

Chapter 1

System Design and Verification

1.1 Introduction

The goal of this project is to determine the pose estimation error of a quadcopter in the outdoors. In this context, the pose is a six-dimensional position and orientation vector given by $P = [T|R]$, where the vector $T = [x \ y \ z]^T$ contains the translation dimensions and the matrix $R = [\theta \ \phi \ \psi]^T$ contains the roll, pitch and yaw dimensions respectively.

In order to determine the pose estimation error of a drone, two sets of data are required, with the first being the quadcopter's own pose estimate as well as its true pose, as given by a reading from an accurate external measurement tool. There are accurate measurement tools available for the outdoor environment, which include laser and sonar systems. However, these are fairly expensive systems and the necessary equipment was not available at the time this project was executed. It was therefore decided to investigate an alternative method of measuring a quadcopter's true pose in the outdoors.

The system that was investigated, and later implemented, is a computer vision-based system. Computer vision systems (CVS) for measurement is an attractive option: they are simple, cheap, compact and easy to operate. However, the measurement accuracy of these systems differ between platforms. Therefore, a method of reliably determining a CVS's pose measurement accuracy, must also be developed before it can be used to make pose measurements.

This chapter sets out to provide detail on the CVS. First, the design and layout of the CVS is discussed. Then, the method employed to determine the pose measurement accuracy is provided, finally followed by the results of the entire process, from testing to data processing.

1.2 Computer Vision System Layout

Introduction

The CVS has both hardware and software components, both of which are important to making accurate measurements. In this section, the design of the CVS, including the hardware and software aspects, is discussed, providing a broad overview of what the system looks like and how it functions.

System Overview

Lys requirements

The CVS is to be used to make pose measurements of a quadcopter in flight. Its measurements are to be used as accurate as possible pose data which can be used as reference pose data when comparing it to the quadcopter's on-board pose data, thereby determining the pose accuracy of a quadcopter. The CVS has to fulfil the following objectives:

- Make six-dimensional pose measurements.
- Provide measurements with an accuracy better than the quadcopter's on-board estimate.
- Make reliable, consistent pose measurements.
- Nog?? meme

Lys metingsproses

To meet these requirements, a CVS was designed and implemented. The proposed computer vision measurement system makes use of the following processes and procedures:

- Camera calibration.
- Automated two-dimensional image feature extraction.
- Pose data extraction and derivation.
- Data noise elimination.
- Nog?? Meme

Weet nie van hierdie deel nie...

Software

The CVS is computer vision system, where a computer extracts information of interest from an arbitrary image. Thus, there is a strong software aspect to the CVS, where the CVS relies on well-researched and understood computer vision techniques and algorithms. To this end, the open source and widely-used OpenCV library¹ and its sub-components was used to perform the computer vision tasks. This library was used since its free, readily available, has a wide support network on the internet and come pre-packaged with a large variety of up-to-date and powerful computer vision functions. To perform the numerical operations, the open-source NumPy² library was used.

The OpenCV library was used to extract feature data from a video stream, calibrate the CVS's camera and determine the pose of a calibration pattern by solving the principle n-points (PnP) problem. The entire process takes place as follows.

First, the camera is calibrated by following the procedure discussed in Chapter memo. Here, a 5×6 square chessboard pattern was used for calibration. To summarise the procedure, the OpenCV functions ‘findChessboardCorners()’ and ‘findConrsersSubPix()’ are used to detect and extract two-dimensional coordinate data of the corners on the chessboard pattern from a still image. These coordinates, along with their three-dimensional world coordinates, are fed to the ‘calibrateCamera()’ function to produce the camera matrix. The calibration procedure produced the camera matrix C , as presented in Equation ?? and repeated again in Equation 1.1 for convenience.

$$C = NP \quad (1.1)$$

In Equation 1.1, the matrix N is given by Equation 1.2 and matrix P by Equation 1.3.

$$N = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1.2)$$

$$P = [R|T] = \begin{bmatrix} r_{11} & r_{21} & r_{31} & t_1 \\ r_{21} & r_{22} & r_{32} & t_2 \\ r_{31} & r_{23} & r_{33} & t_3 \end{bmatrix} \quad (1.3)$$

With the calibrated camera matrices, N and P , it is now possible to perform camera pose estimation. This procedure is based on the relation given in Equation 1.4, which relates the three-dimensional world coordinate system X_w to its two-dimensional image projection x_c .

$$x_c = CX_w \quad (1.4)$$

¹OpenCV v2.4.8

²NumPy v1.8.2

Using the relation in Equation 1.4 and the camera matrix C 's intrinsic parameters, OpenCV's ‘solvePnPRansac()’ function can be employed to extract the pose matrix P of the camera relative to the calibration board. The PnP solver used is the robust PnP solver, employing Leverberg-Marquart optimisation, as described by Schweighofer and Pinz (2006). It was found that a 5×6 chessboard does not have enough features to guarantee accurate results from the ePnP method of Lepetit *et al.* (2009).

Hardware

The hardware requirements for the CVS are minimal and its setup fairly simple. To allow the software to make six-dimensional pose estimations, the camera hardware needs to capture image data at an appropriate resolution for the OpenCV algorithms to be able to detect and extract two-dimensional feature coordinate data. The camera should preferably not have a lens zoom function to keep the focal length constant throughout a measurement session. A computer running the OpenCV and data-processing scripts, as well as a calibration board, are also required.

The camera that was implemented, is a single Microsoft LifeCamHD webcam, capable of capturing 720p high definition (HD) video data at a frame rate of 30 frames per second (FPS). This camera has a variable zoom feature, however, this was controlled and set to a constant value by using the ‘uvcdynctrl’ webcam control library³, thereby keeping the focal length constant. A stereo camera setup can also be used, however, the accuracy between the single and stereo camera in most respects is rather negligible, except in the depth dimension where the stereo vastly outperforms the single camera. This is to be expected and is very akin to a person’s depth perception abilities being compromised with one eye closed. For this implementation, the single camera variant was selected. This makes the system simpler to set up and use. However, care must be taken with the depth estimation when using a single camera. In the future, the project can be expanded to a stereo camera system if it is found that the inaccurate depth estimate compromises the measurement accuracy.

The calibration board used, was a flat, A1-sized, 5×6 -square chessboard pattern calibration board. See Figure 1.4 for an image of the calibration board. This size was selected, since the squares and the board were large, affording the camera a good view, improving the data extraction performance. The large board size also allowed for a fairly large white border around the squares, as per OpenCV’s recommendation.

A laptop running Linux Mint 17.1 ‘Rebecca’ was used as a ground control station for the quadcopter in subsequent measurement tests. In addition, it was used as a recording device along with the webcam, and performed some of the data processing tasks, though a more powerful desktop PC was used to perform the image processing tasks.

³uvcdynctrl v0.2.4, bit.ly/1E4ARv3

1.3 Measurement Test Design

Introduction

Before the CVS could be used to measure the true pose of a quadcopter in flight, the accuracy of the CVS's measurements was first determined. Since the PnP solving algorithm is, at its core, an optimisation problem and produces estimates of the pose, determining the accuracy of the CVS is an important step in the system design phase.

To determine the measurement error of the CVS, a measurement test was performed in an indoor environment where another external measurement device could pose data, whose accuracy is high enough that its pose measurements can be taken as ground-truth values. The error could then be determined by comparing the CVS's measurements with that of the external measurement system's pose measurement data. Both systems were set to record the pose of a flat chessboard calibration pattern that was moved and orientated by hand.

This section describes the test layout, including the external measurement device and its details, as well as the CVS's details for the test. Then, the measurement procedure is presented, followed by the steps taken to process the data during the post-processing phase.

Test Layout

External Measurement Device Layout

The external measurement system used to record the ground-truth data that was used to compare the CVS's measurements to, is a Vicon indoor motion capture system. It is a widely-used commercial system with applications in the film, medical and sporting industries and can reach sub-millimetre accuracy in its measurements, as found by Windolf *et al.* (2008). It works by tracking a set of infrared markers stuck to a surface, with at least two infrared cameras and sophisticated proprietary motion tracking software and does this at a rate of 300Hz, or 300 FPS. The Vicon only measures the translation vectors. Therefore, to get a complete pose vector including the orientation, trigonometric relationships between the different translation vectors were used to determine the angular orientation. Given its well-documented measurement accuracy [Windolf *et al.*], the measurement results from this system was taken as ground-truth.

The Vicon system used for the test is located in the 3D Human Motion Laboratory on Stellenbosch University's Tygerberg medical campus. It consists of eight infrared cameras arranged around a square on the floor in a configuration that maximises the number of infrared markers visible to each camera at any given point in time. Figure 1.1 shows a diagram of the Vicon system layout.

Before the test commenced, the Vicon system was calibrated using a calibration procedure similar to the CVS's, but the calibration object used is a special calibration 'wand', whose dimensions are pre-programmed into the Vicon system. Then, infrared markers were placed on both the calibration

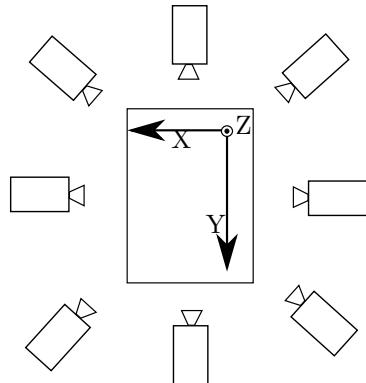


Figure 1.1: Layout of the Vicon motion capture system. Note that this is not drawn to scale.

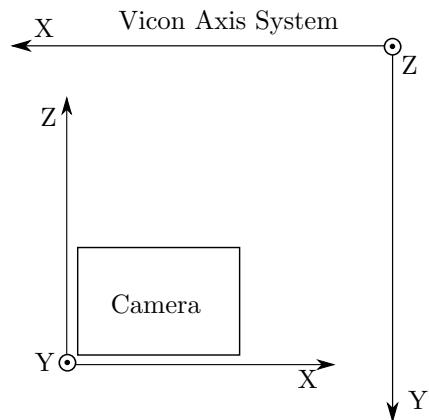


Figure 1.2: The axis orientations of the Vicon and CV systems.

board and the CVS's camera to provide pose data on both. Since the Vicon and CVS camera each have their own coordinate systems, having the position and orientation of the CVS camera available will allow the Vicon's measurements to be related back to the CVS's camera coordinate system during the data processing phase. The markers were placed in such a way that they will produce axes that more or less coincides with the Vicon's axis system, slightly reducing the work load during the data processing phase. Figure 1.2 shows the axis system's for both the CVS and Vicon systems.

Only three markers are required to be visible for the Vicon system to to

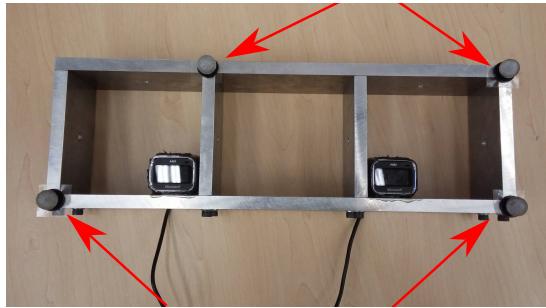


Figure 1.3: Infrared marker placement on the camera.

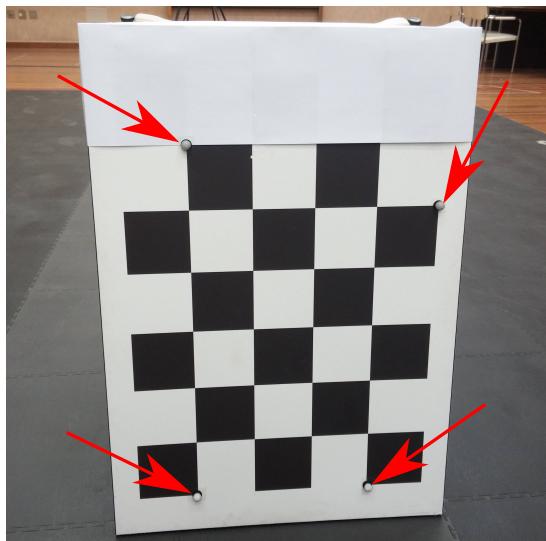


Figure 1.4: Infrared marker placement on the chessboard.

track an object's translation, thereby producing a six-dimensional pose vector, however, a fourth asymmetrical auxiliary marker was placed to provide fail-safe orientation data during the post-processing phase. Figures 1.3 and 1.4 show the marker placements for the camera and the calibration board. These markers were carefully placed in line with one another by hand, but some placement error is inevitable. This placement error offset is taken into account during the data processing phase.

Nodig?? One aspect of the Vicon system to note is that the infrared markers have some high-frequency noise associated with them, which is apparent when inspecting the raw data. This noise is inherent to the marker and can be safely filtered out with a zero-lag, second order Butterworth filter. However, the raw, unfiltered data is used throughout this project.

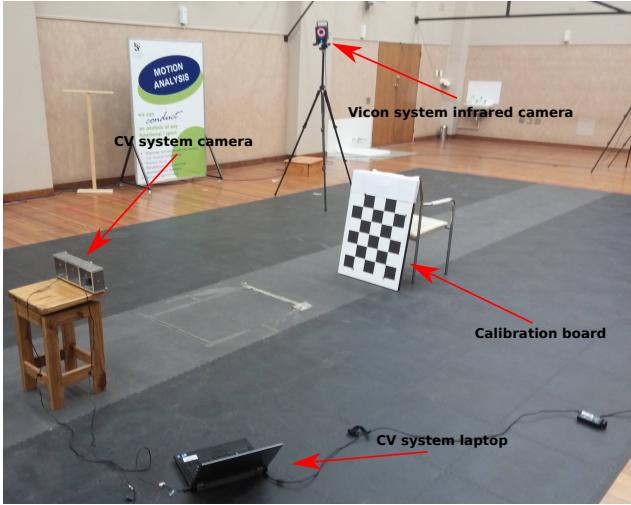


Figure 1.5: Picture of the test layout.

Computer Vision System Setup

The CVS configuration used during the test, is the exact same as the one described in Section MEME of this chapter.

Before the test, the camera was calibrated to determine its camera matrix and focal lengths in the x and y directions. Calibration was done with the same board and camera that was used during the test, against a white, well-lit background. The board was moved to different positions and orientations within view of the camera, and roughly 15 still images in standard resolution (640×480 pixels) were taken. The camera was then calibrated with this set of images and OpenCV's camera calibration module to produce a camera matrix that gives a reprojection error of approximately 0.21 pixel units. It was found that capturing data at a higher resolution did not produce significantly better results and slowed the data extraction process down.

In the test, the camera was placed in a stable aluminium frame and was left untouched throughout the test. The laptop was set to only capture video at 640×480 pixels, while zoom and autofocus was disabled to keep the camera's lens focus constant to the focal length that was found during calibration. The focal length was found to be approximately 700 pixel units in both the x and y direction. The data extraction and pose estimation took place off-line.

Test Procedure

Figure 1.5 shows a picture of the complete test setup. As described in Section 1.3, the camera is placed on a chair and the board is held facing the camera. Both of these are covered with four infrared markers with the eight infrared cameras of the Vicon system surrounding both the board and camera.

Data capture for both systems started when the Vicon system started recording. At the same time, the chessboard was tilted forward, allowing for both measurement data sets to be synchronised to a common timeframe during the data processing phase. During the data capturing phase, the board was moved by hand to different positions and orientations with respect to the CVS's camera axis system to generate measurement a wide variety of pose data vector combinations, with both the CVS and Vicon systems. To produce a diverse cloud of data vectors, the board was moved around the full field of view and at different distances to the camera, while simultaneously varying the board's rotation angles. Each video is approximately 90 seconds long, which equates to close to 2700 pose vector samples per test, though it was decided that only 2400 of these points will be used to compensate for any invalid or inconsistent readings.

The test produced two sets of measurement data: one ground-truth pose measurement data set from the Vicon system, and one data set from the CVS. These data sets allows for the determination of the CVS's measurement error by comparing the two sets of measurement. These sets were also used to provide training and validation data sets for a error prediction model, which is discussed in Chapter MEMEMEMEME.

Data Processing

Introduction

During the test, the CVS only captured video data, leaving the feature data extraction and pose estimation to be done off-line during the data processing phase.

The Vicon system data requires very little processing, since most of the data is generated in real-time and is optimised by the Vicon software. However, some work was done to fix invalid measurements that were introduced when not enough Vicon cameras had a good view of the infrared markers for a few frames. These points were corrected by means of interpolation. Furthermore, because the Vicon is a stationary system, it was convenient to move and rotate the CVS's camera and chessboard pose data to coincide with the Vicon's axis system. Furthermore, the Vicon system's data was downsampled to match the sample speed of the CVS.

Processing the CVS data involves several steps, which include simultaneously optimising the camera's focal length and the infrared marker offsets and then determining the accuracy of the CVS pose estimation system. Each of these aspects are discussed next.

Reducing Vicon Sample Speed

The Vicon system's lowest framerate is 300Hz, whereas the CVS's is 30Hz. Therefore, the Vicon's framerate had to be downsampled to math the CVS's framerate so that the two data sets could be directly related.

The downsampling involved averaging the Vicon data set in intervals of 10 samples, downsampling the Vicon by a factor of 10 and matching the CVS's framerate. This approach was taken to ensure that no data that might contain valuable information and trends, gets discarded.

Rotating the Camera and Chessboard Data

During the test, four markers were placed on the CVS camera frame to provide data on its placement within the Vicon's axis system. These markers were placed along the frame's x and y axes and coincides with the chessboard's axis system.

With the CVS's camera placement and orientation in the Vicon coordinate system known, relocating and reorientating the CVS's camera pose data to be centred around the Vicon's axis system becomes a relatively simple task. Given that the CVS's camera didn't move throughout the test, the CVS's camera pose within the Vicon system is constant and the CVS's measurements are made relative to its camera's axis system. Therefore, relocating and reorientating the camera and calibration board's pose data to coincide with the Vicon's axis system was done by simply subtracting the CVS's camera placement pose with the Vicon axis system from each pose vector of the calibration board within the Vicon axis system.

With the CVS's camera axis system now centred around the Vicon system's origin, the pose data for the chessboard acquired by the CVS and the Vicon are directly relatable to one another.

Camera Parameter Optimisation

Before the test took place, the CVS camera was calibrated using OpenCV's camera calibration toolbox. This procedure provides a fairly good estimate of the intrinsic parameters of the camera, which includes the focal lengths. However, given the lack of reference three-dimensional data, it is only an estimate of the intrinsic parameters where the accuracy of the estimate is not verifiable. Using the ground-truth pose data from the Vicon system will allow the intrinsic parameters to be determined more accurately. At the same time, the distance offset between the true infrared marker centre and the ideal infrared marker centre also needs to be taken into account. This offset is introduced by inaccurately placed infrared markers on the calibration board and camera frames and will also account for any constant measurement bias introduced by the CVS. This presents a circular optimisation problem: the focal length will affect the perceived error offset, while the error offset will affect the CVS pose estimates, which in turn affects the ideal focal length. To find the offset and optimum focal length, a dual optimisation strategy was implemented.

First, the optimisation algorithm is formulated. Suppose \mathbf{P}^* denotes the six-dimensional pose vector of the calibration board (before it is centred around the Vicon's axis system) as produced by the Vicon system, while $\bar{\mathbf{P}}$ is the constant error offset vector and the subscripts c and b represent the camera and

board respectively. ϵ refers to the error between the Vicon and CVSs' measurements that needs to be determined and $\mathbf{F}(f)$ is the pose vector measured by the CVS's camera, as a function of its focal lengths f_x and f_y . The Vicon pose and offset vectors are then given by the Equations 1.5 and 1.6.

$$\mathbf{P}^* = \mathbf{P}_b^* - \mathbf{P}_c^* \quad (1.5)$$

$$\bar{\mathbf{P}} = \bar{\mathbf{P}}_b - \bar{\mathbf{P}}_c \quad (1.6)$$

The equation for the pose estimate given by the CVS is given by Equation 1.7.

$$\mathbf{F}(f) = \mathbf{P}^* - \bar{\mathbf{P}} + \epsilon \quad (1.7)$$

Equation 1.7 can then be simplified to the form given in Equation 1.8, which forms the basis of the optimisation algorithm.

$$\mathbf{F}(f_x, f_y) = (\mathbf{P}_b^* - \bar{\mathbf{P}}_b) - (\mathbf{P}_c^* - \bar{\mathbf{P}}_c) + \epsilon \quad (1.8)$$

The next step is to determine the constant perceived offset for a given focal length. The focal length is initialised with the focal length given by the calibration procedure (approximately 700 pixel units). The offset is then determined by using Equation 1.9.

$$\sum_i \mathbf{F}_i(f) - \sum_i (\mathbf{P}_{b,i}^* - \mathbf{P}_{c,i}^*) = i\bar{\mathbf{P}} + \sum_i \epsilon_i \quad (1.9)$$

At this point, it is assumed that the error vector ϵ is normally distributed around 0, thereby eliminating it's sum and leading to Equation 1.10. This assumption will be verified in Section 1.4.

$$i\bar{\mathbf{P}} = \sum_i \mathbf{F}_i(f) - \sum_i (\mathbf{P}_{b,i}^* - \mathbf{P}_{c,i}^*) \quad (1.10)$$

Using the constant $\bar{\mathbf{P}}$ produced by Equation 1.10, it is now possible to minimise the error ϵ from Equation 1.8 by varying the focal lengths f_x and f_y . The minimum focal lengths were found by setting up a $3 \times n$ error matrix, consisting of the three translation dimensions, $[x \ y \ z]^T$, where n is the number of samples in the data set, and find the optimal focal length combination by minimising the vector's two-norm. The optimisation is based on these three dimensions since it was found that including the orientation angle dimensions added too much variability to the procedure, leading to instability. The error matrix is produced by Equation 1.11.

$$\epsilon = \mathbf{F}(f) - \mathbf{P}^* - \bar{\mathbf{P}} \quad (1.11)$$

Using the $3 \times n$ error matrix found in Equation 1.11, an error vector is generated by summing the matrix along its columns. The optimum focal lengths are

then found by taking the focal length combination that produces the smallest 2-norm of the error vector ϵ . The equation is given by Equation 1.12.

$$\min_{f_x, f_y} \left\| \sum_i \epsilon_i \right\| \quad (1.12)$$

With the optimal focal length now determined, the process starts again from Equation 1.8, finding a new offset for the new focal length. This procedure is iterated a set number of times, minimising the total error norm. After the optimum focal length combination was found, that combination was used in the camera matrix that extracted the pose data from the video data.

In summation, the algorithm is as follows:

1. Initialise focal lengths with values from the calibration procedure.
2. Repeat k times:
 - a) Find $\bar{\mathbf{P}}$ for a constant f_x and f_y (Equation 1.10).
 - b) Determine the error vector ϵ with a new $\bar{\mathbf{P}}$ (Equation 1.11).
 - c) Update focal lengths f_x and f_y with the focal length combination that minimises $\|\epsilon\|$ (Equation 1.12).
3. Use the new focal lengths to estimate the chessboard's pose.

1.4 Results

Introduction

Several sets of data were produced during the data processing phase and are discussed in this section. These include the proofs of error convergence for the optimisation procedure, as well as proofs that the error ϵ is indeed normally distributed around zero. Lastly, the new focal length combination is given and discussed, followed by the pose measurement accuracy results of the CVS.

Proof of Convergence

To prove that the optimisation procedure works as expected, minimising the error in each pose dimension, the error after each iteration of the procedure is plotted. These plots are given in Figure 1.6 and shows the errors over 50 iterations of optimisation.

The graphs in Figure 1.6 show that the error gets reduced in the x and z dimensions, while it gets worse in the y and orientation angle dimensions.

The errors in the rotation angle dimensions may be growing, since it carries no weight in the optimisation cost function. As for the growing error in the y dimension, closer inspection of the graph shows that the y dimension initially has the smallest error. It may be argued that since all three translation dimensions carry equal weight in the optimisation cost function, the x and z dimensions are being improved at the cost of the error in the y dimension.

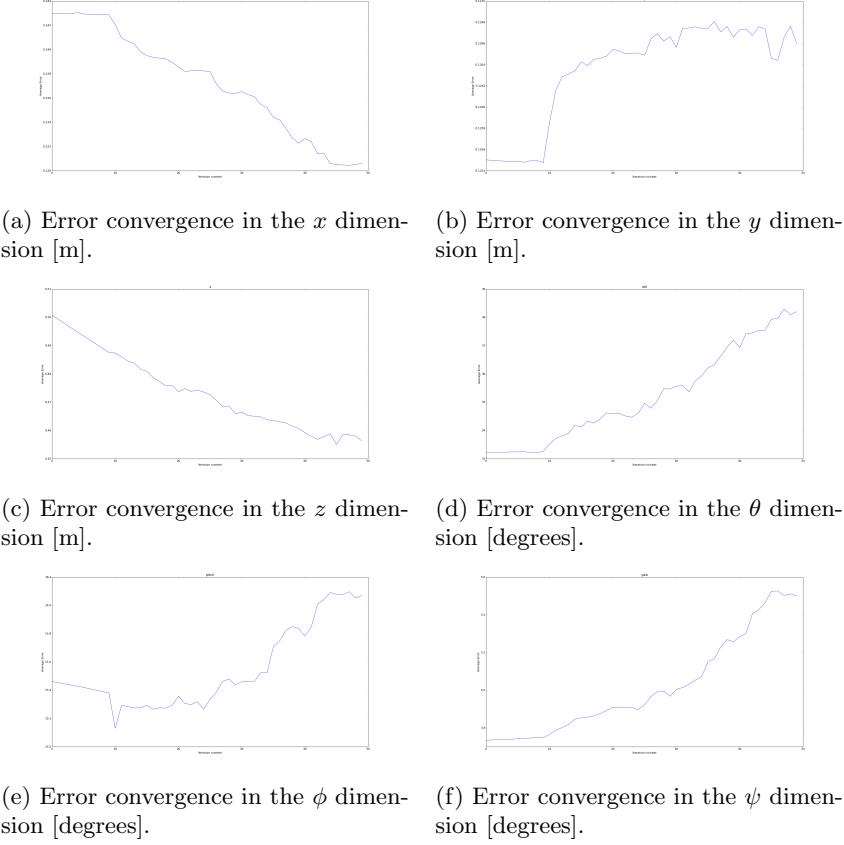
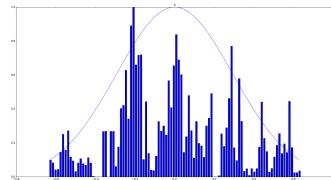


Figure 1.6: Plots showing the error in each dimension during for each iteration of the error minimisation procedure.

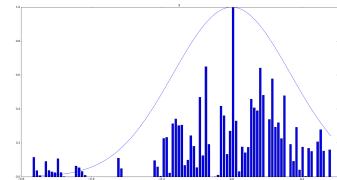
Test for Normality

To check if ϵ is indeed normally distributed as assumed in Equation 1.10, a frequency histogram of the error matrix ϵ in all six dimensions are plotted along with a normal distribution drawn using each dimension's mean and standard deviation. A χ^2 test could also have been used. However, it was found that the sample size of the data set was too large and it is known that some skewness in the data can have a large impact on the χ^2 probability estimate CITE???. Therefore, a graphical approach was taken to determining if the errors are normally distributed around 0.0.

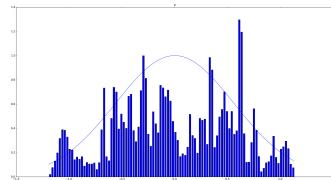
Figure 1.7 shows the frequency histogram plots of the CVS measurement errors in the six dimensions. A normal distribution, using the averages and standard deviations from the data set, are superimposed to illustrate the normal distribution of the data.



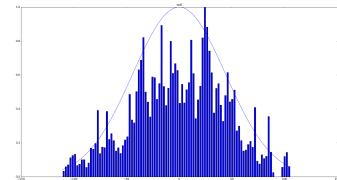
(a) Histogram of the error in the x dimension with a mean of $-5.17\mu\text{m}$ and a standard deviation of 293mm.



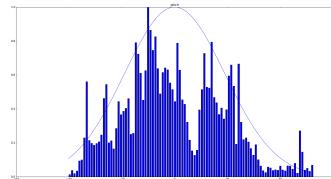
(b) Histogram of the error in the y dimension with a mean of $-20.2\mu\text{m}$ and a standard deviation of 167mm.



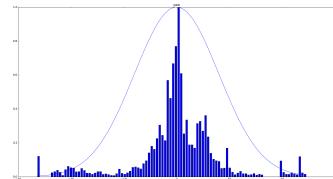
(c) Histogram of the error in the z dimension with a mean of 1.17mm and a standard deviation of 560mm.



(d) Histogram of the error in the θ dimension with a mean of 4 mili-degrees and a standard deviation of 44.0.



(e) Histogram of the error in the ϕ dimension with a mean of meme and a standard deviation of meme.



(f) Histogram of the error in the ψ dimension with a mean of meme and a standard deviation of meme.

Figure 1.7: Plots of the frequency histograms of the error data in each dimension. To demonstrate normality, the histograms are superimposed with a normal distribution plot generated from the data's mean and standard deviation.

It can be seen that all the plots are roughly centred around zero and are look very nearly normally distributed. The z dimension displays the largest standard deviation of approximately 560mm. The reason behind this is the bad depth estimates that a single camera system provides.

In all, the frequency histogram and normal distribution plots of the errors in Figure 1.7 show that the assumption made in Equation 1.10, that the errors are normally distributed around zero, is a valid one.

Optimum Focal Lengths and Offset

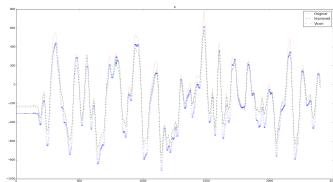
As a result of the focal length and offset optimisation procedure, the optimal focal lengths, f_x and f_y were found to be 628 and 535 after 50 iterations of the algorithm, which is considerably less than the optimum of 700 produced by OpenCV's calibration toolbox. Note that these units are given in camera pixel units and not millimetres. The optimal offset $\bar{\mathbf{P}}$ was found and is given in Equation 1.13.

$$\bar{\mathbf{P}} = [240.8\text{mm} \quad 40.09\text{mm} \quad 375.9\text{mm} \quad 184.3^\circ \quad -2.697^\circ \quad -179.9^\circ]^T \quad (1.13)$$

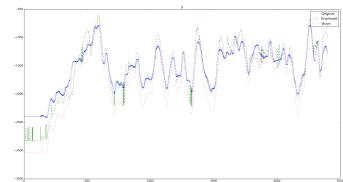
The values in $\bar{\mathbf{P}}$ contain any constant error bias, as well as marker placement errors, that may have been introduced to the Vicon test during the measurement process. The large offsets of $\pm 180^\circ$ in the θ and ψ dimensions can be explained by the differing axis orientations of the CVS camera and Vicon systems, where the axes needed rotating to coincide with one another. This indicates that the offset was correctly calculated and is working as expected. The large offset in the z dimension is to be expected, since single camera's are known to produce very bad depth estimates. However, the relatively large offset of 240mm in the x dimension is slightly surprising.

All of the above produce an error two-norm of approximately 0.81 (normalised), compared to the original 1.15 (normalised), showing an overall reduction in the error vector magnitude. Figure 1.8 shows the results of the Vicon measurements compared to the original and improved CVS measurements in all six dimensions.

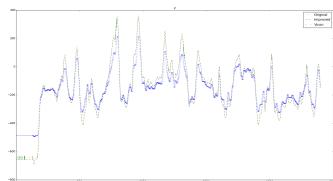
It can be seen that there is some improvement in all the dimensions. However, in some cases the improvement in one section of the data set, is negated by worse estimates in another. This can be attributed to the optimisation process, where an improvement at timeframe t_i in the x dimension, for example, may lead to a worse estimate at time t_i in the ϕ dimension. However, as the reduction in the two-norm magnitude of the error proves, there is an overall reduction in the error with the optimised data set. It can therefore be concluded that the optimisation procedure did indeed function as expected, producing a better focal length combination, which led to a more accurate pose estimate from the CVS.



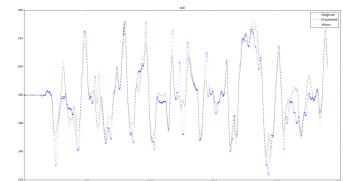
(a) The ground-truth Vicon pose estimate, versus the original and improved CV pose estimates in the x dimension.



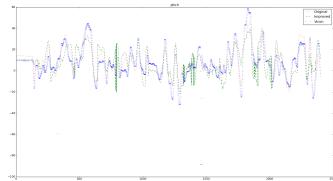
(b) The ground-truth Vicon pose estimate, versus the original and improved CV pose estimates in the y dimension.



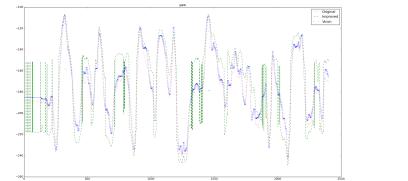
(c) The ground-truth Vicon pose estimate, versus the original and improved CV pose estimates in the z dimension.



(d) The ground-truth Vicon pose estimate, versus the original and improved CV pose estimates in the θ dimension.



(e) The ground-truth Vicon pose estimate, versus the original and improved CV pose estimates in the ϕ dimension.



(f) The ground-truth Vicon pose estimate, versus the original and improved CV pose estimates in the ψ dimension.

Figure 1.8: Plots comparing the Vicon's pose measurements with the CVS's original estimate, as well as the estimate with the optimal focal length combination.

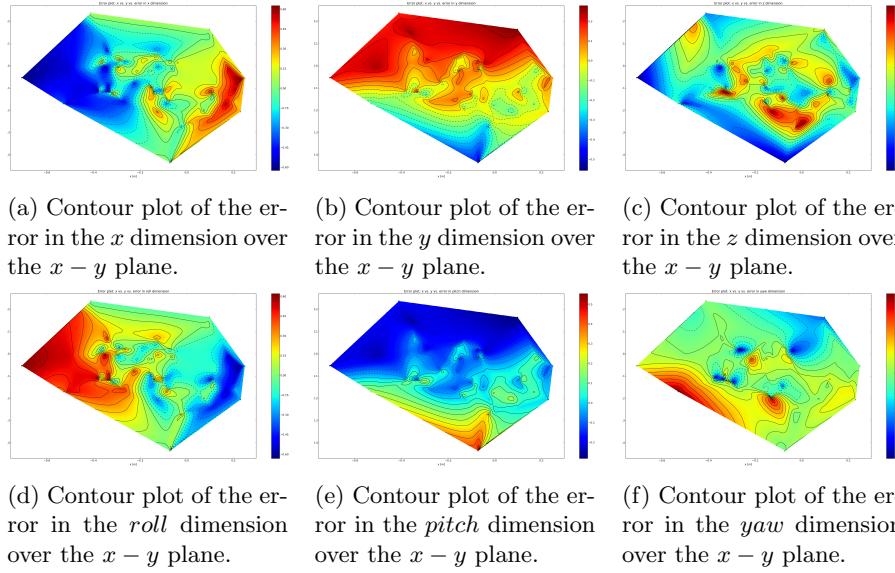


Figure 1.9: A collection of contour plots of the translation and rotation vectors relative to one another. The units for the translation figures are in metres, and in degrees for the rotation figures.

Computer Vision System Accuracy

Determining the accuracy of a multi-dimensional model is often a complex task, but since it was found that the error ϵ is normally distributed about zero, it is possible to use the covariance matrix to check the interdimensional variance and dependence. If the off-diagonal elements of the covariance matrix is sufficiently small enough relative to the diagonal elements, it can be deduced that the dimensions are strong enough independent of one another. The covariance matrix Σ is given in Equation 1.14.

$$\Sigma = \begin{bmatrix} \mathbf{3131.7} & 2255.2 & 98.227 & 94.371 & 98.830 & 106.85 \\ 2255.2 & \mathbf{40924} & 4038.2 & 197.46 & 30.631 & 1953.7 \\ 98.227 & 4038.2 & \mathbf{5592.5} & 241.75 & 106.86 & 385.23 \\ 94.371 & 197.46 & 241.75 & \mathbf{84.939} & 10.303 & 13.792 \\ 98.830 & 30.631 & 106.86 & 10.303 & \mathbf{110.54} & 63.381 \\ 106.85 & 1953.7 & 385.23 & 13.792 & 63.381 & \mathbf{318.17} \end{bmatrix} \quad (1.14)$$

The matrix Σ shows that there are large off-diagonal elements, indicating that there is strong interdimensional dependence. This dependence is further demonstrated when examining the change in the error in a dimension with respect to the other dimensions. This is demonstrated in the contour plots of Figure 1.9.

From Figure [meme](#) it can be seen that the error in a dimension varies when the two other pose dimensions changes. These plots, as well as the covariance matrix Σ indicates that there is no clear indication on the accuracy of the CVS, since its a function of the pose of the calibration board relative to the CVS's camera.

1.5 Conclusion

In this chapter, the design and layout of a computer vision measurement system (CVS) was discussed. The system consists of hardware and software components and the details of both were discussed. The system was tested in an indoor measurement facility to determine its measurement accuracy. With the ground-truth measurements produced by this indoor system, it was possible to optimise the intrinsic camera parameters and improve the pose data from the CVS.

It was shown that the error data is a function of the pose of the calibration board relative to the CVS's camera. This constant error variation makes difficult to determine the measurement accuracy of a sample pose vector from the CVS. The measurement accuracy of the CVS is very important and cannot be ignored. Therefore, another approach to determine the measurement accuracy was taken and was investigated.

Bibliography

- Lepetit, V., Moreno-Noguer, F. and Fua, P. (2009). Epnp: An accurate o(n) solution to the pnp problem. *International journal of computer vision*, vol. 81, no. 2, pp. 155–166.
- Schweighofer, G. and Pinz, A. (2006). Robust pose estimation from a planar target. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 12, pp. 2024–2030.
- Windolf, M., Gotzen, N. and Morlock, M. (2008). Systematic accuracy and precision analysis of video motion capturing systems, exemplified on the vicon-460 system. *Journal of biomechanics*, vol. 41, no. 12, pp. 2776–2780.