Proc. 2002 Australasian Conference on Robotics and Automation
Auckland, 27-29 November 2002

Problems Encountered in the Implementation of Tsai's Algorithm for Camera Calibration

Michael Tapper and **Phillip J. McKerrow** and **Jo Abrantes**
School of Information Technology and Computer Science,
Faculty of Informatics, University of Wollongong, North Wollongong 2510
{mjt09,phillip,jo}@uow.edu.au

Abstract

Camera calibration is an important step in the initialisation of many stereo vision systems. The accuracy of camera parameters has a marked effect on the data obtained from a stereo vision system. In this paper, we describe some difficulties met in an attempt to build a calibration system based on Tsai's Algorithm [Tsai, 1987]. We also offer some methods to overcome these problems. An overview of calibration, and a brief description of Tsai's Algorithm are included.

1 Introduction

If accurate camera calibration methods are used, the problem of recovering depth information from stereo image pairs is significantly simplified. Given sufficiently reliable calibration data, stereo vision can be used to determine useful three-dimensional data about scenes presented in stereo pairs.

The work discussed in this paper pertains to early stages of an attempt to develop an active stereo vision system for the stereo vision head shown in figure 1. This head has a single pan axis and a single tilt axis which is shared between the two cameras. Each camera has its own pan (or vergence) axis. The cameras shown are autofocus CCD cameras, capable of both optical and digital zoom, which can be controlled separately for each camera. Calibration is performed on one camera at a time, however both cameras must be calibrated before stereo vision algorithms can be used.

The parameters determined by the calibration process fall into two categories. One category contains the *intrinsic* parameters, which define the properties of the camera. *Extrinsic* parameters define the position and orientation of the camera within an arbitrarily defined three dimensional coordinate system.

It is possible to use stereo vision to discover useful geometric information about a scene without knowledge of the extrinsic parameters, however this will result in information that is accurate relative to an unknown scaling factor. In fact, it is possible to recover stereo information without any knowledge of the intrinsic parameters as well, subject to an unknown projective transformation. An introduction to algorithms for stereo 3D vision can be found in Trucco and Verri [1998].

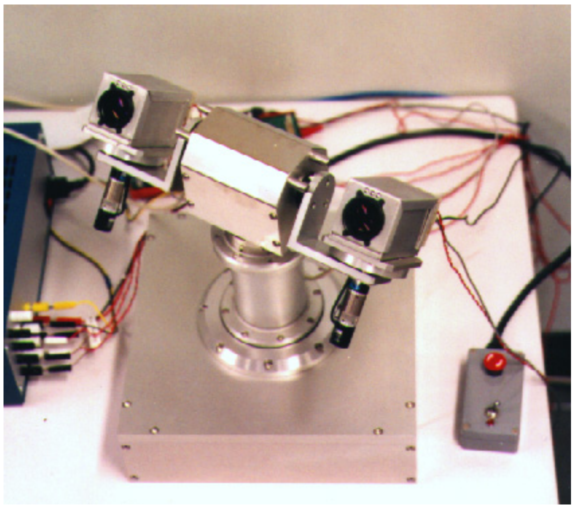

Figure 1: The active stereo vision head, with two Sony digital cameras

While it is tempting to rely on algorithms for uncalibrated stereo vision, or those calibration algorithms which provide only intrinsic parameters, there are shortcomings to this approach. Without information about the relative positions of the two cameras, stereo vision becomes less useful and more complicated. The distance between the cameras is crucial to the recovery of accurate depth information. Perhaps more importantly, the relative positions of the two cameras are necessary for rectification of the images.

Rectification is a transformation of the coordinate systems of the two cameras, such that the image planes of the cameras are made coplanar, and the scanlines in each reprojected image are parallel to the corresponding scanlines in the other image. This simplifies the problem of finding correspondences between pairs of pixels in the two images, since, in order to find a pixel in one image which corresponds to a pixel in the other, it is only necessary to search along a single scanline. Without extrinsic parameters, the information required to rectify the images must be estimated.

This paper primarily refers to the algorithm presented in Tsai [1987]. This algorithm recovers the camera parameters using direct calculation based on the relationship between points in a calibration pattern and the same points in the image plane of the camera. Many other calibration techniques exist, including techniques that rely on a mix of estimation and non-linear search, such as the algorithm described in Zhang [2000]. A popular area of research involves methods of calibration which do not require a purpose-built calibration pattern, and use multiple images taken using a moving camera as input data. Discussion of one of these techniques can be found in Faugeras et al. [1992].

1.1 The Camera Model

Tsai uses a pinhole camera model to describe the transformation of points in 3D space to pixels in the computer's frame buffer. The model contains the following parameters:

R	A 3×3 rotation matrix
T	A translation vector
f	The focal length of the camera
s_x	An uncertainty factor introduced by the image capture hardware
k_1	The radial lens distortion coefficient
(C_x, C_y)	The image centre

In addition, the system requires some information about the CCD array:

d_x and d_y	The horizontal and vertical distances between centres of adjacent cells in the CCD array
N_{cx}	The number of elements in a single row in the CCD array
N_{fx}	The number of pixels in a row in the computer's frame memory

Of these, R, T, f, s_x and k_1 are to be determined using Tsai's algorithm for calibration. C_x and C_y can be determined beforehand, and will not need to be recalibrated.

The first step in the transformation of a point in 3D space (x_w, y_w, z_w) to 2D image coordinates (x_c, y_c) is to apply R and T to convert the 3D coordinates to a coordinates system centred on the focal point of the camera, with its Z-axis passing through the image centre. This step is performed using the equation:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T$$

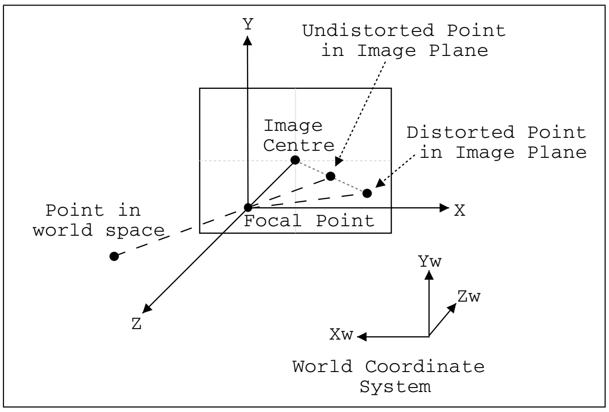

Figure 2: The Pinhole Camera Model

The next step is to transform the 3D coordinates to 2D image coordinates with radial distortion applied:

$$x_d = f\frac{x_c}{z_c} - D_x$$

and

$$y_d = f\frac{y_c}{z_c} - D_y$$

where D_x and D_y are the adjustments required due to radial distortion. These are calculated by:

$$D_x = x_d k_1 r^2$$

and

$$D_y = y_d k_1 r^2$$

where r is the distance from (x, y) to the image centre. Note that in these equations, the coordinates refer to the location of the point in the image plane after radial distortion has been applied. The radial distortion factor k_1 refers to the first coefficient in an infinite series, but Tsai states that further elements of the series have minimal effects, and may even add instability to the process.

The final step is the conversion to a location in the computer's frame buffer (x_f, y_f):

$$x_f = s_x d_x^{-1} x_d + C_x$$
$$y_f = d_y^{-1} y_d + C_y$$

1.2 Tsai's Algorithm

The algorithm given by Tsai is a two stage process designed to be performed without operator assistance. It calibrates the R, T, f, k_1 and s_x parameters from the above camera model. The algorithm executes quickly on PC hardware due to the absence of large non-linear searches.

A calibration pattern is required by this algorithm, and Tsai provides different versions for coplanar and non-coplanar calibration patterns. It is a single view algorithm, however it can be adapted to be used with multiple views of the calibration pattern.

The first stage of the process determines the extrinsic parameters s_x, R and the first two components of the translation vector, T_x and T_y. The focal length f and the z component of the translation vector T_z are also estimated at this stage. This is achieved by solving a system of linear equations whose input is the coordinates of points in the calibration pattern, both in the image and in the real world. The various parameters are then recovered from the solution to this system.

The second stage of the process involves a steepest-descent search. This is used to determine the radial distortion factor k_1 which cannot be determined from the calibration pattern. f and the T_z are also adjusted during the search.

2 Implementation of Tsai's Algorithm

We attempted to implement Tsai's algorithm using National Instruments LabView software. During this project, we learned a number of valuable lessons about various aspects of camera calibration.

2.1 Building a Calibration Pattern

If the uncertainty factor s_x is unknown, it is necessary to use the version of Tsai's Algorithm whose calibration pattern involves non-coplanar points. To this end, we built a simple, two-plane pattern using a pair of wooden boards fixed at right angles to each other. From this experiment, we learned that accurate construction is vital, and small errors in the measurement of the pattern may lead to serious errors in calibration.

Tsai's experiments with non-coplanar calibration patterns used an adjustable platform with accuracy measured in micrometers to take multiple images of the pattern varied over a distance of less than an inch. The minute distances used would indicate that errors of very small distances in measurement are sufficient to affect the accuracy of the system.

It is obviously much easier to build accurate calibration patterns when only a single plane is used. These cannot be used to determine the uncertainty factor in Tsai's Algorithm, unless multiple images are taken and combined, using a method similar to the one used by Tsai.

The distance at which the pattern can be placed without causing the detection of the calibration points to fail also appears to be a factor. In tests using an early version of the corner detector, it was necessary to place the pattern extremely close to the camera for the corner detector to work. When the corner detector was improved, the pattern could be placed as far as 1.5 metres from the camera without causing the corner detector to fail. When performed with the calibration pattern placed at this distance, some improvement was noted in the stability and accurate estimation of several camera parameters. Placement of the pattern at

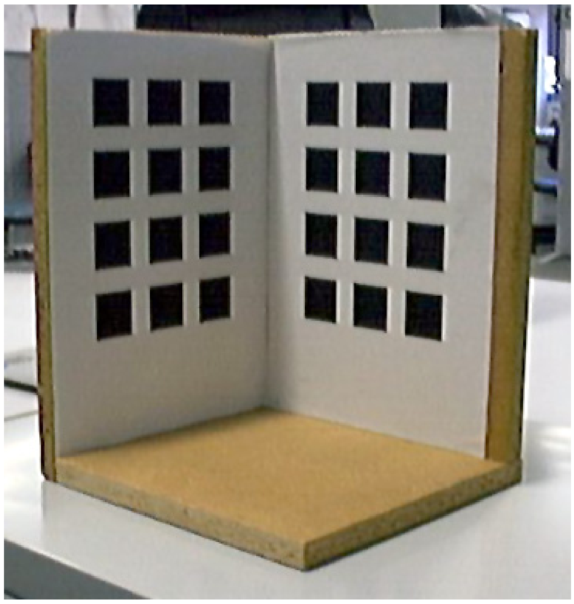

Figure 3: A calibration pattern. The corners of the black squares will be detected and used as calibration points.

this distance impeded accurate detection of the calibration points, however, as the amount of data available for the estimation of the lines in the pattern was greatly reduced. Adjusting the zoom level of the camera could eliminate this problem, however this method would limit the calibration to a reduced set of zoom levels. Instead, it may be necessary when using zoom lenses, to build a number of different calibration patterns, to be used at different zoom and focus settings.

2.2 Detecting Calibration Points

The calibration patterns used in both Tsai's paper and Trucco and Verri [1998] make use of a plane covered in a regular grid of black squares separated by white space. The corners of the squares are used as the calibration points. For this reason, an accurate corner detector is needed.

The process of corner detection involves the following steps:

1. Apply a binary threshold to the image to disambiguate the edges of the pattern.
2. Apply an edge detector. Most edge detectors should have little difficulty in finding the edges of a black-and-white image made up of squares.
3. Find the lines defining the squares using the Hough Lines Transform. We implemented the HLT from a description in Trucco and Verri [1998].
4. Intersect the lines to find corner points.

Our initial experiments with the above method achieved poor results. Despite the arrangement of the squares in a regular grid, the sides of the squares that were incident upon a single line in the pattern were not colinear in the captured image, due to perspective distortion. This resulted in multiple points being found at each corner.

In order to solve this problem, it was necessary to find a bounding box around each individual square in the pattern, which included no other squares. This could be achieved using a blob growing algorithm. The corner detector was then applied to each square individually. This eliminated the problem of multiple points being found at each corner.

It also enabled another enhancement to the corner detector. Since each square has exactly four corners, it is possible to adapt a threshold variable in the Hough Lines Transform to ensure that the number of corners returned is exactly four. In cases where the image is too indistinct for four corners to be found accurately, it is assumed that the problem rests with the lighting or position of the pattern. In most cases it was the position of the pattern that needed to be adjusted. Due to this enhancement, it was found that the algorithm was effective enough to be used under ordinary fluorescent office lights, without no direct lighting needed.

The method we used was restricted to 1 pixel accuracy, since we did not apply any of the techniques that are available for detecting edges with sub-pixel accuracy. Obviously the use of sub-pixel techniques will increase the accuracy of the detector.

2.3 The Image Centre

The image centre in the pinhole camera model is the point in the image plane at the base of the line that is perpendicular to the image plane, passing through the focal point. Tsai initially recommended using the centre of the frame buffer as a reasonable estimate. After further experimentation, he found that the image centre in modern CCD cameras often varies so widely from this estimate that accurate calibration is impossible.

A common method for finding the true image centre is given in Trucco and Verri [1998]. It involves finding the orthocentre of the vanishing points of three orthogonal sets of parallel lines. This can be performed using the edges of a cube, or lines in the two-plane calibration pattern we built. After applying this method, we assumed that the results we obtained were sufficiently accurate.

Unfortunately, it seems this assumption may not be justified. According to Willson [1993], different techniques which rely on different properties of the camera will find different image centres. That is, cameras may have different image centres depending on which property is being measured. For example, the centre of perspective projection, found using the above method can vary by over twenty pixels from the centre of radial distortion. Since Tsai's Algorithm is based on radial distortion properties of cameras, this indicates that the centre of perspective projection may not be the ideal image centre for this algorithm.

Estimation method	Coordinates
Centre of frame buffer	191, 142
Centre of perspective projection	186, 142
Laser collimation	191, 155

Table 1: Results of image centre estimation

Willson's method for finding the centre of radial distortion is not particularly useful either, as it requires first applying Tsai's Algorithm, before adjusting the image centre coordinates in a final steepest descent search. In order to use this method, a good estimate will still be needed. If Willson's experimental results may be used as a guide for other camera systems, it would appear that laser collimation is the best technique for estimating the image centre closest to the centre of radial distortion.

Laser collimation simply involves directing a laser at the lens of the camera through a through a small hole in a white screen. The camera's position is adjusted until the beam of the laser that is reflected back from the lens passes through the hole in the screen. At this point, the laser should be falling upon the image centre.

When the laser collimation method was employed, we found that the weak laser we used caused so much lens flare that it was impossible to reasonably estimate which pixel the beam was falling upon. We resorted to using the location in the image of the pinhole through which the laser had been directed. The results of

Cameras with zoom lenses present additional problems. Since the positions of the elements of the lens are changed during zoom, the image centre is unlikely to remain constant. This may also be a problem when focus is adjusted, either manually or by auto-focus systems. Willson's paper recommends methods for determining how varying the parameters of the lens affect the image centre.

2.4 The Calibration Algorithm

The calibration algorithm itself is well-documented, and implementation is straightforward. If proper attention is paid to the algorithm specification, and the required libraries for matrix operations and optimization are available, there should be no difficulties in programming the algorithm.

One important problem is that of finding values for the parameters that must be provided to the system. In particular, the d_x and d_y values, which define the horizontal and vertical distances between the centres of neighbouring sensor elements in the CCD array. This information can usually be obtained from the product documentation.

If the product information is not available, or incomplete, another method may be used. According to Willson [1995a], these values may be estimated in order to build a working system. To do this, a black square, preferably big enough to fill most of the frame, is shown to the camera. The ratio of the number of pixels along the horizontal side of the square to the number of pixels along the vertical side will provide a usable aspect ratio. It will still be necessary to obtain accurate values for these parameters at some point, or the scale of f and k_1 will be wrong.

Another problem is that of ensuring that the rotation matrix R is orthonormal. This can be achieved in a number of ways. Willson [1995b] extracts the angles of rotation from R based on the definition of R given in Tsai [1987], which describes each element in R in terms of trigonometric functions of the three angles of rotation. With these angles determined, it is simple to rewrite the matrix according to those trigonometric functions.

2.5 Limitations

Tsai designed and tested the algorithm using a 28mm fixed focal length lens. Many modern video cameras feature zoom lenses with auto-focus and automatic aperture adjustment. These lenses are far more versatile than manual focus prime lenses, however they pose problems for calibration, since few, if any of the intrinsic camera parameters can be expected to remain constant over the full range of focus, zoom and aperture settings. Tsai's algorithm provides no way to cope with this problem.

As a result, there has been much work done recently on the problem of calibrating cameras with zoom lenses. As a rule, this work has focused on determining the properties of the camera system across a range of different settings by applying algorithms such as Tsai's algorithm to a variety of configurations. Papers written on these methods include Willson [1994].

3 Summary

Good camera calibration can be an important element in a stereo vision system. Many methods exist for performing calibration, ranging from techniques that use various forms of large scale optimization to find camera parameters, to methods that do not require calibration patterns, and find the camera parameters using data obtained from camera movement. It is likely that direct calibration techniques using calibration patterns, such as Tsai's Algorithm, will offer some of the more useful methods for situations where fast and accurate calibration is desired.

Problems in implementing Tsai's Algorithm are most likely to occur in earlier stages. Accurately building and measuring the calibration pattern, finding the calibration points in the image plane and finding the image centre are likely to provide more difficulty and more sources of error than problems with the algorithm itself.

Finally, it is most important to remember that modern zoom lenses do not function in quite the way Tsai's Algorithm expects. In order to use these lenses, it is necessary to find methods of calibration that are capable of dealing with the variable parameters of the camera system.

References

[Faugeras *et al.*, 1992] Olivier D. Faugeras, Quang-Tuan Luong, and Stephen J. Maybank. Camera self-calibration: Theory and experiments. In *European Conference on Computer Vision*, pages 321–334, 1992.

[Trucco and Verri, 1998] Emanuele Trucco and Alessandro Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice-Hall, Inc., 1998.

[Tsai, 1987] R. Tsai. A versatile camera calibration technique for high-accuracy 3-D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–344, August 1987.

[Willson and Shafer, 1993] Reg. G. Willson and Steven A. Shafer. What is the Center of the Image? Technical Report CMU-CS-93-122, April 1993.

[Willson, 1994] R.G. Willson. Modeling and calibration of automated zoom lenses. In *CMU-RI-TR*, 1994.

[Willson, 1995a] Reg Willson. Tsai camera calibration FAQ. Available online at <URL:http://www.ien.it/is/vislib/README.Tsai_FAQ>, May 1995.

[Willson, 1995b] Reg Willson. Tsai camera calibration software, 1995. C code for Tsai calibration available online at <URL:http://www-2.cs.cmu.edu/afs/cs.cmu.edu/user/rgw/www/TsaiCode.html>.

[Zhang, 2000] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.