the comparison of marker derived IMU data with actual measurements from the instrumented cluster. Initially, six 10-second acquisitions were performed using both measurement systems, consisting of multiple yaw-pitch-roll type motions using the tripod setup. After labelling and synchronisation, the optical data was processed via the sensor model. Figures 9.5 and 9.6 show a comparison of sensor and model output for a single test sequence. Dotted lines represent measured IMU data and solid lines show simulation results. In the case of the Accelerometer data a thicker orange line indicates the estimated contribution of gravity to the model:

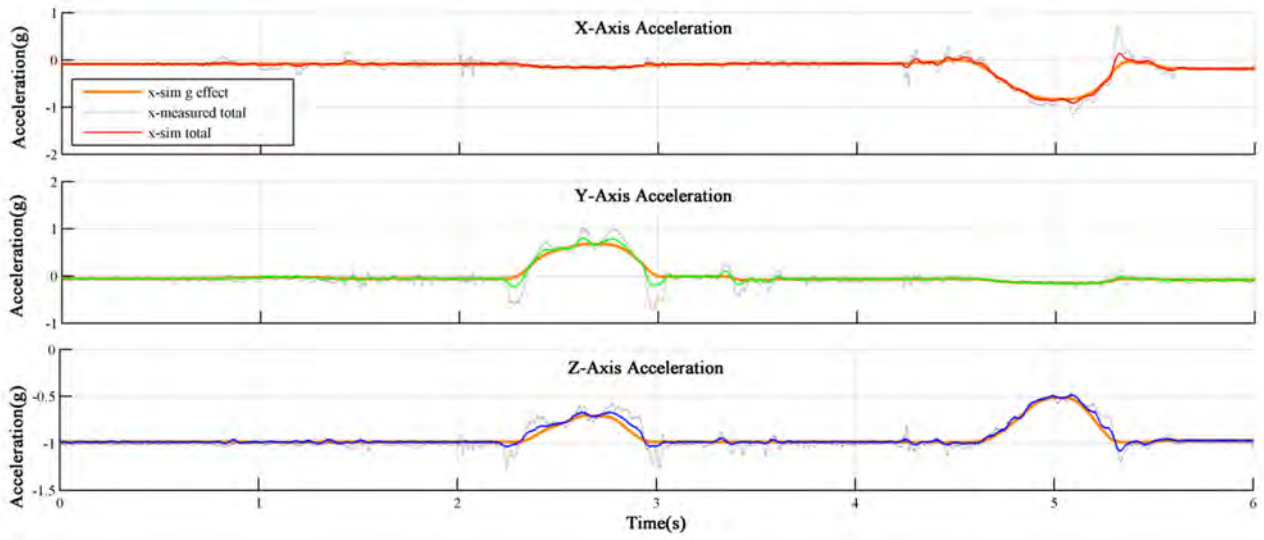


Figure 9.5 – Comparison of simulated acceleration

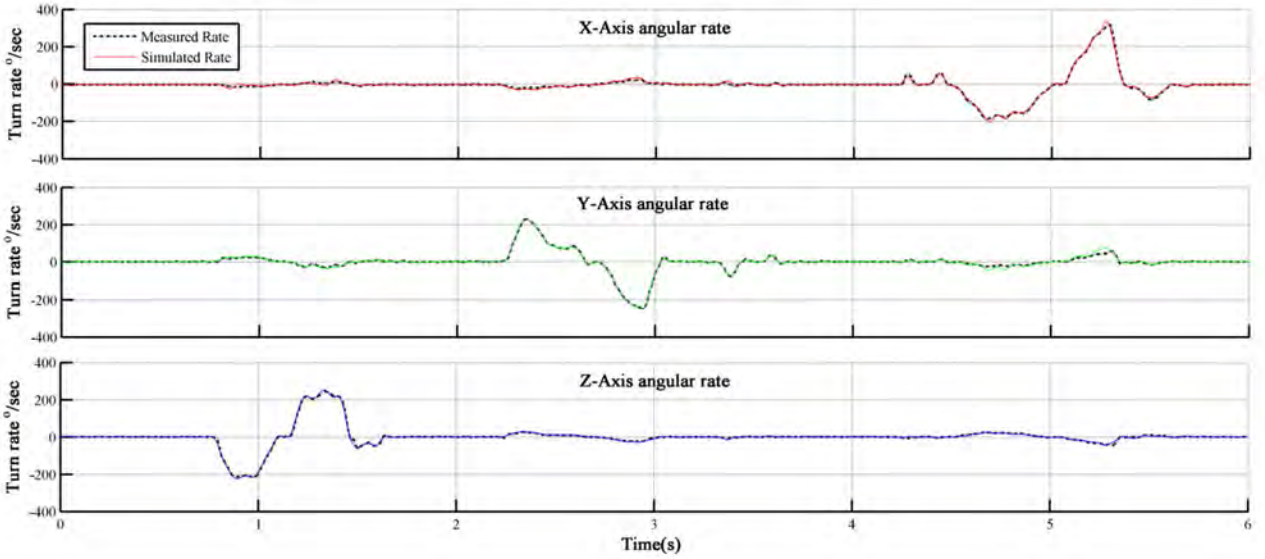


Figure 9.6 – Comparison of simulated rate of turn

It is noted that the model provides good correspondence to measurements, particularly in the case of the gyro where simulated data is virtually indistinguishable from measurements. Deviation in the higher frequency components of the accelerometer signal are due to the extensive low pass filtering of the trajectory data. The overall performance of the sensor model for the tripod test data is summarised in the table below. Mean, maximum and standard deviation are given for the absolute differences between corresponding values for the six sensing axes.

	x-axis	y-axis	z-axis
Mean Error(g)	0.035	0.044	0.028
Error SD(g)	0.046	0.053	0.025
Max Error(g)	0.628	0.528	0.248

Table 9.2 – Performance of accelerometer model

	x-axis	y-axis	z-axis
Mean Error($^o/s$)	2.55	1.94	1.76
Error SD($^o/s$)	3.67	3.21	2.67
Max Error($^o/s$)	33.39	29.54	27.88

Table 9.3 – Performance of gyroscope model

9.3 Simulation of full body inertial data

Extending the single sensor model to the remaining makers of the body is straightforward due to cluster-wise trajectory, reference and orientation data already being provided by the body model (6.1). To account for the alignment error of simulated IMUs, the estimated alignment for the actual instrumented cluster is perturbed by a random angular offset in the range 0.5-1^{o}. Parameters for simulated drift and measurement noise are derived by the methods described in the previous section.

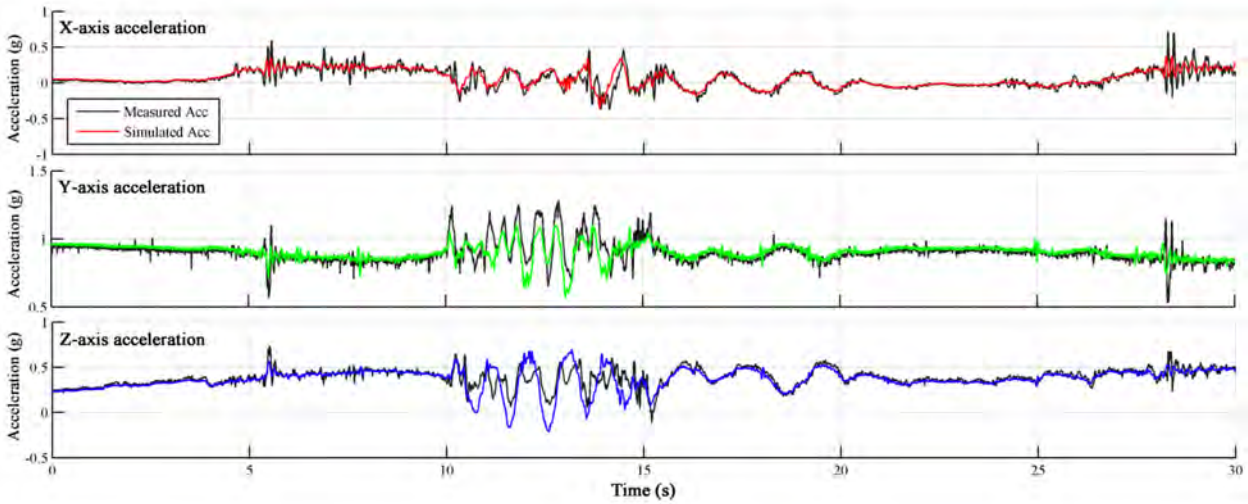


Figure 9.7 – Comparison of simulated/measured lower leg acceleration

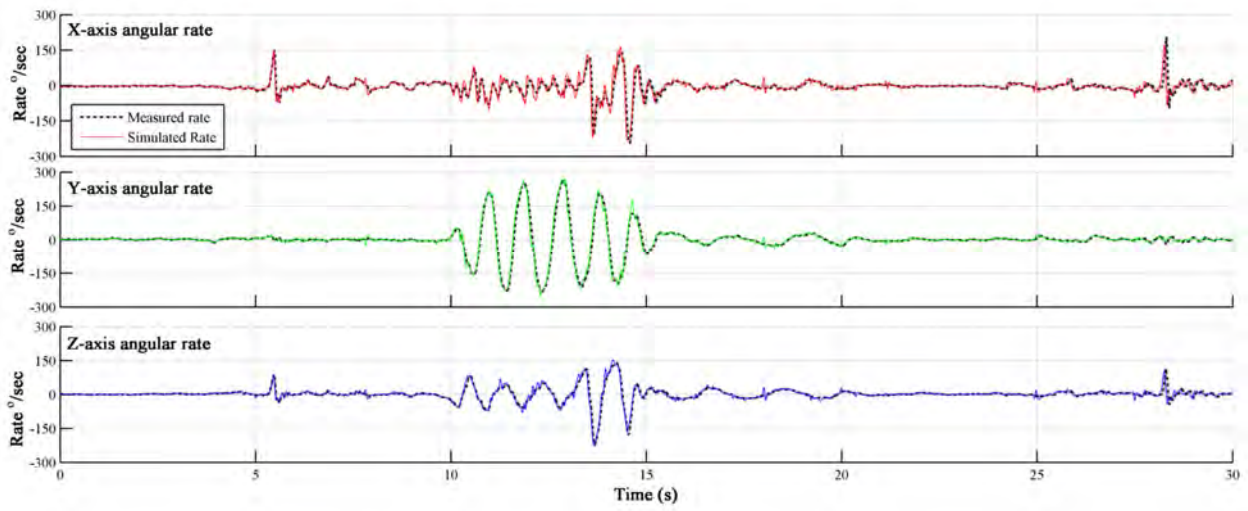


Figure 9.8 – Comparison of simulated/measured lower leg rate of turn

For confirmation of model performance figures 9.7 and 9.8 show the simulated acceleration and angular velocity series for the right lower arm cluster (*r_ll_seg*) during the range of motion sequence of the primary data set (10.1.1). Measured IMU data is included for comparison, signified by the black line while the simulated data is coloured. It should be noted that there is a considerable disparity between acceleration and simulation in the *z*-axis. However, as only gravitational acceleration is utilised in the final reconstruction algorithms, this was not considered to present a serious problem. In the case of the gyroscope model, acceptable performance is noted throughout this sequence and the remainder of the primary data set.

For clusters where real acceleration data was unavailable, verification of the alignment of simulated IMU data was performed by visual inspection. This process was aided by the superposition of graphical devices upon the marker coordinate data representing the direction and magnitude of IMU parameters. Acceleration due to gravity and translation was represented as a series of arrows, indicating direction and magnitude originating from the each cluster's centroid. Rate of turn information was filtered and converted to axis angle representation (Appendix 9), before being plotted via a disc shaped device. The central axis of this disc indicates the axis of rotation and its sector the angular rate. Typical output for a single frame of the primary data set is shown in figure 9.9.

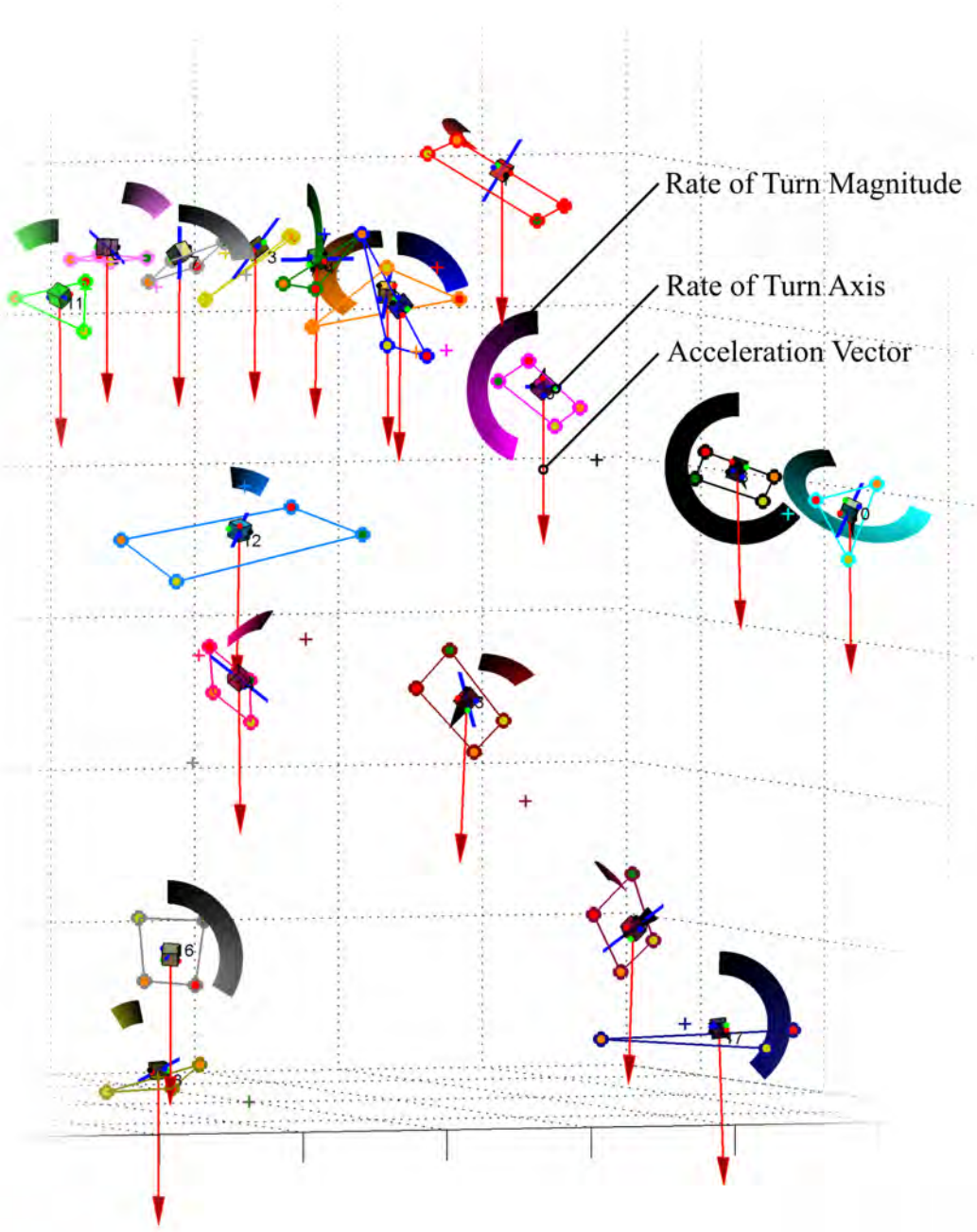


Figure 9.9 – Plot of full body simulated inertial data

9.4 Summary

This chapter has described a method, which allows simulation of accelerometric and gyroscopic measurements using trajectory data acquired by optical motion capture. The method utilised both translational and rotational measurements of a marker cluster to model the effects of translational acceleration, angular moment, and gravity upon the accelerometer. The gyroscope model used the Euler angle decomposition of the relative change in orientation between frames to provide the rate of turn information for each axis of the inertial sensor.

In addition to the sensor models, a novel method of calibrating sensor alignment was described, which allowed the rotational transform between the inertial sensor and optical marker cluster to be estimated.

The validation of the sensor models was conducted using real inertial measurement data acquired from both a test apparatus and human movements. The gyroscope model gave excellent performance, producing results virtually indistinguishable from measurements. In the case of the accelerometer model, differences in dynamic acceleration were observed with respect to the measurements. However, as gravitational acceleration was reliably predicted, the model was considered acceptable for the purposes of this project.

Finally, the sensor models were applied to the full body marker set, allowing for inertial measurements to be simulated over the whole body. A novel visualisation method was implemented for verification of the simulation’s results.

10. PRIMARY EXPERIMENT

To facilitate the collection of inertial measurements alongside optical data it was necessary to conduct a second experiment. This chapter discusses the changes made to the protocol of the pilot experiment to implement this. The following sections detail the inclusion of the inertial measurement hardware, changes to the full body marker set and the derivation of an improved camera layout. The chapter closes with a description of the execution of the actual experiment and initial data processing.

10.1 Modification of experimental protocol

10.1.1 Adjustment of protocol for inclusion of IMU

The protocol used in the primary experiment was identical to that of the pilot, with four sequences in total comprising one general range of motion exercise and three Tai Chi patterns. These sequences were repeated four times each with the instrumented cluster (a cluster with an IMU attached) located upon the subject's right lower leg and right upper arm. Following each change in IMU position upon the body, a short calibration trial was performed to establish alignment. Only two locations were chosen, as the purpose of the inertial measurements was to validate the results of the sensor model. Therefore, if comparable results were obtained in a single location for real and simulated data it would be assumed that similar performance could be replicated over the whole body. The lower leg and arm segments were found to exhibit a greater range of motion than more proximal segments and hence gave more varied test data. However, the hands and feet were avoided to ensure that the maximum rate of turn lay within the measurement range of the IMU. The right hand side of the body was chosen to simplify routing of the IMU tether cable to the acquisition hardware, allowing for the participant to face the experimenter. Using the sequence names established in (5.2.2) for

each Tai Chi pattern and the two IMU locations, the trial structure of the primary experiment is given in the table below:

	Block 1	Block 2	Block 3	Block 4
	IMU Right Lower Leg	IMU Right Lower Leg	IMU Right Lower Arm	IMU Right Lower Arm
Trial 0	IMU Alignment		IMU Alignment	
Trial 1	Range of Mot.	Range of Mot.	Range of Mot.	Range of Mot.
Trial 2	Single Whip	Single Whip	Single Whip	Single Whip
Trial 3	Toe kicks	Toe kicks	Toe kicks	Toe kicks
Trial 4	Lotus Kick	Lotus Kick	Lotus Kick	Lotus Kick

Table 10.1– Trial structure of primary experiment

To synchronise inertial and optical measurement systems the TTL frame signal generated by the camera system, was connected to the trigger input of the IMU controller. This allowed for inertial data acquisition to be triggered upon the first frame of the optical acquisition, ensuring temporal correspondence of measurements.

Figure 10.1– Participant with IMU

The tether cable from the IMU presented practical problems due to the additional weight it imposed upon the marker cluster as well as the tripping hazard it created during the experiment. To reduce the influence of cable weight, Velcro straps were applied along the limb, and a short

loop of cable over each joint prevented cable tension being transferred to the cluster. A video frame detailing this attachment is shown in figure 10.1.

10.1.2 Adjustments to Marker Set

To address the problems noted in the pilot data it was necessary to adjust the layout of the full body marker set. These changes are discussed below:

Removal of sternum markers:

The sternum markers were frequently occluded in the pilot data by the arms. Additionally, chest movement during breathing caused this marker to exhibit very poor translational stability relative to other markers on the torso. Due to these limitations the sternum markers failed to fulfil their function of providing an redundant measure of torso pose and were removed from the set. Additionally, it was noted that a considerably more stable secondary estimate of torso pose could be obtained through the c7 and inner clavicle markers.

Extension of hand clusters:

The distance between the distal markers upon the hand was extended to improve accuracy of orientation estimation and eliminate marker swapping observed during hand motions. Due to the limited area available on the hand it was necessary to use a pylon mounted marker to extend the markers away from the carpel. In addition to the imporved reliability this affords, the mounting plate to which these pylons are attached improves rigidity of the cluster.

Shoulder cluster:

It was decided that the marker configuration would be adapted to include a lower scapula marker. The addition of this extra marker brought the marker set closer to the standard for biomechanical measurement [80]. This makes the adoption of a more anatomically relevant shoulder model possible should it be required.

Simplification of foot cluster:

As the central foot markers (*l_foot_up* and *r_foot_up*) were prone to frequent occlusions due to collisions with clothing, they were removed from the marker set. However, the reduction to three markers increased the potential for perturbation of pose estimates by toe movements.

Following these changes the marker set has been reduced to 65 markers from the original 68. It is depicted in its entirety in figure 10.2, superimposed on the segments of the kinematic body model.

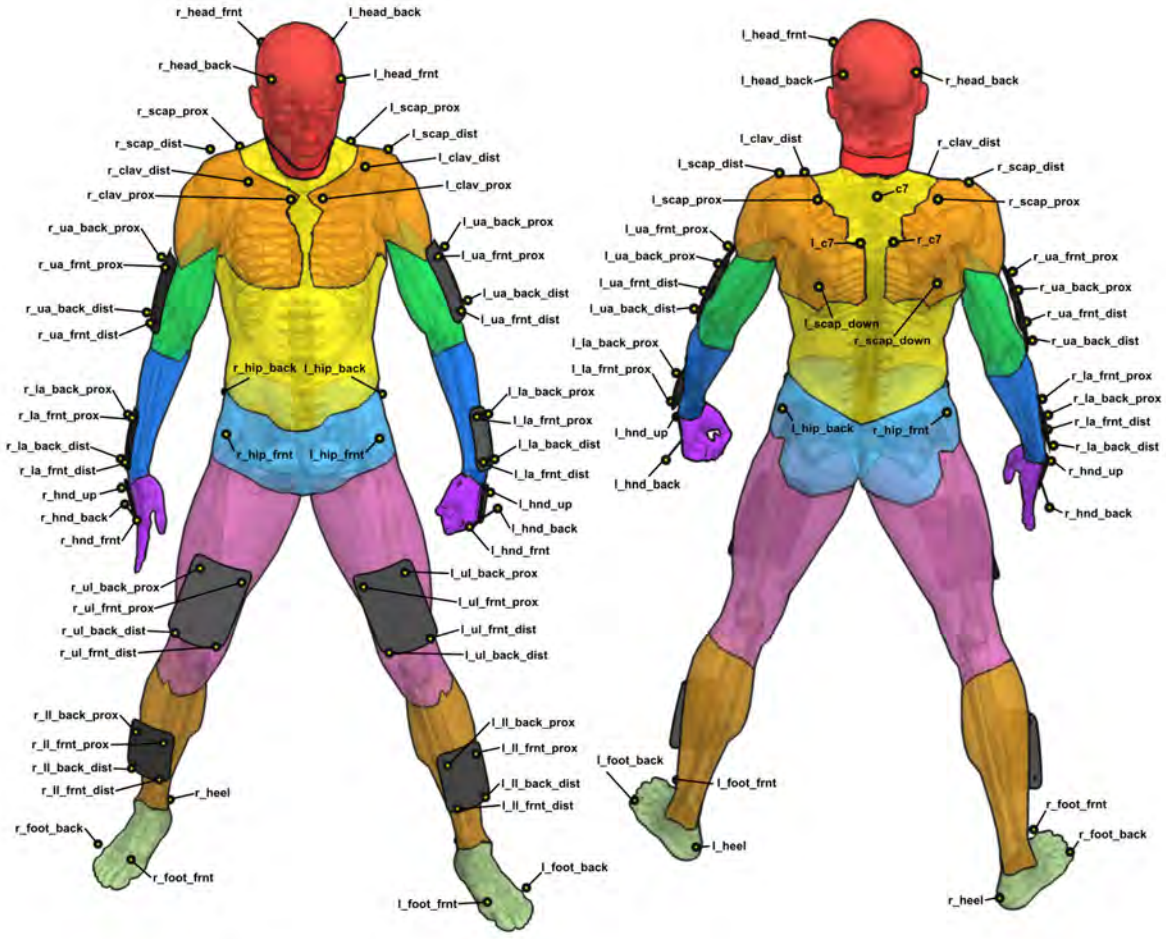


Figure 10.2 – Amended full body marker set for primary experiment

10.2 Selection of camera layout

The poor occlusion performance observed in the pilot experiment was due to difficulty in achieving adequate camera coverage using the available camera mountings. To address this limitation an optimised camera configuration was devised and implemented in the laboratory.

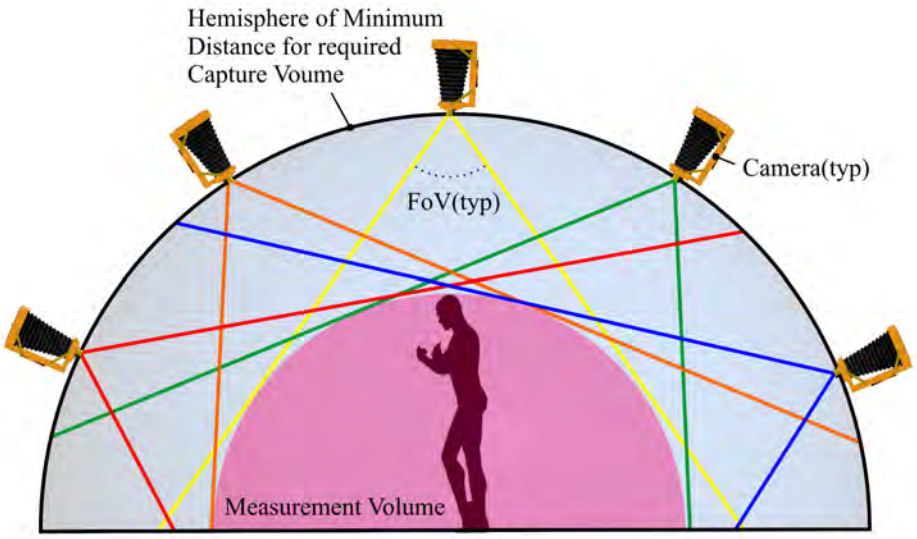


Figure 10.3– Measurement volume defined by intersection of multiple fields of view

10.2.1 The effect of camera placement upon occlusion

Occlusion resistance of a camera system can be optimised by the adjustment of two parameters. These are the number of cameras in the system and their positions about the measurement volume. In an unconstrained measurement scenario, where a subject is free to assume any posture, occlusion can be considered to be equally likely to occur regardless of the position of an observing camera. Therefore, the camera configuration, which covers the greatest solid angle with minimum overlap around the measurement subject will give optimal occlusion resistance [14]. The configuration, which satisfies these requirements is the even distribution of cameras about a hemisphere of fixed radius. This radius is defined as the

minimum distance at which the subject can be framed within each camera's FoV (Figure 10.3).

An approximation to the solution of equidistant points upon the surface of a hemisphere was achieved through division of a hemisphere, both horizontally and vertically into sectors of equal angle. To simplify the generation of a camera configuration for the primary experiment, placements were restricted to two horizontal circles (or tiers) separated by a fixed vertical angle of 45^{o}. Due to the possibility of interference between camera strobes, it was required that each camera did not have a direct view of any other. This was achieved by raising the tiers considerably above the focal point of the cameras (measurement origin). The height of the lower tier was chosen as $2.2m$, which was found to be the minimum height giving adequate clearance of the FoV above a subject of average height ($1.8m$+$0.2m$). The spherical radius was chosen to be 4m in line with the maximum effective camera range and space constraints of the laboratory. Figure 10.4 shows plots of several configurations for varying numbers of cameras designed around these requirements.

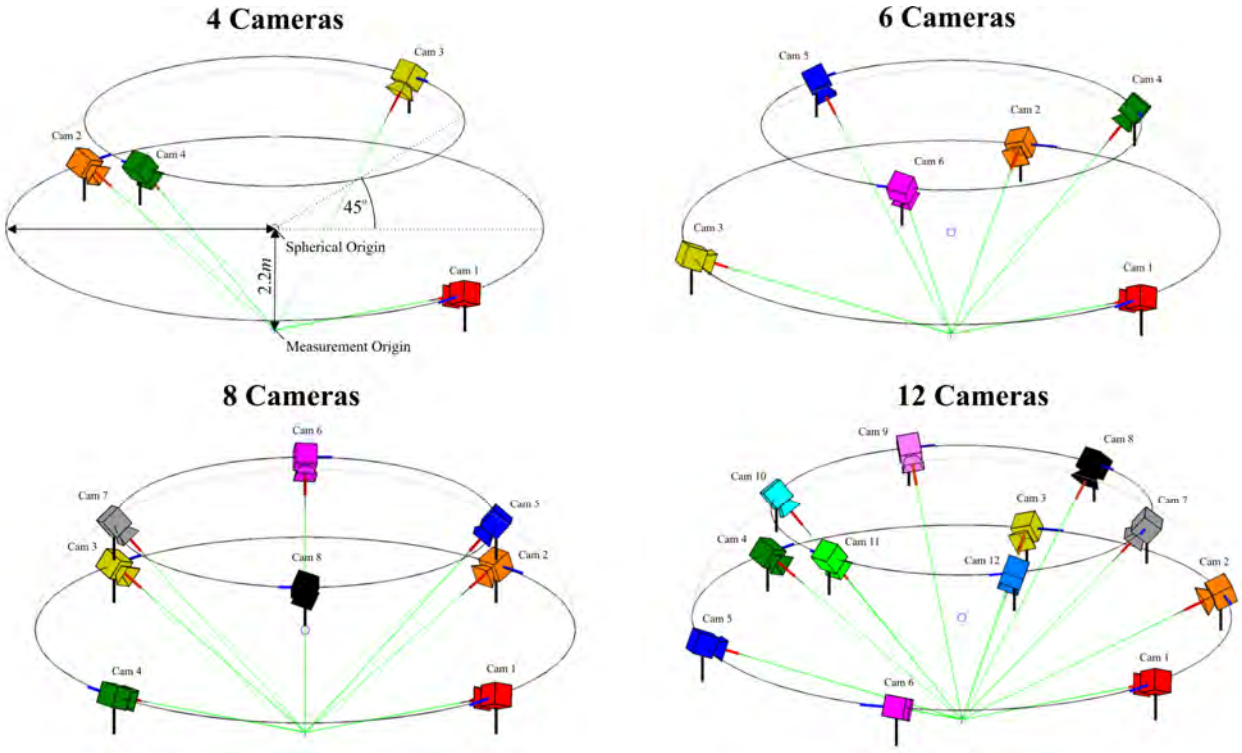


Figure 10.4 – Four Evenly Spaced Camera Configurations

10.2.2 Construction of camera support frame

Although a hemispherical camera configuration was originally planned, this was simplified to a rectangular configuration of similar dimensions to minimise complexity of fabrication and maximise stability. This was implemented as an aluminium structure retrofitted to the walls and ceiling of the laboratory. The structure of this support frame was verified via a three dimensional model of the laboratory created from floor plans and direct measurements. The optimal camera locations given by the 12 camera hemispherical configuration was superimposed upon this model as a guide for the positioning of frame members. A layout was decided, which fulfilled these requirements for location while being both structural stable and compatible with the exiting fittings about the lab (lighting, patient hoist etc.). The design was realised in type 7 (~43*mm* OD) aluminium scaffolding tube with cast steel fittings at joints and abutments with the walls of the lab. To ensure safety and freedom from vibration each

section of the frame was doubly braced. Attachment of the frame to the concrete super-structure of the building was achieved using M6 masonry anchor bolts. The final frame design is shown in figure 10.5 with insets detailing an initial test beam and the completed installation.

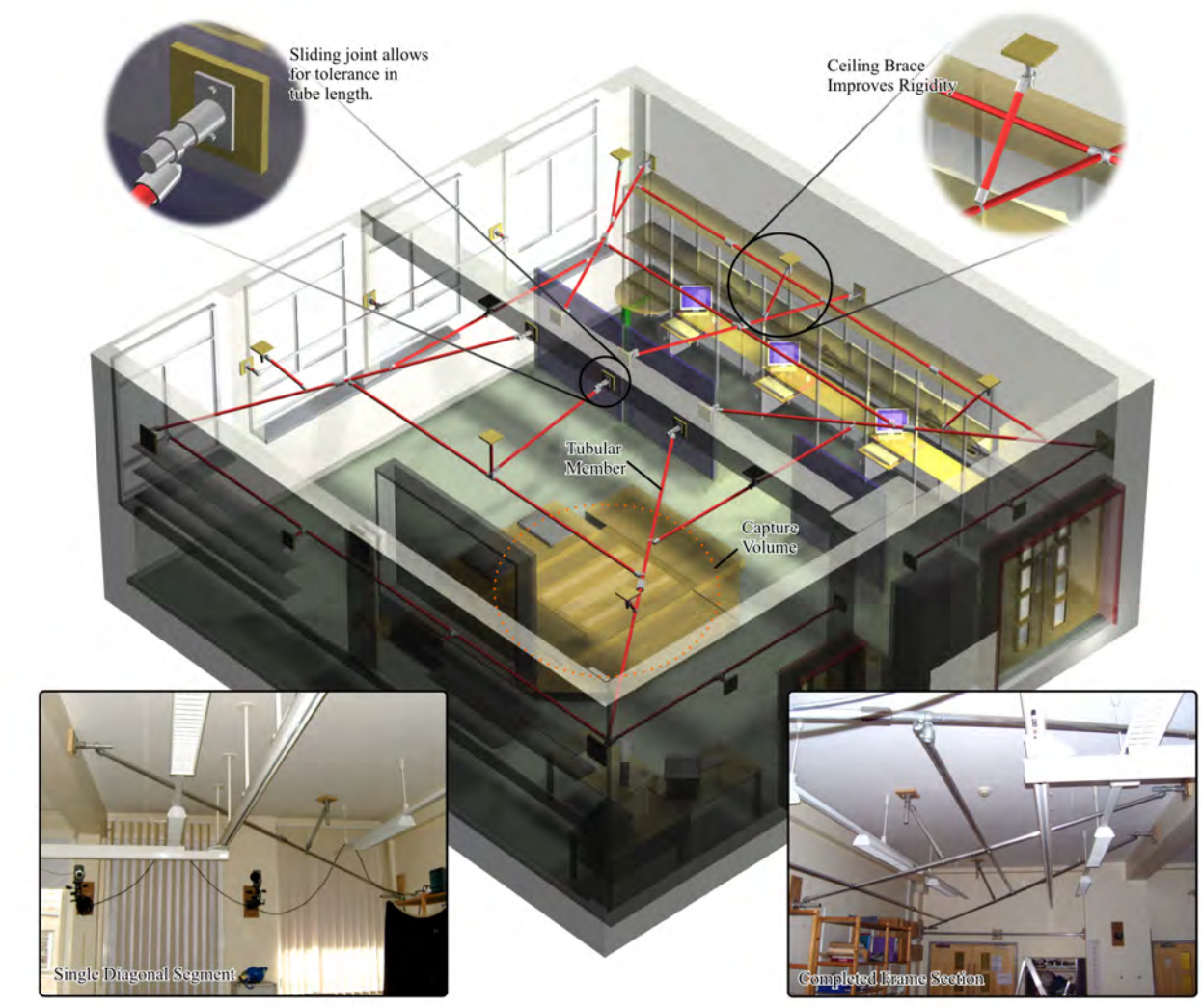


Figure 10.5 – Detail of camera support frame

10.2.3 Camera placement

Several of the planned cameras locations would have been obscured by laboratory fixtures and were adjusted to accommodate for this. Camera fields of view were centred upon the volume origin and pitched vertically where necessary, to ensure that the full height of the subject was visible. Attachment to the frame was achieved via an arrangement of photographic clamps (Manfrotto Super Clamp) and 3-way tripod heads (Manfrotto 804RC2). The resultant camera configuration derived from the camera calibration data is shown in figure 10.6. The configuration was tested to verify coverage over the desired capture volume with the full body marker set prior to the primary experiment.

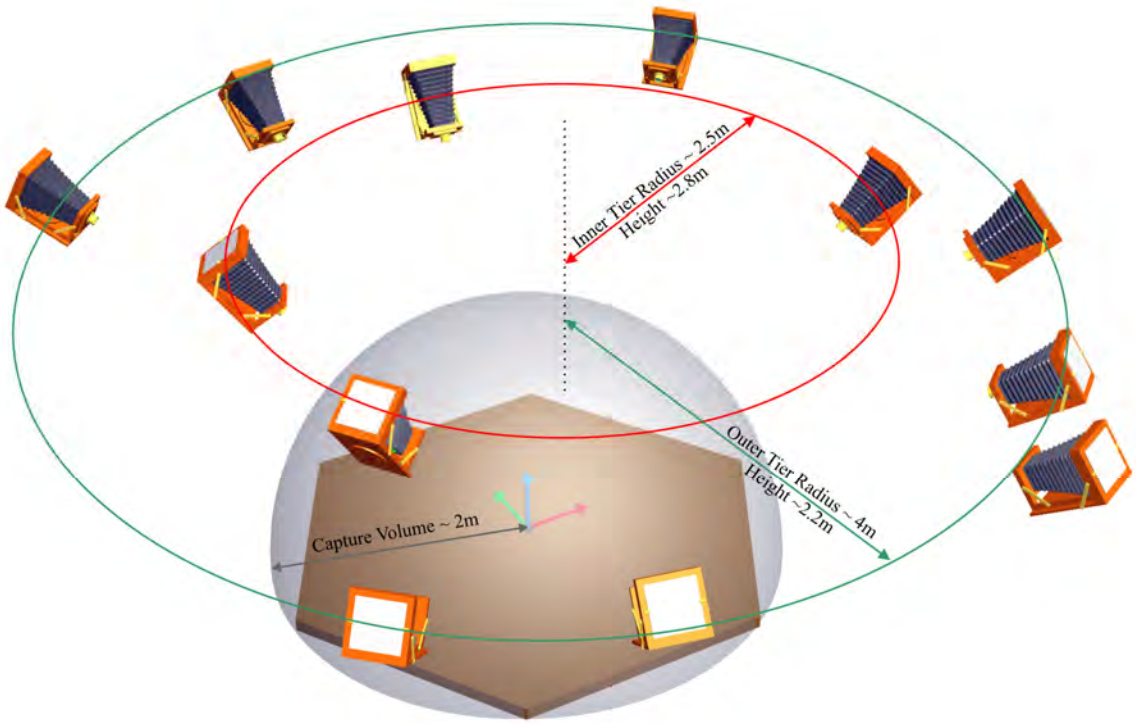


Figure 10.6 – Camera Configuration for Experiment

10.3 Running of Experiment & Data Processing

10.3.1 Subject setup

Following arrival of the participant, the purpose of the experiment was explained and informed consent was obtained. Before markers were applied, the motion sequences were verified against the pilot data and a short practice session performed. The torso and arm markers were applied to bare skin to eliminate clothing movement artefacts. In the case of the legs, elasticated strapping was applied to reduce artefact movement and occlusion due to clothing. The clusters for the hands were bound with Micropore tape to ensure good correspondence with movement of the carpels. Finally, the instrumented cluster was attached as described in (10.1.1). Camera calibration was conducted immediately preceding the subject measurement using the standard calibration fixtures. A spirit level was used to ensure good correspondence between the *z*-axis of the measurement system and gravity, as required by (9.1).

10.3.2 Data acquisition and initial processing

During the measurement session, optical data was acquired at a frame rate of 200*Hz* and inertial data at the hardware limit of ~546*Hz*. The increase in optical frame rate from that used in the pilot experiment (120*Hz*) allowed for improved estimation of acceleration and rate of turn by the inertial sensor model. Capture duration was set to 30 seconds for the range of motion trials and 60 seconds for the TCC movements. After acquisition of each sequence, a test reconstruction of the raw data was performed to verify its integrity. Unfortunately, only two complete blocks of the intended four were run due to equipment problems and the participant's limited availability. As a result, complete inertial data was only available for the lower leg. As inertial data was only required to provide validation of the sensor model, a

single instrumented body segment was considered to provide adequate test data for the continuation of the project.

Raw marker data was processed and labelled in a manner identical to (5.3.2). It was noted that over all trials the mean level of marker visibility over all trajectories and frames was approximately 99.4%, representing an improvement over the pilot experiment (96.1%). Missing trajectory segments were filled with their nearest neighbour co-ordinate values to avoid the introduction of artefacts by spline interpolation. The raw data set was imported into Matlab for orientation estimation and body model calibration.

The IMU measurements were zero delay low-pass filtered at 100*Hz* to prevent aliasing before being downsampled to the 200*Hz* acquisition frequency of the motion capture system.

11, RECONSTRUCTION OF MARKER TRAJECTORIES BY INERTIAL MEASUREMENT

In this chapter trajectory reconstruction is performed for the case of a single cluster, whose constituent marker trajectories are known both before and after an occlusion event. This provides the basis for the implementation of the full body trajectory reconstruction discussed in the remainder of this thesis.

This single cluster reconstruction utilises estimates of orientation provided by the integration of inertial measurements. These orientation estimates are used in combination with a model of the cluster's kinematics to permit the extrapolation of marker trajectories during an occlusion event. The method is verified using an apparatus with known kinematics to establish its accuracy and identify its primary sources of error.

11.1 Estimation of cluster orientation from inertial data

11.1.1 Strapdown integration

A body's absolute orientation can be estimated in relative terms, where each instantaneous orientation is described as a local rotation of its pose at an earlier time. A practical example of this, is the estimation of orientation via inertial sensing, where the angular velocity measurements provided by a gyroscope are used to repeatedly rotate a body's frame of reference from one instant to the next. This process is conventionally referred to as 'strapdown integration' [38] (as the gyro is commonly strapped to a physical body being measured) and is the basis for most inertially based orientation estimation techniques. In this thesis the method has been applied to data produced by the instrumented cluster (8.4.2) to

provide an approximation of body segment orientation independent of the optical measurement system.

In the case of the instrumented marker cluster, angular velocity measurements are provided via a rate gyroscope, as the average rate of turn over the sample period for the *x*, *y* and *z* axes of the device. To estimate the absolute orientation from these measurements, it is necessary to obtain the rotation matrix corresponding to the change in orientation given by this angular velocity over the sample period.

A method by which this differential rotation can be obtained is given in Bortz [38], where the matrix form of the cross product represents the difference between the rotation matrix at two time points ($\mathbf{R}_{A_t}$ and $\mathbf{R}_{A_{t-1}}$) for the angular velocity vector $\boldsymbol{\omega}_t$. The angular velocity vector $\boldsymbol{\omega}_t$ is composed of the individual scalar angular velocities ω_{x_t}, ω_{y_t} and ω_{z_t}, which represent the rates of turn about each cardinal axis of the gyroscope. The difference in the rotation matrix between time *t-1* and *t* is written as:

$$\frac{d\mathbf{R}_{A_t}}{dt} = \begin{bmatrix} 0 & -\omega_{z_t} & \omega_{y_t} \\ \omega_{z_t} & 0 & -\omega_{x_t} \\ -\omega_{y_t} & \omega_{x_t} & 0 \end{bmatrix} \mathbf{R}_{A_{t-1}} = \left[\boldsymbol{\omega}_t\right]_{\times} \mathbf{R}_{A_{t-1}}$$

The $\times$ notation signifies the matrix form of the cross product.

Applying this differential rotation recursively to estimate the orientation from one time point to the next gives an expression for a single step of the strapdown integration. Where $\mathbf{R}_{A_{t=0}}$ is a measure of initial orientation provided by an external reference:

$$\mathbf{R}_{A_t} = \mathbf{R}_{A_{t-1}} + \left[\boldsymbol{\omega}_t\right]_{\times} \mathbf{R}_{A_{t-1}}$$

Although this presents a numerically valid means by which relative rotation can be expressed, in application, rounding errors over successive iterations will cause the matrix $\mathbf{R}_{A_t}$ to drift from orthonormal form (both orthogonal and unit magnitude) over each successive time step. Although this problem can be avoided by periodically renormalizing the matrix, this introduces further error in the orientation estimation. Therefore, it is preferable to use an alternative method of integration that is intrinsically stable for rotations.

An example of such a method is given in [38, 20], where Rodrigues' rotation formula is adapted for the purpose of strapdown integration. Via this method the angular velocity vector ($\boldsymbol{\omega}_t$) is considered analogous to an axis-angle transformation (Appendix 9), where the direction of $\boldsymbol{\omega}_t$ defines the axis about which rotation occurs and its magnitude gives the angle of rotation. Rodrigues' equation provides a means by which this axis-angle representation can be converted directly to a rotation matrix. As this equation produces matrices that are inherently orthonormal, a more stable strapdown integration can be formulated.

By substituting the normalised angular velocity vector $\boldsymbol{\omega}_t / \left|\boldsymbol{\omega}_t\right|$ and its magnitude $\left|\boldsymbol{\omega}_t\right|$ for the axis and angle terms of Rodrigues, an expression for the Rotation matrix $\mathbf{R}_{\omega_t}$ is obtained:

$$\mathbf{R}_{\omega_t} = \mathbf{I} + (1 - cos(|\boldsymbol{\omega}_t|)) \left[\frac{\boldsymbol{\omega}_t}{|\boldsymbol{\omega}_t|} \right]_\times^2 + sin(|\boldsymbol{\omega}_t|) \left[\frac{\boldsymbol{\omega}_t}{|\boldsymbol{\omega}_t|} \right]_\times$$

Rewriting the expression for strapdown integration as a matrix multiplication, the final expression for the orientation estimate $\mathbf{R}_{B_t}$ is given as:

$$\mathbf{R}_{B_t} = \mathbf{R}_{B_{t-1}} \mathbf{R}_{\omega_t}$$

The similarity between the axis-angle and quaternion representations allows this operation to be implemented using quaternion arithmetic [84]. This eliminates the trigonometric operations required by Rodrigues' and improves computational efficiency.

11.1.2 Influence of bias error on integration

A disadvantage of strapdown integration is that small errors in angular velocity measurements accumulate over successive time steps. With reference to the sensor noise model of (9.2.3), the slowly changing zero bias in the gyroscope signal is the dominant source of these errors. Each successive integration step produces a change in orientation in the direction of this bias. Therefore, the estimate of absolute orientation will drift in this direction. Over time, the estimated orientation will gradually depart from the true orientation with a rate dependent on the magnitude of this bias. An additional influence on integration error is the coupling between gyroscope bias and linear acceleration. However, this effect is considered to be negligible with respect to other errors and is not considered further.

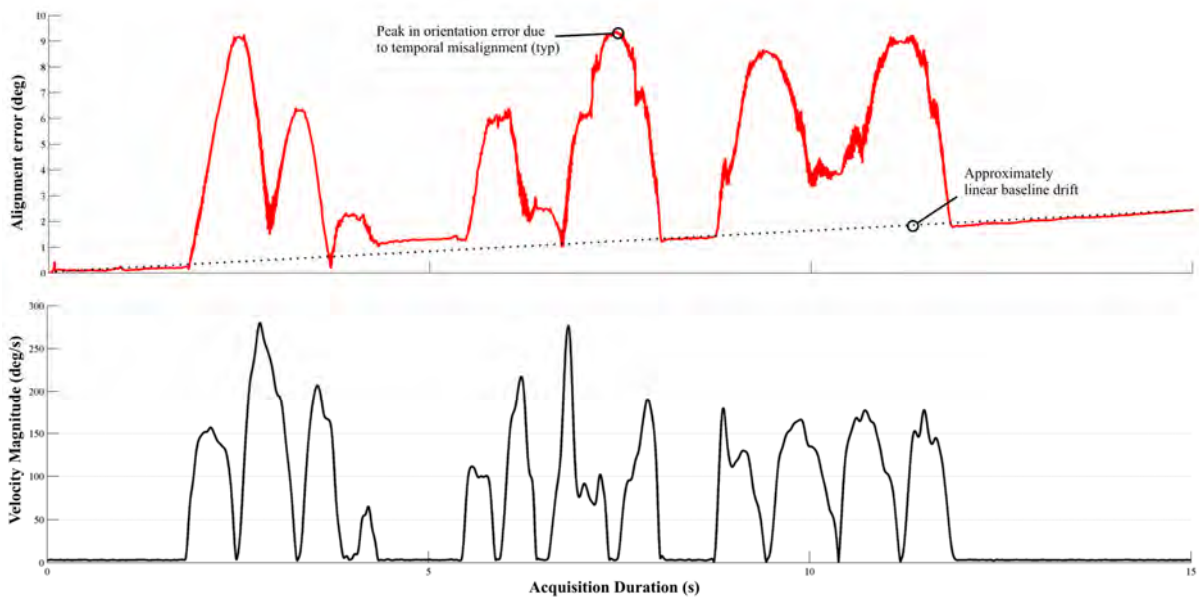


Figure 11.1 – Increase in orientation estimation error during strapdown integration

Figure 11.1 shows the errors in orientation during strapdown integration of the yaw-pitch-roll calibration data obtained using the instrumented cluster (8.4.2). Initial sensor bias was calculated by averaging gyro measurements while stationary. Reference orientation ($\mathbf{R}_{B_0}$) was provided by the positions of optical markers. The upper graph of the figure shows the orientation error between the strapdown estimate and optical orientation measurements. The lower graph shows the magnitude of angular velocity for reference. Sudden increases in error can be observed at peaks of angular acceleration. This is due to a slight temporal misalignment between optical and inertial data, which resulted from a manual start of the two measurement systems (8.4.2). The effect of gyroscope bias is observed as the linear change in the baseline over the course of the measurement. As this error significantly limits the long term usefulness of strapdown integration, a method for compensation of bias error is needed.

11.1.3 Common methods of correction by sensor fusion/filtering

Although the rate of orientation drift is approximately linear, the bias's effect on the orientation estimate is considerably more complicated than this trend suggests. This is predominately due to the 'walking' behaviour of the gyroscope bias shown in figure 11.2. Due to this slow variation, the direction of angular velocity error will vary continuously over time as the bias drifts away from the initial average estimate. As the accuracy of the orientation estimate deteriorates, the angular velocity in the sensor frame will no longer correspond to that of the global frame, leading to an increase in the rate of error accumulation.

A method by which sensor bias can be eliminated is through high pass filtering of the angular velocity measurements. However, such an operation does not discriminate between drift of bias and low frequency components of the actual movement signal. As human movements exhibit great variation in frequency, selection of an appropriate cutoff frequency is problematic.

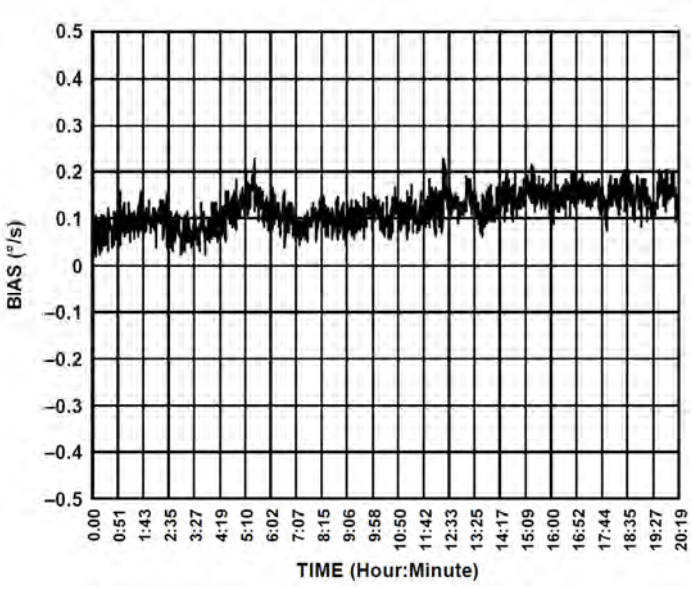


Figure 11.2 – Bias drift over time [36]

A more effective mechanism for compensation of bias is to correct the strapdown integration using an alternative source of orientation measurement. Figure 11.3 shows an example of such a correction scheme, which uses the long term trend of accelerometer measurements to

produce a reference orientation from the acceleration due to gravity [20]. From the normalised instantaneous acceleration vector, an estimate of the global –*z*-axis (down) in the local co-ordinate frame of the IMU is defined. By also estimating the direction of the same axis by strapdown integration, a pair of vectors is obtained. The smallest angle between these vectors infers the error in orientation between the gyroscope orientation estimate and the fixed gravitational axis. By calculating the axis of rotation between these vectors, an axis-angle rotation is obtained relating the strapdown orientation estimate to the gravitational –*z* axis. This rotation can be used as a correcting factor for the effect of sensor drift.

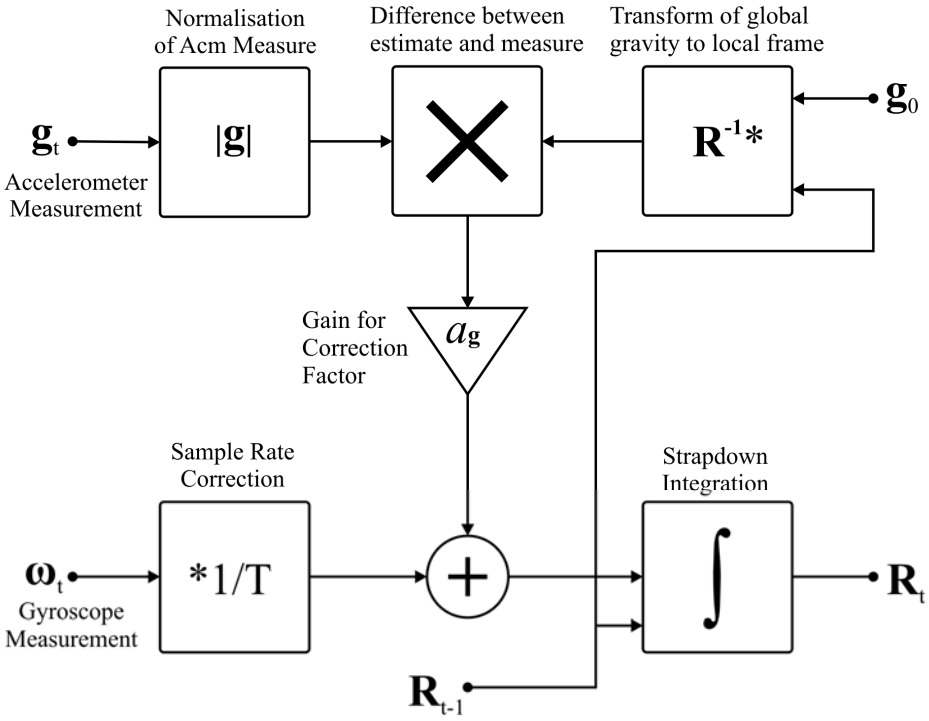


Figure 11.3 – Gravity based orientation correction

If the accelerometer only registers acceleration due to gravity, this scheme will reduce drift in the final estimate of orientation. However, in practice translational acceleration is unavoidable. Some improvement in performance can be achieved through the addition of a proportional gain factor a_g to the rotational correction. This allows the rate at which the

orientation estimate approaches the gravitational estimate to be controlled. By tuning the value of a_g it is possible to minimise orientation drift while simultaneously maximising rejection of artefacts introduced by movement.

Figure 11.4 shows the performance of this method against the uncorrected strapdown orientation estimation ($a_{g=0}$) for the yaw-pitch-roll data. A minimal reduction was seen in the orientation drift, at the cost of a noticeable coupling of linear acceleration (spikes at periods of peak acceleration).

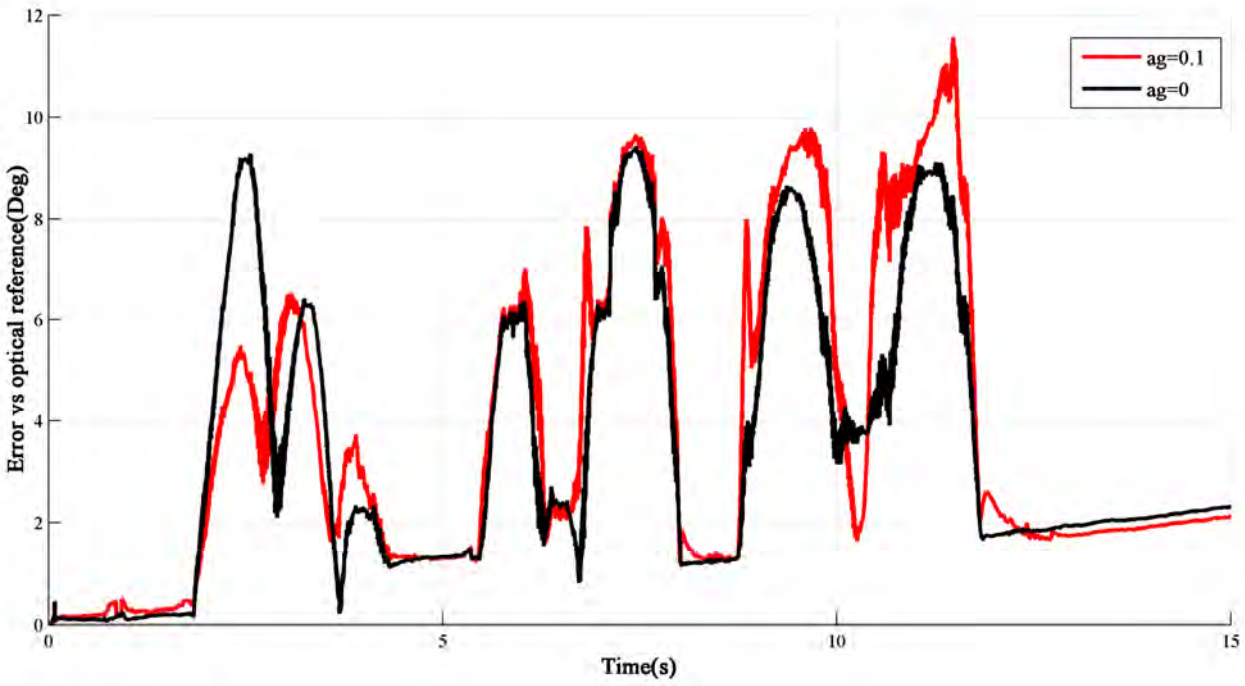


Figure 11.4 – Poor improvement in orientation drift despite gravity compensation

As the gravity vector only constrains orientation about two axes, the error due to sensor bias is free to accumulate around the third axis (parallel with gravity). This results in minimal compensation for drift and hence poor performance of the method.

This failing can be addressed through the use of an additional sensing modality to constrain the free axis of rotation. Examples of such sensors are magnetic compasses [16] and machine vision systems [81]. Typically these methods follow the same principle as the example above, but with additional controls over the correcting signal gain. Example of these schemes are based on Proportional-Integral-Derivative (PID) control [82], modulating gain by angular velocity [83] and the complementary Kalman filter [16, 18]. Among these methods the Kalman filter based technique is considered the most reliable and tends to give better rejection of disturbances. This is due to its ability to account for relationships between gyroscope signals and correcting measurements through analysis of their covariance. However, the limited frequency response of these methods renders them incapable of providing adequate correction for drift over the typically short durations of occlusion events.

11.1.4 Gyroscope bias tracking

As it is only possible to obtain a measurement of gyroscope bias at the start of a motion sequence (when the IMU is stationary), a means to continuously estimate the change in bias during movement was required. This was achieved by the application of the sensor model given in (9.2.2) to estimate angular velocity directly from optical marker measurements during periods of marker visibility. By calculating the difference between this estimate and the gyroscope measurements over each time step of the strapdown integration gyroscope bias error was estimated.

This difference was prone to disturbance by temporal sample misalignment and marker position noise, making it an unreliable indicator of actual bias. To address this problem a variable filter system was devised based on the long term window average of the bias estimate. To reduce the influence of transient changes in marker position, each window's mean average was inversely weighted by its standard deviation using the reciprocal function $1/(1-x)$. This weighting was used to establish a balance between the mean of the new window and the previous estimate of sensor bias. The final average value for bias ($\mathbf{d}_t$) at time *t* was calculated via the mean of a second series (**B**) of equal length to the averaging window to which the new weighed value was appended and the oldest value discarded. This algorithm followed steady low frequency changes in bias but rejected sudden impulses in favour of its current value. Sensitivity to variability was tuned by multiplication of the standard deviation by a factor (s_w). Variation in the rate of bias drift was accommodated by adjustment of window length. A schematic of the operation of the algorithm is shown in figure 11.5.

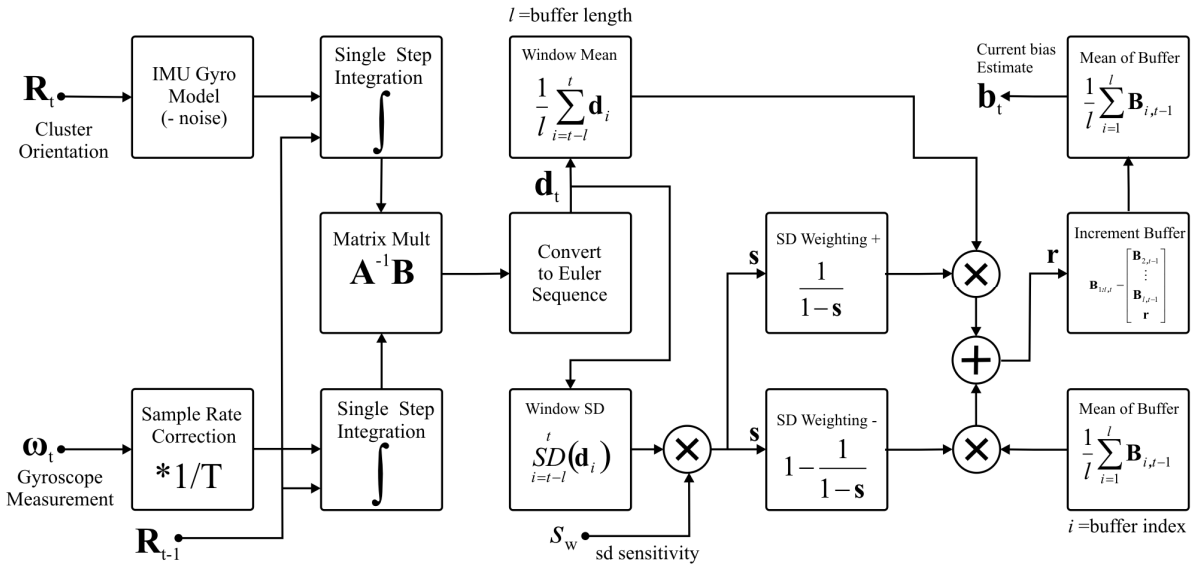


Figure 11.5 – Optical bias tracking scheme

Figure 11.6 illustrates the results of testing this filter using the yaw-pitch-roll dataset. The upper plot shows the raw angular velocity difference $\mathbf{d}_t$, and the lower plot shows the window averaged values and the resultant bias estimate. Best results were found using a window length of 500 samples and s_w weighting of 80.

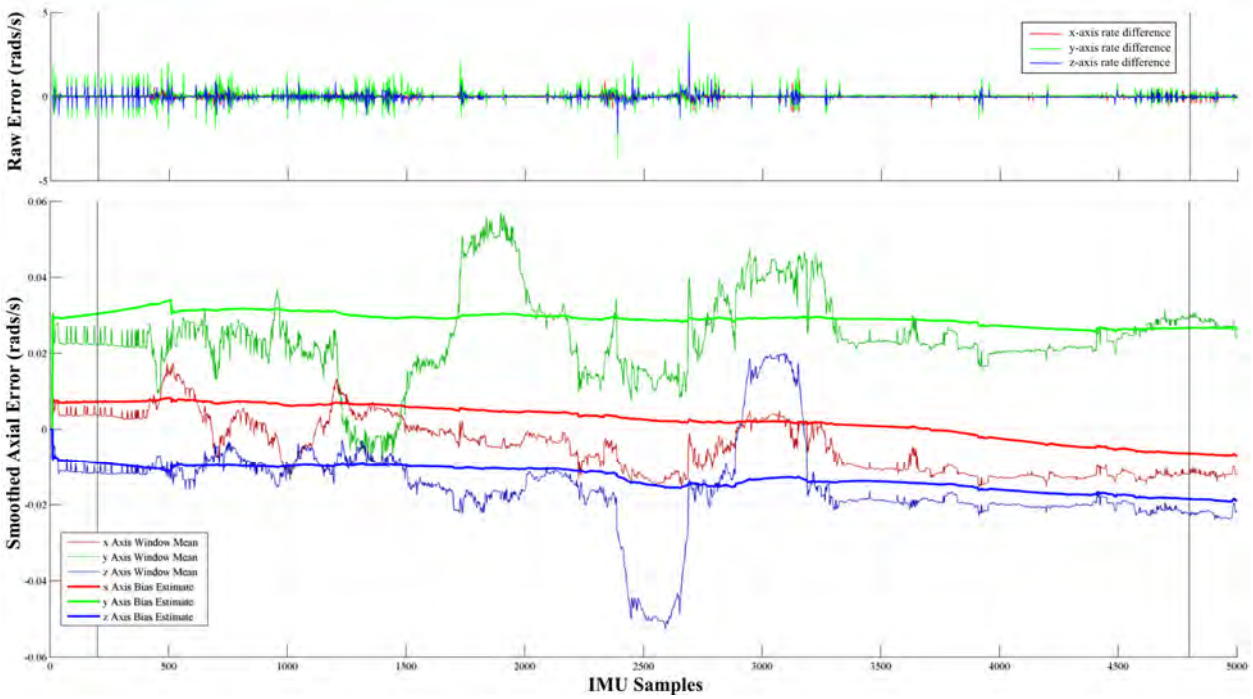


Figure 11.6 – Response of bias tracker

Applying this filter to the trajectory and gyro data allowed the bias of the gyro to be estimated up to the point where an occlusion event occurs. In combination with the optically measured orientation at the point of disappearance, this provided the initial parameters for estimation of orientation across the occlusion gap by strapdown integration.

11.1.5 Approximation of drift across occlusion events

Although the bias tracker reduced the influence of bias error on the orientation estimation, drift still occurred due to the difference between estimated and actual gyroscope bias. By assuming that the gyro bias is constant over the duration of an occlusion event, the drift error

increases in a specific direction of rotation at a constant angular velocity. This allowed the drift error across the gap to be approximated by the linear interpolation of the rotation due to orientation error at the occlusion end (Figure 11.7). This was achieved by converting the orientation error to axis-angle representation (Appendix A) and scaling the angle component by the ratio of the time since the start of the occlusion to its total duration.

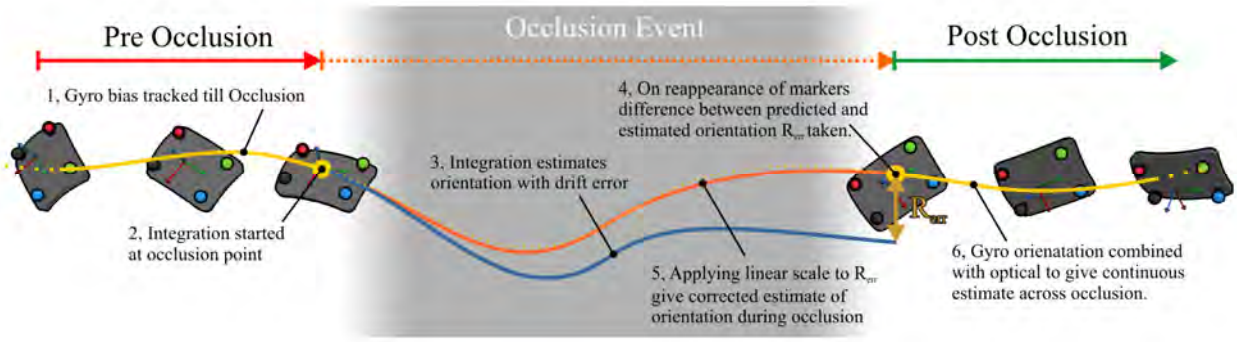


Figure 11.7 – Proposed method of drift error correction

Figure 11.8 shows the results of this method applied to four artificially generated occlusions of the test data obtained in (11.2.1). Orientation is plotted as the Euler rotations about the *x*, *y* and *z* global axes were obtained from the decomposition of the rotation matrix. The durations of the occlusion events are indicted by the shaded areas. The three continuous traces on each graph, coloured red, green and blue represent the ground truth orientation derived from the optical marker data. The two traces visible during the occluded segments represent the inertially derived orientation estimates, coloured magenta for the initial estimate from strapdown integration alone and black for the drift corrected values.

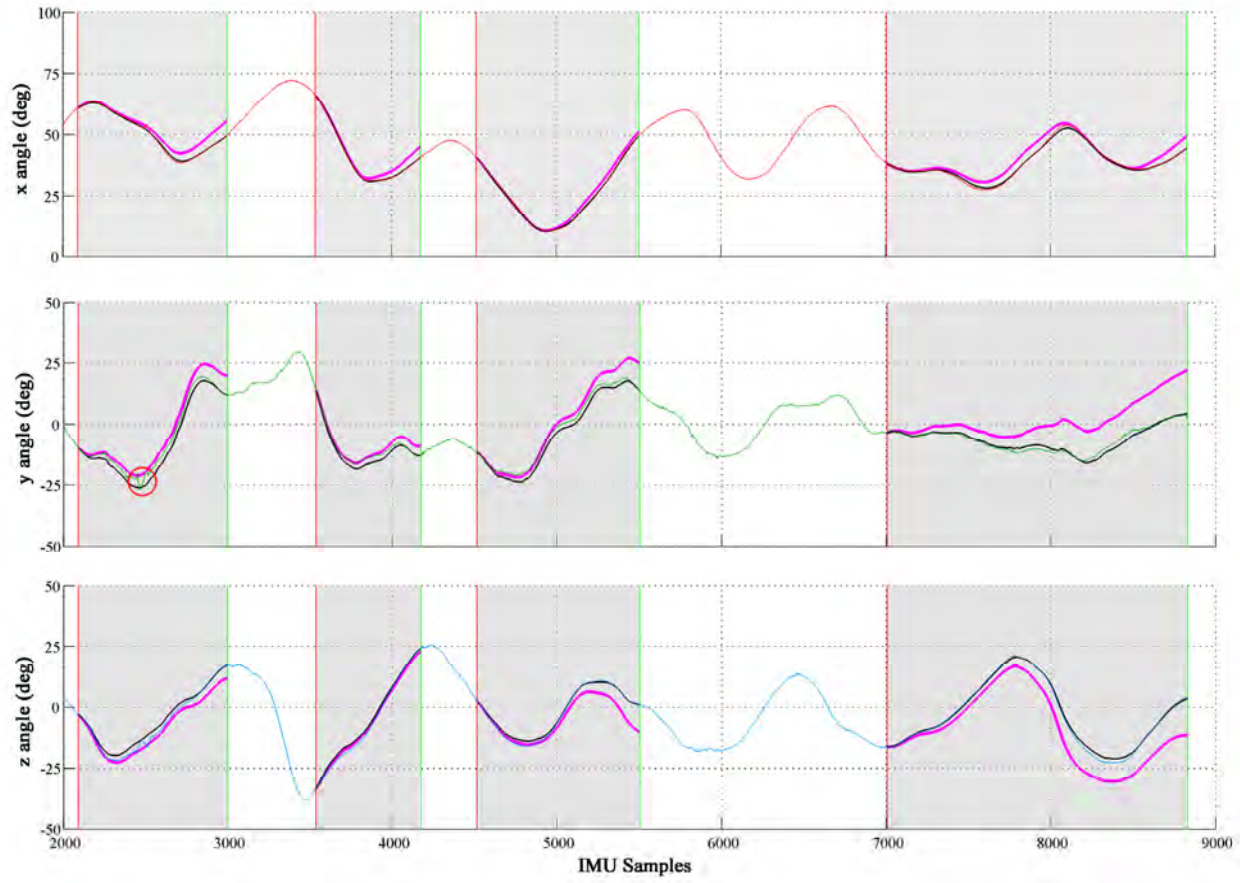


Figure 11.8 - Drift error corrected orientation estimate

The accuracy of orientation estimation was improved considerably by the addition of the linear correction, which reduced the orientation error to a few degrees. This performance was reliable even over longer events, such as the final occlusion of the sequence (~4 seconds). It was noted that some of the higher frequency content in the optical orientation data (see red circle) was absent from the inertial estimates. This was attributed to a partial occlusion of a single marker during measurement. Figure 11.9, shows the magnitude of angular error between the optical and drift corrected inertial estimates of orientation.

The error in orientation typically did not exceed 4^o, with a mean of 1.9^o (excluding points outside occlusion events). This level of error was of the order found in other inertial orientation estimation methods commonly applied to the measurement of human movement [41, 83, 18]. Therefore, this level of error is considered sufficient for the purposes of this project. Furthermore, this value of angular error was in agreement with the limits established for optical measurement artefacts in (4.1.4).

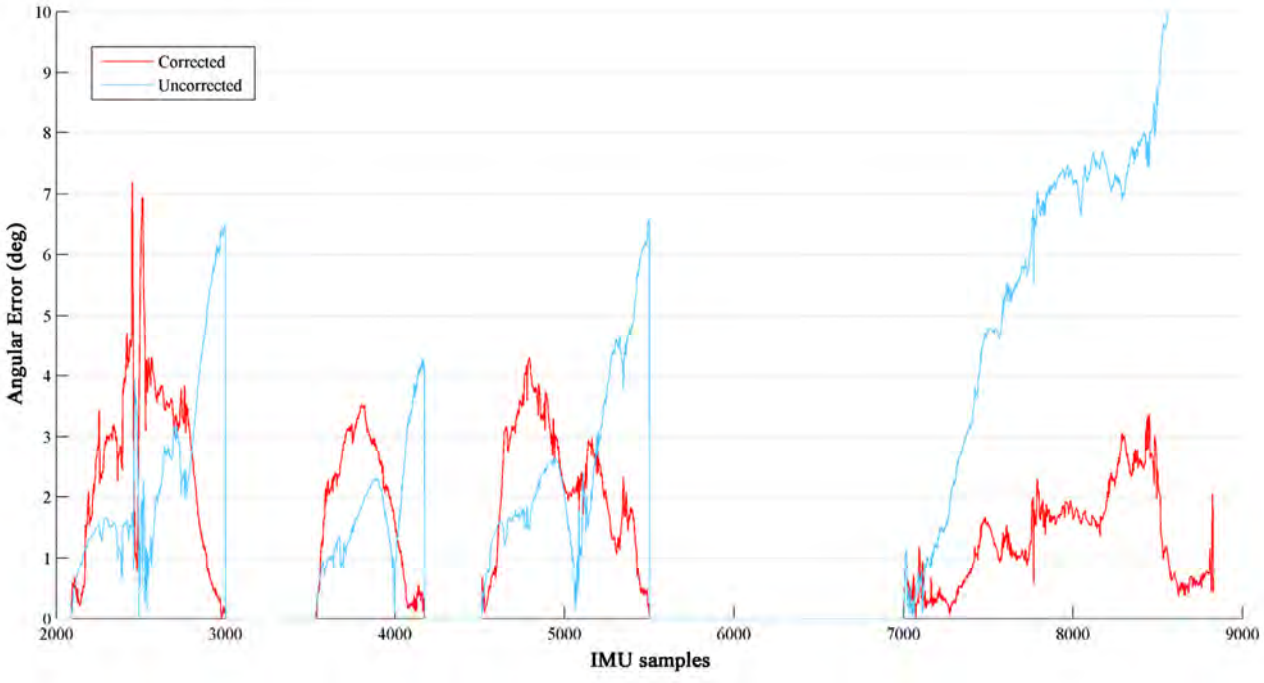


Figure 11.9 - Drift error magnitude for un/corrected data

11.2 Basic reconstruction of marker trajectories for pre-identified markers

11.2.1 Experiment to test trajectory reconstruction

The method of orientation correction described above can be applied to the problem of trajectory reconstruction during occlusion events. Due to the complexity of the primary dataset, it was preferable to obtain a smaller and more controlled range of trajectories for initial development of the reconstruction method. Therefore, a rigid mechanical armature was assembled for this purpose (Figure 11.10). This armature consisted of two 1.2*m* long wooden beams joined by a single axis rotational joint. A marker cluster was attached to each of the beams at a distance of 1*m* from the pivot (one of which was the instrumented cluster). The location of the bearing was indicated by two additional markers placed along its axis. Using the cluster markers the position and orientation of each beam were measured. Four sequences of 30 second duration were acquired in which the armature was moved around an approximately spherical orbit. During this orbit the angle of the hinge was repeatedly varied between approximately 20 and 120 degrees. As in the primary experiment, the frame rate of optical capture was set to 200*Hz* and the IMU sample rate to 546*Hz* with synchronisation achieved via a TTL trigger pulse. Following acquisition, data was processed and labelled, using the convention $\mathbf{a}_1$-$\mathbf{a}_4$ for the passive cluster markers, $\mathbf{b}_1$-$\mathbf{b}_4$ for the instrumented cluster and $\mathbf{p}_1$-$\mathbf{p}_2$ for the centre of rotation markers. The motion data was exported to Matlab, where up-sampling was performed and inertial and optical measurements were temporally aligned. The best quality of the 4 sets in terms of marker visibility was chosen for further processing.

The presence or absence of markers due to occlusion was modelled as a Boolean array of equal dimension to the trajectory dataset (markers x frames). Using this method four blocks of occlusion were defined with random position and length ranging between 200 and 2000

frames. Finally, pose estimation for the cluster trajectories was performed using the quaternion technique described in (6.1.4).

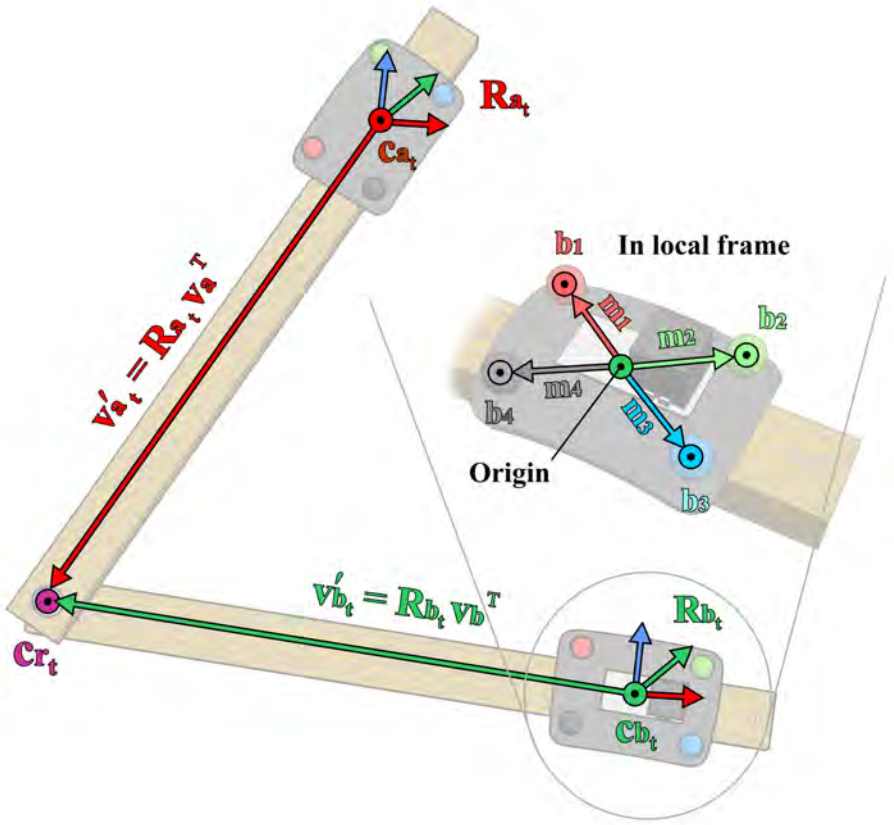


Figure 11.10 – Armature for experimental testing

11.2.2 Simple extrapolation of marker trajectory from IMU and CoR.

For orientation estimates to be of use in inferring the position of markers in space, it was first necessary to quantify the relationship between the two beams of the armature in terms of their mechanical constraints. In the manner of the arm model given in (6.4), the two segments of the armature were defined by two 'segment vectors', $\mathbf{v_a}$ and $\mathbf{v_b}$. These described the relative position of the centre of rotation with reference to each cluster centroid $\mathbf{c}_{a_t}$ and $\mathbf{c}_{b_t}$ (where t equals sample index). The corresponding orientations of each segment with reference to the global frame were described by the matrices $\mathbf{R}_{a_t}$ and $\mathbf{R}_{b_t}$ which yielded the transformed

segment vectors $\mathbf{v'}_{a_t}$ and $\mathbf{v'}_{b_t}$ when applied to $\mathbf{v_a}$ and $\mathbf{v_b}$. These transformed vectors described the translational relationship between the centroids of the two clusters via the centre of rotation ($\mathbf{c}_{r_t}$). These parameters and their relation to the physical structure of the armature are depicted in figure 11.10.

Approximate values for the two rotations $\mathbf{R}_{a_t}$ and $\mathbf{R}_{b_t}$, and centroids $\mathbf{c}_{a_t}$ and $\mathbf{c}_{b_t}$ were provided by the pose estimates from the processed motion data. Additionally, by averaging the positions of the two axial markers $\mathbf{p}_1$ and $\mathbf{p}_2$ for each frame it was possible to obtain the common centre of rotation. To define the local segment vectors $\mathbf{v_a}$ and $\mathbf{v_b}$, it was necessary to transform each vector between the centroids and centre of rotation into its corresponding segment's local co-ordinate system. In an ideal system, this vector should have been equal at every frame of the trajectory. However, measurement noise and deformation of the armature led to the variation of these vectors over time. To address this, values of $\mathbf{v_a}$ and $\mathbf{v_b}$ were approximated as the average of the locally rotated vectors for the duration of the dataset. This method is identical to that discussed in (6.4.1) for estimation of the pose of the arm, and can be written in terms of the centroid and cluster orientation:

$$\mathbf{v_a} = \frac{1}{N}\sum_{t=1}^{N}\left(\mathbf{R}_{a_t}^{-1}(\mathbf{c}_{r_t} - \mathbf{c}_{a_t})^T\right) \quad \text{and} \quad \mathbf{v_b} = \frac{1}{N}\sum_{t=1}^{N}\left(\mathbf{R}_{b_t}^{-1}(\mathbf{c}_{r_t} - \mathbf{c}_{b_t})^T\right)$$

Where N is the total number of frames to be averaged.

To estimate the position of the instrumented cluster's centroid ($\tilde{\mathbf{c}}_{\mathbf{b}_t}$) during an occlusion event the centre of rotation at $\mathbf{c}_{\mathbf{r}_t}$ was traversed by the addition of the transformed segment vectors to the centroid $\mathbf{c}_{\mathbf{a}_t}$:

$$\tilde{\mathbf{c}}_{\mathbf{b}_t} = \mathbf{c}_{\mathbf{a}_t} + \dot{\mathbf{v}}_{\mathbf{a}_t} + \dot{\mathbf{v}}_{\mathbf{b}_t}$$

Substituting rotational transforms this becomes:

$$\tilde{\mathbf{c}}_{\mathbf{b}_t} = \mathbf{c}_{\mathbf{a}_t} + \mathbf{R}_{\mathbf{a}_t}\mathbf{v}_{\mathbf{a}_t}^{T} + \mathbf{R}_{\mathbf{b}_t}\mathbf{v}_{\mathbf{b}_t}^{T}$$

To complete the cluster reconstruction, the coordinates of its constituent markers were reconstructed around the estimated centroid. This was achieved by the rotation of the segment marker vectors $\mathbf{m}_1$-$\mathbf{m}_4$ (Figure 11.10) by $\mathbf{R}_{\mathbf{b}_t}$ and translation to the estimated centroid $\tilde{\mathbf{c}}_{\mathbf{b}_t}$. In this case, the inverse transform operation was avoided as the required vectors were extracted from the reference cluster used in optical orientation estimation (6.1.3). This gives the estimated position of a single marker in the cluster $\tilde{\mathbf{b}}_{fn_t}$ as:

$$\tilde{\mathbf{b}}_{fn_t} = \tilde{\mathbf{c}}_{\mathbf{b}_t} + \mathbf{R}_{\mathbf{b}_t}\mathbf{m}_n^{T} \qquad \text{where } n\text{=1 to 4 depending on marker}$$

By substituting $\mathbf{R}_{\mathbf{b}_t}$ with the corrected inertial orientation estimate of (11.1.5), the final estimation of marker position was obtained. This estimate only involved historical optical measurements of the instrumented cluster **b** and was therefore independent of marker visibility.

By applying this estimation to simulated occlusion gaps in the armature dataset, the marker coordinates of the instrumented cluster were extrapolated past the start of each occlusion event. In figure 11.11 the results of this extrapolation for a single marker (dashed black line) are shown alongside the actual measured trajectory data (solid coloured line). Occlusion events are signified by the shaded areas. Any difference between estimate and measured position at the start of an occlusion is subtracted from the extrapolated trajectory.

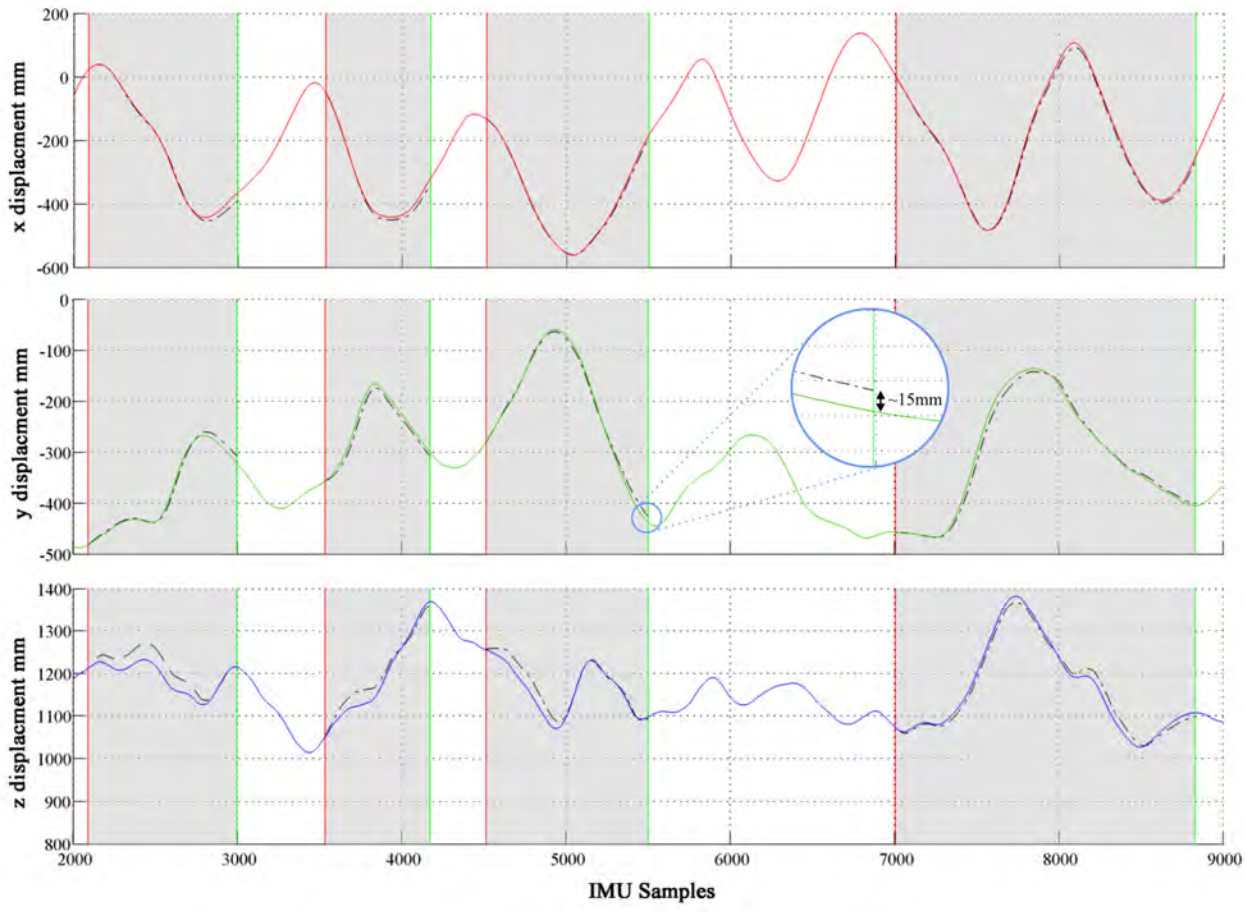


Figure 11.11 – Comparison of extrapolated and measure trajectory data for a single marker

11.2.3 Error correction in trajectory extrapolation

The majority of the reconstructed trajectory data was almost indistinguishable from the measured data. However, discrepancies were seen in several locations, most notably when a marker emerged from occlusion. These errors stemmed from inaccuracies in the inertial orientation estimates and the kinematic model. Through comparison of the level of positional error and angular error, it was inferred that the largest contribution to extrapolated positional error was due to the drift in the orientation estimate. Evidence for this is given in figure 11.12 where angular and positional errors are overlaid. This reveals a correspondence between initial peaks of error for all but the first occlusion event (due to a tracking error).

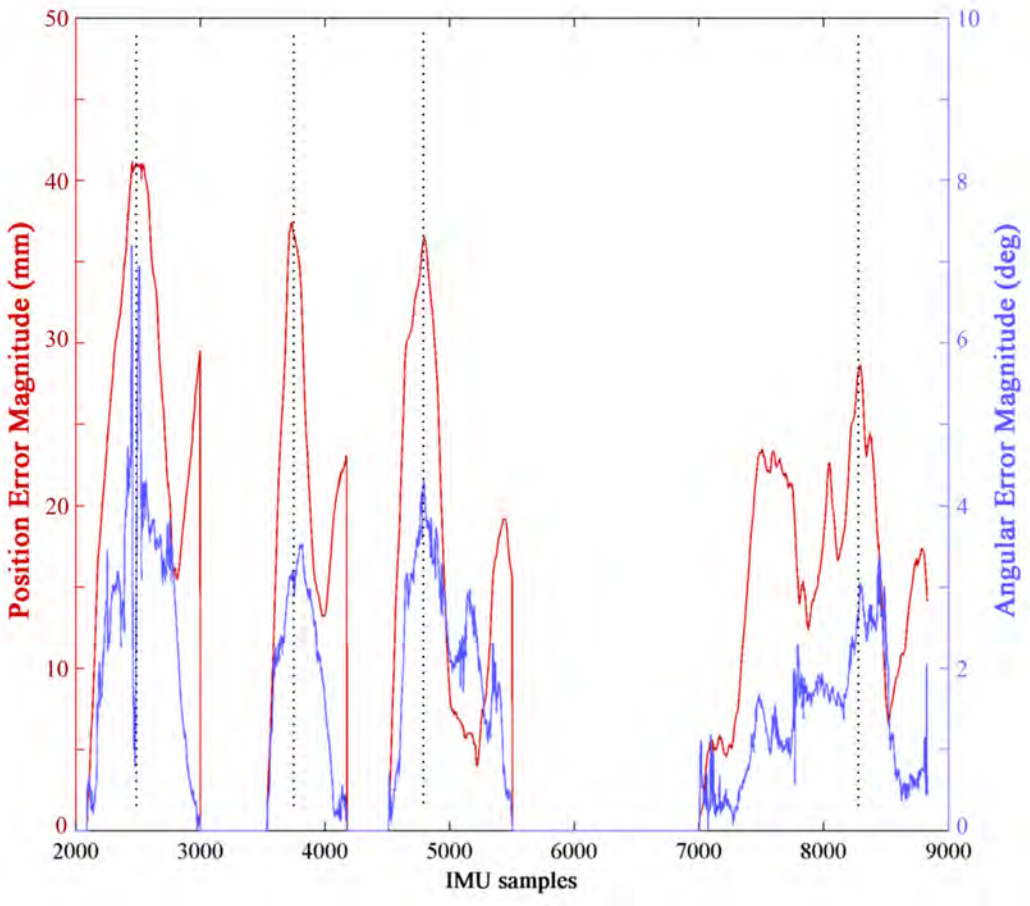


Figure 11.12 – Trajectory error against orientation error

Further examination of figure 11.12 reveals that during each occlusion event a secondary peak occurred with little correspondence to orientation error. It is concluded that such errors are due to deviation of the mechanics of the armature from the ideal of the model. To test this hypothesis the model was driven completely via the orientation estimates obtained from optical marker data. As this eliminated the error due to inertial orientation estimation, any remaining positional error was attributed to the model. Figure 11.13 shows the results of this test against the original inertial extrapolation. The optical data replicates the final peak in the graph of the error in estimated position (shown by the dotted lines) confirming the deviation of the model from the measured trajectory data.

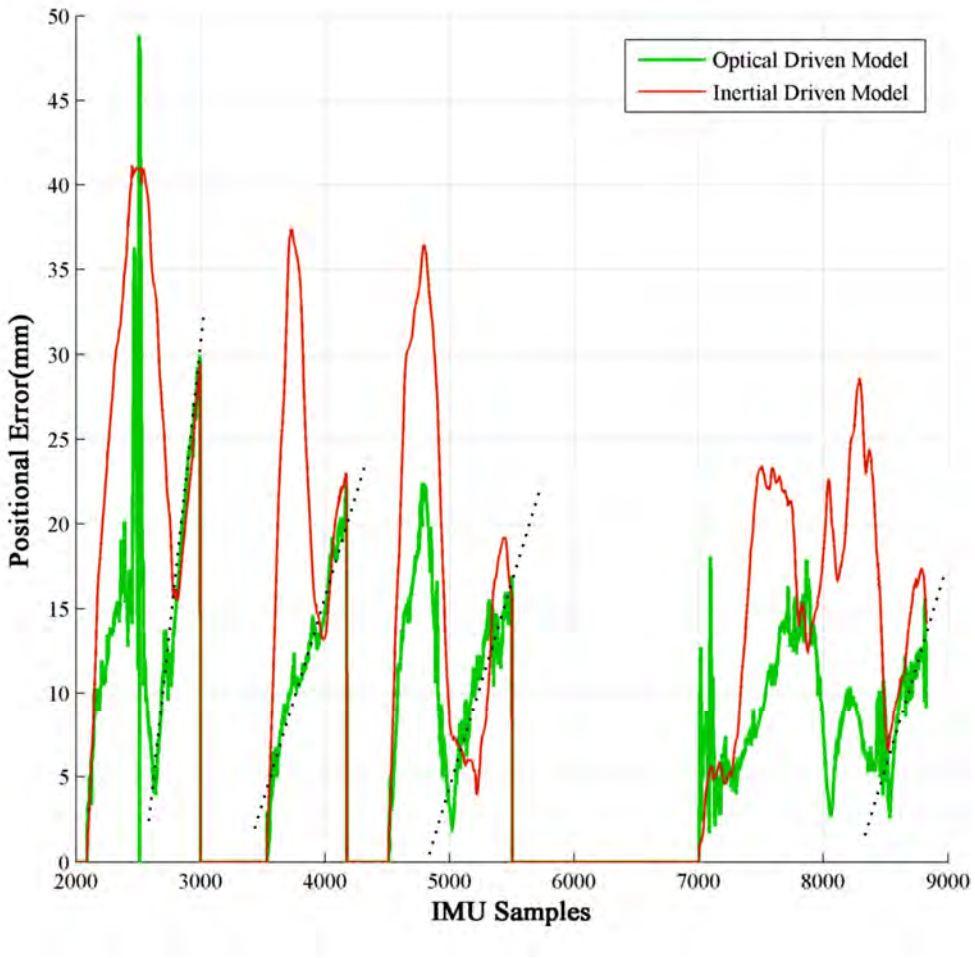


Figure 11.13 – Optically driven test reconstruction

A linear interpolation method was devised to allow for smooth trajectory transitions between inertially extrapolated and measured marker trajectories. This was achieved in a manner resembling the technique described in (10.1.5), where the difference between the final position estimate and the first optical measurement was scaled by the ratio of the occlusion duration. This correction was found to give considerably improved performance when applied to the experimental data. This can be seen in figure 11.14, where mean positional error is reduced by approximately half compared to the uncorrected case, giving a mean value of approximately 10*mm* (excluding non occluded periods) and 35*mm* peak.

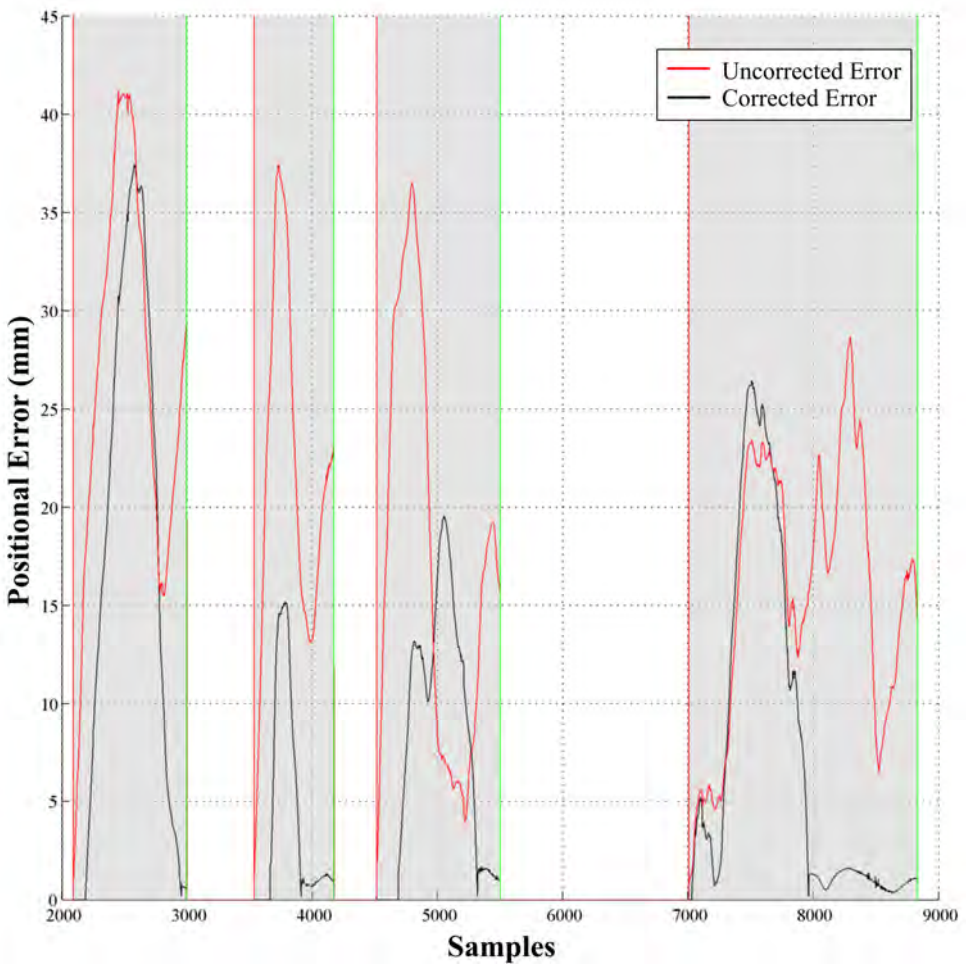


Figure 11.14 –Linearly corrected trajectory extrapolation

Despite these positive results this method has no capacity to maintain either the relative geometry of marker clusters or the kinematic constraints of the body model. Therefore, it will produce corrections which break the correspondence of markers within their cluster or to their centre of rotation. This effect is mitigated by the relatively small positional errors involved and the similarity of positional deviations experienced by markers common to a particular cluster.

11.3 Conclusion

A method has been devised by which occluded marker trajectories can be recovered using inertial measurement, with accuracy appropriate for human motion measurement (~10*mm*). However, the method is intrinsically reliant upon the availability of future trajectory data for the provision of error correction. This limits its usefulness to situations where trajectory measurements are known both before and after an occlusion event. For these criteria to be met it is required that tracking across occlusions is maintained and the identities of trajectories are preserved. However, under practical measurement conditions, occlusion events interrupt tracking leading to loss of marker identity and failure of the method. This is discussed further in the next chapter and modifications to the method are proposed to address this problem.

12. INERTIAL MEASUREMENT ASSISTED TRACKING ACROSS OCCLUSION EVENTS

12.1 Reconstruction of fragmented marker trajectories

The destructive effects of occlusion on passive optical motion capture extend beyond the loss of trajectory information, having detrimental effects on identification of the sources of trajectories and their tracking. This results in ‘trajectory fragmentation’, or the separation of a single trajectory into multiple unidentified pieces. This chapter will detail the development of an extension of the inertial extrapolation technique, allowing marker identity to be maintained across occlusion gaps in the presence of interference from distracting trajectories. Combined with inertially assisted reconstruction, this technique will form the basis for the final trajectory reconstruction function, which is adapted for use with human motion data in the penultimate chapters of this thesis.

12.1.1 Collection of armature data with distracters

Considering an occluded four marker cluster (or target cluster), any unidentified marker which emerges from occlusion (trajectory fragment) represents a possible solution for continuation of the trajectory of one of the cluster’s constituent markers. In the case where the missing cluster is the only source of trajectory data in the capture volume, it is certain that each fragment will correspond to a marker in the occluded cluster. Therefore, the problem of trajectory identification is simplified to finding which marker in the target cluster corresponds to each fragment. However, in a practical capture scenario many other clusters will usually be present, leading to alternative sources for the fragment, many of which have no association to the cluster being tracked. Under this condition, the tracking problem is extended to include the rejection of spurious trajectory fragments from other clusters. In cases where many such

clusters are visible, often with similar geometries and trajectories (such as in the full body marker set), finding a tracking solution can be extremely challenging, often requiring manual assistance to resolve.

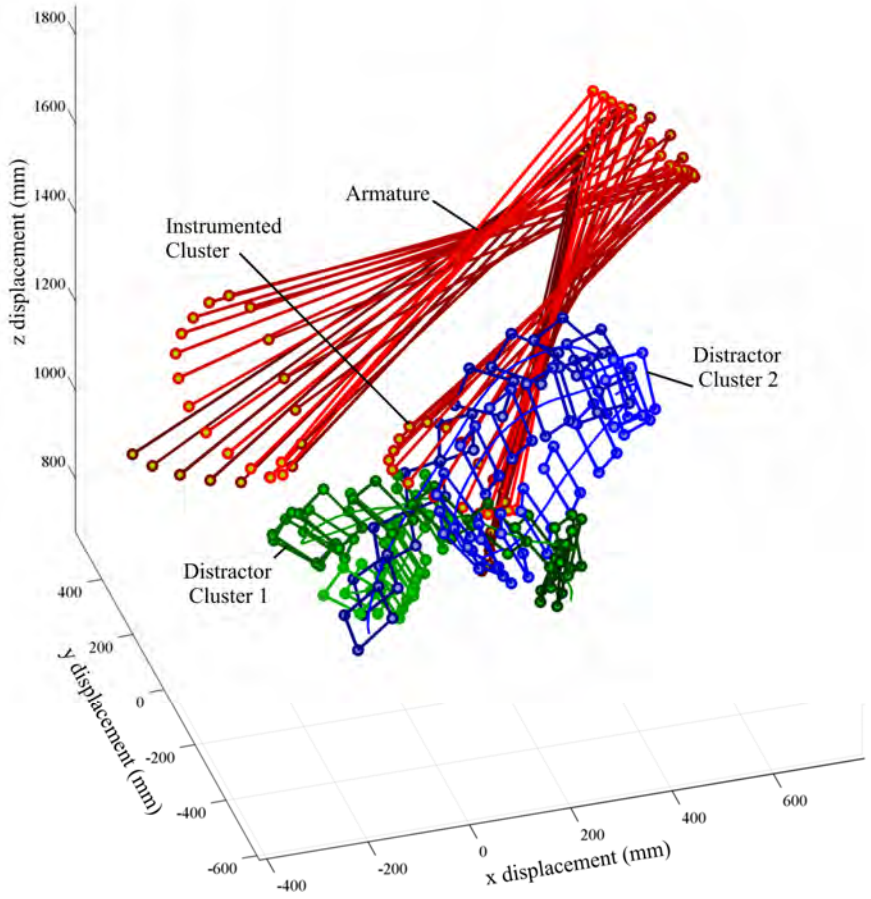


Figure 12.1 – Armature trajectory with distracters

To facilitate development of a tracking solution, it is first necessary to obtain a dataset exhibiting multiple distracter trajectories alongside the target cluster. Additionally, it is desirable that kinematic parameters (CoR, orbit radius etc.) of the structure to which the cluster is attached are well defined and verifiable by measurement. Due to its complexity and the lack of a directly defined kinematic structure, the primary dataset is unsuitable for this

purpose. Therefore, a reduced dataset was acquired by repeating the armature experiment (11.2.1) with the addition of distracter marker trajectories. This distracter data was obtained via two additional marker clusters of similar geometry to the instrumented cluster, approximately following the armature's spherical movements. An example of the trajectory data is shown in figure 12.1 for a 500 frame section of the acquisition, illustrating the relative locations of the four clusters. Two acquisitions were performed and processed via the tracking and labelling operations (11.2.1).

12.1.2 Preparation of artificially fragmented trajectory data

In the examples of trajectory reconstruction given previously (11.2.2), the model of occlusion is insufficient to represent the problems introduced by trajectory fragmentation. Therefore, it is necessary to devise a means by which a single source trajectory may be artificially subdivided into multiple 'anonymous' fragments based upon the timings of simulated occlusion events. This trajectory division process operates in the manner depicted in figure 12.2, where the raw trajectories of two markers are considered over the duration of their measurement and each occlusion event. During this operation each marker's co-ordinate value is copied column-wise from the original raw trajectory array to a secondary 'anonymous' array and compared to the original list of occlusion times and durations. When a marker is noted to have entered occlusion, for the duration of this occlusion a null co-ordinate value is written to the array. This continues until the specified duration of occlusion has elapsed, where copying of marker data resumes at the next available row in the array. This effectively results in each occlusion event creating an independent trajectory (or fragment) in the output data set. This process of trajectory duplication and fragmentation is repeated till all frames have been tested against all markers. With regard to the actual sources of the occlusion

timings, this can be the output of any of the methods of occlusion simulation discussed in (7.1 and 7.2).

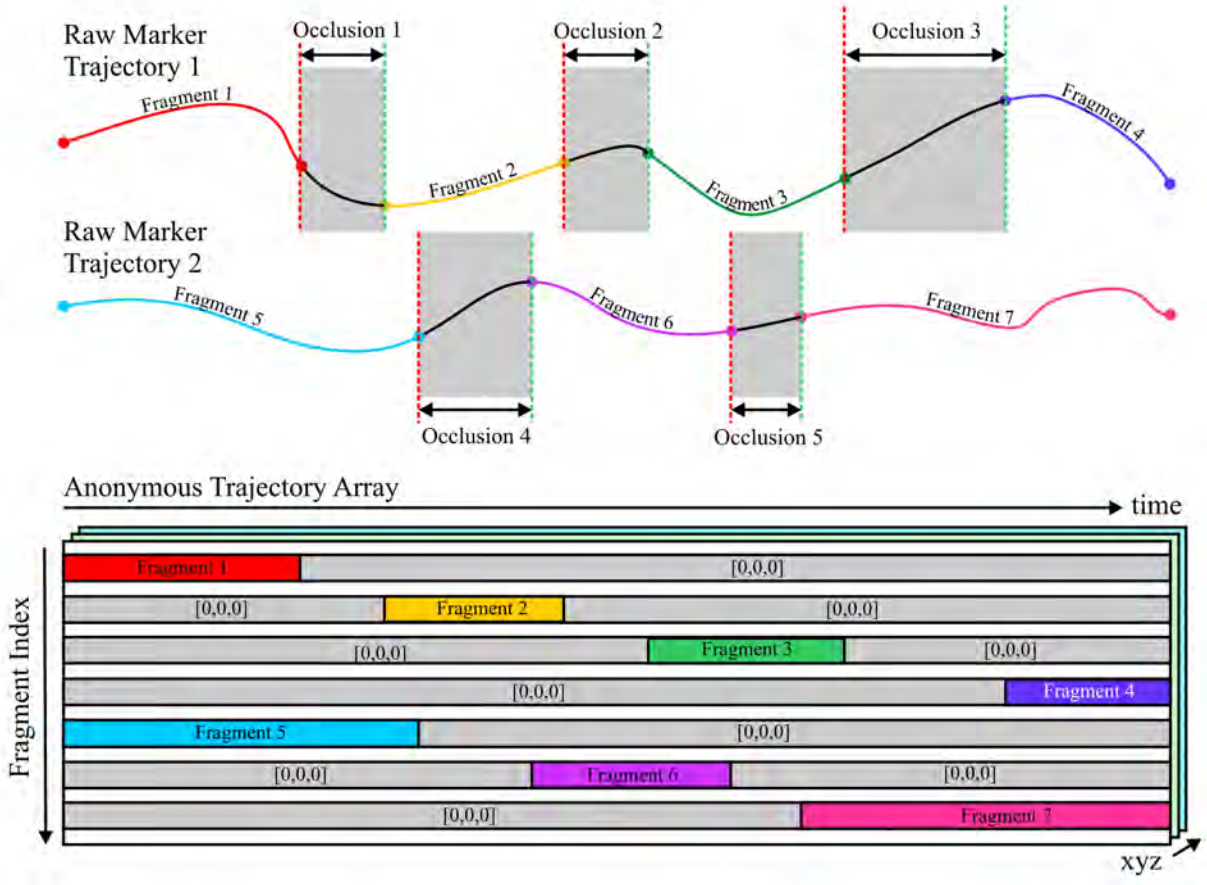


Figure 12.2 – Artificial trajectory fragmentation

12.2 Generation of tracking hypotheses

12.2.1 Resolving marker identity by geometric correspondence

In addition to the calculation of relative orientation between two sets of points, the estimation of absolute orientation (6.1.4) facilitates the assessment of geometric error between two point sets. This method provides the best fit rotation which minimises the average distance between two sets of origin centred points ($\mathbf{p}_{\mathbf{r}_n}$ and $\mathbf{p}_{\mathbf{l}_n}$). Therefore, with rotation represented by the matrix $\mathbf{R}_l$, the average minimum distance *e* between the two sets is obtained by the function:

$$e = \frac{1}{N}\sum_{n=1}^{N}\left\|\mathbf{p}_{\mathbf{r}_n} - \mathbf{R}_l\mathbf{p}_{\mathbf{l}_n}^{T}\right\|$$ where *N* is the number of markers in the set.

In cases where both point sets are identical, regardless of orientation, the value of *e* will be zero indicating perfect correspondence. However, if the two sets are geometrically dissimilar, then *e* will be found to increase proportionally with the mean Euclidean distance between corresponding pairs of points. Therefore, this 'correspondence error' provides a metric for geometric similarity between sets of points independent of gross position and orientation.

It is possible to apply this metric to establish whether a given set of trajectory fragments represent the continuation of the target cluster's trajectories following occlusion [30]. However, due to the errors present in real measurement data perfect correspondence and value of *e=0* will never occur. Therefore, it is necessary to seek the minimum value below a threshold defined by the measurement error level for the target cluster. This technique of discriminating fragments by geometric correspondence forms the basis for the tracking methods described in the remainder of this chapter.

12.2.2 Basic trajectory hypotheses generation

To utilise geometric correspondence for the identification of a target cluster it is necessary to test its last known co-ordinate values against every possible combination of visible fragments. To this end, a set of cluster "hypotheses" ($\mathbf{P}_c$) is assembled through the calculation of each permutation for each combination of all visible trajectory fragments. This set is tested against a planar reference obtained from the target cluster (6.1.3) and the geometric error is recorded for each hypothesis. By thresholding these values, permutations of fragments with superficial resemblance to the target cluster are rejected. Repeatedly applying this operation to each frame of data, it is possible to search for the hypothesis which bears closest resemblance to the target cluster and reestablish tracking from this point. The stages of this process are illustrated in figure 12.3 using data from the armature data set.

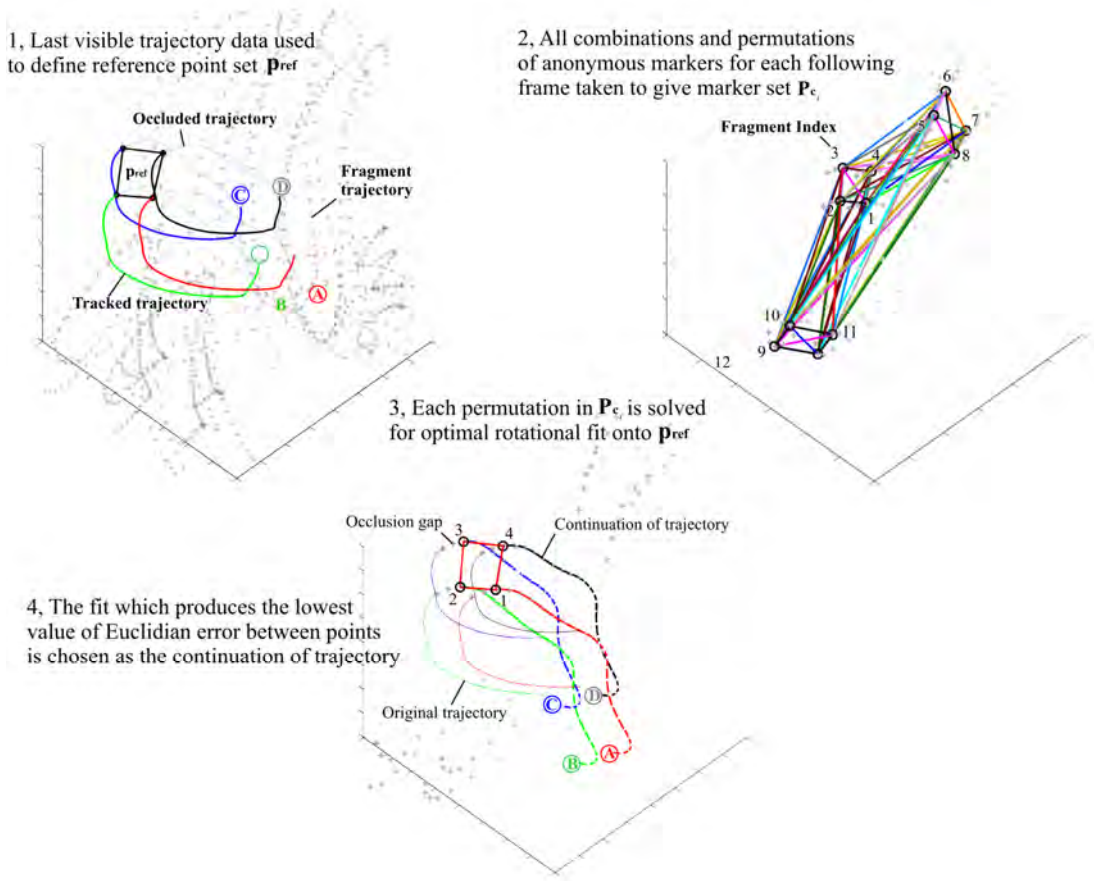


Figure 12.3 – Recovery of trajectory by geometric error method

The values of correspondence error generated by the fragmented armature data set can be seen in figure 12.4 where geometric error is plotted for each hypothesis. Insets detail the positions of the local minimum values of error corresponding to the target cluster (**b**) and distracter clusters (**d1** and **d2**). Labelling of the horizontal axes indicates the order in which fragments were tested, for example 1, 2, 3, 4 is the first to forth trajectory fragment being tested against markers 1-4 in the target cluster. Only marker fragments from the target cluster and distracters were considered, giving a maximum of 12 fragment trajectories.

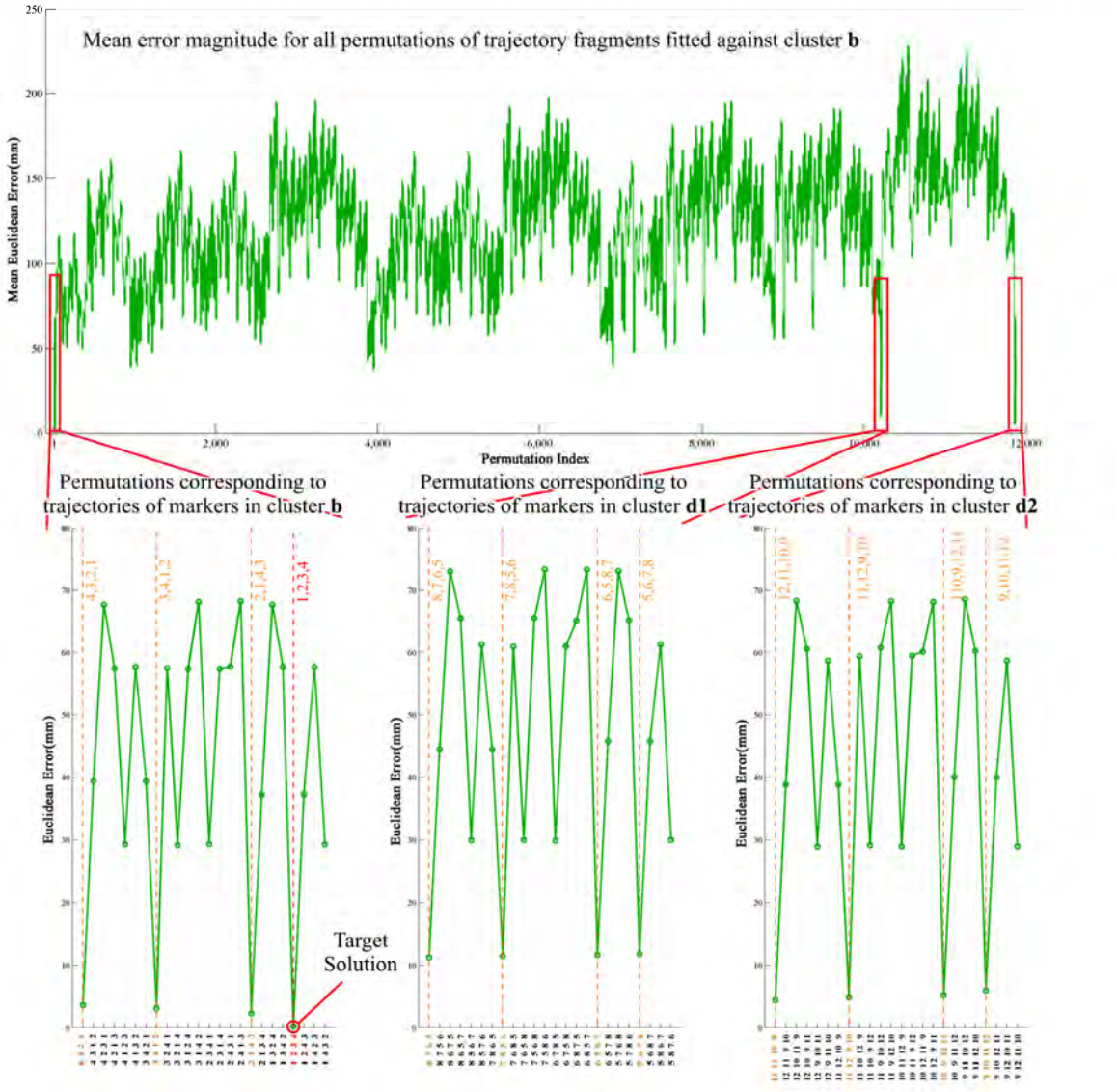


Figure 12.4 – Geometric Error Magnitude for multiple hypotheses

From these results the geometric technique appears to present a workable solution for cluster identification, delivering a minimum error value for the correct fragment hypothesis (1, 2, 3, 4). However, further examination of this data reveals major drawbacks of the technique.

For increasing numbers of fragments the number of hypotheses generated follows a factorial function resulting in large numbers of calculations. For example, in the case of the 12 fragment example, the extent of the hypothesis set is 11.9k, consisting of 495 individual combinations of fragments multiplied by the 24 possible arrangement of each combination. However, in the case of the 65 marker set this increases to approximately 16.2M hypotheses (assuming 100% fragmentation). In this case a 200 frame occlusion gap will require 3.2×10^9 orientation estimations. Given the implementation of the orientation estimation requires approximately 600μs to execute (3.4GHz Intel Pentium 4) this results in a processing time of ~581 hours. As a typical capture session may contain hundreds of such events, albeit for reduced numbers of fragments, this is clearly an unacceptable situation.

A critical flaw of the geometric method is its sensitivity to perturbation of cluster co-ordinates. Although the great majority (>99%) of hypotheses yield large geometric errors (Figure 12.4) a small number come extremely close (<10mm) to the error level of the measurement. This makes confusion of clusters and their orientations

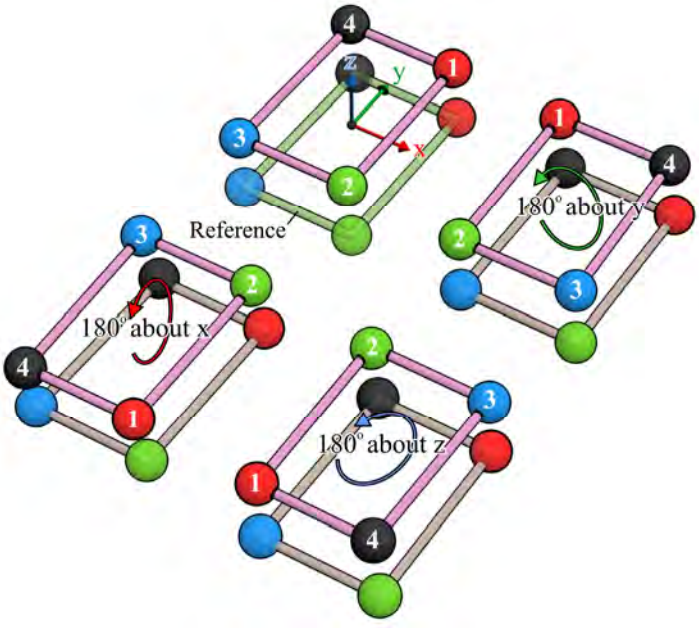


Figure 12.5 – Rotationally symmetric aliases

unavoidable. These close contenders to the solution occur via symmetries between clusters over the four possible 180^o rotations which provide closest correspondence between their constituent markers (Figure 12.5). Due to these inter and intra cluster symmetric 'aliases', a relatively small perturbation of a marker's co-ordinate value can influence the value of geometric error. This can lead to the selection of a spurious hypothesis for trajectory continuation.

12.2.3 Hypothesis reduction by trajectory extrapolation

To provide a reduction in the number of hypotheses, a spatial constraint is applied to the initial selection of trajectory fragments. This constraint is derived through the adaptation of the trajectory extrapolation described in the previous chapter (11.2.2). By this technique the body model and inertial measurements are used to extrapolate an occluded marker's position throughout the course of an occlusion event. However, due to the absence of correction provided from optical measurements, this estimate will deviate from the true trajectory with increasing duration of occlusion. To address this limitation, a spherical volume is defined around the extrapolated marker position forming a 'prediction volume', which encompasses the true marker position. The extent of this volume in space is used to restrict which fragments will be considered for geometric correspondence. This limits the hypotheses tested to those in close proximity to the true marker location.

The dominant source of error in trajectory reconstruction is gyroscope bias (11.2.3), which results in an approximately linear increase in orientation error over short periods of integration (<~10 *sec*). Modelling this error as proportional to the duration of cluster occlusion (t_{occ}) and the maximum drift rate of the orientation estimate (θ_d) it is possible to

calculate the worst case angular error. The maximum positional error (p_d) can then be estimated as the sine of magnitude of orientation error multiplied by the distance between target marker and centre of rotation (d_{cor}). In practical application the value of θ_d is derived from the bias tracking algorithm (11.1.4) and d_{cor} from the armature markers p1 and p2 (or later the body model CoR). Finally the value of p_d is scaled by a 110% safety margin to ensure actual drift will always lie within the prediction volume.

The prediction volume for a single marker is shown in figure 12.6. The scaling of the prediction volume over time is depicted as an undulating cone, whose axis extends from the start of occlusion to the point of reappearance.

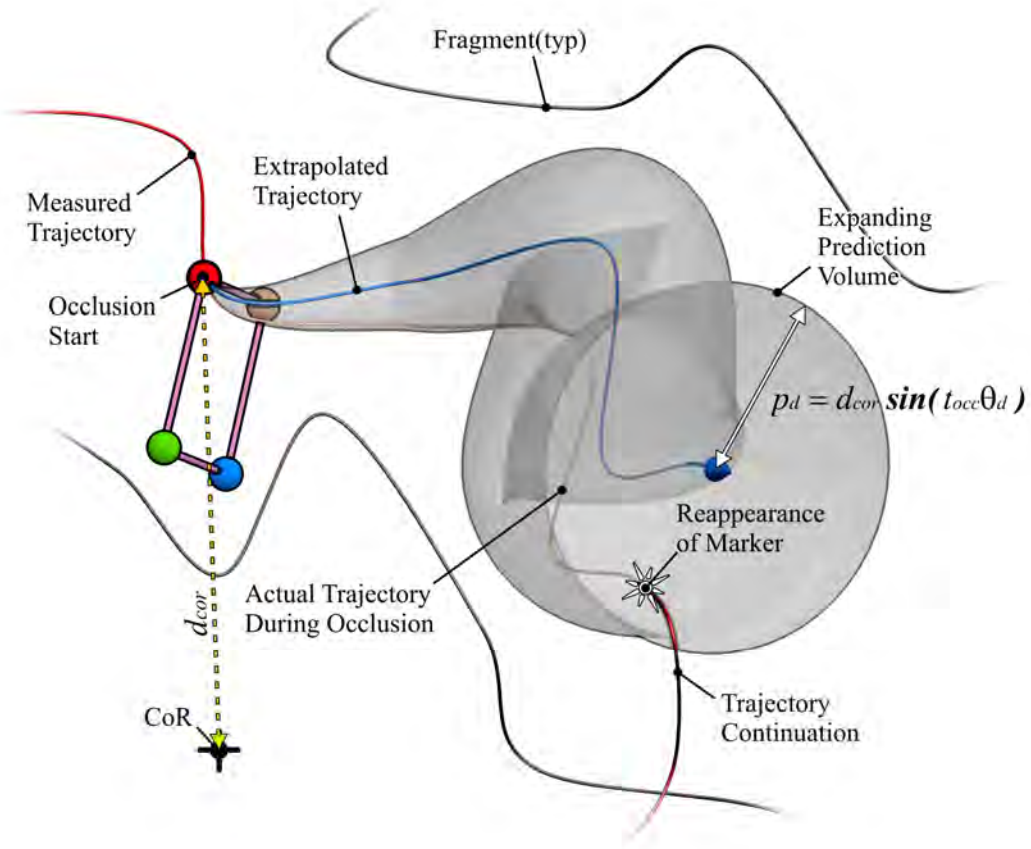


Figure 12.6 – Expansion of prediction volume

Combining this technique with the geometric correspondence method, each fragment within the prediction volume is marked as a contender for trajectory continuation. This is repeated over successive frames for the extrapolated location of each marker in the target cluster until at least one valid cluster hypothesis is found. These hypotheses are in turn tested for geometric correspondence.

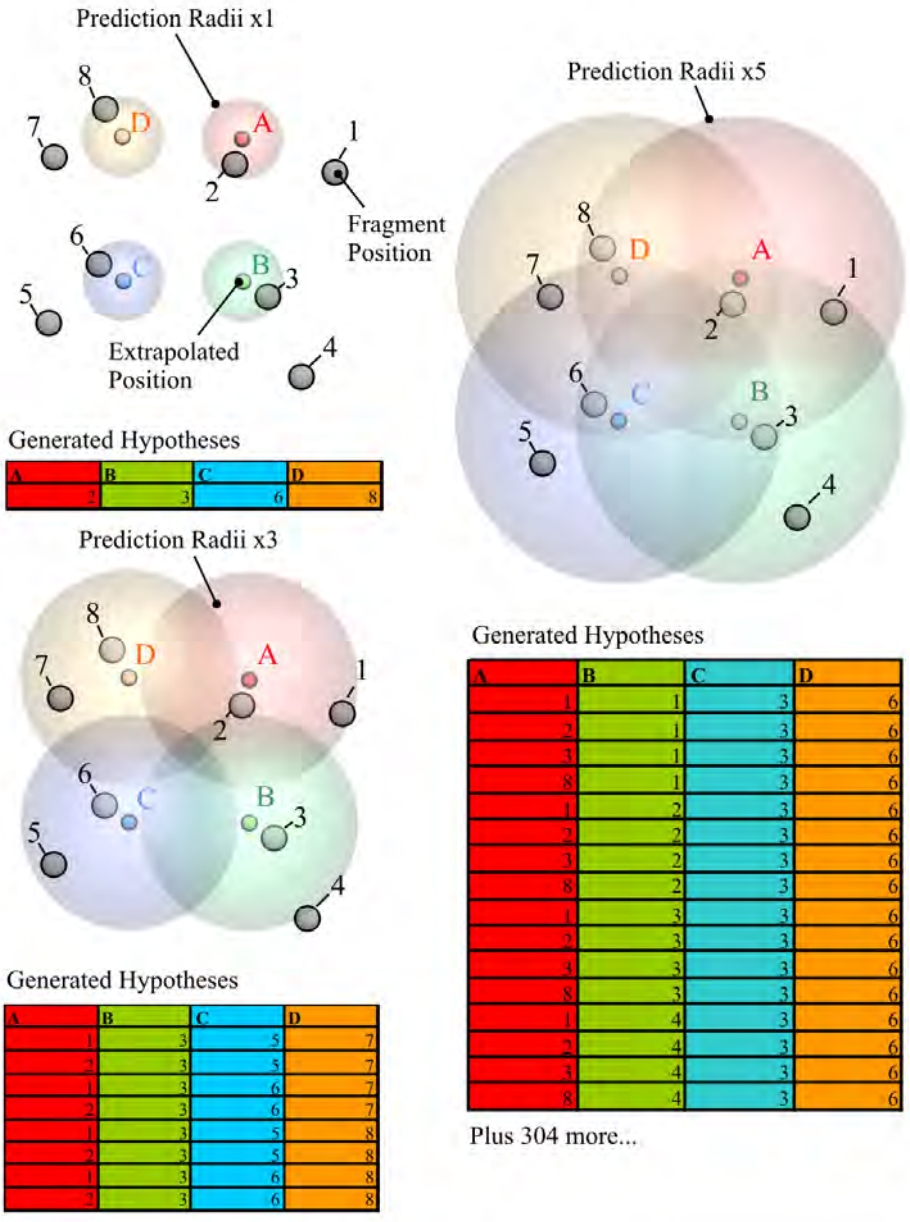


Figure 12.7 – Hypotheses increased by volume radius

This prediction volume method is detailed in figure 12.7 for three different sizes of the four prediction volumes of a typical cluster. It is shown that increasing the prediction radius results in a sharp increase in hypotheses. For larger radii, fragments will appear in multiple volumes. This leads to the return of symmetric aliases and the problems they entail. Furthermore, should this expansion progress unchecked, each prediction volume will eventually enclose the entire measurement volume. This increase could be prevented by imposing a hard limit on the maximum radius of the prediction volume. However, this will limit the useful duration of occlusion recovery.

12.2.4 Hypothesis reduction by intra-fragment distance

By employing a secondary positional constraint for hypothesis rejection, it is possible to restrict the number of fragments tested by the prediction volume set, without resorting to limiting. This is achieved by calculation of the distances between pairs of fragments and comparing them to the inter-marker distances in the target cluster [30]. In the case of a pair lying outside the maximum and minimum distances found in the target cluster, the pair is eliminated from further hypotheses. Practically, these thresholds for rejection are increased slightly to reflect the expected measurement error. This inter-marker rejection method dramatically reduces the number of hypotheses to be considered by the prediction volume method.

This technique is particularly effective in eliminating hypotheses involving markers between clusters. Figure 12.8 shows the method applied with the right upper arm marker cluster designated as the target. To distinguish results, marker pairs within the acceptance range are signified by red lines, whereas those outside the range are coloured grey. An excellent

reduction in the number of possible marker pairs is achieved (2016 to 63 pairs), almost all of which lie within the same clusters. Additionally as testing only requires half the total number of fragments squared, computational requirements are reasonable.

Due to its similarity with the geometric correspondence method, this method offers no improvement in terms of trajectory discrimination. Therefore, the advantage of this method lies solely in its ability to efficiently reduce the number of initial hypotheses.

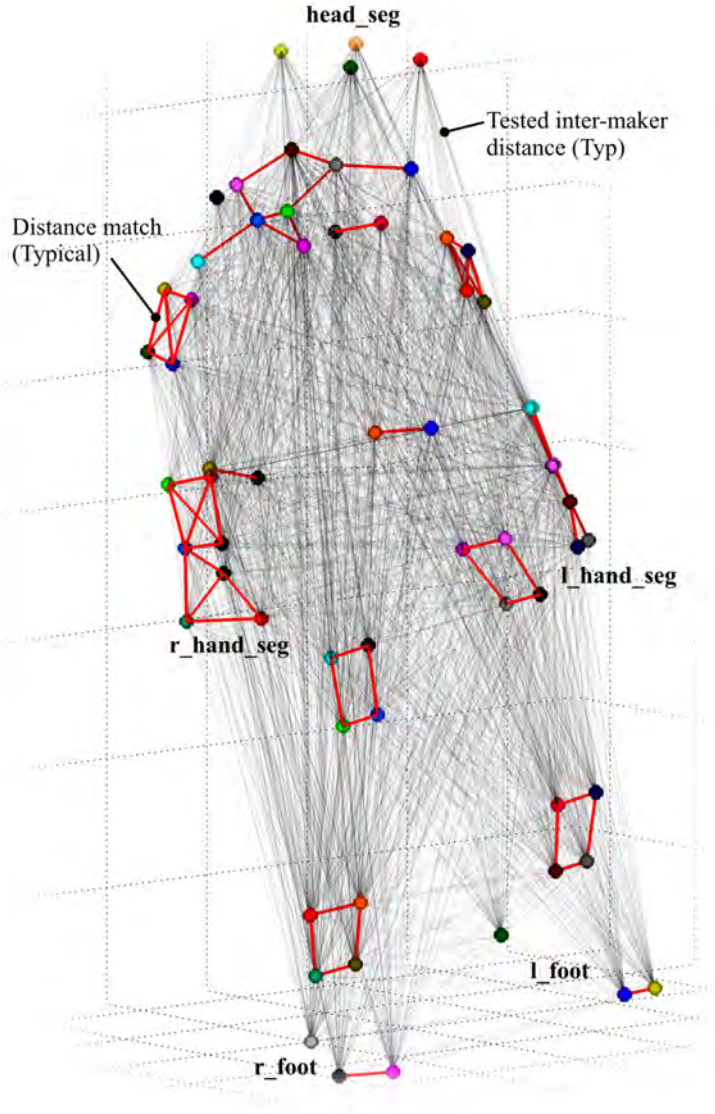


Figure 12.8 – Reduction by inter-fragment distance

12.3 Temporal correspondence of fragments

The methods described up to this point rely exclusively upon information obtained from a single frame of fragment data. To further improve rejection of erroneous hypotheses, particularly the symmetric aliases, it is necessary to take into account variation of fragment position over time. This section presents three temporal correspondence metrics, which when combined give greater discriminatory capability than the geometric method alone.

12.3.1 Temporal correspondence of position

By continuous calculation of average Euclidean distance between each fragment and inertially extrapolated trajectory over time it is possible to derive a coefficient of trajectory similarity for a particular hypothesis. However, this average must take into account the increasing positional error with duration of extrapolation (11.2.2). By utilising the error prediction scheme given in 12.2.3, Euclidean distance is normalised against worst case extrapolation error for each frame. Averaging this normalised distance results in value in the range 0 to 1, where 1 signifies perfect correspondence and zero the worst case error. Applying this operation over a predefined sample window (or the length of the shortest fragment in the hypothesis) yields a single value by which relative similarity of trajectories is assessed.

As this method favours groups of fragments bearing the closest relation to the extrapolated trajectories, it gives good discrimination against distracters with similar geometric configurations but different position. However, when the prediction radius exceeds the average inter-marker distance, reduced sensitivity to cluster orientation will result in poor discrimination of intra-cluster aliases.

12.3.2 Orientation Correlation

To address the problem of symmetric solutions (Figure 12.5), a correlation coefficient of orientation between extrapolated and hypothesised cluster trajectories is calculated over the sample window. This correlation is performed between each pair of column vectors comprising the two rotation matrices resulting from estimation of extrapolated and hypothesised cluster orientation. Before correlation, each vector component is high pass filtered at $0.2Hz$ to discount any constant orientation error and drift. The correlation coefficient of each pair of vectors gives a value which will represent the similarity of rotation over the three orthogonal axes defined by each rotation matrix. Averaging these three values gives a gross measure of similarity in change of orientation between extrapolated and hypothesised cluster trajectories.

Although insensitive to absolute position, the orientation correlation provides excellent discrimination of marker order within hypotheses. This sensitivity is due to the three symmetric aliases of marker co-ordinates effectively constituting 180^o rotations. As such a change in marker order will lead to the reversal of at least one local axis of a cluster then any rotation of the cluster about that axis will be seen to be reversed, resulting in a negative correlation. Alternatively, in the case where a correct hypothesis is proposed, all correlations will be positive. This effect is shown in figure 12.9, where correlation coefficients and marker order are plotted for three incorrect and one correct permutation (far left) over a one second window. The graphs show the correspondence of each column vector of the two rotation matrices. Red, green and blue lines represent *x*, *y* and *z* axial components respectively, solid lines show hypothesised, while dotted lines represent extrapolated, orientation.

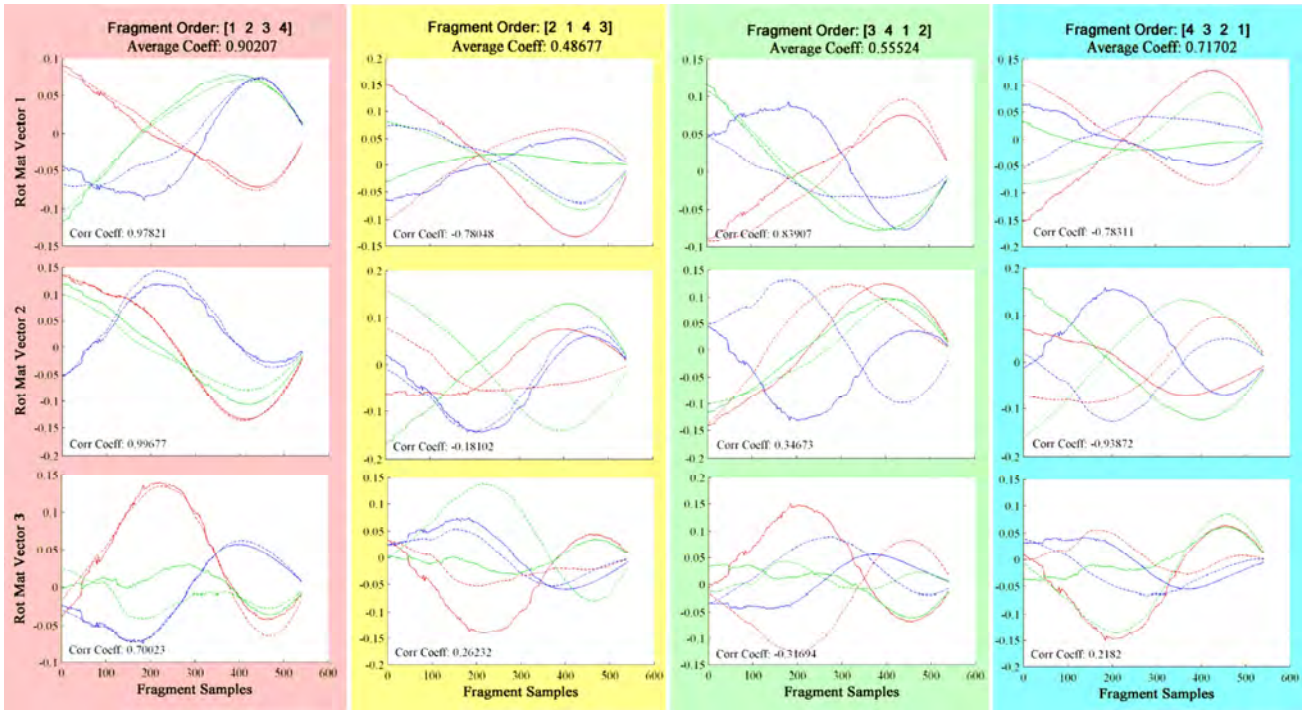


Figure 12.9 – Cluster orientation correlation for four hypotheses

As symmetric aliases have been indentified as the primary cause of tracking failure, this correlation performs an essential function in the assisted tracking algorithm. However, when applied to human data, mechanical coupling between neighbouring segments leads to high values of orientation correlation between neighbouring clusters. As this results in false positive correspondences it is essential that this method be used in combination with the position metric.

12.3.3 Geometric correspondence

Through averaging each value of orientation correspondence over the duration of the sample window a more reliable measure of similarity between hypothesis and target cluster is obtained. In particular, the effects of transient perturbations of marker position (such as measurement noise and soft tissue artefacts) are reduced. To permit comparison with the two other temporal metrics this average is normalised by the estimated peak geometric error

obtained via the bias tracker (11.1.4). As an additional safety measure against false positive solutions, the standard deviation over the window is also taken. By thresholding this value, hypotheses which do not exhibit good rigid body correspondence may be detected and eliminated from further consideration.

12.3.4 Combined classifier

Combination of the three temporal correspondence metrics is achieved through the composition of their results into a single vector. Selection of the hypothesis for continuation of cluster trajectory is determined by thresholding the magnitude of this vector. However, to prevent any single metric dominating this result the angle of the correspondence vector to the unit diagonal vector is utilised as a second thresholding parameter. Therefore, only when a supra-threshold magnitude is accompanied by an appropriately small angle is the current hypothesis considered valid. However, when this angle is large, we deduce that at least one of the metrics exhibits poor correspondence and reject the hypothesis accordingly. In cases where simultaneous hypotheses exist within this range of acceptance, the hypothesis which results in the greatest magnitude is given priority.

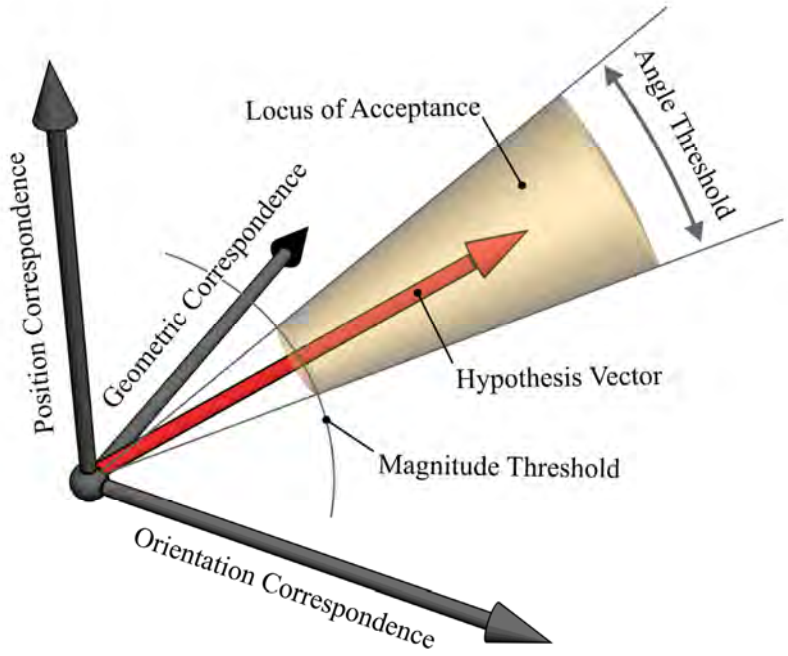


Figure 12.10 – Combined vector classifier

This comparison scheme is depicted in figure 12.10, where the thresholds for acceptable hypotheses are marked and each axis is shown in terms of its correspondence metric. Actual values for threshold angle and magnitude are derived experimentally in the next chapter (13.2.2).

12.3.5 Blending and Backtracking

Following selection of a solution for trajectory continuation, it is necessary to ensure continuity between extrapolated and measured data. As discussed in (11.2.3), this is achieved by a linear blending function. However, as the markers comprising an entire cluster will rarely be occluded and reappear simultaneously, some means of accommodating trajectories that overlap the start/end of cluster occlusion events is required. This is addressed by tracing each trajectory fragment backward in time, from the point the cluster is re-acquired to the fragment's appearance (shown in figure 12.11). Linear blending of the trajectory is then performed from this point. An identical operation is performed in forward direction to account for trajectory data extending beyond the start of the cluster occlusion.

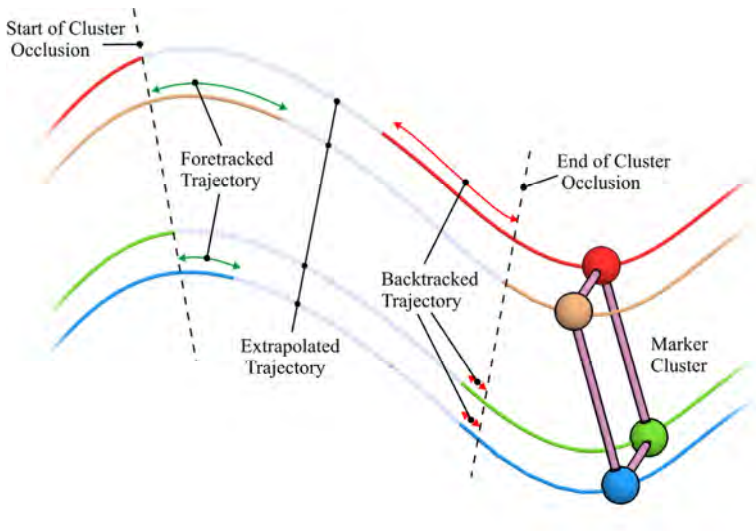


Figure 12.11 – Fore/backtracking of trajectories

12.4 Final algorithm and Conclusion

Combining the hypothesis generation and trajectory classification operations with the established orientation estimation and trajectory reconstruction functions, we obtain the final trajectory recovery algorithm. The block diagram shown in figure 12.12 illustrates its major features and intermediate calculation steps. This algorithm is capable of maintaining cluster tracking across a single occlusion event given initial trajectory data, the segment CoR data and inertial measurements. On resumption of tracking, the errors accumulated during extrapolation are corrected and blended with the optical data selected for trajectory continuation.

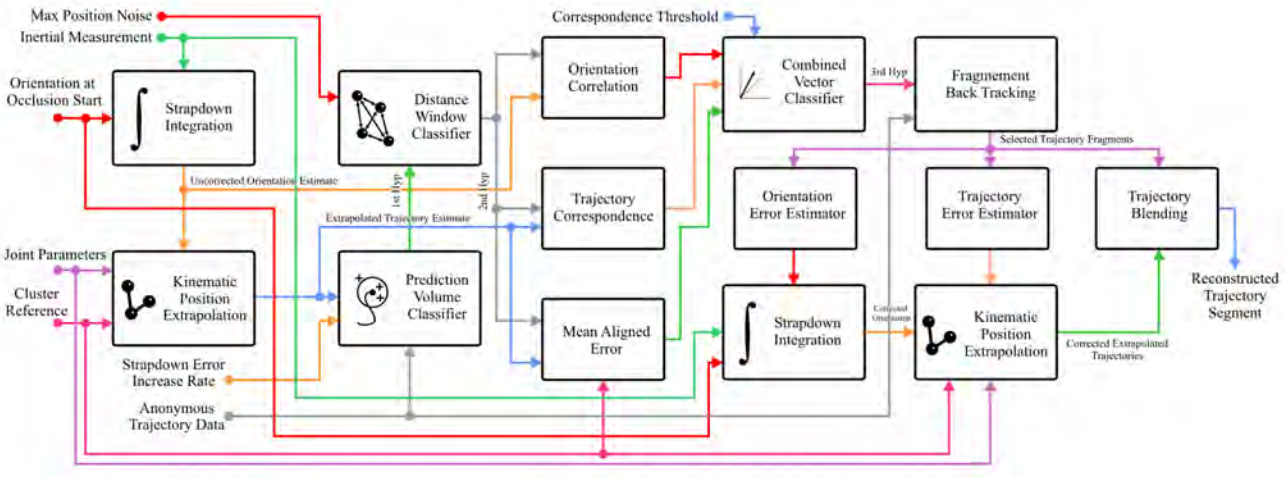


Figure 12.12 – Schematic of final reconstruction algorithm

To conclude, in this chapter an initial test of trajectory reconstruction is performed. For this, the final algorithm is applied individually to four randomly generated occlusions of the armature dataset collected in (12.1.1). In figure 12.13, the Euclidean distance between reconstructed and optically measured trajectory is shown for each marker over the duration of acquisition. A good correspondence between predicted and measured trajectories is observed, producing mean and peak values ~15*mm* and ~40*mm* respectively. For reference. the

recovered 3D trajectory data is shown in the four isometric plots of figure 12.14. Each reconstructed trajectory segment is shown for 200 frames each side of its occlusion events. The recovered and measured trajectories are signified by solid and dashed lines respectively, with points of occlusion and reacquisition marked.

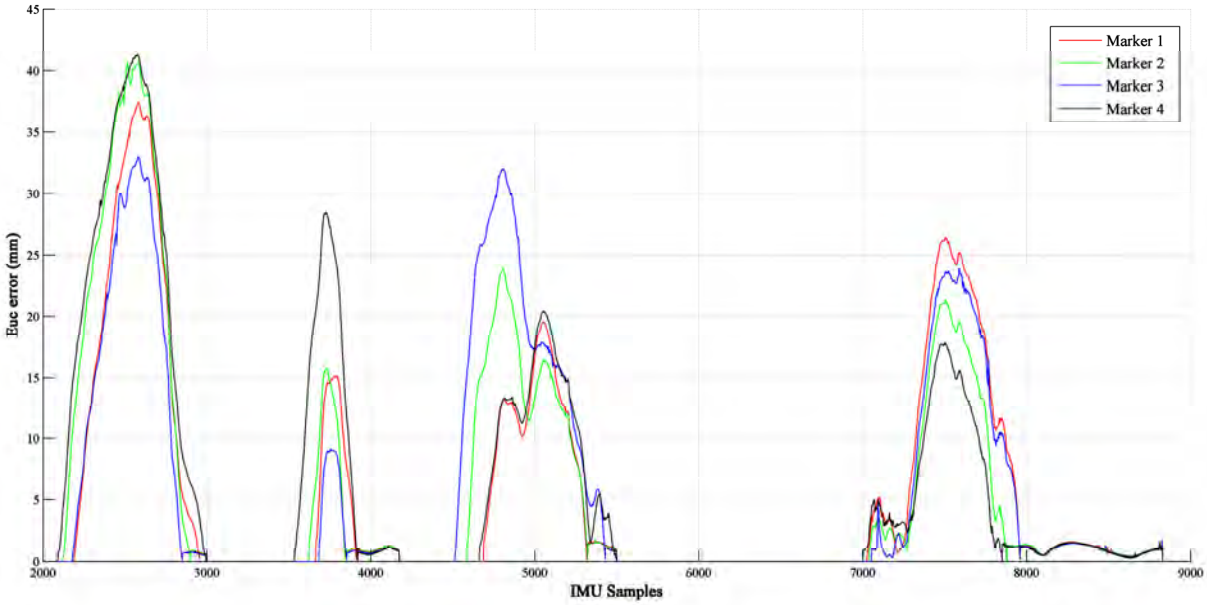


Figure 12.13– Reconstruction error for each marker in cluster

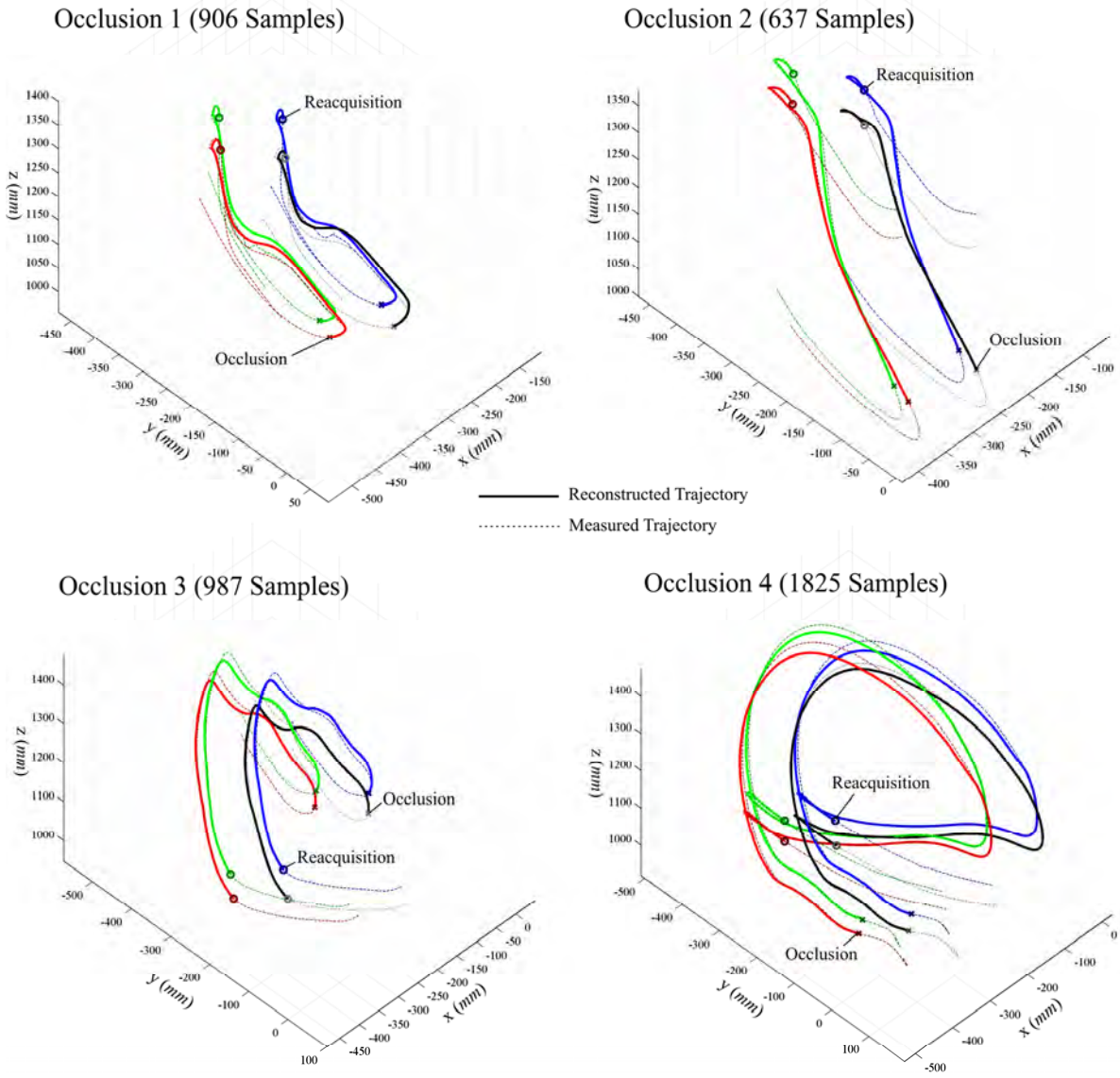


Figure 12.14 – Reconstructed trajectory fragments

13. RECONSTRUCTION AND TRACKING FOR WHOLE BODY DATA

In this chapter, the final full body tracking and reconstruction algorithm is implemented and tested against the primary data set. In the first section, the combination of the single cluster reconstruction with the output from the kinematic body model is described and tuning of this reconstruction function is discussed. The second section of the chapter focuses on the pre-processing of the primary data set through occlusion simulation and application of the reconstructor. Finally, results of this reconstruction are presented and the performance of the method is validated.

13.1 Adaptation of inertial reconstruction to primary data set

13.1.1 Final implementation of reconstructor

To apply the tracking and reconstruction method to the primary data set it is necessary to obtain the parameters required for reconstruction (CoR location etc) from the kinematic body model (Chapter 6). However, before the algorithm can be applied to real data, it should be capable of handling repeated occlusion events over a measurement sequence. This is achieved by division of the algorithm into 3 separate functional states, which are sequentially executed during tracking. This forms a repeated measurement-prediction-correction cycle, which will be referred to as the ‘reconstructor’ and is shown in figure 13.1.

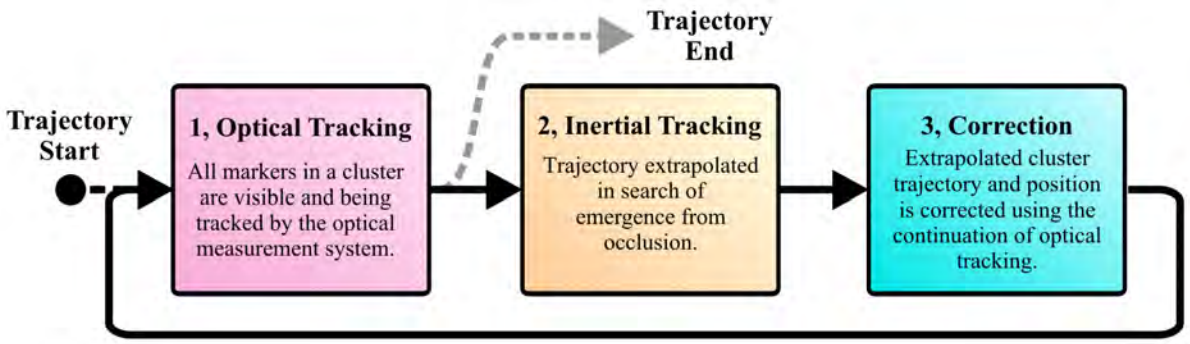


Figure 13.1 – Progression of states of the reconstructor during operation

Assuming initial marker visibility, a typical trajectory reconstruction will start in state 1 and progress through visible trajectory data. Over this period the gyroscope bias estimation will be maintained until an occlusion event occurs. At this point state two will be entered, where cluster trajectories are extrapolated and the trajectory classifier is applied to any visible fragments. Once a fragment hypothesis is found which satisfies the classifier, the third state will be entered and correction and blending of extrapolated data will be performed. The algorithm then returns to state one and optical tracking continues until the next occlusion event. Should the end of the measurement sequence occur during the extrapolation phase, any uncorrected trajectory data will be appended to the reconstructed trajectories. The Matlab implementation of this algorithm is given in Appendix 6 for reference.

13.1.2 Automatic derivation of order of traversal of kinematic skeleton

As has been shown in (11.2.2) for the reconstruction of a target cluster it is necessary that its CoR is defined as a function of a neighbouring cluster's pose. This dependency prevents reconstruction of marker trajectories in isolation from neighbouring clusters. It is therefore necessary to develop some means by which an individual cluster may be associated with its neighbours in the context of the kinematic hierarchy.

To facilitate calculation of the required CoRs an algorithm was devised to automatically traverse the entire hierarchy of the body model given a particular starting point or ‘root’. Practically, this was achieved by recursively searching the kinematic lookup table, described in (6.4.2). The resultant kinematic ‘tree’ is encoded as a list of segment indices, describing the progression through the body model hierarchy in each possible direction from the root. Examples of two such trees are presented in figure 13.2, showing the model originating at the ‘c7_seg’ and ‘r_hand’ segments.

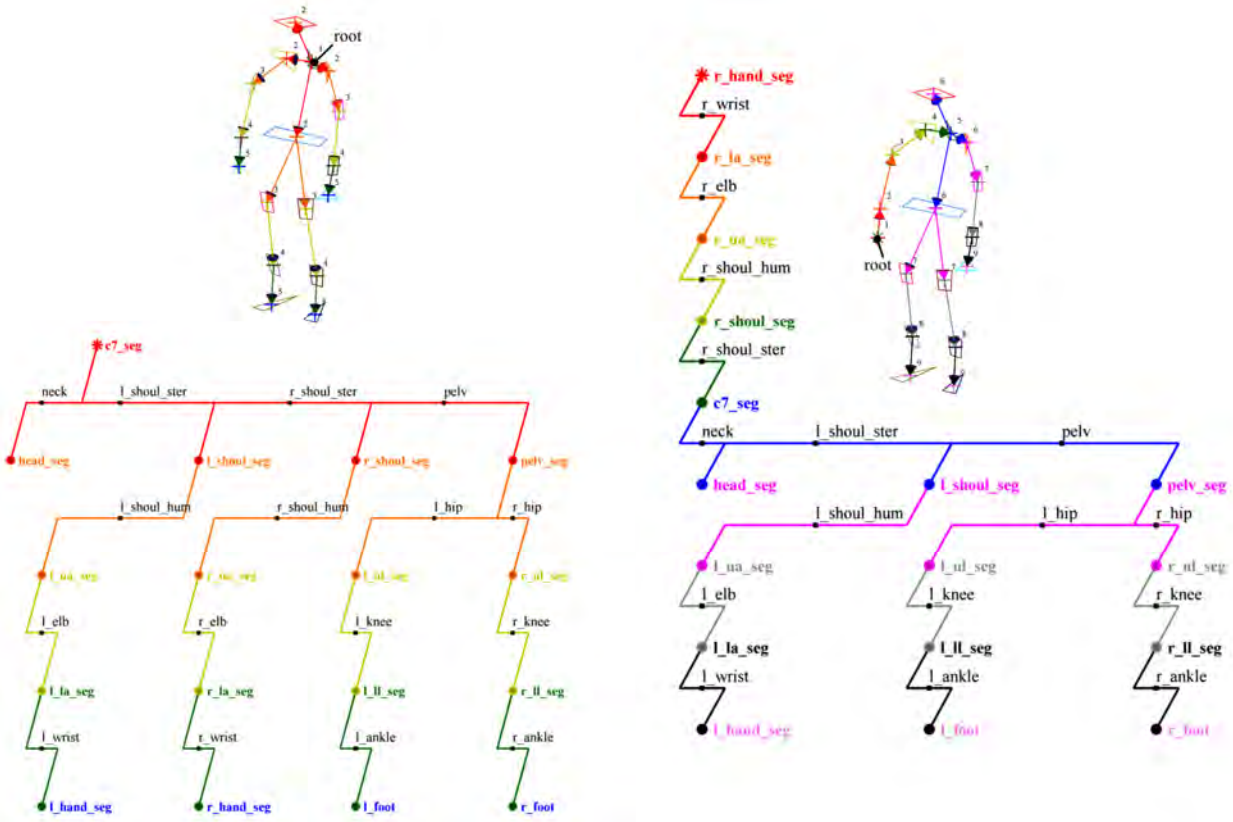


Figure 13.2 – Order of traversal through two kinematic ‘trees’ with different root segments

13.1.3 Recalculation of CoRs via neighbouring segments

Via the kinematic tree, the translation ($\mathbf{p}_c$), orientation ($\mathbf{R}_n$) and local CoR vector ($\boldsymbol{v}_m$) of a segment neighbouring the target cluster is retrieved from the locked data set (6.4.2). As shown in figure 13.3, transforming $\boldsymbol{v}_m$ by neighbouring rotation $\mathbf{R}_n$ then translating by $\mathbf{p}_c$ the CoR of the target cluster is obtained. This is then used to drive the reconstructor as required. In this example, depending on the direction of traversal of the hierarchy the target cluster could alternatively be defined via the hand segment and wrist CoR.

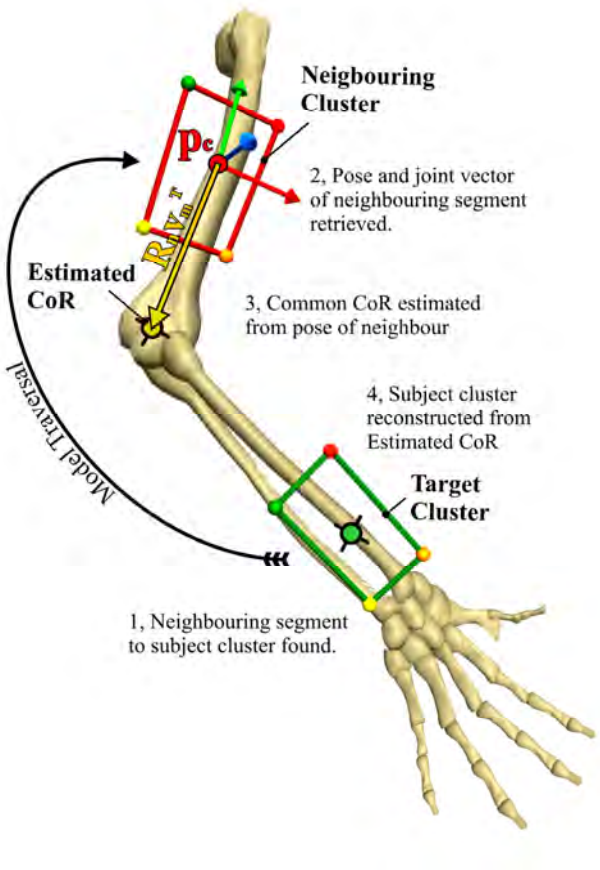


Figure 13.3 – Recalculation of CoRs by neighbouring segments

13.1.4 Fixed order reconstruction with permanent root

The dependence on CoR data from neighbouring clusters leads to dependence on the visibility of neighbouring clusters. Therefore, the reconstructor can be considered to shift this dependence from the target cluster to its neighbours. This renders target cluster reconstruction impossible while neighbouring clusters are occluded.

Despite this dependence, it is not essential that neighbouring trajectory data be provided by optical measurements. Instead, it is possible that the neighbouring trajectory data is itself provided by the reconstructor. Therefore, by ensuring a single cluster in the marker set is visible for the entire capture duration, it is possible to traverse the kinematic hierarchy using each reconstructed cluster to define the CoR of the next. By repeating this process until the terminals of the hierarchy are reached the entire motion sequence is recovered.

In practice this reconstruction sequence is implemented using the kinematic traversal algorithm described in the previous section, with the root cluster elected to be continuously visible. This visibility is artificially introduced via adjustment of the occlusion model. As continuous visibility cannot be guaranteed practically, this method of recovery is not viable for use with real measurements. However, for the purposes of testing the reconstructor via simulated data it provides a simplified means by which complete reconstruction can be ensured.

A drawback of a fixed reconstruction order is that large errors are possible when multiple segments are occluded simultaneously. This is due to the propagation of position errors via the cascading dependence on CoR. In the case of several clusters being occluded simultaneously in their hierarchical order, increased error in reconstruction will affect the ability of the prediction volume classifier to enclose a solution (12.2.3). To address this problem a modification to the reconstructor is made, whereby CoRs based upon extrapolated data are flagged and the prediction volume expanded accordingly. Additionally, the correction mechanisms applied to extrapolation error (11.2.3) will act to limit accumulation of CoR errors.

13.2 Preparation for reconstruction of primary data set

13.2.1 Application of anonymisation functions to trajectory data

The artificially degraded dataset to test full body reconstruction was generated via application of the occlusion models to the primary dataset. The four motion sequences were processed by both statistical (7.2) and volume occlusion models (7.1), with the camera/occlusion rate configurations described in (7.1.8 and 7.2.2). This resulted in a total of 28 fragmented sequences totalling 25 minutes of motion data. Processing of the labelled trajectory data was achieved using the trajectory fragmentation technique described in 12.1.2. Over both models and all conditions a total of 6794 occlusion events were generated with a mean duration of 729 frames (1.56 seconds), SD of 516 frames(1 *sec*) and maximum and minimum durations of 5940 (11.88 *sec*) and 4(8 *msec*) frames respectively. Over all trajectories and acquisitions this equates to an average trajectory loss of approximately 30%.

13.2.2 Tuning of trajectory classifier

To achieve reliable performance of the reconstructor, it is necessary to adjust thresholds of the vector classifier (12.3.4) for optimal discrimination of hypotheses. This is achieved by applying the classifier to gold standard trajectory data, where cluster membership is verifiable. The scatter plot in figure 13.4 shows values of angle against magnitude generated by the vector classifier for a single motion sequence. To distinguish between hypotheses, points which represent a valid continuation of a trajectory are marked in red, while erroneous hypotheses are marked in blue. As correct tracking solutions are collected in the top left hand corner of the graph, thresholds of 0.85 for magnitude and angles below 10^{o} represent appropriate choices for acceptance. However, with these thresholds, the reconstructor was found to occasionally reject correct hypotheses, particularly in the case of clusters exhibiting

poor rigidity. Although the graph indicates a quite narrow separation between correct and incorrect hypotheses, the majority of incorrect hypotheses occur due to rotationally symmetric aliases (12.2.2). As these aliases involve the same trajectory fragments that constitute the correctly orientated solution, they always appear in conjunction with this solution. Given the solution should always produce a higher classifier magnitude than its alias, it will always be selected in preference to it. Therefore, the acceptance range is increased to include the range of alias magnitudes, with the effect of eliminating false rejections with no increase in the occurrence of false positives. For the final testing of the reconstructor, threshold values were chosen of 0.75 and 20^{o} for magnitude and acceptance angle respectively.

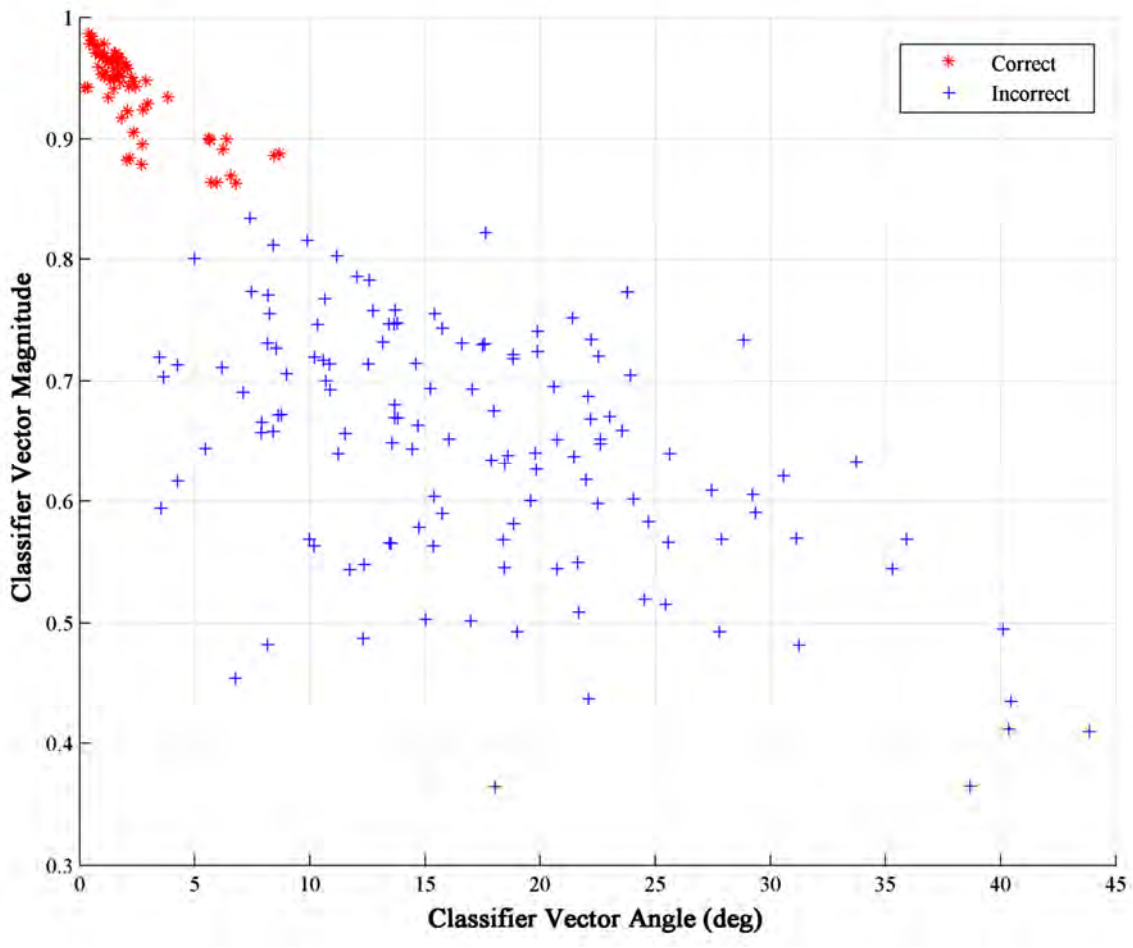


Figure 13.4 – Classifier vector angle and magnitude sorted by correct incorrect hypothesis

With regard to the choice of global sample window length, this was set through trial and error to 500 frames with a minimum of 30 frames. This range gave a number of trajectory points large enough for trends in motion to be established, but low enough for rejection of extremely short fragments. Both window length and correspondence thresholds are maintained at these values for all following tests of the reconstructor.

13.2.3 Initial test reconstruction of markers and trajectories

For initial testing of the reconstructor, the body model was first obtained from the gold standard range of motion sequence by the technique given in (6.4). Following this a single fragmented sequence was submitted to the reconstructor in fixed hierarchical order with the 'c7' segment as root. Simulated inertial data is provided via the sensor model (9.2) applied to the gold standard data. Occlusions occurring in the vicinity of the shoulder clusters and c7 were suppressed for the duration of the measurement. This was due to the difficulty in defining the centre of rotation of the clavicle (6.5.4). The reconstruction of 70 occlusion events in the 24k frames of the motion sequence, took approximately 1 hour on a 3.4*GHz* Intel Pentium 4 PC with 2Gb of RAM.

On completion of the reconstruction, 100% of occlusion gaps were recovered, with no failures in tracking. Error in trajectory reconstruction was calculated as the Euclidean distance between corresponding pairs of markers in the gold standard and reconstructed data. Excluding non-occluded trajectories, a mean error of 5.0*mm* was obtained with a peak value of 30.8*mm*. This error was plotted against the duration of measurement for each individual marker (Figure 13.5). Each peak in this graph represents an occlusion event; errors in excess

of 5*mm* are labelled with the name of their source cluster. Marker identity is signified by colour of trace.

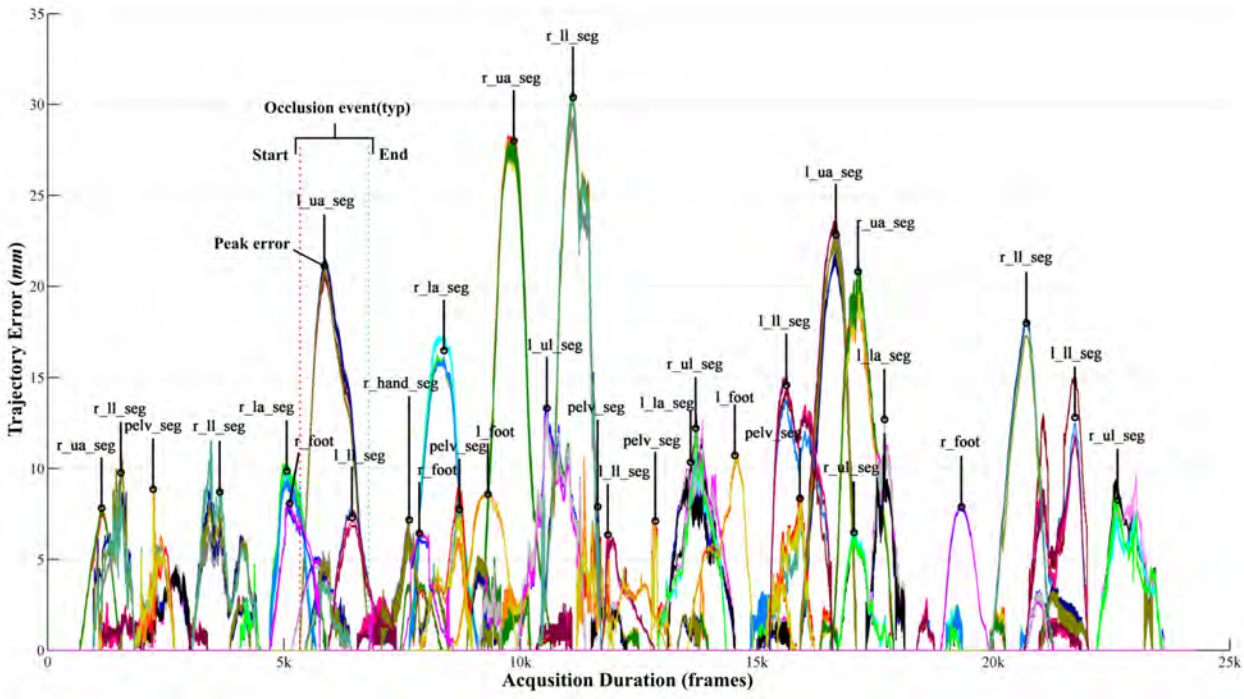


Figure 13.5 – Trajectory errors per marker

This graph indicates that the magnitude of trajectory error is in accordance with the findings in (11.2.3), with peak error occurring at the mid point of each occlusion event. The 1-2*mm* high frequency component visible throughout all series is due to the residual noise of the optical measurement.

Figure 13.6 shows the mean, standard deviation and maximum of error per cluster. The average error over all clusters lies within the 10*mm* range considered acceptable for trajectory recovery. This level of performance is stable across all clusters. Peak values of error occur due to problems with the shoulder model leading to poor estimation of the upper arm CoR (6.4.2), and a transient occlusion of the lower leg cluster in the gold standard data.

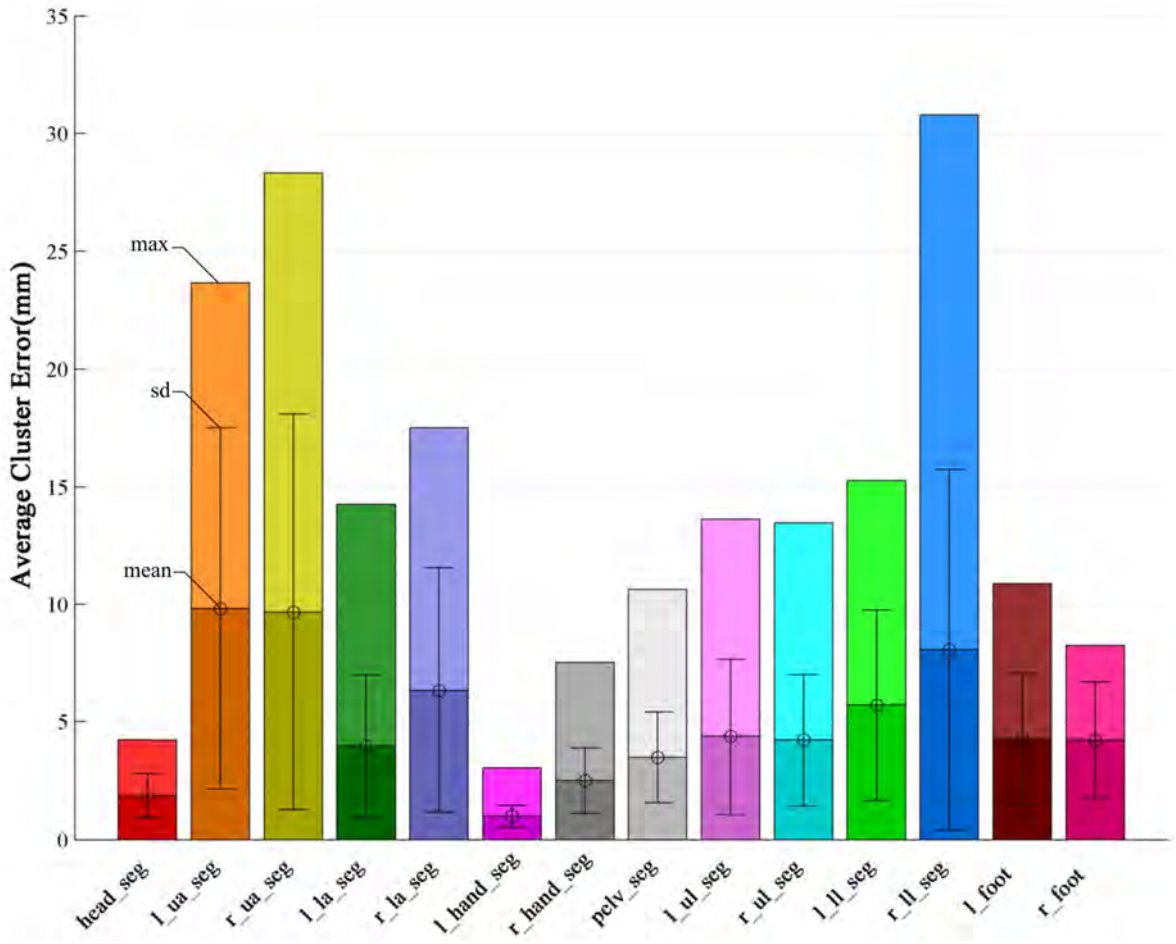


Figure 13.6 – Average of reconstruction errors

Through visual inspection of animated sequences rendered from the reconstructed trajectory data, correct reproduction of motion was verified. Figure 13.7 shows a multiple exposure plot indicating the progression of subject motion over a section of the reconstructed sequence. Although little quantitative information can be gathered from this figure, it illustrates the preservation of trajectory continuity and nuance of motion.

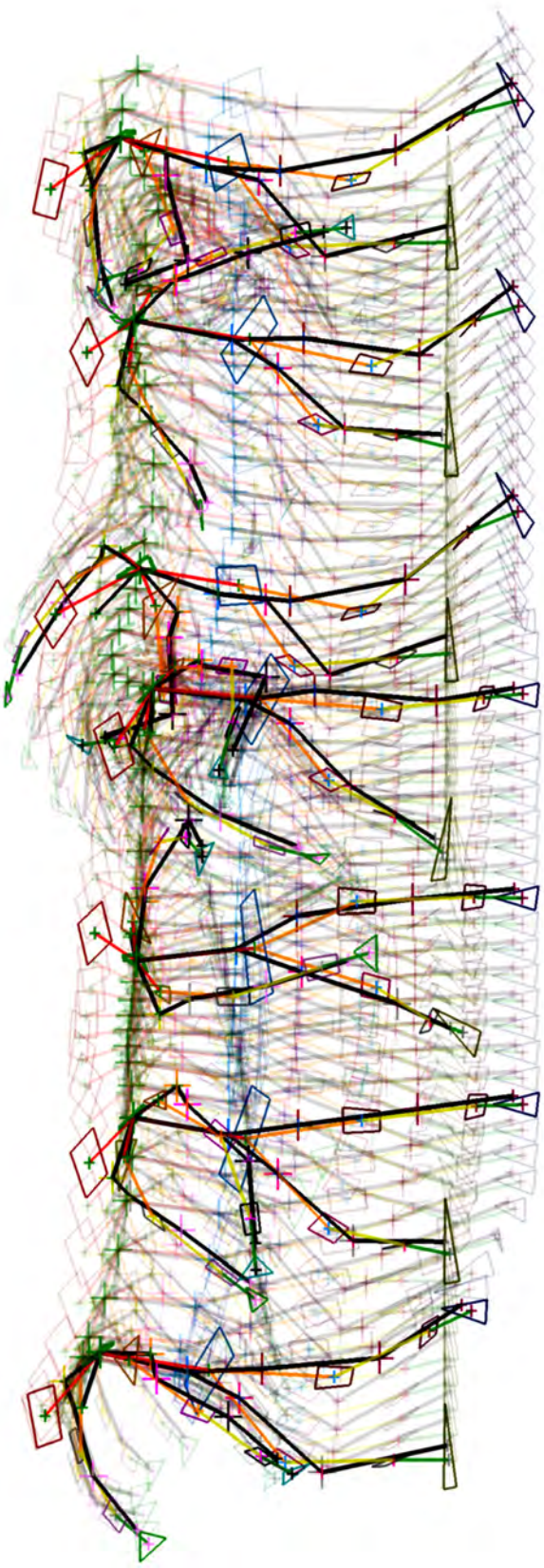


Figure 13.7 – Subject motion during reconstruction

13.3 Final reconstruction of primary data set

In this section the reconstructor is applied to the full data set. Due to the multiple occlusion models and sources of inertial data this testing was conducted over several different operational conditions using both simulated and measured data. As consistent performance of the reconstructor is demonstrated across these conditions, reliability of the method is confirmed. Finally, the reconstruction accuracy is compared to an alternative means of occlusion recovery and the advantages of the two techniques discussed.

13.3.1 Comparison of reconstruction between statistical and volume model

The primary data set is separated into two sets of simulated occluded trajectory data, one generated by application of the statistical model (20 sequences) and the other via the volume occlusion method (8 sequences). The two sets were processed as in (13.2.3), using simulated inertial data. It was found that 100% trajectory recovery was achieved in all cases. Assessment of reconstruction accuracy was again performed by calculation of distance between corresponding trajectories. These results were sorted by cluster and their distribution is shown in figure 13.8 for volume and statistical occlusion simulations.

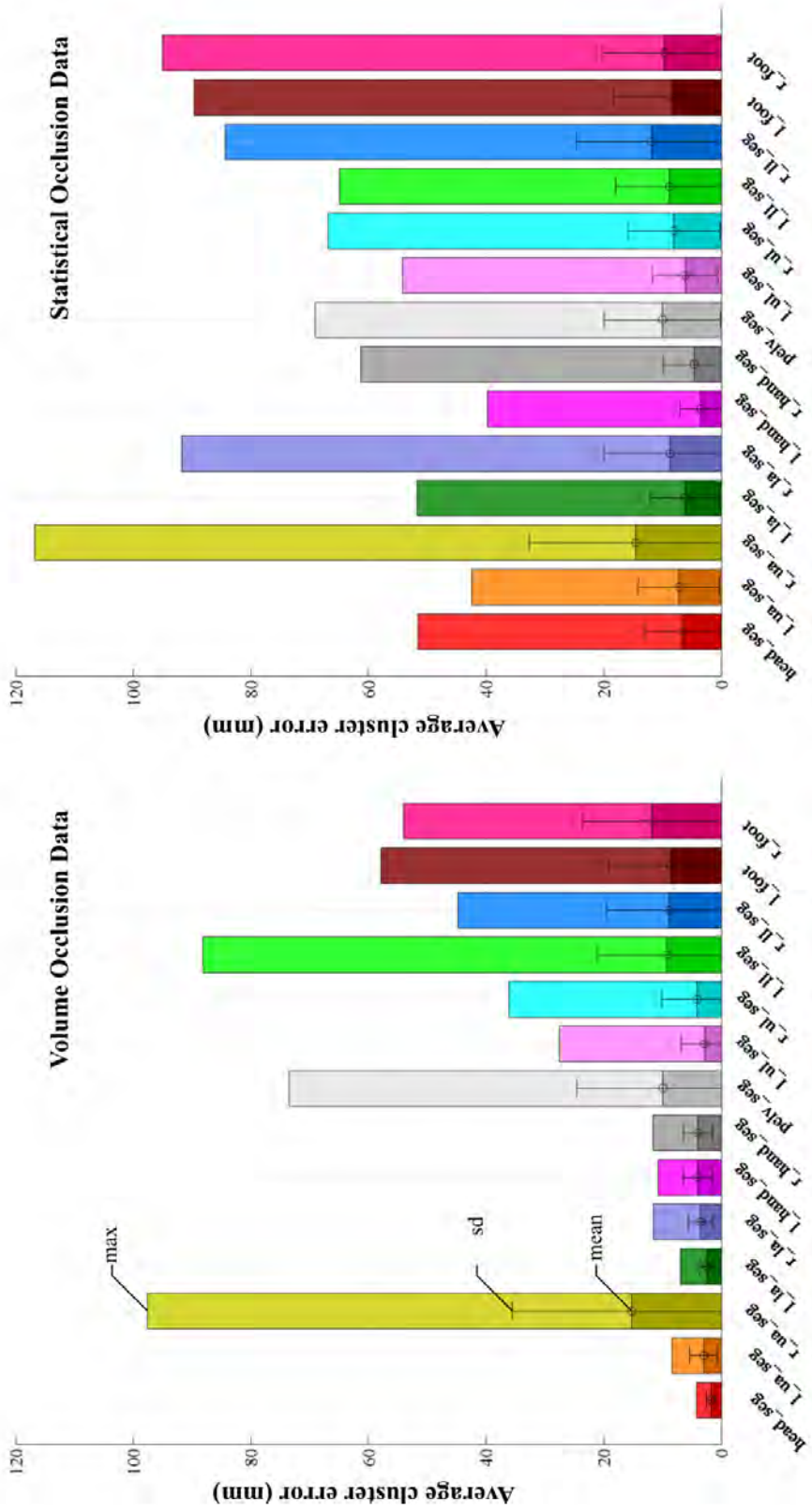


Figure 13.8 – Reconstruction performance by occlusion simulation method

Comparing these two figures, it is noted that error levels for the first six clusters in the volume model, excluding ‘r_ua_seg’, are considerably lower than the equivalent values given by the statistical method. This variation is explained by a bias in occlusion frequency produced by the volume method. In figure 13.9 the ratio of frames reconstructed to the total number of frames processed is plotted for each method. It is shown that the statistical model results in an even distribution of occlusion relative to the volume method. This is to be expected as unlike the statistical method where probability of occlusion is constant, the volume model is dependant on the physical locations of markers and cameras. Therefore, in some locations occlusion events are shorter and less frequent leading to reduced extrapolation error. This is confirmed by the correspondence between position error and proportion of occlusion. This inconsistency makes the volume method an inefficient means of testing the reconstructor. Therefore, all further analysis will be conducted upon data obtained from the statistical model, unless otherwise stated.

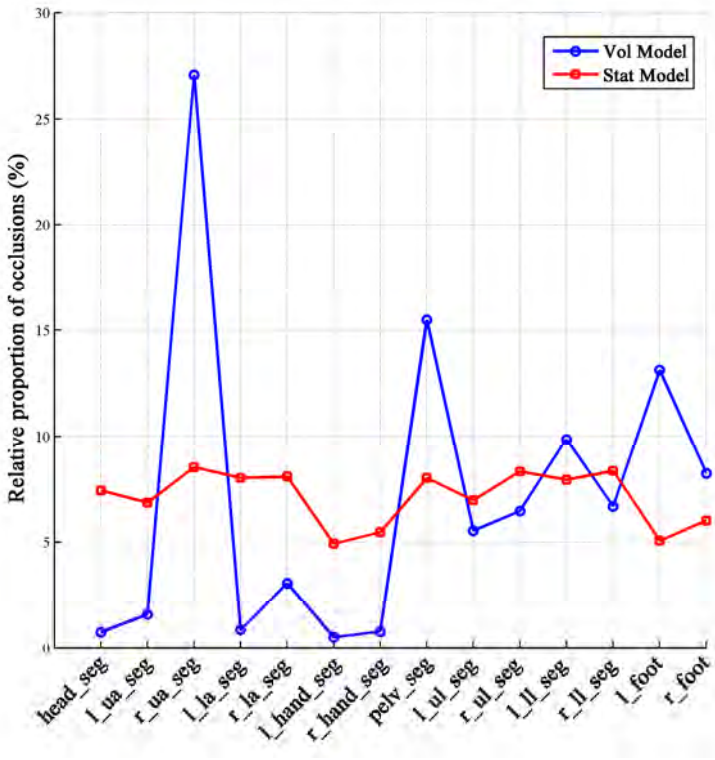


Figure 13.9 – Relative proportions of occlusion for volume and statistical models

13.3.2 Reconstruction error against occlusion duration

As can be seen from the results of (11.1.5), duration of occlusion has an influence on the scale of reconstruction error. Figure 13.10 shows a scatter plot of mean reconstruction error against occlusion duration. Each point represents the mean average error over a single occlusion event and its colour signifies the source cluster. The blue line represents a 100 sample binned average for the range of values on each side of the marker, with standard deviation indicated by the error bars.

This series shows a trend of increasing average error with increasing duration, as well as an increase in the spread of error values. This result stems from the assumption made by the reconstructor of a linear increase in orientation error with duration of extrapolation (11.1.5). However, the random walk of gyroscope bias causes deviation from this approximation and increases uncertainty in estimated extrapolation error over time. This makes definition of an accurate model for single case errors impossible without precise knowledge of the gyro bias. This situation is compounded by similar difficulties in assessing accuracy of joint centres in the body model.

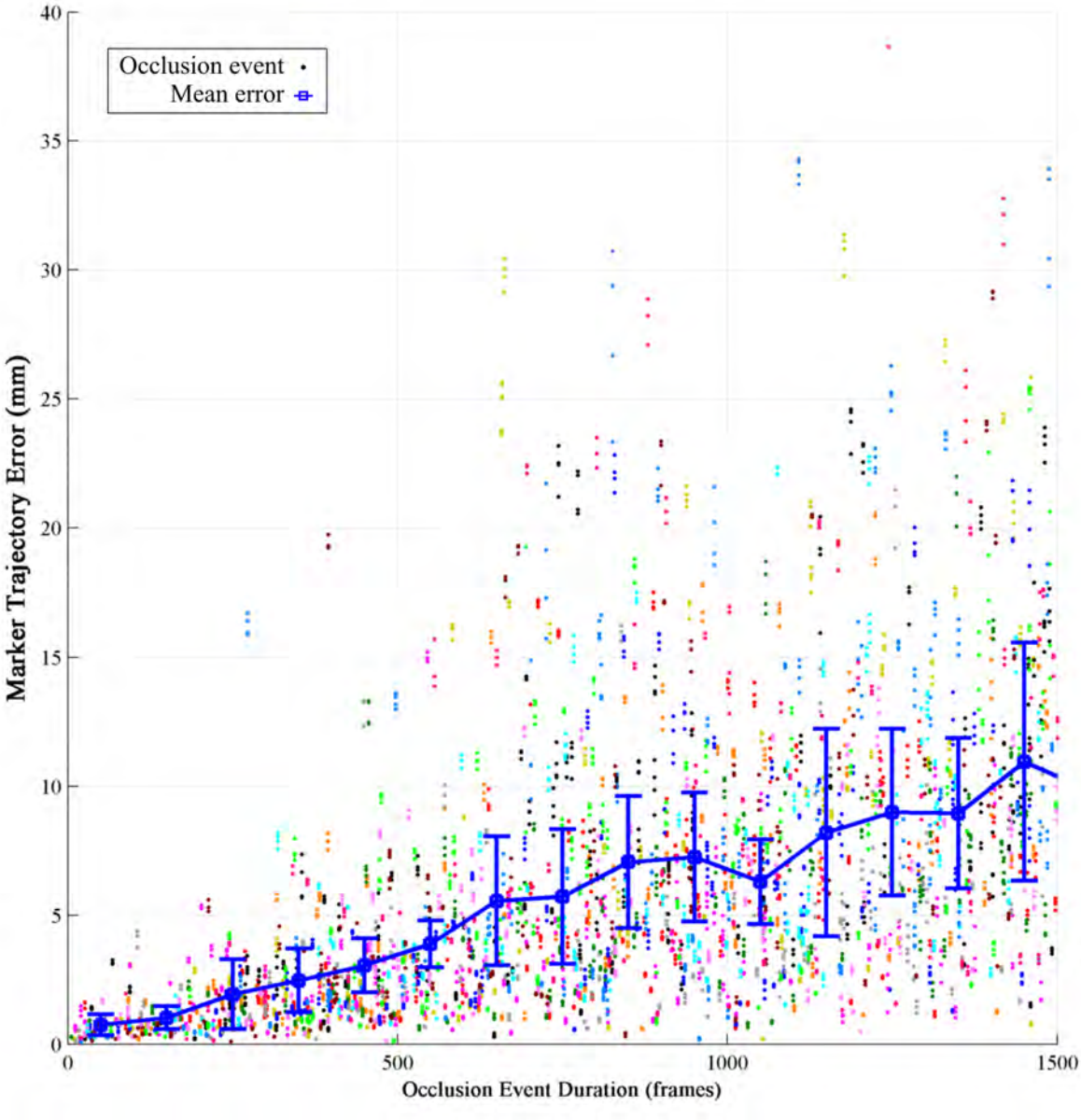


Figure 13.10 – Trajectory error against occlusion duration

13.3.3 Overlap of neighbouring segment occlusions

In comparison with previous tests of the reconstructor (12.2.3 and 11.2.3) a considerable increase in maximum error is encountered for the reconstruction of the primary data set (100*mm* vs 30*mm*). The cause of this increase is the propagation of reconstruction errors which occur during simultaneous occlusion of neighbouring clusters (12.1.4). This is

confirmed in figure 13.11, where occlusion events are segregated based upon whether they occur in isolation (red) or simultaneously with a proximal neighbouring cluster (green). Clusters with 100% visibility and their neighbours are excluded (c7 and shoulders).

Note that the peak error for the proximal condition is more than twice the single occlusion error, confirming the sensitivity of reconstruction to CoR perturbation. As the occurrence of proximal occlusions is dependant upon the order in which the kinematic hierarchy is traversed, this error may be attributed to the fixed order used in testing. Although a more flexible traversal method could significantly reduce these errors (14.3.1), increased prevalence of occlusion will eventually make proximal occlusions unavoidable. This represents an inherent weakness in the inertial reconstruction method.

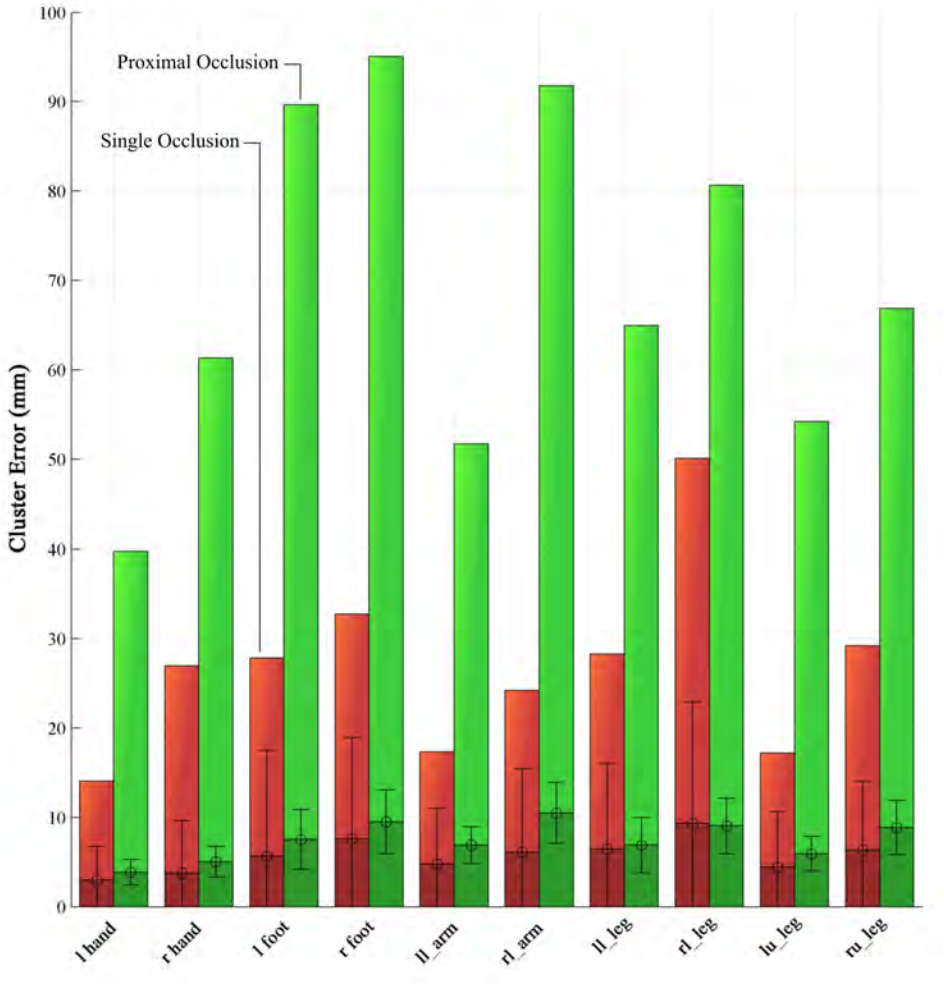


Figure 13.11 – Comparison of error for occlusions reconstructed using CoRs derived from measured cluster trajectories (single) and trajectories recovered by reconstructor (proximal).

13.3.4 Reconstruction performance of real vs. simulated inertial data

To validate the sensor model in the context of full body trajectory data, it is necessary to compare the simulated reconstruction results to those obtained from actual inertial measurements. As inertial measurements were only collected from a single cluster this comparison is limited to the trajectory of the right lower leg (rl_leg). However, given the stability of optical measurements of other non-instrumented clusters, it is assumed that similar performance would be obtained in other locations.

Four simulated occlusion sequences were generated using the statistical occlusion model over the four motion sequences of the primary experiment. This resulted in 17 occlusion events over 4700 frames. After application of the sensor model these sequences were duplicated and inertial measurements were substituted to yield a second dataset. Both sets of data were processed via the reconstructor, and the trajectory error magnitude was calculated against the gold standard data (Figure 13.12).

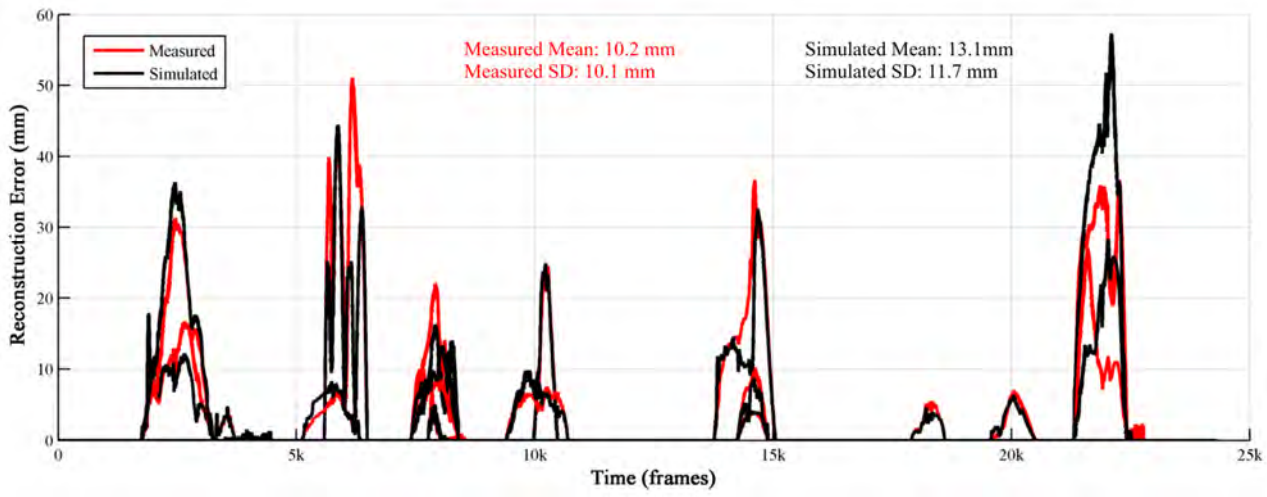


Figure 13.12 – Error magnitude for measured and simulated inertial data

Each occlusion event gives comparable levels of error across datasets indicating similar reconstruction accuracy for both simulation and measurement. To assess similarity of trajectory estimates the two data sets were compared directly. This was achieved by calculation of the average Euclidean distance between trajectories and the angle between cluster orientation estimates. The distribution of these distances and angles is shown in figure 13.13.

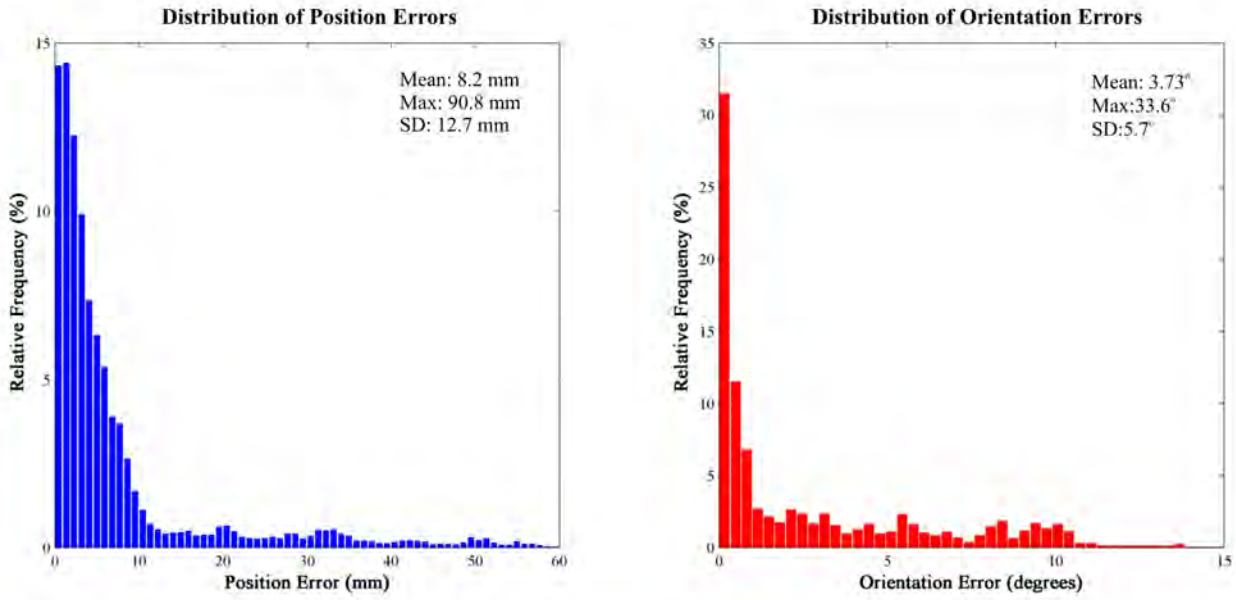


Figure 13.13 – Trajectory error against occlusion duration

The mean difference between the two sets of trajectory estimates lies around the predetermined limits for acceptable reconstruction error (4.1.4). However, as values of gyroscope bias and noise will inherently be different for measured and simulated data identical results were not expected. Therefore, this result constitutes validation of the sensor model.

13.3.5 Comparison with spline reconstruction

To put the performance of the reconstructor in context, a comparison with an alternative means of trajectory reconstruction was performed. This comparison was conducted against the method of spline interpolation typically implemented in commercial motion capture software [7,8].

The spline interpolation was applied to the fragmented primary data set, across all motion sequences and upon each occlusion event in turn. The fitting function used was the internal Matlab cubic spline interpolation [86]. Control points for the start and end of the curve were defined via sections of trajectory data 200 samples in length. As the spline interpolation provides no capability for tracking across occlusions, identities of marker fragments were obtained from the gold standard data set. In figure 13.14 a typical spline interpolated trajectory is shown by the thick coloured lines with the equivalent inertially assisted reconstruction plotted in black. Duration of occlusion events are indicated by the shaded boxes. Although smooth trajectory continuations are obtained across each occlusion, large differences between the two reconstructions are noted. Measured trajectory data is omitted from this plot as it closely resembles the reconstructor data at this scale.

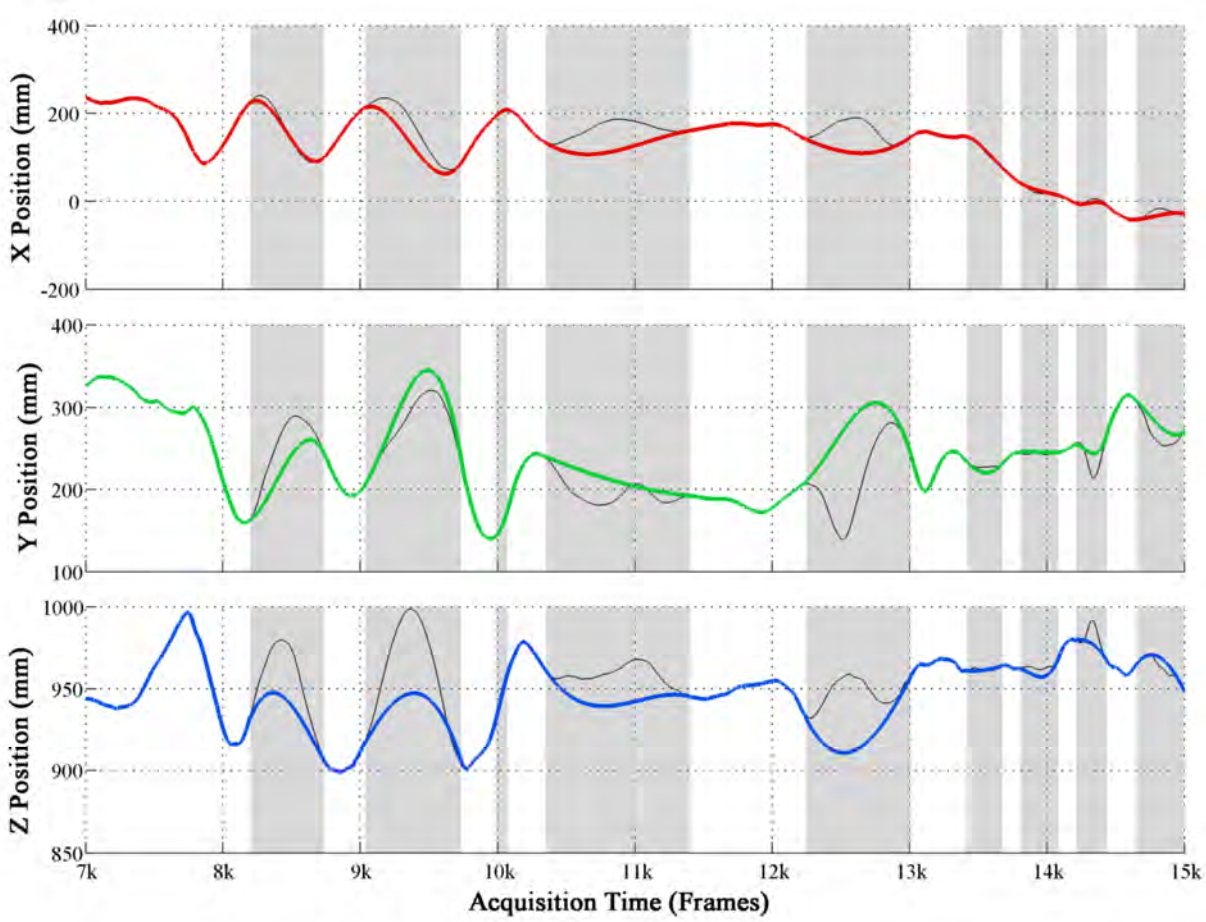


Figure 13.14 – Typical spline interpolation

The scatter plot in figure 13.15 shows each occlusion event as a single point. Error against the gold standard data is plotted upon each axis. The colour of each point indicates length of the occlusion event, with the range 1-2000 frames signified by colours green to red. The threshold of 1:1 equivalence of error is indicated by the black diagonal line.

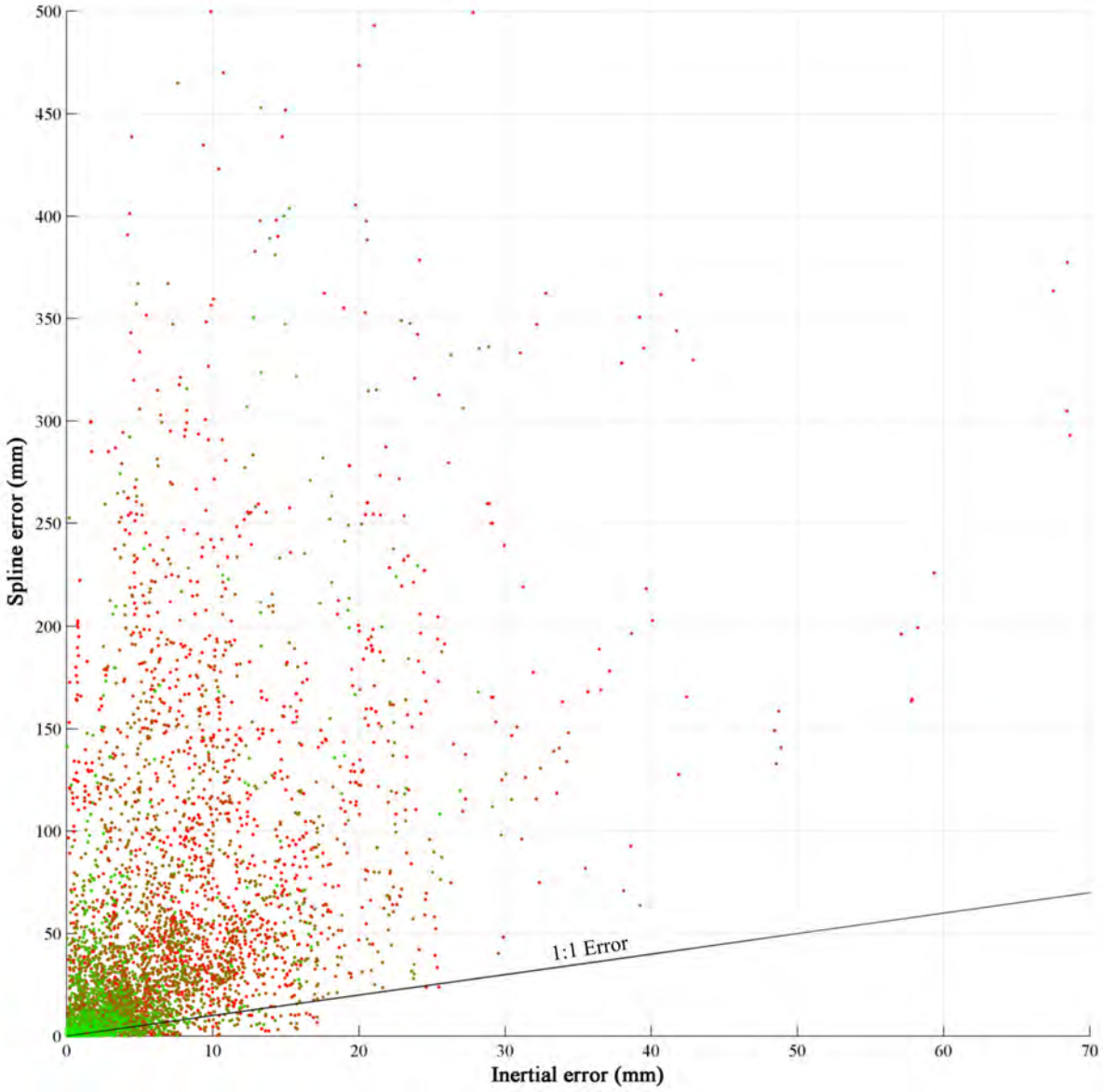


Figure 13.15 – Comparison of reconstructor errors

In 87% of cases the inertially assisted method outperforms spline interpolation in terms of accuracy. Mean and maximum values are found to be 223*mm* and 738 *mm* for the spline filled data (against 7.9*mm* mean and 116*mm* maximum for the inertial reconstructor). Furthermore, where spline interpolation yields superior accuracy, the difference in performance is slight (<=10*mm*) with error for both methods less than 20*mm*. Finally, the inertial method is more stable, having reduced variability of error with increasing occlusion duration.

13.3.6 Overall performance

For the final analysis of reconstruction performance, all trajectories for both occlusion models and real and simulated inertial data are combined to produce a single dataset. This set comprises a total of 7818 discrete marker occlusion events over approximately 3.86M frames. Position and orientation errors are calculated against the gold standard data by the usual methods. The complete distribution of these errors ordered by cluster, is shown by the radial histogram in figure 13.16. The outer ring corresponds to relative frequency of position errors and the inner ring orientation errors. This method of visualising the data was chosen as it allowed for the comparison of the complete set of distributions upon a single axis, avoiding overlaid data or multiple sub-axes. Figures for mean, maximum and standard deviation of error are shown in table 13.1.

Cluster	Mean Pos(*mm*)	Pos SD (*mm*)	Max Pos(*mm*)	Mean Orient (deg)	Orient SD (deg)	Max Orient (deg)
head_seg	6.4	6.6	51.6	0.2	0.3	3.3
l_ua_seg	6.9	6.9	42.5	0.7	1.5	15.7
r_ua_seg	13.1	17.9	116.8	1.2	3.7	50.8
l_la_seg	6.0	5.7	51.7	0.5	0.7	9.9
r_la_seg	8.7	11.5	91.8	0.8	2.3	28.0
l_hand_seg	3.5	3.6	39.8	0.5	0.7	12.7
r_hand_seg	4.6	5.2	61.3	0.5	1.0	34.2
pelv_seg	8.7	8.3	56.9	0.5	1.3	20.1
l_ul_seg	5.6	5.5	54.2	0.5	0.9	11.9
r_ul_seg	7.5	7.2	48.3	1.1	3.7	159.3
l_ll_seg	8.4	9.1	59.7	1.6	3.7	68.3
r_ll_seg	10.5	11.7	84.4	1.6	3.3	118.5
l_foot	8.2	10.0	89.7	1.1	3.6	90.0
r_foot	8.8	9.1	56.8	1.6	3.3	89.7

Table 13.1 – Mean, standard deviation and maximum of pose errors for primary data set

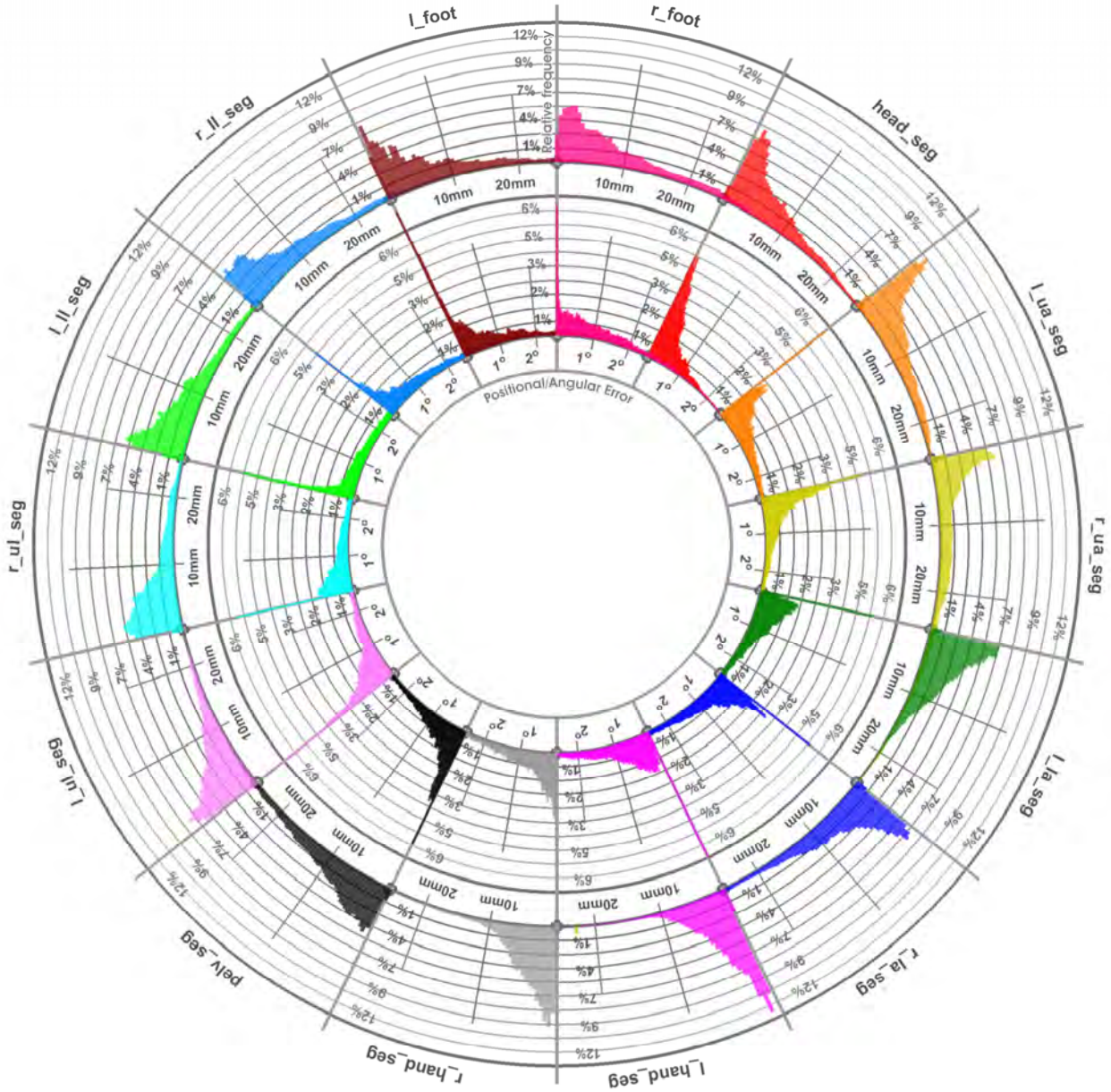


Figure 13.16 – Distribution of pose errors for primary data set

The grand average values for the entire data set are found to be 7.9 *mm* (SD 9.7 *mm*) for position error and 0.9^{o} (SD 17.9^{o}) for orientation error. Therefore, the average performance lies within the desired 10*mm* range for positional error and 3^{o} limit for orientation, satisfying the requirements established in (4.1.4). Several clusters present unusually high values of maximum orientation error, which are attributed to the correction of extrapolation error

breaking rigid body correspondence (11.2.3). However, this condition occurs rarely and is easily detected via geometric correspondence methods like those described in (12.2.1).

13.4 Summary

In this chapter the single cluster trajectory extrapolation was extended to provide a means of reconstructing full body trajectory data.

The trajectory reconstructor was tested using 8 sequences of simulated occlusion events produced by the volume model, and 20 sequences produced by the statistical model. Inertial sensor data was also provided by simulation, using the sensor model (9.2). The volume model provided uneven levels of occlusion across the dataset and was therefore eliminated from further testing.

Occlusion duration and the occlusion of neighbouring body segments were identified as primary sources of error in the reconstruction process. In the case of occlusion duration, an approximately linear relationship between duration, mean error and error standard deviation was observed. For multiple neighbouring occlusions, accuracy of the reconstructor was compromised, with a peak error more than twice the value found in isolated occlusions. However, over all tested conditions the mean error provided by the reconstructor was within the proposed limits defined in (4.1.4).

To validate the performance of the inertial sensor model, IMU measurements were substituted for the simulated inertial data. Good correspondence was observed between measured and simulated results with a mean distance between cluster trajectories of ~8*mm* (SD 12*mm*) and a mean orientation difference of ~4^{o}(SD 6^{o}).

The accuracy of the reconstructor was compared to that obtained by spline interpolation of identical trajectory data. For 87% of cases the reconstructor gave improved results over spline interpolation. The mean error for spline interpolation was found to be ~220*mm* versus ~8*mm* for the reconstructor.

Finally, the errors for full body trajectory reconstruction were assessed for the entire dataset. The grand average error across all conditions was 7.9 *mm* (SD 9.7 *mm*) for position error and 0.9^{o} (SD 17.9^{o}) for orientation error. This lay below the desired 10*mm* limit for positional error and 3^{o} for orientation.

14. GENERAL DISCUSSION

This thesis has presented a method by which two human motion measurement techniques, inertial and optical motion capture, are combined. This composite method was shown to give superior performance to either method applied alone. In the first and second sections of this chapter, the project's achievements are outlined in relation to the primary and secondary goals. The third section discusses future directions of development and how the limitations of the method could be addressed. Finally, the forth section details the dissemination of work beyond this thesis.

14.1 Primary aims

The primary research question of this thesis was whether sources of inertial measurement can be used in tandem with passive optical motion capture data to obtain improvement beyond conventional measures in tracking and occlusion recovery of human motion data. To investigate this, an algorithm to combine these two measurement modalities was developed and tested. This section discusses the effectiveness of this algorithm and its limitations.

14.1.1 Reconstruction Accuracy

It has been demonstrated that inertially assisted optical motion tracking results in a great improvement in occlusion resistance, working accurately and reliably over a variety of occlusion conditions. Unlike previous studies (discussed in 2.3.3), which have either relied solely upon simulation or conducted no formal testing of accuracy [11, 15] its performance for experimentally derived data is clearly demonstrated. In the final analysis, error in reconstruction was reduced to a third of that obtained by spline reconstruction. It was also

found to satisfy the criteria established in (4.1.4) and (7.2.2) for average accuracy of 10-30*mm* for occlusion gaps up to 2000 frames (4 secs). Events of up to 6000 samples (12 secs) were also reconstructed but with increased error. Although this error is at the limit of what is considered acceptable for biomechanical analysis [53], it is appropriate in applications where precise anatomical detail is of reduced importance, such as in computer animation [87].

Under simultaneous occlusion of neighbouring clusters, the accuracy of the method is compromised (13.3.3). Although this is an unavoidable occurrence in real measurement, its effects are exaggerated by the fixed traversal of the body model hierarchy (13.1.4). An alternative means, by which this problem could be addressed is through the adaptive traversal of the hierarchy described in (14.3.1). However, for the purposes of demonstrating the basic capability of the reconstructor, the fixed method was preferred due to its relative reliability and efficiency in comparison to the adaptive method.

14.1.2 Tracking Performance

In relation to tracking performance, excellent results were obtained, with no failures observed over the entire primary data set with ~30% marker loss and fixed configuration of the reconstructor (13.2.2). This represents a great improvement over standard tracking techniques, which typically fail beyond a few frames of occlusion (2.1.3).

A unique advantage of the method is its ability to reliably recover marker identity after long term occlusion events (several seconds) regardless of trajectory. This is an advantage over techniques based upon probabilistic or kinematic methods, where reliability of tracking

deteriorates considerably with increasing time, sudden changes in velocity or occlusion of neighbouring markers [10, 15, 16].

The dependence of the reconstructor upon locally linear approximations of non-linear errors indicates a limit to its stability. However, the point of tracking failure lies beyond the ranges tested experimentally due to the excellent performance observed with the primary data set. Although a larger data set could determine this point, such an investigation would go beyond the scope of this project. The aim of this project was to demonstrate the feasibility of the technique, rather than perfect its application. However, a number of improvements, which could improve stability of tracking are discussed in section (14.3).

A limitation of the reconstructor is the necessity for all marker fragments comprising a cluster, to be visible before it is considered to have emerged from occlusion. Therefore, trajectory fragments, which do not start or end in a visible cluster will be abandoned by the tracking algorithm. This represents an inefficient use of available data and could contribute to large reconstruction errors, where occlusion is extremely prevalent. However, even in cases where trajectory identification is performed manually, isolated trajectories of this type can also prove unclassifiable.

14.1.3 Practical implementation

The aim of this thesis was to devise a novel reconstruction method and demonstrate its capability. In addition to this aim being met, a practical solution applicable to real measurements has been implemented. Movement data can be automatically processed from raw co-ordinate data to reconstructed trajectories. However, in its current form the method is

computationally inefficient, dependent upon calibration of the body model by gold standard data and manual initial marker identification. Although this limits the usefulness of the method, potential solutions to these issues are discussed in section (14.3).

Another limitation of the method's practicality is its high complexity for the moderate performance advantage it offers in conventional measurement scenarios. In addition to the extensive data processing and calibration requirements, the physical requirement for 18 IMUs is of particular concern. Although this limitation makes conventional methods of addressing occlusion appear more desirable (2.2.3), in some measurement scenarios such strategies are unfeasible or ineffective. Examples of such situations are, where only limited camera coverage is possible, the cost of additional cameras is prohibitive, or other objects and people clutter the capture volume. The latter is particularly pertinent, as ensemble motion capture is a common requirement in computer animation. In these conditions, the advantages presented by inertially assisted optical motion capture could far outweigh the inconveniences it incurs. Additionally, as MEMS sensor technology continually improves in performance due to intense commercial interest, a more compact inertial sensing setup could be developed (14.3.4).

As the reconstructor depends upon future trajectory information for correction of extrapolation errors, the method is restricted to processing of previously acquired data. As well as preventing recovery of occlusions which overlap the start or end of a measurement sequence, this limitation also eliminates the possibility of real-time operation. As the effects of occlusion, are particularly undesirable in real-time applications of motion tracking, for example Augmented/Virtual Reality [81], this represents a missed opportunity for the method.

14.2 Secondary Aims

The development of body, sensor and occlusion models constituted the secondary aims of the project. The capabilities and limitations of each of these are discussed below.

14.2.1 Body Model

As actual CoR locations and limb lengths cannot be measured directly the performance of the body model is difficult to assess. However, its reliability may be inferred from the analysis of CoR estimation given in (6.3.3), as well as the satisfactory performance of the reconstructor in the final analysis. Additionally, the comparable results obtained for both pilot and primary datasets indicate the method is robust against variation across capture sessions. The inability to reconstruct complex joints such as the shoulder represents a failing of the model. However, this is the result of the required simplification of the body model (4.1), and performance could be improved through substitution of more sophisticated joint models (14.3.2).

Dependence upon the gold standard calibration trial is a major limitation of the current implementation. For preliminary testing of the inertial reconstruction method, this is acceptable. However, in a practical measurement scenario, this requirement would be inconvenient, if not impossible, due to the occlusion or limitations of the subject's movement (i.e. in medical settings). Although it is possible that the calibration process could be integrated with the operation of the reconstructor (14.3.3), it cannot be dispensed with entirely. Therefore, some degree of external assistance in marker identification and regimented subject movement is unavoidable.

14.2.2 Inertial measurement system and sensor model

The development of the IMU system and associated model represents a highly successful phase of this project. The hardware was shown to facilitate measurement of body motion using a state-of-the-art inertial sensor. The choice of serial digital communication rather than conventional analogue interfacing provided excellent immunity to noise and signal degradation during sensor operation. Additionally, Ethernet/TCP provided an extremely effective means of communication with the host PC allowing for scalability, platform independence and remote operation in future applications. The restriction of tethered operation was noted to be a problem due to the tangling hazard it presented. However, wireless operation was considered an unnecessary complication to the design (14.3.5). With regard to precision of the actual sensing device, the data given in (13.3.4) suggests adequate performance, approximately equivalent to simulated results. Although the gyroscope dynamic range was less than the established requirement (8.1.2), and superior bias stability would have reduced orientation drift, these were limited by the currently available technology.

Validation of the sensor model was performed by comparison of simulated and measured inertial data (9.2.4). In these tests, although good correspondence was found between simulated and measured gyroscope data, the effect of dynamic acceleration was not adequately predicted by the model. However, as acceleration is not utilised beyond the initial calibration of IMU-optical alignment (9.1.1) this was not deemed to be problem. The spatially similar trajectory data produced by measured and simulated inertial data, confirmed this performance (13.3.4).

The filtering applied to suppress differentiation noise results in poor frequency response when compared to real inertial measurements. However, due to the slow and controlled motions of the primary dataset this has minimal effect on the performance of the reconstructor. Another limitation of the model is the sequence dependant errors introduced by the Euler angle decomposition at high rates of turn. Although this is also addressed by the relatively low velocities of the test data, a more stable means of estimating axial components of rotation would be desirable.

14.2.3 Controlled degradation of passive optical motion capture data

The least successful aspect of this project has been the implementation of the volume occlusion model. This was due to the assumption that occlusion is solely caused by the gross obstruction of markers by body segments (7.3). The impact of other factors such as distance/angle of markers, marker reflectivity and movement of clothing was underestimated. As these effects were difficult to quantify, a deterministic approach to modelling occlusion would inherently bear little resemblance to actual observations. The poor computational efficiency of the volume model further limited its usefulness. Additionally, occlusions in the simulated data were heavily biased toward visibility of certain clusters (13.3.1). As a result, insufficient occlusion data was available to test the reconstructor.

Although simulated data obtained via the statistical model does not resemble that obtained experimentally (7.3), it presents a far more useful distribution of occlusion for the purposes of testing the reconstructor. Additionally, due to the simplicity of its implementation, the statistical model is much more efficient in terms of computational effort. A possible extension of the occlusion modelling method would involve using the output of the volume model in

combination with the statistical method. This would allow probability of occlusion to be modulated in terms of the body and camera pose. Although this would result in a spatially more accurate occlusion model, the problem of cluster bias would still exist. Additionally, the efficiency of the method would be further reduced by the increased complexity. These factors made further development of the combined method unattractive.

14.3 Further Work

This thesis has thoroughly explored the potential of inertially assisted optical motion tracking for the recovery of occluded trajectories. However, during the course of the project several alternative avenues of investigation were encountered, which were disregarded to avoid detracting from the primary aims of the project. This section aims to briefly discuss the potential of exploring these directions, giving an indication how work could progress beyond this thesis:

14.3.1 Adaptive kinematic traversal

A more sophisticated method of traversing the kinematic hierarchy would reduce errors due to neighbouring occlusions and eliminate the requirement for a continuously visible 'root' cluster (13.1.4). Such an alternative traversal could be performed by dynamically changing the position of the root to reflect the nearest visible cluster to an occluded target cluster. By using trajectory data from neighbouring clusters to define the CoRs required for reconstruction, propagation of extrapolation errors could be minimised. This adaptation was intended to be included in the final implementation of the reconstructor. However, the fixed root method was found to predictably demonstrate the reliability and accuracy of the reconstructor in line with the aims of the thesis.

14.3.2 Improvement in body model accuracy

The simplified kinematic model used in this thesis leads to inaccuracy in trajectory reconstruction and limits the reliability of tracking. This is a particular problem in the case of complex joints, such as the shoulders, where the model is inadequate for trajectory recovery and tracking (6.5.4). An obvious solution to this problem would be the substitution of the

simplified model with a more anatomically accurate one. Examples of such models are the shoulder and spine models given in [71] and [73]. An alternative option would be to supplant the entire body model with a method not based upon locally derived solutions, such as a global optimisation method [47,88], which may give improved results for limited ranges of motion. Another possibility would be to adopt a proprietary body model, such as those provided by ‘Anybody’ [2] or ‘Visual 3D’ [89]. However, given these models are typically “closed source” it may not be possible to perform the modifications necessary for use with occluded data.

14.3.3 Improved analysis of IMU measurements

As the simple linear correction of orientation drift gave adequate results, alternative correction methods have received little consideration (10.1.3). However, by utilising an estimator such as the Kalman filter [16, 19, 41], it would be possible to reduce long term bias drift. This reduction would allow for more reliable estimation of orientation and hence, improved accuracy of trajectory reconstruction. Furthermore, as such methods inherently provide a measure of confidence, expected trajectory error may also be estimated. Extending this concept, it may be possible to implement the whole reconstruction operation as a single probabilistic filter. A solution of this type is presented in [15], although its application to inertially assisted tracking is only suggested.

Errors introduced by bias and the effect of gravity on measurement of translational acceleration make long term dead reckoning of position via accelerometry unfeasible [84]. However, short term estimates of position could be useful in the calibration of the body model. An example of this would be the estimation of a short section of the trajectory of a

body segment's orbit. By applying the CoR estimation algorithm to several of these sections, the body segment lengths and centres could be extracted. Although such a method could prove extremely useful in both optically assisted and conventional inertial motion tracking, the potential levels of error involved make its implementation extremely challenging.

Finally, inertial measurements could automate the initial identification of unlabeled marker clusters, eliminating the need for manual initialisation of the reconstructor. Through modification of the trajectory classifier, the correlation between trajectory hypotheses and inertial measurements could associate groups of markers with individual sensors (12.3.2). As the locations of inertial sensors upon the body would be predefined, this association could be used to indirectly identify marker clusters. Furthermore, the exaggerated range of movement required by the initial model calibration would provide ideal conditions for application of this method.

14.3.4 Expansion of IMU system

The development of a full body inertial sensing system was considered beyond the scope of this thesis. However, since the primary experiment was completed the size and cost of MEMS sensor technology has become much more favourable. Combined with similar improvements in wireless communication and battery technology, the implementation of a full body inertial measurement system is now more feasible [90, 91]. It is the author's opinion that using currently available technology, a wireless inertial sensor with a form factor similar to the passive markers of the optical measurement system, could be implemented. This would make inertially assisted tracking no more cumbersome than conventional optical methods, greatly increasing the practically of the techniques developed in this thesis.

14.3.5 Further testing

The restriction to a single measurement session and subject makes assessment of robustness against variation in marker layout impossible. The body model calibration was successfully applied to both pilot and primary datasets. However, to identify the limitations of the method as a whole, further measurement sessions and subjects would be necessary. The requirement for a larger data set has already been identified in the discussion of tracking failure (14.1.2) and is therefore a priority for continued investigation.

The majority of clusters in the full body marker set exhibit redundancy through the use of four markers rather than three (4.3). Although this improved accuracy of orientation estimation in some cases, these markers were found to quantify redundant or constrained degrees of freedom. By removing or repositioning these markers the subject setup could be simplified and load upon the measurement system reduced.

14.3.6 Augmented measurement bandwidth

A potential extension of the inertially assisted tracking method is the synthesis of measurement bandwidth. The superior frequency response of inertial sensors allows inertially reconstructed data to preserve details of trajectories beyond the measurement range of optical motion capture. Therefore, by combining optical and inertial measurements, total measurement bandwidth could be increased while simultaneously preserving spatial accuracy. This method would be desirable in orthopaedic and sporting applications as it would allow for improved estimation of dynamic loadings upon joints, particularly during impacts [5]. However, as the simulated inertial data produced by the sensor model has the same bandwidth

restrictions as the optical system, a full body inertial measurement system would be necessary to further develop this technique.

14.4 Dissemination

The methods and results discussed in this thesis have been disseminated to the wider academic community via the following means:

Cardiff University, March 2008:

A talk was given describing the kinematic modelling processes and basis for the volume model at the Cardiff School of Engineering, University of Cardiff. UK

Munster University, April 2008:

A presentation was given describing advanced applications of optical motion capture for the Qualisys User group meeting at the Institute of Sports Science, WWU Munster, Munster, Germany.

3DMA, October 2008:

A conference paper entitled ‘A software framework for quantitative analysis of occlusion in optical motion capture’ was included in the proceedings of ‘3D motion analysis 08’ (3DMA) in Amsterdam, the Netherlands. A poster was given detailing the inertially assisted reconstruction (Appendix 7).

Midlands Sensorimotor Group meeting, November 2009:

A presentation discussing the preliminary results of trajectory reconstruction and related methods was given for the first Midlands Sensorimotor Group meeting at the School of Psychology, University of Birmingham, UK.

Paper on inertially assisted optical measurement, Pending:

Following submission of this thesis an overview of the method and results will be prepared as a journal paper in late 2011. Proposed target journal is ‘Computer Methods in Biomechanics and Biomedical Engineering’.

18, APPENDICES

Appendix 1: Matlab implementation of orientation estimation function:

MATLAB code listing on accompanying CD, path:

`cddrive:\matlab\sub_functions\quat_attitude.m`

Appendix 2: Matlab Implementation of CoR Algorithm and notes

MATLAB code listing on accompanying CD, path:

`cddrive:\matlab\sub_functions\multipoint_cor_1_2.m`

Appendix 3: Matlab Ray Polygon Intersection

MATLAB code listing on accompanying CD, path:

`cd_drive:\matlab\geometry\quad_intersect.m`

Appendix 4: Volume model construction code.

Code listing on accompanying CD, path:

`cddrive:\matlab\main_functions\extract_volume1_1.m`

Appendix 5: Implementation of Occlusion testing functions (including C)

MATLAB code listing on accompanying CD, path:

`cddrive:\matlab\main_functions\occlusion_test_breakdown\do_occlusion_tests6.m`

C MEX DLL code on accompanying CD, path:

`cddrive:\matlab\c\mex occ_proc_opt7.c`

Appendix 6: Final trajectory reconstruction and tracking algorithm.

MATLAB code listing on accompanying CD, path:

`cddrive:\Matlab\opto_imu_tracking\cor_hierarchy_tracking\cor_heirarchy_tracking_15_func_final.m`

Appendix 7: 3DMA Poster

Appendix 8: Euler rotations and gimbal lock

Three dimensional rotations can be represented by three axial rotations occurring clockwise about the x, y and z axes. By aeronautic convention [92] these rotations are individually referred to as pitch, roll and yaw (or α, β and γ) as shown in figure A8.1.

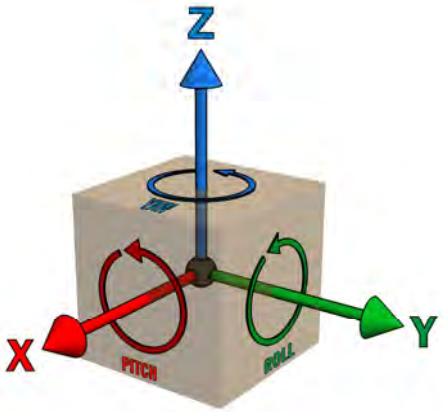


Figure A8.1 – 3D Rotational

With reference to the identity matrix the rotation matrix $\mathbf{R}_{\alpha,\beta,\gamma}$ corresponding to the sequential rotation by the three components is defined by the multiplication of the three matrices describing the transformation of two axes of the co-ordinate system by rotation about the third. As shown in figure A8.2.

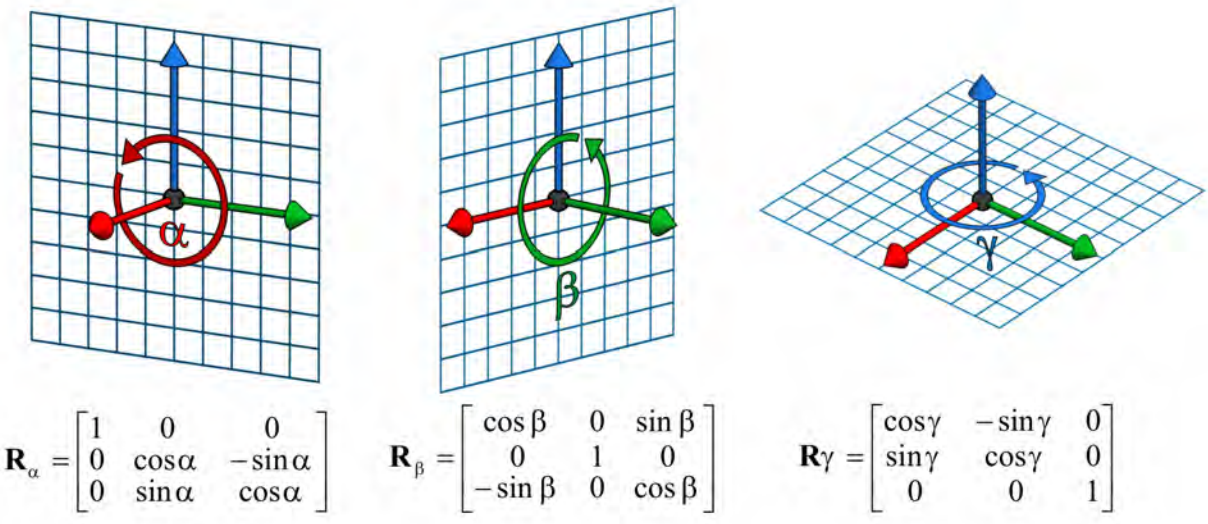


Figure A8.2 – Principle Euler Rotations

Combining yaw, pitch and roll transforms in order and multiplying out gives the expression for the rotation matrix $\mathbf{R}_{\alpha,\beta,\gamma}$:

$$\mathbf{R}_{\alpha,\beta,\gamma} = \mathbf{R}_{\alpha}\mathbf{R}_{\beta}\mathbf{R}_{\gamma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\beta\cos\gamma & \cos\gamma\sin\alpha\sin\beta - \cos\alpha\sin\gamma & \cos\alpha\cos\gamma\sin\beta + \sin\alpha\sin\gamma \\ \cos\beta\sin\gamma & \cos\alpha\cos\gamma + \sin\alpha\sin\beta\sin\gamma & \cos\alpha\sin\beta\sin\gamma - \cos\gamma\sin\alpha \\ -\sin\beta & \cos\beta\sin\alpha & \cos\alpha\sin\beta \end{bmatrix}$$

A drawback of the Euler system of representing rotation is 'gimbal lock' which result in the loss of a degree of rotational freedom. Considering the system of Euler rotations as representing the angles of the rings (or gimbals) of the mechanical gyroscope shown in figure A8.3. Gimbal lock occurs when a sequence of rotations is performed which results in the alignment of two of these rings. In the particular case of the gyroscope, this condition occurs when the angle of β reaches 90^{o}. At this point changing either of the other axes will effectively result in the same rotation rendering rotation in global z axis impossible.

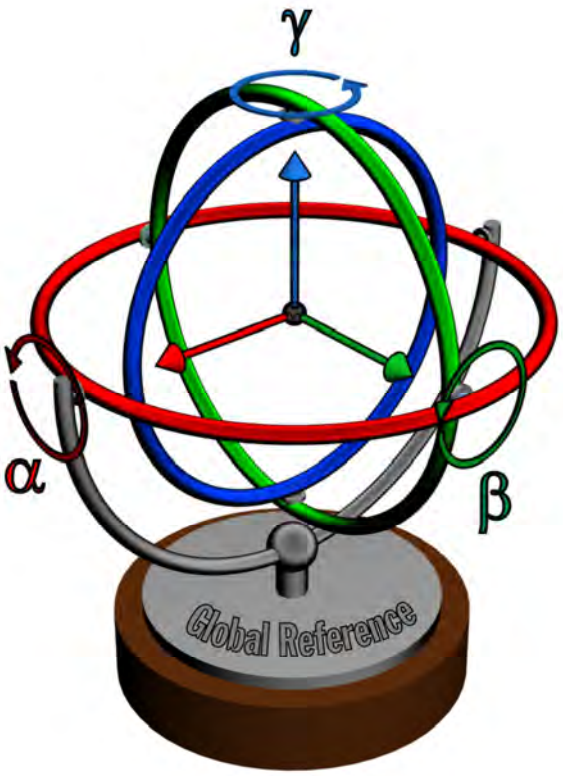


Figure A8.3 – Euler Gyroscope Analogy

Expressing this problem in term of the rotation matrix $\mathbf{R}_{\alpha,\beta,\gamma}$ and substituting $\cos\beta = 0$ and $\sin\beta = 1$ for $\beta = 90^{o}$ we obtain:

$$\mathbf{R}_{\alpha,\beta,\gamma} = \mathbf{R}_{\alpha}\mathbf{R}_{\beta}\mathbf{R}_{\gamma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}\begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \sin\alpha\cos\gamma - \cos\alpha\sin\gamma & \cos\alpha\cos\gamma + \sin\alpha\sin\gamma \\ 0 & \cos\alpha\cos\gamma + \sin\alpha\sin\gamma & \cos\alpha\sin\gamma - \sin\alpha\cos\gamma \\ -1 & 0 & \cos\alpha \end{bmatrix}$$

Finally, substituting appropriate trigonometric identities:

$$= \begin{bmatrix} 0 & \boldsymbol{\sin(\alpha - \gamma)} & \boldsymbol{\cos(\alpha - \gamma)} \\ 0 & \boldsymbol{\cos(\alpha - \gamma)} & \boldsymbol{\sin(\alpha - \gamma)} \\ -1 & 0 & \boldsymbol{\cos\alpha} \end{bmatrix}$$

This working shows an identical result to the mechanical analogy with rotation constrained about z-axis (first column of matrix) and equivalence between the α and γ terms.

Although the condition forbids any transform where β = +/-90^{o} the Euler system is still capable of representing all possible rotations through the non-commutativity of rotation. This means that changing the order in which rotations are multiplied leads to a different final transforms. Therefore, by changing the sequence in which the axial transforms are combined, the point at which gimbal lock occurs can be moved to a different axis. Given the three axial transforms about the x,y and z axes there are six possible permutations each with a different . In addition to these there is a second set of six sequences, designated Carden rotations, where a single axial transform is applied twice. Together, these two sets of sequences cover the entire space of 3D rotations and are shown in the table below:

Cardan	x-z-y, x-y-z, y-x-z, y-z-x, z-x-y, z-y-x
Euler	x-y-x, x-z-x, y-x-y, y-z-y, z-x-z, z-y-z

By switching sequences it is therefore possible to avoid the problem of gimbal lock. However, mixing several Euler/Cardan sequences in a single application can often result in confusion as well as difficulties relating corresponding transforms between each sequence. Therefore, sequences are selected dependent upon application to place the singularity away from the range of expected orientations. This is a popular strategy in biomechanics, where limited ranges of motion about joints can be exploited [48].

Appendix 9: Axis-angle representation of rotation

The system of axis angle representation allows the entire space of 3D rotation to be encoded as a single 3-vector. This is achieved by defining any change of orientation as a rotation about a vector (axis) by an angle proportional to the vectors magnitude. This is shown by the two transforms in figure A9.1.

This means of representation provides a compact and intuitive means of defining rotation, allowing for interpolation of rotation and greatly simplified calculation of the angle between two coordinate systems. However, unlike the rotation matrix, it cannot be applied directly to the calculation of rotational transforms. For this to be achieved it is first necessary to apply the Rodrigues' rotation formula [93]. This formula describes the rotation of a vector $\mathbf{p}_a$ about an axis defined by $\mathbf{v}_a$ by an angle θ.

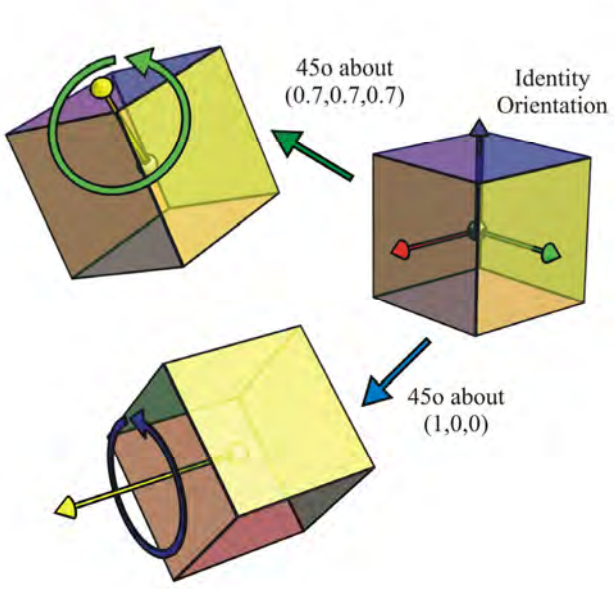


Figure A9.1 – Axis-Angle Transform

$$\mathbf{p'}_a = \mathbf{p}_a + (cos(\theta) - 1)((\mathbf{v}_a \times \mathbf{p}_a) \times \mathbf{v}_a) + sin(\theta)(\mathbf{v}_a \times \mathbf{p}_a)$$

By extension of this method it is possible to obtain the relation between the axis-angle representation and the rotation matrix with relative ease. By substituting the vector $\mathbf{p_a}$ with the Identity matrix **I** it is possible to provide the matrix $\mathbf{R_p}$ describing the rotation from the null transform achieved rotation by θ about $\mathbf{v_a}$:

$$\mathbf{R_p} = \mathbf{I} + (\cos(\theta) - 1)((\mathbf{v_a} \times \mathbf{I}) \times \mathbf{v_a}) + \sin(\theta)(\mathbf{v_a} \times \mathbf{I})$$

By substituting the cross product terms with their skew symmetric matrix form. The expression becomes:

$$\mathbf{R_p} = \mathbf{I} + (\cos(\theta) - 1)([[\mathbf{v_a}]_\times \mathbf{I}]_\times \mathbf{v_a}) + \sin(\theta)([\mathbf{v_a}]_\times \mathbf{I})$$

Where:

$$\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_\times \mathbf{b} = \begin{bmatrix} 0 & -a3 & a2 \\ a3 & 0 & -a1 \\ -a2 & a1 & 0 \end{bmatrix} \begin{bmatrix} b1 \\ b2 \\ b3 \end{bmatrix}$$

The double cross product term may now be rearranged by the use of the anti-commutative property of the cross product (ie: $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$) to give:

$$\mathbf{R_p} = \mathbf{I} + (1 - \cos(\theta))([\mathbf{v_a}]_\times [\mathbf{v_a}]_\times \mathbf{I}) + \sin(\theta)([\mathbf{v_a}]_\times \mathbf{I})$$

Which collecting terms and eliminating redundant identity matrices gives our final expression for the rotation matrix $\mathbf{R_p}$:

$$\mathbf{R_p} = \mathbf{I} + (1 - \boldsymbol{cos}(\theta))[\mathbf{v_a}]_\times^2 + \boldsymbol{sin}(\theta)[\mathbf{v_a}]_\times$$

Finally expanding to 3x3 rotation matrix with the scalar components of $\mathbf{v_a}$ given as v_x, v_y and v_z,:

$$\mathbf{R_p} = \begin{bmatrix} 1+(1-\cos(\theta))(v_x(v_x-1) & -v_z\sin(\theta)+(1-\cos(\theta))v_xv_y & v_y\sin(\theta)+(1-\cos(\theta))v_xv_z \\ v_z\sin(\theta)+(1-\cos(\theta))v_xv_y & 1+(1-\cos(\theta))(v_y(v_y-1) & -v_x\sin(\theta)+(1-\cos(\theta))v_yv_z \\ -v_y\sin(\theta)+(1-\cos(\theta))v_xv_z & v_x\sin(\theta)+(1-\cos(\theta))v_yv_z & 1+(1-\cos(\theta))(v_z(v_z-1) \end{bmatrix}$$

Although this provides a means by which the axis angle representation can be used to define a workable rotational transform, a much simpler implementation is possible by the use of unit quaternions. Unlike the rotation matrix, the unit quaternion bears a more direct correspondence to the axis-angle representation.

Appendix 10: Unit quaternion representation of rotation

An alternative method by witch rotation can be represented is by the four dimensional system of unit quaternions. It provides several advantages over the rotation matrix in terms of stability, size and computational efficiency. By this system the entire space of 3D rotations is represented by the surface of a 4-dimesional sphere of unit radius. The quaternion system of numbers extends the concept of complex numbers beyond a single definition of $\sqrt{-1} = i$ to give three separate imaginary quantities i, j and k, each sharing the same relationship to the negative square root [94]. In relation to one other, the product of any two imaginary terms will always result in the third. This is shown below with a full map of the relations between terms given in figure A10.1:

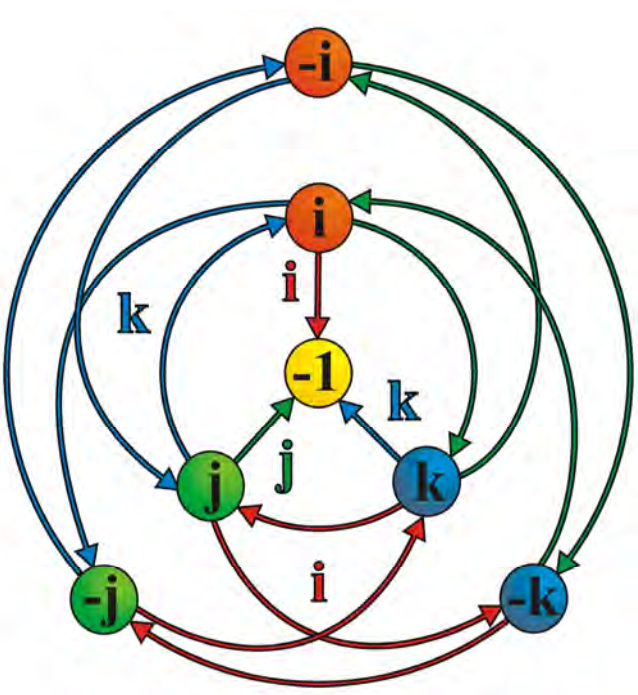


Figure A10.1 – Quaternion Multiplication Chart

$$ij = k,\ jk = i, ki = j$$

Also for the reverse combinations:

$$ji = -k, kj = -i, ik = -j$$

While their relation to minus -1 is:

$$i^2 = -1, j^2 = -1, k^2 = -1$$

This behaviour lends the quaternions the property of non-commutativity. Therefore, in a manner similar to rotation, the result of a quaternion multiplication depends upon its order. For example:

$$iji = j \qquad jii = -j \qquad iij = -1 + j$$

When performing arithmetic upon a quaternion it can be considered to have two distinct parts, a real scalar component *w* and an imaginary vector **q**. Therefore, the quaternion $\dot{\mathbf{q}}$ can be written as:

$$\dot{\mathbf{q}} = w + iq_x + jq_y + kq_z$$

where $\mathbf{q} = [q_x \; q_y \; q_z]$.

In the case of the addition and subtraction of quaternions, the result is the combination of like terms:

$$\dot{\mathbf{p}} + \dot{\mathbf{q}} = (p_w + q_w) + i(p_x + q_x) + j(p_y + q_y) + k(p_z + q_z)$$

For multiplication, to preserve the rules stated above the working is more complex:

$$\dot{\mathbf{p}}\dot{\mathbf{q}} = (p_w q_w - p_x q_x - q_y p_y - q_z p_z) + i(p_w q_w + p_x q_w + p_y q_z - p_z q_y) + j(p_w q_x - p_x q_z + p_y q_w + p_z q_x) + k(p_w q_z + p_x q_y - p_y q_x + p_z q_w)$$

Due to the rule of non-commutativity $\dot{\mathbf{p}}\dot{\mathbf{q}} \neq \dot{\mathbf{q}}\dot{\mathbf{p}}$.

It is possible to express this multiplication as a matrix – vector multiplication by expanding one quaternion in the operation to a 4x4 skew symmetric matrix [95]. Depending on the order of multiplication then this matrix may be partially transposed down its central diagonal:

$$\dot{\mathbf{p}}\dot{\mathbf{q}} = \begin{bmatrix} p_w & -p_z & -p_y & -p_z \\ p_x & p_w & -p_z & p_y \\ p_y & p_z & p_w & -p_x \\ p_z & -p_y & p_x & p_w \end{bmatrix} \dot{\mathbf{q}} = \mathbf{P}\dot{\mathbf{q}} \text{ or } \dot{\mathbf{q}}\dot{\mathbf{p}} = \begin{bmatrix} p_w & -p_z & -p_y & -p_z \\ p_x & p_w & p_z & -p_y \\ p_y & -p_z & p_w & -p_x \\ p_z & p_y & -p_x & p_w \end{bmatrix} \dot{\mathbf{q}} = \overline{\mathbf{P}}\dot{\mathbf{q}}$$

To allow for representation of rotation a quaternion must have unit magnitude, i.e.:

$$q_w^{\,2} + q_x^{\,2} + q_y^{\,2} + q_z^{\,2} = 1$$

A single vector or point in 3D space $\mathbf{p}_{xyz}$ may be encoded as a quaternion with a purely imaginary vector component and a scalar rotation of zero, as follows:

$$\dot{\mathbf{q}}_{xyz} = 0 + ip_x + jp_y + kp_z$$

To rotate this encoded vector by the unit quaternion the following operation is performed:

$$\mathbf{p}_{xyz} = \dot{\mathbf{q}}\dot{\mathbf{p}}_{xyz}\dot{\mathbf{q}}^{*}$$

Where $\dot{\mathbf{q}}^{*}$ is the conjugate of $\dot{\mathbf{q}}$, given reversing the signs of the imaginary components:

$$\dot{\mathbf{q}}^{*} = w + i(-q_x) + j(-q_y) + k(-q_z)$$

It can be noted that the conjugate represents the inverse transform, therefore the opposite rotation to $\dot{\mathbf{q}}$ is given by:

$$\dot{\mathbf{p}}_{xyz} = \dot{\mathbf{q}}^{*}\dot{\mathbf{p}}'_{xyz}\dot{\mathbf{q}}$$

This presents a simpler method of obtaining the inverse transform than is possible using rotation matrices.

Expanding the transform above in to skew symmetric matrix form:

$$\dot{\mathbf{p}}`_{xyz} = (\mathbf{Q}\dot{\mathbf{p}})\dot{\mathbf{q}}^{*} = \overline{\mathbf{Q}}^{T}(\mathbf{Q}\dot{\mathbf{p}}) = (\overline{\mathbf{Q}}^{T}\mathbf{Q})\dot{\mathbf{p}}$$

Now multiplying out the expression for $\overline{\mathbf{Q}}^{\mathrm{T}}\mathbf{Q}$ we obtain the final expression for the rotation of $\dot{\mathbf{p}}_{xyz}$ by $\dot{\mathbf{q}}$:

$$\dot{\mathbf{p}}`_{xyz} = \begin{bmatrix} \dot{\mathbf{q}} \cdot \dot{\mathbf{q}} & 0 & 0 & 0 \\ 0 & (q_w{}^2 + q_x{}^2 - q_y{}^2 - q_z{}^2) & 2(q_x q_y - q_w q_z) & 2(q_x q_z + q_w q_y) \\ 0 & 2(q_y q_x + q_w q_z) & (q_w{}^2 - q_x{}^2 + q_y{}^2 - q_z{}^2) & 2(q_y q_z - q_w q_x) \\ 0 & 2(q_z q_x - q_w q_y) & 2(q_z q_y + q_w q_x) & (q_w{}^2 - q_x{}^2 - q_y{}^2 + q_z{}^2) \end{bmatrix} \dot{\mathbf{p}}_{xyz}$$

The unit quaternion representation can be envisaged as being analogous to the axis-angle system of representation, which allows for straightforward conversion between the two systems. Defining $\dot{\mathbf{q}}$ as and rotation of θ about an axis **a**, with components a_x, a_y, a_z, results in the expression:

$$\dot{\mathbf{q}} = \boldsymbol{cos}\frac{\theta}{2} + \boldsymbol{sin}\frac{\theta}{2}\boldsymbol{(}ia_x + ja_y + ka_z\boldsymbol{)}$$

The inverse operation gives θ and **a** from $\dot{\mathbf{q}}$:

$$\theta = 2\boldsymbol{cos}^{-1}\boldsymbol{(}q_w\boldsymbol{)} \qquad \mathbf{a} = \left[\frac{q_x}{\sqrt{1-q_w{}^2}} \quad \frac{q_y}{\sqrt{1-q_w{}^2}} \quad \frac{q_z}{\sqrt{1-q_w{}^2}} \right]$$

The equivalent operation for deriving the from the quaternion $\dot{\mathbf{q}}$ from the rotation matrix $\mathbf{R_q}$ is:

$$\dot{\mathbf{q}} = \frac{\sqrt{1 + \mathrm{R}_{1,1} + \mathrm{R}_{2,2} + \mathrm{R}_{3,3}}}{2} + i\left(\frac{R_{3,2} - R_{2,3}}{2\sqrt{(1 + \mathrm{R}_{1,1} + \mathrm{R}_{2,2} + \mathrm{R}_{3,3})}} \right)$$
$$+ j\left(\frac{R_{1,3} - R_{3,1}}{2\sqrt{(1 + \mathrm{R}_{1,1} + \mathrm{R}_{2,2} + \mathrm{R}_{3,3})}} \right) + k\left(\frac{R_{2,1} - R_{1,2}}{2\sqrt{(1 + \mathrm{R}_{1,1} + \mathrm{R}_{2,2} + \mathrm{R}_{3,3})}} \right)$$

Where $\mathbf{R_q} = \begin{bmatrix} R_{1,1} & R_{2,2} & R_{1,3} \\ R_{2,1} & R_{2,2} & R_{2,3} \\ R_{3,1} & R_{3,2} & R_{3,3} \end{bmatrix}$

Finally, transformation between the quaternion representation and rotation matrix is achieved by the rotation of the three column vectors of the Identity matrix by the quaternion multiplication.

REFERENCES

[1] A. Cappozzo, U. Della Croce, A. Leardini, L. Chiari. “Human movement analysis using stereophotogrammetry: Part 1: theoretical background”, *Gait & Posture*, vol. 21, pp. 186-196, 2005

[2] J. Rasmussen, M. Damsgaard, E. Surma, Søren T. Christensen, and M. de Zee. “Designing a general software system for musculoskeletal analysis”, *Proc. of IX International Symposium on Computer Simulation in Biomechanics*. Sydney, Australia, 2003

[3] J. P. Holden, J. A. Orsini, K. L Siegel, T. M. Kepple, L. H. Gerber, S. J. Stanhope. “Surface movement errors in shank kinematics and knee kinematics during gait”, *Gait and Posture*, vol. 5, pp. 217–227, 1997

[4] M. P. Kadaba, H. K. Ramakrishnan, M. E. Wootten. “Measurement of Lower Extremity Kinematics During Level Walking”, Journal of Orthopaedic Research, vol. 8, pp. 383-392, 1990

[5] K. O'Leary, K.A. Vorpahl, B. J Heiderscheit. “Effect of cushioned insoles on impact forces during running”, *Am Podiatr Med Assoc*., vol 98, pp. 36-41

[6] N. Apuzzo, R. Plankers, P. Fua , N. D’apuzzo A , R. Plänkers B , P. Fua, A. Gruen, D. Thalmann. “Modeling Human Bodies from Video Sequences” *Zurich : Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology*, 1999

[7] Oqus Motion Capture Camera Specification. Qualisys Motion Capture Systems AB. http://qualisys46.kaigan.se/archive/product_information_pdf/PI_Oqus.pdf

[8] Vicon MX Motion Capture System. Vicon Motion Systems Ltd. http://www.vicon.com/products/viconmx.html

[9] Optotrack Certus. Northern Digital Inc. http://www.ndigital.com/lifesciences/certus-techspecs.php

[10] L. Herda, P. Fua, R. Plankers, R. Boulic, D. Thalmann. “Skeleton-based motion capture for robust reconstruction of human motion”, *Computer Animation*, Philadelphia, PA, 2000

[11] J. G. Richards. “The measurement of human motion: A comparison of commercially available systems”, *Human Movement Science*, vol. 18, pp 589-602. 1999

[12] A. Ansar, K. Daniilidis. "Linear Pose Estimation from Points or Lines", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25. pp. 578-589, 2003

[13] G. Liu, L. McMillan. “Estimation of missing markers in human motion capture”, *The Visual Computer*, vol. 22, pp. 721-728, 2006

[14] X. Chen, J. Davis. “An occlusion metric for selecting robust camera configurations”, *Machine Vision and Applications*, vol. 19, pp. 217–222.

[15] E. Todorov. "Probabilistic Inference of Multijoint Movements, Skeletal Parameters and Marker Attachments From Diverse Motion Capture Data", *Biomedical Engineering, IEEE Transactions on*, vol. 54, pp.1927-1939, 2007

[16] H. J. Luinge, P.H. Veltink. “Measuring orientation of human body segments using miniature gyroscopes and accelerometers”, *Med. Biol. Eng. Comput.*, vol. 43, pp. 273-282, 2005

[17] T. Molet, R. Boulic, D. Thalmann. “Human motion capture driven by orientation measurements”, *Presence: Teleoperators & Virtual*, vol. 8, pp. 187-203, 1999

[18] D. Roetenberg, H. Luinge, and P. Slycke, “Xsens MVN: Full 6-DoF human motion tracking using miniature inertial sensors,” Xsens Technologies Technical Whitepaper, 2009

[19] D. Roetenberg. “Inertial and Magnetic Sensing of Human Motion”, *PhD thesis, Universiteit Twente*, 2006

[20] D. Campolo, F. Keller, E. Guglielmelli. “Inertial/Magnetic Sensors Based Orientation Tracking on the Group of Rigid Body Rotations with Application to Wearable Devices”, *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China, 2006

[21] T. Cloete, C. Scheffer. Benchmarking of a full-body inertial motion capture system for clinical gait analysis, *Proc of 30th Annual International IEEE EMBS Conference*. Vancouver, British Columbia, Canada, 2008

[22] J. Neter, W. Wasserman, M. H. Kutner. *Applied linear statistical models, 2nd Ed.* Homewood, IL. 1985

[23] R. Shapiro. “Direct linear transformation method for three-dimensional cinematography”. *Res. Quart*, vol. *49*, pp. 197-205, 1978

[24] P. Cerveri, N. A. Borghese, A. Pedotti. “Complete calibration of a stereo photogrammetric system through control points of unknown coordinates”, *Journal of Biomechanics*, vol. 31, pp. 935-940, 1998

[25] R. Tsai. "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *Robotics and Automation, IEEE Journal of* , vol. 3, pp. 323-344, 1987

[26] P. J. Figueroa, N. J. Leite, R. M. L. Barros. “A flexible software for tracking of markers used in human motion analysis”, *Computer Methods and Programs in Biomedicine*, vol. 72, pp. 155-165, 2003

[27] N. A. Malik, T. Dracos and D. A. Papantoniou. “Particle tracking velocimetry in three-dimensional flows”, *Experiments in Fluids*, vol. 15, pp. 279, 1993

[28] Proreflex MCU. Qualisys Motion Capture Systems AB. http://qualisys46.kaigan.se/archive/product_information_pdf/AN_ProReflex.pdf

[29] M. Han, W. Xu, H. Tao, Y. Gong. “An Algorithm for Multiple Object Trajectory Tracking,” *Proc. Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 864-871, 2004

[30] T. Pintaric, H. Kaufmann. “Affordable infrared-optical pose-tracking for virtual and augmented reality”, *Proceedings of Trends and Issues in Tracking for Virtual Environments Workshop* (IEEE VR 2007), Charlotte, NC, USA, 2007

[31] accuTrack 500 product information. Atracsys LLC. http://atracsys.com/_products/accuTrack.php

[32] L. Chiari, U. Della Croce, A. Leardini, A. Cappozzo. “Human movement analysis using stereophotogrammetry: Part 2: Instrumental errors”, *Gait & Posture*, vol. 21, pp. 197-211, 2005

[33] R. Cassinis, F. Tampalini, R. Fedrigotti. “Active markers for outdoor and indoor robot localization” in *Proc. of DEA-Unibs*, 2005

[34] Q, Yu, Q, Li, Z. Deng. “Online motion capture marker labeling for multiple interacting articulated targets”, *Computer Graphics Forum (Proceedings of Eurographics 2007)*, pp. 477–483, 2007

[35] N. Yazdi, F. Ayazi, K. Najafi, “Micromachined inertial sensors” *Proc. IEEE*, vol. 86, pp. 1640–1659, 1998

[36] ADIS16350: High Precision Tri-Axis Inertial Sensor. Analog Devices Inc. http://www.analog.com/static/imported-files/data_sheets_obsolete/OBSOLETE%20WATERMARK/ADIS16350_16355.pdf

[37] M. Brodie, A. Walmsley, W Page. “Fusion motion capture: a prototype system using inertial measurement units and GPS for the biomechanical analysis of ski racing”, *Sports Technology*, vol. 1, pp. 17–28, 2008

[38] J. E. Bortz. “A New Mathematical Formulation for Strapdown Inertial Navigation. IEEE trans. serospace and electronics systems, vol. 7, pp. 61-66. 1971

[39] L. Chang and N. Pollard. “Constrained least-squares optimization for robust estimation of centre of rotation”, *Journal of Biomechanics*, vol. 40, pp. 1392–1400, 2007

[40] B. K. P. Horn. “Closed-form solution of absolute orientation using unit quaternions”, *Journal of the Optical Society of America*, vol. 4, pp. 629, 1987

[41] C. Rose, B Guenter, B. Bodenheimer, M. F. Cohen, M. F. “Efficient generation of motion transitions using spacetime constraints”, *Proceedings of ACM SIGGRAPH 1996*, pp. 147-154. 1996

[42] A. Cappozzol, F. Catan, U. Della Crocel, A. Leardini. “Position and orientation in space of bones during movement: anatomical frame definition and determination”, *Clinical Biomechanics*, vol. 10, pp. 171-178, 1995

[43] T. M. Kepple, A. S. Arnold, S. J Stanhope, K. L. Siegel. "Measurement of musculoskeletal motion from surface landmarks: A three-dimensional computer graphics approach." *Journal of Biomechanics*, vol. 27, pp. 365-371, 1998

[44] A. Reinbolt, J. F. Schutte, B. J. Fregly, B. Il Koh, R. T. Haftka, A. D. George, K. H. Mitchell. “Determination of patient-specific multi-joint kinematic models through two-level optimization”, *Journal of Biomechanics*, vol 38, pp. 621-626, 2005

[45] M. C. Silaghi, R. Plankers, R. Boulic, P. Fua, D. Thalmann. “Local and global skeleton fitting techniques for optical motion capture”, *International Workshop on Modelling and Motion Capture Techniques for Virtual Environments*, Geneva, Switzerland, 1998

[46] K. Dorfmüller-Ulhaas. “Optical Tracking - From User Motion To 3D Interaction”, *Ph.D. thesis, Vienna University of Technology*, 2002

[47] K. Aoki, M. Kouchi, M. Mochimaru. “A kinematic estimation of functional joint rotation centers of whole body”, *9th International symp. 3-D analysis of human movement, Valenciennes, France*, 2006

[48] H. J. Woltring. “Representation and calculation of 3-D joint movement”, *Human Movement Science*, vol. 10, pp. 603-616, 1991

[49] D. T. Harryman, J. A. Sidles, J. M. Clark, K. J. McQuade, T. D Gibb, F. A Matsen. “Translation of the humeral head on the glenoid with passive glenohumeral motion”, *J Bone Joint Surg Am.*, vol. 72, pp. 1334-43, 1990

[50] K. Manal, I. McClay, J. Richards, B. Galinat. “Knee moment profiles during walking: errors due to soft tissue movement of the shank and the influence of the reference coordinate system”. *Gait Posture*, vol. 15, pp. 10–17. 2002

[51] C. Reinschmidt, A. J. van den Bogert, B. M. Nigg, A. Lundberg, N. Murphy. “Effect of skin movement on the analysis of skeletal knee joint motion during running”, *J. Biomech*, vol. 30. pp. 729–732, 1997

[52] A. Cappello, A. Cappozzo, P. F. La Palombara, L. Lucchetti, A. Leardini. “Multiple anatomical landmark calibration for optimal bone pose estimation”, *Human Movement Science*, vol. 16, pp. 259-274, 1997

[53] E. J. Alexander, T. P. Andriacchi. “Correcting for deformation in skin-based marker systems”, *Journal of Biomechanics*, vol. 34, pp. 355-361. 2001

[54] C. Goodvin, E. J. Park, K. Huang, K. Sakaki. “Development of a real-time three-dimensional spinal motion measurement system for clinical practice”, *Medical and Biological Engineering and Computing*. vol 44, pp. 1061-1075, 2006

[55] N. Miyata, M. Kouchi, T. Kurihara, M. Mochimaru. "Modeling of human hand link structure from optical motion capture data", *Proc of Intelligent Robots and Systems 2004 (IROS 2004)*, vol. 3, pp. 2129-2135, 2004

[56] A. Cereatti, U. Della Croce, A. Cappozzo. “Reconstruction of skeletal movement using skin markers: comparative assessment of bone pose estimators”. *Journal of NeuroEngineering and Rehabilitation*, vol. 3, 2006

[57] A. Cappozzo, A. Cappello, U. Della Croce, F. Pensalfini. “Surface-marker cluster design criteria for 3-D bone movement reconstruction”, *IEEE Trans Biomed Eng.*, vol. 44, pp. 1165-1174. 1997

[58] A. Angeloni, A. Cappozzo, F. Catani, A. Leardini. “Quantification of relative displacement of skin- and plate-mounted markers with respect to bones”, *Journal of Biomechanics*, vol. 26, pp. 864, 1993

[59] Plug in gait maker set. Vicon polygon user manual. Vicon Motion Systems Ltd. http://www.vicon.com/products/polygon.html

[60] L. Jones, C. A. Holt, A. Bowers. “Movement of the shoulder complex: the development of a measurement technique based on proposed ISB standards”, *Proceedings of the Eighth International Symposium on 3D motion analysis (International Society of Biomechanics)*, 2006

[61] T. D. Collins, S. N. Ghoussayni, D. J. Ewins and J. A. Kent. “A six degrees-of-freedom marker set for gait analysis: Repeatability and comparison with a modified Helen Hayes set”, *Gait & Posture*, vol. 30, pp. 173-180, 2009

[62] Vicon 512 System. Vicon Motion Systems Ltd. http://users.aber.ac.uk/hoh/CS390/512ViconSWManual.pdf

[63] G. I. Zahalak, S.J Heyman. “A quantitative evaluation of the frequency response characteristics of active human skeletal muscle in vivo”, *J. Biomech. Eng.*, vol. 101, pp. 28-37, 1979

[64] T.C. Koh. “Tai Chi Chuan”, *The American Journal of Chinese Medicine (AJCM)*, vol 9, pp. 15-22, 1981

[65] Kai Ming Association for Tai Chi Chuan, West Midlands, UK. http://www.kaiming.co.uk/

[66] A. Lorusso, D.W. Eggert, R.B. Fisher. “A comparison of four algorithms for estimating 3-D rigid transformations”, *Technical Report 737, Department of Artificial Intelligence, University of Edinburgh, Edinburgh, Scotland*, 1995

[67] C. Nester, R. K. Jones, A. Liu, D. Howard, A. Lundberg, A. Arndt, P. Lundgren, A. Stacoff, P. Wolf. “Foot kinematics during walking measured using bone and surface mounted markers”, *J. Biomech*, vol. 40, pp. 3412–3423, 2007

[68] J. W. Fernandez, M. Akbarshahi, H. J. Kim, M. G. Pandy. “Integrating modelling, motion capture and x-ray fluoroscopy to investigate patellofemoral function during dynamic activity”, Computer Methods *in Biomechanics and Biomedical Engineering*, vol. 11, pp. 41-53, 2008

[69] S. J. Piazza, N. Okita, P. R. Cavanagh. “Accuracy of the functional method of hip joint center location: effects of limited motion and varied implementation”, *Journal of Biomechanics*, vol. 34, pp. 967-973, 2001

[70] V. Pratt. “Direct least-squares fitting of algebraic surfaces”, *Computer Graphics (SIGGRAPH ’87 Proceedings)*, vol. 21, pp. 145–152, 1987

[71] A. E. Engin, S. T. Tumer. “Three-Dimensional Kinematic Modelling of the Human Shoulder Complex Part I: Physical Model and Determination of Joint Sinus Cones”, *J. Biomech. Eng.*, vol. 111, pp. 107-112. 1989

[72] Thomas B. Moeslund, Claus B. Madsen, and Erik Granum. “Modelling the 3D pose of a human arm and the shoulder complex utilising only two parameters”, *Integr. Comput.-Aided Eng.* vol. 12, pp. 159-175, 2005

[73] G. Monheit, N. I. Badler. “A kinematic model of the human spine and torso”, *Computer Graphics and Applications, IEEE*, vol. 11, pp. 29-38, 1991.

[74] B. Allen, B. Curless, Z. Popovic. “Articulated body deformation from range scan data”, *ACM SIGGRAPH 2002*, pp. 612—619, 2002

[75] T. Moller. “Fast, minimum storage ray/triangle intersection”, *J. Graphics Tools*, vol. 2, pp. 25–30, 1997

[76] P. Bourke. “Minimum Distance between a Point and a Line”, 1988. http://local.wasp.uwa.edu.au/~pbourke/geometry/pointline/

[77] B. Schmuland. “Shark attacks and the Poisson approximation”, 2001 http://www.stat.ualberta.ca/people/schmu/preprints/poisson.pdf

[78] SCB65EC Single Board Computer, Modtronix Engineering http://www.modtronix.com/products/sbc65ec/sbc65ecr3.pdf

[79] PIC18F6621 Data Sheet. Microchip Technology Inc. http://ww1.microchip.com/downloads/en/DeviceDoc/39612.pdf

[80] G. Wu, F. C. T. van der Helm, H. E. J. Veeger, M. Makhsous, P. Van Roy, C. Anglin, J. Nagels, A. R. Karduna, K. McQuade, X. Wang, F. W. Werner, B. Buchholz. “ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion-Part II: shoulder, elbow, wrist and hand”. *Journal of Biomechanics*, vol. 38, pp. 981-992, 2005

[81] S. You, U. Neumann. “Fusion of vision and gyro tracking for robust augmented reality registration”, *Proc. IEEE Virtual Reality Conference (VR 2001)*,Yokohama, Japan, pp. 71–78, 2001

[82] S. K. Hong. “Fuzzy logic based closed-loop strapdown attitude system for unmanned aerial vehicle (UAV)”, *Sensors and Actuators*, vol. 107, pp. 109-118, 2003

[83] R. Mahony, T. Hamel, J. Pflimlin. “Nonlinear Complementary Filters on the Special Orthogonal Group”, *IEEE trans. on automatic control*, vol. 53, pp. 1203-1218 , 2008

[84] D. Giansanti, V. Macellari, G. Maccioni, A. Cappozzo. "Is it feasible to reconstruct body segment 3-D position and orientation using accelerometric data?", *IEEE Trans. on Biomedical Engineering*, vol. 50, pp. 476-483, 2003

[85] A. Kim, M. F. Golnaraghi. “A quaternion-based orientation estimation algorithm using an inertial measurement unit”, *Position Location and Navigation Symposium (PLANS 2004)*, pp. 268-272, 2004

[86] C. de Boor. “A Practical Guide to Splines”, *Springer-Verlag*, 1978

[87] J. Tiesel, J. Loviscach. “A Mobile Low-Cost Motion Capture System Based on Accelerometers”. *Advances in Visual Computing, Springer-Verlag*, 2006

[88] T. W. Lu, J. J. O'Connor. “Bone position estimation from skin marker co-ordinates using global optimisation with joint constraints”. *Journal of Biomechanics*, vol. 32, pp. 129-134, 1999

[89] Visual 3D biomechanical analysis and modelling software. C-Motion Inc. http://www.c-motion.com/products/visual3d.php

[90] Motionpod IMU. Movea Inc. http://www.movea.com/index.php/products/motionic-hardware/motionpod

[91] IMU tag v1.0 - Inertial/Magnetic Data Logger. x-io Technologies Ltd. http://www.x-io.co.uk/node/3

[92] R. F. Stengel. “Flight Dynamics”, *Princeton University Press*, 2004

[93] O.A. Bauchau, L. Trainelli. “The vectorial parameterization of rotation”, *Nonlinear Dynamics*, vol. 32, pp. 71-92, 2003

[94] W. R. Hamilton. “On quaternions, or on a new system of imaginaries in algebra”. *Philosophical Magazine*. vol. 25, pp. 489-495, 1844

[95] J. B. Kuipers. “Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality”. *Princeton University Press*, 2002