Multi-spectral Imaging System Calibration and Data Fusion for Applications in the Assessment of Rheumatoid Arthritis

by

Susan Twelves

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfilment of the requirements for the degree of

Master of Applied Science

in

Biomedical Engineering

Carleton University
Ottawa, Ontario

The undersigned recommend to
the Faculty of Graduate and Postdoctoral Affairs
acceptance of the thesis

**Multi-spectral Imaging System Calibration and Data Fusion for Applications in the
Assessment of Rheumatoid Arthritis**

submitted by

Susan Twelves

in partial fulfillment of the requirements for

the degree of Master of Applied Science in Biomedical Engineering

_____
Chair, Dr. Howard Schwartz, Department of Systems and Computer Engineering

_____
Thesis Co-Supervisor, Dr. Monique Frize

Carleton University

_____
Thesis Co-Supervisor, Dr. Pierre Payeur

University of Ottawa

April 2012

Abstract

A multi-spectral stereo calibrated and synchronised system has been developed. It comprises one IR and one optical camera and an innovative calibration procedure and target which can obtain feature points from both IR and optical images. The target consists of a conventional chessboard pattern with corner LEDs. The LEDs remove the need for external heating of the pattern prior to IR imaging which eliminates the problem of the chessboard pattern diminishing over time, allowing indefinite IR feature point capture. This target provides an improvement over external heating: the IR calibration technique most frequently cited in the literature. Software algorithms, which improve the accuracy of the feature point locations, and the camera synchronisation hardware are described. Successful registration and fusion has been carried out on IR and optical images of a human hand, combining visual detail with temperature distribution which would allow the assessment of rheumatoid arthritis.

Acknowledgements

Firstly, I would like to thank both my supervisors Dr. Monique Frize and Dr. Pierre Payeur for their constant support and advice during the completion of this work. I am also extremely grateful for the opportunity they gave me, as a new immigrant, to carry out a preliminary investigation which was a forerunner to this project. Without that experience, it would have been much harder for me to re-enter the workforce in a technical field.

Secondly, I want to recognise my husband, Chris for his encouragement and understanding. Particularly for putting up with the hours I have spent working in the lab and office during evenings and weekends whilst I completed this project.

Finally, I want to dedicate this thesis to my Dad, who typed up the whole of my first undergraduate thesis on his portable typewriter for me, but who, sadly, will never get to read this one.

Table of Contents

List of Illustrations

List of Tables

1 Introduction

1.1 Context

The temperature distribution on a human body can be used as an indicator of muskuloskeletal diseases such as rheumatoid arthritis (RA) [1] [2]. Clinical studies have demonstrated that analysis of heat distribution (thermography) on patients can provide a quantitative measure of disease activity and inflammation in RA [3], [4], [5], [6], [7]. The work reported in [8], suggests that synovitis[1] can be detected through observing variations in thermal patterns which could, therefore, be used for early diagnosis. This is because painful areas tend to have different temperature distributions to healthy ones.

Although thermographic imaging has been investigated previously, earlier work was hampered by the relatively low resolution and large size of infra-red (IR) sensors and the limited power of computers for image analysis. With the advent of new technology these latter problems have been greatly reduced which, as will be shown, allows the development of a practical system which can carry out accurate calibration on both IR and intra-optical (IO)[2] cameras. Thus, it is now technically feasible to build a diagnostic system based on thermography and optical images that will allow quantitative assessment of arthritis. The combination of IR and IO images provides users with the benefit of visual detail which makes it easier to recognise e.g. which particular area of the body has a higher temperature than another.

---

[1] Inflammation of a synovial membrane which often causes pain on movement of the joint it is in

[2] Note that both the terms optical and IO are used throughout this thesis to refer to images or equipment that are pertaining to the visible band of the electro-magnetic spectrum.

This project involves the automatic registration and fusion of IR and more detailed IO images by using stereo calibration techniques. The motivation for this work is to provide a key component of a more complex system which could subsequently be used in the quantitative assessment of RA. The research carried out for this project is a stepping stone to producing a relatively inexpensive multi-spectral imaging system that could be made widely available to physicians with patients who suffer from RA. The fully developed system could be used to track the evolution of RA disease using only IR and IO imaging by quantifying the extent to which a patient's heat distribution changes to allow comparison over time as the disease is treated.

One of the key aims of this project was to build and test an innovative multi-spectral calibration approach which can be used to simultaneously calibrate IR and optical cameras in a composite network of cameras. For this purpose, an original calibration target has been developed that is based on a conventional chess board pattern with LEDs added at the corners. The system which was built for this thesis work consists of only one IR and one optical camera whereas it is envisaged that the extended system would contain significantly more cameras. Nevertheless, the work carried out here on the dual camera system will pave the way to developing the larger system as, for instance, the new target could be used to calibrate more than two cameras once it has been proven viable, which this report intends to show.

1.2 Objectives

The main challenges that were met whilst tackling multi-spectral calibration were:

- Synchronisation of the IR and IO cameras.

- Ensuring that there were shared feature points visible in both IR and IO spectra.

- Computation of calibration parameters for drastically different devices, i.e. the IR and IO cameras.

- Achieving alignment and correspondence of IR and IO images which have an inherently different nature from each other.

These challenges led to the formulation of six key objectives that the work undertaken for this project set out to achieve which are as follows:

Objective 1

Synchronise the frame capture between an IR and optical camera by constructing a hardware interface and then capture a series of synchronised images.

Objective 2

Develop a new type of calibration target which is, not only relatively simple and cheap to make, but also does not require the use of external heating and can be simultaneously detected by IR and optical cameras.

Objective 3

Use captured IR and optical images of the target to simultaneously calibrate both the IR and optical cameras to obtain their intrinsic and extrinsic parameters.

Objective 4

Successfully register captured IR and optical images of a body part by using the parameters obtained during the calibration process.

Objective 5

Determine the correspondence between IR and optical images of the body part to produce a fused image which accurately combines useful IR and optical information whilst preserving the information from both spectra.

Objective 6

Analyse the effectiveness of various camera system arrangements where:

a. The optical axes of the cameras are parallel but placed different distances apart.

b. The cameras are placed the same distance apart but the angle between their optical axes is varied: effectively placing the cameras around circles with different radii.

c. The cameras are placed at different distances apart and the angle between their optical axes is varied: effectively placing the cameras around a circle with the same radius, pointing in towards the centre.

1.3 Thesis Organisation

Chapter 2 includes the theory behind the stereo calibration process used here. It then continues to explore the current applications and techniques used to carry out multi-spectral image analysis with an emphasis on IR and optical imaging. In particular, it will highlight the pros and cons of using different methods such as machine vision stereo imaging techniques and intensity based image registration techniques. It distinguishes the work between those that are concerned with medical applications and other areas.

Chapter 3 is concerned with the system hardware. It includes a section on the design of the proposed calibration target which will describe the process used to determine the optimum number of feature points together with the layout of the features. This is followed by a section on the acquisition platform which will describe the IR and optical cameras as well as how the hardware and cameras are connected together. It will explain how the synchronisation between the two cameras is achieved. It also provides the system specifications which includes the limits in cable lengths between the various devices. Finally, there is a section that provides a step-by-step guide as to how to use the developed system to capture images.

Chapter 4 describes the various system software algorithms in detail. More specifically, there are descriptions of:

- How the feature points are detected in IR images of the target and subsequently ordered so that each feature point in any image of the target can be recognised individually. This subsection includes a discussion of the accuracy achieved using this method.

- The optical image corner feature point detection and the subsequent processing of these points.

- The experiments that were carried out on a smaller prototype target and their results. As a result of the experiments on the prototype target, a new larger target was designed.

- The improved target design. In addition to this, there is an extra layer of processing carried out on the IR images which is explained here.

- The optical feature point correction which is occasionally necessary if the number of points found falls below the expected number.

- How the stereo calibration data are used to rectify images of body parts. Images of hands are used to provide a specific example to illustrate how the algorithm works.

- The algorithm used to produce fused images of the IR and optical images along with the results associated with the analysis.

Chapter 5 provides an extensive set of results together with a critical analysis of them. More specifically, it analyses three different arrangements of the cameras: in parallel; inclined around different circles; inclined around the same circle. After examining the outputs from the various experiments, this chapter concludes by making recommendations on how to properly setup the system.

Chapter 6 concludes the thesis by providing a summary of how well the system works. It highlights the contributions that have been made, how well the objectives have been met and gives suggestions for future work in this area.

Appendix A provides a technical report on how to use the FLIR A320 camera.

2 Theory and Literature Review

This Chapter will give a detailed explanation for the theory behind stereo camera calibration. It will then present a review of current literature in the area of multi-spectral calibration and data fusion.

2.1 Theory

The theory used for carrying out the separate calibration of the individual IR and optical cameras is based on the approach introduced by Zhang [9] which built on the work described in Tsai's now classic paper [10]. With reference to [11], an explanation of this theory, and how it relates to this project, is given below.
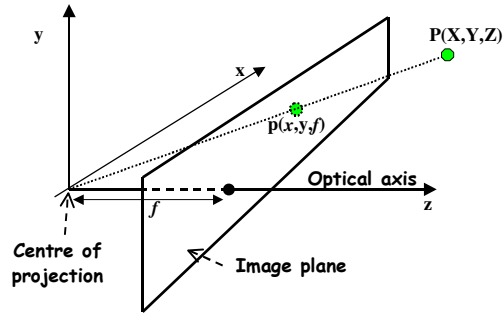
Figure 2-1: The perspective projection

Figure 2-1 shows the perspective projection of a point in the world reference frame, P(X,Y,Z), on to the image plane at point, p(x,y,f), where f is the focal length of the camera. This projection is based on that of a pinhole camera, except that rather than have the image plane at a distance f behind the centre of projection, it has been located at f in front of the centre of projection. This has the effect of preventing the images from being inverted, consequently simplifying the mathematical formulation. The optical axis is a line passing through the centre of projection which is also

orthogonal to the image plane. Observing similar triangles in the diagram, gives the following relationship:

$$x = f_x \left(\frac{X}{Z} \right) + c_x \qquad (2\text{-}1)$$

Where f_x is the focal length in pixels obtained with feature points in the x direction and c_x is the offset of the centre of the camera sensor from the optical axis, in the x direction. There is a similar equation for the y axis parameters. Although there is only one physical focal length, two results for focal lengths are output by the OpenCV *cvCalibrateCamera2* function because pixels are sometimes rectangular on the camera sensor rather than square. For instance, the focal length in e.g. mm is given by:

$$f_{mm} = s_x f_x = s_y f_y \qquad (2\text{-}2)$$

where s_x and s_y are mm/pixel in the x and y direction respectively. Note that the focal length is defined in the z axis direction and the x and y subscripts here are just referring to which coordinates in the feature points are being used to obtain the focal length. However, the cameras used with this project both have camera sensors with square pixels, therefore, f_x and f_y should always be equal.

These relationships allow us to write an equation for point q, on the image plane, in terms of point P, on the object plane:

$$q = MP \qquad (2\text{-}3)$$

Where M contains the so-called intrinsic parameters:

$$\begin{bmatrix} x' \\ y' \\ w \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad (2\text{-}4)$$

Here, point q is in homogenous image plane co-ordinates and $w = Z$, from equation 2-4. Therefore, to recover the definition of x in equation 2-1, it is necessary to divide through by w (or Z) as follows:

$$x = \frac{x'}{w} = \frac{x'}{Z} \qquad\qquad (2\text{-}5)$$

However, in the case of equations 2-2 and 2-3, the object point P has been described in the same frame of reference as the camera frame. A more general case is where P is described with respect to an object frame of reference. In order to express it in the camera frame, it has to undergo a rotation and translation as shown in Figure 2-2. These rotation and translation parameters are referred to as the extrinsic parameters. Here r_1, r_2 and r_3 are the columns of the rotation matrix and are referred to individually as a rotation vector.
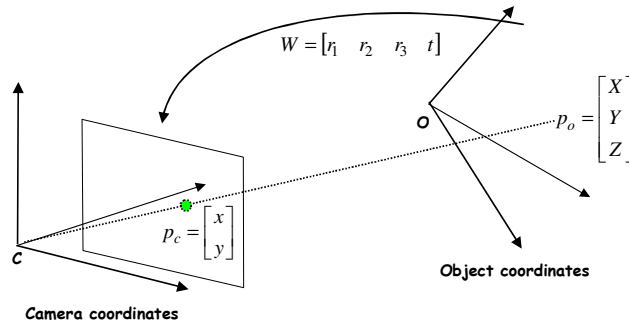
Figure 2-2: The extrinsic parameters

If only a planar target is used, as is the case here, then the Z parameter can be set to 0 with no loss of generality. In addition to this, r_3 in Figure 2-2 can also be omitted because it just represents the rotation of the Z parameter around the three axes and, once Z is set to 0, it will have no effect. However, in order to allow the use of the 3x3 matrix, M in equation (2-3), the points are written in homogenous co-ordinates, which add an extra dimension to each of them. Thus, after setting $Z = 0$, the equation to describe Figure 2-2 is:

9

$$p_c = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = sM\begin{bmatrix} r_1 & r_2 & t \end{bmatrix}\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \qquad (2\text{-}6)$$

where s is a scaling factor, r_1 and r_2 are the rotation vectors around the x and y axes respectively and t is the translation vector. The rotation and translation vectors effectively move the origin of the object co-ordinate frame, O, to the origin of the camera reference frame, C.

By collecting several images of the same planar calibration target, with known X, Y values for a set of feature points on the target, there is a closed form solution to calculate the intrinsic parameters in equations (2-4) and (2-6). This solution is based on two constraints and the fact that the rotational matrix is an orthonormal one. If the so-called homography matrix, H, is defined as follows:

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} = sM\begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \qquad (2\text{-}7)$$

Then the two constraints are:

$$h_1^T M^{-T} M^{-1} h_2 = 0 \qquad (2\text{-}8)$$

$$h_1^T M^{-T} M^{-1} h_1 = h_2^T M^{-T} M^{-1} h_2 \qquad (2\text{-}9)$$

The first constraint arises because the rotation vectors r_1 and r_2 are orthonormal and the second because of this orthornormality and the fact that the magnitudes of the rotation vectors are equal.
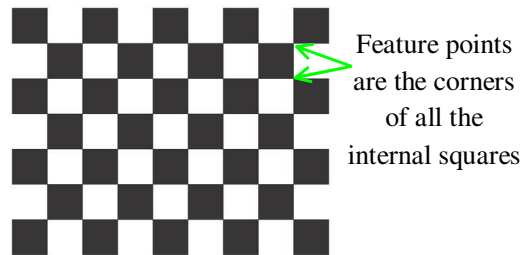
Figure 2-3: Conventional optical calibration target

Each image of a planar calibration target can give 8 independent equations because the feature points are the corners of squares, see Figure 2-3. This is a consequence of the fact that only up to 4 corners of the target provide independent information and there is a pair of x and y value for each of those corners. This is because four points gives the information related to anything a planar perspective view can do, i.e. they can stretch a square into any shaped quadrilateral.

Ignoring distortion, there are 6 extrinsic parameters (3 rotational and 3 translational) and 4 intrinsic parameters (2 focal lengths and 2 image centre offsets) associated with each image of the target. Whereas the 6 extrinsic parameters change with each image, the 4 intrinsics remain the same. Therefore, out of each set of 8 independent equations for each image, 6 of the equations are needed to obtain the 6 extrinsic parameters but this leaves only 2 equations for the 4 intrinsics. Thus, at least 2 images of the target are required to obtain the closed form solution to all 10 parameters.
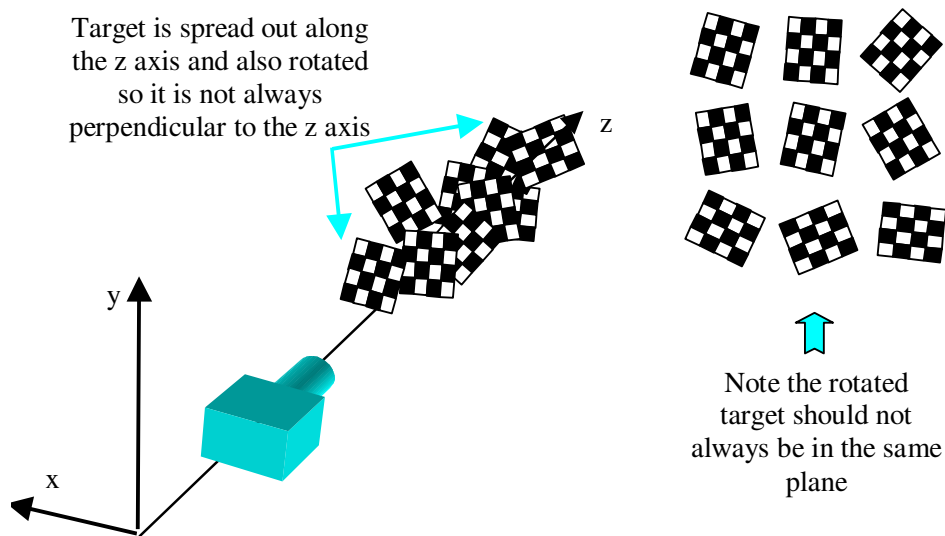
Figure 2-4: Multiple target capture is required to find the extrinsic and intrinsic parameters

However, in addition to this, there are distortion and noise factors to be taken into account, which cause numerical instability. These parameters are found using the closed form solution as a starting point before carrying out a least mean squared error minimisation process on many images of the target, see Figure 2-4. It is particularly important to move the target sufficiently between each image capture to ensure that the points are not all co-planar, to give independency between the feature points. During the experiments for this project, the target was rotated in both directions about the y axis as well as the z axis. It was also moved to different distances from the cameras along the z axis. The actual path moved by the target during the experiments is described in detail in Section 4.3.

Once the intrinsic and extrinsic parameters for the individual cameras have been obtained, stereo calibration can be carried out where the geometrical relationship between the two cameras in space can be computed, as depicted in Figure 2-5.
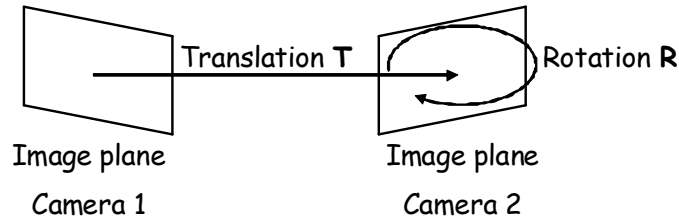
Figure 2-5: Rotation and translation vectors relating the two camera planes

Stereo calibration gives the rotation and translation vectors between the two cameras where e.g. a point P in each camera's co-ordinates are related by equation (2-10).

$$P_l = R^T \left(P_r - T \right) \qquad\qquad (2\text{-}10)$$

where P_l is the point in the left camera's co-ordinates, P_r is the point in the right camera's co-ordinates, R is the rotation vector between the two cameras and T is the translation vector between them. These vectors are related to the extrinsic parameters

obtained from the calibration of the individual cameras by equations (2-11) and (2-12).

$$R = R_r (R_l)^T \tag{2-11}$$

$$T = T_r - RT_l \tag{2-12}$$

where the subscripts r and l denote vectors relating to the right and left cameras respectively. Thus, once the rotation and translation parameters have been found for each separate camera, they can be substituted into equations (2-11) and (2-12) to obtain the rotation and translation vectors that relate the left and right camera planes to one another. An iterative minimisation technique is used to find the best estimate of the vectors.

Stereo rectification can now be carried out which is the mathematical process of aligning the cameras so that their principal rays are in parallel and every pixel row in one camera is exactly aligned with the pixel row in the other camera. Effectively it is the process of correcting individual images from each camera so that they appear to have been taken with two cameras with pixel row aligned image planes in a frontal parallel configuration, i.e. parallel to the *x-y* plane, as shown in Figure 2-6.
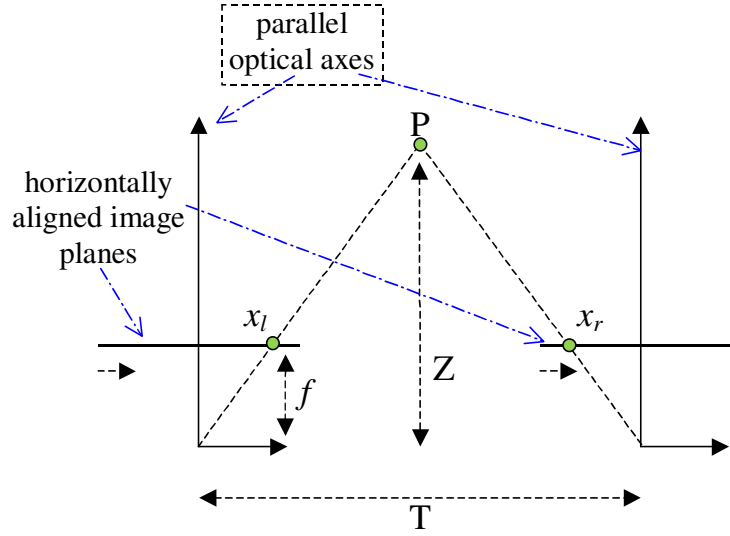
Figure 2-6: Rectified stereo camera planes with parallel principal rays and aligned pixel rows

Once one of the camera planes has been rotated by R, the two cameras will have their image planes in parallel. However, this does not give a unique rectification between the two cameras because there are an infinite number of possible frontal parallel planes to choose from or, more simply put, whilst the planes are aligned in parallel to the *x-y* plane, they can still be shifted to any position with respect to one another in the *y* or *z*-axis direction. In order to align the images row-by-row in this thesis project, Bouguet's algorithm [12] was used which applies the epipolar constraint, as will now be explained.

In a stereo system, an epipolar point on the left image plane is the image of the centre of projection (O_r in Figure 2-7 (a)) of the right camera on the left plane. Thus, in Figure 2-7 (a), e_l is the left camera image plane epipole and e_r is the right camera image plane epipole. The epipolar line for a point P in either image plane, is the line connecting the image of P in the camera plane (P_l or P_r in Figure 2-7 (a)) and the epipoles. The significance of, for example, the epipolar line in the left camera plane,

is that it represents the image of the line between the centre of projection in the right camera, O_r, and the point, P, taken by the left camera.

The point P_r in the right image plane could have been produced by any point on the line between the image plane and the point P. Similarly, the point P_l could have been formed by any point on the line joining it to point P. The image of this line P_lP in the right hand plane is the right epipolar line e_rP_r. In order to find corresponding points on the two image planes, it is therefore, only necessary to search along the corresponding epipolar line in one image plane for each point in the other image plane. This reduces the computational time enormously and is known as the epipolar constraint.
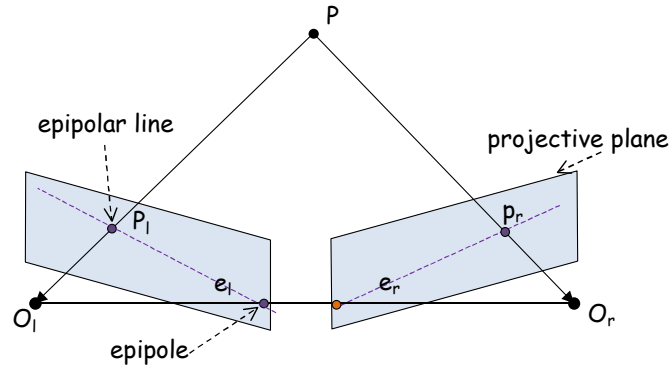
(a)
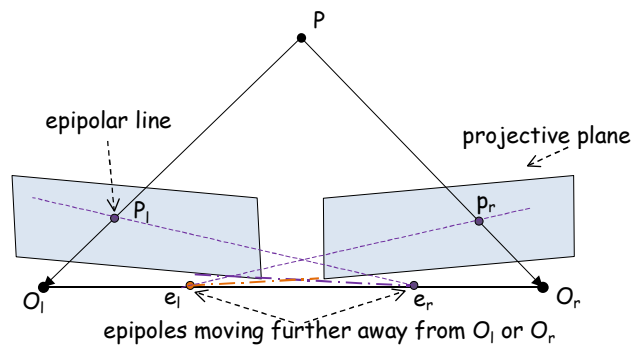
(b)
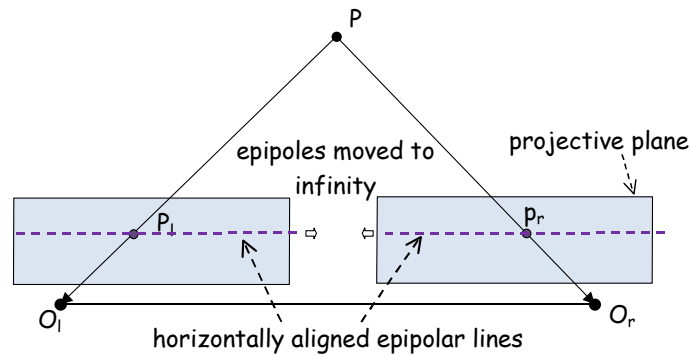
(c)

Figure 2-7: Epipoles and alignment of epipolar lines

In addition to this, if the image planes can be rotated so that the epipolar points move to infinity, the epipolar lines in both images will become horizontally aligned with each other, as shown progressively in Figure 2-7(a) to (c). Thus, the search for

correspondence of one point in, e.g. the left-hand image will be reduced to searching the same row in the right-hand image.

Bouguet's algorithm [12] provides a way of mathematically rotating the image planes to horizontally align the epipolar lines. This algorithm carries out the alignment by minimising the image reprojection distortion whilst maximising the overlap between the left and right camera viewing areas. Instead of using the rotation matrix R on one of the cameras to give it the same orientation as the other one, the algorithm splits the rotation R equally between the two cameras into two matrices r_l and r_r. This splitting of R reduces the reprojection distortion and places the cameras in coplanar alignment but the rows are still misaligned. The next step in the algorithm calculates the rotation matrix which is necessary to move both epipoles to infinity. Using the left camera as an example, three orthogonal unit vectors are needed to define the epipole at infinity frame of reference. The first epipole's direction is along the translation vector T between the two centres of projection. Thus, the first unit normalised vector can be taken as this direction vector, e_1 and can be calculated as shown in equation (2-13).

$$e_1 = \frac{T}{\|T\|} \tag{2-13}$$

The second epipole's direction is orthogonal to e_1, by definition, and also is taken to be along the image plane. This is found by normalising the cross product of e_1 and the principal ray and is given in equation (2-14).

$$e_2 = \frac{\left[-T_y \quad T_x \quad 0\right]^T}{\sqrt{T_x^2 + T_y^2}} \tag{2-14}$$

Finally, the third normalised vector has to be orthogonal to both e_1 and e_2 so can simply be obtained by finding the cross product between them, see equation (2-15).

$$e_3 = e_1 \times e_2 \qquad (2\text{-}15)$$

These three normalised vectors can be used to obtain the rectification rotation matrix, R_{rect}, see equation (2-16). This matrix can be used to take both epipoles to infinity so that the two cameras are row aligned.

$$R_{rect} = \begin{bmatrix} (e_1)^T \\ (e_2)^T \\ (e_3)^T \end{bmatrix} \qquad (2\text{-}16)$$

R_{rect} can be combined with the two 'half' rotation matrices r_l and r_r to give one rotation matrix for both cameras, as shown in equations (2-17) and (2-18).

$$R_l = R_{rect} r_l \qquad (2\text{-}17)$$

$$R_r = R_{rect} r_r \qquad (2\text{-}18)$$

Thus, the result of applying Bouguet's algorithm gives the ideal alignment which is depicted in Figure 2-6. For a more comprehensive explanation about the stereo calibration process, see [9] and [11].

2.2 Literature Review

This section will explore the current applications and techniques used to carry out multi-spectral image analysis with an emphasis on IR and IO imaging. Much has already been written about registration between IR and IO images. There are two broad approaches:

- Feature based computer vision stereo imaging techniques which usually involve a calibration target.

- Intensity or region based image registration techniques.

There follows a discussion of these two approaches which will highlight their pros and cons.

The feature based stereo imaging approach is based on the theory that has already been described in Section 2.1. The work undertaken for this thesis falls into this first approach category. It also relies on the epipolar geometry constraint to solve the correspondence problem. However, this constraint is often used in both approaches.

Intensity based techniques rely on finding correlation between regions of an image in e.g. the left stereo camera and those in the right stereo camera. Current research in the field has been focussing on so-called maximisation of mutual information techniques, which is another way of utilising the inherent entropy of the images. This is related to the well known Shannon communications theory for entropy [13].

A good overview of multi-sensor data fusion is available in [14], which covers a wide variety of different applications for multiple sensor data fusion, including medical imaging.

The following review has been split into two main sections, the first is concerned with papers for medical applications and the second section is for all other papers concerned with multi-spectral analysis. Both these sections have been split into two sub-sections, one which is concerned with work using feature based computer vision techniques, which usually incorporate a calibration target, and those that involve intensity or region based techniques.

2.2.1 Multi-spectral Imaging in Medical Applications

2.2.1.1 Feature based techniques

Askenov et al. [15] describe a system where two IO cameras and an IR camera were combined to produce a 3D thermogram, which could then provide quantitative monitoring of inflammation for clinicians. To do this they combined a 3D scanner with a thermal camera where all the optical axes were aligned in parallel and located

19

adjacent to one another, with the IR camera sandwiched between the two IO cameras. The IO cameras were treated as a stereo pair, each producing very high resolution images with 16 megapixel sensors. The IR camera was an Indigo Systems Merlin camera. A purpose built heated calibration target was used which consisted of some 14 circular holes aligned in rows. Although the target material was not specified, it was presumably metallic. The stereo matching process used during the calibration was that of Zhengping's [16] and the calibration process was carried out on all three cameras. The optical image information produced a range map, which was then used to obtain a 3D surface in the form of a triangulated polygon mesh. Finally, the IR image pixels were mapped on to this 3D surface.

The work included a method for standardising the thermogram images by registering the sequence of 3D IR images with one another so that they were effectively viewed from the same observer point. Thus, subtraction of the images could be undertaken to measure differences between the thermograms which provided a means for a quantitative analysis of inflamed areas over time. Some experimental examples of facial and leg thermograms were given where the images were subtracted and then had standard image processing techniques applied to them to obtain the inflamed regions of interest.

Other attempts have been made to combine thermographic imaging with 3D surface reconstruction. Ju et al. [17] opted for a combination of traditional stereoscopic approaches to obtain 3D reconstruction of limbs over which heat distribution, collected via an IR sensor, is mapped after registration between the sensors. The resulting multi-modal models are used in the diagnosis of equine inflammatory lameness.

20

Cornejo and Algorri [18] were concerned with image-guided neurosurgery. In particular, their work was about the development of the motion tracking and capture part of a neurosurgical system, using two IR cameras. In this work, the IR cameras were off-the-shelf CCD sensor webcam cameras where the IR blocking filters, normally included in these cameras, were replaced with visible light blocking filters instead, which is an interesting way of obtaining an IR camera at low cost. The system consisted of two of these cameras connected via a PC. Each surgical instrument had 5 IR emitting marker LEDs attached to it – 4 of these provided position and orientation information whereas the last one was just an identifier.

In [18], a very simplified set of assumptions were made: the focal lengths of both cameras were taken to be equal; the right and left y-axis image values were assumed to be equal; the translation x distance between the two cameras was measured manually; the two cameras were manually aligned in parallel and it was subsequently assumed that the optical axes of both cameras were also in parallel. The work presented in this thesis does not rely on any of these simplified assumptions. However, sub-pixel accuracy was obtained in [18] by finding the centre of mass of the IR images of the LEDs, which is the same method used for this project, see Chapter 4.

2.2.1.2 Region based techniques

Multi-spectral imaging has already been used to good effect within the medical community. The idea of superimposing thermal and visual images has already been explored for 2-D static medical images [19], and other applications [20]. Schaefer et al. [19] carried out skin detection prior to maximisation of mutual information. A step by step approach to an automatic detection of the skin in human visual images is presented using a log opponent colour transformation. For image registration a gradient descent optimiser was used where mutual information was the similarity

metric. A simple weighted sum of 0.8*IO + 0.2*IR was used to obtain overlaid IR/IO images.

Gomez et al. [21] developed an integrated imaging system to obtain accurate and reproducible multi-spectral dermatological images. This system can collect up to 10 different spectral bands (UV to IR) and was developed to allow comparative studies of time series of images. This system has already shown that the multi-spectral approach can be used successfully to segment lesions easily and to track the evolution of dermatological diseases. Similar multi-spectral imaging techniques have been used to classify melanoma [22]. In the latter, it is shown that the multi-spectral images add lesion depth and structural information to the imagery.

2.2.2 Multi-spectral Imaging in Other Applications

2.2.2.1 Feature based techniques

Varshney [14] gives an example of fusing two images from an IR sensor and a millimetre wave (MMW) sensor to show a concealed weapon. Registration of the images is carried out through pyramidal decomposition using a wavelet transform. A feature selection rule is used to fuse the transformed pyramid before carrying out an inverse transform to get the fused image.

Adopting a different approach, Hilsenstein [23] developed a thermographic stereo imaging device that combines two IR sensors from which images are processed to extract and match feature points. Here a specially designed calibration target was developed by milling into a printed circuit board to produce a chessboard pattern alternating between copper squares and the base material of the circuit board. The target was heated with a hairdryer before imaging to produce high contrast IR images between the copper and base material squares, because of their different emissivity

properties. The application was 3D modeling of wavy water surfaces given the optical properties of water that create more visible texture in the IR spectrum than in the IO spectrum.

Wang and Sung [24] superimposed visible face images on to their IR image counterparts after calibrating the IR and optical cameras to find their geometrical relationship. The calibration target used here was a grid consisting of two orthogonal planes. By selecting calibration points on each of these planes, non-singular calibration was ensured. The target grid could be used by the optical and IR cameras by placing holes in every grid point. The target was then illuminated from behind which gave the holes a different temperature to the rest of the target. This temperature variation could then be captured by an IR camera, highlighting the holes in the image. In [24], the line joining the corners of the eyes was viewed as being parallel to the line joining the corners of the mouth. These lines were then used to estimate the head pose with respect to the optical camera.

The technique worked for this facial application, because the working volume was smaller than that required for more general body part images, e.g. legs and arms, which are typical examples of images that would be analysed for the broader diagnostic aim following on from the work presented in this thesis. This rear illumination method would not be as convenient as the separate LEDs in the planar target described in Section 3.2, because in order to achieve non-singular calibration, the target must be moved into different planes along with the illumination. Separate LEDs allow for this required manoeuvrability.

Ababsa et al. [25] calibrated IR and optical cameras as part of their augmented reality (AR) tracking system. AR is where real world objects are merged into virtual space, in order to give virtual objects more realistic behaviour. This was a real-time system

which effectively continually updated the camera extrinsic parameters, that is, how a camera is orientated with respect to a world co-ordinate reference frame. It was a system which required the user to wear 3D vision glasses and fiducial markers. The calibration was carried out using a combination of the known fixed fiducial markers and IR target positions with respect to the IR camera frame. They used Zhang's method [9] to obtain the camera intrinsic parameters (see Section 2.1) for the optical vision system. They also made use of a pose estimation algorithm on joystick images, which also had markers on it, to obtain the relative position of the IR and optical tracking systems. Therefore, this system relied on markers to obtain the required calibration. However, one of the aims of the project presented in this thesis is to carry out calibration without the use of external markers.

To remove the inconvenience of attaching markers to a subject, so called markerless gait analysis has been investigated for biometric applications such as surveillance as the overview given by [26] shows. This overview emphasises that markerless gait analysis is in its infancy and still has many obstacles to overcome. In particular, this article discusses the various methods used for markerless gait analysis such as model based versus holistic. In addition to this, explanations are provided as to how feature extraction can be used to simplify gait recognition along with the benefits of carrying out Fourier analysis.

Vidas et al. [27] compared a conventional chessboard target, which was irradiated with a heat lamp in order to calibrate an IR target, with their technique of using a mask with holes of 2cm squares. The mask was placed over a heated background which provided an IR camera with contrast between the open squares of the mask and the surrounding unheated square frames. They used OpenCV calibration functions in order to carry out the calibration with their mask. They showed an improvement in

calibration accuracy of up to 78% over the conventional method. However, one of the drawbacks both with the new mask and the irradiated chessboard grid techniques was that the IR image of the chessboard diminishes over time. In this paper even the mask image was somewhat degraded after only 30 seconds whereas with the target discussed in this thesis, it is possible to obtain IR images of it over an indefinite period.

Yahyanejad et al. [28] use a conventional paper-based chessboard pattern as their target together with a heat lamp to irradiate the pattern. Since the black squares absorb the heat more than the white ones, an IR image of the irradiated chessboard will reveal a pattern of alternating squares, very similar to that obtained with a conventional optical camera. Uniform irradiation of the target is difficult to achieve, therefore, they take several images of the target with the lamp concentrating its radiation on a different part of the target each time. They then obtain a more uniform image of the target by stitching the parts that have been irradiated most from all the images. The centroids of the squares are then used as the feature points as the squares produced are still not that well defined at the edges. However, since the pattern of the chessboard is known in advance, the centroids of the squares can be used to produce accurate chessboard corner positions. Once they have obtained the corner feature points, Bouguet's Matlab calibration tool from Caltech [12] is used to carry out the calibration. In particular, they were interested in improving the measure of the second order radial distortion over conventional methods. They quote a reprojection error of 0.36 and 0.38 pixels in the x and y direction which is an improvement over the 0.49, 0.51 pixel error quoted for using just the automatic Matlab corner detection method with their target. They also quote an average 17% improvement in the mean squared error (MSE) during registration of IR and red-green-blue (RGB) chromatic images

when the images are undistorted. However, this technique also relies on irradiating a chessboard target which requires access to a heat lamp. The additional requirement to stitch several images together to obtain a uniformly irradiated image is more complex than the one proposed in this thesis.

In [29], Ng and Du were trying to map the 3D temperature surface of a car using stereo calibration techniques. Their target comprised a metal net over a plastic board where the net was a pattern of squares. A similar "net and board" approach was also used by Xiao et al. [30]. They blew hot air over the target and the metal heated up more than the plastic so that the IR images showed a pattern of squares. The Matlab calibration tool box [12] was then used to carry out stereo calibration on images captured by placing the target on a turntable and moving the turntable through 20 degrees between image capture. Their quoted reprojection error was somewhat better than [28] – 0.177 and 0.175 pixels. In this case, however, their setup was more sophisticated and complex than [28].

Cheng et al. [31] were interested in tracking the hand and head movements of a driver using both 4 IR cameras and 4 conventional cameras. However, they did not combine the IR and IO images but analysed the different modalities separately. They used a simple grammar to work out the driver action at any particular time. Their target was a conventional chessboard which was irradiated for a few seconds by a flood lamp, as described in [28]. Again the Matlab calibration toolbox was used to carry out the calibration after image capture. The head and hand segmentation was carried out using skin segmentation with the IR images using an upper and lower temperature threshold.

For a tele-immersive system, Johnson and Bajcsy [32] developed a multicamera multimodal system using IR and IO cameras. As with the system described in this

thesis, they used Point Grey IO cameras and a FLIR IR camera. They were interested in segmenting images into foreground and background. They used the information from the IR camera to detect people and the IO cameras to detect foreground objects. Unfortunately, although the paper describes the way in which they align the IO/IR images, it only states that the extrinsic and intrinsic parameters are used to do this but not how they obtained the parameters themselves. There is also no information about the calibration target that they used.

Svoboda et al. [33] modified a conventional laser pointer to include a coloured transparent plastic cap. This device is used as a calibration tool by producing many feature points for a system which incorporates 12 cameras. The constraint with this system is that at least 3 cameras are required. However, as the project for this thesis is a forerunner for a system that will eventually use at least 3 cameras, this constraint is not considered a problem here. However, unlike the work for this thesis, [33] only considers optical cameras. The device is moved through the working volume and a set of feature points are extracted from the captured images. These feature points are then used to compute the nonlinear distortion.

2.2.2.2 Region based techniques

Both Pohl et al. [34] and Klonus et al. [35] give good overviews of image data fusion techniques. Other examples of multi-sensor data fusion are where 3D Colour and IR imagery have been combined to develop classifiers of objects in the field of autonomous off-road robotic navigation [36]. Another classification system is used for locating and tracking a human in the vicinity of a robot [37]. This system combines 2 colour sensors, multiple passive infrared sensors and a pair of microwave sensors. Here, a data fusion algorithm has been developed, based on Dempster-Shafer theory [38] [39] to decide whether or not a human exists in a particular area.

However, in these latter two cases, unlike the work presented in this thesis, the cameras have been treated as individual sensors.

A remote sensing example is given in [40], where a classifier is produced by using a series of CCD cameras to obtain images in the UV through near infrared regions and uses Maximum Likelihood algorithms for clustering.

In [41] two separate systems were tested for mapping 3D surface temperatures on to objects using computer vision techniques. Neither system used a purpose built calibration target. The first experimental setup included two visual cameras with one IR camera mounted between them. The system was calibrated so that the position of the IR camera was known precisely with respect to the visual cameras. Two objects were imaged during the testing: one was a cylindrical chessboard pattern which was heated from within the cylinder and the other was a powered laptop. The correspondence between the two visual camera images was obtained using correlation techniques and then triangulation was used to reconstruct the 3D surface. The known geometry of the IR camera position with respect to the visual cameras was used to back project the IR image temperatures on to the 3D surface.

The second experimental setup consisted of two IR cameras mounted in parallel. Again a correlation technique was used to find the correspondence between the images from each camera. This time, they used isotherms in the images together with the epipolar constraint so that they searched for e.g. the 30°C points along an epipolar line in the right hand image to correspond with a 30°C point in the left hand image.

The two IR camera system had the advantage that it was independent to external lighting conditions. However, the surface reconstruction only provided a thermal map of the object and gave less detail than the visual surface because of the more diffuse

nature of IR images. The disadvantage of the combined visual and IR camera system was that the position of the IR camera had to be precisely calibrated from a geometric point of view in order to be able to accurately back project the IR images on to the surface. This geometric calibration would have been time consuming and highly error prone for the application considered in this thesis.

In [42], Yang et al. use the wavelet transform to carry out image fusion between IR and optical cameras. They use different fusion rules for the approximation subband to the detailed subband. They carry out fusion based on different methods: 1) take the minimum value, 2) take the maximum value 3) take the average value. Then the inverse wavelet transform is carried out. An objective measure is used to decide which of the fused image results is best. This objective measure is obtained by firstly applying the Sobel edge operator to the fused image. This gives the orientation and magnitude associated with each pixel. By comparing the input images and the fused image, a value is obtained of how well the edge magnitude and orientation values have been preserved which is used as a performance measure. By carrying out several fusion methods for each image, and then selecting the best one, this system can automatically adapt to different lighting conditions which is important for some applications, such as a military environment. In the example results that they gave, using the average value for fusion of the approximation coefficients and the maximum value for the detailed coefficients gave the highest performance measure in both cases of an outdoor scene. However, they did not provide details as to how the IR and visible images were registered before fusion took place.

The work presented in this thesis falls into the category of multi-spectral registration of images through the use of feature based computer vision techniques with a purpose built calibration target. What distinguishes this work from that described in the above

review is that the target used here is relatively simple to make and does not have a time limit associated with it so that IR target images can be captured indefinitely. Even though the target is simple and cheap to make, it will be shown below how the system still achieves good accuracy through being able to use feature based computer vision calibration techniques as opposed to region based ones.

2.3 Summary of Chapter 2

Chapter 2 provides a detailed explanation of the theory behind the stereo calibration process using feature based computer vision techniques. This is followed by an in depth review of the current literature covering multi-spectral imaging using combinations of IR and IO cameras. This review is divided into two parts, the first of which examines medical applications and the second part considers all other types of application. In particular, the review highlights the short falls of current multi-spectral calibration targets which motivates a significant contribution brought by this thesis in the design of a new calibration target.

3 System Description

This chapter describes the system hardware. A system overview includes a discussion about its characteristics. The overview is followed by the following sections:

- A section on the calibration target which will describe the process used to determine the optimum number of feature points together with the layout of the features.

- A section on the acquisition platform which will describe how the hardware and cameras are connected. It will explain how the synchronisation between the two cameras is achieved.

- A section on the image capturing process.

3.1 System Overview

The system used to carry out the stereo calibration is shown in Figure 3-1. An IR camera (see Section 3.3.1) and an IO camera (see Section 3.3.2) are aligned so that their optical axes are approximately parallel. Images are taken with both cameras of a checker board calibration target which has had LEDs added to it at alternating corners of the squares. The IR camera can detect the light from the LEDs. The two cameras are connected together via the camera synchronisation hardware (see Section 3.3.3).
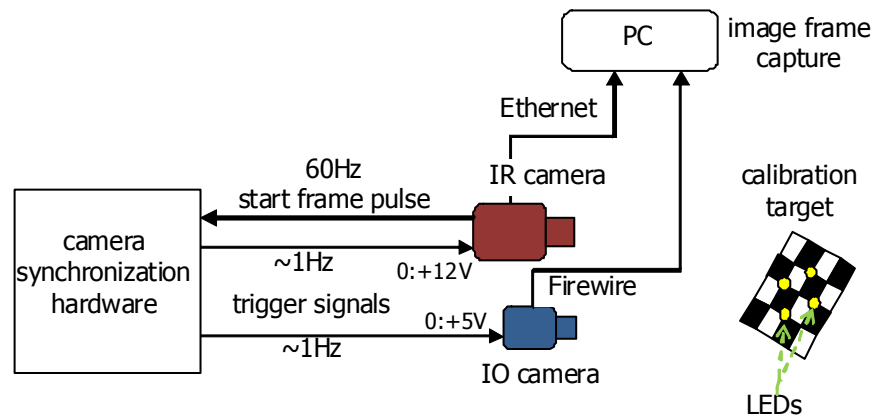
Figure 3-1: IR/IO camera calibration system

In order to carry out stereo calibration with the system, and later on data fusion, it is necessary to capture a series of frames from each camera which have been synchronised together. Due to network bandwidth limitations, the successful transmission and capture of the IR camera's frames as JPEG images was limited to a maximum of 1Hz. Even at this rate, some IR frames are dropped. Therefore, the IO camera was synchronised to the IR camera to transmit images at the same 1Hz rate. The start of each internal frame pulse in the IR camera is used to trigger the transmission of two JPEG images which are the current frames captured in the IR and IO cameras. The output frame pulse from the IR camera has a frequency of 60Hz, however, the synchronisation hardware divides this by 64 to obtain a 15/16 Hz square wave signal which is used to trigger both cameras. There is a slight delay between the IO and IR camera triggers because the IR trigger signal has to be pulled up to a +12V voltage rail which is typically 20ns but can be as high as 30ns. This delay is considered to be insignificant for the purposes of IR/IO image calibration and fusion given that the calibration target or human limb do not move at high speed.

The rest of this chapter describes this system in detail commencing with the calibration target. An explanation follows to show how the two cameras are

interconnected and synchronised. The specifications of the system are also given, including constraints associated with the various system cables. Finally, the image capturing process is described.

3.2   The Calibration Target

The first part of the calibration process carries out the individual calibration of an IR and an optical camera. A new approach has been implemented, where the individual calibration of both types of camera is carried out simultaneously using the same calibration target. The system developed here has built on the existing computer vision functions provided by OpenCV [11].

In order to carry out simultaneous intrinsic and extrinsic calibration of optical and IR cameras, it is necessary to have a calibration target that allows feature point detection with both types of camera sensor. The prototype target that was built and tested to see if it could be used to calibrate the cameras with sufficient accuracy, is shown in Figure 3-2a. Here it can be seen that 26 LEDs have been placed at the corners of the chess board which can be detected by the IR camera. After carrying out a series of experiments, it was decided to increase the number of LEDs to 50, as can be seen in Figure 3-2b, in order to improve the accuracy of calibration. The reason for this change and the details of these preliminary experiments, together with the description of the software algorithms used to carry out feature detection, are given in Section 4.4.
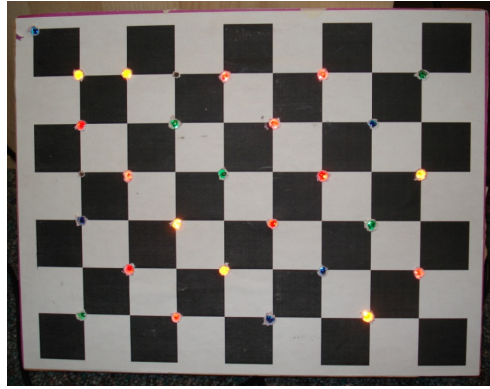
a) Prototype target
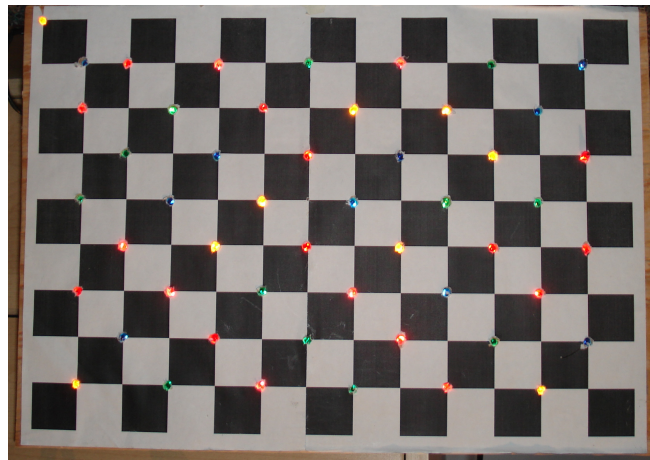
b) Final target

Figure 3-2 Combined IR/Optical calibration target

The chess board pattern can also be detected by a conventional optical camera. The LEDs have been placed at every other square corner apart from the four LEDs at the top left of the board, where three of these are placed at the corners of the same square and the fourth is an outlier placed at the extreme edge of the board. These latter LEDs are used to help determine the order of the LEDs, as will be described in Chapter 4.

The main reason for using LEDs rather than incandescent bulbs is because LEDs are cooler. This means that they are not only easier to handle by the person carrying out the calibration but also they do not create such large blobs in the binary image created during the feature point detection process, see Section 4.2.1. This reduces the

possibility of the blobs overlapping one another which would complicate feature point detection.

The final target was constructed out of a plywood board with dimensions of 28×40cm. The chessboard pattern was printed on a laser printer and stuck to the board with glue. The holes for the LEDs were hand-drilled at every other corner of the chessboard pattern. The LEDs were off-the-shelf consumer decorative lights which were powered off the 120V mains supply. Two small handles were attached to the back of the board to allow the user to move the target about in front of the cameras without his/her hands causing spurious high temperature areas to appear in the IR images. Such spurious objects in the IR images would have otherwise increased the complexity of the IR feature point detection algorithm, see Section 4.2.1.

3.3 The Acquisition Platform

3.3.1 FLIR A320 IR Camera

The IR camera is a FLIR A320 which can stream images over 100MB Ethernet. The built-in alarm feature is triggered externally causing the camera to automatically transmit the current frame via FTP as a JPEG image. More details of the triggering mechanism are given in Section 3.3.3 where the synchronisation hardware is described.

Although the IR camera is capable of producing 16 bit thermal images with a resolution of 320 x 240 pixels, they are saved here as interpolated 8-bit 640 x 480 pixel images so that they are the same size as those produced by the IO camera. It is possible to overlay the IR images with isotherms which will map out lines of constant temperatures. This feature could be used, for example, in a diagnostic system to find a region of interest where the body part has an elevated temperature above a threshold,

which could be set by the user. The A320 software development kit (SDK) can be used to inversely deduce the temperatures from the IR images directly. The SDK does this conversion via the *ToTemperature* method which converts a given absolute pixel value to temperature in Kelvin. It is also possible to set the image type to consist of absolute temperature pixels which means that every pixel is effectively the temperature in Kelvin.

The following is a summary of the camera's technical data which is relevant to this project [43]:

Thermal sensitivity	70mK @ +30°C
Max frame rate	30Hz
Close focus limit	0.4m
Object temperature range	-20°C to +120°C
Accuracy	±2% of reading
Spatial resolution	1.36 mrad
Field of view (FOV)	25° x 18.8°
Focal length	18mm
Size	170x70x70mm
Weight	0.7kg

Table 3-1: FLIR A320 IR camera specifications

As can be seen, in Table 3-1, the device is relatively light and compact compared to other IR cameras. This makes it ideally suited to mounting on a lightweight gantry.

3.3.2 Point Grey Flea2 IO Camera

The intra-optical (IO) Flea2 camera has an 800Mb/s interface via the IEEE-1394b, so-called Firewire B standard. This high speed interface allows transmission of RGB images. A summary of the camera's technical data which is relevant to this project is listed in Table 3-2.

dimensions	29 x 29 x 30 mm
weight (excluding optics)	58g
Max frame rate	60Hz
Focal length	6mm
Resolution	640 x480
Signal to noise ratio	> 60dB
Operating temperature	0 to 45°C
Operating relative humidity	20 to 80%

Table 3-2: Point Grey Flea2 IO camera specifications

As can be seen in Table 3-2, the camera is relatively small and light, like the IR camera, which makes it well suited for mounting on a gantry of lightweight construction.

3.3.3 Synchronisation hardware

The schematic for the synchronisation hardware can be seen in Figure 3-3. The hardware requires a 12V dc supply. However, some of the circuit uses CMOS ICs, so a +5V voltage regulator (LM7805) is employed to provide a +5V voltage rail.
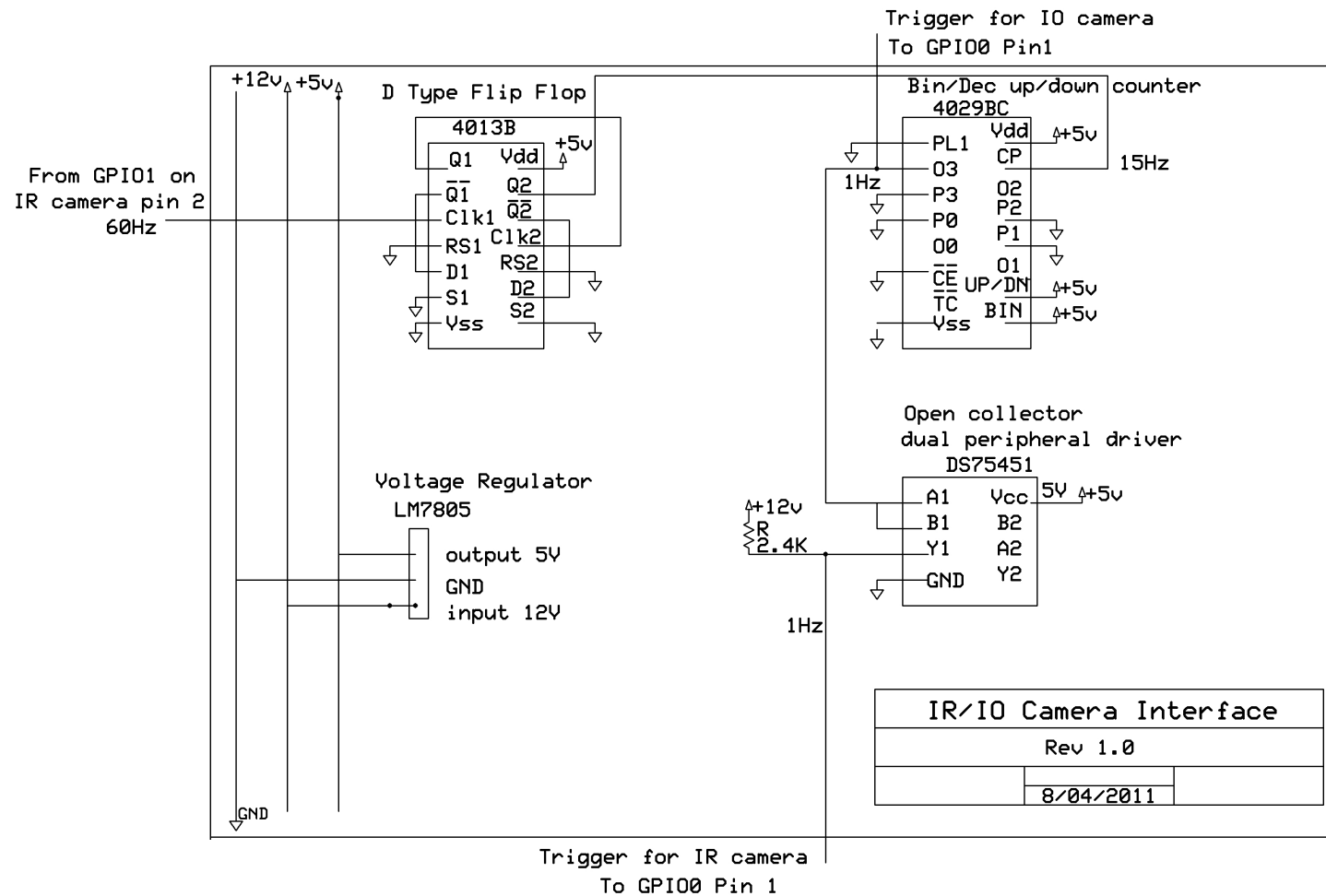
Figure 3-3: Schematic for the IR/IO camera synchronisation hardware
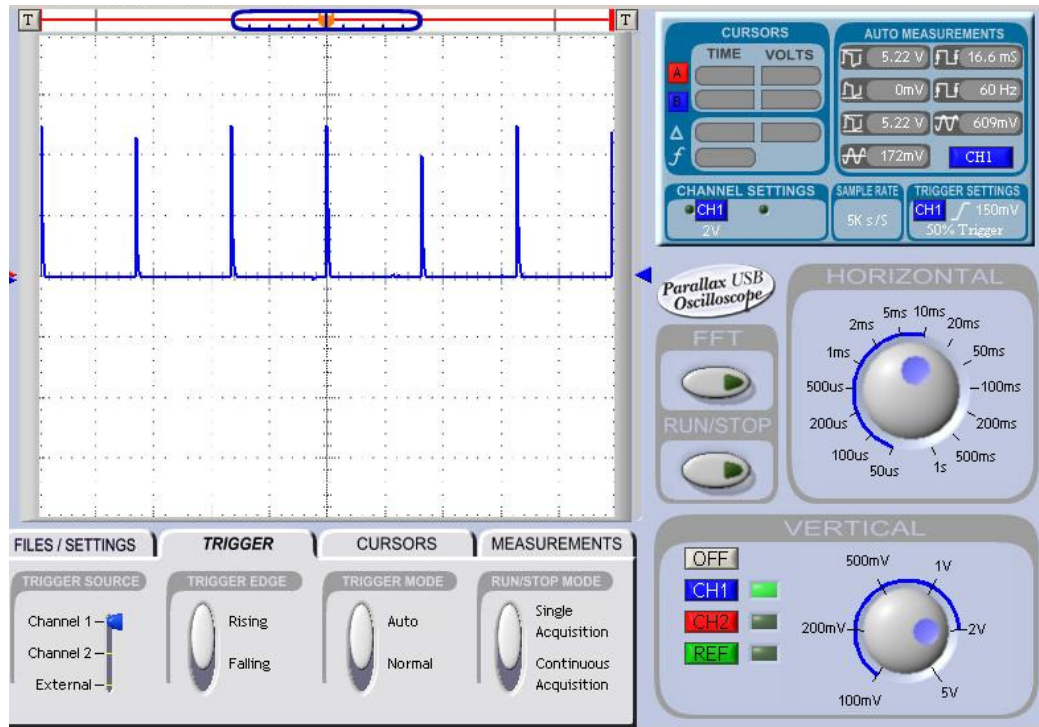
Figure 3-4: 60Hz Frame synchronisation pulse from the IR camera

The frame synchronisation pulse from the IR camera is the 60Hz input signal to the hardware from the A320 general purpose input output (GPIO) 1 pin 2, see Figure 3-4. This signal is divided by four using the CMOS D-type flip-flop (4013B) to produce a 15Hz signal that is then passed to a binary up/down counter (4029BC) where it is reduced to a frequency of ~1Hz (actually 15/16Hz), as shown in Figure 3-5. This signal is then output directly to the IO camera's GPIO0, pin1 because this camera has a CMOS interface.
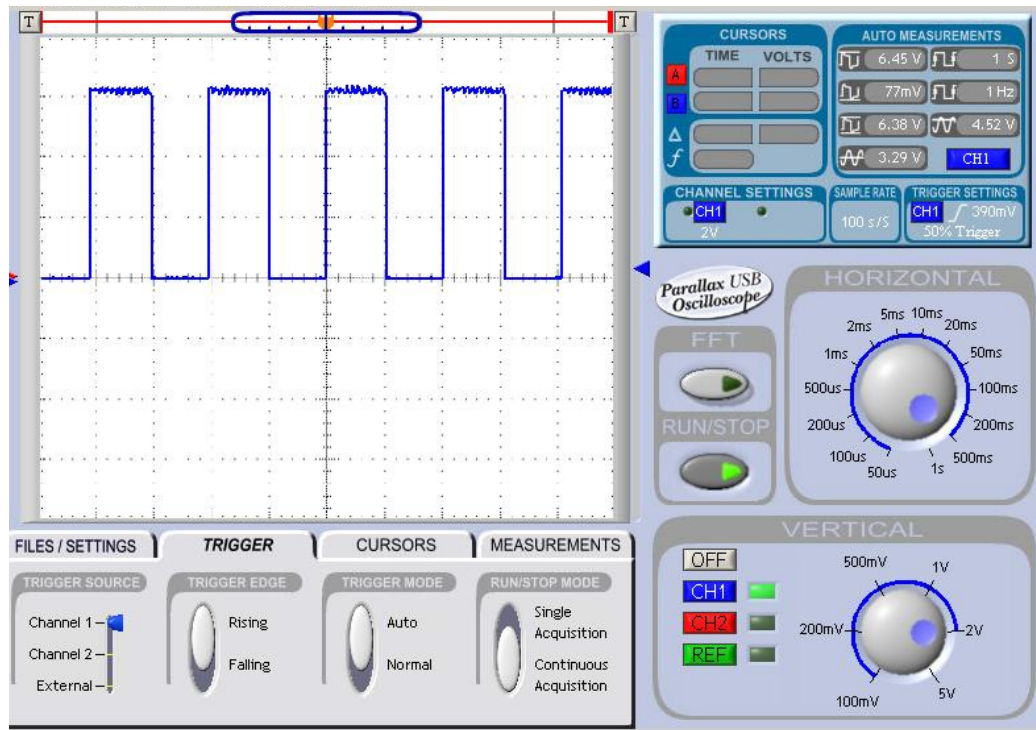
Figure 3-5: The 1Hz trigger signal

However, the IR camera requires a higher voltage pulse for its input trigger signal. Therefore, the signal is passed to an open collector peripheral driver (DS75451) where the output is pulled up to the 12V rail by a 2.4KΩ resistor.

A photograph of the complete system can be seen in Figure 3-6.
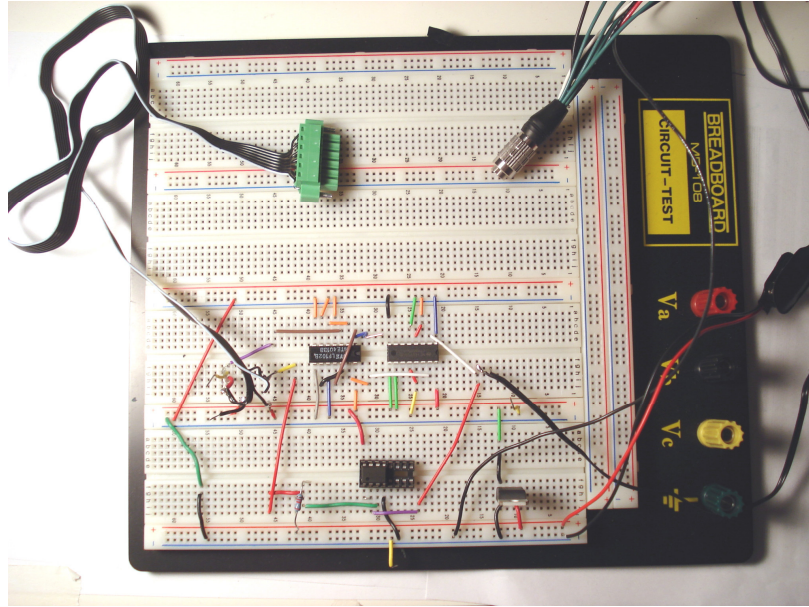
Figure 3-6: The synchronisation hardware

3.3.4 System Specifications

This section gives the system specifications together with some recommended sizes for the cable lengths which interconnect the various components.

The maximum recommended length for the CAT-6 Ethernet cable that connects the IR camera to the image frame capture workstation is 100m. However, as it is envisaged that the final system will not be operated remotely, but rather in the confines of one room, this constraint is not a problem. It is also possible to connect the camera to the workstation indirectly via a network hub.

The recommended maximum cable length for the IEEE-1394b, that connects the IO camera to the image frame capture workstation, is 4.5m [44]. However, it is possible to extend this maximum length to 10m with the use of 9-pin shielded twisted pair copper cable. If repeaters are used, distances of more than 100 meters can be achieved [45].

The CMOS 4029 up/down counter output can drive up to 50 CMOS or TTL inputs. The Point Grey GPIO input pin is a TTL 3.3V pin which has been designed so that it can also be driven by 5V logic. Therefore, the synchronisation circuit has no practical limit here for the proposed application.

Although the cable for a digital I/O port that connects the IR camera to the synchronisation hardware can theoretically have a maximum length of 100m [46], the limit for line length for TTL/CMOS logic is generally known to be ~3m [47]. Therefore, the IR and IO cameras can only be placed this distance from the synchronisation hardware. This has implications for the extension of the network to multiple IO cameras. However, the Point Grey 'MultiSync' software can be used to synchronise other cameras to the camera which is connected to the IR/IO synchronisation hardware via a 1394b bus [48].

3.4 The Image Capturing Process

The process for using the system to carry out stereo calibration, using images of the target, and then fusion of the IR and IO images of a body part is described below.

In order to start triggering the two cameras to transmit their frames as images, the following process needs to be undertaken:

1. Firstly, connect the two cameras to the PC, using a Firewire cable for the IO camera and an Ethernet cable for the IR camera, possibly via a network hub.

2. Start up the FTP server software. FileZilla FTP software was used in this thesis project [49].

3. Connect the camera synchronisation hardware into a +12V DC supply.

4. Use the FLIR 'ipMonitor' software to setup an alarm for the camera triggered from the input to GPIO0, pin 1. This software allows the alarm to cause the camera to transmit an image via the file transfer protocol (FTP).

5. Start the Point Grey Flycapture software and use it to set up the GPIO0, pin 1 port for the input trigger signal. Use the same software to decide how many frames are to be captured once the triggering starts. Typically 50 frames were captured each session for this project.

6. Find out the IP address for the IR camera using the FLIR ipConfig software.

7. Telnet to the IR camera using its IP address.

8. At the Telnet prompt, type 'rset .power.settings.digOutFunc1 "startFrameRise" '. This causes the IR camera to output its frame pulse signal on GPIO1, pin 2. It will also start triggering the IR camera alarm every second to transmit IR images to be saved on the workstation via FTP. This trigger will also activate the Point Grey Flycapture software to capture the e.g. 50 IO frames.

9. Once all the IO frames have been captured, turn off the IR camera alarm, by deselecting the activate alarm check box, to stop the IR images from being transmitted and saved.

3.5 Summary of Chapter 3

Chapter 3 is devoted entirely to describing the physical multi-spectral calibration system including all its off-the-shelf equipment as well as its purpose built hardware. There is a whole section on the multi-spectral calibration target which depicts both the smaller prototype 26 LED target and the final 50 LED design.

After this, each major component of the acquisition platform is described separately. Firstly, the IR and IO cameras used in the system are shown. This is followed by a detailed explanation as to how the synchronisation hardware was designed and built.

Some recommendations for the maximum lengths of various system cables are then listed which dictate the maximum size of the system set-up. The most limiting factor is the ~3m separation between the two cameras and the synchronisation hardware, due to the TTL/CMOS logic line length limit. However, this is not considered a significant problem for the medical assessment application being considered since the planned multi-spectral examination platform will remain fairly small to make it affordable and easy to use. However, the project described in this thesis was carried out with the IR and IO camera separation limited to ≤ 50cm due to the length of the synchronisation hardware connecting cables.

Finally, a step-by-step guide is presented as to how the image capturing process is carried out using a PC connected to the calibration system.

At this point, images from both cameras have been captured of the target image. This process must be repeated for capture of the body part images. The next steps to achieve fusion involve post processing the captured images. The following chapter describes the software algorithms that have been incorporated into the post processing to carry out the necessary image processing.

4   Multi-spectral Calibration, Registration and Data Fusion

This chapter describes the various system software algorithms in detail, it comprises:

- A summary of the processing carried out on the multi-spectral camera system.

- An explanation as to how the feature points are detected in images of the target and subsequently ordered so that each feature point in any image of the target can be recognised individually.

- An explanation as to how the stereo calibration data are used to rectify images of body parts. Images of hands will be used to provide a specific example to illustrate how the algorithm works.

- A description of the algorithm used to produce an overlay of the IR/IO images.

4.1   Summary of the Multi-spectral Camera System Processing

Figure 4-1 provides an overview of the calibration and fusion process which includes the following steps being carried out on the IR and IO camera system.

1. The target is moved through a variety of positions whilst the IO and IR cameras simultaneously capture images of it.

2. Chessboard corner feature point detection is carried out on the set of IR and IO image pairs.

3. If possible, missing feature points are filled in on the known target grid pattern.

4. The feature points are sorted so that they can be represented by the same integer co-ordinates in the corresponding IR and IO images.

5. If necessary, the feature points of both types of image are more accurately repositioned to reflect the known ideal regular grid pattern of the chessboard with LEDs target.

6. Individual camera calibration and stereo calibration is carried out with the OpenCV function *cvStereoCalibrate* using Zhang's algorithm [9].

7. Rectification between the two cameras is carried out with the OpenCV function *cvStereoRectify* using Bouguet's algorithm [12].

8. Finally, the rectified IR and IO images are merged together using data fusion techniques.
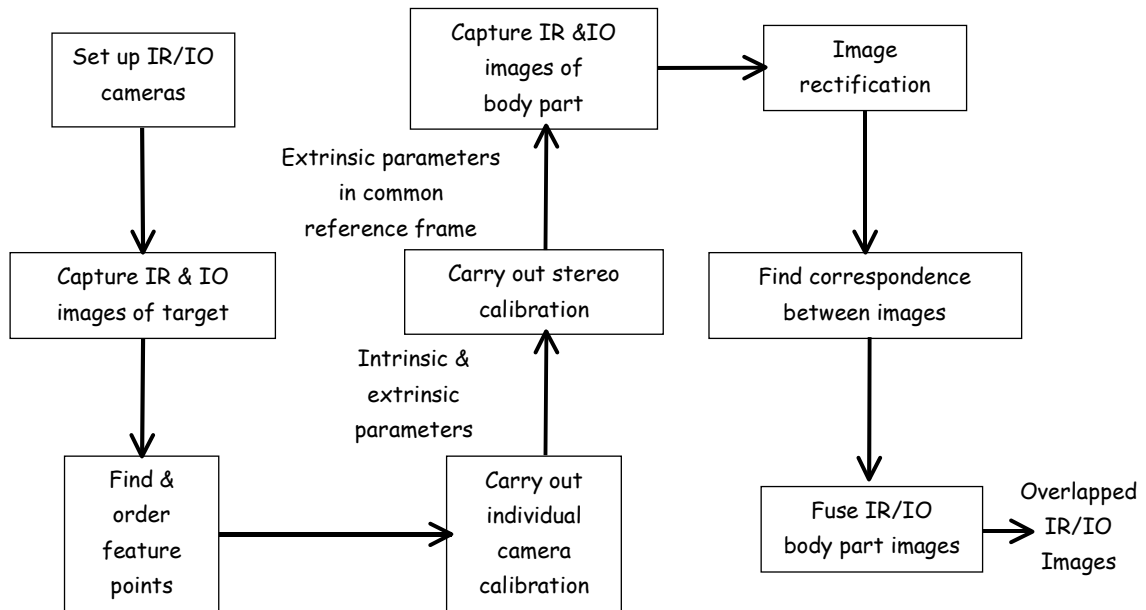
Figure 4-1: Calibration and image fusion process

4.2 Calibration Target Feature Point Detection

The target design makes use of LEDs at the corners of a conventional chessboard calibration pattern. The LEDs allow the corners of the chessboard pattern to be detected by an IR camera so that the target can be used to simultaneously calibrate both IR and

optical cameras. The use of LEDs provides a novel alternative to preheating the chessboard pattern for IR camera calibration. A preheated chessboard pattern is a disadvantage because it cools down during the calibration process and makes the IR image less well defined as time passes. The preheated chessboard is, however, often used to calibrate IR cameras, as reported in the literature [15][29][30].

Two generations of target were designed:

- The first prototype had 26 LEDs on it which corresponded to 24 IR feature points plus two additional location LEDs, see Figure 4-2.

- The second target design had doubled the number of IR feature points to 48 plus two additional location LEDs, making a total of 50.
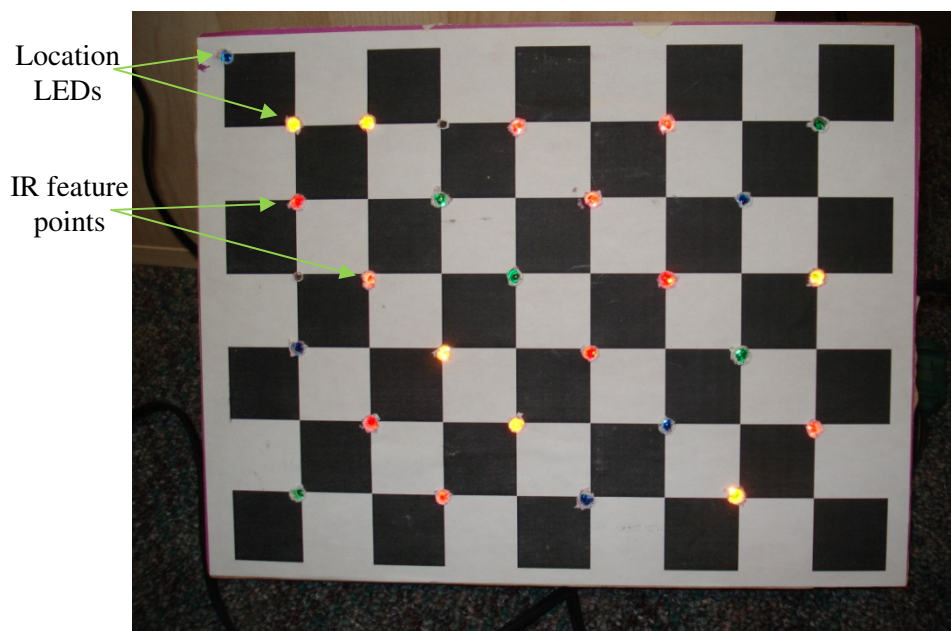
Figure 4-2: The first prototype target with 24 LED feature points plus two location LEDs

Initially, a large number of tests were carried out on the first prototype target. These tests are now described in order to explain to the reader how the second target design was developed.

4.2.1 IR feature point detection and processing

OpenCV provides a function which carries out camera calibration, *cvCalibrateCamera2*. However, this function requires regular feature points from an image to be used effectively, for instance the corners of a chess board target. In order to obtain such feature points with the IR camera, the LEDs must be detected within the IR images. Each LED appears as a bright area in the IR images and the centre of each of these areas corresponds to a chess board corner feature point. This allows these LED feature points to be used in conjunction with *cvCalibrateCamera2*. Figure 4-3 shows a typical IR image of the LED target where the 24 regularly distributed LEDs (in a 4×6 pattern where the rows have been alternately staggered) can be seen together with the two outlying location LEDs, whose use will be explained below.
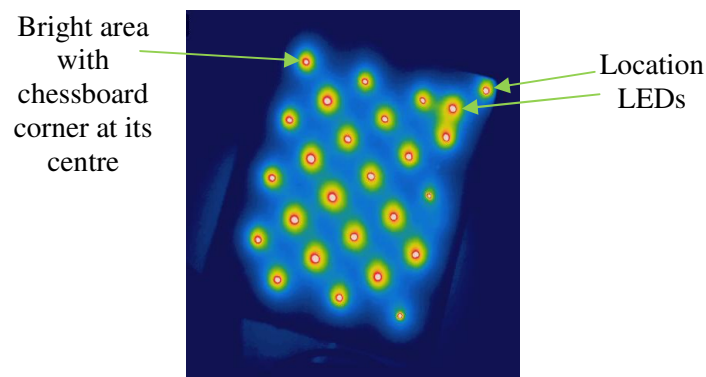
Figure 4-3: Typical IR image of the first calibration target prototype

The first step in processing the IR images is to obtain a binary image with 26 separate LED blobs. This was achieved by applying a temperature threshold to the image. After

many trials, it was found that using a threshold equal to the mean temperature plus 3 times the standard deviation gave the most consistent results of a binary image highlighting the LEDs: an example is given in Figure 4-4(a). However, a problem with this approach is that the LEDs which have been placed closer together on the chess board can produce binary image blobs which overlap, as can be seen in the top right of Figure 4-4(a). Therefore, a process of LED blob separation has to take place so that each binary image results in 26 distinct target LEDs. The result of such LED blob separation can be seen in Figure 4-4(b).

Overlapping blobs with large contour length

(a) Binary image with overlapped blobs   (b) Binary image after blob separation
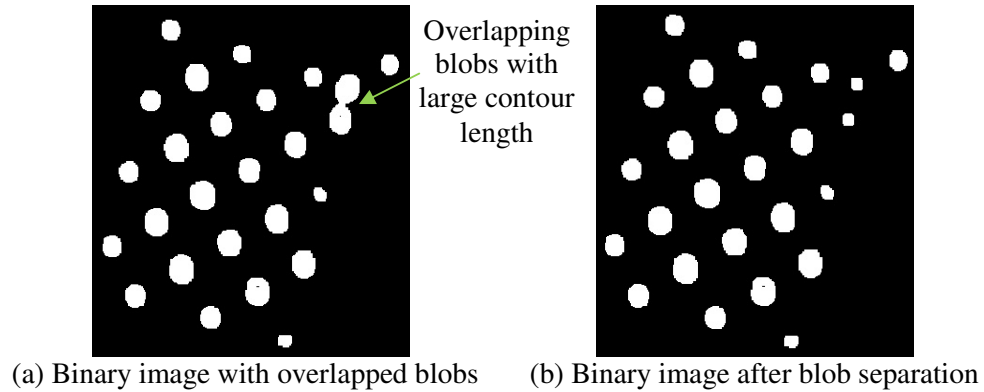
Figure 4-4: IR image of LED target after temperature threshold has been applied

Disconnected blob
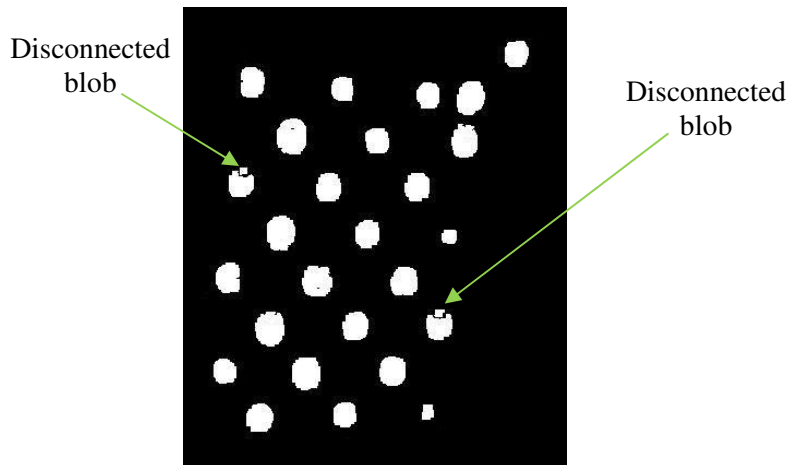
Disconnected blob

Figure 4-5: Binary image before dilation

The process for obtaining the binary images of Figure 4-4 is given in the flow diagram of

Figure 4-6. After the temperature threshold has been applied, there are often small noisy disconnected blobs near a main LED blob, see Figure 4-5. This noise is removed by dilating the image using a 3x3 square structuring element. The resulting binary image is then analysed to see if there are any neighbouring LED shapes that have been joined together, as in the example of Figure 4-4(a). This analysis entails finding the contours of the LED blobs and then checking to see if any of the contour lengths are longer than a threshold which is set to equal the dynamically calculated mean contour length of all blobs plus 3.5 times the standard deviation. If any contour falls into this category, its bounding box is found which is then used to demarcate the border of a region of interest (ROI) in the original IR image, an example is shown in Figure 4-6. This ROI is then subjected to a temperature threshold that is set at a higher value than for the rest of the image. After carrying out several experiments with different IR images of the target, it was found that the mean temperature plus 3.5 times the standard deviation gave the best results. Lower temperature thresholds, did not produce separation and higher thresholds tended to either remove blobs completely or produce very small ones which were difficult to process further. Once the conjoined blobs have been separated, they are dilated using a 3x3 square structuring element. The ROI of the bounding box, which now contains the separated blobs, is then added back to the rest of the binary image in its original position.

Temperature Threshold → Dilation → Find largest contour → Contour > threshold?

Image

Bounding box

YES

NO

Higher temperature threshold for ROI

Dilate ROI

Find centres of blobs using first spatial moments

Order corners
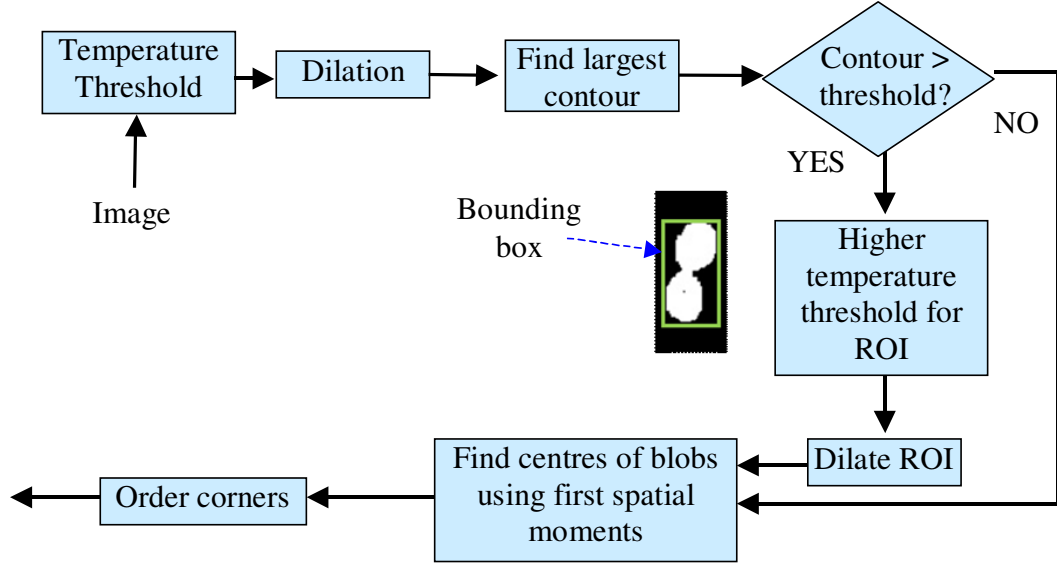
Figure 4-6: LED detection and processing

The centre of each LED blob is then found using the first spatial moments of each contour in the binary image which are obtained using the OpenCV function *cvContourMoments* [11].

Spatial and central moments are properties of an image, see [50] for a good introduction. The spatial moment of an image is given by:

$$M(m,n) = \sum_{j}\sum_{k} x_k^m y_j^n P_{j,k} \qquad (4\text{-}1)$$

where, the first summation is over all image rows; the second summation is over all image columns; $P_{j,k}$ are binary pixel values; x_k and y_j are pixel co-ordinates; m and n are integers that give the moment order.

In this thesis, the spatial moments are used to obtain the 'centre of gravity' of the contours surrounding LED blobs in the binary IR images of the calibration target. In these cases, a centre of gravity coincides with the centre of a contour and, consequently,

the centre of a blob. A contour's centre of gravity (x_0, y_0) is defined by the following two equations:

$$x_0 = \frac{M(1,0)}{M(0,0)} = \frac{\sum_j \sum_k x_k P_{j,k}}{\sum_j \sum_k P_{j,k}}$$ (4-2)

$$y_0 = \frac{M(0,1)}{M(0,0)} = \frac{\sum_j \sum_k y_j P_{j,k}}{\sum_j \sum_k P_{j,k}}$$ (4-3)

where, these equations have been obtained through substitution of the various values for m and n into equation (4-1).

Since, the LED blobs constitute a binary image, each contour is also a binary image so that the double summation of M(0,0) for each separate contour is effectively just the total number of pixels in the contour because each non-contour pixel is equal to zero. In this case, the double summation of M(1,0) is just the sum of all the x coordinates for the pixels in the contour and M(0,1) is the sum of the y coordinates. This is because any co-ordinate not on the contour will be set to zero because it is weighted by the value of the pixels and all non-contour pixels are zero.
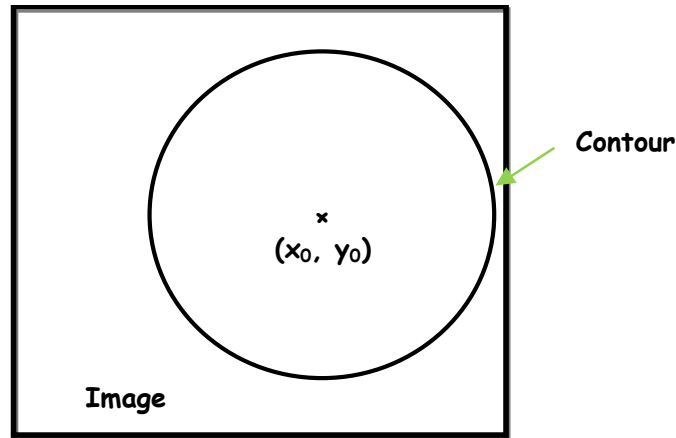
Figure 4-7: Spatial moments give the centre of a circular contour in a binary image

Thus, equation (4-2), in this binary image case, is just an equation to find the mean x_0 co-ordinate value for a contour and, similarly, equation (4-3) gives the mean y_0 co-ordinate value. Thus, if the contour was a perfect circle, then the co-ordinates of the 'centre of gravity', (x_0, y_0), will be exactly in the centre of this circle, see Figure 4-7. In practice, the contours are not perfectly circular, however, the moments will still give the average of the contour co-ordinates so that (x_0, y_0) will be the 'centre of gravity' of the LED blob, which is effectively its centre.

If the feature point detection has worked, the result of this step should be 26 (for the smaller prototype target) or 50 (for the final enlarged target design) separate points on the image which correspond to corners of the chess board square pattern. If any of the points are not detected, that image frame is rejected from the rest of the calibration process. However, at this point, the corners are not ordered and for the function *cvCalibrateCamera2* to work properly, the corners have to correspond to the correct known integer co-ordinates of the chess board's corners. The algorithm used to correctly order the corner feature points is illustrated in Figure 4-8.
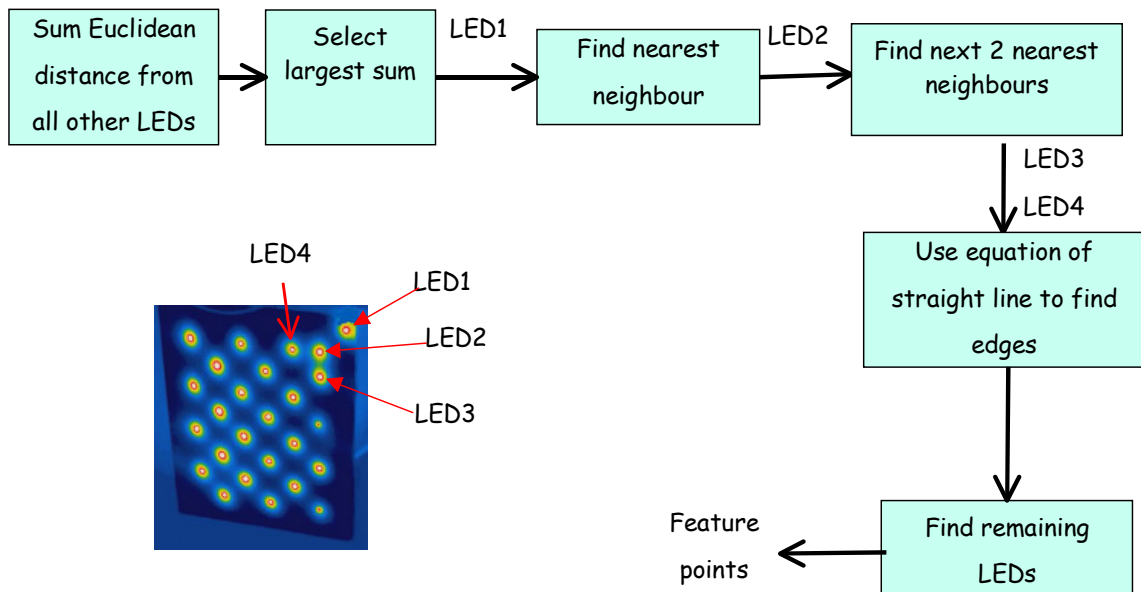
Figure 4-8: Corner ordering

As the inset IR image highlights in Figure 4-8, four LEDs are key to the locating and ordering algorithm. These LEDs are the two location LEDs of Figure 4-2 along with the two LEDs closest to them. The algorithm is as follows:

1. Firstly, LED1 is an outlier compared to all the other LEDs which are located on a square grid. Therefore, the sum of the Euclidean distances between each LED and all the other LEDs is calculated. LED1 produces the largest sum because it is furthest from all other LEDs, because it is an outlier.

2. Then it is a simple matter to find LED2 because this is the nearest neighbour to LED1. This is done by calculating the Euclidean distances between LED1 and all the other LEDs and selecting the LED situated the smallest distance away from it.

3. LED3 and LED4 can then be found in a similar manner by using LED2 as the starting point for the Euclidean distance calculation. This is because they are the

closest neighbours to LED2 as they have only been placed a single chess board square away, as opposed to all other LEDs which are located at the corners of every other square.

4. The equation of the straight line and Pythagoras' theorem is used to find all the other LEDs on the same edge as LEDs 2 and 3 as shown in Figure 4-8 and explained below.
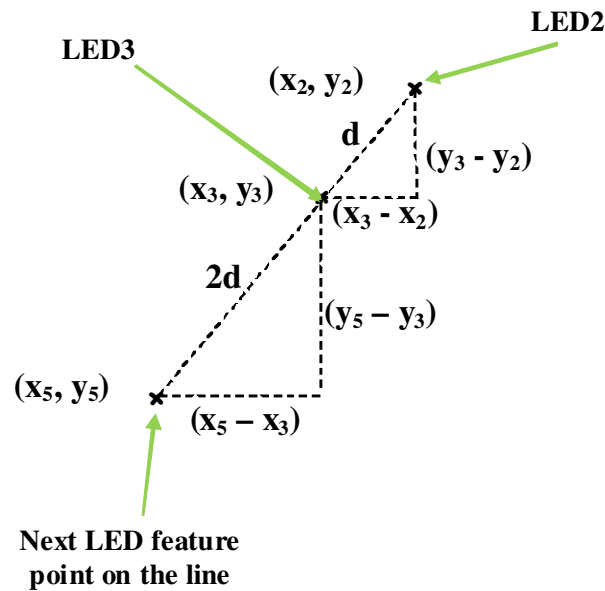
Figure 4-9: Geometry for calculating the ideal position of the next LED on a row

The distance from LED3 to the next LED on this edge should be twice the distance from LEDs 2 to 3 because apart from the four LEDs highlighted in Figure 4-8, all the others have been placed at every second corner of the chessboard. An example of this situation is shown as the distance d and 2d in Figure 4-9. Using Pythagoras' theorem for both the separation of LEDs 2 and 3 and the separation of LEDs 3 and 5 gives:

$$(y_3 - y_2)^2 + (x_3 - x_2)^2 = d^2 \qquad (4\text{-}4)$$

$$(y_5 - y_3)^2 + (x_5 - x_3)^2 = 4d^2 \qquad (4\text{-}5)$$

Dividing equation (4-5) by (4-4) and rearranging gives:

$$(y_5 - y_3)^2 + (x_5 - x_3)^2 = (2(y_3 - y_2))^2 + (2(x_3 - x_2))^2 \qquad (4\text{-}6)$$

Therefore, by inspection:

$$(y_5 - y_3) = 2(y_3 - y_2) \qquad (4\text{-}7)$$

and

$$(x_5 - x_3) = 2(x_3 - x_2) \qquad (4\text{-}8)$$

Therefore, point LED5 is calculated by adding twice the x distance between LEDs 2 and 3 to get its ideal x location and similarly for the y location. Then the nearest LED feature point to this position is found using Euclidean distances, as described earlier. Similarly this newly found LED is used as the starting point to find the next LED on the line. The process stops when the correct number of LEDs for a row have been found: 4 for the first prototype and 6 for the second, not including the corner 'location' LED2.

To allow for the case where the target is tilted away from the plane parallel to the camera image planes where equation 4-8 may no longer hold in all cases, the slope of the line is also used. For example, in Figure 4-9, the points at LED2 and LED3 define a straight line which will pass through all the points on the same edge whether or not the target has been tilted away from the cameras. Therefore, when the nearest LED feature point to the $2(x_3 - x_2)$ position is found, the slope of the line between it and LED3 is calculated. If this slope is not the same as the

56

slope between LED2 and LED3, this feature point is rejected and the next nearest LED feature point is tested. Once a feature point is found that is on a line joining it to LED3 with a slope that matches the slope of the LED2 to LED3 line, the search stops.

5. In the same manner as step 4, the LEDs on the same edge as LEDs 2 and 4 are found using the x and y differences between them and extrapolating along the straight line defined by them.

6. All the remaining LEDs can then be found using the same method by repeating steps 4 and 5 from each of the remaining edge LEDs, i.e. other than LED3 and LED4.

Note that once the order and location of all the LEDs has been found, LEDs 1 and 2 play no further part in the calibration process. This is because the OpenCV chessboard corner finding function does not detect the equivalent outlier LED1 feature point in the optical images. In addition to this, only the non-LED corner points are used for the optical camera calibration, as they are more accurately located in the optical images than the LED corners. Therefore, the optical feature point at the same location as LED2 is not used. For consistency, the IR feature points actually used in the calibration process are also the ones corresponding to the IO feature points on non-LED corners.

An example of the feature points obtained with this technique is given in Figure 4-10(a) where they are ordered from top to bottom and left to right in the diagram. Note that the zig-zag ordered pattern will always move along the LEDs on the long side of the target before moving to the next line irrespective of its orientation.

Feature point 1; co-ordinate (0,1)

Feature point 2; co-ordinate (0,3)

Feature point 24; co-ordinate (5,6)

Interpolated point

(a) Ordered feature points   (b) Interpolated feature points
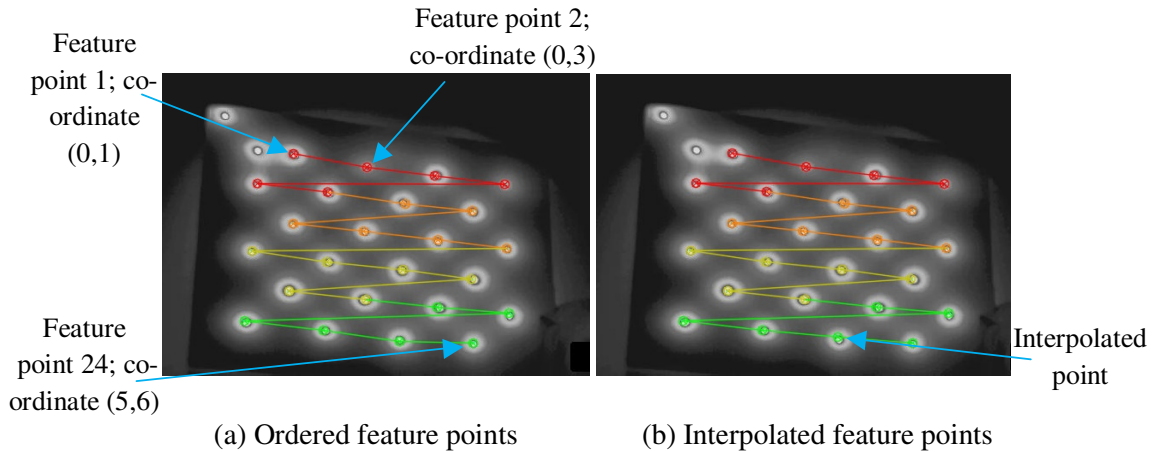
Figure 4-10: IR image ordered feature points

At this point, 24 feature points (all the points highlighted in Figure 4-10(a)) can be used with the function *cvCalibrateCamera2* to obtain the intrinsic and extrinsic parameters for the IR camera. The corresponding integer object points needed for this function are set to alternating odd and even positions where the first row and column is numbered zero. Therefore, instead of numbering the feature points from 1 to 24 consecutively, the co-ordinates of the points become: (0,1) (0,3) .. (1,0) (1,2)......(4,1) (4,3)..(5,4) (5,6) see Figure 4-10. This is because the OpenCV function assumes there are feature points at every chessboard corner on the underlying conventional chessboard pattern whereas with the IR feature points, there are chessboard pattern corner points missing from every other chessboard square, e.g. corner point (0,2) falls midway between the first two feature points (0,1) and (0,3) shown in Figure 4-10. This co-ordinate numbering scheme is used to label the IR points so that they match the same co-ordinates for the IO equivalent feature points. For instance, feature point (0,1) in an IR image corresponds to the same target location as feature point (0,1) in an IO image. The difference is that the IO image provides feature points at all the corner point locations, e.g. there will be an IO feature

point at co-ordinate (0,2), whereas there are only IR feature points at every other corner point, as shown in Figure 4-10.

It is evident from Figure 4-10 that one of the feature points in the bottom row is out of alignment. This has occurred because the hole for the LED was drilled slightly below the corner position on the chess board. In order to test how much error this misalignment made, an experiment was undertaken where calibration was carried out after replacing this feature point with a linear interpolated value calculated from its two nearest neighbours on the same row, see Figure 4-10(b). The interpolation was carried out by programming the software to always realign this misaligned point using linear interpolation between the two points on either side of it. The results from this experiment are discussed in Section 4.3.2.

4.2.2 Optical image corner detection and processing

For convenience, a typical optical image of the prototype target is presented again in Figure 4-11
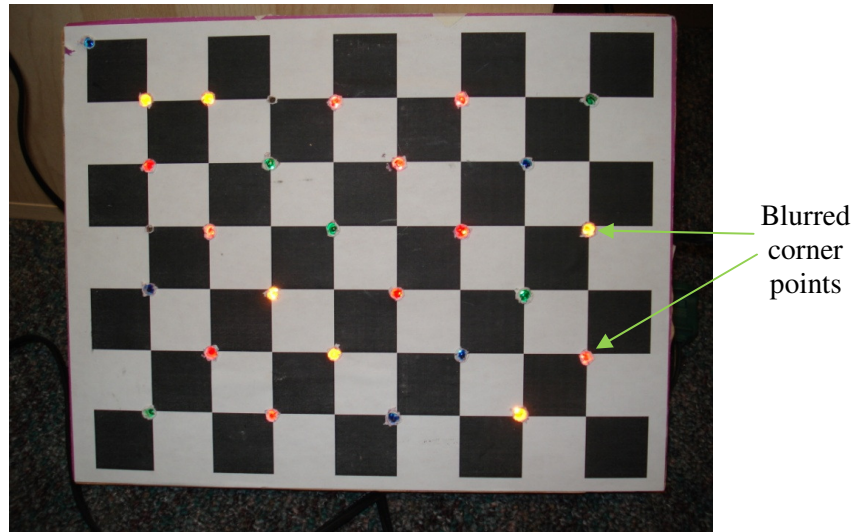
Figure 4-11: The first prototype target showing blurred corner points coinciding with the LEDs

As can be seen, there is some blurring of the black and white corner points on the optical target image which is caused by the LEDs. OpenCV has a function, called *cvFindCorners,* which finds corner feature points in an image and this is used to obtain all 48 internal corner points in the optical target images. Note that there are twice as many corner feature points in the optical images, that use the black and white chessboard pattern, than available in the IR images because the LEDs are only placed at every other corner.

Once the corners are found, they are passed to *cvCalibrateCamera2* to obtain the intrinsic parameters of the IO camera.

4.3   The Prototype Target Tests

This section starts by describing the specific tests carried out with the prototype target before continuing to discuss the results obtained from the tests.

4.3.1 The experiments

In order to assess the suitability of the prototype LED target for calibrating the IR and optical cameras, four different experimental equipment setups were used to obtain intrinsic parameters:

1. The IR camera was used to capture images of the prototype target. In this case, the intrinsic parameters were calculated using: a) no correction of the misaligned point; b) correction of the misaligned point using interpolation.

2. The optical camera was used to capture images of the prototype LED target. In this case, the intrinsic parameters were calculated using all target chessboard corners both with and without LEDs on them;

3. The optical camera was used to capture images of a conventional optical target i.e. a chessboard pattern without any LEDs on it. This test was undertaken to try to measure the error associated with the blurred corners due to the LEDs.

4. The optical camera was again used to capture images of the prototype LED target but this time only target chessboard corners without LEDs on them were used in the calibration calculations. This was partly to test the optical calibration but also to verify that the LEDs were correctly aligned with the square corners for the IR calibration.

This last test entails a halving of the available feature points from 48 to 24, as every other corner has an LED on it. An example showing the feature points obtained with each of the IO camera experiments can be seen in Figure 4-12. The feature points obtained with the combined IR/optical target are shown in Figure 4-12 (a). The feature points obtained

61

using a conventional optical target, with no LEDs, can be seen in Figure 4-12 (b). For examples of the even and odd feature points, where only the corners without LEDs are used, see Figure 4-12 (c) and (d).
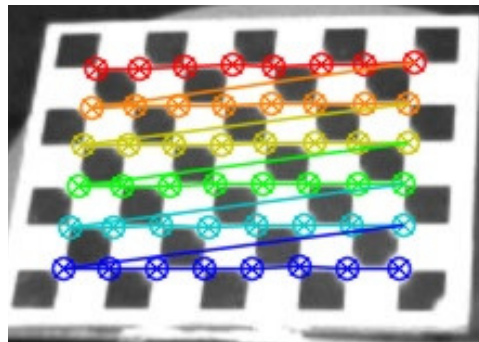
(a) Full combined target	**(b) Conventional target**
(c) No LED even points	**(d) No LED odd points**
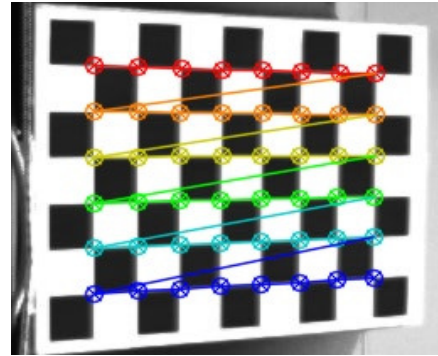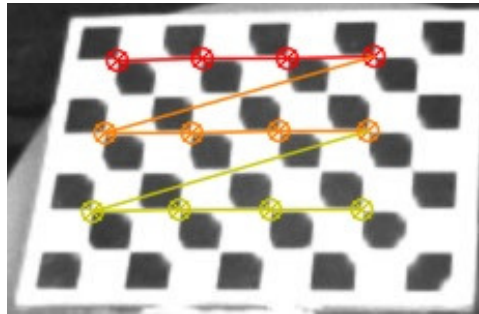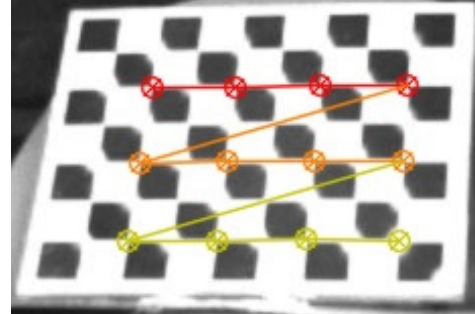
Figure 4-12: Optical target with feature points for various experiments

The reason for splitting the even and odd feature points in this way was so that the OpenCV function would not receive alternating lines of staggered points for processing. However, it was subsequently found out that the function can process alternating lines of staggered points, as shown for the IR feature points in Figure 4-10, as long as the co-ordinates are numbered as described previously in Section 4.2.1. Therefore, all subsequent experiments did not split even and odd feature points in this way, see Chapter 5 for a description of these experiments.

It is a relatively easy matter to determine the non-LED corners in an image taken with the optical camera. Once the first ordered corner LED has been found using the techniques

described in Section 4.2.1 for an IR image, the equivalent corner can be found in the list of optical image chessboard corners. Then every other corner is taken as a feature point.

Once the IR and optical feature points have been obtained, the OpenCV function *cvCalibrateCamera2* was used on each set of points separately to obtain the intrinsic parameters for both types of cameras.

The experiments with the prototype target were carried out prior to the cameras being synchronised. Therefore, a series of still images of the target was obtained where the target was moved between each image capture.

Each set of images for each experimental setup was obtained by placing the target on a swivel chair on a box. For successful calibration, it is important to capture images that do not all have co-planar feature points. The swivel chair allowed the target to be easily moved into different planes. The box allowed the target to be raised to about the same height as the cameras for ease of image capture. It also blocked out any spurious hot areas behind the target which would have otherwise appeared in the IR images, e.g. the electric socket.

For each experiment, the series of images captured includes the following positions:

1. The target is placed on its long side see Figure 4-13.
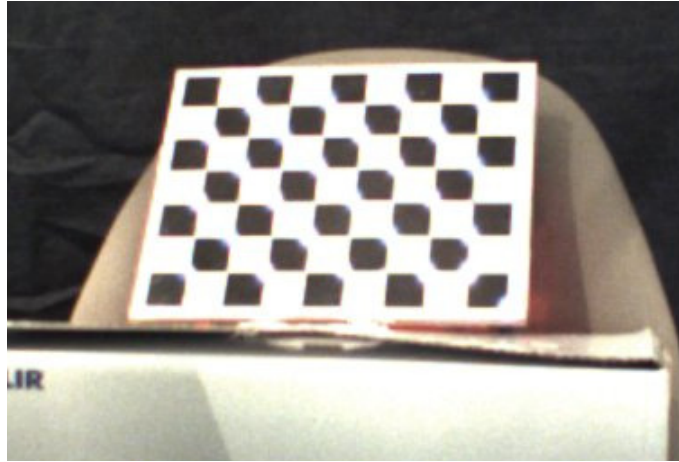
Figure 4-13: Long side target position

2. The target is then raised at one side so that it is oriented at approximately 20° to the horizontal along the box, see Figure 4-14.
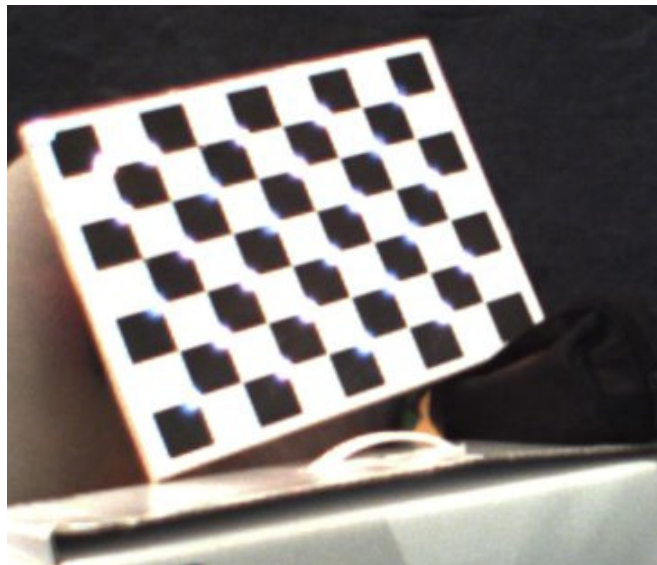
Figure 4-14: Raised target position

3. The target is then raised at its opposite side, see Figure 4-15.
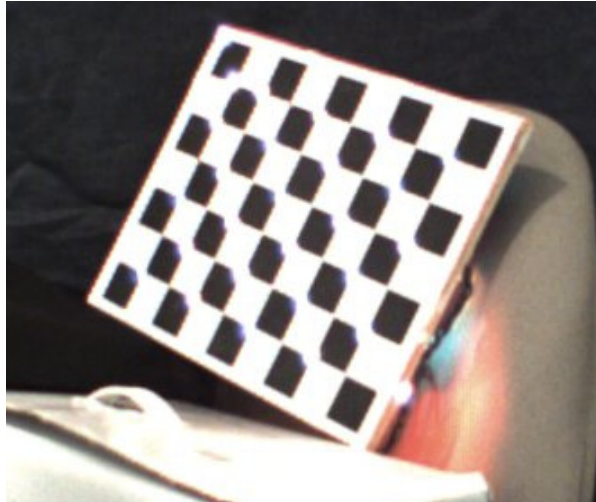
Figure 4-15: Angled target position

4. The target is then placed on its short side, see Figure 4-16.
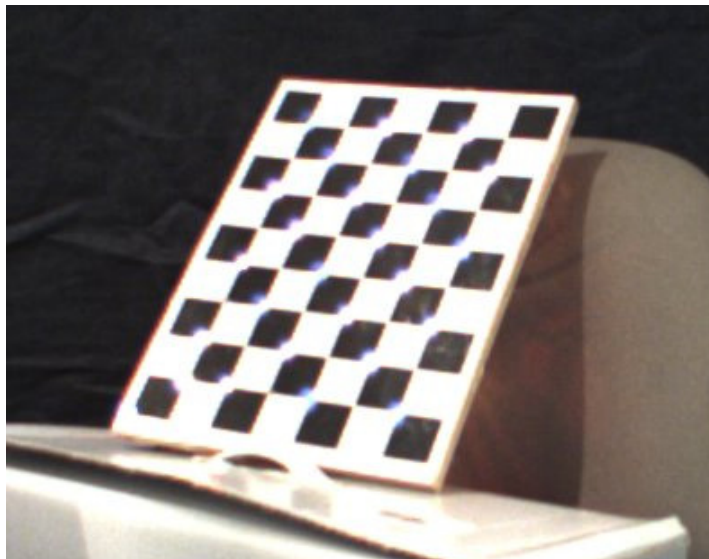
Figure 4-16: Short side target position

For each target orientation, the swivel chair is then used to move the target through a range of angles to the optical axis with a limit of approximately 45° to each side. This movement ensures that the feature points are not all co-planar. The chair is then moved

forwards ~40cm towards the cameras from its original position of 120cm from the cameras. This movement increases the working volume that is calibrated by the process.

The next section gives a detailed analysis of the calibration results from these various experiments with the prototype target.

4.3.2 Results from prototype experiments

For the purposes of testing the combined calibration of IR and optical cameras with the prototype target, the focus has been on obtaining and analysing the intrinsic parameters. An analysis of the extrinsic parameters obtained with the final target design is given in 5.4.2. The results obtained with the prototype target for the intrinsic parameters are summarised in Table 4-1 where they have been split into 2 sections:

- The first three rows are for the optical camera results using a conventional 'no LED' target compared to results obtained with a target which includes LEDs. With the LED target, rows 2 and 3, two sets of results are given, firstly where the feature points include corners with and without LEDs and secondly where the corners are the ones which do not have LEDs on them.

- Finally, the last two rows present the results from using the IR camera with and without any position correction using interpolation.

The fourth column, entitled '#images', lists the number of images collected, that successfully produced the expected number of feature points per image, which is listed in the third column. Only these images' feature points were used to calibrate the camera. The radial distortion coefficients are in column 5 and the focal length results are in the columns for f_x and f_y.

Camera	Target	#points	#images	$|k_1|$	f_x Image width (mm)	f_y	Max % Diff. f	Ground truth f (mm) (error)
IO	conventional	48	12	0.37	1.31 (6.2)	1.31 (6.2)		6.0 (3.3%)
IO	prototype with LEDs	48	12	0.36	1.31 (6.2)	1.32 (6.3)	0.8	6.0 (5.0%)
IO	prototype with LEDs: using only feature points with no LEDs	12+12	12	0.30	1.29 (6.1)	1.28 (6.1)	2.3	6.0 (1.6%)
IR	prototype with LEDs	24	12	0.04	2.23 (17.8)	2.25 (18.0)		18.0 (1.1%)
IR	prototype with LEDs: interpolated point	24	12	0.26	2.29 (18.3)	2.32 (18.6)	3.1	18.0 (3.3%)

Table 4-1: The intrinsic parameters obtained from the prototype experiments

The values given for the intrinsic parameters are quoted in units of image width and physical distance in millimetres. The image widths are used to examine the relative differences between the experimental results. The physical distance values are used to compare the results to the ground truth which are the focal lengths provided by the camera manufacturers. In the IR camera case, a nominal focal length of 18mm is quoted and the detector pitch (that is distance between pixels) is given as 0.025mm. This means that the sensor or internal image width in the camera is 0.025 * 320 = 8mm because the IR camera image has a width of 320 pixels. In the case of the IO camera, the manufacturers quote a focal length of 6mm and a distance between pixels of 0.0074mm. This means that the sensor width in the camera is 0.0074*640 = 4.7mm because the IO camera image has a width of 640 pixels. In order to convert the experimentally obtained image widths to mm, it is just a matter of multiplying them by 8, for the IR results, or 4.7,

for the IO results. However, it should be noted that the camera manufacturers do not provide any indication of the errors associated with these various parameters.

The column entitled 'Max % Diff. f' refers to the largest difference between the focal lengths values obtained between the different experiments. In the case of the optical camera, this is referring to the differences obtained using the conventional target without LEDs, row 1 in Table 4-1, and the other verification methods that were discussed previously, rows 2 and 3. In the case of the IR camera, this is referring to the biggest difference between rows 4 and 5 in Table 4-1.

The final column in Table 4-1 gives the ground truth focal lengths for both IR and IO cameras. The difference from the ground truth obtained with each experiment is also given here as a percentage error in brackets.

The recommended number of images to use is at least 8 to 10 with a chessboard of size 7x8 [11]. In all the experiments reported on here, a 7x9 chessboard was used for both the conventional target and that with LEDs. For these prototype experiments, the number of images was fixed at 12. Note that this is the number of images that successfully return the expected number of feature points for each image.

The intrinsic parameters for the optical camera were measured using the conventional chessboard target without LEDs. The focal length was found to be 1.31 image widths, which is equivalent to 6.2mm, in both x and y directions, see the first row of results. Together with the ground truth focal lengths, these results are used as a benchmark against which to compare the results using the prototype target with LEDs.

As can be seen in Table 4-1, second row, the results from using the prototype LED target gives the same f_x = 1.31 (6.2mm) and f_y = 1.32 (6.3mm) image widths. This is a difference of 0.8% from the results using the conventional target and only 5% larger than the ground truth. This shows that it is possible to use the prototype LED target to calibrate an optical camera. A significant part of the deviation can be explained by imprecise machining operations used to insert the LEDs at the exact corners of the chessboard pattern.

Examination of the results on the following row, using only the feature points which fall on corners without LEDs, show that the results differed from the benchmark by a maximum of 2.3%. However, the result of 6.1mm focal length is closer to the ground truth than the other results. The advantage of not using the corners with LEDs on them, is that the corner points can be found more accurately because they are not blurred by the LED light. Conversely, the disadvantage with this technique is that there are half the number of feature points available (24 as opposed to 48) when the LED corners are not used. This latter problem is also associated with the IR camera calibration results, because again only 24 feature points are available in the images captured with the IR camera compared with the 48 found for each image from the optical camera.

The IR camera calibration results are shown in the last two rows of Table 4-1. The first of these gives f_x = 2.23 (17.8mm) and f_y = 2.25 (18.0mm) image widths which were obtained using all 24 LED feature points in one call to *cvCalibrateCamera2*. These results are either very close (~1.1% difference) or equal to the ground truth value of 18mm.

69

The effect of replacing the LED which is out of alignment in the final row, by using linear interpolation, gives results which differ by as much as 3.1% and f_x = 2.29 (18.3mm) and f_y = 2.32 (18.6mm) image widths. There is a larger difference from the ground truth too: ~3.3% difference. However, in addition to the misalignment, this relatively large change could be due to the fact that the LEDs provide an insufficient number of feature points to give stable results. After all, [11] does recommend, that as many as 42 feature points should be used for each image.

4.4 Final target design and processing

This section discusses the design of the final IR/optical camera calibration target used for this research project.

The results from the experiments with the first prototype target, presented in Section 4.3.2, show that the optical camera calibration accuracy is only minimally affected by the presence of LEDs in the target (~0.8% difference to results obtained from a conventional target and the prototype LED target), when all 48 black and white target corners are used as feature points. However, accuracy has been shown to be very much dependent upon how many feature points are used in the calibration process. Ten images, each with 42 feature points, are recommended in [11] and ten images with 64 points were used in [9]. With the first prototype experiments, twelve images were used for both cameras with 48 points during optical camera calibration but only 24 points during IR camera calibration. It has also been shown that even one slightly misaligned LED can cause large differences (~3.3%) in measured intrinsic parameters. This suggests that there are insufficient IR feature points in the first prototype calibration target to obtain levels of accuracy appropriate for a medical assessment application.

As a direct consequence of these experiments with the first prototype calibration target, and in order to improve the calibration accuracy, a new IR/IO calibration target with 48 LEDs was built and used for all subsequent experiments described here. This target required an increase in the number of chessboard squares from 7×9 to 9×13 in order to accommodate the extra LEDs. The physical size of the target also increased from 21.9×28.2cm to 28×40cm, although the individual squares remained 2.9×2.9cm in size. It was also decided only to use the corners of the chessboard that do not have LEDs on them for the optical camera calibration, see Section 4.2.2 for an explanation of how this is accomplished. An example of the non-LED corners obtained using the improved target with the IO camera is shown in Figure 4-17.
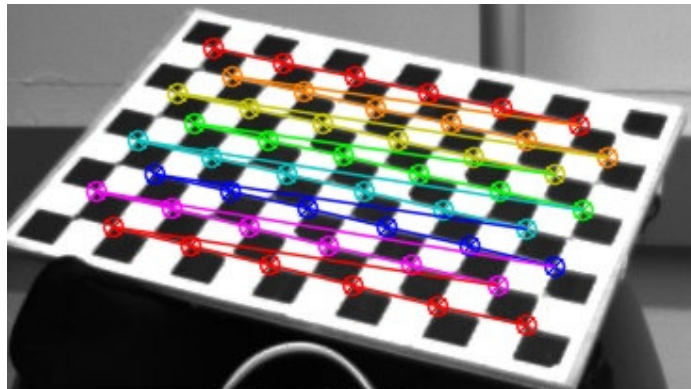
Figure 4-17: Larger IR/IO calibration target with IO feature points on non-LED corners

However, the problem associated with any misalignment of the LEDs still needed addressing for images captured with the IR camera. The approach used to solve this problem is the subject of the next section.

4.4.1 IR feature point processing
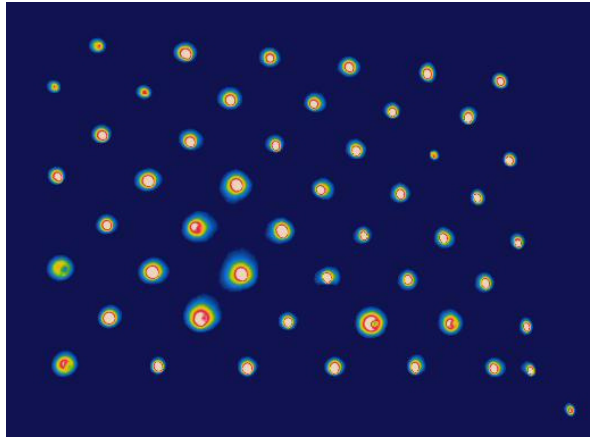
Figure 4-18: Example of variable size LED IR features

Figure 4-18 shows a typical example of an IR image of the target showing the locations of the LEDs. As can be seen, the size of the 'blobs' associated with the LEDs is variable and some of them are significantly larger than others. This variability is mainly due to slight differences in the size of the holes in the target which alters how far through a hole an LED protrudes and also how tightly fitting an LED is inside a hole. Both these differences affect the heat dissipation characteristics surrounding the LEDs, which leads to different temperature distributions around each LED.

The process used to determine a feature point location is to convert the colour image into a binary image before calculating the centre of each 'blob' using the central moments associated with its contour. However, the centre of a 'blob' may not coincide exactly with the corner of the chessboard square because of construction imprecision.

Thus, it was decided that a more robust approach needed to be taken than just correcting a particular LED with simple linear interpolation, as described in Section 4.2.1.

The LED feature points should lie on a rectangular grid pattern. Thus, it is known in advance where the ideal feature points should be once the first three feature point positions have been located. The target's regularity has been exploited by using lines of best fit to adjust the captured feature points so that they fall on the rectangular grid pattern. The adjusted feature points used for the rest of the calibration process lie on the points of intersection of the grid lines which do not have LEDs on them. These alternative non-LED points have been selected as the feature points because they coincide with the optical image feature point positions which coincide with corners of the black and white chess board pattern that do not have LEDs on them. This adjustment process involves calculating the lines of best fit for each horizontal and vertical row of IR image LED feature points. The lines of best fit are found using the OpenCV function *cvFitLine* which minimises the L2 distance metric across all points.
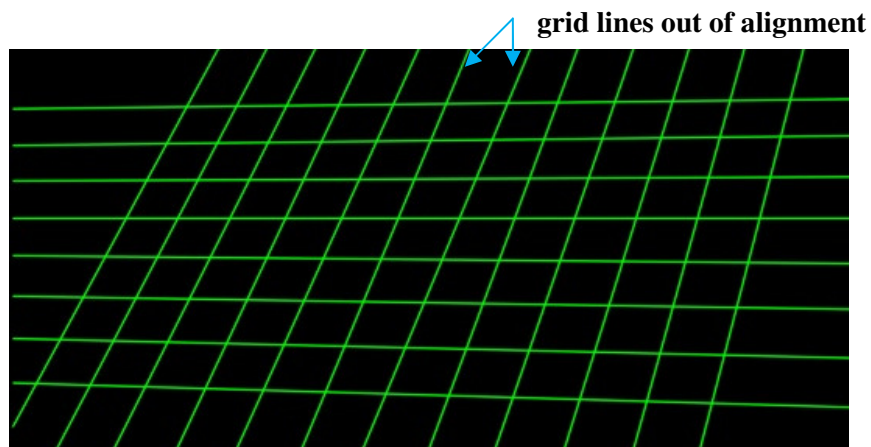
Figure 4-19: Original lines of best fit

An example of the lines of best fit, prior to adjustment, can be seen in Figure 4-19. To correct for this misalignment, the feature points which fall on lines at the edges of the grid are analysed in turn. For example, the distances between each feature point on the top horizontal line are compared and if the distance between two points falls outside the mean by more than 1.5 standard deviations, it is flagged. If this distance is less than the

73

mean, the next gap in the line is analysed to see if it is larger than 1.5 times the standard deviation. If this is the case, the feature point that is common to both these adjacent gaps is moved to fall exactly half way between the other two points, as shown in Figure 4-20.
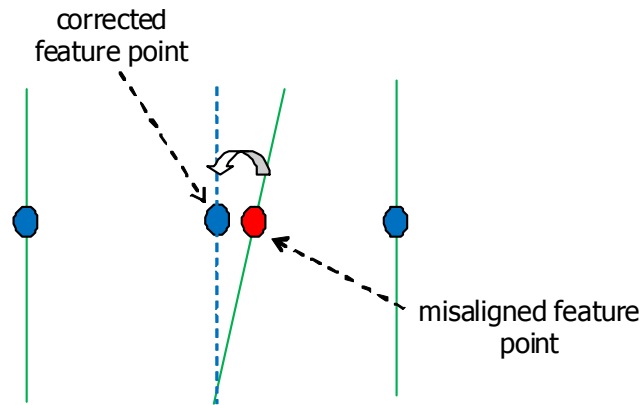
Figure 4-20: Correction of a misaligned feature point

An example of the corrected grid lines, after this readjustment process has been carried out, can be seen in Figure 4-21, where the adjusted lines are highlighted in blue.
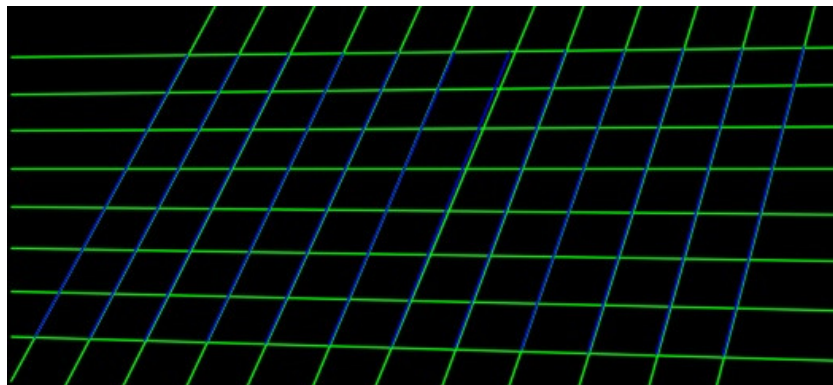
Figure 4-21: Adjusted lines of best fit

Note that the horizontal grid lines did not need adjustment as the LEDs were accurately aligned in this direction. Therefore, this correction process was only carried out on the vertical grid lines in this example.

The original unadjusted corners associated with the LED feature points for the above example, can be seen in Figure 4-22. Figure 4-23 shows the corrected feature points on the non-LED chess board corners which were obtained by finding the intersection of the readjusted lines of best fit, as described earlier.
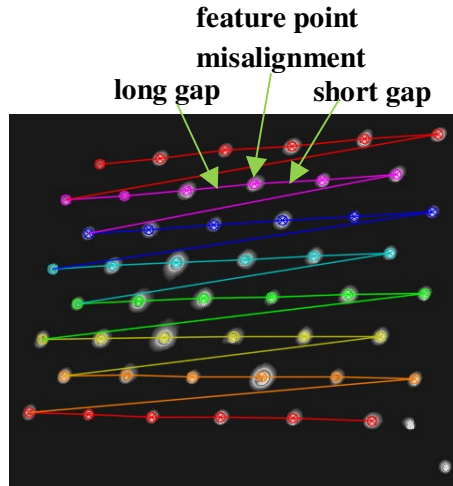
Figure 4-22: Original IR feature points on LED corners
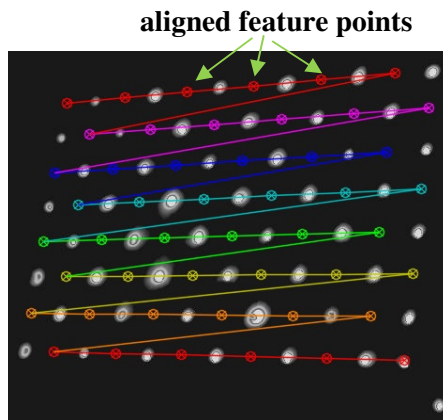
Figure 4-23: Adjusted IR feature points on non-LED corners

4.4.2 IO Feature Point Correction

For the IO feature point detection, the OpenCV function *cvFindChessboardCorners* was used before extracting the even 'non-LED' points for the stereo calibration. Occasionally,

however, the OpenCV chessboard corner detection algorithm failed to find all the corners in the IO image, an example of this failure can be seen in Figure 4-24.
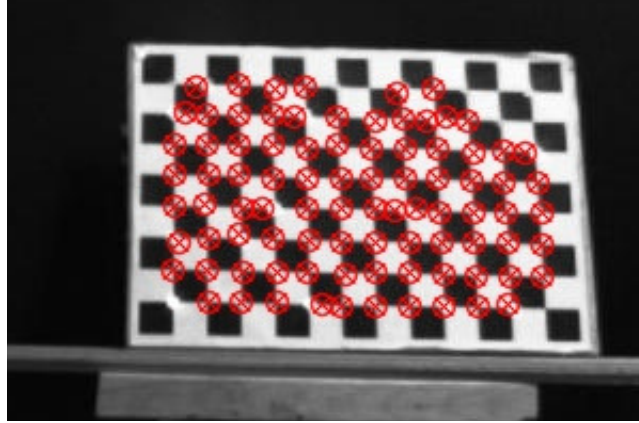
Figure 4-24: IO feature point detector has missed a few of the chessboard corners
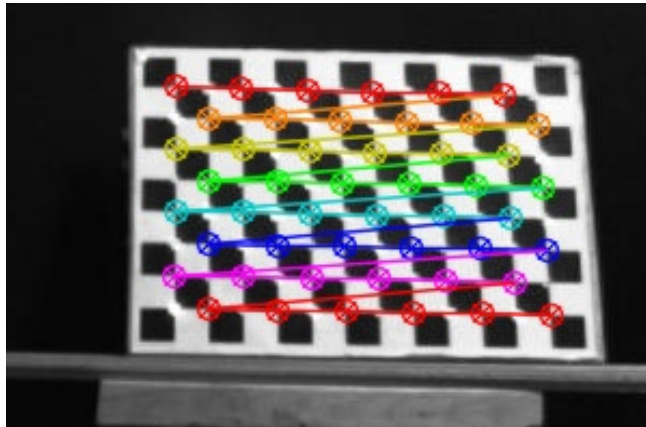
**Figure 4-25: Non-LED feature points
retrieved from the feature points depicted in Figure 4-24**

If the feature points of the target in the image are approximately aligned horizontally, the missing corners can be retrieved by ordering the existing feature points by using their coordinate values. If not aligned horizontally, the set of incomplete feature points will be rejected from the calibration process. The process of missing corner retrieval is as follows:

1. The feature points are grouped according to their *y* values (which is why this correction only works with a horizontally aligned target).

2. Then the top left feature point is used as a starting point and the row for each group is ordered by finding the next largest *x* value in its group. These rows are used to find the horizontal grid lines.

3. Once the rows have been ordered separately, as long as at least three rows are complete (that is 12 black and white chessboard corner feature points are on each row for the final target design), these rows can be used to find the vertical grid lines by using the feature points on the complete horizontal rows, together with Pythagoras' theorem, as described in Section 4.2.1.

4. After all the grid lines are in position, like those shown for the IR image in Figure 4-21, a process of realignment of the lines is carried out as described in Section 4.4.1.

5. Finally, the missing IO feature points can be filled in by using the points of intersection of the grid lines and the required non-LED feature points can be obtained, see Figure 4-25.

Once the feature points have been obtained for a set of corresponding pairs of IR and IO images, the OpenCV function *cvStereoCalibrate* is used to carry out stereo calibration between the two cameras and *cvStereoRectify* is used to carry out the rectification using Bouguet's algorithm, as described in Section 2.1.

4.5 Object of Interest Image Capture

Once the IR and IO cameras had been stereo calibrated, a series of synchronised pairs of images of the object of interest could be captured. As will be explained further in Chapter 5, there were some experiments carried out when the IO and IR cameras were not synchronised. In this case, the object that was used for image capture was a small lamp which could be statically imaged by both IR and IO cameras, see Figure 4-26. The lamp was needed because a human limb could not be held sufficiently still to allow both cameras to capture an image of the limb in the same position.

Figure 4-26: Lamp used for obtaining static IR and IO images

However, for the purposes of medical assessment of RA, the object of interest is invariably a body part. For this thesis, the body part used to test the system was the author's hand. For each experiment, the hand was moved through a sequence of positions and the synchronised IR and IO image frames were captured before being fused together, as shown below, in Figure 4-27.

Figure 4-27: The sequence of synchronised and fused IR/IO images

This particular image sequence was obtained when the cameras' optical axes were in parallel with a separation baseline of 10cm between the two cameras.

4.6   IR/IO Image Synchronisation

Although the frame capture from the IR and IO cameras is synchronised by a dedicated piece of hardware, developed as part of this research and detailed in Section 3.3.3, and the IR images are time stamped, the Point Grey multiple image capture for the IO camera does not currently provide a time stamp with the file names. This causes problems with establishing the corresponding IR and IO images after frame capture has been completed. However, Point Grey is planning to release a version of its firmware that will provide timestamps on continuously captured images, so this problem will be removed in the future. In the mean time, the process of finding corresponding IR/IO images relies on the correct time alignment of the first frame for each camera.

Currently, the PC time stamp for the initial creation of the image files is used as a guide to find the first corresponding images. Once the first pair of corresponding frames from the IO and IR cameras are identified, the subsequent frames are easily matched by using the time stamp on the IR images. Unlike the IO images, sometimes an IR image is not captured every second, due to network bandwidth restrictions. In these cases, the IR time stamp is used to calculate how many dropped frames have occurred and that number of IO frames is skipped to make sure the IR/IO frames are aligned once again.

If the wrong initial images are chosen as corresponding ones, all subsequent images will be misaligned. In these cases, the overlays are incorrect too as can be seen in Figure 4-28 where the IR image is of a closed hand whereas the IO image is that of a splayed one.

Figure 4-28: Incorrect selection of initial IR/IO image frames for overlay

During the work reported on in this thesis, the PC time stamp only failed to align the first IO and IR frames in a sequence for about 7% of the experiments. However, the problem does need to be addressed to produce a fully automated system suitable for use by medical practitioners. Currently, the only remedy for this, until Point Grey releases its new time stamping software, is to run the images through the process again after a reselection of different initial corresponding images.

See the next Section 4.7 for an explanation as to how IR/IO image fusion is carried out.

4.7   IR/IO Image Fusion

Objective 5 for this thesis is to overlay the IR and IO images to obtain a fused multi-spectral image. The reason for this objective is because such fused images could be used by doctors in a clinical setting to assess muskuloskeletal diseases, such as RA. The pain associated with RA can be 'seen' by the areas of elevated temperatures in an IR image of the patient's joint [2]. The IO component of the fused IR/IO images provides a higher definition of the shape of the limb which allows a medical practitioner to associate the higher temperature areas in the IR image with the physical area on a patient's joint. For

instance, by collecting a sequence of fused IR/IO images at each doctor's visit, effectiveness of the RA treatment can be quantitatively assessed by measuring e.g. the temperature differential around the joint. If the treatment is successful, each subsequent doctor's visit should show a reduction in the temperature differential.

It should be noted that the value of the maximum temperature for a patient's joint will probably vary between doctor's visits, even after a period of acclimatisation has taken place, depending on how cold the patient is feeling on a particular day. However, this change of maximum temperature does not prevent the fused images from being compared between visits because it is the relative temperature distribution across the diseased joint that is of interest here rather than actual values. In other words, even if the maximum temperature observed on a fused image is lower at one visit compared to another, the size of the temperature differential will still allow an assessment of the extent of RA being experienced by the patient and can be quantitatively compared to the temperature differential on a fused image that was captured during a different doctor's visit.

Once the feature points for both IR and IO images have been obtained for at least 10 synchronised pairs of images, the intrinsic and extrinsic parameters are found out using the calibration techniques that have already been described in Chapter 2. Then stereo rectification is carried out using the OpenCV *cvStereoRectify* function to obtain a one to one correspondence between the IR and IO image feature points. Once rectification has taken place, it is possible to fuse the IR and IO images to give a composite image of a body part, as per the objective of this thesis, an example of which is shown in Figure 4-29. Note that the stereo calibration and rectification process, described in Section 2.1, in combination with the correspondence algorithm, described below, allows alignment

and correspondence of the two multi-spectral images even without depth information. This is because the stereo calibration ensures that the two images have scaled the body part to the same size within each image and the rectification algorithm ensures a row-by-row alignment between the two images.

The overlay image in Figure 4-29 highlights the areas of the hand which have the highest temperature. As discussed in Section 1.1, the motivation for highlighting higher temperature areas is that, if someone has RA, they are an indication of inflammation [2] and such images could be used to assess RA. This assessment could be done either by looking at the changes in temperature distribution between doctor's appointments or by comparing them to a database of images of the same joint from people both with and without the disease.

Figure 4-29: Fused IR/IO image of a hand

Note that any pixel in the IR image red channel not exceeding the value of 96, which is equivalent to a temperature threshold of ~26°, is not included in the fused image because, for assessment of RA, it is probably the regions of elevated temperature that are of most

interest. This temperature threshold accounts for the missing IR content in the fused image. The method used to obtain such image fusion is the subject of the rest of this section.

With optical stereo images, the OpenCV function *cvFindStereoCorrespondenceBM* can be used to obtain the disparity map between the two images, which gives the horizontal shift between them. However, this function relies on finding the correlation between corresponding pixels in the left and right images on a row-by-row basis. Thus, it cannot be used here because there is no simple correlation between IR and IO images.

If the object position along the z-axis was known accurately, the disparity map could be found by scaling the translation baseline which is given by the extrinsic parameters, as in equation (4-9).

$$(x_l - x_r) = \frac{Z}{fT} \tag{4-9}$$

where Z is the depth, or z-axis distance between the object being imaged and the cameras, f is the focal length, T is the baseline distance between the two cameras along the x-axis, x_l and x_r are the x-axis pixel locations in the left and right images respectively, see Figure 2-6. However, this equation cannot be used here to accurately calculate a disparity map between the two images because the distance of the object from the cameras is only known approximately.

The stereo calibration process together with the application of the epipolar constraint during rectification means that the object depth information is not necessary for achieving accurate rectification between the multi-spectral images. This process ensures that the images from both cameras are the same size.

As the OpenCV function cannot be used to find the correspondence, a different technique was used to carry out the data fusion between the IR and IO images, as will now be described. It should be noted that the method of image fusion described here was necessary to combine images from different parts of the spectrum (i.e. IO and IR) because the more conventional methods for image fusion do not work in this case, see [34] and [35] for overviews of image fusion techniques. For instance, averaging or Bayesian probabilistic methods do not work here because the pixel levels in the optical and IR images are independent of one another. This means that the average value of the optical image pixels is completely different and unrelated to the average value of the IR image pixels. Therefore, any statistically based technique is inappropriate for image fusion in this case.

Firstly, the correspondence between the IR and IO images is calculated using Canny edge detection on the IO image together with a binary fill operation on the IR image. The binary fill operation is carried out using a simple temperature threshold level to segment the image after conversion of the IR image to greyscale where the threshold was set to 64. This threshold was found to work well with the images captured during the work for this project.

A typical example of the result from each operation is shown in Figure 4-30 and Figure 4-31. The left most edge of the IR silhouette in Figure 4-31 is detected then the corresponding left most edge is found on the same image row in Figure 4-30.

The difference between the x co-ordinate values for the same row is calculated as the required lateral shift in each rectified image to obtain an overlaid IR/optical image. The difference is averaged over several rows of the images to obtain a more accurate shift

value. Note that the epiline alignment process during image rectification means that there should never be a need for vertical realignment, see Section 2.1.
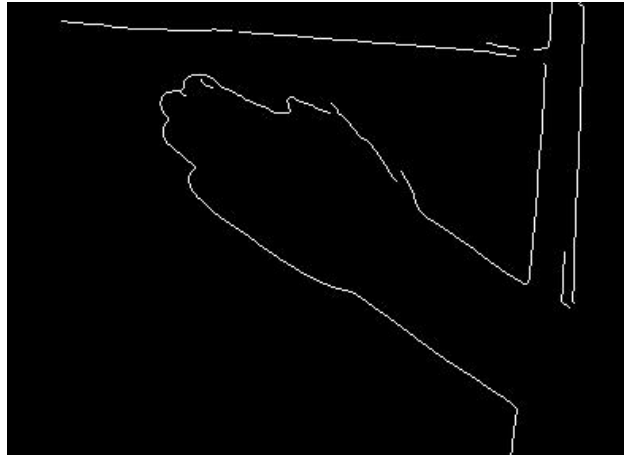
Figure 4-30: Canny edge detection on a typical IO image of the hand
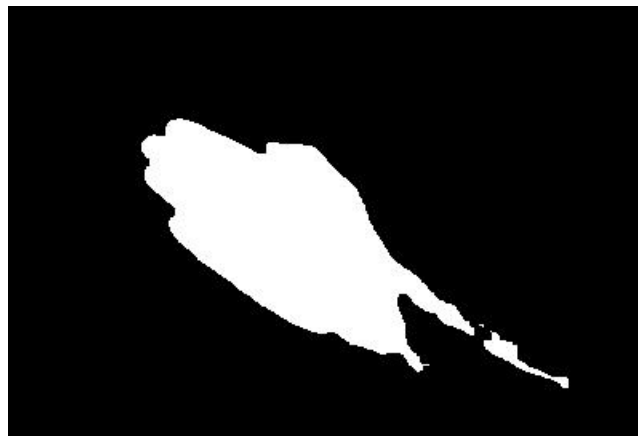
**Figure 4-31: Binary fill operation on a typical IR image
synchronised to the IO image in Figure 4-30**

There are 2 parameters that are used to change the IR/IO composition of the fused image: a) the ratio of IR pixel weighting to optical pixel weighting; b) the IR pixel level threshold in the IR image red channel below which only the optical pixel is added to the fused image. The typical overlay shown in Figure 4-29 occurs if the red channel in the IR image is greater than 96 and a combination of 0.8 times the temperature colour map from the IR image is combined with 0.2 times the colour map information from the optical

image. The IR pixel level threshold means that some areas of the fused image show only the optical image. The relatively large weighting of 0.8 given to the IR pixels compared to the optical image pixel weighting of 0.2 means that the areas where the IR pixels do contribute to the fused image, the optical pixel information is almost completely masked by the IR pixel. This is a greater weighting for the IR pixels than in [42], where an average value was used in the approximation band of their wavelet transform. However, more discussion about these parameters is given in Section 5.5 where the stereo calibration experimental results are presented.

4.7.1 Misalignment and solution

Occasionally, the proposed shifting algorithm, introduced above, does not work because the edge of the target limb or object is discontinuous, see Figure 4-32. In this case, the lateral shift will be miscalculated and the overlaid images will be misaligned, see Figure 4-33.

In these cases, an interactive function in Matlab, *cpselect*, is used which allows the user to manually select equivalent points on the IR and IO images using a mouse. The function returns the x and y co-ordinates for the selected points. The required horizontal shift is then calculated from the average of the differences between the x values for corresponding points. Similarly, if necessary, the vertical shift can be found from the average difference between the y values. Whilst this method works well and consistently, it is not as convenient as the automatic shift algorithm described earlier at the beginning of this section because it must be manually applied.
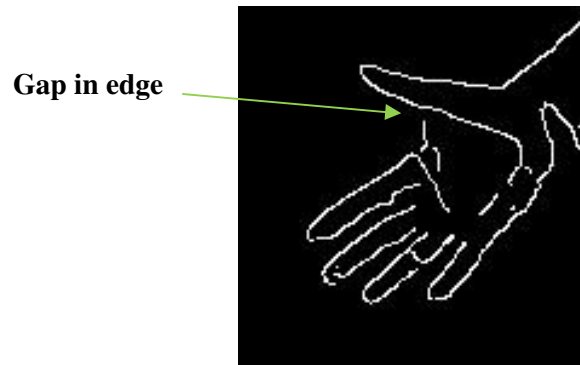
Figure 4-32: Edge of hand is discontinuous
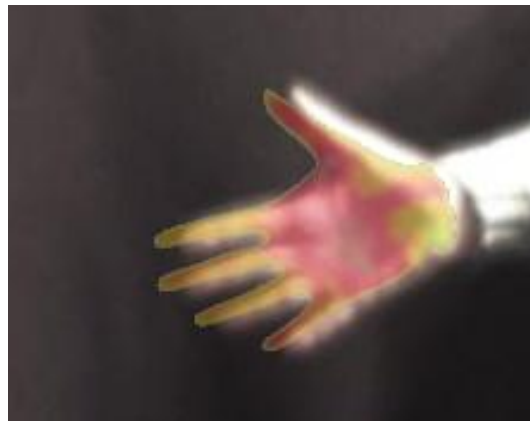
Figure 4-33: Misaligned overlay due to discontinuous edges

4.8 Summary of Chapter 4

Chapter 4 describes both the IR and IO feature point detection processes used on a prototype target which has 26 LEDs on the corners of its chessboard pattern. The details about various tests carried out on the prototype target are given along with the results from the tests. In particular, it gives a comparison between the intrinsic calibration results obtained with a conventional black and white chessboard calibration target and the prototype LED one. These results indicate that whilst the LED prototype target gives consistent results for both IR and IO feature points, the target would be improved by increasing the number of LEDs to 50. Thus, Chapter 4 contains a section about this improved target design and how the feature point processing is carried out with it. This

discussion includes sections on how both the IR and IO feature point positions are corrected through software algorithms that rely on the fact that the feature points should lie on a regular grid pattern.

After the stereo calibration and rectification of the system has been described, there is an explanation as to how the object or body part of interest is imaged with the system. This explanation includes a discussion of how corresponding pairs of IO and IR images are matched up. A description of the image fusion process then follows along with the approaches used to overcome the various associated problems.

5 Experiments and Results

In order to evaluate the accuracy of the calibration that can be achieved and also to experimentally determine the limitations of the proposed multi-spectral calibration technique, three different sets of experiments were carried out using the system described in the earlier chapters. In all cases, the results have been obtained using 10 pairs of IR and IO images, with a total of 480 IR and 480 IO feature points for each separate experiment.

The target was moved between image captures so that each separate target image's feature points were in different positions relative to those feature points in the other target images. The calibration path the target was moved through was the same as steps 1, 2 and 3 described in Section 4.3.1 where the target was placed on its long side and then tilted at 20° for different positions of the swivel chair with respect to the cameras' optical axes. However, this time, the target was not placed on its short side (as in step 4 of Section 4.3.1). This change was to ensure that the target was placed nearer to the horizontal on its long side which allows the correction algorithm described in Section 4.4.2 to work properly. The difference between placing the target at a 20° tilt and then on to its short side did not change the number of co-planar feature points significantly so it was decided to drop the short side position from the movement sequence.

Both the intrinsic and extrinsic parameters are presented as results separately for each different experimental setup. However, the analysis of all these experimental results is presented together in Sections 5.4.1 and 5.4.2 for the intrinsic and extrinsic parameters

respectively. For consistency, the focal lengths are given in multiples of image width along with their physical values in millimetres. A typical experimental setup can be seen in Figure 5-1. These experiments and their setups are described below.
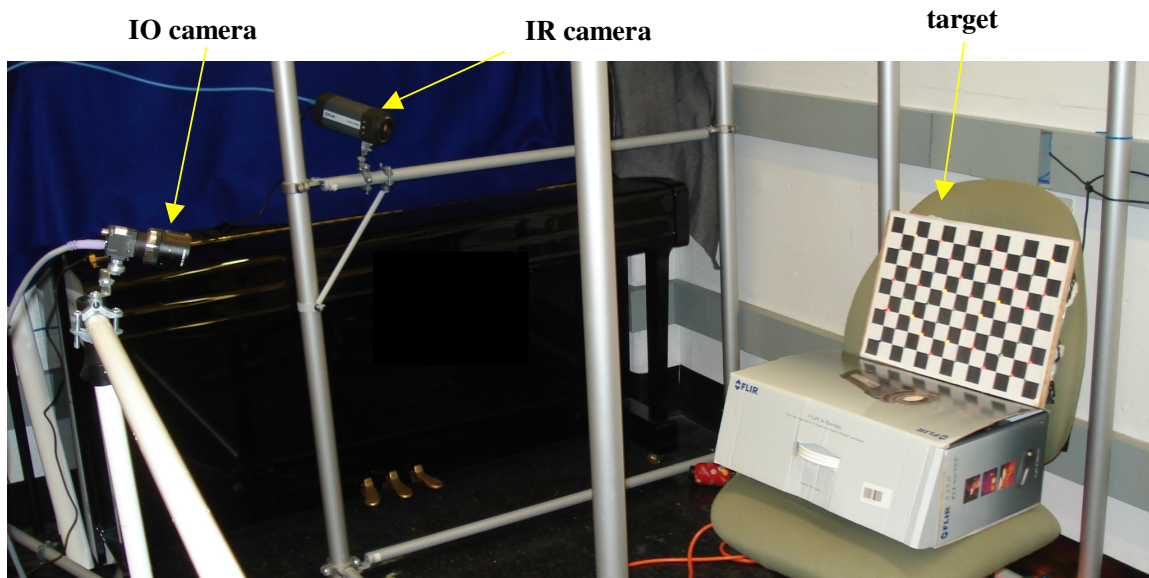
Figure 5-1: A typical setup for the calibration experiments

5.1 Experiment 1: parallel cameras

In the first experiment, the IO and IR cameras were placed so that their optical axes were in parallel. The calibration process was repeated with the cameras separated by different distances. The setup for this experiment is shown in the diagram in Figure 5-2.
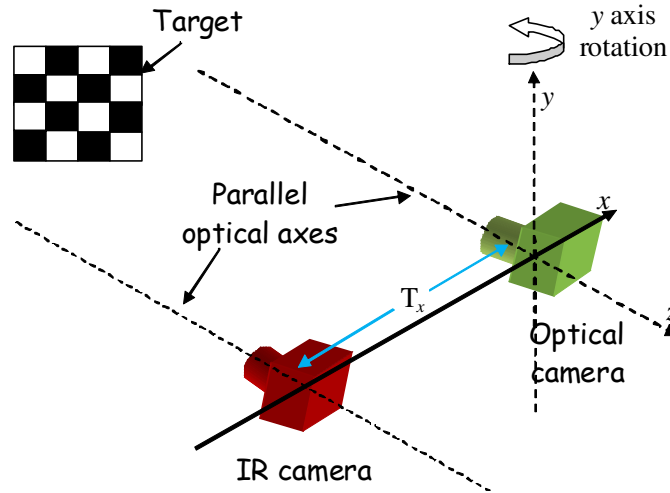
Figure 5-2: Setup for Experiment 1 – optical axes in parallel

The intrinsic and extrinsic parameters obtained experimentally for experiment one are given in Table 5-1 and Table 5-2 respectively for separation between the cameras varying from 10cm to 50cm in steps of 10cm. Note that the labelling of the axes is consistent throughout all the experiments described here.

Distance between IO & IR Cameras (cm)	Avg. Error in Epilines (pixels)	Radial distortion coeff.		IR (Ground truth f = 18mm)		IO (Ground truth f = 6mm)	
		$\|k_1\|$		f_x	f_y	f_x	f_y
		IR	IO	Image width (mm)		Image width (mm)	
10	0.81	0.09	0.11	2.38 (19.0)	2.40 (19.2)	1.37 (6.5)	1.37 (6.5)
20	1.03	0.16	0.13	2.34 (18.7)	2.34 (18.7)	1.34 (6.4)	1.34 (6.4)
30	0.77	0.05	0.16	2.46 (19.7)	2.46 (19.7)	1.31 (6.2)	1.31 (6.2)
40	0.95	0.04	0.29	2.56 (20.5)	2.51 (20.1)	1.34 (6.4)	1.34 (6.4)
50	0.95	0.15	0.02	2.50 (20.0)	2.60 (20.8)	1.36 (6.4)	1.36 (6.4)

Table 5-1: Intrinsic parameters from experiment 1

Note that the 'epiline' error in the third column is calculated as follows. The epiline from the left image drawn through a particular feature point is extended through to the right image. Then the distance of this epiline from the equivalent feature point in the right image is measured, giving an error for the alignment of these corresponding points. This process is repeated for all the feature points in the images in order to calculate the mean error. The significance of the epiline error will be discussed further in Section 5.4.1.

Parallel cameras distance (cm)		R		|T| (cm)	Rotation around axis (degrees)	
						Ground Truth	Results
10	x	0.9953	-0.0570	-0.0779	10.2	0	6.1
	y	0.0648	0.9926	0.1023	0.7	0	-4.1
	z	0.0715	-0.1068	0.9917	3.9	0	3.7
20		0.9970	-0.0381	-0.0679	20.4	0	3.2
		0.0418	0.9977	0.0527	0.1	0	-3.8
		0.0657	-0.0554	0.9963	2.3	0	2.4
30		0.9999	-0.0127	-0.0018	31.7	0	2.9
		0.0128	0.9986	0.0508	0.0	0	-0.1
		0.0012	-0.0509	0.9987	6.5	0	0.7
40		0.9970	0.0024	0.0778	38.3	0	-0.5
		-0.0017	1.0000	-0.0094	0.9	0	4.5
		-0.0778	0.0092	0.9969	6.3	0	0
50		0.9979	-0.0150	-0.0633	49.4	0	1.2
		0.0163	0.9997	0.0206	0.1	0	-3.6
		0.0630	-0.0215	0.9978	2.3	0	0.9

Table 5-2: Extrinsic parameters from experiment 1

In all tables presenting the extrinsic parameters, e.g. Table 5-2, the three rows from top to bottom are associated with the x, y and z axes shown in Figure 5-2. The first column of

each table provides the ground truth for the experiment: either the distance of separation between the two cameras (experiment 1) or the angle of inclination between them (experiments 2 and 3). The next three columns give the 3×3 rotation matrices for each separate experiment, as shown in equation 5-1, where the x, y and z rotations correspond to ϕ, θ and ψ angles respectively.

$$
\begin{bmatrix}
\cos\theta\cos\psi & \cos\phi\sin\psi + \sin\phi\sin\theta\cos\psi & \sin\phi\sin\psi + \cos\phi\sin\theta\cos\psi \\
-\cos\theta\sin\psi & \cos\phi\cos\psi - \sin\phi\sin\theta\sin\psi & \sin\phi\cos\psi - \cos\phi\sin\theta\sin\psi \\
-\sin\theta & -\sin\phi\cos\theta & \cos\phi\cos\theta
\end{bmatrix}
\qquad (5\text{-}1)
$$

The translation vectors are given in the fifth column of the extrinsic results tables. The other ground truths in the sixth column are the rotations about the y-axis in degrees for experiment 1 and the translation vector for experiments 2 and 3. The right hand column presents the rotation around the axis in degrees, as calculated from the rotation matrices results and equation 5-1.

After calibration, images were captured of a human hand for subsequent rectification and overlay after registration using the parameters experimentally estimated via the calibration procedure.

5.2 Experiment 2: inclined cameras around different circles

The intrinsic and extrinsic parameters given in Table 5-3 and Table 5-4 respectively were obtained during an experiment where the IR and IO cameras were placed 50cm apart. The two cameras were then inclined towards one another so that the axes of projection were no longer parallel. This angle between the cameras was varied during a series of experiments between 10 to 40 degrees with respect to one another. Consequently, in this

experiment, the radius to the centre of the projected circle changed with each different angle of inclination, as can be seen in Figure 5-3.

Note that the ground truth values for the translation matrix are obtained by projecting the 50cm on to the *x*-axis for the *x* component and the *z* component is found along the *z*-axis as $z = 50\sin\left(\frac{\theta}{2}\right)$ where θ is the angle between the two cameras and the 50 is the 50cm baseline between the two cameras. The same calculations were carried out for experiment 3 ground truths.
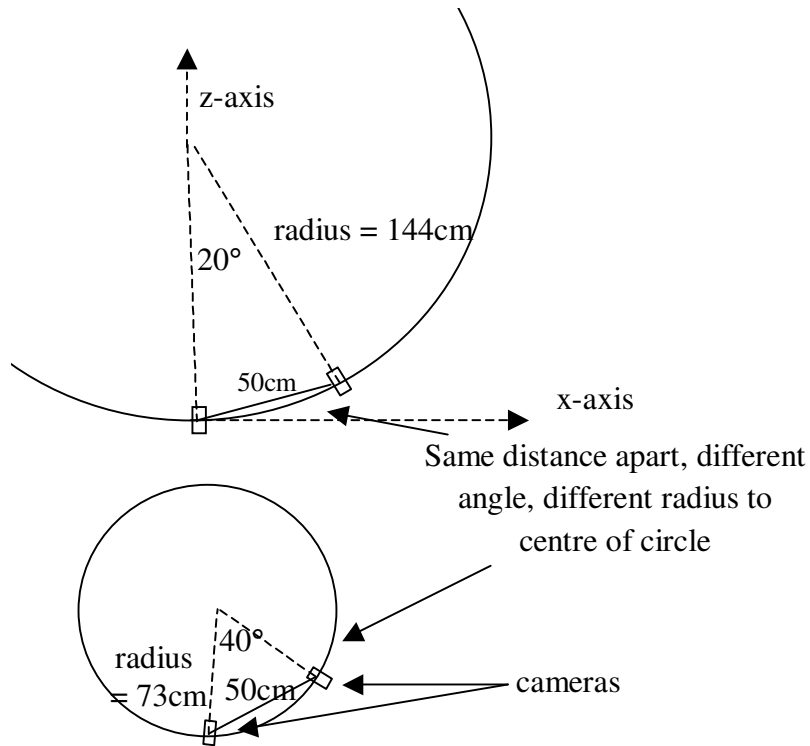
Figure 5-3 : Setup for Experiment 2 - optical axes inclined towards centres of circles with different radii

Angle between IO & IR Cameras (deg)	Radius to centre of projected circle (cm)	Avg. Error in Epilines (pixels)	Radial distortion coeff. $\lvert k_1 \rvert$		IR (Ground truth $f = 18$mm) f_x	f_y Image width (mm)	IO (Ground truth $f = 6$mm) f_x	f_y Image width (mm)
			IR	IO	f_x	f_y	f_x	f_y
10	286.8	0.93	0.90	0.20	2.56 (20.5)	2.55 (20.4)	1.44 (6.8)	1.41 (6.7)
20	144.0	0.77	0.05	0.07	2.43 (19.4)	2.42 (19.4)	1.34 (6.4)	1.35 (6.4)
30	96.6	1.14	0.12	0.11	2.39 (19.1)	2.42 (19.4)	1.41 (6.7)	1.39 (6.6)
40	73.1	1.05	0.03	0.25	2.40 (19.2)	2.38 (19.0)	1.36 (6.4)	1.35 (6.4)

Table 5-3: Intrinsic parameters from experiment 2

Angle between IO & IR Cameras (deg)		R			$\lvert T \rvert$ (cm) Results	Ground Truth	Rotation around axis (degrees)
	x	0.9818	-0.0149	-0.1894	54.6	49.8	3.8
10	y	0.0269	0.9978	0.0609	1.7	0	-10.8
	z	0. 1881	-0.0649	0.9800	2.2	4.4	1.6
		0.9284	-0.0756	-0.3639	52.6	49.2	2.6
20		0.0860	0.9962	0.0126	3.1	0	-21.2
		0.3615	-0.0430	0.9314	6.7	8.7	5.3
		0.8437	-0.0820	-0.5305	50.8	48.3	6.5
30		0.1203	0.9920	0.0380	5.9	0	-31.5
		0.5231	-0.0959	0.8468	13.9	12.9	8.1
		0.7058	-0.0315	-0.7077	47.2	46.9	6.6
40		0.0797	0.9962	0.0352	3.4	0	-44.7
		0.7039	-0.0813	0.7057	22.0	17.1	6.4

Table 5-4: Extrinsic parameters from experiment 2

5.3   Experiment 3: inclined cameras around the same circle

The intrinsic and extrinsic parameters given in Table 5-5 and Table 5-6 respectively were obtained during an experiment where the IR and IO cameras were placed variable distances apart. The two cameras were inclined towards one another so that the axes of projection met at the centre of a circle with radius 153cm. The angle between the cameras was varied during a series of experiments between 10 to 60 degrees with respect to one another by moving the IO camera away from the IR camera along the circumference of the circle. Note that the sign of the rotations around the x and y axes have been reversed when compared with the results from experiment 2. This is simply because the cameras have been swapped around in experiment 3 so that now the IR camera is to the left of the IO camera which is the opposite to their relative positions in experiment 2.

Figure 5-4 is a diagram of the experimental setup and an example of this experiment can be seen in Figure 5-1. Here the IO camera is mounted on a frame with wheels which is moved around a horizontal circle drawn on the floor. Consequently, in contrast with experiment 2, the radius to the centre of the projected circle remained fixed at 153cm with each different angle of inclination.

Different distance apart, different angle,
same sized circle
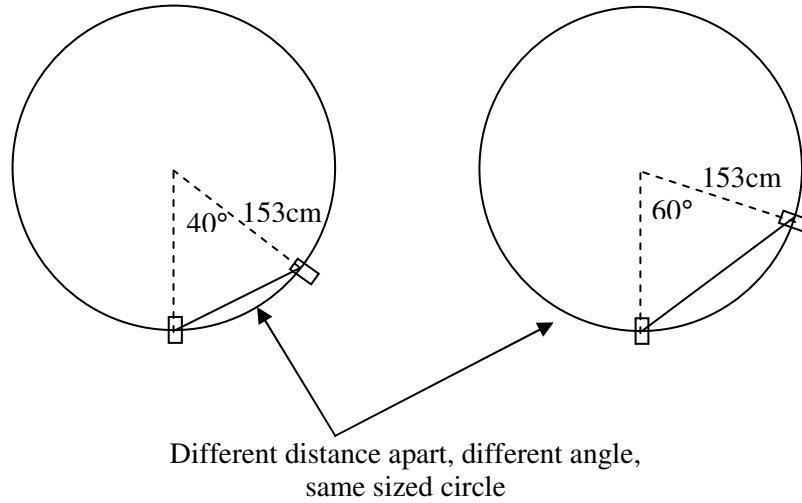
**Figure 5-4: Setup for Experiment 3 - optical axes inclined
towards the centre of the same circle**

Angle between IO & IR Cameras (deg)	Radius to centre of projected circle (cm)	Avg. Error in Epilines (pixels)	Radial distortion coeff. $\|k_1\|$		IR (Ground truth $f = 18$mm)		IO (Ground truth $f = 6$mm)	
			IR	IO	f_x	f_y	f_x	f_y
					Image width (mm)		Image width (mm)	
10	153	0.80	0.07	0.09	2.46 (19.7)	2.46 (19.7)	1.36 (6.4)	1.36 (6.4)
20	153	0.83	0.49	0.11	2.45 (19.6)	2.41 (19.3)	1.26 (6.0)	1.28 (6.1)
30	153	0.77	0.11	0.15	2.42 (19.4)	2.40 (19.2)	1.32 (6.3)	1.33 (6.3)
40	153	0.59	0.40	0.01	2.53 (19.2)	2.52 (20.2)	1.34 (6.4)	1.34 (6.4)
50	153	0.65	0.57	0.07	2.52 (20.2)	2.53 (20.2)	1.34 (6.4)	1.33 (6.3)
60	153	0.67	0.20	0.20	2.38 (19.0)	2.36 (18.9)	1.32 (6.3))	1.31 (6.2)

Table 5-5: Intrinsic parameters from experiment 3

However, for this experiment, it was not possible to use the synchronisation hardware because the connection cables on the prototype were too short to connect the cameras

when they were separated around the circumference of the relatively large circle. Therefore, when the system was tested after calibration, a small lamp was used to provide a static 'hot' object rather than a moving human hand, as introduced in Section 4.5, see Figure 4-26. The new static object position was manually set up between each image capture rather than the continuous set of images captured in the previous two experiments.

Angle between IO & IR Cameras (deg)	R			|T| (cm)		Rotation about axis (degrees)
				Results	Ground truth	
x	0.9911	0.0192	0.1317	31.1	26.5	-0.1
10	y	-0.0188	0.9998	-0.0044	3.3	0	7.6
z	-0.1317	0.0019	0.9913	0.9	2.3	1.1
	0.9502	-0.1166	0.2890	61.6	52.2	3.8
20	0.0930	0.9912	0.0944	4.5	0	17.3
	-0.2975	-0.0628	0.9527	14.2	9.2	5.6
	0.8534	-0.1674	0.4936	82.7	76.3	-5.9
30	0.1899	0.9818	0.0047	7.6	0	29.0
	-0.4854	0.0897	0.8697	25.8	20.4	12.5
	0.7584	-0.2114	0.6165	106.9	98.3	-5.8
40	0.2148	0.9742	0.0698	11.9	0	38.0
	-0.6154	0.0795	0.7842	31.8	35.7	15.8
	0.5852	-0.2444	0.7732	125.0	117	-10.4
50	0.2406	0.9629	0.1222	18.9	0	50.7
	-0.7743	0.1145	0.6223	53.2	54.6	22.4
	0.5361	-0.2190	0.8152	144.4	132.5	-14.2
60	0.2428	0.9650	0.0995	21.5	0	53.9
	-0.8085	0.1445	0.5705	68.5	76.5	24.4

Table 5-6: Extrinsic parameters from experiment 3

5.4 Analysis

The results from the three sets of experiments have been analysed and will now be discussed. Firstly, the analysis of the intrinsic parameters will be presented and this will be followed by the extrinsic parameter analysis.

5.4.1 Intrinsic parameter analysis

A summary of the results obtained for all experiments is presented in Table 5-7.

The radial distortion coefficient for the radius squared term is given in the fourth column. Here it can be seen that the overall average for this coefficient is 0.13 with a standard deviation of 0.08 for the IO camera and a mean of 0.23 with a standard deviation of 0.25 for the IR camera. Although there was no information available about the distortion coefficients from the camera manufacturers, calibration of the same IO camera was carried out by Bériault [51] with a conventional chessboard calibration target, without LEDs, that gave the same 0.13 result. In addition to this, as the distortion coefficient for the IR camera is the same order of magnitude as the one for the IO camera, the results for both camera calibration coefficients are considered reasonable.

Experiment	Parameter	Average error in epilines (pixels)	Radial distortion coeff. (IO)		IR		IO	
			$|k_1|$		f_x	f_y	f_x	f_y
			IR	IO	Image width (mm)		Image width (mm)	
1	mean	0.90	0.10	0.14	2.45 (19.6)	2.46 (19.7)	1.34 (6.4)	1.34 (6.4)
1	standard dev.	0.11	0.06	0.10	0.09	0.10	0.02	0.02
2	mean	0.97	0.28	0.16	2.45 (19.6)	2.44 (19.5)	1.39 (6.6)	1.38 (6.5)
2	standard dev.	0.16	0.42	0.08	0.08	0.07	0.05	0.03
3	mean	0.72	0.31	0.11	2.46 (19.7)	2.45 (19.6)	1.32 (6.3)	1.33 (6.3)
3	standard dev.	0.10	0.21	0.07	0.06	0.07	0.03	0.03
All	**mean**	**0.85**	**0.23**	**0.13**	**2.45 (19.6)**	**2.45 (19.6)**	**1.35 (6.4)**	**1.34 (6.4)**
All	**Standard dev.**	**0.16**	**0.25**	**0.08**	**0.07**	**0.08**	**0.04**	**0.03**

Table 5-7: A summary of the intrinsic parameters

An example of the transformed images, obtained via the process described in Section 4.1, can be seen in Figure 5-5 with some of the horizontal epilines superimposed on top. As explained previously, the 'epiline' error in the third column is referring to the distance of a feature point in one image, from the epiline drawn through the equivalent feature point in the other image and vice versa. Thus, as can be seen in Figure 5-5, the error is referring to the vertical alignment of the two images because the rectification process involves vertically aligning the horizontal epilines, as described earlier in the Section on theory, Section 2.1.

IR feature point   **Corresponding IO feature point**

(a) Frontal view

IR feature point   **Corresponding IO feature point**
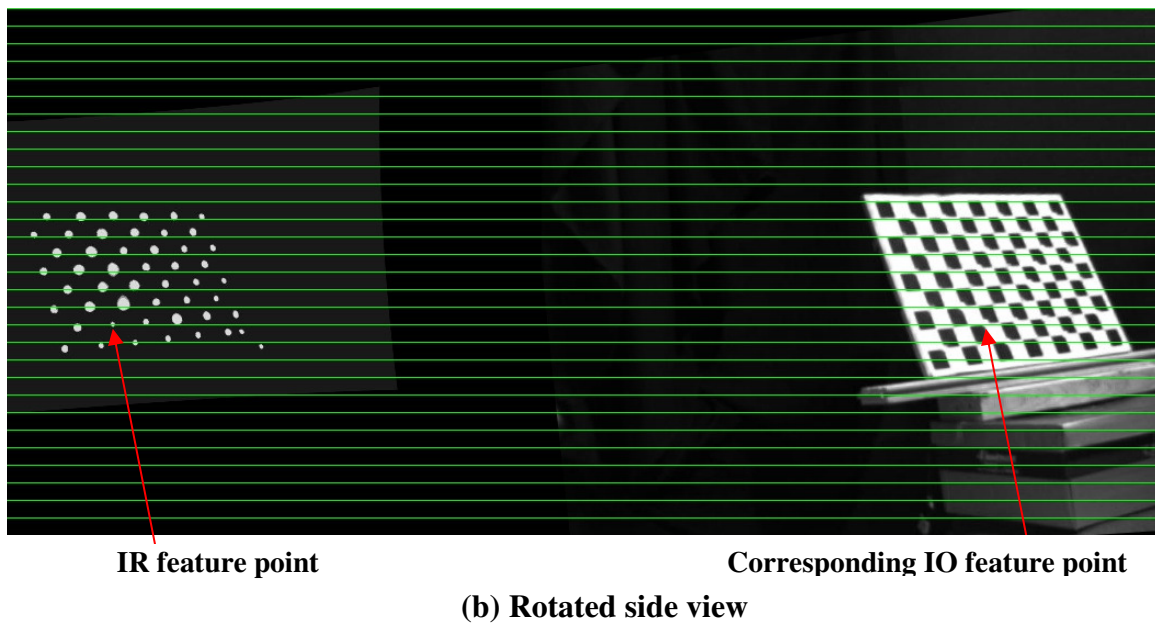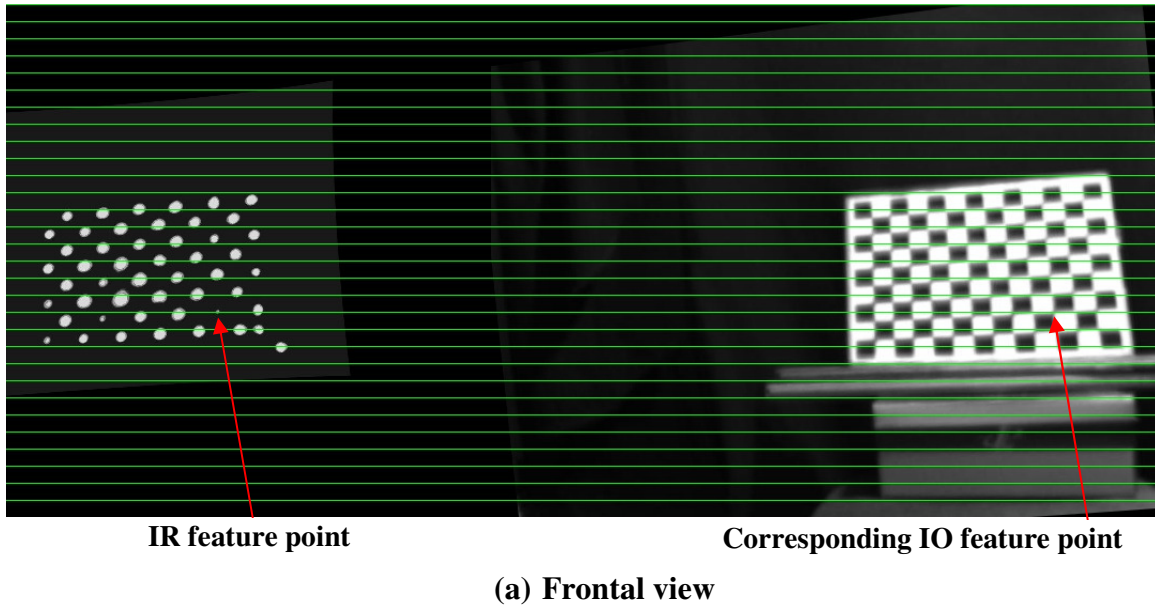
(b) Rotated side view
Figure 5-5: Rectified target images from experiment 1
with a separation between cameras of 20 cm

As listed in Table 5-7, the overall mean error in the epilines is 0.85 pixels with a standard deviation of 0.16 pixels. This error is calculated automatically by the calibration software. The error quoted here is based on the combined results from all three experiments, each of which consisted of between 4 and 6 different camera positions (see

Sections 5.1 to 5.3) where each position had ten separate images associated with it and each of those images contained 48 IR and 48 IO feature points.

Two examples are shown in Figure 5-5 to illustrate the fact that the rectification works with images of the target that are viewed directly from the front (see (a)) and also with those where the target has been rotated to the side by ~30° (see (b)). The example shown here is for the 20 cm separation from experiment 1 because this had an epiline error greater (1.03 pixels) than the mean (0.85 pixels) to illustrate that even the examples with the larger errors do not show a significant misalignment to the naked eye.

The statistics for the results of each separate experiment are also presented in Table 5-7. These statistics indicate that whilst the mean error for the distance from the epilines is comparable for the first two experiments (0.90 and 0.97 pixels), experiment 3 has a mean error of only 0.72. However, this latter value is misleading because the larger angles provided little overlap of the image between the two cameras. This meant that it was not possible to vary the position of the target as widely as with the other experiments, so that the 10 images used in the calibration process displayed feature points that had not moved as significantly between images as with the earlier experiments. Although experiment 2 also involved changing the cameras' angle of inclination, because the bigger angles lead to the target being closer to the cameras, even relatively small movements in the target gave relatively large changes in pixel values between the feature points. In contrast, experiment 3 had the target the same distance away from the cameras at the centre of the constant radius circle and this distance was 153cm compared with 73cm for the 40° angle of inclination case. Further discussion of the consequences of these results will be given shortly.

The results for the focal length when measured in the *x* and *y* direction separately are also given in Table 5-1 to Table 5-7 under the columns labelled f_x and f_y. In the IR case, the results for all experiments are the same to 1 decimal place for both directions and overall the result is 2.5± 0.2 image widths, or 19.6±1.6 mm. This latter figure overlaps with the ground truth of 18mm IR focal length but is a higher value than the result obtained with the smaller prototype target.

In the IO case, the results for experiments 1 and 3 are in agreement to within 1 decimal place: 1.33 ± 0.06 image widths, equivalent to 6.3 ± 0.5 mm. However, experiment 2 gives a slightly higher result of 1.39 ± 0.10 image widths, equivalent to 6.6 ± 0.5mm where these results are a summary of both the *x* and *y* results. However, the error associated with each experiment means that there is overlap between all the results and, after combining the results from all three sets of experiments, the overall f_x and f_y means are given as 1.35±0.08 and 1.34±0.06 pixels respectively, which are equivalent to 6.4±0.4 mm in both cases. This latter result agrees with the ground truth of a 6mm IO camera focal length.

Thus, all the experiments give results that are in agreement with the ground truth for the intrinsic parameters for both cameras. Thus, the use of the final IR/IO combined target in the multi-spectral stereo calibration process appears reliable.

5.4.2 Extrinsic parameter analysis

The extrinsic parameters are presented for each separate experiment in Table 5-2, Table 5-4 and Table 5-6 respectively.

For convenience, the key values for this analysis are given in Table 5-8. That is, the translation along the x axis (T_x) which gives the baseline between the two cameras in accordance with Figure 5-2 and the rotation around the y axis which gives the inclination of the cameras towards one another.

Experimental setup: Ground truth	$\|T_x\|$ (cm)	Rotation around y-axis (degrees)	Difference from ground truth	
Exp. 1: Parallel cameras distance (cm)			cm	degrees
10	10.2	4.1	0.2	4.1
20	20.4	3.8	0.4	3.8
30	31.7	0.1	**1.7**	0.1
40	38.3	4.5	**1.7**	4.5
50	49.4	3.6	0.6	3.6
Exp. 2: Inclined cameras; Diff radius (degrees)				
10	54.6	10.8	4.8	0.8°
20	52.6	21.2	3.4	1.2°
30	50.8	31.5	2.5	1.5°
40	47.2	44.7	0.3	**4.7°**
Exp. 3: Inclined cameras; Same radius (degrees)				
10	31.1	7.6	4.6	2.4°
20	61.6	17.3	9.4	2.7°
30	82.7	29.0	6.4	1.0°
40	106.9	38.0	8.6	2.0°
50	125.0	50.7	8.0	0.7°
60	144.4	53.9	11.9	**6.1°**

Table 5-8: Selected extrinsic parameter results

It should be noted that there are errors associated with the so-called 'ground truth' values presented in the first column of Table 5-8. For instance, the error associated with the

second and third experimental setup arises because these experiments entailed inclining the cameras around a circle and it is difficult to align the cameras accurately at a particular angle considering that at least one of the cameras is mounted on a moving frame. The error associated with the 'ground truth' values for these experiments is estimated to be ±3°. Conversely, experiment 1 was relatively easy to set-up since it only involved measuring the baseline separation between the two cameras with a ruler. The error associated with these ground truths is estimated to be ±0.7cm. This error is partly due to the fact that it is not known exactly where the centres of the sensors are located within the cameras and partly due to measurement error.

As can be seen in the last two columns of Table 5-8, the differences between the results and the ground truth values are given. As highlighted in bold in these columns, there are only a small number of cases where the differences fall outside the error range of the ground truth values: experiment 1 where the baselines are 30 and 40cm; experiment 2 where the inclination is 40°; experiment 3 where the inclination is 60°.

For the parallel camera cases, the difference from the ground truth is 1.7cm for both the 30cm and 40cm cases. This 1.7cm error does not take into account that there is an error associated with the ground truth itself. In reality, the accuracy of the calibration is probably somewhat better than ±1.7cm since the ground truth may actually have been as much as 30.7cm in the 30cm case and as little as 39.3cm in the 40cm case, if the maximum error for setup occurred with both these cases. In this case, the difference associated with the 30cm and 40cm cases would be only 1cm. However, quoting an error of ±1.7cm combines the effects of both the error of experimental setup and the error in the calibration process which gives an overall error associated with the system. For

completion, the other ground truth is a zero angle of rotation around all three axes because the cameras should be aligned in parallel. It can be seen in Table 5-2 that this is the case within 6° in all cases.

For the inclined cameras in experiment 3, the 60° case gives a difference of 6.1° from the ground truth which is outside the ±3° error associated with the ground truth itself. Similarly, the 40° case in experiment 2 gives a difference of 4.7°, which is also outside the ground truth error. However, later in this thesis, see Section 5.5.4, it will be argued that the system should not use IR and IO cameras which are inclined more than 30° towards one another. Therefore, if only the cases with camera inclinations of less ≤ 30° are considered in Table 5-8, an error of ±3° is reasonable as the differences given for all these cases fall within this range. For completeness, the ground truth of the translation vectors are also given in Table 5-4 and Table 5-6 for experiments 2 and 3 respectively.

For experiment 2 results, it can be seen in column 6 of Table 5-4 that the baseline between the two cameras is kept to within 5cm of the ground truth for all the angles of inclination. The ground truth of zero translation along the y-axis is also maintained to within ~6cm. The results for translation along the z-axis also agree well with the ground truth, differing by less than 3cm for all angles of inclination between the cameras. These differences were not considered a significant problem as the main parameter of interest in experiment 2 was the angle of inclination between the cameras.

For experiment 3 results, the ground truth in Table 5-6, column 6, shows reasonable agreement for the x and z axes translation with no more than 12cm of difference with any of the camera inclination angles. The ground truth of zero cm translation along the y-axis

is deviated from, however, by as much as ~21cm in the 60 degree camera inclination case. This relatively large deviation was not considered a major problem as it only represents a difference in vertical height between the two cameras and the main parameter of interest was the angle of inclination between the cameras for this experiment.

In Table 5-2, Table 5-4 and Table 5-6, the divergence on less critical extrinsic parameters, T_y, T_z, R_x and R_z are minor for all experiments. The observed small variations are related to imprecise positioning of the cameras.

Other observations in Table 5-8 are:

- In experiment 1, the cameras' optical axes are not perfectly in parallel as indicated by the rotation around the y-axis in the third column which varies between 0.1° and 4.5°.

- In experiment 2, as described in the experimental setup in Section 5.2, the cameras were placed approximately 50cm apart. Column 2 of Table 5-8 shows that this was indeed the case within ±4.6cm.

- In experiment 3, the cameras were placed increasingly further apart for each increment of 10° of inclination because they were being situated around the circumference of the same circle. Thus, as expected from the theoretically calculated values in Table 5-6, the baseline separation distance results, T_x, in column 2 of Table 5-8, show an increase with degrees of inclination.

5.5   *Experimental Evaluation of IR and IO Data Fusion*

Before examining the experimental results obtained on data fusion, two important points need to be made about the parameters used to obtain the fused images:

- During the experiments, the IR camera was set to detect both narrow and wide temperature ranges in order to allow a qualitative assessment of how useful the fused IR/IO images would be to a doctor in a clinical setting.

- As a result of the fused images obtained during early experiments, see Section 4.7, it was decided to try different IR to IO weighting ratios from 0.8:0.2 to a more evenly balanced ratio. In particular, the weighting ratios of 0.5:0.5 and 0.4:0.6 were tested. A more evenly balance ratio means that the IR pixel contribution to the fused image does not mask the IO pixel contribution so significantly. This has the advantage that, whilst the temperature distribution is still presented to the viewer, the detail of the IO image is also still visible, thus combining the advantages of both images. With the earlier experiments, where the 0.8:0.2 weighting was used, the detail of the IO image was lost wherever the IR pixels were included in the fused image.

The output from the various experiments will now be presented.

5.5.1 Fusion of IR and IO images in experiment 1
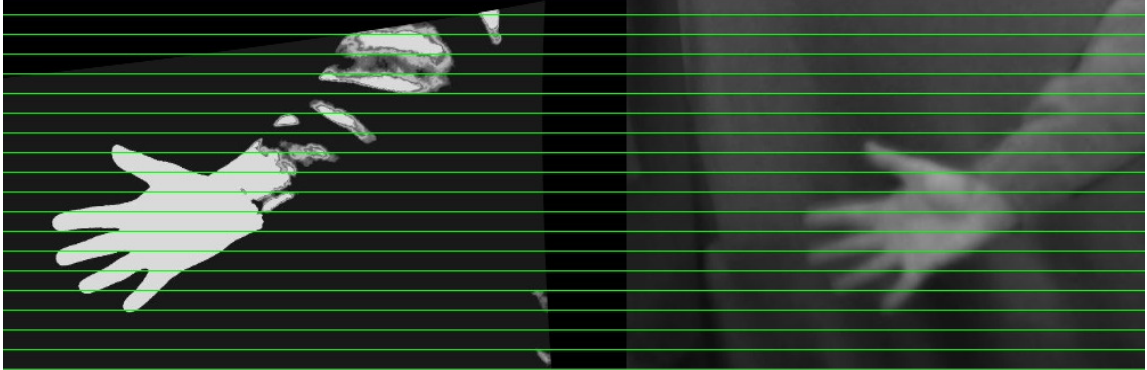
Figure 5-6: Rectified hand from experiment 1 with a camera separation of 10cm

Figure 5-6 shows the rectified IO and IR images from experiment 1 where the cameras were separated by 10cm and have already been subjected to the stereo calibration process. In this case, the IR camera was set to detect a narrow temperature range of 22° to 24°C.

The result of overlaying the two images with a weighting of IR to IO of 0.4:0.6 can be seen in Figure 5-7 for both camera separations of 10 and 30cm. In these cases, the narrow temperature range means the IO image of the hand is almost completely concealed by the IR overlay as this latter has a saturated pixel level over the entire region of the hand because the hand actually has a range of temperatures that exceeds the 24°C. Nevertheless, Figure 5-6 and Figure 5-7 show that the IR and IO images have been both successfully registered and fused.

24°C

22°C

(a) Camera separation of 10cm

(b) Camera separation of 30cm

Figure 5-7: Overlay of parallel camera IR & IO images with a weighting of 0.4:0.6

For the clinical diagnostic application of interest here, it is considered beneficial to be able to distinguish between areas of the body that have high and low temperatures. Therefore, the IR camera was readjusted to recognise a greater temperature range, typically 22° to 33°C, although this was varied according to the subject's temperature at the time of the experiment.

Other examples of overlays from experiment 1 with the parallel camera case are shown in Figure 5-8. In these cases, whilst the weighting IR/IO ratio has remained at 0.4:0.6, the temperature range has been increased. The overlays in Figure 5-8 clearly distinguish

between regions of high and low temperatures. For example, the knuckle area of Figure 5-8(a) is cooler than the thumb and in Figure 5-8(b) the palm of the hand is hotter than the finger tips.

33°C

22°C
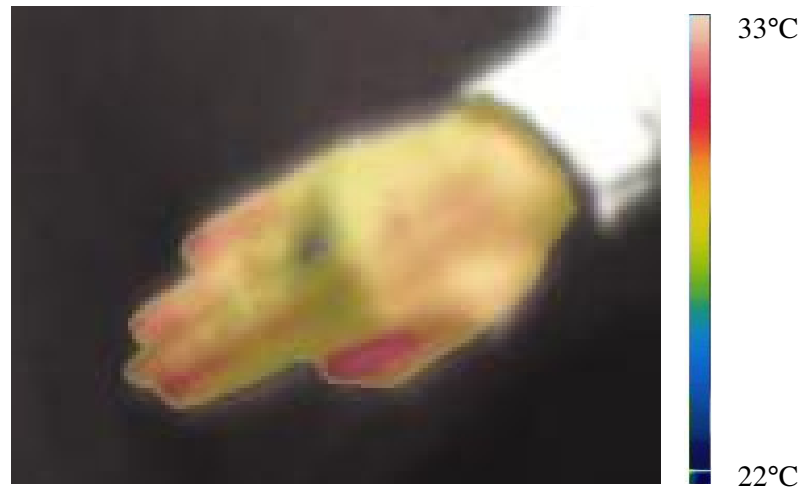
(a) Camera separation of 20cm

(b) Camera separation of 50cm

Figure 5-8: Examples of overlay with 0.4:0.6 IR:IO weighting from experiment 1

5.5.2 Fusion of IR and IO images in Experiment 2

An example of the rectified and aligned (see Sections 2.1 and 4.4) IR and IO images from experiment 2 is given in Figure 5-9, where the angle between the cameras has been set at 40°. As can be seen the epilines are properly aligned across the images.
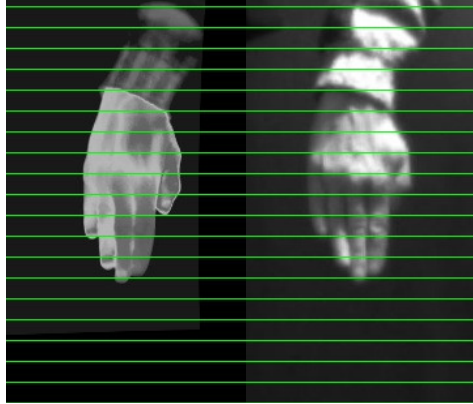
Figure 5-9: Rectified hand from experiment 2 with angle between cameras of 40°

Examples of overlays from experiment 2, with various angle values between the cameras are shown in Figure 5-10. Note that the experiments were not all carried out on the same day. The significance of this is that the minimum and maximum of the temperature range exhibited by the hand changed between the periods when the experiments were undertaken, even though at least 20 minutes of room acclimatisation was undergone prior to image capture for all the experiments. However, even though the range of temperatures was shifted between experiments, the actual size of the range between the minimum to maximum remained approximately the same, as would be expected for a healthy body part. This is important if the system is to be used to assess RA disease, as discussed in Section 4.7.

In Figure 5-10, the temperature range was the same as for experiment 1 that is 22° to 33°C. In (a), the angle of inclination between the cameras was 20° but the ratio between the IR and IO overlay was set at 0.5:0.5, i.e. equal amounts from each image contributed to the picture. In this case, some of the detail has been lost from the IO image.

(a) Camera inclination of 20°

(b) Camera inclination of 30°

(c) Camera inclination of 40°

**Figure 5-10: Examples of overlaid images from experiment 2
where the angle between the cameras has been varied**

In contrast, the overlays shown in figures (b) and (c) had the ratio of 0.4:0.6 set so that

more detail of the underlying hand is visible – note in particular the area of the thumb

joint in figure (b). This extra detail is essential if a doctor is going to be able to use the system to quantitatively compare the temperature distributions over exactly the same area of the body during different visits by the patient.

Figure (c) shows the problem with using a large angle of inclination between the two cameras. The region of overlap has become much smaller because the radius of the circle, defined by where the two cameras' axes of projection intersect, is much shorter. Thus, in this case, only the fingers of the hand were captured in the IR image. There is much less room for manoeuvre for the patient in this particular experimental setup. This was why experiment 2 only tested up to an angle of 40° between the cameras. Anything larger than an angle of 40° between the IR and IO camera principal axes did not allow a sufficiently varied number of target images to be captured during the calibration process.

Figures (a), (b) and (c) also change the temperature threshold that is used to decide whether or not to add the IR image pixels to the IO image as described in Section 4.7. If a pixel value in the image is lower than the threshold, it is not added to the IO image. With image (a) the threshold for IR addition was set to 64, i.e. at a quarter of the 256 possible levels in the 8 bit image. Thus, any pixel with a value above 64 is added to the fused image. With image (b) the threshold was set at 128, i.e. at half of the 256 levels. The effect of increasing this threshold shows that the IR information is missing at the edge of the hand along with the tip and edge of the forefinger because many more pixels have values which fall below this higher threshold. This could be useful for diagnostics if the user of the system wishes to concentrate on the regions of elevated temperature only. However, the consequences of raising the threshold too far can be seen in figure (c)

where the IR threshold was set at 192 and a large part of the IR information has been lost in the fingers area which is the only region of overlap for both cameras.

5.5.3 Fusion of IR and IO images in experiment 3

An example of the rectified and aligned IR and IO images from experiment 3 is given in Figure 5-11, where the cameras have been placed around the circumference of the circle with their optical axes pointed towards its centre so that the angle between the cameras is 10°. Here the 'hot' object has been used for stability of the subject due to lack of synchronisation, as discussed earlier. As can be seen the epilines are properly aligned across the images, thanks to the calibration procedure.

Figure 5-11: Rectified IR & IO images for experiment 3
with a relative camera angle of 10°

The overlaid IR&IO images shown in Figure 5-12 were obtained whilst carrying out experiment 3 with a weighting ratio of IR to IO of 0.4:0.6. Firstly, note that overlays have been successfully obtained for all the cases where the angle between the two cameras has been varied from 10° to 50° in steps of 10°. There are some differences between the various images which will now be discussed.

44°C

44°C

28°C

26°C
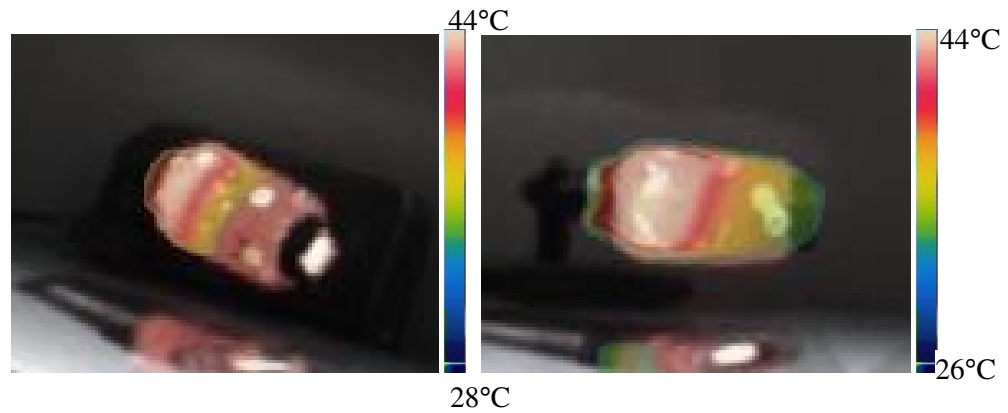
(a) Relative angle of cameras = 10° **(b) Relative angle of cameras = 20°**

44°C

26°C
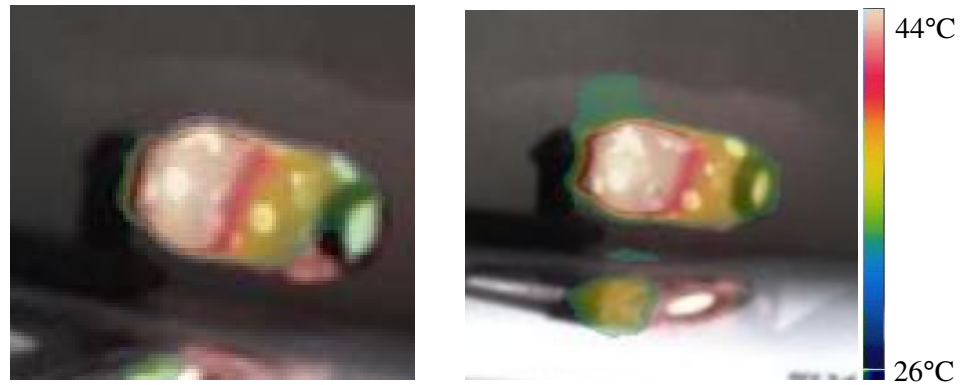
(c) Relative angle of cameras = 30° **(d) Relative angle of cameras = 40°**

44°C

26°C

(e) Relative angle of cameras = 50°

Figure 5-12: Examples of IR & IO overlaid images of the object in experiment 3

The image in (a) was obtained with the temperature range for capture with the IR camera set to 28° to 44°, whereas for all the other images the range was widened to 26° to 44°. The reason for this increase in range is that only the left hand side of the object was

visible in the IR image without the lower temperatures. In all cases, the threshold for whether an IR pixel was included in the overlay image was set to 128.

Another difference can be observed in figures (b) and (c) where the whole object is overlaid with the IR image. This difference arises purely because the lamp had been switched on for longer when these particular images had been taken compared with the other ones and the tip of the lamp had risen further in temperature.

In several of the images, the box that the lamp is sitting on is warm near the base of the lamp after it has been switched on for a while. This explains the patch of green-yellow hue just below the lamp in images (b) to (e). In fact, in image (d), the lamp has warmed up areas above it as well which has lead to a blue tinge just above the lamp as well.

Overall, however, the experiments have shown that, for all the tested camera inclinations, as long as the object of interest is placed so that it crosses the optical axes of the cameras symmetrically, proper estimation of the intrinsic and extrinsic parameters has been achieved which led to:

- Adequate rectification and alignment of IR and IO images, as shown by the coherent contours and edges location throughout all experiments.

- Correct registration, which allowed accurate fusion of IR and IO which can be of use for assessment of muskuloskeletal disease, such as RA, that produces temperature variations of the surface of the skin.

More discussion about the object positioning will be discussed in the following Section 5.5.4.

5.5.4 System limitations

However, although these images show that an accurate overlay of the IR and IO images can be obtained using the proposed method of calibration and registration. There remain some problems with the technique for the larger angles between the cameras.

Whilst the results of Table 5-7 suggest that the error between feature points and epilines is smaller for experiment 3 than in the other two experiments, this is misleading. The mean error is lower here because the error given in Table 5-5 for the angles between the cameras of 40° to 60° is also lower than all the other results. One of the reasons these errors are so low is that, because there is a relatively small region of overlap between the cameras' field of view when reaching 40° to 60° angular difference between the optical axes, the target cannot be moved as much between each image capture during the calibration process. This means that the feature points are not as differently positioned with respect to one another as they are with the other experimental cases. Even experiment 2 does not suffer so much from this problem because the target is moved closer to the cameras when the angle between them is increased. Thus, the difference in feature points is further in pixel measurements relative to the same angle of inclination in images captured in experiment 3. The consequence of this limitation on target movement is that the calibrated working volume is smaller for the 40° to 60° angles than that of the other experiments. Such angles of inclination would, therefore, reduce the effectiveness of the system in a clinical assessment application because the patient would have to confine the movement of his body part to this limited working volume.

Thus, the target in experiment 3 cannot be twisted at such a large angle towards or away from the vertical plane parallel to the two cameras, and passing through the centre of the

119

calibration circle, because one of the cameras would not be able to see it. This means that the effective calibrated volume is smaller for these larger angles and at 40°, the calibration and, consequently, the rectification break down if the object is rotated too much to one side away from this parallel plane, as can be seen in Figure 5-13. Here the IR and IO images are clearly not overlaid correctly and insufficient rotation of the IR image has taken place with respect to the IO image, as a result of erroneous calibration parameters.

Figure 5-13: Object is rotated out of the calibrated volume for the 40° angle case

With the larger angles of inclination, the problem arises at even quite small object rotations. An example for the 50° case can be seen in Figure 5-14.

End of lamp in IO image

End of lamp in IR image

Figure 5-14: A small object rotation causes a misalignment at an angle of 50°

As was shown in Section 5.5.3, if the object is aligned with the parallel plane through the centre of the calibration circle and parallel to the two cameras' image planes, an overlaid IR and IO image can be obtained. In this case, the horizontal epilines for the two images are in alignment in the x direction, as can be seen in Figure 5-15.

**Figure 5-15: An example of epiline alignment for the 50° angle case
when the object is aligned directly in front of the two cameras.**

However, there will be cases where it is inconvenient to hold the body part that is being analysed directly in front of the cameras like this. For instance, a doctor may wish to examine a joint from several different viewpoints, such as a wrist or knee in the bent and extended position.

5.6 System Design Recommendations

As a result of the extensive experimentation conducted with the first prototype of the system, here are some recommendations as to how the system could be enhanced:

1. Unlike now, the system must have more than one IR camera in it so that an IO camera is not placed at an angle more than 30° with respect to the IR camera. If the system is to cover a semi-circle around the body part, this would require an IR camera to be placed at every 60° around the semi-circle. This means a minimum of 3 IR cameras per system.

2. Another possibility is to use the 48 LED target described in Section 4.4 to carry out only the intrinsic calibration. A second, smaller and more flexible target could then be used for the extrinsic calibration part of the process. This proposed target could be something like that shown in Figure 5-16 where one end of the calibration stick has a red and a blue LED and the other end has a single blue LED. A similar approach with a laser pointer is described in [33] and a 2-LED stick is used with IO cameras in [51].

   The stereo calibration process would then involve recognising the three LEDs as feature points. The two LEDs at one end of the stick would be easily distinguishable from the single LED at the other end in the IR images. The red and blue LEDs will be distinguishable in an RGB image captured with the IO camera.

Figure 5-16: Alternative more manoeuvrable stereo calibration stick

A slight modification of this stereo calibration stick would be to mount the LEDs so that they are sticking out from the side of the wooden bar, rather than protruding through the middle of the bar, as shown in Figure 5-16. This would give the additional advantage that the feature points would not be occluded from one of the cameras by moving it around a large volume. This removes the problem with the existing target that when twisted away from the plane facing a camera, it becomes occluded from it.

3. The IR image could just be overlaid on one view of the IO image, that obtained by the IO camera immediately adjacent to the IR camera where the two optical axes are almost parallel. Then this IR/IO stereo camera pair could be moved to different locations around the body part and a new stereo calibration procedure could be carried out each time. Alternatively, the body part being analysed could be moved to new orientations and a new series of image frame capture could be

carried out for each new position. This latter alternative has the advantage that the stereo calibration process would not have to be repeated for each new position.

4. Once Point Grey releases a new version of software which includes time stamping of a sequence of captured images, the system software should be updated to enable automatic time correspondence between the IR and IO images to be found.

5. The length of cable connecting the synchronisation hardware to the cameras should be extended to its maximum of 3m to allow more flexibility with camera separation, see Section 3.3.4.

5.7   *Summary of Chapter 5*

Chapter 5 provides a detailed description of the three different experimental set-ups that have been used to test the multi-spectral stereo calibration approach. The results for both intrinsic and extrinsic results are presented and an analysis of the data is also given. Successfully fused IR and IO images, which were obtained from the experiments, are shown. However, some system limitations are also highlighted here which have led to some recommendations for improvement to the system being made.

However, it has been shown through these sets of results that the 50 LED calibration target can be used to obtain accurate overlays of IR and IO images which can then be used as part of a system to assess RA disease, where such a system is based on observing changes to the temperature distribution, provided by the IR camera, over a body part of interest whilst simultaneously providing the clear body part definition from the IO camera.

6 Conclusion

This final chapter commences with a summary of the principal findings from the project. This is followed by a discussion as to how well the objectives, initially laid out in Section 1.2, have been achieved. A list of the contributions made to research as an outcome of the work undertaken for this thesis is then presented before giving suggestions as to how the work might progress in the future.

6.1 Summary of Findings

The proposed combined IR and optical target worked well for simultaneously calibrating the multi-spectral cameras. An overall mean error of 0.85 pixel was measured for the distance of a feature point in an image captured with one of the cameras, from the epiline in the image captured with the other camera. The standard deviation of this error was 0.16 pixel.

Successful stereo calibration was carried out between the two cameras for all three of the different system configurations that were trialled. If the object of interest was symmetrically positioned across the optical axes of the cameras, successful registration and correspondence between the cameras was achieved. However, if the cameras are placed widely apart and at an angle of inclination greater than 30° to one another, it was not possible to successfully fuse the IR and IO images if the subject was orientated so that it presented a different face to one camera from the other. These results suggest that if a more complex RA assessment system is required, where a 360° 3D surface is

reconstructed with combined IR and optical images from all angles, then at least six IR and six IO cameras would be necessary.

6.2 Contributions

There are three main contributions attributable to the work described in this thesis:

1. The successful construction and testing of a new type of IR and optical calibration target has been completed. The new target, which comprises a black and white chessboard pattern and LED combination, allows both IR and optical cameras to be calibrated simultaneously and interactively. In addition to this, the target can be used for an indefinitely long period as it remains visible by an IR camera as the heat signature of the target does not diminish over time, unlike the popular technique reported in recent literature where a chessboard target is irradiated with a heat lamp prior to calibration. As in [41], an advantage to using this target is that it also makes the camera system independent of external lighting conditions.

2. An investigation into possible larger system geometries by analysing the calibration accuracy with just one IR and one optical camera placed at various distances between each other and at various inclinations towards each other. In particular, as a result of this investigation, a recommendation has been made for the number of IR and IO cameras to be included in a more complex system which could then be used as an assessment tool for RA. Some other suggestions for system improvement have also been made as a result of this investigation, as reported in Section 5.6.

3. The successful fusion of registered IR and optical images has allowed an investigation to be undertaken to help determine the most effective way of overlaying IR and optical images with a view to using the fused images as a diagnostic aid. Two parameters have been considered:

    o   The weighting ratio given to IR image pixels and their corresponding optical pixels.

    o   The threshold level in an IR image which determines whether or not each pixel is included in the fused image. If a particular IR pixel value is less than the threshold, it is discarded and the fused image will only have the optical pixel included at this point.

    Although only qualitatively assessed, the preferred results for the fused multi-spectral images was where the ratio of the IR to IO images was set to 0.4:0.6 with pixels below a red channel level of 128 of 256 being discarded from the IR image during fusion. These preferred parameters are aimed at a diagnostic application where a medical practitioner would want to concentrate on the higher temperature regions of the body. However, this is very subjective and a user would probably want to have the ability to change these parameters interactively. This would require the addition of a user interface to the system.

6.3 Meeting the Objectives

Objective 1

Synchronise the frame capture between an IR and optical camera by constructing a hardware interface and then capture a series of synchronised images.

Section 3.3.3 describes the hardware interface that was developed between the IR and optical cameras which synchronises the frame capture from each of them. Chapter 5 provides examples of synchronised IR and IO images, e.g. Figure 5-6.

Objective 2

Develop a new type of calibration target which is, not only relatively simple and cheap to make, but also does not require the use of external heating and can be simultaneously detected by IR and optical cameras.

Section 4.4 describes the final calibration target design, see Figure 4-17. As has been shown in Chapter 5, the addition of LEDs to the corners of a conventional black and white chessboard calibration target allows the corner feature points to be detected by both IR and optical cameras. It also means that the target does not have to be externally heated which is the most common technique quoted in the current literature, for instance [14][23][27][28][29][31][41].

Objective 3

Use captured IR and optical images of the target to simultaneously calibrate both the IR and IO cameras to obtain their intrinsic and extrinsic parameters.

Calibration of the IR and optical cameras has been successfully carried out. The intrinsic and extrinsic parameters have been obtained and analysed. A summary of the analysis of the intrinsic and extrinsic parameters is presented in Table 5-7 and Table 5-8 respectively. The experimentally measured focal length intrinsic parameter has been found to be in agreement with ground truth values quoted by the camera manufacturers. The rotation and translation extrinsic parameters which were obtained experimentally

have been compared with the ground truth values of the geometry of the equipment set-up for the *x* axis translation, i.e. the baseline between the two cameras and the *y* axis rotation (angle of inclination between the two cameras). The baseline experimental results were found to be in agreement with the ground truth values to within ±1.7cm. There were only two experimental cases found to fall outside the ±3° range for the *y* axis rotation. These were the maximum angles of inclination in both experiments 2 and 3, see Sections 5.2 and 5.3, that is, 40° and 60° respectively. However, it is recommended that angles of inclination between the cameras should not exceed 30°, see Section 5.5.4, therefore, the 40° and 60° cases are not considered important for this medical application.

Objective 4

Successfully register captured IR and optical images of a body part by using the parameters obtained during the calibration process.

The accurate measurement and calculation of the extrinsic parameters is not only evident through their agreement with the ground truth values but also through the successful registration of the IR and optical images shown in Section 5.5 see Figure 5-8, Figure 5-10 and Figure 5-12.

Objective 5

Find the correspondence between IR and optical images of the body part to produce a fused image which combines useful IR and optical information.

Once the images have been rectified, there is only a translational difference remaining between the IR and optical images, see Section 2.1. The technique used for finding the correspondence between the rectified images has been described in Section 4.7.

Successfully fused IR and optical images can be seen in Figure 5-8, Figure 5-10 and Figure 5-12.

Objective 6

Analyse the effectiveness of various camera system arrangements:

1. *The optical axes of the cameras are parallel but placed different distances apart*

2. *The cameras are placed the same distance apart but the angle between their optical axes is varied: effectively placing the cameras around circles with different radii.*

3. *The cameras are placed at different distances apart and the angle between their optical axes is varied: effectively placing the cameras around a circle with the same radius, pointing in towards the centre.*

These three different camera system arrangements have been tested and analysed. The results of this analysis can be found in Section 5.5. In summary, it was found that all three arrangements worked well for the cases where the object was placed so that it cut across both IR and IO optical axes at about the same angle. Whilst the first two system arrangements worked well in all cases tested, for case (3), above, successful registration was not achieved in situations where the object was rotated so that the side facing each camera was significantly different and the cameras were inclined at more than 30°, as shown in Figure 5-13 and Figure 5-14. This was attributed to the fact that the working volume was too small to allow correct registration at these larger rotations. This was because at larger inclinations between the two cameras, it was difficult to capture images

of the target in both cameras through a large working volume, as discussed in Section 5.5.4.

6.4   Future Work

6.4.1 Improving the current system

Adding a timestamp to the optical images

Once Point Grey has released its new software which incorporates a timestamp with a sequence of optical image capture, the alignment software should be updated to automatically align the first two frames from the IR and optical cameras.

6.4.2 Extending the current system

In order to improve the system's medical RA assessment capabilities, it could be extended to provide 3D surface reconstruction. The work to extend the system could be carried out in stages, as described below:

1. Stereoscopic IO cameras plus one IR camera for 3D surface reconstruction

It may be sufficient for some applications to only recreate a 3D multi-spectral surface over a smaller arc rather than the full 360°. The next step to follow on from this work could be just to combine two optical cameras with one IR camera. In effect to reproduce the work carried out in [15] but using the new multi-spectral target that was tested in this project. In this case, the three cameras would be aligned with all their optical axes in parallel and with the IR camera placed between the two optical cameras. The two optical cameras would provide the depth and 3D reconstruction and the IR image information could be fused to this 3D surface.

This stereoscopic multi-spectral camera system could then be used to take different view points of a subject, either by moving it around the circumference of a circle surrounding the subject or by moving the subject itself.

2. 360° 3D surface reconstruction with one multi-spectral view

Work could progress from the previous step by adding more optical cameras so that the system configuration would now comprise one IR camera co-located with one optical camera and five other optical cameras situated evenly around the circumference of the circle which has the subject at its centre. The optical images from the six cameras would be used to obtain the 360° 3D surface reconstruction but the IR images would only be fused to one side of this surface. Thus, with this set-up, 3-D reconstruction could be obtained from the optical image information whereas the thermography analysis would be confined to one view point of the subject.

3. 360° Multi-spectral 3D surface reconstruction

The ultimate multi-spectral RA diagnostic system is one which allows full 360° multi-spectral 3D surface reconstruction with the patient's temperature distribution overlaid on optical imagery. In order to achieve such a system, these preliminary results indicate the need for at least 6 IR cameras which would increase the cost of the system, perhaps making it unattainable for most general practices. Two possible solutions to this problem are now considered.

1. Whilst the thermographic IR cameras from manufacturers such as FLIR are very expensive, GE have recently developed a cheaper, IR sensor technology using carbon nanotubes that have a resolution of 0.02°C, [52]. Therefore, a future

132

project could involve building a system with multiple IR cameras that are less costly than the FLIR A320 used here..

2. As indicated in Section 5.5.4, another possibility is to use the proposed LED based calibration target to carry out the intrinsic calibration but use a more flexible target, with only a few LEDs, for the stereo calibration part of the process which could more easily be moved throughout the required working volume. Thus, work could be done to use a target "calibration stick" with a small number of LEDs like that shown in Figure 5-16 or by using a laser pointer as described in [33].

7 References

[1]   C. Herry, M. Frize, "Digital processing techniques for the assessment of pain with infrared thermal imaging", *Proc. IEEE Conf. EMBS*, pp. 1157 − 1158, Houston, Oct., 2002.

[2]   C. Herry, M. Frize, "Quantitative assessment of pain-related thermal dysfunction through clinical digital infra-red thermal imaging", *Biomedical Engineering Online*, Vol. 3, No.19, pp.1 − 14, 2004.

[3]   M. Aubry-Frize, G.R.C. Quartey, H. Evans and D. LaPalme, "The thermographic detection of pain", *Proc. Canadian Clinical Engineering Conf.*, pp. 82-83, Saskatoon, Sep. 1981.

[4]   A.J. Collins, E.F.J. Ring, J.A. Cosh, P.A. Bacon, "Quantification of thermography in arthritis using multi-isothermal analysis", *Annals of Rheumatic Diseases*, Vol. 33, pp. 113-115, 1974.

[5]   M.D. Devereaux, G.R. Parr, D.P. Page Thomas, B.L. Hazleman, "Disease activity indexes in rheumatoid arthritis; a prospective, comparative study with thermography", *Annals of Rheumatic Diseases*, Vol. 44, pp. 434-437, 1985.

[6]   M. De Silva, V. Kyle, B. Hazelman, R. Salisbury, D.P. Page Thomas, P. Wraight, "Assessment of inflammation in the rheumatoid knee joint", *Annals of Rheumatic Diseases*, Vol. 45, pp. 277-280, 1986.

[7]   N. Ilowite, G.A. Walco, R. Pochaczevsky, "Assessment of pain in patients with juvenile rheumatoid arthritis: relation between pain intensity and degree of joint inflammation", *Annals of Rheumatic Diseases*, Vol. 51, pp. 343-346, 1992.

[8]   R.S. Salisbury, G. Parr, M. De Silva, B. Hazleman, D.P. Page-Thomas, "Heat distribution over normal and abnormal joints: thermal pattern and quantification", *Annals of Rheumatic Diseases*, Vol. 42, pp. 494-499, 1983.

[9]   Z. Zhang, "A flexible new technique for camera calibration", IEEE Trans. on pattern analysis and machine intelligence", Vol. 22, No. 11, pp. 1330 – 1334, Nov. 2000.

[10]   R.Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses", *IEEE journal robotics and automation*, Vol. RA-3, No. 4, pp. 323 – 344, Aug. 1987

[11]   G. Bradski, A. Kaebler, "*Learning OpenCV – Computer vision with the OpenCV library*", O'Reilly, 2008.

[12]   J.Y. Bouguet, "Camera calibration toolbox for Matlab", California Institute of Technology, http://www.vision.caltech.edu/bouguetj/calib_doc/, Apr. 2012.

[13]   C. Shannon, "Communication theory of secrecy systems", *Bell System Technical Journal*, Vol. 28(4), pp. 656–715, 1949.

[14]   P.K. Varshney, "Multisensor data fusion", *Electronics & Communication Engineering Journal*, Vol.9, Iss.6, pp. 245-253, Dec. 1997.

[15]   P. Askenov, I. Clark, D. Grant, A. Inman, L. Vartikovski, J-C Nebel, "3D thermography for quantification of heat generation resulting from inflammation", *Proc. 8th 3D Modelling Symposium*, Paris, 2003.

[16]   J. Zhengping, "On the multi-scale iconic representation for low-level computer vision systems", PhD thesis, The Turing Institute and the University of Strathclyde, 1988.

[17]   N. Ju, J.-C. Nebel, J.P. Siebert, "3D thermography imaging standardization technique for inflammation diagnosis", *Proc. SPIE – Infrared Components and Their Application*, Vol. 5640, pp. 266-273, Jan. 2005.

[18]   A. Cornejo, M.E. Algorri, "Construction of a frameless camera-based stereotactic neuronavigator", *Proc. IEEE Conf. EMBS*, pp. 1864 – 1867, San Francisco, 1-5 Sep. 2004.

[19]   G. Schaefer, R. Tait, S.Y. Zhu , "Overlay of thermal and visual medical images using skin detection and image registration**",** *IEEE Conf. EMBS*, pp. 965-967, Aug. 2006.

[20]   B. Abidi, S. Huq, M. Abidi, "Fusion of visual, thermal, and range as a solution to illumination and pose restrictions in face recognition", *Conf. on Security Technology*, pp. 325 330, Carnahan, Oct. 2004

[21]   D.D. Gomez, J.M. Carstensen, B.J. Ersbell, "Precise multi-spectral dermatological imaging", *IEEE Conf. Nuclear Science Symposium*, Vol.5, pp.3262-3266, Oct. 2004.

[22]   S.V. Patwardhan, A.P. Dhawan, "Multi-spectral imaging and analysis for classification of melanoma", *IEEE Conf. EMBS*, Vol.1, pp.503- 506, 1-5 Sep. 2004.

[23]   V. Hilsenstein, "Surface reconstruction of water waves using thermographic stereo imaging", *Proc. of Image and Vision Computing* , pp. 102-107, New Zealand, Nov. 2005.

[24]   J. Wang, E. Sung, "Facial feature extraction in an infrared image by proxy with a visible face image", *IEEE Trans. on Instrumentation and Measurement*, Vol. 56, No. 5, pp. 2057 – 2066, Oct. 2007.

[25]   F. Ababsa, J.Y. Didier, A. Tazi, M. Mallem, "Software architecture and calibration framework for hybrid optical IR and vision tracking system", *Proc. Mediterranean Conf. on Control and Automation*, Athens, Jul. 2007.

[26]   N.V. Boulgouris, D. Hatzinakos, K.N. Plataniotis, "Gait recognition: a challenging signal processing technology for biometric identification", *IEEE Signal Processing Magazine*, Vol.22, Iss.6, pp.78- 90, Nov. 2005.

[27]   S. Vidas, R. Lakemond, S. Denman, C. Fookes, S. Sridharan, T. Wark, "A mask-based approach for the geometric calibration of themal-infrared cameras", *IEEE Trans. on Instrumentation and Measurement*, 2012.

[28]   S. Yahyanejad, J. Misiorny, B. Rinner, "Lens distortion correction for thermal cameras to improve aerial imaging with small-scale UAVs", *IEEE Symposium on Robotic and Sensors Environments*, pp. 231 – 236, Montreal, Sep. 2011.

[29]   Y. M. Ng and R. Du, "Acquisition of 3D surface temperature distribution of a car body," *Proc. IEEE Conf. on Information Acquisition*, pp. 16–20, 2005.

[30]   H. Xiao, Y. Zhang, A. Wang, "Multi-spectral three-dimensional digital infrared thermal imaging," *Optical Engineering*, Vol. 42, pp. 906–911, 2003.

[31]   S. Y. Cheng, S. Park, M. Trivedi, "Multiperspective thermal IR and video arrays for 3D body tracking and driver activity analysis," *IEEE Conf. on Computer Vision and Pattern Recognition - Workshops*, pp. 3, Jun. 2005.

[32]   M. J. Johnson, P. Bajcsy, "Integration of thermal and visible imagery for robust foreground detection in tele-immersive spaces," *IEEE Conf. on Information Fusion*, pp.18, 2008.

[33]   T. S. Svoboda, D. Martinec, T. S. Pajdla, "A convenient Multi-Camera Self-Calibration for virtual environments", *PRESENCE: Teleoperators and Virtual Environments*, Vol. 14, No. 4, pp. 407–422, Aug. 2005.

[34]   C. Pohl, J.L. Van Genderen, "Review article Multisensor image fusion in remote sensing: Concepts, methods and applications", *International Journal of Remote Sensing*, Vol. 19, No. 5, pp. 823-854, 1998.

[35]   S. Klonus, M. Ehlers, "Performance of evaluation methods in image fusion", *Conf. on Information Fusion*, pp. 1409-1416, Seattle, July 2009.

[36]   C.S. Dima, N. Vandapel, M.Hebert, "Sensor and classifier fusion for outdoor obstacle detection: an application of data fusion to autonomous off-road navigation", *Proc. IEEE Applied Imagery Pattern Recognition Workshop*, 2003.

[37]   Y. Lu, L. Zeng, G. M. Bone, "Multisensor system for safer human-robot interaction", *Proc. IEEE Conf. on Robotics and Automation*, Barcelona, April 2005.

[38]   P. Dempster , "Upper and lower probabilities induced by multivalued mapping", *Annals of Mathematical Statistics*, Vol. 38, pp.325-339, 1967.

[39]   G. Shaefer, "*A mathematical theory of evidence*", Princeton University Press, Princeton, NJ, 1976.

[40]   E. Preston, T. Bergman, R. Gorenflo, D. Hennann, E. Kopala, T. Kuzma, L. Lazofson, R. Orkis, "Development of a field-portable imaging system for scene classification using multi-spectral data fusion algorithms", *IEEE AES Systems Magazine*, Sep. 1994.

[41]   S. Prakash, P. Y. Lee, A. Robles-Kelly, "Stereo techniques for 3D mapping of object surface temperatures", *QIRT Journal*, Vol. 4, No. 1, pp. 63 - 84, 2007.

[42]   G. Yang, Y. Yin, H. Man, "Adaptive selection of visual and infra-red image fusion rules", *IEEE Workshop on Applied Imagery Pattern Recognition*, pp. 1 – 5, Oct. 2010.

[43]   FLIR Systems, "FLIR A320 and FLIR A325 user's manual", FLIR Systems Inc., No. T559007, Apr. 2008.

[44]   Point Grey Research, "Flea 2 technical reference manual", V1.1, Sep. 2006.

[45]   Point Grey Research, "Extending the distance between a camera and the controlling host system", Point Grey Research Inc., Application Note TAN2008009, Feb. 2010.

[46]   FLIR Systems, "FLIR A320 digital input output", FLIR Systems application note, http://support.flir.com/Answers/A311/A311-Dig%20I-O.pdf, Apr. 2012.

[47]   P. Horowitz, W. Hill, "*The art of electronics*", Cambridge University Press, 1988.

[48]   Point Grey Research, "MultiSync user manual", Point Grey Research Inc., V1.0 Alpha 10, Jan. 2007.

[49]   FileZilla, "FileZilla: the free FTP solution", http://filezilla-project.org/, Apr. 2012

[50]   Intel, "Image statistics functions: image moments", http://software.intel.com/sites/products/documentation/hpc/ipp/ippi/ippi_ch11/ch11_image_moments.html, Apr. 2012.

[51]   S. Bériault, "Multi-camera system design, calibration and 3D reconstruction for markerless motion capture", Masters of Applied Science thesis, University of Ottawa, 2008.

[52]   A.D. Pris, Y. Utturkar, C. Surman, W.G. Morris, A. Vert, S. Zalyubovskiy, T. Deng, H.T. Ghiradella, R.A. Potyrailo, "Towards high-speed imaging of infrared photons with bio-inspired nanoarchitectures", *Nature Photonics*, Vol. 6, pp. 195–200, Feb. 2012.

[53]   FLIR Systems, "Basic ICD FLIR IR camera - PC", FLIR Systems application note, http://flir.custhelp.com/app/home, Apr. 2012.

8   Appendix A: Flir A320 Camera Guide

This is a guide on how to use the FLIR A320 camera which includes useful references to FLIR utilities.

8.1   Initial start up

In order to get the camera working with the FLIR *IR Monitor* and *IPconfig Utility* programs, there is a very good step-by-step process given in J. Kleissel's 'FLIR A320 Setup and Operation Instructions' guide[3], the essential points of which are emphasised here for convenience:

1. If there is more than one device to be connected to the computer via Ethernet or Firewire etc., it is better to connect the IR camera first and get it working before connecting anything else.

2. It is very important to connect the camera's Ethernet cable before connecting the power supply. Note that if necessary, the cable can be connected to the workstation via a network hub rather than via a direct connection.

3. Once the power supply is connected, run the IR Monitor program from the FLIR Systems software suite. Note that it often takes a while for the camera view to be shown in the software window, giving the appearance that the software has hung. Be patient, as it often comes up after a few minutes.

---

[3] Obtainable from J. Kleissel's University of California, San Diego website, URL:
http://maecourses.ucsd.edu/kleissl/manuals/FLIRA320SetupOperationInstructions.doc

4. If the camera does not start working, run the IPconfig utility. This usually picks up the IP address of the IR camera automatically. However, if it fails to do this, open *Network Connections* from the control panel and copy the Ethernet IP address from here to the IPconfig utility. The camera should start working now when the IR Monitor program is restarted.

8.2 Accessing the images after they have been captured

According to the FLIR documentation, it should be possible to mount the camera as a network drive on a PC. However, this does not seem to work properly. Fortunately, there are other ways to access the images captured by the camera. Figure 8-1 shows a screen shot of the available IR Monitor camera controls.
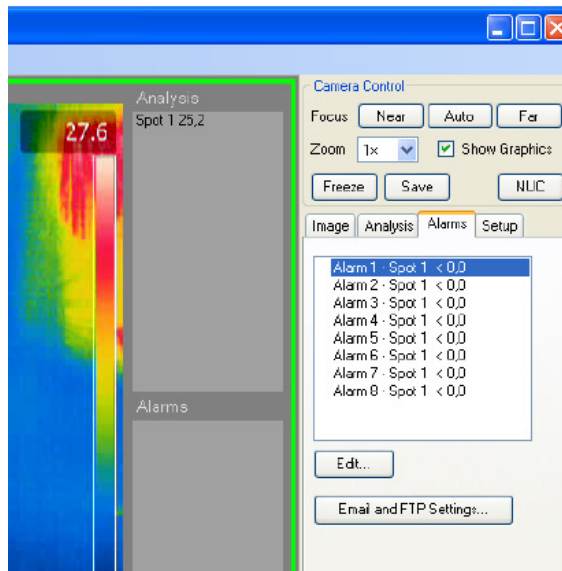
Figure 8-1: FLIR IR Monitor camera controls

8.2.1 Interactive image capture

The first way is the simplest and works well if images can be captured interactively. In this case, the IR monitor interface allows images to be captured by using the save button, as can be seen in Figure 8-1.

8.2.2 File Transfer Protocol (FTP)

The second way allows automatic download of images by using FTP to transfer files in either direction. In other words, either the user can log into the camera by using an FTP connection to the IP address of the camera or the camera can send images automatically to the user's own FTP server. In this thesis, the latter was done, using the free Filezilla FTP server software.

In order to setup an FTP connection to the camera, type *ftp*, followed by the internet address of the camera in a command window, if it is being accessed from a PC. The internet address is the one specified in the IPconfig utility. The response will be a request for the username and password. FLIR have a default username, `flir`, and password, `3vlig`. Once logged in, it is possible to use all the standard FTP commands for carrying out the image file transfer to the parent computer. Type help at the FTP prompt to get a list of possible commands. The images are usually saved in the camera's memory at `\\ipaddress\Root$\tmp\images`.

In order to setup an automatic file transfer from the camera to the parent computer using FTP, it is necessary to use the IR monitor software. Select the *Alarms* tab in the IR Monitor window, see Figure 8-1. Then pressing the Email and FTP settings button will

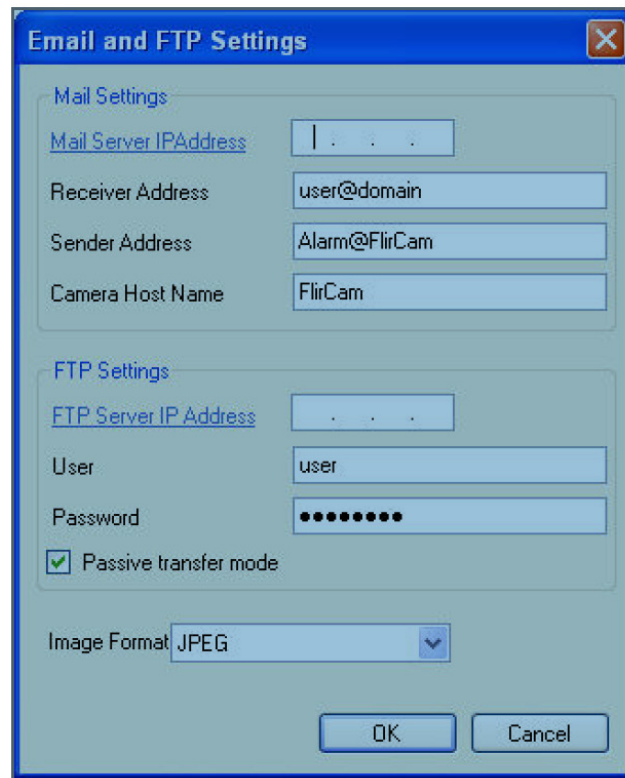bring up the screen shown in Figure 8-2 where the FTP parameters can be entered for the computer.

Figure 8-2: IR Monitor Email and FTP settings window

There are various ways of triggering the camera to save an image using the alarms settings. In the case of this project, the synchronisation hardware was used to send a trigger signal to the camera's GPIO pins which triggered an alarm. Other examples are by scheduling a regular time or setting a temperature threshold to trigger image capture, both of which can be set by using the *Edit* button, see Figure 8-1.

8.2.3 Telnet

The third way makes use of telnet to access the internal memory of the camera directly. This method also allows various settings on the camera to be set, the details of which can

be found in the interface communication document (ICD) for the camera. Although there is a A320 specific ICD, there is a generic IR camera ICD document [53] that contains a fuller set of instructions as to how to initialise the various parameters such as those described above. The two relevant ICDs for the A320 are called:

- ICD A320 Camera – PC

- Basic ICD FLIR IR Camera – PC

The second one contains useful details in the 'Camera Resource Reference Guide' that are missing from the first one. For instance, the digital input/output resources section in the second document includes the resource branches that can control how the digital input/output can be controlled via alarms, for example typing the following command at the Telnet prompt:

```
rset .power.settings.digOutFunc1 "startFrameRise"
```

configures the camera's `digital input 1` to cause an alarm trigger on the rise of the frame pulse, which was the configuration used for frame capture in this project. There are many other configurations which can be setup in this manner.

Resource monitoring constitutes another set of useful branches in the resource tree. It is possible to set parameters such as the image format and the directory where the images are to be stored here. It is also possible to disable or enable a particular alarm in this branch, e.g. typing `rset .resmon.items.1.active false` at the Telnet prompt, turns Alarm 1 (see Figure 8-1) off. See the ICD [53] for a full set of changeable parameters.