QUANTIFICATION AND COMPARISONS OF MOTION METRICS FOR FINGER TAPPING PERFORMED BY CHILDREN WITH CEREBRAL PALSY BEFORE AND AFTER THERAPY WITH TWO MOTION TRACKING CAMERA SYSTEMS

by

ANKITA H CHAINANI

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN BIOMEDICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2012

Acknowledgements

When I first came to USA, research was just in my imagination. I had this passion to conduct a research and contribute to lives of many people. When I first started working on this project, it was really hard but things went on getting easy with guidance of my mentor Dr. George Alexandrakis. I really appreciate the time and suggestions he gave for strategic planning of my project. The project I worked on is entirely based on Computer vision. I am grateful for the course taught by Dr. Gian Luca, which helped me understand the basics of working and calibration of cameras that helped me to initiate the project. I am really thankful to Dr. Gian for his advice and many insightful suggestions and discussions throughout this project. I am thankful to my colleague Nathan Hervey. His suggestions helped me a lot in my coding. I would like to acknowledge Bilal Khan, Laura Shagman, Chester Wildey, Duncan Macfarlane for their help.

I am obliged to my parents and my family for so many things but most of all, for the confidence they showed in me. I am thankful to my mother, Mrs. Rachna Chainani for her love and blessings and my father Mr. Harilal Chainani for his words of wisdom which have helped me a lot in practical life. Their belief and faith in me gave tremendous encouragement. I want to thank my elder sister Swati Chainani for being there in my highs and lows, for supporting me when I was stressed. Swati has taught me many things and her moral support always gave me boost to do my work more enthusiastically. I am deeply indebted to my friends for their constant motivation and support. They say life is merrier with friends. I am glad that I have friends who despite of being in the same phase as I was while writing, made things look simpler with their humor and jokes. I am thankful to my roommate as she saw me in the worst of times and gave encouragement for completing my work.

November 26, 2012

Abstract

QUANTIFICATION AND COMPARISONS OF MOTION METRICS FOR FINGER TAPPING PERFORMED BY CHILDREN WITH CEREBRAL PALSY BEFORE AND AFTER THERAPY WITH TWO MOTION TRACKING CAMERA SYSTEMS

Ankita Chainani, M.S.

The University of Texas at Arlington, 2012

Supervising Professor: George Alexandrakis

The quantification of any improvements in the use of a paretic arm after children with cerebral palsy (CP) undergo rehabilitation is currently done by occupational therapists who administer a series of manual ability tests and assign a subjective score for each test task performed. In this work the hypothesis was explored of whether quantification of hand motion metrics by motion tracking while children with CP performed a hand tapping task could be used as a more quantitative surrogate for assessing how well these children control the use of their hand. To that end, two children with CP, both classified as level one in the Manual Ability Classification Scale (MACS), both age eleven, were measured during finger tapping just before and immediately after two weeks of Constraint Induced Movement Therapy (CIMT). In addition, two age-matched controls were measured twice during the same time interval. Motion tracking was performed by attaching retroreflective targets on the nails, knuckles and wrists of both hands for all subjects measured. The motions of these targets were measured simultaneously by two motion tracking systems, a lower cost two-camera one and a higher cost six-camera one. The data from both camera systems were analyzed to quantify three simple motion

metrics per subject, namely the tapping amplitude, average tapping velocity and the instantaneous velocity profile of the paretic hand of subjects with CP and the dominant right hand of controls. These hand performance metrics were quantified before and after CIMT for the subjects with CP, and the measurements on control subjects were used to assess the level of motion metric variability between the two measurement sessions that were also two weeks apart. It was found that some of the motion metrics showed changes after CIMT that follow qualitatively the trends seen in clinically administered manual ability tests, though a lot more subjects would need to be measured to verify the significance of these trends. In addition, the accuracy of the two-camera system to track hand motions was compared relative to the six-camera one for all subjects measured. The purpose of this latter comparison was to assess the extent to which the lower cost two-camera system could be used for quantifying finger tapping motion metrics accurately. It was found that the two-camera system's performance was inferior, for technical reasons that are discussed, which made its use in this work much more challenging than using the six-camera system.

Table of Contents

List of Illustrations

List of Tables

Chapter 1

Introduction

1.1 Background

Cerebral Palsy (CP) is one of the most common congenital childhood disorders that affect 1.5 to 4 out of 1000 live births [1]. The creation of a prenatal or perinatal lesion could lead to brain damage or abnormal development of the brain that subsequently manifests itself under a variety of developmental impediments that are all classified under the umbrella term of CP. Specifically; this condition affects muscle tone, motor skills and the ability to move in a coordinated manner. Patients with CP may also have neurodevelopmental impairments that can affect the learning process, communication, reflexes, balance and behavior. This study is focused on patients with spastic hemiplegia, i.e. they only have one affected arm. In this disorder some of the muscles on one side of body remain in state of perpetual contraction (spasticity), which results in low functional abilities as the patient cannot perform motor tasks at full capacity.

Cerebral palsy cannot be cured, but a variety of treatments can help people with CP to maximize their strengths and abilities. Specific treatments vary with respect to the symptoms observed in the patients. The treatment methods developed focuses on ways to maintain a person's quality of life. There are several methods that are employed for the treatment of CP. One of these methods focuses on helping patients attain muscle control by performing basic tasks that they repeat in everyday life. A functional task oriented treatment approach has been developed in neurodevelopmental therapy, which is being used now as a preferred method of treatment [3]. There are also treatment methods that focus on performing different structured tasks repetitively as per the patients learning strengths and needs [4]. No matter the choice of tasks for the patients to practice, the outcome could be better if the patient was forced to use their paretic arm more and thus

develop more self-confidence in using it. Constrained Induced Movement Therapy (CIMT) was designed to address this need and is currently used for the treatment of spastic hemiplegia. As the CIMT name indicates, this therapy encourages intensive use of the affected limb by restraining the unaffected hand for a prolonged period of time. This method of treatment has been observed to be one of the most effective to date and moreover, it does not involve medications that could possibly have side effects [5].

Motion tracking has been used for assessing a patient's condition in diverse patient populations diagnosed with movement disorders and developmental disabilities like Parkinson's disease, stroke, brain or spinal cord injuries [10] [11]. Motion analysis can also be used for providing quantitative measures of the effectiveness of post rehabilitative treatment [12] though it has not been applied to the monitoring of children with CP to date.

1.2 Motivation

There is no standard treatment that will work for every individual with CP. Therefore, it became necessary to analyze the functional state of every patient with CP so as to give the most effective treatment possible. For the purpose of assessing the functional performance of patients with CP different classification systems have been developed that discriminate and categorize them on the basis of their performance and limitations. One of the systems used is Manual Ability Classification System (MACS) [7] which represents the child's manual ability while handling objects in daily activities. MACS classifies the patients with CP into five levels from 1-5, ranging from low level of severity to high level of severity. Level 1 is assigned to the patients who can handle objects easily but have limitations on speed and accuracy with which the task is performed, while Level 5 is assigned to patients who cannot handle objects and require complete assistance.

A few recent studies have shown that MACS is a reliable classification technique, although it has not been used widely [7] [6]. One of the issues with the MACS scale is that it is crude and as a result, patients with visually perceived improvements after rehabilitation do have any changes in their MACS score. Also, the MACS score is not predictive of rehabilitation outcomes as it is frequently the case that between two children that have the same impairment severity classification, as quantified by MACS, one of them responds to a given treatment and the other one does not [8][9]. Other manual ability assessment scales such as Assisted Hand Assessment (AHA) [13] and Melbourne test. In this work the hypothesis of whether hand motion tracking can provide reliable variables to assess improvement after therapy was explored. This was only a preliminary study to assess the feasibility of this hypothesis that, if proven true, could lead to future follow-up studies on assessing its sensitivity, specificity and treatment outcome predictability.

1.3 Purpose

The aim of this thesis was to derive objective measures of upper limb function more quantitatively in children with upper extremity hemipleagia. Camera based motion tracking was employed to measure different motion metrics for subjects performing a finger tapping task. The functional performance metrics quantified were the average tapping amplitude, average tapping velocity and the instantaneous velocity profile of the index finger of the tapping hand. These metrics were measured in two eleven year old subjects with CP (MACS-1) before and after two weeks of CIMT and in two age-matched controls that were measured twice during the same time interval to obtain baseline measurements of performance variation between measurement sessions.

Two motion tracking systems, a low cost two-camera one and a high cost six-camera one, were used to measure the same motion metrics on each subject so that

comparison could be made between the two systems. The purpose of this part of the work was to assess whether the lower cost system could do as good a job in motion tracking as the higher cost system. The average tapping amplitude, average tapping velocity and instantaneous velocity profiles were calculated from both systems to perform this comparison.

Chapter 2 gives the description of these two camera systems used, and the detailed theory of the calibration algorithm used for the two-camera system. The latter had to be set up and calibrated manually in the lab, while the higher cost six-camera system, located in the Motion Analysis Lab of the Texas Scottish Rite Hospital for Children at Dallas, was automated in how it was calibrated and deduced the 3D co-ordinates of targets. The later part of that chapter describes how the 3D co-ordinates of retroreflective targets on the moving finger were estimated for the two-camera system by a triangulation algorithm and how a coordinate transformation was performed from the two-camera system to the six-camera system so as to perform motion metric comparisons. Chapter 3 describes the experiments performed to check the performance of the two-camera system for different camera geometry setups that were implemented in the laboratory before measurements were performed on children with the setup deduced as optimal. The later part of this Chapter shows the motion metric results computed for both camera systems, describes all the changes found in motion metrics after CIMT for the subjects with CP and the baseline variations in controls.. The last section of this Chapter shows the comparison between the two camera systems so as to identify the strengths and limitations of the lower cost two-camera system. Finally, a brief Conclusions Chapter summarizes the findings of this work and suggests possible future work.

Chapter 2

Setup And Calibration Procedures For A Two-Camera And A Six-Camera Motion Tracking System

2.1 Setup Optimization and Calibration of a Two-Camera Motion Tracking System

In computer graphics, many advanced camera systems are being used in order to get 3D co-ordinates for a vast variety of applications like manipulation of objects by robotic arms, autonomous navigation of vehicles and developing a map of the environment.[23][24][25][26]. A camera performs the mapping of a 3D object (world space) onto 2D image plane (image space). There are three main components of the camera that take part in this 3D-2D mapping, these parts being the lens, which helps in focusing the light that enters the camera onto the image plane, the aperture that controls the amount of light entering camera and the image plane itself, which is the CCD array.

Motion tracking with a two-camera system entails the transformation of 2D coordinates of retro-reflective targets registered by each camera to estimates of the 3D coordinates of those targets. A camera calibration procedure was first performed in order to get the pose (location and orientation) of each camera with respect to the coordinate system of the calibration target. Using the deduced camera pose and 2D pixel co-ordinates of each target for both camera systems, a transformation was computed to transform, through a triangulation procedure, the 2D coordinate values registered by each camera to the 3D co-ordinates of targets on the subjects' hands. These sequences of procedures, from camera calibration to target triangulation, are explained sequentially here below.

2.1.1 Instrumentation

Three dimensional co-ordinates were obtained by using two cameras from Vision Components (VC), model VC4438E. Each camera has 64MB DRAM, 4MB Flash EPROM for data storage, and a 1/3-inch SONY CCD sensor with a pixel resolution of 640×480. The maximum frame rate attainable is 63 frames per second. Attached to the top side of each camera there is an Infrared (IR) illuminator with 140 Light emitting diodes (LEDs) that emit IR light at approximately 850 nm. IR high pass filters for 720 nm and longer wavelengths. 5mm in diameter, were attached onto each camera. The filters were employed in order to remove any extra reflections originating from room lighting that has a lot of its spectral power in the visible wavelength range.

The VC camera system has a high speed trigger input down to 5 µsec. Both cameras were connected to a common manual trigger to avoid delays between them in order to synchronize the temporal sampling of both cameras. Each camera runs on-board software that is optimized for the image acquisition process so that after the user presses the trigger (start) button the hardware can handle exposure control, readout and image processing at same time. The software used in these cameras deduces the number of retro reflective targets by calculating the centroid of each target. The camera gives four variables in output, time (msec), area of target and X, Y pixel co-ordinates of each target. The video out interface was connected to a monitor display. The sampling rate of cameras was 30Hz, but for the experiments performed on subjects, the sampling rate was 25Hz as with increased number of targets, the sampling rate reduced.

2.1.2 Calibration target

In order to retrieve accurately the 3D coordinates of retroreflective targets attached onto a person one first needs to derive the camera system matrix. This is achieved by imaging a collection of targets with precisely known locations and use algorithms like the one described in subsection 2.4.1.2 below to compute what the system matrix must have been in order to attain those known target coordinates. For purposes of improved reproducibility of the calibration procedure a set of reflective targets is often constructed in a solid 2D or 3D structure, so that the distance between targets remains fixed. There are many possible arrangements for these calibration targets. However, it has been observed that 3D calibration objects with reflective targets arranged in a coplanar manner produce less stable and precise system calibration estimates compared to non-coplanar target arrangements [31]. This happens due to inaccurate coupling between intrinsic and extrinsic parameters with respect to minimization of the reprojection error. Even if planar targets are used, multiple views are required to generate sufficient point correspondences in both cameras; this makes the calibration procedure more tedious. It has also been observed that symmetric targets like the ones that are circular in shape are detectable with sub-pixel accuracy [30][32]. Perspective projection of circle on an image plane will result in a circle and detection of its centroid is thus more accurate [20].

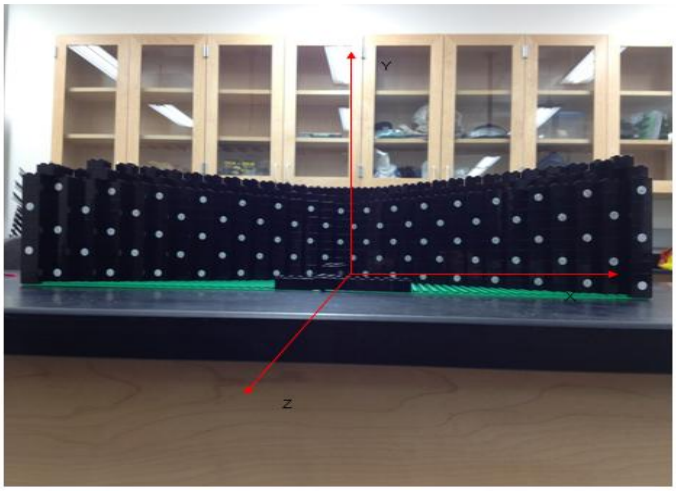


Figure 2-1: 3D Calibration target made from Legos with circular retro reflective targets

In this work a 3D rigid frame calibration target was made from Legos as these are known to be sufficiently stable for this purpose [19]. Circular retroreflective targets of 2 mm diameter were punched from tape and were affixed to Lego block. One target was attached on two Lego blocks. The vertical distance between two Lego blocks was 38.2mm. The step size for X was 15mm and Z was 15mm. The camera software performed raster scanning of each processed image frame and recorded the 2D pixel co-ordinates of each target in the order that they were encountered during the raster scan. The resulting output was written on disk as a text file with the format of a vertical column. The first column indicated the time in msec; the second column represents area of the target, the third column X pixel co-ordinate and the fourth column the Y pixel co-ordinate. The frame rate was 25 Hz, which provided a value for the 2D co-ordinates of each target pixel every 40 msec .For the purpose of sorting the targets pixel positions with respect to time, two and three targets were arranged alternatively on the columns of Lego as seen

in Figure 2-3. In all, there were 80 targets arranged in 3D fashion. Black color Lego columns were used as they were less reflective and gave a background of high contrast. The farthest column of Lego in the Z direction was the origin of the World or the Calibration co-ordinate system and the X, Y, Z axis orientation was chosen using a right hand co-ordinate system.

2.1.3 Camera Setup Optimization

In order to determine the placement of both cameras that could give the least amount of errors during motion tracking of finger tapping, different arrangements of camera set-up were physically implemented in the laboratory set up and errors were estimated. More specifically, average error in 3D and RMS errors in X, Y and Z were estimated.

Although the above was not an exhaustive search of all possibilities, from the ones tested it was found that best position for taking measurement was if both cameras were placed on a pole at height difference of 1 ft 4" with horizontal difference to be approximately 2-3" between both cameras as seen in Figure 2-2. The distance of the bottom camera from the table was 2ft 10". These distance measurements were taken from the approximate centre of each camera with a measuring tape. The horizontal distance from the cameras to the table was 87".

Figure 2-2: Optimized Position of two cameras for 3D tracking of targets

2.1.4 Camera Calibration

Camera calibration is a technique used to obtain information about the projection from 3D co-ordinates to 2D image co-ordinates and to help find the parameters internal to the camera that affect how images are rendered. Many techniques of camera calibration have been used in computer vision for various applications [14][15][16][17]. There are two sets of important parameters that completely describe the relation between images and the camera systems from which these were captured. These are called intrinsic and extrinsic camera parameters. Intrinsic parameters determine how pixel coordinates might be calculated given prior information of 3D point targets with respect to the camera. These parameters are focal length, scale factors, skew factor, and the camera principal point. Extrinsic camera parameters like Rotation and Translation give information about orientation and position of the camera with respect to a reference frame. Extrinsic and Intrinsic camera parameters, put together in the form of a matrix, form an image projection matrix (P-matrix) which can be determined with a calibration process. The

subsections below describe how the P-matrix was obtained from the calibration procedure of the two-camera motion tracking system.

2.1.4.1 Co-ordinate Systems

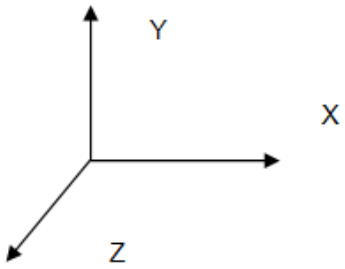


Figure 2-3: Right hand Co-ordinate System

The different coordinate systems that are used for the derivation of the camera pose equations are the World coordinate system, the Camera coordinate system and the Calibration coordinate system. The World co-ordinate system is the one where an object is defined in 3D space. The origin of the Camera co-ordinate system is at the geometric center of the camera. The Calibration coordinate system is defined by the 3D calibration Lego target developed for the calibration of the cameras with the origin as shown in Figure 2-1. All three coordinate systems are right handed with the positive X axis pointing towards the right, the positive Y axis in the upward direction and the positive Z axis pointing forward. A positive Z axis can be determined by pointing the right hand fingers in the positive X direction and curling them around the positive Y axis, so that the thumb points in the upward Z direction. For the work in this thesis the World and Calibration coordinate systems are same.

Camera matrices that relate the 3D world co-ordinates to the 2D image co-ordinates can be easily explained by keeping the optical centre at the origin of The

Camera co-ordinate system. As seen in Figure 2-4, the optical axis of the camera is considered to be collinear with the camera's Z axis. The intersection of the optical axis with the image plane is called the principal point. The image plane is in front of the optical centre at Z=f, where f is the focal length of camera, to avoid image inversion. World and Camera co-ordinate systems are considered to be coincident for this explanation.

2.1.4.2 The Camera Projection Matrix

Consider a 3D world point W with co-ordinates (*X, Y, Z*). As seen in Figure 2-4, point W will be imaged as the intersection of the line passing from (0,0,0) to (*X,Y,Z*) and the Z=f plane. By means of similar triangles, the image co-ordinates (*x, y*) are given as,

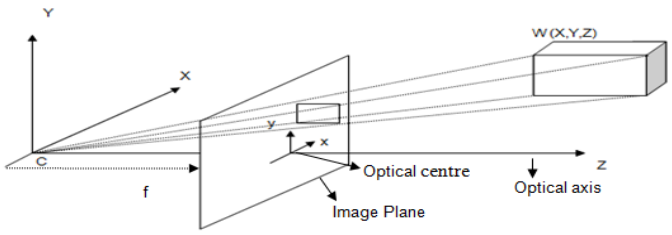


Figure 2-4 :Point W with co-ordinates (X,Y,Z) on a 3D World object whose projection is seen in the image plane. Point C is the optical centre of the camera and the origin of the camera co-ordinate system.

$$X = \frac{fX_c}{Z_c} \qquad y = \frac{fY_c}{Z_c} \tag{2.1}$$

The division in the above equation is not linear in Z. To avoid this non-linearity, homogeneous co-ordinates are used for this coordinate representation, where the dimensionality of the transformation matrix is increased by one to enable homogeneous coordinate representation. In matrix notation:

$$(\lambda x, \lambda y, \lambda)^T = (f * X_c, f * Y_c, Z_c)^T \tag{2.2}$$

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \approx \begin{bmatrix} f\ X_c \\ f\ Y_c \\ f\ Z_c \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{2.3}$$

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{2.4}$$

λ=Z is the homogeneous scaling factor.

The image sensor, CCD is made up of a square grid of pixels. The Image plane co-ordinate system is defined by taking its origin at the centre of the image plane at the principal point, such that the image x and y axes are parallel to camera's x, y axes. The co-ordinates (x, y, 1) seen in Equation (2.4) have units in millimeters. The imaging surface has its origin at the top left corner of the pixel array. So, in order to transform from real physical position to pixel position, we need to add the distance from that point to left corner of the pixel array. Also, by using scaling factors for the x and y directions (pixels /mm), the position of a point on the image plane can be transformed from physical units to pixels.

$$u = u_o + k_u * x \tag{2.5}$$

$$v = v_o + k_v * y \tag{2.6}$$

Where k_u, k_v are scaling factors in number of pixels per mm, and (u_o, v_o) is the position of principal point in pixels.

In a homogeneous co-ordinate system we thus have:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_u & 0 & u_o \\ 0 & k_v & v_o \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{2.7}$$

Combining equation (2.4) and (2.7),

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_u & 0 & u_o \\ 0 & k_v & v_o \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{2.8}$$

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_u f & 0 & u_o & 0 \\ 0 & k_v f & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{2.9}$$

The camera pixels are not necessarily square and the pixel aspect ratio might not always be 1:1. Also, the pixel grid might be skewed which means that u, v might not be orthogonal to each other. In order to account for non-rectangular pixels, a skew parameter is introduced in the projection matrix. The equation can now be updated as:

$$\lambda\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_u & s & u_o & 0 \\ 0 & \alpha_v & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{2.10}$$

$M_{intrinsic} = K = \begin{bmatrix} \alpha_u & s & u_o & 0 \\ 0 & \alpha_v & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ is called the intrinsic matrix.

Consider a system in which the Camera and World co-ordinates are not coincident. In order to get transformation from Camera co-ordinates to World co-ordinates, extrinsic parameters are defined. If the point W in Camera frame is expressed as $(X_c, Y_c, Z_c)^T$ and in World co-ordinates as $(X, Y, Z)^T$, we need to find a translation vector that maps the camera's origin to the World co-ordinate frame and then perform a rotation that will align the camera's axes to World's axes:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} * \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} \tag{2.11}$$

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2.12}$$

$M_{extrinsic} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$ is the extrinsic matrix.

The matrix for rotation around the X axis is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix}$, for rotation around the Y axis is $\begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix}$ and for rotation around the Z axis is $\begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$. These are the elementary rotation matrices which can be combined to form any 3D rotation. Depending upon the order of rotations, first rotation around X axis (α) then around Y axis (β) and later Z axis (γ), the composite rotation matrix in Equation (2.12) has the following parameters:

$$R = R_\gamma * R_\beta * R_\alpha \tag{2.13}$$

$$R = \begin{bmatrix} \cos\beta\cos\gamma & \cos\alpha\sin\gamma + \sin\alpha\sin\beta\cos\gamma & \sin\alpha\sin\gamma - \cos\alpha\sin\beta\cos\gamma \\ -\sin\beta\sin\gamma & \cos\alpha\cos\gamma - \sin\alpha\sin\beta\sin\gamma & \sin\alpha\cos\gamma + \cos\alpha\sin\beta\sin\gamma \\ \sin\beta & -\sin\alpha\cos\beta & \cos\alpha\cos\beta \end{bmatrix} \tag{2.14}$$

The rotation matrix is non commutative and therefore the order of multiplication is very important. Other properties of the rotation matrix are that rows and columns of R are orthogonal. Determinant of R is equal to 1.

Combining Equation (2.10) and (2.12), we get:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_u & s & u_o & 0 \\ 0 & \alpha_v & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2.15}$$

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = M_{intrinsic} * M_{extrinsic} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.16)$$

In the above equation the intrinsic and extrinsic matrices can be combined to form a 3*4 projection matrix:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.17)$$

2.1.4.3 Solving the P-matrix

Camera calibration helps in determining the extrinsic and intrinsic camera matrices. The method used in this project for calibration is based on the Direct Linear Transform (DLT) [18]. This method uses sets of targets which are fixed on a rigid frame of known geometry so that its 3D World co-ordinates are known. It estimates the mapping between calibration frame 3D co-ordinates and their Image co-ordinates to give the projection matrix.

The pinhole camera model shown in equation (2.15) can be written as:

$$\lambda * \overline{u_i} = K * \left[R|T\right] * \overline{\left[X_i\right]} \quad (2.18)$$

Where $\overline{X_i}$ = World co-ordinate

$\overline{u_i}$ = Image co-ordinate

The scaling term in above the equation will change with different cameras and as a result the number of unknown variables increases. In order to eliminate the scaling term, we can perform the cross product of each term in above equation:

$$\overline{\left[u_i\right]} \times \left(\lambda \overline{u_i}\right) = \overline{\left[u_i\right]} \times \left(K * \left[R|T\right] * \overline{\left[X_i\right]}\right) \quad (2.19)$$

where $[\bar{u}_i]\times(\lambda\bar{u}_i)=0$ and $(K^*\ [R|T])= P_{3*4}$

$$0 = \left[\overline{u_i}\right]\times\left(P_{3*4} * \left[\overline{X_i}\right]\right) \tag{2.20}$$

$$\left[\overline{u_i}\right]\times\left(\begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} * \left[\overline{X_i}\right]\right) = 0 \tag{2.21}$$

$$\left[\overline{u_i}\right]\times\left(P_{3*4} * \left[\overline{X_i}\right]\right) = \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \times \left(\begin{bmatrix} P_1^T \overline{X_i} \\ P_2^T \overline{X_i} \\ P_3^T \overline{X_i} \end{bmatrix}\right) = 0 \tag{2.22}$$

where P_i^T is the i_{th} row of the Projection matrix.

Consider an example of the vector cross multiplication $\overline{C} = \overline{A}\times\overline{B}$. This equation can be represented in matrix transform as $\overline{C} = [\sim A]\overline{B}$ where $[\sim A]$ is a skew symmetric matrix derived from expanding the terms of vector cross multiplication:

$$\begin{aligned} C_x &= A_y * B_z - B_y * A_z \\ C_y &= A_z * B_x - B_z * A_x \\ C_z &= A_x * B_y - B_x * A_y \end{aligned} \tag{2.23}$$

Rearranging in matrix form:

$$\overline{C} = \begin{bmatrix} 0 & -A_z & A_y \\ A_z & 0 & -A_x \\ -A_y & A_x & 0 \end{bmatrix}\overline{B} \tag{2.24}$$

Similarly, from equation (2.22), the vector $\bar{u}_i$ can be shown in skew symmetric form:

$$\begin{bmatrix} 0 & -1 & v_i \\ 1 & 0 & -u_i \\ -v_i & u_i & 0 \end{bmatrix} * \left(\begin{bmatrix} P_1^T \overline{X_i} \\ P_2^T \overline{X_i} \\ P_3^T \overline{X_i} \end{bmatrix} \right) = 0 \tag{2.25}$$

$$\begin{bmatrix} -P_2^T \overline{X_i} + v_i P_3^T \overline{X_i} \\ P_1^T \overline{X_i} - u_i P_3^T \overline{X_i} \\ -v_i P_1^T \overline{X_i} + u_i P_2^T \overline{X_i} \end{bmatrix} = 0 \tag{2.26}$$

$$\begin{bmatrix} 0^T & -\overline{X_i^T} & v_i \overline{X_i^T} \\ \overline{X_i^T} & 0^T & u_i \overline{X_i^T} \\ -v_i \overline{X_i^T} & -u_i \overline{X_i^T} & 0 \end{bmatrix} * \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} = 0 \tag{2.27}$$

In the above matrix, only the first two rows are linearly independent as the third row is a linear combination of the first two. For each point correspondence between 3D and 2D image co-ordinates in homogeneous coordinate representation, there are two linearly independent equations in P:

$$\begin{bmatrix} 0^T & -\overline{X_i^T} & v_i \overline{X_i^T} \\ \overline{X_i^T} & 0^T & u_i \overline{X_i^T} \end{bmatrix} * \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} = 0 \tag{2.28}$$

The above equation represents two linear equations for one target's 3D coordinates and its corresponding image points. In order to solve for the Projection matrix, we need more data points, so as to have enough equations to solve for the unknowns. All the linear equations from the different target points can be combined in matrix form as:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & -X_1 & -Y_1 & -Z_1 & -1 & v_1*X_1 & v_1*Y_1 & v_1*Z_1 & v_1 \\ X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & u_1*X_1 & u_1*Y_1 & u_1*Z_1 & u_1 \\ 0 & 0 & 0 & 0 & -X_2 & -Y_2 & -Z_2 & -1 & v_2*X_2 & v_2*Y_2 & v_2*Z_2 & v_2 \\ & \vdots & & & \vdots & & & \vdots & & & \vdots & \\ & \vdots & & & \vdots & & & \vdots & & & \vdots & \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & u_n*X_n & u_n*Y_n & u_n*Z_n & u_n \end{bmatrix} * \begin{bmatrix} P_{11} \\ P_{12} \\ P_{13} \\ P_{14} \\ P_{21} \\ \vdots \\ \vdots \\ P_{32} \\ P_{33} \\ P_{34} \end{bmatrix} = 0 \quad (2.29)$$

$$A_i p = 0 \quad (2.30)$$

For one image co-ordinate we have two linear equations so for n 3D-2D point correspondences, we will have A (2nx12). p is a 12 vector containing unknowns $P_{ij.}$ With 12 unknowns in p and each point correspondence giving 2 linear equations, at-least 6 point correspondences are required for its estimation. The obvious solution for the above equation is if p=0, which is not significant. In order to prevent the vector P from becoming zero, we need to add a constraint. Abdel-Aziz and Karara [18] used the constraint-like norm ||p|| =1. Due to the presence of noise in the image co-ordinates, the rank of matrix A will become 12 which lead to an over-determined system. An over-determined system is a system which has no unique solution that satisfies all the equations. We can find an approximate solution by minimizing ||Ap||=0 using the linear least squares method subject to the constraint ||p||=1.The above problem can be formulated as the minimization of, $f= \|Ap\|^2$ subject to the constraint, $g= \|p\|^2-1=0$.

Using the concept of Lagrange multiplier, we need to minimize the function:

$$L = f - \lambda g = \|Ap\|^2 - \lambda\left(\|p\|^2 - 1\right) \quad (2.31)$$

where λ is the nonzero Lagrange multiplier. Equivalently:

$$L = (Ap)^T (Ap) - \lambda\left(p^T p - 1\right) \tag{2.32}$$

Differentiating L with respect to p, and equating it to 0:

$$\frac{\partial L}{\partial p} = A^T Ap - \lambda p = 0 \tag{2.33}$$

$$A^T Ap = \lambda p \tag{2.34}$$

Differentiating L with respect to λ, and equating it to 0:

$$\frac{\partial L}{\partial \lambda} = p^T p - 1 = 0 \tag{2.35}$$

$$p^T p = 1 \tag{2.36}$$

Premultiply on both sides of equation (2.34) with p^T gives:

$$p^T A^T Ap = \lambda p^T p \tag{2.37}$$

$$(Ap)^T (Ap) = \lambda \tag{2.38}$$

$$\|Ap\|^2 = \lambda \tag{2.39}$$

This means that to minimize the above function (2.39), we need to minimize λ. Also, from equation (2.34), we observe that p is the Eigenvector that corresponds to the

least Eigen value λ of matrix A^TA . p can be computed by using Singular Value Decomposition (SVD):

$$SVD(A^TA) = UDV^T \tag{2.40}$$

where, U is orthogonal matrix whose columns are Eigenvectors of $AA^{T,}$ V is an orthogonal matrix whose columns are Eigenvectors of A^TA. D is a diagonal matrix with singular values in descending order on the diagonal, and P is a column vector which corresponds to last column of V= least singular value.

2.1.5 3D Co-ordinate Estimation Using Triangulation

In order to estimate the 3D co-ordinates of moving targets in World space, at least two cameras are required. One camera will determine the line from its optical center to the object point in World space while the second camera will determine the line passing from its optical center and intersecting with the first line in 3D space to give the exact position of the target. This method of determining 3D co-ordinates by using correspondences between two camera image points is called triangulation [21][22].

Once the projection matrix is determined for both the cameras by the method stated in Section 2.1.4.3 above, we have the following equations for the left and right cameras:

$$\begin{aligned} \lambda\overline{u_L} &= P_L\overline{X_i} \\ \lambda\overline{u_R} &= P_R\overline{X_i} \end{aligned} \tag{2.41}$$

$$\left[\overline{u_L}\right]\times\left(\lambda\overline{u_L}\right) = \left[\overline{u_L}\right]\times P_L\left[\overline{X_i}\right] \tag{2.42}$$

For the left camera we have:

$$\left[\overline{u_L}\right]\times P_L\left[\overline{X_i}\right]=0 \tag{2.43}$$

$$\begin{bmatrix} -P_L(2,:)^T + v_L P_L(3,:)^T \\ P_L(1,:)^T - u_L P_L(3,:)^T \\ -v_L P_L(1,:)^T + u_L P_L(2,:)^T \end{bmatrix} * \left[\overline{X_i}\right]=0 \tag{2.44}$$

Similarly, for right camera:

$$\left[\overline{u_R}\right]\times P_R\left[\overline{X_i}\right]=0 \tag{2.45}$$

Only first two rows are used from equation (2.44), Combining equations from left and right cameras in matrix form,

$$\begin{bmatrix} -P_L(2,:)^T + v_L P_L(3,:)^T \\ P_L(1,:)^T - u_L P_L(3,:)^T \\ -P_R(2,:)^T + v_R P_R(3,:)^T \\ P_R(1,:)^T - u_R P_R(3,:)^T \end{bmatrix} * \left[\overline{X_i}\right]=0 \tag{2.46}$$

which can be written as,

$$A_i \overline{X_i} = 0 \tag{2.47}$$

Unknown World co-ordinate $\overline{X}_i$ can be determined by decomposing matrix A via SVD.

As explained in Section 2.1.4.3, we need to minimize the above equation so that the solution is the eigenvector corresponding to least Eigen value of matrix A, and $\overline{X_i}$ is column vector which correspond to last column of V= least singular value.

2.2 Setup Optimization and Calibration of a Six-Camera Motion Tracking System

The six camera system used in the Motion Laboratory of Texas Scottish Rite Hospital was from Vicon. These cameras had built in software for estimation of 3D

location of targets. The 3D targets for this camera system were obtained from the hospital. The target triangulation algorithms described above was thus used only for the two-camera system. The Vicon cameras used has 16 megapixels of resolution and a full frame capture speed of120 frames/sec. These cameras sent 2D tracking information to the computer for reconstruction and labeling of the markers. They have 320 high power LEDs around the camera which gives an even spread of light across camera's field of view and 252 (NIR+IR) LED's having wavelength of 780nm and 850nm Sensor resolution was 4705 × 3456. After getting the 3D co-ordinates from the six-camera system, data was filtered using a Woltring filter

2.3 Co-ordinate Transformation between Two Camera Systems

For the purpose of comparison between two and six camera systems, it is extremely important that both systems are in same reference frame. Co-ordinate transformation was achieved by using an algorithm [27] [29] which is explained in more detail below.

Four targets were placed at the corners of table where subjects performed finger tapping, with two targets placed approximately 4 cm in front of hand and two at about 10cm distance from front two and 10 cm to the side so that these targets were not obscured by the subjects' arms. In this arrangement all four targets were visible in the field of view of both camera systems. 3D co-ordinates were estimated from both systems and least square fitting was performed on these data sets to get the Rotation and Translation matrices that transformed data from the VC camera system's reference frame to Vicon's reference frame.

This calculation starts by supposing that two camera systems see N targets with corresponding 3D co-ordinates that are each represented by 3*1 column vector V_i and

P_i for VC camera system and Vicon camera system respectively, i=1,2,3....N targets, such that they are related by:

$$P_i = R\,V_i + T + N_i \tag{2.48}$$

where R= 3*3 rotation matrix, T= 3*1 is the Translation vector, andN_i is a noise vector.

We need to find transformation R and T that will map 3D co-ordinates from two camera system to the Vicon system. This can be achieved by minimizing the least square error as shown in the following equation:

$$\sum^2 = \sum_{i=1}^{N} \left\| P_i - (R\,V_i + T) \right\|^2 \tag{2.49}$$

For the least squares solution, the data points V_i and P_i should have the same centroid. Using this constraint [28], new equations can be generated:

$$\begin{aligned} V &= \frac{1}{N} \sum_{i=1}^{N} V_i \qquad & V_{Ci} &= V_i - V \\ P &= \frac{1}{N} \sum_{i=1}^{N} P_i \qquad & P_{Ci} &= P_i - P \end{aligned} \tag{2.50}$$

where V and P are mean vectors of data points V_i and P_i , and V_{Ci} and P_{Ci} are variances around mean vectors V and P.

Equation (2.49) can be rewritten as:

$$\sum^2 = \sum_{i=1}^{N} \left\| P_{Ci} - \left(\hat{R} V_{Ci}\right) \right\|^2 \tag{2.51}$$

Expanding above equation:

$$\sum^2 = \sum_{i=1}^{N} \left(P_{Ci} - \hat{R}V_{Ci}\right)^T * \left(P_{Ci} - \hat{R}V_{Ci}\right) \tag{2.52}$$

$$\sum^2 = \sum_{i=1}^{N} \left(P_{Ci}{}^T P_{Ci} + V_{Ci}^T \hat{R}^T \hat{R} V_{Ci} - P_{Ci}^T \hat{R} V_{Ci} - V_{Ci}^T \hat{R}^T P_{Ci}\right) \tag{2.53}$$

The last term in above equation is same as $P_{Ci}^T \hat{R} V_{Ci}$:

$$\sum^2 = \sum_{i=1}^{N} \left\| P_{Ci}^T P_{Ci} + V_{Ci}^T V_{Ci} - 2 P_{Ci}^T \hat{R} V_{Ci} \right\|^2 \tag{2.54}$$

Minimizing above equation is equivalent to maximizing trace ($\hat{R}H$):

$$\sum^2 = \sum_{i=1}^{N} P_{Ci}^T \hat{R} V_{Ci} \tag{2.55}$$

where H is a correlation matrix defined by:

$$H = \sum_{i=1}^{N} V_{Ci} P_{Ci}^T \tag{2.56}$$

Let the singular value decomposition of the above equation be $H = UDV^T$. Then the optimal rotation matrix will correspond to $X = VU^T$. In order to obey with properties of rotation as described in Section 2.1.4.2, the determinant of X needs to be checked.

If determinant (X) =1; the solution is the Rotation matrix; $\hat{R} = X$.

If determinant (X) ≠1; X is a reflection.

When determinant of X is -1, it indicates the presence of noise in corresponding data sets, Therefore reflection is computed instead of rotation [28]. In such cases, rotation can be found out by $\hat{R} = V' U^T$ where V' is formed by [v1, v2, -v3]' v3 column corresponds to the singular value of H that is zero.

Once the rotation matrix is determined, translation will align the centroid of set P_i with the rotated centroid of set V_i, which can be determined by:

$$\hat{T} = P - \hat{R}V \tag{2.57}$$

2.4 Human Subjects and Motions Performed

Retroreflective targets were placed on the fingers and wrist of both hands of each subject. In all, there were 8 targets on fingers, two for each finger, excluding the thumb. The placement of the targets on each finger was at the nail and at the upper knuckle, nearest to the palm of the hand. Two targets were also placed on wrist of each hand, one on the side of the thumb and the other on the side of the pinkie. Motion tracking data were collected at the Texas Scottish Rite Hospital for Children (TSRHC) at Dallas as part of an ongoing study to assess the improvement in hand use of children with Cerebral Palsy (CP) after Constraint Induced Motion Therapy [33]. This study was approved under the UT Southwestern Institutional Review Board (IRB) protocol number STU 032012-001. A subset of the data collected in this study for motion tracking was analyzed in this work. More specifically, two healthy children age of 11 were measured twice while performing finger tapping with about two weeks having elapsed between the first and the second measurement. In addition, three children with CP were measured before and after two weeks of CIMT. These children were previously assigned a Manual Ability Classification Scale (MACS) score. Two of the children were MACS I and was MACS II with increasing impairment severity being implied by higher MACS numbers.

2.4.1 Protocol for Finger Tapping

A PowerPoint presentation was made which instructed an adult subject to tap on the table while ensuring that the wrist was on the table at all times. As per the instructions from a PowerPoint presentation, subjects performed finger tapping at their own pace

which was near a frequency of 1 Hz for all subjects. The protocol included eight epochs of tapping for 25 seconds with a rest period of 15 seconds in between the tap and a 3 minute period of rest before just before the taping began. The arrangement of two camera systems is shown in Figure 2-5 and Figure 2-6:

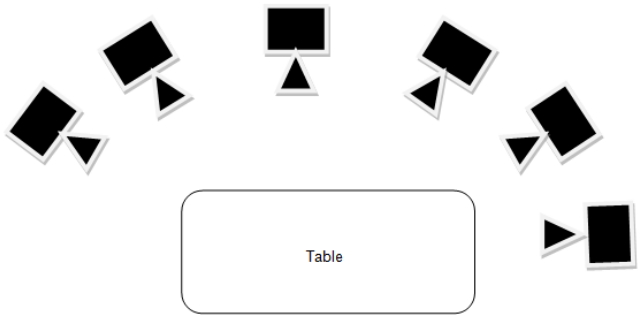


Figure 2-5: Top view showing the arrangement of the six-camera setup placed closer to the ceiling at the Motion Lab at Texas Scottish Rite Hospital

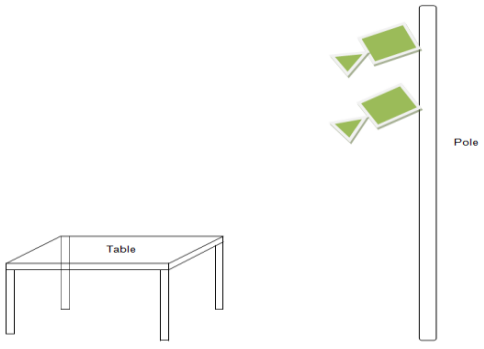


Figure 2-6: Side view showing the arrangement of the two-camera setup in green color

2.4.2 Motion Metrics Measurements

Amplitude and velocity measurements were taken from each subject's affected hand. Only one target from the affected person's hand, the one on the index nail, was selected for computing motion metrics with both camera systems. The index nail target of the right hand was also used for computing the motion tracking metrics of the control subjects, who were both right-handed. One epoch which had 15 taps were used to determine the amplitude for one target. Average velocity was determined by calculating the velocity over an epoch, 15 taps. Thus, average tapping amplitude and average velocity were calculated for the two measurement times and for both the systems. The instantaneous finger velocity was calculated at four different points in time corresponding to the beginning, ascending midpoint, descending midpoint, and end of the tap (points V1, V2, V3 and V4, respectively. In Figure 2-5 the instantaneous velocity at each of these four time points was computed only for one tap of the index finger.

In this work the hypothesis was tested that motion metrics can be used as a surrogate measure of improvement in performing a finger taping task after CIMT treatment. To that end, T-tests were performed between the pre-and post-treatment measurements of the average amplitude, average velocity and instantaneous velocities during finger tapping. Specifically, the two tail T-test was performed between the two measurements to test the Null hypothesis which states that the magnitude of a given motion metric has not changed between the two measurements.

A difference between the pre- and post-treatment metric values was considered significant for p-values of less than 0.05.

Table 2-1: Summary of Data for which the improvement of performance was calculated from amplitude, average velocity and instantaneous velocity for both camera systems. T indicates the Trials

Two camera system			Six camera system	
Subject	T1	T2	T1	T2
MACS-1	M-1a	M-1a	M-1a	M-1a
			M-1b	M-1b
NORMAL	N1	N1	N1	N1
	N2	N2	N2	N2

2.4.2.1 Instantaneous Velocity

As seen in Figure 2-7, the entire tap profile was divided into four time intervals and the instantaneous velocity was calculated between those tap intervals using Equation (2.58):

$$V_{Inst} = \frac{\Delta s}{\Delta t} = \frac{V_2 - V_1}{t_2 - t_1} \tag{2.58}$$

where Δs represents the displacement. V_2 and V_1 are the final and initial 3D co-ordinates in the Y direction. Δt is the time difference between two data points V_2 and V_1

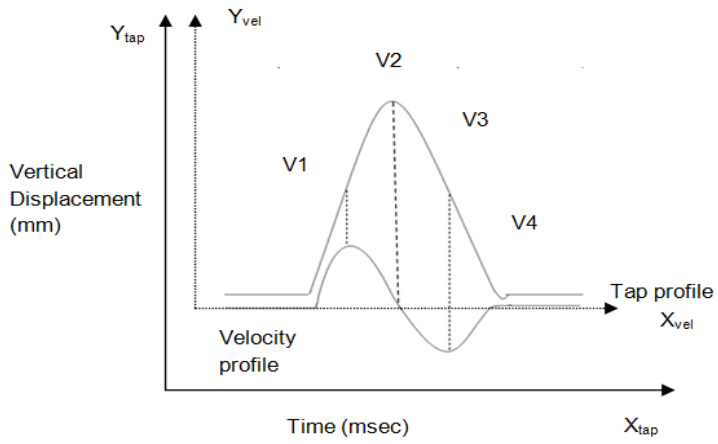


Figure 2-7: Tapping amplitude profile (top) and velocity profile (bottom), i.e. the derivative with respect to time, showing the four time points at which instantaneous velocity was measured

For the purpose of comparisons between the two camera systems, the average velocity, average amplitude was calculated for individual taps with both camera systems. The mean and standard deviation values for each motion metric were calculated from 15 taps, i.e. one tapping epoch, which was the first of eight that each subject performed during each measurement session.

For comparison purposes, instantaneous velocities were estimated at four time instants V1 was selected as the point of maximum peak in positive direction. V3 was selected as point of maximum peak in negative direction. V2 was estimated as the positive data point just before the zero crossing in the descending tap and V4 was estimated as the positive data point just after the zero crossing during the ascending part of tap. These calculations were performed after fitting a second order spline to the

instantaneous velocity data using Matlab function csaps for both the camera systems as seen in Figure 2-8

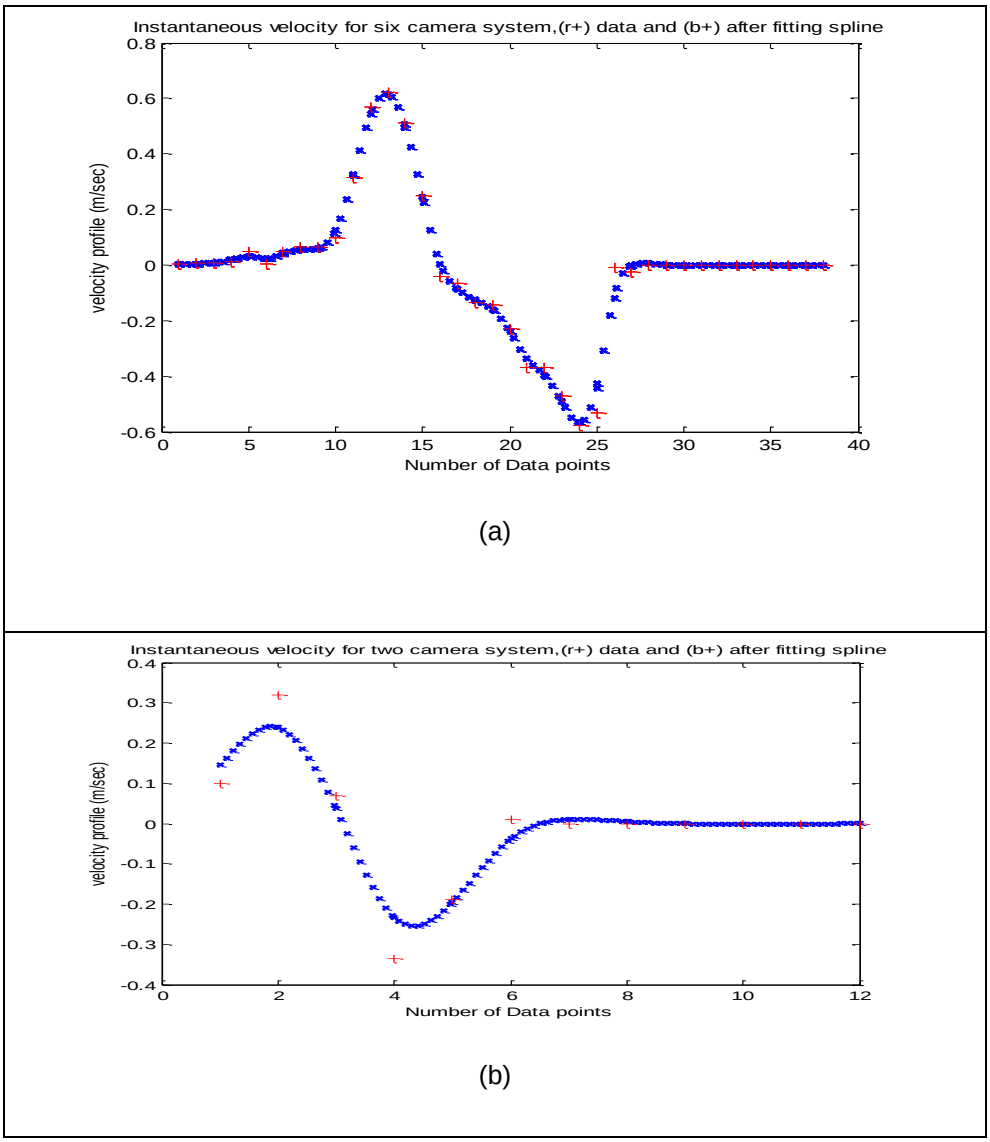


Figure 2-8: : Instantaneous velocity profile for (a) the six-camera system and (b) the two-camera system, with blue crosses indicating the velocity data after spline fitting and the red cross symbols indicates the raw velocity data profile

2.4.2.2 Calculation of Absolute and Percent Error of the Motion Metrics Determined by the Two-Camera System

Seven data sets were used in order to compare the performance of the two-camera system relative to the six-camera system. These data consisted of two trials for subject M-1a, two trials for N-1, two trials for N-2 and one trial for M-1b. For each subject trial, the average amplitude, average velocity and instantaneous velocity were calculated for one target on the index finger over 15 taps with both systems

The percent error relative to the six-camera system was calculated by using Equation (2.59):

$$\text{Percent error} = \left(\frac{S_A - T_A}{S_A} \right) * 100 \tag{2.59}$$

where S_A was the tapping amplitude of Index finger target measured with the six-camera system and T_A the tapping amplitude of index finger target measured by the two-camera system.

Though as we will see further below, a large contributor to errors in motion metric measurements by the two-camera system was its lower data sampling rate and its increased susceptibility to target occlusion relative to the six-camera system, it was worthwhile to investigate if additional errors contributed to these differences even if the targets were static. There were two main sources of error that could contribute to this comparison between the two camera systems: (a) their 3D target triangulation accuracy, and (b) any errors in the co-ordinate transformation from the two-camera to the six-camera system.

The accuracy of 3D target localization was determined by using Equation (3.1). The error in estimating the transformation parameters was calculated by using equation (2.60):

$$E_{transformation} = \| X_{six} - (R X_{two} + t) \| \quad (2.60)$$

where X_{six} was the 3D vector corresponding to X,Y,Z co-ordinates of the four targets placed on the table where the subjects were tapping in the six-camera system's frame of reference, X_{two} was the 3D vector corresponding to X,Y,Z co-ordinates of those four table targets in the two-camera system's frame of reference, and R and t were rotation and translation matrices used to transform the 3D co-ordinates from the two-camera to the six-camera system's frame of reference.

Table 2-2: Summary of subjects measured per trial for the comparison between the two camera systems

Two v/s six camera comparisons		
Subject	T1	T2
MACS-1	M-1a	M-1a
	M-1b	
NORMAL	N1	N1
	N2	N2

Chapter 3

Experiments and Results

The total error in motion analysis can accumulate from several different sources that relate to the performance specifications of a particular camera setup, its calibration procedure, and the choice of retroreflective marker placements for tracking specific body motions. The nature of these errors can vary quite a lot depending on the camera systems, calibration procedures, and tracking algorithms used. In this chapter we discuss the main sources of error that were identified for the two-camera system used for our experiments. Firstly, the accuracy of localizing a retroreflective marker in 3D is a function of the relative position and angle of the cameras [34]. For example, for our two-camera system if the two cameras are too close to each other the parallax angle of a target may be too small, which will reduce the localization accuracy in the depth dimension, z. Even for a perfectly stationary target the accuracy of localization can be further degraded by stochastic image noise, which affects the center of gravity estimation for the target, and any non-linear mapping response of the camera [36]. Secondly, accuracy localization depends heavily on a good camera calibration procedure [35]. In this work a multiple target setup was used where the known 3D coordinate of each target was used to estimate camera parameters. Clearly, any differences between the actual and assumed positions of each target could translate to localization errors in subsequent experiments relying on that calibration. Finally, the frame rate of the cameras used in this work decreased with increasing numbers of targets in the field of view and could not exceed 30 Hz. This image capture rate may not be sufficient to capture the trajectory details of finger and arm motions with enough accuracy to quantify metrics like the instantaneous displacement, amplitude and velocity profile of a tapping finger. In the sections below we quantify the relative magnitude of each one of these localization accuracy errors for our

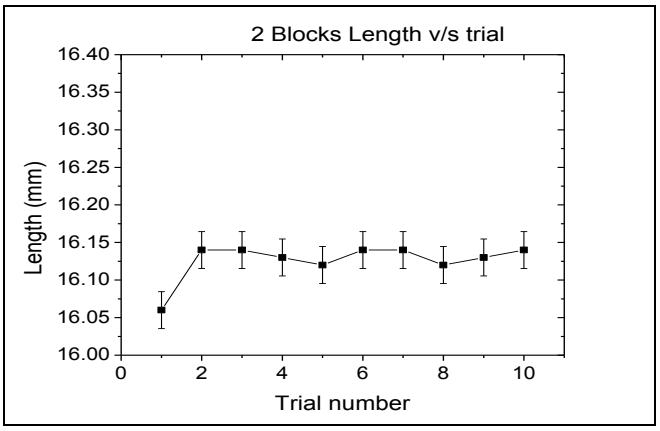
two-camera motion tracking system. Then in the next Chapter we explore how these errors affect the quantification of hand motion metrics for two children with CP and two age-matched controls.

3.1 Experimental Tests for the Two-Camera System Performance

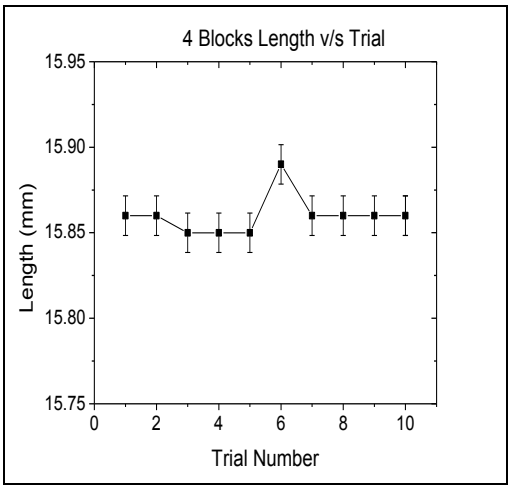
In order to estimate the target localization accuracy of our two-camera (VC) system various tests were performed at UTA before it was transported to a hospital environment for measurements on children with CP. Firstly, the VC system was calibrated using a 3D Lego calibration model along with the P-matrix estimation described above. Secondly, a number of different camera arrangements were explored for their relative performance in localizing accurately the calibration targets. These experimental steps are described here below.

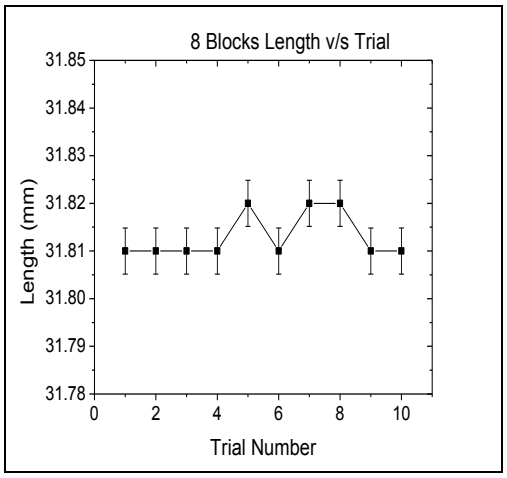
3.1.1. Lego Block Measurement Accuracy

A calibration frame was made with blocks of Lego. A Pittsburgh digital Vernier caliper was used for measuring the length of Lego blocks. Its accuracy was 0.03 mm. In order to quantify the repeatability of error in using the calipers three sets of Lego block measurements were repeated ten times. In the first set of measurements, two blocks were kept close to each other and the length of joined blocks was measured. In the second and third sets of measurements, four and eight blocks were joined, respectively, and their combined lengths were measured. While taking measurements, both the jaws of the calipers and the Lego block measuring faces were cleaned so as to remove any dust particles from them. The calipers were then set to zero and then expanded to bracket the length of the Lego block while ensuring that it is not tilted, which could affect measurements by up to 0.01mm



(a)





(b)

(c)

Figure 3-1: Length measurements of (a) two blocks of Lego (b) four blocks and (c) eight blocks, connected to each other on their short side, repeated for ten trials

A square shaped Lego building plate was glued onto a wooden plank. The wooden plank was affixed to a platform by use of double sided tape. Length measurements of the two, four and eight block Lego assemblies were performed while these were affixed horizontally onto the building plate .The mean and standard deviation of the ten trials for each data set was calculated and shown in Table 3-1 and are also

plotted in. Figure 3-1 (a)-(c) for measurements on two, four and eight Lego blocks, respectively.

Table 3-1: Lego blocks measurements

Measurement of Lego using Vernier (mm)			
	2 blocks	4 blocks	8 blocks
1	16.06	15.86	31.81
2	16.14	15.86	31.81
3	16.14	15.85	31.81
4	16.13	15.85	31.81
5	16.12	15.85	31.82
6	16.14	15.89	31.81
7	16.14	15.86	31.82
8	16.12	15.86	31.82
9	16.13	15.86	31.81
10	16.14	15.86	31.81
Mean	16.13	15.86	31.81
Std dev	0.03	0.01	0.01

The standard deviations shown in Table 3-1 are a combination of subjective error and caliper measurement error. These measurements were carried out to determine the accuracy with which Legos could be measured. In order to avoid this error, one block was measured horizontally and vertically for 20 times and the average of those lengths was determined. The average length and width were used to set the 3D coordinates of the calibration frame.

3.1.2. Stationary Target Localization Accuracy for the Two-Camera System

The noise limited and intrinsic linearity error of the two-camera system was measured using a static target. The error estimated in this procedure was in pixels. In order to convert it to mm, Horizontal Field of View (HFOV) and Vertical Field of View (VFOV) mm/pixel conversion factors were calculated. The camera was placed in front of

a single Lego block with a retroreflective target attached to it at a distance of 80" from camera's centre to table. Using a laser level (STRAIT LINE X-3), the camera was aligned with the X and Y axes of the target. In order to check for linearity of camera system, it was essential that CCD's X and Y axis are aligned with target's X and Y axis. As a result of this alignment, it was possible to compare the x pixel value of camera and the horizontal distance by which the target was moved by using caliper. Using double sided adhesive tape, the laser leveler was fixed above the camera. The camera was tilted in horizontal X direction until the horizontal bubble vial over the level showed that the bubble was centered. This process helped in aligning camera's Y axis. For alignment along X axis, the camera was tilted up and down until the bubble vial along Z axis showed that the bubble was centered. The overall setup for this alignment process is shown in Figure 3-2.

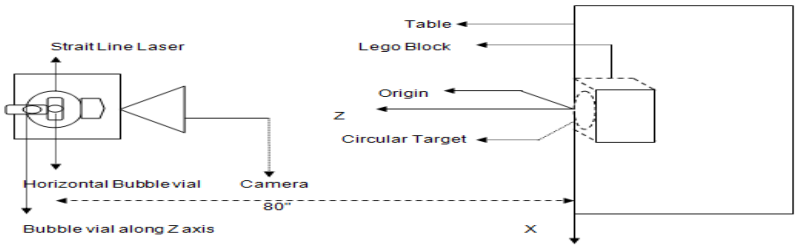


Figure 3-2: Top View of the camera and of the Circular Target on the Lego. The Y axis of the Circular Target is coming out of the paper from origin. Strait Line laser shows the horizontal bubble in the centre indicating that the camera was aligned to the Y axis of target and the bubble along the Z axis indicated that camera plane was aligned to the X axis

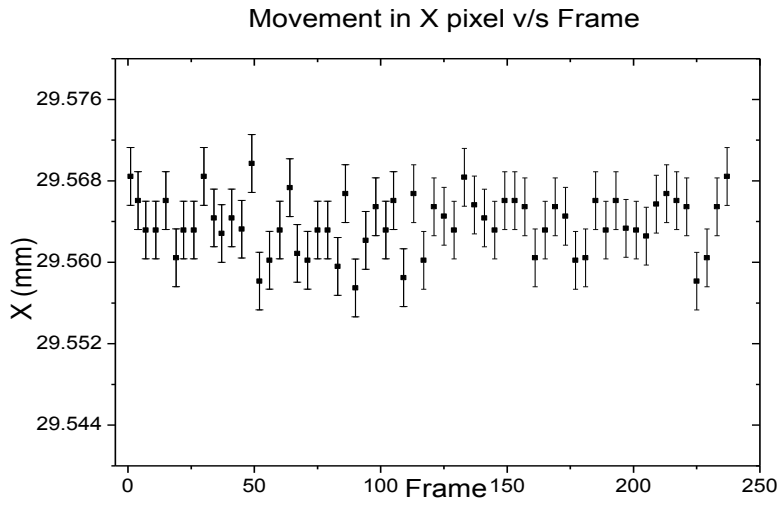
Aligning these axes helped in determining the conversion factors just by placing two targets at a known horizontal distance in front of camera. Two targets on Lego blocks were affixed onto the Lego building plate at a horizontal distance of 66 mm (2.6") and at a distance of 87" from the camera. Using the X pixel values of two targets and the known distance of 66mm, the HFOV conversion factor was calculated as 0.99 mm/pixel. Two targets were then kept at a vertical distance of 101mm (4"), and by using the Y pixel values reported by the camera and the known distance of 101mm; the VFOV conversion factor was calculated to be 0.899 mm/pixel.

3.1.2.1 Target Localization Error due to Stochastic Image Noise

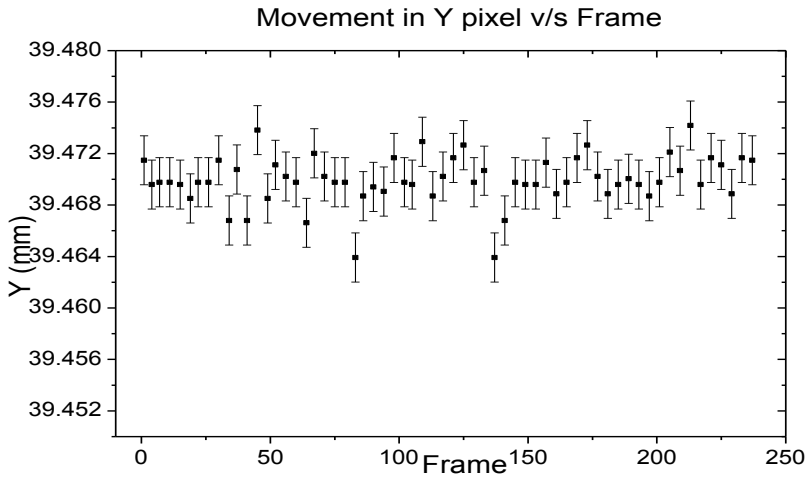
For determining the noise limited localization error of the camera, only one retro-reflective target on a Lego block was kept in front of the camera at 80" as described above. This subsection is concerned with the localization error in the projection image of each camera. Subsequent subsections describe the localization error in 3D after information from the two cameras is used for triangulation.

For the noise limited error in projection images it is assumed that the center of gravity of the target is affected by the number of photons recorded per frame. Due to stochastic variation in the number of detected photons the resulting center of gravity of the target may vary from frame to frame. It is possible that minute vibrations of the floor that propagated to the calibration target and the cameras could increase this variability. The amplitude of such vibrations was not measured though. Measurements of the single stationary target ware taken for 250 frames. The measurement of same target was made for 1 minute and was repeated for 8 times, with a rest period of 5 minutes in between. The cameras reported X and Y pixel co-ordinates of the target at 33 millisecond intervals, i.e. at a rate of 30 Hz. Using the pre-calculated conversion factors from the two-target procedure described above, pixel values were converted to mm. The X pixel value had

mean of 29.56 mm and standard deviation of 0.03 mm (data plotted in Figure 3-3). The Y pixel movement over 250 frames had a mean of 39.47 mm and standard deviation of 0.002 mm.



(a)



(b)

Figure 3-3: Scatter plot of static target for (a) pixel movement in the X and (b) Y directions pixel over 250 frames

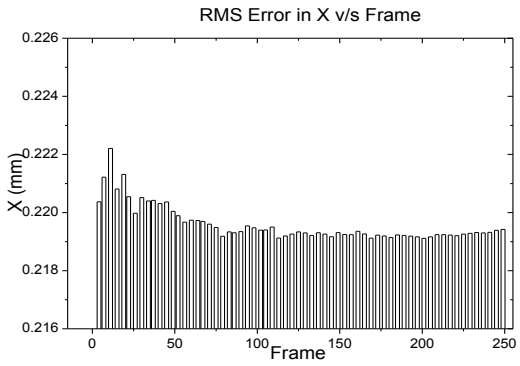
The measurements were performed for both the cameras of the two-camera system. The results obtained with each camera were pretty much the same. As seen from the above Figures, the errors associated with the two-camera system are very small as they correspond to sub-pixel dimensions.

3.1.2.2 3D Localization Error for a Static Target

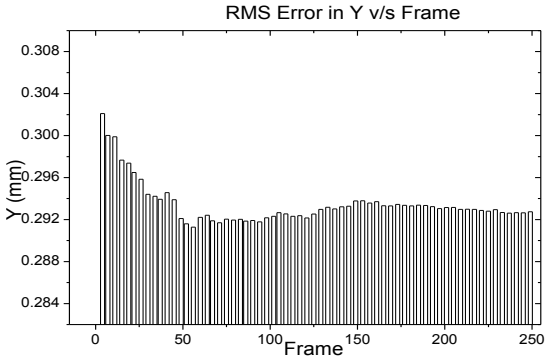
The two camera system was employed in order to estimate the target localization error in 3D World co-ordinates using the triangulation algorithm described above. The calibration frame of 80 targets described in the Methods Section was used for estimating the difference between calculated and measured 3D co-ordinates at different locations within the field of view of the two cameras. The two cameras were both attached on a pole at height difference of 1 ft 4". The pole was at a distance of 80" from the origin of calibration frame. The lower camera was at height of 3ft 10" from the floor. The calibration frame was centered to the axis of the pole holding the cameras and was facing the cameras head on so that it was visible in the field of view of both cameras. After calibration was performed and the 3D co-ordinates of all 80 targets were estimated with triangulation the RMS error in X, Y and Z with respect to the known World co-ordinates of each target was computed. The estimated mean and standard deviation was calculated as a function of increasing number frames being integrated to a single time-averaged image as shown in Figure 3-4. The RMS error and its standard deviation for 250 frames was calculated by starting with two frames and augmenting the values for each additional frame with the accumulated value of previous frames

Table 3-2: RMS error and standard deviation calculated for 3D co-ordinates for 250 frames

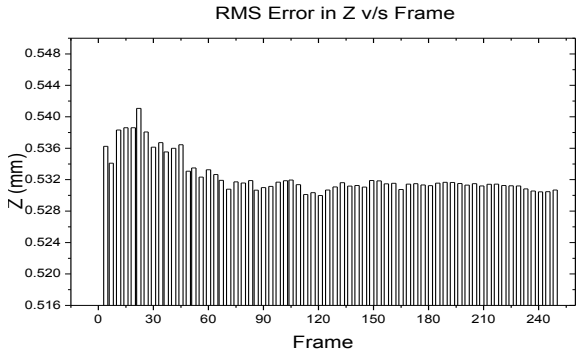
(mm)	X	Y	Z
Mean RMS error	0.219	0.293	0.532
Mean standard deviation	0.22	0.295	0.536



(a)



(b)



(c)

Figure 3-4: Bar plot of 3D World co-ordinate RMS error in (a) the X axis, (b) the Y axis, (c) the Z axis

As seen in Figure 3-4, as the number of frames increases the error reduces initially and the plateaus. It should also be mentioned that for the above plots, data was taken for 250 frames over 14 seconds. Thus, when tracking 80 targets, the frame rate was reduced to17 frames/sec from 30 frames/sec for a single target. This reduction in speed is attributed to limitations in the computational capacity of the camera hardware to perform real-time computation of the target center of gravity 2D coordinates before streaming the information to the computer disc for storage.

3.1.2.3 3D Localization error as a Function of Target position in the Field of View.

Using the same setup described in above section, the 3D Euclidean error was measured for each of the 80 target points using the formula:

$$E = \sqrt{\left(\left(X_i - X_i^{'}\right)^2 + \left(Y_i - Y_i^{'}\right)^2 + \left(Z_i - Z_i^{'}\right)^2 \right)} \tag{3.1}$$

where (Xi, Yi, Zi)= Wcalculated represents the World co-ordinates that were obtained after triangulation from the projection matrices of both cameras, and (Xi', Yi', Zi')= Wmeasured are the World co-ordinates measured by using the Vernier calipers and tape measure.

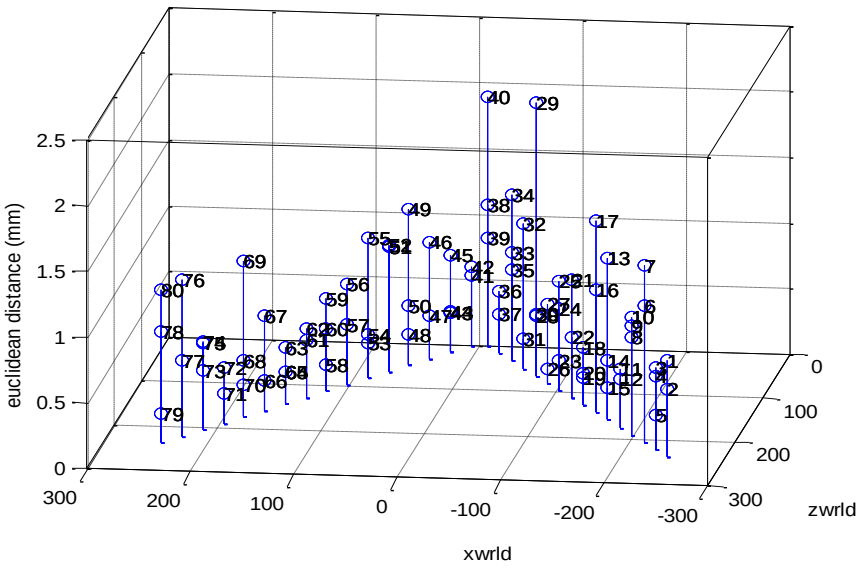


Figure 3-5: Stem Plot for Euclidean errors plotted with respect to X and Z World co-ordinates. The target number is showing next to the 'o' symbol and the height of the stem indicated the magnitude of the error.

As seen in Figure 3-5 the largest error magnitudes of 1.90 mm and 2.08 mm were measured for the marker positions 40 and 29, respectively, that were furthest away from the cameras indicating that 3D localization error increases as the depth of field increases due to the reduced parallax angle for deeper targets. Many targets located closer to the centre of the field of view with X positions from about 100 to -150, had considerably less error.

3.1.2.4 Checking the Co-ordinate Mapping Linearity for the Two-Camera System

Some camera systems may have nonlinear mapping between X, Y Word co-ordinates and X, Y pixel data. In order to check the linearity of response of our two-camera system, we mapped the response of the cameras when a target was shifted incrementally by calibrated amounts. Specifically, a single target on a Lego block was

mounted onto a sliding caliper jaw with double sided adhesive tape and movement of caliper allowed the target to be traversed through distance of approximately 50mm with an accuracy of 0.03mm. The caliper was mounted on the Lego platform with horizontal lines were drawn to ensure that the caliper axis, while lying flat on the table, was aligned with the X axis of the World co-ordinate system. The caliper jaw was moved incrementally to fourteen different distances as the separation between the moving and stationary jaws of the caliper was increased.

At each distance, measurements were taken for 3 minutes with the two-camera system at a frame rate of 30Hz resulting in 6000 frames captured per distance. These measurements were then used to compute mean and standard deviation values for the X, Y pixel values (Table 3-3). A best fit line was plotted on the plot of caliper distance moved in mm versus the X pixel value that the camera assigned to that position. The R-squared value of >0.99 indicates high linearity in the mapping of the X coordinates with this two-camera system (Figure 3-6).

Table 3-3: Caliper distance moved horizontally along with the X, Y pixel values and error from a linear fit

Dist (mm)	Mean X (pixels)	Mean Y (pixels)	Std X (pixels)	Std Y (pixels)	Predicted X (pixels)	Difference (pixels)
17.61	225.93	412.7	0.03	0.03	225.92	<0.01
18.78	224.70	412.78	0.02	0.03	224.76	-0.06
20.89	222.65	412.84	0.03	0.03	222.66	-0.01
22.79	220.73	412.91	0.04	0.04	220.77	-0.04
24.77	218.78	413.00	0.04	0.04	218.80	-0.02
28.21	215.40	413.05	0.03	0.03	215.38	0.01
30.57	213.00	413.09	0.02	0.04	213.04	-0.04
39.27	204.39	413.08	0.03	0.03	204.39	<0.01
47	196.71	413.05	0.03	0.05	196.71	<0.01

Table 3-3- continued

54.95	189.31	413.09	0.03	0.02	188.80	0.50
64.39	179.37	413.10	0.02	0.02	179.42	-0.04
68.34	175.39	413.24	0.02	0.02	175.5	-0.11
71.5	172.29	413.23	0.02	0.03	172.36	-0.06
73.38	170.37	413.28	0.01	0.02	170.49	-0.12
					mean	0.002
					std	0.15

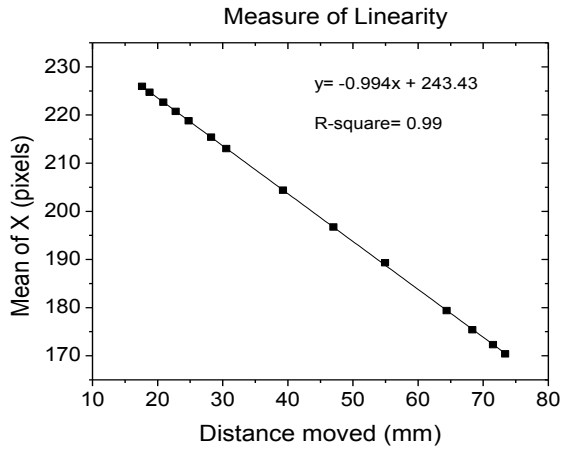


Figure 3-6: Linear plot of the horizontal distance that the target was moved with calipers versus the X co-ordinate pixel value that the camera produced after averaging over a three minute acquisition. The R-square value shows goodness of fit

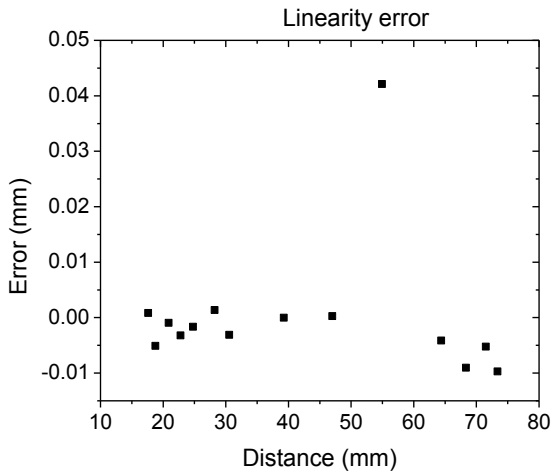


Figure 3-7: Mean of Residual error from to be linear fit was found to be <0.01 mm

The high level of linearity of camera response is also seen in Figure 3.7 where, with the exception of one outlier the linear fit error was <0.01 mm. The response for the other camera was found to be linear too.

3.2. Two-Camera Setup Geometry Optimization

In order to determine the 3D position of targets accurately with our two-camera system during finger tapping, six different arrangements for this setup were implemented in the laboratory. Before each set of measurements, the calibration frame with 80 targets was used to compute the P-matrix (intrinsic and extrinsic parameters) of each camera setup. In all six arrangements the two cameras were attached to a single pole, at different horizontal and vertical displacements relative to each other, which was placed at a horizontal distance of 87" from the front edge of the table where calibration frame was kept. All distance measurements were taken with the help of a 25 foot measuring tape. The specification of the six different two-camera setups is shown in Table 3-4. The

vertical distance between the two cameras was measured relative to the top surface of the heat sink of each camera.

After performing the calibration for each camera setup, finger tapping data measurements were performed following the finger tapping protocol described in the Method subsection. Also, for all measurements with the different setups the same control subject performed the finger tapping. This helped in comparing any target occlusions that occurred for different setups as the variability of finger tapping was reduced since it was always the same subject.

Although there was no standard measurement for comparison and determination of errors or accuracy of two-camera system, the optimal two-camera setup was selected on the basis of the magnitude of the calibration error for the 80 stationary targets and also comparing the frequency of occurrence of target occlusions while performing finger tapping. From use of the calibration target frame, an average error in 3D co-ordinates for X, Y and Z was estimated (Table 3-5). Also, changes in the number of X, Y and Z co-ordinates recorded per image frame were used to determine occlusions.

Table 3-4: Different arrangements of the two-camera setup with specific distances where the two cameras are indicated as A and B

	Setup 1	Setup 2	Setup 3	Setup 4	Setup 5	Setup 6 (existing)
B-A height difference (Vertical Distance)	3 ft	3 ft	3ft	2 ft	2ft	1 ft 4"

Table 3-4- continued

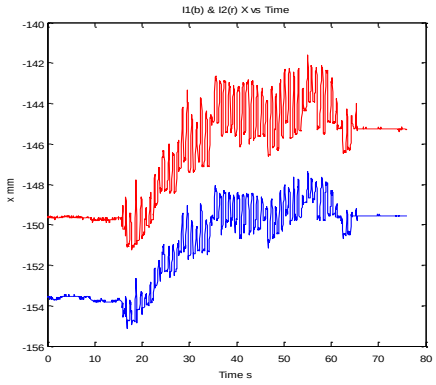
B-A Horizontal Distance	~ 2-3" (almost inline)	~ 7"	1ft 10"	~ 2-3" (almost inline)	11"	~ 2-3" (almost inline)
Lower camera to Table distance	1 ft 2"	1 ft 2"	1 ft 2"	2 ft 2"	2 ft 2"	2 ft 10"

The purpose of keeping the cameras on a pole was to ensure that cameras were well above the level of the table where finger tapping would occur so that all the targets were within field of view of both the cameras, while reducing the probability of target occlusions that could happen if those cameras were near the level of the tapping hand. The different setups were then designed to keep cameras at varying vertical distances from the table where the subject would perform finger tapping. Different setups were implemented starting from the least distance between the table and the lower camera to the highest distance between the two. Cameras positions were varied in the horizontal direction too, though this was not an exhaustive search of all possible two-camera setups, which would be prohibitively time consuming.

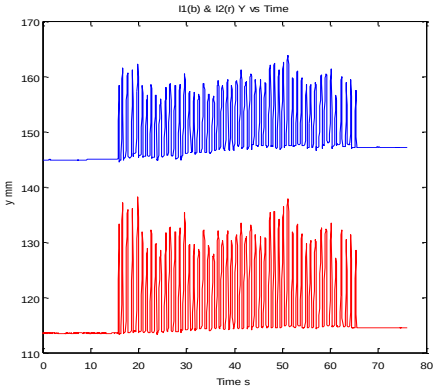
Figures 3-8 to 3-13 show the camera setup pictures and a sample X, Y, Z co-ordinate of the index finger nail and knuckle targets with respect to time.

Setup 1

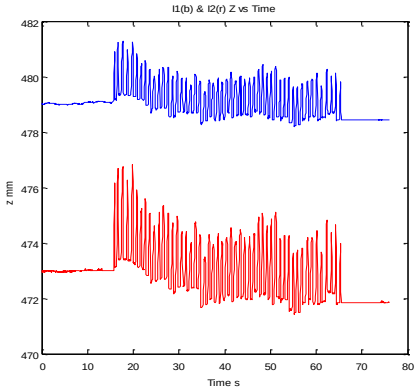
(a)



(b)



(c)

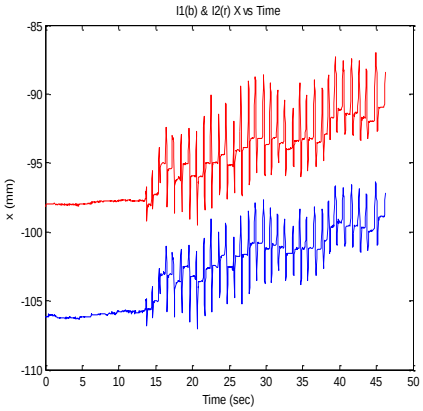


(d)

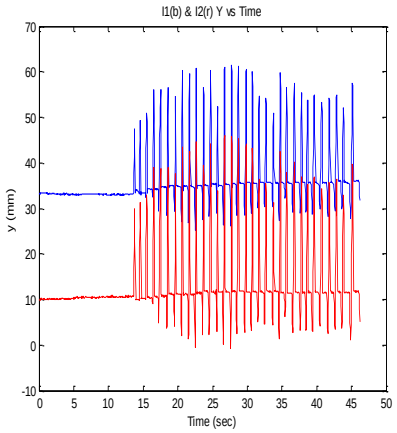
Figure 3-8: Arrangement of cameras in Setup 1 showing (a) the picture of the camera setup (b) the X co-ordinate for the index finger nail target shown in blue and the knuckle target shown in red with respect to time. (c) The Y co-ordinate and (d) the Z co-ordinate for the same targets

Setup 2

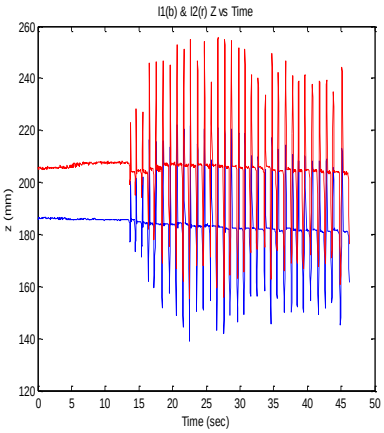
(a)



(b)



(c)

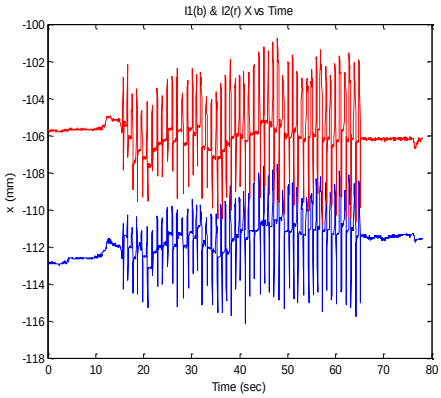


(d)

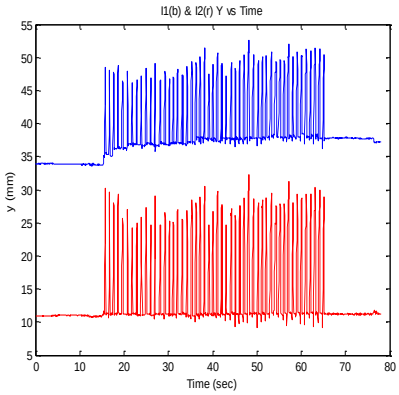
Figure 3-9: Arrangement of cameras in Setup 2 showing (a) the picture of the camera setup (b) the X co-ordinate for the index finger nail target shown in blue and the knuckle target shown in red with respect to time. (c) The Y co-ordinate and (d) the Z co-ordinate for the same targets

Setup 3

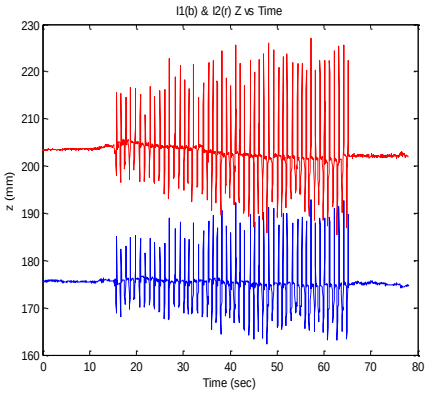
(a)



(b)



(c)

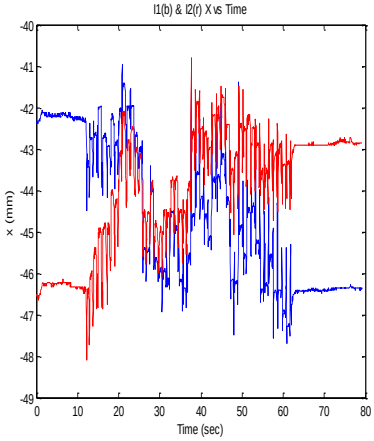


(d)

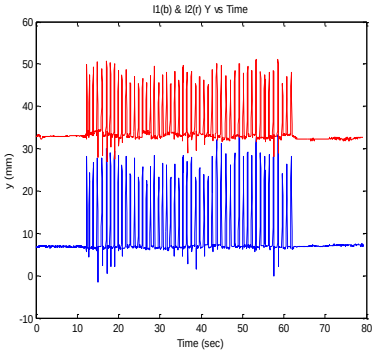
Figure 3-10: Arrangement of cameras in Setup 3 showing (a) the picture of the camera setup (b) the X co-ordinate for the index finger nail target shown in blue and the knuckle target shown in red with respect to time. (c) The Y co-ordinate and (d) the Z co-ordinate for the same targets

Setup 4

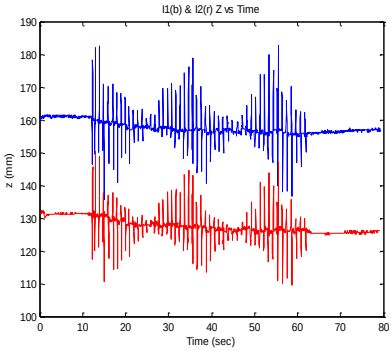
(a)



(b)



(c)

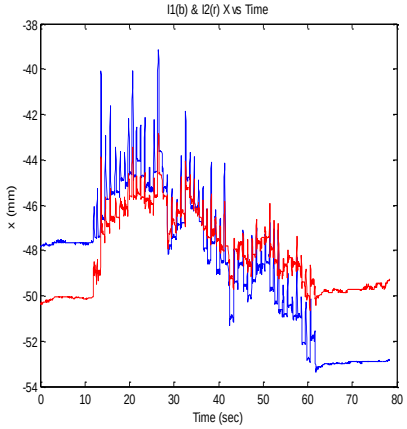


(d)

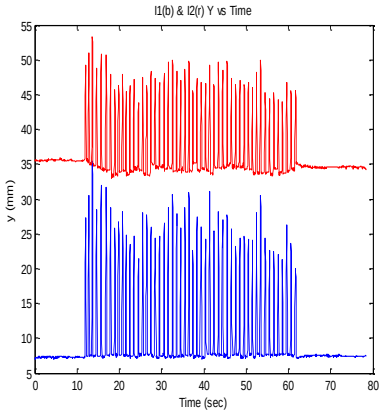
Figure 3-11: Arrangement of cameras in Setup 4 showing (a) the picture of the camera setup (b) the X co-ordinate for the index finger nail target shown in blue and the knuckle target shown in red with respect to time. (c) The Y co-ordinate and (d) the Z co-ordinate for same targets

Setup 5

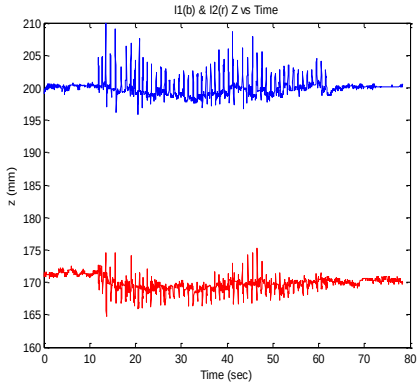
(a)



(b)



(c)

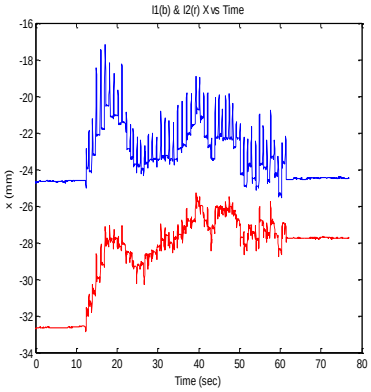


(d)

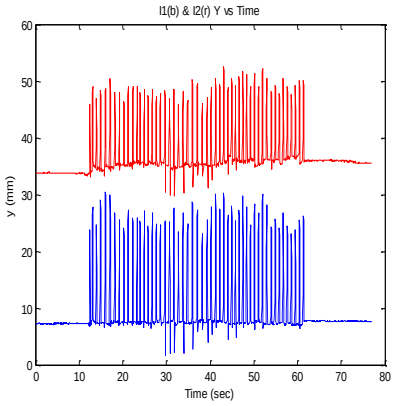
Figure 3-12: Arrangement of cameras in Setup 5 showing (a) the picture of the camera setup (b) the X co-ordinate for the index finger nail target shown in blue and the knuckle target shown in red with respect to time. (c) The Y co-ordinate and (d) the Z co-ordinate for the same targets

Setup 6

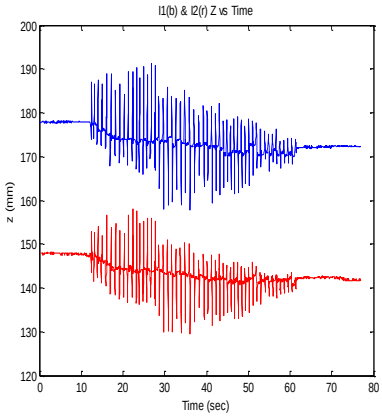
(a)



(b)



(c)



(d)

Figure 3-13: Arrangement of cameras in Setup 6 showing (a) the picture of the camera setup (b) the X co-ordinate for the index finger nail target shown in blue and the knuckle target shown in red with respect to time. (c) The Y co-ordinate and (d) the Z co-ordinate for the same targets.

Table 3-5: Average 3D Error in X, Y, Z co-ordinates of the 80 targets in the stationary calibration frame for the six different two-camera arrangements tested in this work

Average Error (mm)	Setup1	Setup2	Setup3	Setup4	Setup5	Setup6
X	0.16	0.18	0.18	0.14	0.15	0.15
Y	0.19	0.19	0.20	0.31	0.30	0.21
Z	0.30	0.35	0.29	0.46	0.46	0.38

Table 3-6: RMS error in 3D X, Y, Z co-ordinates of the 80 targets in the stationary calibration frame for the six different two-camera arrangements tested in this work

RMS error (mm)	Setup1	Setup2	Setup3	Setup4	Setup5	Setup6
X	0.2	0.24	0.23	0.17	0.18	0.19
Y	0.25	0.24	0.26	0.41	0.37	0.26
Z	0.38	0.44	0.35	0.61	0.58	0.47

As seen from Figure 3-8(b) and Figure 3-9(b), the values for the X co-ordinate become more negative with respect to time because the subject shifted their hand in left direction for the first two setup trials. However, this subject was more stable during rest of the setups.

As seen from Table 3-5, Setups 1-3 and 6 had similar errors, corresponding to RMS values of 0.2mm in X, 0.25 mm in Y and 0.3 mm in Z. For this one adult subject performing finger tapping no two targets were running together in Setup 1 as seen from Figure 3-8. Setup 1 was found to be accurate but it was way too close in distance with the table and this is not applicable in the clinical environment where children feel more comfortable with the cameras further away. If we compare Setup 2 and Setup 3 which

are at same vertical distance, but at different horizontal distances with Setup 3 being more apart, we observed that as the horizontal distance between cameras increased, finger and knuckle targets were running together as seen in Figure 3-9(d) and Figure 3-10(d).

Setups 4 and 5 had somewhat higher RMS errors of 0.61 mm and 0.58 mm in Z, 0.41 mm and 0.37 mm in Y respectively. Also, from Figure 3-11(b) and Figure 3-12(b), it was seen that the X coordinate was not resolved for the two targets at the nail and knuckle positions.

Selection of optimal camera position was a tradeoff between stationary 3D error and the way finger tapping was performed. Though in all these measurements a control subject performed the tapping task we needed to take into consideration the way that some children with Cerebral Palsy would perform finger tapping. Some of these children would curl their fingers while tapping which merged the knuckle and nail targets into the same line of sight if the cameras were near the level of the table on which the tapping hand rested. In order to avoid these instances cameras need to be placed way higher than the level of the table. To improve the anticipated capacity to resolve knuckle from nail targets when children with CP were tapping, Setup 6 was selected as the optimal arrangement for our two-camera system even though it had a slightly larger RMS error for the stationary targets when compared to Setups 1-3.

3.3 Quantification of Finger Tapping Motion Metrics

Motion tracking by use of retroreflective targets placed on the arms and fingers of pediatric subjects was performed during a finger tapping task. Measurements were performed at the Texas Scottish Rite Hospital for Children on two children with CP classified as MACS-1, two as MACS-2 and on two age-matched controls. The details of

the data acquisition protocols, such as retroreflective target placement, camera setups, motions performed, the age of children etc were describe in the Methods Section 2.4. The motion tracking measurements were performed twice for each subject with approximately two weeks elapsed in between the two measurements. For subjects with CP one measurement took place before the treatment (CIMT) and the second measurement after the treatment. Control subjects were measured with a very similar spacing in time between first and second measurements, but with no intervention occurring in-between. Before measuring each subject, both camera systems (the two and six camera setups) were calibrated as described in Methods Section.2.1.4. The 3D coordinates of the retroreflective targets were then processed and three variables were calculated from each subject's finger tapping data, namely the average amplitude, average velocity and instantaneous velocity of one target on the finger tap trajectory, as described in the Methods Section 2.4.3.

3.3.1 Finger Motion Metrics Quantified by the Two-Camera System

3.3.1.1 Amplitude

The amplitude calculated for each subject's affected hand, and the right hand for the controls, for the two measurement times (before treatment – T1; after treatment – T2; T short for Trial) for the two-camera system is shown in Table 3-7. The columns M-1a and M-1b indicate the two MACS-1 subjects (letter M for MACS), and the columns N1 and N2 indicate the corresponding measurements for the two controls (letter N for Normals).

Table 3-7: Two-camera system amplitude measurements for MACS-1 subjects ,M-1a, M-1b and two normals N-1, N-2 along with a T-test for the two trials

Amplitude (mm)							
Tap number	M-1a		N-1		N-2		M-1b
	T 1	T 2	T 1	T 2	T 1	T 2	T1

Table 3-7- continued

1	28.39	24.76	7.44	20.36	19.6	39.22	20.33
2	12.12	21.26	10.86	26.58	25.86	14.53	15.51
3	11.11	21.62	31.47	28.95	22.85	17.02	15.2
4	12.54	12.88	19.49	25.11	22.85	10.66	12.18
5	6.89	20.73	31.94	25.05	25.27	19.41	8.76
6	9.56	27.26	25.95	20.59	84.95	25.52	9.94
7	11.85	24.36	16.69	34.28	15.8	19.16	8.77
8	7.71	23.15	11.41	29.58	35.24	16.46	8.02
9	16.83	25.42	18.34	31.18	93.47	13.73	8.71
10	15.19	25.99	22.62	30.17	32.84	14.66	10.07
11	12.32	24.09	25.72	28.37	29.42	10.35	9.53
12	14.62	18.36	31.25	28.16	23.48	11.61	9.97
13	13.74	21.71				15.4	6.89
14	15.48	19.41					
15	14.81	7.27					
p-value	0.0003		0.04		0.03		

For M-1b only one trial was calculated since in T-2, each of the two cameras did not see the same number of retroreflective targets, which resulted in time sampling differences that made it very difficult to find what data points in one camera corresponded to those of the other camera.

As seen from Table 3-7, subject M-1a showed p value of 0.0003 for the finger tapping amplitude, which at first glance could imply that CIMT had a significant effect on how this subject tapped. However, when comparing corresponding data for N-1 and N-2 we see that both of these subjects also had significant changes in tapping amplitude between trials (the significance threshold is set at 0.05). From these results it can be concluded that average amplitude of tapping cannot be used as a measure that determines the effectiveness of treatment as the variability in how a subject taps between trials is very significant.

3.3.1.2 Instantaneous Velocity

Using equation (2.58) in (Methods Section 2.4) the instantaneous velocity was calculated at four distinct time points (V1 – V4; see Figure 2-5) during a finger tap for one MACS-1 (M1a), and the two Normal subjects for both trials and a T-test was performed between the two trials.

. Table 3-8: Instantaneous velocity measurements using the two-camera system for subject M-1a (MACS-1) along with a T-test for the two trials

M-1a								
Number of points	V1		V2		V3		V4	
	T1	T2	T1	T2	T1	T2	T1	T2
1	0.07	0.03	0.07	0.06	-0.13	-0.15	-0.03	-0.02
2	0.03	0.07	0.01	0.01	-0.07	-0.14	-0.01	-0.04
3	0.02	0.07	0.03	0	-0.06	-0.14	-0.01	-0.02
4	0.03	0.02	0.03	0.01	-0.06	-0.08	-0.02	0
5	0.02	0.06	0.01	0.02	-0.04	-0.12	-0.01	-0.04
6	0.03	0.08	0	0.04	-0.05	-0.19	-0.01	-0.08
7	0.04	0.08	0.01	-0.01	-0.04	-0.16	-0.03	-0.05
8	0.03	0.02	0.01	0.04	-0.04	-0.14	0	-0.03
9	0.05	0.07	0.01	0.04	-0.09	-0.15	-0.01	-0.04
10	0.04	0.04	0.02	0.08	-0.06	-0.18	-0.03	-0.02
11	0.02	0.07	0.03	0.04	-0.06	-0.13	-0.01	-0.07
12	0.04	0.06	0.02	0.01	-0.06	-0.12	-0.04	-0.02
13	0.03	0.04	0.02	0.07	-0.06	-0.1	-0.03	-0.06
14	0.05	0.04	0.01	0.03	-0.08	-0.12	-0.01	-0.01
15	0.03	0.02	0.03	0.01	-0.07	-0.03	-0.03	-0.03
16			0.02	0.03	-0.06	-0.13	-0.02	-0.03
p-value	0.04		0.22		3.7723E-06		0.02	

As seen in Table 3-8 the p values in the subject's starting velocity, where he just began to tap (V1), the velocity at which the hand was coming down after reaching the

highest peak (V3) and at the point when he was about to rest the hand on the table after one complete tap (V4) were significantly different between the two trials. The instantaneous velocity at the ascending midpoint of the tap (V2) did not change significantly. As observed in the recorded video of tapping, in Trial 1, the subject was performing slow tapping with his fingers being more flat on the table. However, in Trial 2 the fingers were curled very slightly and subject's starting velocity and the point at which he was about to rest the hand back on the table was faster in Trial 2 than in Trial 1.

Table 3-9: Instantaneous velocity measurements using the two-camera system for control subject 1 (N-1) along with a T-test for the two trials.

N-1 (m/sec)								
Number of points	V1		V2		V3		V4	
	T1	T2	T1	T2	T1	T2	T1	T2
1	0.01	0.07	-0.01	0.02	-0.05	-0.1	0	-0.05
2	0.04	0.09	0	0.03	-0.07	-0.14	-0.04	-0.13
3	0.13	0.1	-0.09	0.04	-0.19	-0.16	0	-0.07
4	0.08	0.08	0.04	0.02	-0.14	-0.13	-0.08	0
5	0.06	0.08	0.03	-0.03	-0.14	-0.13	-0.14	-0.02
6	0.06	0.06	0.05	0.06	-0.16	-0.12	-0.11	-0.04
7	0.04	0.09	0.02	0.06	-0.18	-0.17	-0.03	-0.05
8	0.02	0.1	0.01	0.01	-0.02	-0.14	-0.07	-0.11
9	0.07	0.08	0.01	0.05	-0.13	-0.08	-0.08	-0.14
10	0.01	0.1	-0.02	0.02	-0.08	-0.15	-0.08	-0.11
11	0.09	0.01	0.01	-0.01	-0.11	-0.09	-0.13	-0.11
12	0.09	0.08	0.01	0.05	-0.27	-0.09	-0.11	-0.13
p-value	0.13		0.11		0.88		0.7	

Table 3-10: Instantaneous velocity measurements using the two-camera system for control subject 2 (N-2) along with a T-test for the two trials

N-2								
Number of points	V1		V2		V3		V4	
	T1	T2	T1	T2	T1	T2	T1	T2
1	0.11	0.06	0.02	-0.03	-0.74	-0.1	-0.25	-0.04
2	0.01	0.08	0.03	-0.01	-0.1	-0.14	-0.05	-0.02
3	0.02	0.06	0.01	0.04	-0.11	-0.11	-0.01	0.01
4	0	0.05	-0.01	0.03	-0.07	-0.12	0	-0.1
5	0.03	0.07	0.02	-0.04	-0.11	-0.12	-0.06	-0.06
6	0.05	0.08	0.02	-0.02	-0.18	-0.13	-0.07	-0.06
7	0.04	0.26	0.03	0.02	-0.09	-0.43	-0.06	-0.02
8	0.02	0.05	0.03	0	0	-0.08	0	0
9	0.02	0.09	-0.02	-0.01	-0.08	-0.16	0	-0.11
10	0.01	0.08	0	0.03	-0.09	-0.18	0	-0.11
11	0.01	0.21	0	0.05	-0.06	-0.4	-0.01	-0.18
12	-0.03	0.04	-0.04	0.03	-0.08	-0.17	0.01	-0.03
13	0.03	0.06	0	-0.02	-0.11	-0.17	-0.07	-0.06
14	0.01	0.02	0	0.04	-0.05	-0.1	-0.03	-0.09
15	0.02	0.05	0.01	0.04	-0.11	-0.17	-0.04	-0.02
16	0.11	0.05	0.02	0.04	-0.74	-0.06	-0.25	-0.14
p-value	0.009		0.606		0.951		0.737	

As seen from Tables 3-9 and 3-10 instantaneous velocities were not significantly different between the two trials for N-1 and N-2 as p value was greater than 0.05 for all four time points, except for V1 for subject N-2. Observation of the video recorded while this subject was tapping indicated a significant difference in tapping style between the two trials, with the subject in T-1 performing higher amplitude tapping, which forced him to tap faster in order to keep up with the tapping pace indicated by the PowerPoint presentation.

3.3.1.3 Average velocity

For average velocity measurements, just one target on the affected hand was measured and the average velocities over 15 taps were calculated for both trials.

Table 3-11: : Average velocity measured for 15 taps for the index nail target along with the standard deviation and T-test between the two trials for one MACS-1 and two Normal subjects

	Average Velocity (m/sec)			SD (m/sec)	
	T1	T2	p-value	T1	T2
M-1a	0.043	0.096	< 0.001	0.09	0.05
N-1	0.128	0.125	0.837	0.09	0.12
N-2	0.149	0.083	0.716	0.154	0.16

The T-test was performed between two trials for M-1a, N-1 and N-2. A p- value of 0.000004 indicated shows significant difference between the two trials for M-1a while the p values of 0.837 and 0.716 indicated no significant differences in average velocity between the two trials for the two control subjects.

.

Table 3-12: Summary of the two camera system's p-values for one MACS-1 and two Normal subjects between the two trials for amplitude and average velocity

	M-1a	N1	N2
Amplitude	0.0003	0.04	0.03
Average velocity	< 0.001	0.84	0.72

Table 3-13: Summary of the two camera system's p-values for one MACS-1 and two Normal subjects between the two trials for instantaneous velocity

	M-1a	N1	N2
V1	0.04	0.13	0.01
V2	0.22	0.11	0.61
V3	< 0.001	0.88	0.95
V4	0.02	0.7	0.73

3.3.2 Finger Motion Metrics Quantified by the Six-Camera System

3.3.2.1 Amplitude

The tapping amplitude was measured for two MACS-1 (M-1a, M-1b) and two control subjects (N1, N2) for the six-camera system and a T-test was performed between the two trials..

Table 3-14: Amplitude measurement of one target for two MACS-1(M-1a, M-1b) and two normal subjects (N-1, N-2) for two trials along with a T-test

Amplitude (mm)								
Tap number	M-1a		N-1		N-2		M-1b	
	T 1	T 2	T 1	T 2	T 1	T 2	T1	T2
1	28.45	23.93	4.83	22.26	20.08	13.89	19.87	16.35
2	12.04	21.48	7.95	23.41	24.61	16.66	14.64	16.29
3	11.94	21.76	23.54	29.36	19.17	9.36	13.99	14.09
4	12.75	12.84	11.28	27.39	21.08	18.01	11.41	12.33
5	7.27	20.41	16.57	17.71	24.97	24.47	8.16	14.65
6	10.1	26.6	21.77	20.81	23.95	16.93	9.14	17.36
7	12.87	24.19	16.33	33.64	36.39	15.93	8.29	13.13
8	8.52	21.77	27.77	29.33	14.92	12.43	7.23	12.67
9	16.68	24.9	22.67	34.91	28.28	13.31	8.56	5.36
10	15.42	26.03	12.99	33.07	34.88	9.63	9.28	7.74

Table 3-14- continued

11	12.39	24.73	9.94	29.75	42.5	10.01	8.48	11.34
12	14.78	18.87	16.52	31.31	31.92	14.23	8.89	6.62
13	14.37	21.82	16.82	37.01	27.02	8.99	6.55	12.64
14	15.82	19.84	23.27	31.17	22.24			
15	15.27	7.5	27.79	33.12	28.28			
p-value	0.0004		0.00003		0.00006		0.185	

As seen from Table 3-14, the p-value for M-1a shows a marginally significant difference between the two trials. However, for M-12, no significant differences were found. Also, for the two control subjects, the average amplitude differences were significant between the two trials. These results are in accordance with the corresponding tapping amplitude results found with the two-camera system, with the exception that subject M1b could also be measured here as the six-camera system was lot less sensitive to target occlusion than the two-camera system.

3.3.2.2 Instantaneous Velocity

For the six-camera system, 3D co-ordinates were obtained for two MACS-1 subjects and two controls. Instantaneous velocity measurements were performed at same four time points and T-test results between two trials are tabulated below.

Table 3-15: Six-camera system instantaneous velocity measurements for MACS-1 subject M-1a at four time instants along with a T-test for the two trials

M-1a (m/sec)								
Number of points	V1		V2		V3		V4	
	T1	T2	T1	T2	T1	T2	T1	T2
1	0.06	0.07	0.05	0.08	-0.1	-0.07	-0.03	-0.13
2	0.02	0.09	0.03	0.04	-0.02	-0.06	-0.04	-0.12
3	0.02	0.08	0.02	0.05	-0.02	-0.07	-0.03	-0.11

Table 3-15- continued

4	0.03	0.03	0.02	0.05	-0.02	-0.03	-0.03	-0.07
5	0.01	0.07	0.02	0.05	-0.03	-0.06	0	-0.11
6	0.02	0.1	0.02	0.07	-0.04	-0.09	-0.01	-0.14
7	0.02	0.09	0.03	0.06	-0.04	-0.06	-0.01	-0.13
8	0.02	0.08	0.01	0.04	-0.03	-0.01	-0.01	0
9	0.02	0.1	0.04	0.05	-0.03	-0.06	-0.05	-0.14
10	0.03	0.09	0.03	0.05	-0.02	-0.08	-0.05	-0.14
11	0.02	0.06	0.03	0.09	-0.04	-0.07	-0.02	-0.13
12	0.02	0.06	0.03	0.05	-0.04	-0.05	-0.03	-0.1
13	0.03	0.08	0.02	0.05	-0.05	-0.07	-0.02	-0.11
14	0.03	0.04	0.02	0.08	-0.05	-0.05	-0.02	-0.11
15	0.03	0.03	0.02	0.01	-0.04	-0.02	-0.03	-0.04
p-value	0.000001		0.000054		0.02		0.000001	

Table 3-16: Six-camera system instantaneous velocity measurements for MACS-1 subject M-1b at four time instants along with a T-test for the two trials

M-1b (m/sec)								
Number of points	V1		V2		V3		V4	
	t1	t2	t1	t2	t1	t2	t1	t2
1	0.05	0.04	0.02	0.02	-0.02	-0.04	-0.08	-0.07
2	0.03	0.03	0.02	0.04	-0.01	-0.07	-0.07	-0.04
3	0.03	0.03	0.02	0.03	-0.02	-0.04	-0.06	-0.06
4	0.01	0.03	0.03	0.02	-0.04	0	-0.02	-0.08
5	0.02	0.04	0.01	0.02	-0.02	-0.07	-0.03	-0.04
6	0.02	0.04	0.01	0.04	-0.03	-0.09	-0.02	-0.04
7	0.02	0.03	0.01	0.02	-0.03	-0.06	-0.02	-0.03
8	0.02	0.03	0.01	0.03	0	-0.04	-0.03	-0.05
9	0.01	0.01	0.01	0.02	-0.03	-0.01	-0.02	-0.03
10	0.02	0.01	0.01	0.02	-0.01	-0.02	-0.04	-0.03
11	0.02	0.02	0.01	0.01	-0.02	-0.01	-0.03	-0.03
12	0.02	0.03	0.01	0.02	-0.02	-0.04	-0.03	-0.04
13	0.02	0.01	0.01	0	-0.02	-0.01	-0.02	-0.03
p- value	0.28		0.03		0.05		0.31	

As seen from above Table 3-15, subject's M-1a instantaneous velocity measurements showed significant differences between the two trials as p-values were less than 0.05. Also, for subject M-1b (Table 3-16), significant differences were observed in p-values that correspond to V2 and V3 which indicated that ascending tap velocity while reaching the peak and the descending tap velocity after attaining the peak tapping amplitude was different. This implies that the velocity at which the subject began tapping was not a lot different between two trials. The reason for significant differences that were observed in V2 and V3 was because in Trial 1 the subject raised his hand for tapping and after a brief pause brought back his hand to the table. As a result the tapping velocity profile was choppier in Trial 1, while in Trial 2 he performed regular tapping with no pause in between. Also, the amplitude of tapping was very low in Trial 1.

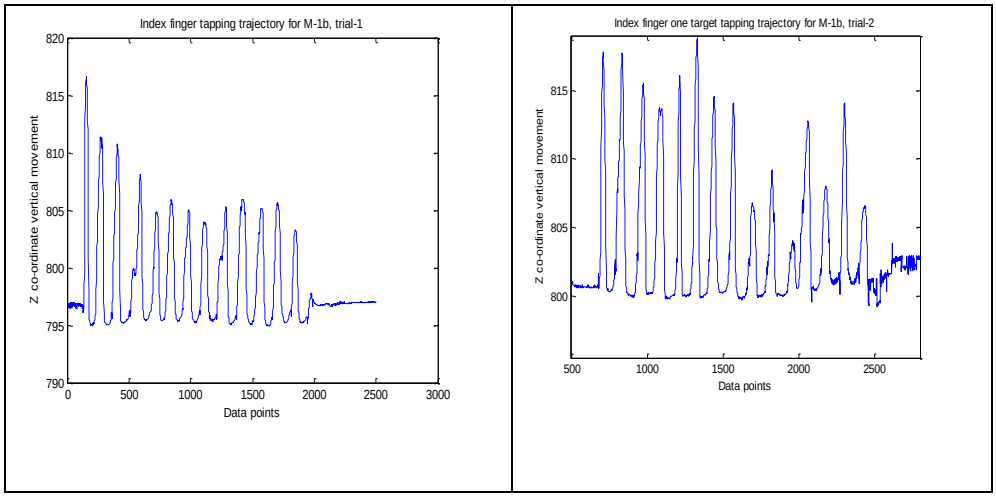


Figure 3-14: Tapping profile of index finger nail target for M-1b from (a) Trial 1 and (b) Trial 2

Table 3-17: Six-camera system instantaneous velocity measurements for control subject N-1 at four time instants along with T-test for the two trials.

N-1 (m/sec)								
Number of points	V1		V2		V3		V4	
	T1	T2	T1	T2	T1	T2	T1	T2
1	0.013	0.057	0.01	0.042	-0.008	-0.057	-0.024	-0.089
2	0.015	0.036	0.017	0.015	-0.033	-0.051	-0.048	-0.101
3	0.057	0.057	0.037	0.066	-0.07	-0.085	-0.165	-0.107
4	0.032	0.059	0.016	0.05	-0.04	-0.007	-0.074	-0.009
5	0.044	0.01	0.022	0.002	-0.04	-0.051	-0.126	-0.069
6	0.069	0.058	0.018	0.026	-0.056	-0.049	-0.16	-0.08
7	0.044	0.077	0.024	0.06	-0.041	-0.055	-0.115	-0.16
8	0.062	0.006	0.047	0.003	-0.099	-0.044	-0.167	-0.146
9	0.046	0.028	0.042	0.026	-0.057	-0.068	-0.164	-0.161
10	0.033	0.044	0.021	0.094	-0.008	0	-0.014	0
11	0.027	0.079	0.012	0.025	-0.026	-0.006	-0.067	0.001
12	0.045	0.058	0.019	0.026	-0.002	-0.067	0.001	-0.142
p-value	0.427		0.166		0.66		0.845	

Table 3-18: Six-camera system instantaneous velocity measurements for control subject N-2 at four time instants along with T-test for the two trials.

N-2 (m/sec)								
Number of points	V1		V2		V3		V4	
	T1	T2	T1	T2	T1	T2	T1	T2
1	0.04	0.01	0.08	0.07	-0.32	-0.06	-0.27	-0.12
2	-0.01	0	0.02	0	-0.14	-0.16	-0.03	-0.07
3	0	0.03	0.02	0.05	-0.09	-0.14	-0.12	-0.03
4	-0.01	0.02	0.01	0.06	-0.06	-0.15	-0.06	-0.05
5	0.02	0.06	0.03	0.03	-0.05	-0.08	-0.14	-0.13
6	0.02	0.06	0.04	0.01	-0.14	-0.18	-0.16	-0.01
7	0.04	0.09	0.08	0.04	-0.31	-0.2	-0.28	-0.11
8	0	0.03	0	0.03	-0.1	-0.12	-0.07	0
9	-0.01	0.07	0.02	0.02	-0.14	-0.14	-0.07	-0.11

Table 3-18- continued

10	0	0.07	0.01	0.04	-0.04	-0.12	-0.07	-0.18
11	0.02	0.1	0.03	0.05	-0.03	-0.13	-0.16	-0.19
12	0.03	0.04	0.04	0.05	-0.1	-0.22	-0.2	-0.01
13	0.08	0.05	0.08	0.03	-0.04	-0.06	-0.44	-0.17
14	0	0.01	0.03	0.05	-0.12	-0.16	-0.07	-0.01
15	0	0.04	0.02	0.05	-0.08	-0.07	-0.12	-0.15
16		0.05		0.04		-0.05		-0.15
p-value	0.003		0.705		0.707		0.088	

As seen from Table 3-17 no significant differences were observed for N-1 between trials for the instantaneous velocity measurements. However, N-2 showed a marginal difference in V1, these results are in accordance with two camera system results and it was observed because subject performed fast tapping in Trial 1.

3.3.2.3 Average Velocity

Table 3-19 shows the average tapping velocity calculated for two controls and two MACS-1 subjects for the index finger target. This was the average of 15 taps.

Table 3-19: Average velocity for the six-camera system calculated for two MACS-1 (M-1a, M-1b) and two control (N-1, N-2) subjects for the two trials

	Average Velocity (m/sec)			SD (m/sec)	
	T1	T2	p-value	T1	T2
M-1a	0.12	0.15	0.0003	0.05	0.11
M-1b	0.03	0.05	0.002	0.05	0.04
N-1	0.07	0.09	0.75	0.09	0.12
N-2	0.19	0.13	0.16	0.21	0.18

In the above table the p-values show that the average velocity was significantly different for the MACS-1 subjects. For normal subjects the p-value was greater than 0.05 which indicates no significant differences in the two trials.

Table 3-20: Summary of Six camera system's p-values for two MACS-1(M-1a, M-1b) and two control subjects (N-1,N-2) between two trials for amplitude and average velocity.

	M-1a	M-1b	N1	N2
Amplitude	0.0004	0.00003	0.00006	0.185
Average velocity	0.0003	0.002	0.75	0.16

Table 3-21: : Summary of Six camera system's p-values for two MACS-1(M-1a, M-1b) and two control subjects (N-1,N-2) between two trials for instantaneous velocity

	M-1a	M-1b	N1	N2
V1	0.000001	0.28	0.43	0.03
V2	0.00005	0.03	0.17	0.705
V3	0.02	0.05	0.66	0.707
V4	0.000001	0.31	0.85	0.088

The conclusion is the opposite of what was initially hypothesized: We expected that controls would not be significantly different and subjects with CP would change. We found instead that the controls were very different between the two trials and the CPs sometimes were more similar between trials with some metric differences been significant and others not. If anything, the only thing that can be concluded is that the tapping protocol used in this works allowed for too much variability in how subjects could tap, which precluded any quantitative motion metric comparisons to be made between trials.

3.4 Comparison of the Two Camera Systems

In this work the six-camera system was considered as the standard for testing the motion tracking performance of the two-camera system. This was because the six-camera system was much less sensitive to target occlusion, due to the many more camera views available, and had a much higher time sampling of 120 Hz versus about 25 Hz for the two-camera system. However, the six camera system is much more costly, making it worthwhile to explore if the much lower cost two-camera system could still track finger tapping motions adequately. The metrics of comparison for deciding if the two-camera system performed adequately were the amplitude, average velocity and instantaneous velocity of finger tapping. These metrics were calculated from finger tapping data for one MACS-1, M-1a both the trials and other MACS-1, M-1b with one trial along with two normal subjects, N-1 and N-2. Before measuring each subject, both the camera systems (two- and six- camera setup) were calibrated and a system transformation matrix was computed, using four static table targets seen by both systems, so that the coordinates of the two camera systems were aligned, as described in methods Section 2.3

In the subsections below, it was first validated that the co-ordinate transformation did not alter the trends in statistical significance for changes in tapping amplitude between trials, Then both static target errors (3D triangulation and coordinate transformation errors) and moving target errors (movements of the index nail) were quantified for each subject individually. Specifically, the mean and standard deviation of the errors for amplitude, average velocity and instantaneous velocity are tabulated in the subsections below for each subject.

3.4.1 Validation of the Coordinate Transformation from the Two-Camera to the Six-Camera System using the Tapping Amplitude Metric

One target from affected hand's Index Nail was selected and amplitude was calculated for the first epoch, 15 taps. In order to validate that the transformation from the two-camera to the six-camera coordinate system (Methods Section 2.3) did not affect the two-camera results significantly, the tapping amplitudes for two measurement sessions were compared after this coordinate transformation had occurred, just as was done for the two-camera system in Table 3-7 above. The finger tapping amplitude per tap for the index finger is shown in Table 3-22 below for these co-ordinate transformed data. The T-test was performed between the first and second measurement sessions for a MACS-1, M-1a and two normal subjects, as was previously performed in Table 3-7 for the two-camera co-ordinate data.

Table 3-22: Amplitude measurements for MACS-1 ,M-1a and two normal subjects N-1, N-2 along with T-test for two trials and MACS-1, M-1b with one trial calculated from the two-camera system 3D co-ordinates after transforming them to the six-camera system's reference frame

Amplitude (mm)							
Tap number	M-1a		N-1		N-2		M-1b
	T 1	T 2	T 1	T 2	T 1	T 2	T1
1	27.9	24.79	7.52	20.37	19.62	14.54	20.26
2	11.66	21.29	10.91	26.8	25.88	17.03	15.39
3	10.64	21.65	31.74	29.08	21.35	10.67	15.09
4	12.06	12.91	32.09	25.1	22.85	19.42	12.05
5	6.4	20.76	26.01	25.44	22.83	25.52	8.64
6	9.06	27.29	16.72	20.69	25.26	19.15	9.83
7	11.34	24.4	11.43	34.32	38.08	16.49	8.67
8	7.22	23.19	18.34	29.66	15.79	13.75	7.9

Table 3-22- continued

9	16.36	25.45	22.68	31.18	31.22	14.67	8.62
10	14.71	26.01	25.76	30.25	35.2	10.37	9.96
11	11.84	24.12	31.29	28.37	42.59	11.64	9.44
12	14.14	18.4	10.1	28.15	32.83	15.42	9.87
13	13.26	21.75			29.38	9.94	6.79
14	14.99	19.44					
15	14.34	7.34					
T-Test	0.0002		0.0265		0.0001		

As seen from Table 3-22, the amplitude calculated for the two-camera system after transforming its co-ordinates to the six-camera system's reference frame gave very similar results to the ones found for the two-camera reference frame. The p-values for M-1a were 0.0004 in Table 3-7 compared to 0.0002 for the two-camera system after the coordinate transformation. Both results show a significant difference between the two trials for the tapping amplitude.

For normal subjects N-1 and N-2 the amplitude shows significant difference between trials, as seen from Table 3-7. Statistically significant differences were maintained after the coordinate transformation to the six-camera system’s reference frame with p-values of 0.0265 and 0.0001 respectively (Table 3-22). Though not conclusively proving the validity of this reference frame transformation, the above results indicate that statistical significance trends are not altered once transformation has occurred.

3.4.1.1 Errors in Tapping Amplitude for Subject M-1a

The index finger displacement per tap and measurement session (trial number T1, T2) is shown in Table 3-23 below for subject M-1a (MACS-1) along with the root mean square (RMS) errors and their standard deviation (SD).

Table 3-23: Percent error and error in magnitude for amplitude measurements of single target calculated for two trials of M1a

Tap number	M-1a			
	T1		T2	
	Error (%)	Error (mm)	Error (%)	Error (mm)
1	1.919	0.546	3.603	0.862
2	3.183	0.383	-0.86	-0.185
3	10.885	1.299	-0.509	-0.111
4	5.376	0.685	0.536	0.069
5	11.872	0.863	1.734	0.354
6	10.258	1.036	2.625	0.698
7	11.881	1.529	0.844	0.204
8	15.318	1.306	6.542	1.424
9	1.909	0.318	2.185	0.544
10	4.597	0.709	-0.073	-0.019
11	4.371	0.541	-2.47	-0.611
12	4.344	0.642	-2.508	-0.473
13	7.724	1.11	-0.35	-0.076
14	5.253	0.831	-1.998	-0.396
15	6.092	0.93	-2.125	-0.159
RMSE	8.04	0.88	2.49	0.55
SD	4.10	0.354	2.53	0.171

Table 3-24: 3D Error in X, Y and Z co-ordinate triangulation and error in estimating the transformed coordinates along with standard deviation calculated for different trials for subject M-1a.

M-1a	T1	T2

Table 3-24- continued

	X (mm)	Y (mm)	Z (mm)	X (mm)	Y (mm)	Z (mm)
Error in transformation	1.401	1.432	0.242	0.977	4.123	0.168
SD	0.485	1.138	0.221	0.398	0.237	0.02
Triangulation Error	0.156	0.291	0.496	0.171	0.317	0.513
Triangulation SD	0.138	0.212	0.395	0.124	0.25	0.416

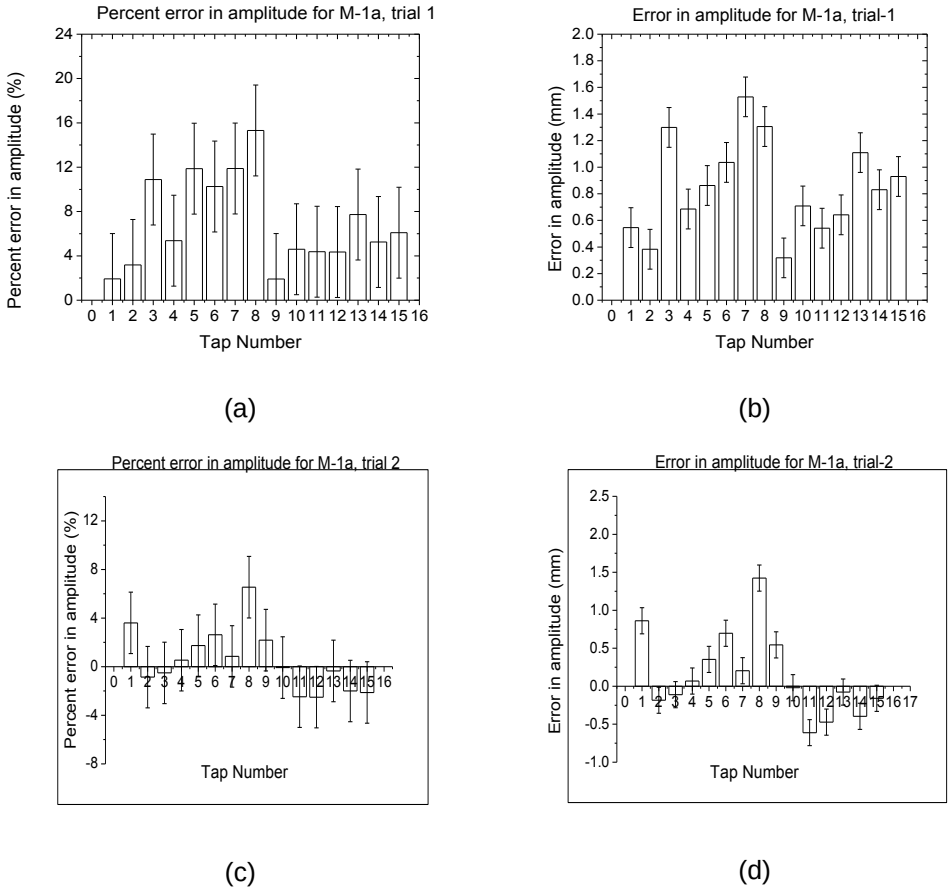


Figure 3-15: (a) Percent amplitude error (%) and (b) Error in magnitude of amplitude calculated for index nail target on subject M-1a, Trial 1, along with corresponding standard deviations. (c) and (d),are as in (a) and (b), respectively, but for Trial 2

.

As seen from Table 3-24, the maximum error between two camera systems for Trial 1 and Trial 2 was around 1.5 mm with RMS error of 0.88mm and standard deviation

of 0.35mm. Therefore the two-camera system gave amplitude measurements that were close to those of the six-camera system. Moreover, for whole trajectory, all the targets were visible by both systems. The average tapping amplitude for one epoch was found to be 13 mm for T1 and 21 mm for T2, which is a noteworthy difference. As for the possible contribution of static error to the tapping amplitude (along what was defined as the Z-axis) the coordinate transformation and triangulation error were both less than 0.5 mm (Table 3-24). Also, the error in transformation was huge in the Y direction for both trials as the left and right pinky finger nail targets and the two targets on the ulna were used for getting the transformation between two systems (no static targets were available during those measurements). The 3D co-ordinates of these targets were selected when the hand was at rest

3.4.2.2 Errors in Tapping Amplitude for Subject M-1b

The index finger displacement per tap and measurement session (trial number T1) is shown in Table 3-25 below for subject M-1b (MACS-1) along with the root mean square (RMS) errors and their standard deviation (SD).

Table 3-25: Percent error and error in magnitude for amplitude measurements of single target calculated for index finger target for M-1b, Trial 1

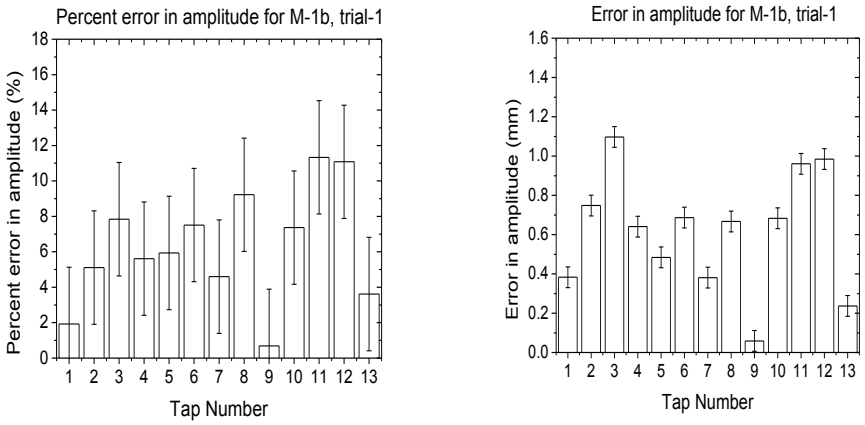
Tap number	M-1b	
	T-1	
	Error (%)	Error (mm)
1	1.928	0.383
2	5.11	0.748
3	7.842	1.097
4	5.614	0.641
5	5.933	0.484
6	7.509	0.687
7	4.597	0.381

Table 3-25- continued

8	9.222	0.667
9	0.688	0.059
10	7.367	0.683
11	11.331	0.96
12	11.079	0.984
13	3.617	0.237
RMS	7.01	0.68
SD	3.22	0.31

Table 3-26:Error in X, Y and Z co-ordinate triangulation and error in estimating the transformed coordinates along with standard deviation calculated for different trials for subject M-1b

M-1b	T1		
	X (mm)	Y (mm)	Z (mm)
Error in transformation	1.228	9.137	0.336
SD	0.879	2.694	0.252
Triangulation Error	0.151	0.262	0.494
Triangulation SD	0.109	0.214	0.389



(a) (b)

Figure 3-16: (a) Percent amplitude error (%) and (b) error in amplitude calculated for the index nail target on subject M-1b, Trial 1, along with corresponding standard deviations

For subject M-1b Trial1, the error in transformation and in 3D co-ordinates was less than 0.4 mm as seen from Table 3-26. As for the tapping data the error was as large as 12% (0.96 mm) for certain taps, which was similar to the numbers found for subject M-1a. Also, for this trial, the two-camera system did not observe all the targets at all times. The index finger knuckle and nail targets were reported as one target by Camera B of the two-camera system. For some time points, the area of targets was greater than 40 (as opposed to 20 for single targets), which indicated that the two targets were overlapped for that time point. Target occlusion occurred because this subject kept his fingers flat during tapping and for Camera B, which was lower and closer in height to the table than Camera A, the two targets on the index finger fell into the same line of sight. The fact that one less target was seen for certain time-intervals meant that the two-camera system had a slightly different data sampled rate at those instances. For the six-camera system, no occlusions were observed.

3.4.2.3 Errors in Tapping Amplitude for Subject N-2

The index finger displacement per tap and measurement session (trial number T1, T2) is shown in Table 3-27 below for subject N-2 (normal control) along with the root mean square (RMS) errors and their standard deviation (SD).

Table 3-27 : Percent error in amplitude measurements of single target calculated for two trials of subject N-2

N-2				
Tap number	T1		T2	
	Error (%)	Error (mm)	Error (%)	Error (mm)
1	2.292	0.46	4.699	0.652
2	-5.155	-1.269	2.264	0.377
3	-11.355	-2.177	13.97	1.308
4	-8.399	-1.771	7.855	1.414
5	8.564	2.138	4.258	1.042

Table 3-27- continued

6	-5.459	-1.307	13.083	2.216
7	-4.643	-1.69	3.551	0.565
8	-5.814	-0.868	10.585	1.316
9	-10.39	-2.938	10.189	1.357
10	-0.909	-0.317	7.706	0.742
11	-0.218	-0.093	16.327	1.634
12	-2.833	-0.904	8.338	1.187
13	-8.719	-2.356	10.518	0.946
14	-5.453	-1.213		
15	-9.458	-2.675		
RMS	6.71	1.66	9.6	1.23
SD	5.12	1.27	4.23	0.02

Table 3-28: 3D Error in X, Y and Z co-ordinate triangulation and error in estimating the transformed coordinates along with standard deviation calculated for different trials for subject N-2

N-2	T1			T2		
	X (mm)	Y (mm)	Z (mm)	X (mm)	Y (mm)	Z (mm)
Error in transformation	1.748	3.09	0.651	0.372	2.538	0.503
SD	0.26	0.778	0.304	0.276	0.785	0.136
Triangulation Error	0.16	0.254	0.428	0.141	0.255	0.461
Triangulation SD	0.111	0.198	0.345	0.097	0.206	0.355

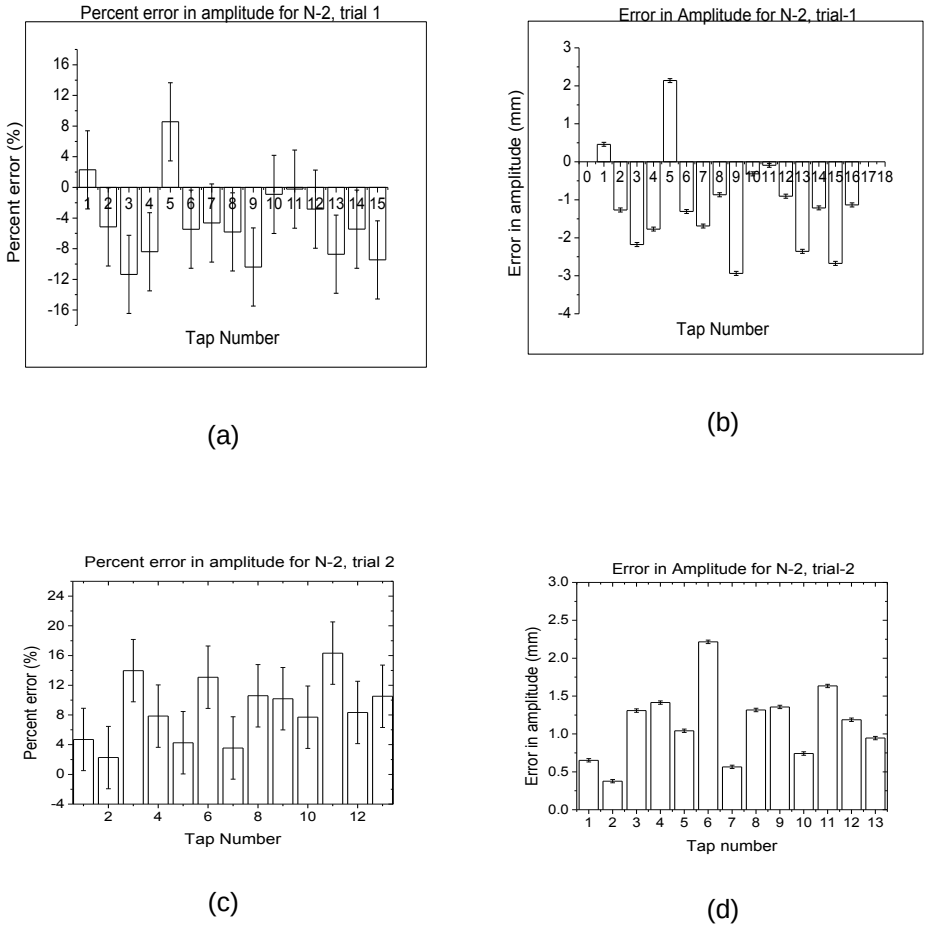


Figure 3-17: (a) Percent amplitude error (%) and (b) error in amplitude calculated for index nail target on subject N-2, Trial 1, along with corresponding standard deviations. (c) and (d), are as in (a) and (b), respectively, but for Trial 2

Overall, absolute magnitude and percent errors for tapping were similar for this subject as for the ones above. The static errors were also similar (Table 3-28). For Trial 1, the index finger target was occluded from one camera of the two-camera system for 11 time points (duration of 440 ms). For the purpose of calculation of the Z co-ordinates of the moving finger from two-camera system, extra X and Y pixel co-ordinates were added. The index finger's X, Y pixel co-ordinates were missing at several time instants and extra co-ordinates for the index finger were added from the time points just before the time

point where the target went missing. As a result, at certain taps in the middle of an epoch, high error of 2.5 mm was observed as seen in Figure 3-17(b).

For Trial 2with the two-camera system, there were four time points (duration of 160 ms) where all four targets from the nails were missing. This was not observed with the six-camera system.

3.4.2.4 Errors in Tapping Amplitude for Subject N-1

The index finger displacement per tap and measurement session (trial number T1, T2) is shown in Table 3-29 below for subject N-2 (normal control) along with the root mean square (RMS) errors and their standard deviation (SD).

Table 3-29: Percent error in amplitude measurements of single target calculated for two trials of subject N-1

N-1				
Tap number	T1		T2	
	Error(%)	Error(mm)	Error(%)	Error(mm)
1	55.648	2.689	8.495	1.891
2	37.234	2.961	-14.482	-3.39
3	34.828	8.199	0.963	0.283
4	184.462	20.81	8.344	2.285
5	56.954	9.438	-43.664	-7.732
6	-23.188	-5.048	0.582	0.121
7	-30.032	-4.905	-2.018	-0.679
8	-33.974	-9.435	-1.144	-0.336
9	0.044	0.01	10.686	3.73
10	98.27	12.767	8.534	2.822
11	214.696	21.344	4.625	1.376
12			10.101	3.162
13			-6.452	-2.388
14			1.068	0.333
15			11.42	3.782

Table 3-29- continued

RMS	96	11.12	13.5	2.99
SD	83.14	10.23	14	3.07

Table 3-30:3D Error in X, Y and Z co-ordinate triangulation and error in estimating the transformed coordinates along with standard deviation calculated for different trials for subject N-1.

N-1	T1			T2		
	X (mm)	Y (mm)	Z (mm)	X (mm)	Y (mm)	Z (mm)
Error in transformation	0.819	2.874	0.359	1.184	2.697	0.144
SD	0.229	1.551	0.168	0.342	0.386	0.066
Triangulation Error	0.149	0.283	0.541	0.136	0.289	0.514
Triangulation SD	0.107	0.247	0.466	0.103	0.218	0.372

Overall, magnitude and percent static errors were similar for this subject as for the ones above (Table 3-29). However, the tapping errors were large at this instance. The error observed for single target amplitude for Trial 1 of subject N1, was higher with a maximum error of 21 mm as seen from the Figure above. However, the glitch in percent error can be explained as the number of taps observed by the six-camera system was 15 and for the two-camera system it was 12. Moreover, in Trial 1 most taps observed by the two-camera system showed abnormally high amplitude which indicates that there were errors in triangulation. The source of the error was that triangulation was performed between the two cameras by using data corresponding to different time points. Suppose, at time point t1, Camera A registered X, Y pixel co-ordinates of a nail target but at same time, Camera B reported recorded it at a shifted time point ($t1+\Delta t$) because the two cameras did not see the same number of targets. This shift in time resulted in erroneous triangulation.

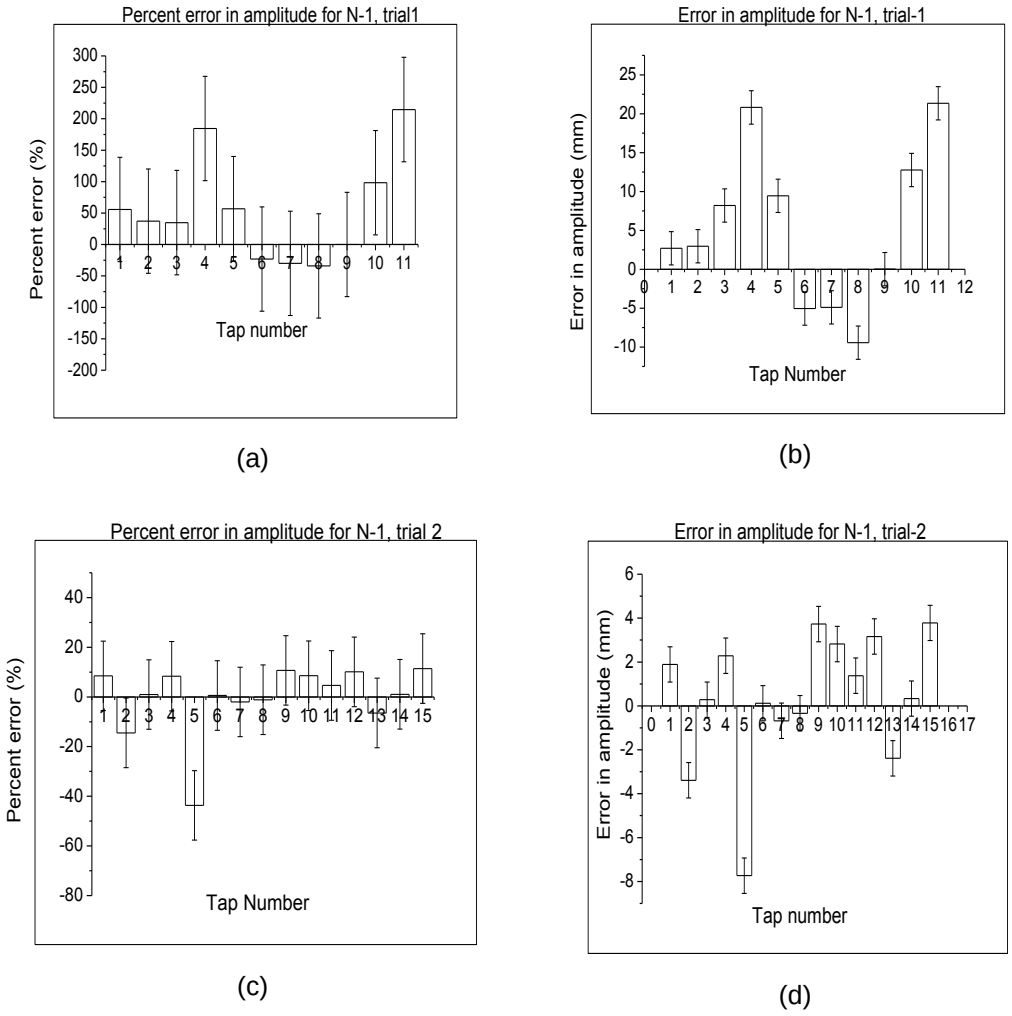


Figure 3-18: (a) Percent amplitude error (%) and (b) error in amplitude calculated for index nail target on subject N-1, Trial 1, along with corresponding standard deviations. (c) and (d) are as in (a) and (b), respectively, but for Trial 2.

For Trial 2, only the fifth tap showed a huge error of 7 mm and the reason for this is seen in Table 3-30. Errors observed in X and Y transformation parameters were large as the targets used for getting transformation parameters were not all perfectly static. The two front targets were on the table, but the two back targets, on the left and right pinky

finger knuckles were not always immobile even when the subject intended to keep the hands at rest

The above comparisons indicate that it is extremely important for the two-camera system not to have occluded targets if it is to track motions properly. The loss of targets can results in ambiguity of what point in one camera corresponds to the same point at the other camera, which can lead to erroneous triangulation. In addition, as the number of points recorded affects the sampling rate of these cameras, If the two cameras do not see the same number of points due to occlusions, they also don't sample images at exactly the same frame rate, which then necessitates interpolation in order to estimate the X, Y coordinates of a target on both cameras at the same instance in time, Therefore, it was found that the two-camera system could track 3D co-ordinates accurately only if both cameras are positioned in such a way that whole field of view covered both the hands and the subject did not tap in a way that could cause occlusions for that setup. It was also seen that the magnitude of static errors was very small in all instances. Hence the main source of error in the comparison between the two-camera systems was due to differences in time sampling. As compared to the six-camera system, which had a sampling frequency of 120 Hz, the two-camera system acquired data at a rate of less than 25 Hz when the number of targets was about 26. It is possible that comparisons between the two systems would be more favorable if a smaller number of targets were involved, which could increase the two-camera system's frame rate and, with appropriate target placement, reduce the probability of occlusions. Moreover, estimation of 3D co-ordinates was time consuming for the two-camera system, as it was done by simple home-grown algorithms that were not general purpose and required modifications with respect to each subject to compensate for occluded targets.

3.4.2 Validation of the Coordinate Transformation from the Two-Camera to the Six-Camera System using the Tapping Instantaneous Velocity Metric

Instantaneous velocity was calculated at four time instants, V1, V2, V3 and V4 along a tap as defined in Section 2.4.Seven data sets from both camera systems were used for these comparisons as previously described in above section. The differences in instantaneous velocity calculated for the two-camera system relative to the six-camera system, which was used as the gold standard for this comparison, were expressed as percent error, error in m/sec, as described in Equation 2.59. All of these errors were calculated for the displacements of the target on the index finger nail during tapping. These errors were plotted individually for each data set, as shown here below.

3.4.2.1 Subject M-1a

Table 3-31: Percent error, error in (m/sec) calculated for instantaneous velocity at four time instants for subject M-1a, Trial1.

Instantaneous velocity (m/sec)		
M-1a,T1	Percent error	Error in magnitude
V1	20.28	0.067
V2	81.82	0.005
V3	22.24	-0.073
V4	33.33	0.001

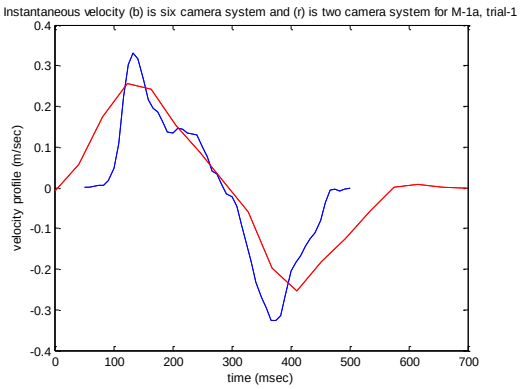
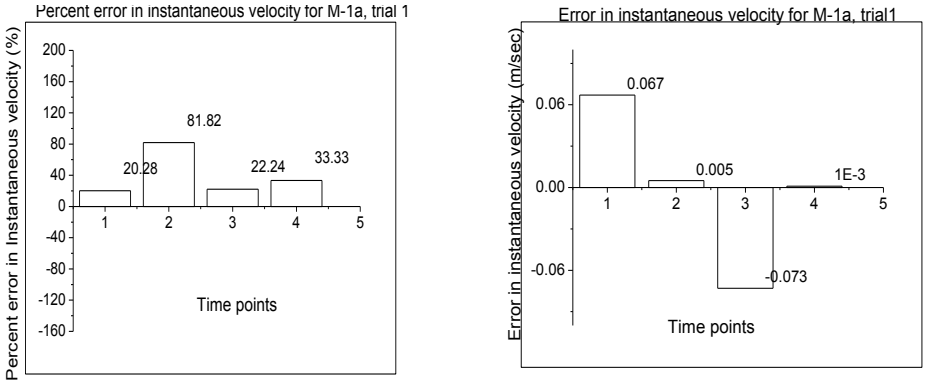


Figure 3-19: Time shifted instantaneous velocity profile of one tap for the index finger nail target for M-1a, Trial-1: blue indicates the velocity profile for the six camera system and red indicates the velocity profile for two camera system

As seen from Table 3-31, the percent error calculated for the two-camera system shows large variations. However, the error in magnitude indicates a seemingly small error in m/sec. In Figure 3-20(a), large errors were observed in percentage for V4, the point where the finger just reaches the table after attaining maximum peak height of tap and in V2, the point where the hand is at maximum peak position. The reason for such huge errors in V2 and V4 is that the six-camera system has more accurate instantaneous velocity data than the two-camera system due to its capacity for higher temporal sampling,

Table 3-32: Percent error, error in (m/sec) calculated for instantaneous velocity at four time instants for subject M-1a, Trial 2

Instantaneous velocity (m/sec)		
M-1a,T2	Percent error	Error in magnitude
V1	6.39	0.009
V2	-900	-0.009
V3	46.08	-0.235
V4	99.52	0.021



(a) (b)

Figure 3-20: (a) Percent error in instantaneous velocity and (b) error in instantaneous velocity (m/sec) calculated for the index nail target on subject M-1a, Trail1

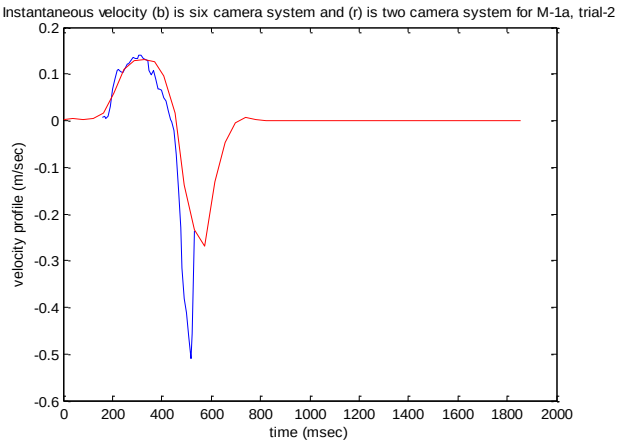
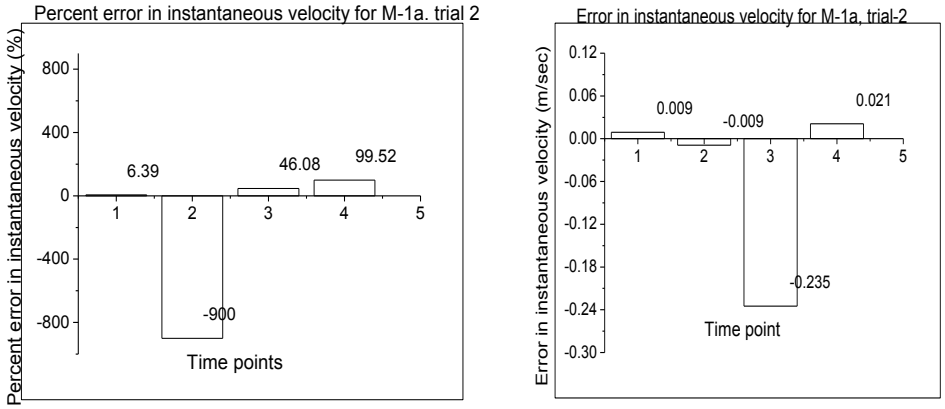


Figure 3-21: Time shifted instantaneous velocity profile of one tap for the index finger nail target for M-1a, Trial-2; blue indicates the velocity profile for the six-camera system and red indicates the velocity profile for the two-camera system



(a) (b)

Figure 3-22:(a) Percent error in instantaneous velocity and (b) error in instantaneous velocity calculated for the index nail target on subject M-1a, Trial2

As seen in Figure 3-21, the six-camera system shows large instantaneous velocity for V3. Also, from Table 3-32, it was observed that percent errors were small for V2, as the six-camera system showed the velocity at V2 to be 0.001 m/sec while the two-

camera system measured 0.01 m/sec. However, for V4, the six-camera system data was 0.021 m/sec while the two-camera system measured 0.0001 m/sec.

3.4.2.2 Subject N-2

Table 3-33: Percent error, error in (m/sec) calculated for instantaneous velocity at four time instants for subject N-2, Trial 1

Instantaneous velocity (m/sec)		
N2-T1	Percent error	Error in magnitude
V1	61.09	0.377
V2	-681.63	-0.033
V3	55.17	-0.312
V4	60	0.003

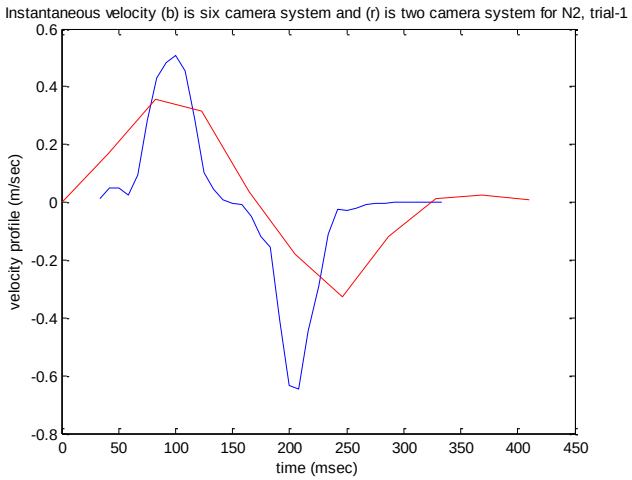
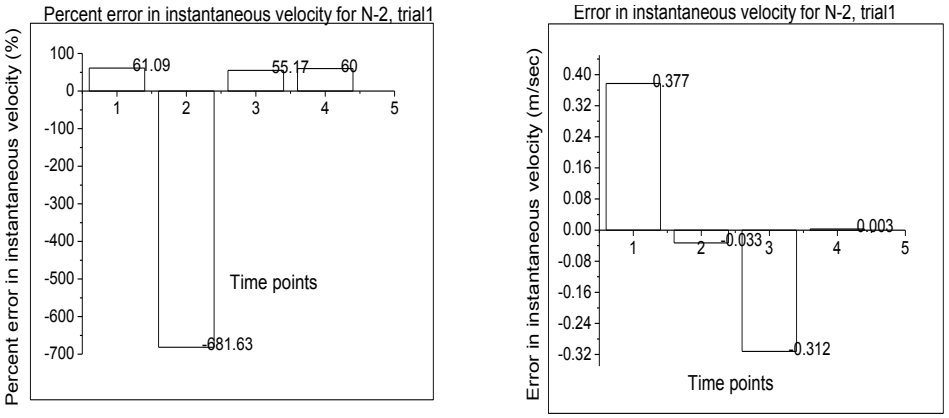
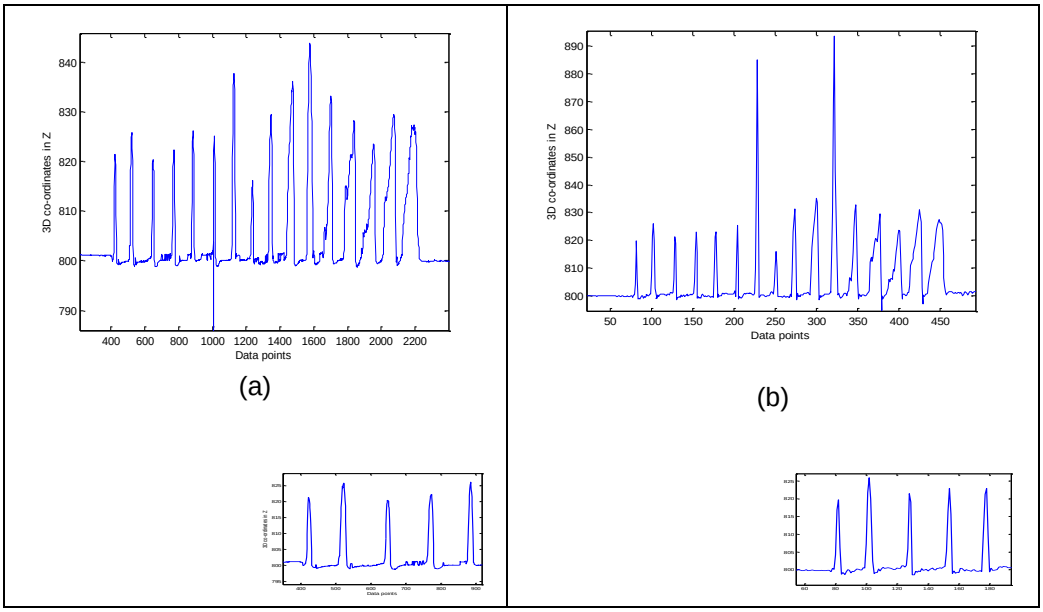


Figure 3-23: Time shifted instantaneous velocity profile of one tap for the index finger nail target for N-2, Trial-1; blue indicates the velocity profile for the six-camera system and red indicates the velocity profile for the two-camera system



(a) (b)

Figure 3-24:(a) Percent error in instantaneous velocity and (b) error in instantaneous velocity calculated for the index nail target on subject N-2, Trial 1



(a) (b)

Figure 3-25: Tapping profiles for the index nail target in (a) the six-camera system and (b) the two-camera system after coordinate transformation to align the two systems' reference frames.

Subject N-2, during Trial 1 was tapping normally. As seen in Figure 3-25, the first 5 taps showed almost the same values for both camera systems. However, the sixth tap

showed a very high value in Z for the two-camera system and the six-camera system reported zero at one time point, the seventh and eleventh tap showed also huge differences in amplitude. This increase in amplitude was due to a target missing from one of the cameras of the two-camera system.

The insets in Figure 3-25(a) and Figure 3-25(b) show the first five taps. If we observe the Z values from these insets, we observe that they were very close. This shows that the large errors are not due to transformation from the two- to the six-camera system's reference frame but rather due to the slower temporal sampling of the two-camera system that resulted in smaller numbers of data points between and during taps. Also, huge percent errors are seen for V2 as the six-camera system recorded a velocity of 0.0049 m/sec and the two-camera system recorded 0.038 m/sec

Table 3-34: Percent error, error in (m/sec) calculated for instantaneous velocity at four time instants for subject N-2, Trial 2

Instantaneous velocity (m/sec)		
N2-T2	Percent error	Error in magnitude
V1	19.7	0.139
V2	-98.31	-0.058
V3	-21.15	0.257
V4	-338.1	-0.071

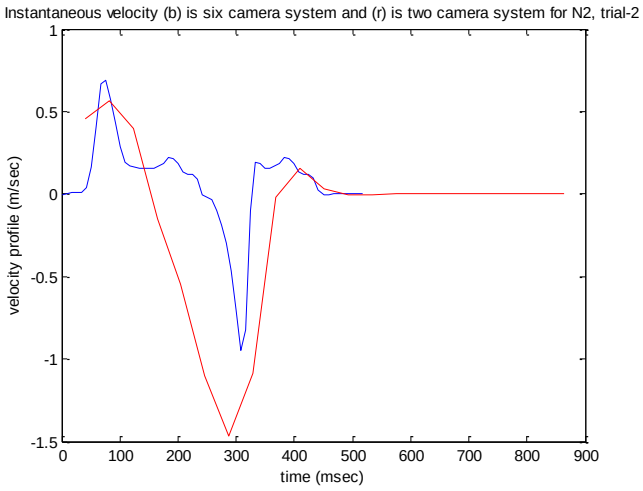
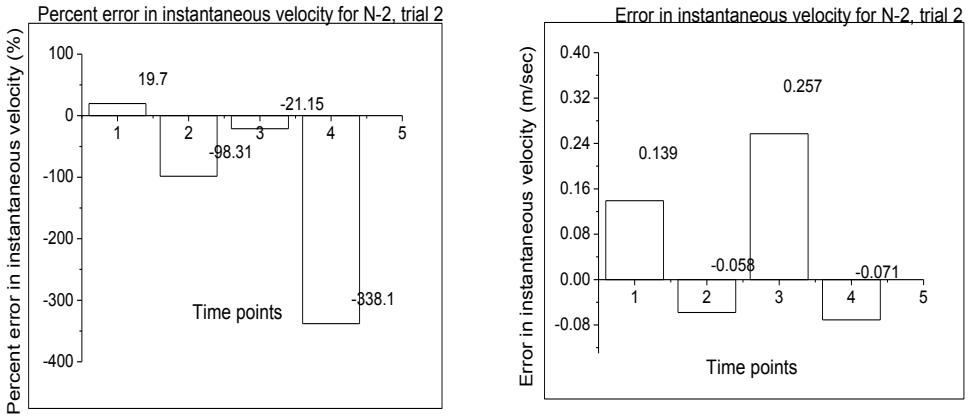


Figure 3-26: Time shifted instantaneous velocity profile of one tap for index finger nail target for N-2, Trial-2; blue indicates the velocity profile for the six-camera system and red indicates the velocity profile for the two-camera system



(a) (b)

Figure 3-27: (a) Percent error in instantaneous velocity and (b) error in instantaneous velocity calculated for the index nail target on subject N-2, trail 2

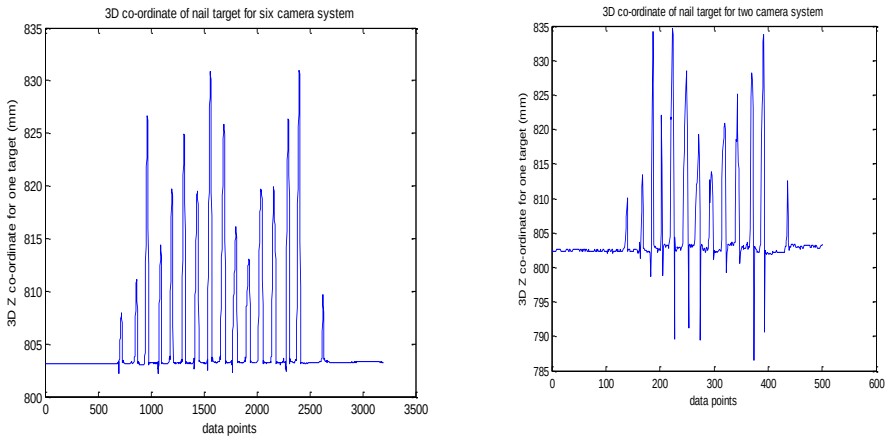
As seen from Table 3-34 above, it was observed that the error was high for V2 and V4. The six-camera system measured velocity of 0.059 m/sec for V2 and the corresponding velocity for the two-camera system was 0.117 m/sec. It can be observed

from Figure 3-26 that the differences between the two tap profiles, just where it goes into negative direction and just when it goes in positive direction is highest. It can be inferred that highest differences were observed in V1, V2, and V3 Figure 3-26 shows choppy velocity profile for two camera system at V1 as targets went missing at four time points.

.3.4.2.3 Subject N-1

Table 3-35: Percent error, error in (m/sec) calculated for instantaneous velocity at four time instants for subject N-1, Trial 1

Instantaneous velocity (m/sec)		
N1-T1	Percent error	Error in magnitude
V1	13.23	0.006
V2	-333.81	-0.014
V3	-74.96	0.074
V4	-16.67	-0.001



(a) (b)

Figure 3-28:Tapping profiles for the index nail target for (a) the six-camera system and (b) the two-camera system for subject N-1,Trial-1 after coordinate transformation to align the two systems' reference frames

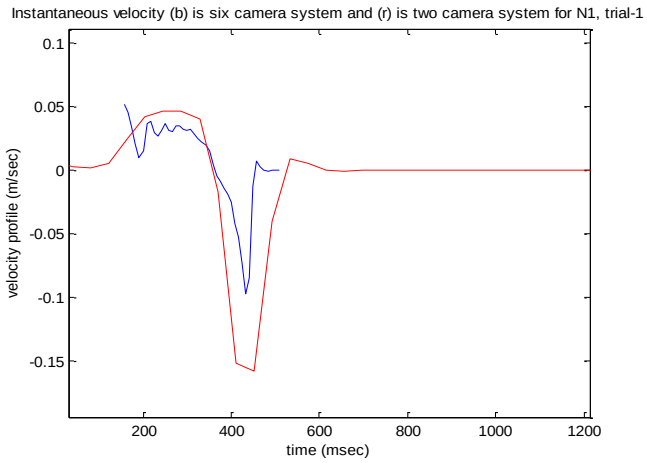
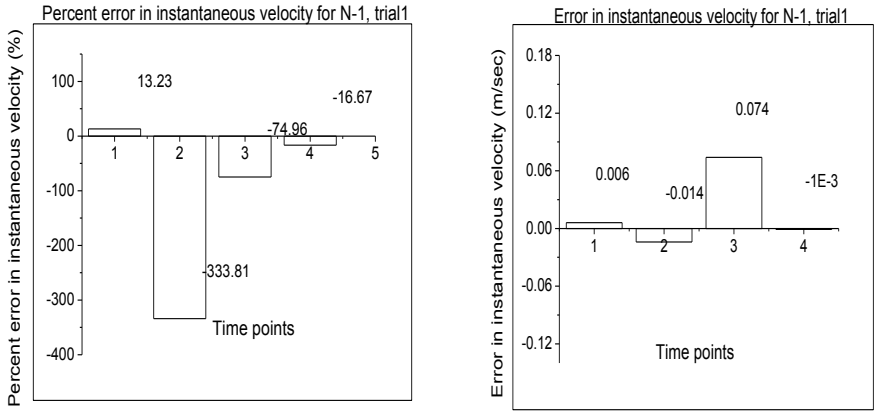


Figure 3-29: Time shifted instantaneous velocity profile of one tap for the index finger nail target for N-1, Trial-1; blue indicates the velocity profile for the six-camera system and red indicates the velocity profile for the two-camera system



(a) (b)

Figure 3-30: (a) Percent error in instantaneous velocity and (b) error in instantaneous velocity calculated for the index nail target on subject N-1, Trial 1

As seen from Figure 3-28, the six-camera system reported 15 taps while the two-camera system reported only 12 taps. The reason is that at many time points target went missing. Also, as seen from Instantaneous velocity profile in Figure 3-29, it was observed that the two-camera system showed higher velocities for V3 and V2 since the taps differed in amplitude. Huge percent error was observed in V2 with the six-camera system velocity being 0.004 m/sec while the two-camera system measured 0.018 m/sec as seen in Figure 3-30

Table 3-36 : Percent error, error in (m/sec) calculated for instantaneous velocity at four time instants for subject N-1, Trial 2.

Instantaneous velocity (m/sec)		
N1-T2	Percent error	Error in magnitude
V1	13.48	0.032
V2	-4.88	0.0002
V3	4.66	-0.013
V4	84.92	0.011

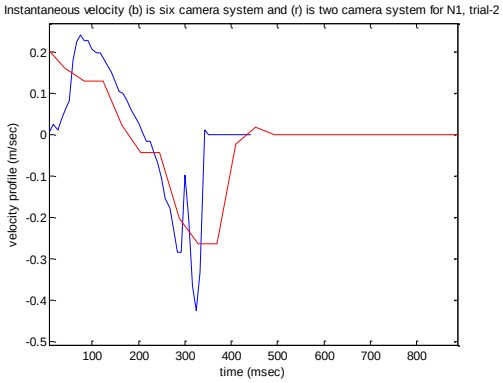
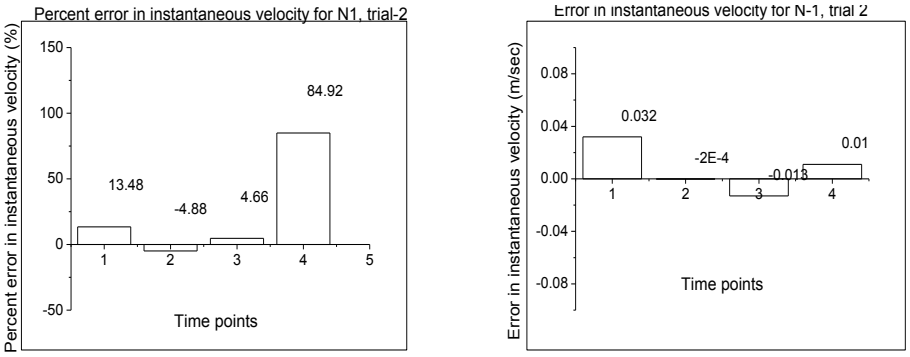


Figure 3-31: Time shifted instantaneous velocity profile of one tap for the index finger nail target for N-1, Trial-2; blue indicates the velocity profile for the six-camera system and red indicates the velocity profile for the two-camera system



(a) (b)

Figure 3-32:(a) Percent error in instantaneous velocity and (b) error in instantaneous velocity calculated for the index nail target on subject N-1, Trial 2

For N-1, Trial 2, huge error was observed in V4 just where the hand comes back to table and in the tap profile just where the tap goes towards the positive direction from the negative direction. As seen in Figure 3-31, the six-camera system showed a steep increase in slope while the two-camera system showed only a slow increase. The velocity data at V4 for the six-camera system was 0.012 m/sec and for the two-camera system it was 0.001 m/sec.

3.4.2.4 Subject M1-b

Table 3-37: Percent error, error in (m/sec) calculated for instantaneous velocity at four time instants for subject M-1b, Trial 1.

Instantaneous velocity (m/sec)		
M-1b,T1	Percent error	Error in magnitude
V1	50.67	0.211
V2	80	0.004
V3	38.69	-0.125
V4	97.78	0.009

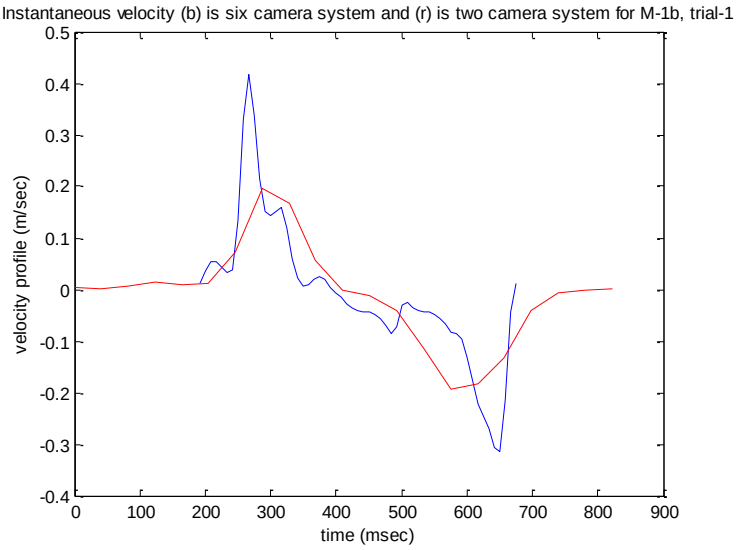
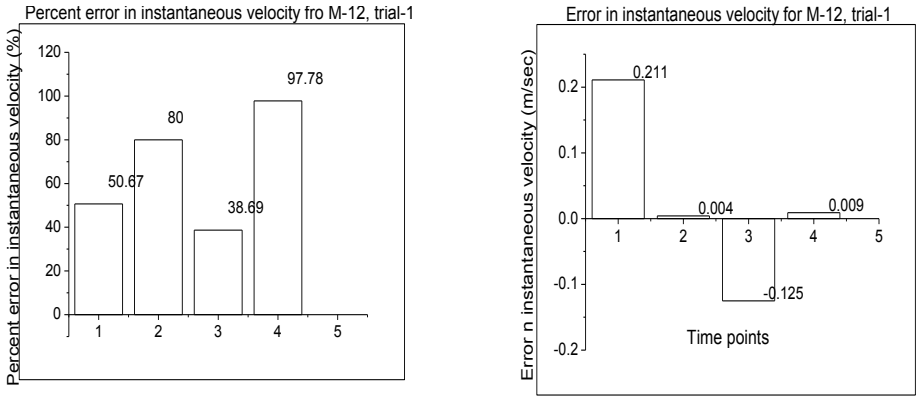


Figure 3-33 : Time shifted instantaneous velocity profile of one tap for the index finger nail target for M-1b, Trial-1; blue indicates the velocity profile for the six-camera system and red indicates the velocity profile for the two-camera system



(a)

(b)

Figure 3-34:(a) Percent error in instantaneous velocity and (b) error in instantaneous velocity calculated for the index nail target on subject M-1b, Trial 1

If we observe all subjects' data, we see that percent errors for the two- camera systems were very large primarily at V2 and V4. Also, the error magnitude in m/sec observed between the two camera systems for instantaneous velocities show less value. Such huge errors in V2 and V4 were due to the higher number of data points available for the six-camera system that showed slower decrease and increase in instantaneous velocities at V2 and V4 respectively. However, the two-camera system had huge intervals in between the same data points and, as a result; it showed faster changes for these data points. Such huge errors were not typically observed at V1 and V3, the point of maximum finger velocity in the positive and negative directions of the tap profile, respectively. It can be concluded that the two-camera system is marginally accurate with respect to six-camera system (barring any target occlusions).The larger errors seen near the ends of the finger trajectory occur because velocity numbers are very small in magnitude (the fingers should be almost immobile) and therefore and small magnitude differences between the two- and six-camera systems are amplified when expressed as percent errors.

3.4.3 Validation of the Coordinate Transformation from the Two-Camera to the Six-Camera System using the Tapping Average Velocity Metric

Average velocity measurements were calculated for two trials for each of the MACS-1, and the two Normal subjects. Percent error and error in magnitude for average velocity of 15 taps index nail target were estimated as shown in Table 3-38.

Table 3-38: Percent error and error in magnitude calculated for average velocity of 15 taps for the index finger nail target.

Average velocity (m/sec)			
Subjects	Trials	Percent error	Error

Table 3-38-contiuned

M-1a	T1	10.67	0.006
	T2	18.4	0.022
N1	T1	-72.2	-0.05
	T2	-35.39	-0.03
N2	T1	21.47	0.04
	T2	41.08	0.06
M-1b	T1	-6.59	-0.002

The percent errors observed for subject N1 was very large for the first trial. However, for all other subjects percent errors were less than 50%. The negative sign in some percent errors indicates that the two-camera system measured higher average velocity values than the six-camera system.

Chapter 4

Conclusions

This thesis had two aims. The first aim was to determine if motion metrics measured while children with CP tapped their hand would show significant differences before and after two weeks of CIMT. At the same time two age-matched controls were also measured during finger tapping and the hypothesis there was that these subjects should not have any significant motion metric differences between measurement sessions that also occurred about two weeks apart. The second aim of this work was to compare quantitatively the extent to which a lower-cost two-camera system could measure motion tracking metrics of finger tapping motions, for both children with CP and controls, as accurately as a much higher cost six-camera system.

Before any measurements were performed on human subjects, experiments were performed on a V-shaped calibration target made of Lego blocks. These measurements on static targets of known 3D coordinates were used to test the accuracy of both camera systems. It was found that the localization error for both systems was sub-millimeter in magnitude, which corresponded to sub-pixel image dimensions. Additional tests were performed to verify the high linearity of localization response across the field of view of both cameras of the two-camera system. Subsequently, finger tapping motion metrics (average tapping amplitude, average taping velocity and instantaneous velocity at different time points across the taping profile) were quantified on pediatric subjects with both camera systems. Due to the higher sensitivity of the two-camera system to target occlusions, it was possible to calculate motion metrics for only one MACS-1 subject (M-1a) before and after the treatment, and two normal subjects (N1, N2) with this system. However, for the less occlusion-sensitive six-camera system, the same

motion metrics could be calculated for two MACS-1 subjects, (M-1a, M-1b) and two control subjects (N1, N2).

For both camera systems it was observed that the tapping amplitude measurements for the MACS-1 (M-1a) subject showed significant differences after the treatment. Also, for the six-camera system it was observed that subject M-1b had a significant difference after treatment. However, for normal subjects too, there were significant differences in tapping amplitude between sessions, for both camera systems. It can be concluded that tapping amplitude is not a very strong metric for determining the effectiveness of treatment as it is influenced by the variability in tapping between measurement sessions.

The average tapping velocity measured for the MACS-1, M-1a, subject showed a significant change after treatment for both camera systems. Measurements performed for subject M-1b with the six-camera system did not show a significant difference. As per our hypothesis, normal subjects did not show significant changes in average tapping velocities with the exception of N2, as measured by the six-camera system. It is concluded that average velocity is a more promising metric than tapping amplitude for noticing changes in how a patient taps after therapy, but nevertheless remains a rather weak metric as the variability of tapping between sessions could still affect its value.

Instantaneous velocities measured for subject M-1a with the six-camera system showed significant differences across all of V1 – V4 time points of the tapping trajectory. However, when instantaneous velocities were calculated from the two-camera system measurements, M-1a showed less clear differences between the two trials with the exception of point V3. On the other hand, for subject M-1b, the six-camera system only showed marginal changes for the fastest points of the finger tapping trajectory. Encouragingly, normal subjects did not show significant differences between the two

trials. It is concluded that measurements pertaining to the instantaneous velocity profile may be useful surrogate measures for quantifying the effect of CIMT on hand use when tapping, though a lot more subjects would need to be measured to validate the currently observed trends in the data.

It should also be mentioned that clinical scores from the Melbourne and AHA tests were collected for all subject with CP at Texas Scottish Rite Hospital by the Occupational Therapists leading the CIMT study. Subject M-1a did not show a significant difference between before and after therapy with scores of 117 and 118, respectively. Also the AHA scores for the same subject were 84 and 88, respectively, which show a small but not clinically significant improvement. Clinical test measures for subject M-1b are known to be not significant, but the exact numbers were not available at the time of this thesis writing. Therefore the measured average and instantaneous velocity measures appear to mirror the pattern of changes seen in the clinical tests for these two subjects with CP, with a greater degree of change for subject M1-a. Interestingly, these motion metrics showed statistically significant changes after CIMT whereas clinical measures did not. It is therefore hypothesized that motion metrics may provide more sensitive measures of change after therapy compared to standard clinical tests. On the other hand clinical tests compare the overall use of the affected arm (Melbourne) or of both arms co-operatively (AHA), whereas motion metrics in this work only measured finger tapping, which is a narrow metric of hand use and cannot therefore be compared directly with the clinical test measures.

For the second aim of this work, the quantitative comparison of motion metrics between the two camera systems, the differences of the two-camera system relative to the six-camera one were computed as error magnitudes and percent errors in average tapping amplitude, average tapping velocity and instantaneous velocity at four time points

of the tapping trajectory. From the percent errors calculated for subjects M-1a, N-1 and N-2, it was observed that the two-camera system could track the amplitudes more accurately if the following conditions were satisfied:

- The positions of both cameras were such that all targets were inside the field of view of both cameras and there were no missing targets due to bending of the fingers that would obstruct viewing of the targets. The two-camera system did suffer from target occlusions which resulted in ambiguities in target localization by the 3D triangulation algorithm used.
- The data sampling of a camera should not depend on the number of targets being tracked. In the current system, if the two cameras saw different numbers of targets due to occlusions, they reported target coordinates at different sampling rates, which then required time interpolation in software to synchronize X, Y target coordinates before triangulation.

While comparing instantaneous velocities quantified by the two camera systems, it was observed that for both normal subjects and ones with CP, the higher temporal sampling, that was independent of the number of targets recorded, and the fewer target occlusions for the six-camera system resulted in very large percent errors with respect to instantaneous velocities measured by the two-camera system. Specifically, the six-camera system could record high velocities at V1, the maximum velocity point in the ascending tap and V3, the maximum velocity point in the descending tap. The two-camera system recorded values that were about half of the six-camera system's values at these time points, due to its inferior time sampling. In some cases the two-camera recorded larger velocities than the six-camera system, but in those instances the trajectories of the two-camera system had missing points due to target occlusions, which led to velocity overestimation

As the magnitude of static errors was determined to be very small in all instances, the main source of error in the comparison between the two camera systems was due to differences in time sampling. As compared to the six-camera system, which had a sampling frequency of 120 Hz, the two-camera system acquired data at a rate of less than 25 Hz when the number of targets was about 26. It is possible that comparisons between the two systems would be more favorable if a smaller number of targets were involved, which could increase the two-camera system's frame rate and, with appropriate target placement, reduce the probability of occlusions. Moreover, estimation of 3D co-ordinates was time consuming for the two-camera system, as it was done by simple home-grown software. The overall conclusion in terms of the comparison between the two systems is that one is better off using the higher cost six-camera system which allows more accurate motion tracking for the tapping speeds relevant to this subject population. If only the lower-cost two-camera system were available, it would be preferable to reduce the number of targets to the three nails of the hand so that sampling speed is as high as possible, but there are more than one targets being tracked as a back-up for possible occlusions.

References

[1] Capute and Accardo's Neurodevelopmental Disabilities in Infancy and Childhood, Third Edition. Edited by Pasquale J. Accardo, MD. 2008, Paul H. Brookes Publishing Co, Baltimore, MD. p17.

[2] Ahl LE, Johansson E, Granat T, Carlberg EB. Functional therapy for children with cerebral palsy: an ecological approach. Dev Med Child Neurol. 2005;47:613–619. doi: 10.1017/S0012162205001210.

[3] Campbell SK. In: Physical Therapy for Children. Campbell SK, Vander Linden DW, Palisano RJ, editor. St.Louis: Saunders; 2006. The Child's Development of Functional Movement; pp. 33–76.

[4] Valvano J. Activity-focused motor interventions for children with neurological conditions. Phys Occup Ther Pediatr. 2004;24:79–107. doi: 10.1300/J006v24n01_04.

[5] T. L. Sutcliffe, W. C. Gaetz, W. J. Logan, D. O. Cheyne, and D. L. Fehlings, "Cortical reorganization after modified constraint-induced movement therapy in pediatric hemiplegic cerebral palsy," *J Child Neurol,* vol. 22, pp. 1281-7, Nov 2007.

[6] Arner M, Eliasson AC, Rösblad B, Rosenbaum PL, Beckung E, Krumlinde-Sundholm L (2008) Manual Ability Classification System for children with cerebral palsy.

[7] Eliasson AC, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Öhrvall A-M, Rosenbaum PL (2006) The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. Dev Med Child Neurol 48:549–554

[8] D. Damiano, M. Abel, M. Romness, D. Oeffinger, C. Tylkowski, G. Gorton, A. Bagley, D. Nicholson, D. Barnes, J. Calmes, R. Kryscio, and S. Rogers, "Comparing functional profiles of children with hemiplegic and diplegic cerebral palsy in GMFCS Levels I and II: Are separate classifications needed?," *Dev Med Child Neurol,* vol. 48, pp. 797-803, Oct 2006.

[9] D. L. Damiano, M. D. Gilgannon, and M. F. Abel, "Responsiveness and uniqueness of the pediatric outcomes data collection instrument compared to the gross motor function measure for measuring orthopaedic and neurosurgical outcomes in cerebral palsy," *J Pediatr Orthop,* vol. 25, pp. 641-5, Sep-Oct 2005.

[10] Szabo, Zoltan, et al. "Objective Assessment of Patients with Parkinson's Disease." *3rd WSEAS International Conference on REMOTE SENSING, Venice, Italy*. 2007.

[11] R. Krupicka, Z. Szabo;M. Jirina " Motion Camera System for Measuring Finger Tapping in Parkinsons Disease" 5th European Conference of the International Federation for Medical and Biological Engineering pg 846-849 2012

[12] Huiyu Zhou; Huosheng Hu; , "Inertial motion tracking of human arm movements in stroke rehabilitation,"*Mechatronics and Automation, 2005 IEEE International Conference* , vol.3, no., pp.1306-1311 Vol. 3, 2005 doi: 10.1109/ICMA.2005.1626742

[13] S. Greaves, C. Imms, K. Dodd, and L. Krumlinde-Sundholm, "Assessing bimanual performance in young children with hemiplegic cerebral palsy: a systematic review," *Dev Med Child Neurol,* Jan 5 2010.

[14] O. Faugeras. Three-Dimensional Computer Vision: a Geometric Viewpoint. MIT Press, 1993

[15] O. Faugeras and G. Toscani. The calibration problem for stereo. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 15–20, Miami Beach, FL, June 1986. IEEE

[16] D. Gennery. Stereo-camera calibration. In Proceedings of the 10th Image Understanding Workshop, pages 101–108, 1979.

[17] Q.-T. Luong and O. Faugeras. Self-calibration of a moving camera from point correspondences and fundamental matrices. The International Journal of Computer Vision, 22(3):261–289, 1997

[18] Abdel-Aziz, Karara., Direct linear transformation into object space coordinates in close range photogrametry. In Proc. Symp. Close-Range Photogrametry, 1971: p. 1-18.

[19] Using Lego pieces for camera calibration: a preliminary study L. Baronti1 and M. Dellepiane1 and R. Scopigno11Visual Computing Lab, ISTI-CNR, Pisa. Italy

[20] J. Heikkilä and O. Silven, “A Four-step Camera Calibration Procedure with Implicit Image Correction” in Proceedings of the 1997 IEEE Conference on Computer Vision and Pattern Recognition, June 1997, pp. 1106-1112.

[21] Triangulation *Richard I. Hartley and Peter Sturm* GE-CRD, Schenectady, NY Lifia-Inria, Grenoble, France

[22] Fast stereo triangulation using symmetry WH Li, L Kleeman Australasian Conference on Robotics and Automation

[23] Natarajan, S.K.; Ristic-Durrant, D.; Leu, A.; Graser, A.; , "Robust stereo-vision based 3D modelling of real-world objects for assistive robotic applications," *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on* , vol., no., pp.786-792, 25-30 Sept. 2011

[24] Eng Swee Kheng; Abu Hassan, A.H.; Ranjbaran, A.; , "Stereo vision with 3D coordinates for robot arm application guide," *Sustainable Utilization and Development in Engineering and Technology (STUDENT), 2010 IEEE Conference on* , vol., no., pp.102-105, 20-21 Nov. 2010

[25] Wientapper, F.; Wuest, H.; Kuijper, A.; , "Reconstruction and Accurate Alignment of Feature Maps for Augmented Reality," *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2011 International Conference on* , vol., no., pp.140-147, 16-19 May 2011

[26] Mun-Soo Park; Ji-Wook Kwon; MinHo Park; Jung Su Kin; Suk-Kyo Hong; Sang Wan Han; , "Experimental Study on Camera Calibration and Pose Estimation for the Application to Vehicle's Wheel Alignment," *SICE-ICASE, 2006. International Joint Conference* , vol., no., pp.2952-2957, 18-21 Oct. 2006

[27] Arun, K. S.; Huang, T. S.; Blostein, S. D.; , "Least-Squares Fitting of Two 3-D Point Sets," Pattern *Analysis and Machine Intelligence, IEEE Transactions on* , vol.PAMI-9, no.5, pp.698-700, Sept. 1987

[28] Shinji Umeyama. 1991. Least-Squares Estimation of Transformation Parameters Between Two Point Patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* 13, 4 (April 1991), 376-380.

[29] D. W. Eggert, A. Lorusso, and R. B. Fisher. 1997. Estimating 3-D rigid body transformations: a comparison of four major algorithms. *Mach. Vision Appl.* 9, 5-6 (March 1997), 272-290.

[30] Heikkila, J.; , "Geometric camera calibration using circular control points," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , vol.22, no.10, pp. 1066- 1077, Oct 2000

[31] E. Shen, P. K. Carr, P. Thomas, and R. Hornsey, “Non-planar target for multi-camera network calibration,” in *Proc. IEEE Sensors*, Oct. 2009,pp. 1410–1414.

[32] H. Zhang, K. Wong, and G. Zhang, “Camera calibration from images of spheres,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 3, pp.499–503, Mar. 2007.

[33] H. H. Huang, L. Fetters, J. Hale, and A. McBride, "Bound for success: a systematic review of constraint-induced movement therapy in children with cerebral palsy supports improved arm and hand use," *Phys Ther,* vol. 89, pp. 1126-41, Nov 2009.

[34] Matthew J. Thornton, Matthew C. Morrissey, Fiona J. Coutts, Some effects of camera placement on the accuracy of the Kinemetrix three-dimensional motion analysis system, Clinical Biomechanics, Volume 13, Issue 6, September 1998, Pages 452-454, ISSN 0268-0033, 10.1016/S0268-0033(98)00001-1.

[35] Kevin J. DeLuzio, Urs P. Wyss, Jian Li, Patrick A. Costigan, A procedure to validate three-dimensional motion assessment systems, Journal of Biomechanics, Volume 26, Issue 6, June 1993, Pages 753-759, ISSN 0021-9290, 10.1016/0021-9290(93)90037-F.

[36] Marks, R. and Karkouti, E. (1996), Evaluation of the reliability of reflective marker placements. Physiother. Res. Int., 1: 50–61. doi: 10.1002/pri.47

Biographical Information

Ankita Chainani was born in Mumbai, India on December 25, 1988. She received her Bachelors of Science in Biomedical Engineering from Mumbai University in May 2010. She began her graduate studies in biomedical engineering from University of Texas at Arlington in Fall 2010. Her research was based on Motion tracking and Imaging. She has worked on diverse projects like Voice controlled Wheel chair prototype. She plans to work in a company profile as testing and design engineer.