Original papers

# Calibrating cameras in an industrial produce inspection system

Andrew Wilson [a,*], Gadi Ben-Tal [a], Jamie Heather [b], Richard Oliver [a], Robert Valkenburg [b]

[a] Applied Sensors, BioEngineering Technologies, Plant & Food Research, Private Bag 3230, Waikato Mail Centre, Hamilton 3240, New Zealand
[b] Compac Sorting Equipment, PO Box 13 516, Onehunga, Auckland 1643, New Zealand

## ARTICLE INFO

## ABSTRACT

We describe a multi-camera calibration method for a produce inspection system with color and monochrome cameras. The method uses a novel spheroidal calibration target that is similar in size to the produce being graded, and features a pattern of large and small dots. This enables us to calibrate the camera system for the localized volume through which the produce moves, where human access is impractical. We describe the detection and localization of the dot centres, and the process for putting dot images into correspondence with 3D points on the target. The calibration parameters are estimated via standard bundle adjustment techniques. The method reliably gives a reprojection error RMS of approximately 0.35 px, and is fully automated. We further validate the method by measuring error in sparse reconstructions of chessboard targets and the spheroid. These objects are reconstructed with approximately 0.2 mm RMS error. Finally, we use the calibrations to build 3D models of fruit and vegetables, achieving volume estimates within 7.3 mL (2.6%) of the true volumes.

© 2017 Elsevier B.V. All rights reserved.

## 1. Introduction

This paper describes a method for calibrating multiple cameras simultaneously in a real-time fresh-produce inspection system, as part of a 3D reconstruction pipeline. The novelty of the method is using a target of similar size to the produce being graded, and using the grading conveyor system to present the target to the cameras. This means the target can be moved through hard-to-access areas, allowing us to calibrate the camera system for the localized volume traversed by the produce.

Computer vision systems are widely used in commercial produce grading machines, and the application of vision techniques to grading remains an area of active research (Zhang et al., 2014; Ma et al., 2016). These systems commonly use multiple cameras to maximize view coverage of the object surface. However, a working model of the underlying 3D geometry of the camera network is required to coherently relate observations from different cameras. Camera calibration is the process of estimating the parameters of this model, given images of a reference object.

The cameras are generally parametrized by a pinhole model, comprising extrinsic (position, orientation) and intrinsic (focal length, sensor offset, lens distortion) components. This model describes the projection of 3D points onto the sensor, allowing pixel measurements to be related to physical quantities. The cameras are used to image a *target* with known geometry and easily detectable markers. By relating the observed pixel coordinates of these markers to the location predicted by the model and target geometry, the camera parameters can be estimated.

Camera calibration has been studied extensively, and established techniques have emerged. The most popular methods are designed to calibrate a single camera, either from a single image of a 3D target such as a box corner (Tsai, 1987), or more commonly, from images of a planar target (e.g. a chessboard) in multiple distinct orientations (Zhang, 2000). Standard planar method implementations are available (Bouguet, 2016; OpenCV, 2016a), with extensions to handle stereo camera pairs.

Planar methods are readily adaptable to more general convergent multi-camera setups, provided the convergence angle is small enough for multiple cameras to view the pattern simultaneously. However, for standard chessboard targets, the entire pattern needs to be visible for it to be oriented correctly. Trying to satisfy this constraint for multiple cameras while moving the target through multiple poses is challenging. Correspondences can be found from partial views by augmenting these patterns with a few special markers; using self-identifying combinatorial markers (Fiala and Shu, 2007); or using feature descriptor-based patterns (Li et al., 2013).

Other non-planar multi-camera methods have also been developed, based on finding outline contours of a spherical target (Agrawal and Davis, 2003), or a rigid network of spheres (Shen and Hornsey, 2011). These methods require accurate segmentation

---

* Corresponding author.
  E-mail address: machinevision@plantandfood.co.nz (A. Wilson).

of target from background, and because pixel intensities often transition gradually at the object boundary, the outline is not well-defined, corresponding to a scale ambiguity in the calibration. Another method uses a laser pointer to create a virtual target (Svoboda et al., 2005).

Most reported work on calibrating industrial inspection systems uses a planar-target method (e.g. Anchini et al., 2009; Adamo et al., 2010; Molleda et al., 2010). In the produce inspection domain specifically, there has been minimal discussion of camera calibration. A few systems use off-the-shelf planar target packages to calibrate stereo pairs (Chalidabhongse et al., 2006; Font et al., 2014).

The requirement for planar targets to be imaged in multiple poses necessitates some means of presenting the target in the required range of poses, i.e. a human operator, or special hardware such as a robotic arm. The availability of such hardware is application-dependent, so this task commonly falls to operators. Layperson operators typically struggle to produce a good set of poses, lacking knowledge of what data properties are required to adequately constrain the underlying model parameters (Richardson et al., 2013).

The system hardware configuration can compound these problems. In our target platform, the 3D volume-of-interest (VOI) containing the imaged produce is enclosed by a cabinet, designed for infrequent maintenance access only (e.g. Fig. 1). Consequently, it is an ergonomic challenge for human operators to place a planar target in the confined VOI with the frequency required to maintain calibration. This makes it even less likely that a good range of poses will be obtained. Similarly, it is difficult to move the target throughout the VOI systematically, which is advisable regardless of target planarity.

Together, these factors impede the design of a repeatable calibration process. The variability in target pose distributions between data sets can lead to variability in the quality of the calibrations. This in turn makes it difficult to develop algorithms that use the calibrations. Additionally, because calibration is an offline procedure, a frequent and protracted manual process leads to significant system downtime.

However, conveyor-based produce inspection systems make affordances not available in general. The VOI is determined by the produce trajectory, and is hence well-approximated by a cylinder on top of the conveyor. This means the conveyor can be used to



**Fig. 1.** An inspection zone enclosed by a cabinet which makes the calibration volume of interest inaccessible. The produce are moved and rotated by the conveyor under the inspection system.

present a suitably shaped target automatically, with full VOI coverage.

A method based on a spheroidal target with marker lines at constant latitudes/longitudes was proposed to exploit these circumstances (Heather, 2014). The camera parameters are estimated by minimizing the error between the observed and predicted target outlines and marker lines. This approach is appealing because it gives data uniformly distributed over the full VOI, while remaining fast and user-friendly.

However, the use of lines as the geometric primitives leads to ambiguities in the process. The two detected boundaries of the line must be resolved to a single line, and since there is a one degree-of-freedom ambiguity along each line, a sampling must be defined for comparisons between lines. Minimizing outline error also leads to the same scale ambiguity described above. Finally, the algorithm lacks a simple extension for lens distortion, as it models the target outline with an ellipse.

In this work, we present an improved multi-camera calibration method for our produce inspection system. The new method is point-based, using a spheroidal target with a novel dot pattern. The method also incorporates a radial distortion model. Due to the geometry of the cameras, VOI, and target, our problem has characteristics that differentiate it from most other camera calibration work. Part of the work described here is addressing the issues that arise as a result. In particular, the various considerations associated with selecting components of the camera model (such as radial distortion and principal point) are described in Section 3.3.

Firstly, the long narrow VOI projects to each image as a horizontal strip, spanning only the central $\frac{1}{4}$ of the image height (shown schematically for a single camera in Fig. 2). This lack of data in large regions of the images influences model selection. Secondly, the target geometry impacts on the geometric constraints exerted by the data. Calibration targets exert *rigidity constraints* during parameter estimation, arising from the fact that the target points move as a rigid body. In our case, the target needs to be sized on the same order as the produce, so it can be held by the existing conveyor rollers. Using a produce-sized target means these rigidity constraints are only exerted over small localities of the VOI. This should not pose a problem as we aim to use the calibrations for localized reconstructions of produce. Nonetheless, we also investigate the effects when reconstructing larger-scale objects.

In Section 2, we describe the system hardware and geometry, and the design of the calibration target. The algorithm for calibrating the cameras is described in Section 3. In Section 4 we explain the methods of evaluating the calibration quality, report the results from these evaluation methods, and discuss the results and limitations of our approach. The evaluations are based on model fit quality and an accuracy measurement based on sparse 3D reconstruction of validation targets. Finally, in Section 5 our conclusions are given.

## 2. Materials and hardware

### 2.1. Calibration target design

The calibration target is a white spheroidal ball with black dots arranged in a regular latitude-longitude grid, shown from several views in Fig. 3. The spheroidal shape ensures that the target is a good fit for the conveyor rollers, and rotates around a stable axis. Each dot centre uniquely defines a 3D point.

Dots can be problematic as markers because their centroid projection is biased by lens and perspective distortion (Mallon and Whelan, 2007). Chessboard corners are a common alternative, and are not subject to these biases, but it is more difficult to imprint them to the required precision on a non-planar target. In