A Novel Machine Vision Algorithm for a Fast Response Quality Control System

Yonghuai Liu and Marcos A Rodrigues
Department of Computer Science
The University of Hull, Hull, HU6 7RX, UK
{*Y.Liu, M.A.Rodrigues*}@dcs.hull.ac.uk
http://www2.dcs.hull.ac.uk/aise/aise.html

Abstract

We are investigating the design of machine vision algorithms for real time inspection of filter components for quality assurance. Filter components are rigid 3D objects with predefined geometry so that a good deal of knowledge can be incorporated into the system design. However, while the objective is to reason about 3D structure, current machine vision techniques only allow the acquisition of 2D object descriptions and this is normally referred to as a 3D-2D correspondence problem. We have investigated geometrical techniques in 2D and in 3D and have developed a novel method to analyse rigid body transformations that have been successfully applied to machine vision calibration problems. In this paper, we first describe a geometrical analysis of image correspondence applied to 2D rigid body transformations. We then develop a novel calibration algorithm that, given a 3D model of an object and a set of 3D-2D image correspondence points allows the calibration of all transformation parameters of interest. The calibrated parameters can then be used to verify the physical dimensions of the object under inspection. For a comparative analysis, we also develop a calibration algorithm based on epipolar geometry applied to the same task. Experimental results have shown that our novel algorithm performs much better than the algorithm based on epipolar geometry and that it is well suited to the real time requirements of the task.

1 Introduction

The research reported in this paper is concerned with an industrial application of 3D machine vision to quality inspection. We are investigating the design of a real time, automatic vision system for inspection of filter components in a manufacturing line. Several types of filter are manufactured by our industrial collaborator, including filters for petrol and diesel engines, pollution control, and other industrial applications. Variations in size and shape for each kind of filter are significant aspects to be incorporated into the system design. The main parameters of interest, which vary widely for different kinds of filter, include physical dimensions such as width, height, depth, and diameter. The task requirements in terms of speed and accuracy are demanding: production levels stand at around 4 million filter units per year where each filter – a 3D object – must be inspected at the production line in just 1.8 seconds, and that includes setting-up time, image acquisition, processing, and flagging out of defective components. Typical accuracies are described in terms of deviation from standard values which are $\pm 0.5\%$ of the expected value for any measurement. While some parameters can be rigorously controlled, others are more problematic such as thermal properties of metallic heating plates. These plates are used to bake a polystyrene compound whose final configuration depend on the baking temperature, amount of compound dispensed, and baking time.

Through analysing the essence of the problem, given that our aim is to verify conformance to specifications in terms of filter dimensions within stringent performance requirements, a good deal of knowledge can be incorporated into the system. For instance, filters can be mechanically driven to predefined positions in a reference coordinate frame and a camera would acquire images from a known, suitable place. Once the camera has been installed, the theoretical orientation and position of the camera in the reference coordinate frame together with the depth of each visible point on the filter in the camera centred coordinate frame can be exactly determined, as we have full knowledge of the 3D structure of an ideal filter. Thus, a real time automatic inspection of filter production would only require the acquisition of 2D image data. This reduces the problem to the determination of which points on the 2D image correspond to points on the 3D filter in the reference coordinate frame. This is called a 3D-2D correspondence problem and landmarks can normally be used to determine such correspondences.

Many algorithms have been proposed to solve 3D-2D problems, such as techniques based on conservation of distance [2], triangular geometry [9], [12], iterative least

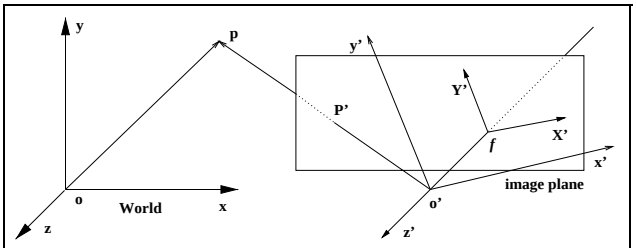

Figure 1: 3D-2D problem: given all information about a scene ($\mathbf{p}$) in $\mathbf{oxyz}$ coordinate frame and an image ($\mathbf{P'}$) of this scene taken from another viewpoint $\mathbf{o'}$, how to estimate the orientation and position of the camera at $\mathbf{o'}$ in $\mathbf{oxyz}$ coordinate frame and the structural information of this scene in camera centred coordinate frame $\mathbf{o'x'y'z'}$? In the figure, f represents the focal length of the camera, and $(\mathbf{p}, \mathbf{P'})$ is called a 3D-2D correspondence.

squares method [6], [10], [11], and iterative method [13], [14], [15], [16]. However, such algorithms do not explicitly use distance and angular information as constraints to calibrate the orientation and position parameters of the camera even though it seems that such information will increase the performance of the algorithms. Second, these algorithms are normally based on iterative methods which have a disadvantage of depending on the initial value of unknowns and that the best solution cannot be guaranteed to be found.

We have analysed geometrical properties of correspondence vectors synthesised into a single coordinate frame. These are described in Section 2. Based on this geometrical analysis, we designed the system layout with the geometry as depicted in Figure 1. We align the optical axis of the camera so that it is parallel to the $\mathbf{z}$ axis of the reference coordinate frame and coincides with the focal point of the camera in the $\mathbf{xy}$ plane of reference. As a result, the image plane will be parallel to the $\mathbf{xy}$ plane of the reference. We then developed a novel algorithm called Simplified Geometrical Algorithm (SGA) to estimate the parameters of interest including the orientation $\hat{\mathbf{R}}$ and position $\hat{\mathbf{t}}$ of the camera, and the filter structural data $\hat{\mathbf{z}}'_{\mathbf{i}}$ describing the depth of each visible point in camera centred coordinate frame. The SGA algorithm provides the closed form solutions to all calibrated parameters making full use of distance and angle information. For a comparative study of the performance of the SGA algorithm, we also developed an algorithm based on the epipolar geometry called Simplified Epipolar Geometry Algorithm (SEGA). A comparative analysis is presented in Section 4.

The rest of this paper is organised as follows. The analysis of 2D correspondence vectors is described in Section 2, the novel SGA algorithm to estimate the orientation and position parameters of the camera and structural data of the filter in the camera centred coordinate frame is described in Section 3, and the validation of the SGA algorithm is described in Section 4. Finally, some conclusions are drawn in Section 5.

2 Analysis of 2D Correspondence Vectors

A survey of machine vision literature indicates that significantly less research effort has been spent into the calibration of rigid body transformation parameters given two sets of 2D correspondences in comparison with calibration given two sets of 3D correspondences (e.g. [7], [8], [17]). Given that 2D is a special case of 3D and the fact that we normally acquire 2D images to reason about 3D, there is scope to further investigate the 2D case aiming at determining and extending calibration methods and concepts from 2D to 3D. Therefore, in this paper we concentrate on the study of geometrical properties of 2D image correspondences which can be used to calibrate the orientation and position of the camera and on the use of such properties to determine the structural parameters of 3D objects. The starting point for our analysis is that we mainly consider the relationships between 2D correspondences that have been synthesised into a single coordinate frame. The general assumptions and constraints of our analysis method are given as follows:

1. all transformations must be rigid body transformations;
2. all correspondences must undergo the same rigid body transformation;
3. all correspondences must be synthesised into a single coordinate frame;
4. because the field of view cannot exceed 180 degrees, the rotation angle θ of the rigid body transformation must be defined as: $0 < \theta < \pi$. As a result, the rotation angle θ can be uniquely determined by its cosine value.

Generally speaking, the relationship between a point $\mathbf{p}$ described in one coordinate frame and its corresponding point $\mathbf{p'}$ described in another coordinate frame after a rigid body transformation can be represented as follows:

$$\mathbf{p'} = \mathbf{R}(\mathbf{p} - \mathbf{t}) \tag{1}$$

where $\mathbf{R}$ represents the rigid body rotation matrix (corresponding to an orthonormal matrix with determinant equal to 1), $\mathbf{t}$ represents the translation vector of the rigid body transformation and $(\mathbf{p}, \mathbf{p'})$ is called a correspondence.

The geometrical considerations and formalisation of our method are described as follows. In 2D, synthesise a set of feature points and their correspondences $(\mathbf{p_i}, \mathbf{p'_i})$ $(i = 1, 2, \cdots, n)$ into a single coordinate frame. It is verified that the perpendicular bisectors of correspondence vectors $\mathbf{CV_i} = \mathbf{p_i} - \mathbf{p'_i}$ $(i = 1, 2, \cdots, n)$ all intercept at a fixed

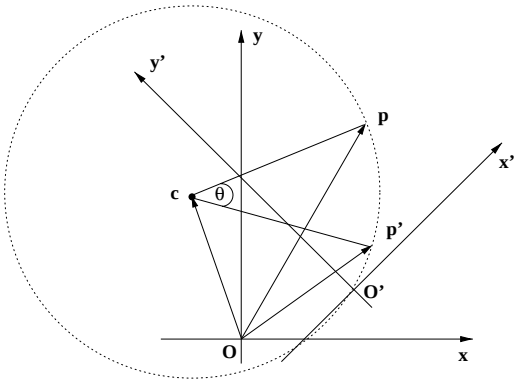

Figure 2: After the rigid body transformation, the points $\mathbf{p}$ and $\mathbf{p'}$ lie on the circle centred at the point $\mathbf{c}$ with the including angle between vectors $\mathbf{p}-\mathbf{c}$ and $\mathbf{p'}-\mathbf{c}$ equal to the rotation angle of the transformation.

point $\mathbf{c}$. Plotting the translation vector $\mathbf{t}$ at the origin of the coordinate frame, it is verified that its perpendicular bisector also intercepts at the same fixed point and that the including angle between the lines passing through any correspondence $(\mathbf{p_i}, \mathbf{p_i'})$ and the fixed point is equal to the rotation angle θ of the transformation. From appearance, it seems that the direction of rotation from vector $\mathbf{p_i}$ to vector $\mathbf{p_i'}$ around this fixed point is opposite to that of the rigid body transformation. The above properties as depicted in Figure 2 are formalising as follows:

Property 1 *There is one and only one point* $\mathbf{c}$ *in 2D which is uniquely determined by the rigid body transformation equidistant to any correspondence* $(\mathbf{p}, \mathbf{p'})$ *subject to the same rigid body transformation and the including angle between vectors* $\mathbf{p}-\mathbf{c}$ *and* $\mathbf{p'}-\mathbf{c}$ *is equal to the rotation angle* θ *of the transformation.*

Proof of Sufficiency: If the transformation $(\mathbf{R}, \mathbf{t})$ is known, then the critical point $\mathbf{c}$ can be constructed as:

$$\mathbf{c} = (\mathbf{I} - \mathbf{R^T})^{-1}\mathbf{t}$$

which is equivalent to:

$$\mathbf{c} = \mathbf{t} + \mathbf{R^T}\mathbf{c}$$

where $\mathbf{I}$ is a 2D identity matrix, the superscript $\mathbf{T}$ means transpose. Thus, we have:

$$\begin{aligned} ||\mathbf{p'} - \mathbf{c}|| &= ||\mathbf{R}(\mathbf{p}-\mathbf{t}) - \mathbf{c}|| \\ &= ||\mathbf{p} - \mathbf{t} - \mathbf{R^T}\mathbf{c}|| = ||\mathbf{p} - \mathbf{c}|| \end{aligned} \quad (2)$$

and

$$\begin{aligned} \frac{(\mathbf{p}-\mathbf{c})^\mathbf{T}(\mathbf{p'}-\mathbf{c})}{(\mathbf{p}-\mathbf{c})^\mathbf{T}(\mathbf{p}-\mathbf{c})} &= \frac{(\mathbf{p}-\mathbf{c})^\mathbf{T}(\mathbf{R}(\mathbf{p}-\mathbf{t})-\mathbf{c})}{(\mathbf{p}-\mathbf{c})^\mathbf{T}(\mathbf{p}-\mathbf{c})} \\ &= \frac{(\mathbf{p}-\mathbf{c})^\mathbf{T}\mathbf{R}(\mathbf{p}-\mathbf{c})}{(\mathbf{p}-\mathbf{c})^\mathbf{T}(\mathbf{p}-\mathbf{c})} \\ &= \cos\theta \end{aligned} \quad (3)$$

■

Proof of Necessity: If there is another point $\mathbf{c'}$ satisfying Equations 2 and 3, then we have $(\mathbf{p}-\mathbf{p'})^\mathbf{T}(\mathbf{c}-\mathbf{c'}) = \mathbf{0}$. Because $(\mathbf{p}, \mathbf{p'})$ is an arbitrary $\mathbf{CV}$, therefore, $\mathbf{c} = \mathbf{c'}$. ■

Because we are interested in analysing geometrical properties from sets of image correspondences $(\mathbf{p_i}, \mathbf{p_i'})$, we must look at the problem from such perspective. From image correspondence, the point $\mathbf{c}$ can be determined which is then used to calibrate the rotation angle θ of the transformation, the rotation matrix $\mathbf{R}$, and the translation vector $\mathbf{t}$. Thus, given two non-parallel correspondence vectors $\mathbf{CV_1}$ and $\mathbf{CV_2}$, their perpendicular bisectors will intersept at the point $\mathbf{c}$ which can be estimated by:

$$\mathbf{c} = \begin{pmatrix} (\mathbf{p_1}-\mathbf{p_1'})^\mathbf{T} \\ (\mathbf{p_2}-\mathbf{p_2'})^\mathbf{T} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\mathbf{p_1^T p_1} - \mathbf{p_1'^T p_1'}}{2} \\ \frac{\mathbf{p_2^T p_2} - \mathbf{p_2'^T p_2'}}{2} \end{pmatrix} \quad (4)$$

If the two correspondence vectors are parallel then the two perpendicular bisectors of $\mathbf{CV_1}$ and $\mathbf{CV_2}$ will coincide and the point $\mathbf{c}$ is undetermined. Once the critical point is known, the rotation angle θ can be estimated by:

$$\cos\theta = \frac{(\mathbf{p}-\mathbf{c})^\mathbf{T}(\mathbf{p'}-\mathbf{c})}{(\mathbf{p}-\mathbf{c})^\mathbf{T}(\mathbf{p}-\mathbf{c})} \quad (5)$$

As a result, the rotation matrix $\mathbf{R}$ of the transformation is uniquely determined by:

$$\mathbf{R} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad (6)$$

Finally, the translation vector can be represented as a correspondence $\mathbf{CV_0} = \mathbf{t} - \mathbf{0}$, which can be estimated by:

$$\mathbf{t} = (\mathbf{I} - \mathbf{R^T})\mathbf{c} \quad (7)$$

The above analysis shows that, from image correspondences $(\mathbf{p_i}, \mathbf{p_i'})$ we have developed a set of explicit expressions making full use of feature vectors and angular information. These equations are useful to the estimation of rigid body transformation parameters (e.g. $\theta, \mathbf{R}, \mathbf{t}$,) and to the design and analysis of a setup for the real time system for automatic inspection of filter components under investigation.

3 Description of the Algorithm

In this Section we describe an algorithm that has been developed taking in consideration the special camera setup as highlighted in Section 1. In the setup, a reference coordinate frame $\mathbf{oxyz}$ describes the ideal position $(\mathbf{p})$ of the filter at a fixed place in the production line. From this setup, it is

known that the camera can only move in the $\mathbf{xy}$ plane of the reference coordinate frame $\mathbf{oxyz}$ and rotate around the optical axis $\mathbf{h} = (\mathbf{0}, \mathbf{0}, \mathbf{1})^{\mathbf{T}}$. We can describe the camera's 3D orientation and position $(\mathbf{R}, \mathbf{t})$ in relation to the reference coordinate frame $\mathbf{oxyz}$ as:

$$\mathbf{R} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{t} = \begin{pmatrix} t_x \\ t_y \\ 0 \end{pmatrix} \tag{8}$$

A 2D image or picture $(\mathbf{P'})$ is acquired through the camera for each filter under inspection. Generally, the relationship between the point $\mathbf{p} = (\mathbf{x}, \mathbf{y}, \mathbf{z})^{\mathbf{T}}$ in the $\mathbf{oxyz}$ coordinate frame and the image point $\mathbf{P'} = (\mathbf{X'}, \mathbf{Y'})$ can be represented as:

$$\mathbf{z'} \begin{pmatrix} \mathbf{P'} \\ 1 \end{pmatrix} = \mathbf{R}(\mathbf{p} - \mathbf{t}) \tag{9}$$

where $\begin{pmatrix} \mathbf{P'} \\ 1 \end{pmatrix}$ represent the homogeneous coordinate of image point $\mathbf{P'}$, $\mathbf{p}, \mathbf{R}$, and $\mathbf{t}$ are as defined in Equation 1, and $\mathbf{z'}$ represents the depth of point $\mathbf{p}$ in camera centred coordinate frame. We assume a unit focal length of the camera $(f = 1)$ for convenience of computation. Equation 9 is equivalent to:

$$\mathbf{zP'} = \mathbf{R}(\begin{pmatrix} x \\ y \end{pmatrix} - \mathbf{t}) \tag{10}$$

where $\mathbf{R}$ is a 2D rotation matrix and $\mathbf{t}$ is a 2D translation vector. Obviously, Equation 10 represents a 2D rigid body transformation where the 2D correspondence $(\mathbf{p}, \mathbf{p'}) = (\begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{zP'})$. Once a number of 3D-2D correspondences are known (through landmark tracking, for instance) their counterpart 2D correspondences are uniquely determined. Thus, our aim here is to estimate the transformation parameters and the 3D structure of the filter given a number n (usually, $n \geq 3$) of 3D-2D correspondences. Making full use of the geometrical properties of 2D image correspondences $(\mathbf{p_i}, \mathbf{p'_i})$ as described in Section 2, we propose a Simplified Geometrical Algorithm (SGA) to estimate the parameters of interest rotation angle, rotation matrix, translation vector, and depth of each feature point $(\hat{\theta}, \hat{\mathbf{R}}, \hat{\mathbf{t}}, \hat{\mathbf{z}}'_\mathbf{i})$, where the hat means estimated. The steps in the algorithm are described as follows.

1. Estimate the point $\hat{\mathbf{c}}$ as:

$$\hat{\mathbf{c}} = (\mathbf{A^T A})^{-1} \mathbf{A^T b}$$

where

$$\mathbf{A} = \begin{pmatrix} (\mathbf{p_1} - \mathbf{p'_1})^{\mathbf{T}} \\ (\mathbf{p_2} - \mathbf{p'_2})^{\mathbf{T}} \\ \vdots \\ (\mathbf{p_n} - \mathbf{p'_n})^{\mathbf{T}} \end{pmatrix}$$

and

$$\mathbf{b} = \begin{pmatrix} (\mathbf{p_1^T p_1} - \mathbf{p_1'^T p'_1})/\mathbf{2} \\ (\mathbf{p_2^T p_2} - \mathbf{p_2'^T p'_2})/\mathbf{2} \\ \vdots \\ (\mathbf{p_n^T p_n} - \mathbf{p_n'^T p'_n})/\mathbf{2} \end{pmatrix}$$

2. Using Equation 5, estimate a set of the cosines a_i of rotation angles corresponding to $(\mathbf{p_i}, \mathbf{p'_i})$. In order to eliminate the effects of noise, the finally calibrated rotation angle $\hat{\theta}$ is estimated by a band pass filter as:

$$\hat{\theta} = \arccos\{average\ of\ band\ pass\ filtered\ a_i\}.$$

3. The 2D rotation matrix $\hat{\mathbf{R}}$ can be estimated by Equation 6.

4. The 2D translation vector $\hat{\mathbf{t}}$ of the camera can be estimated by Equation 7.

5. The depth $\hat{\mathbf{z}}'_\mathbf{i}$ of each point in camera centred coordinate frame can be estimated by:

$$\hat{\mathbf{z}}'_\mathbf{i} = (\mathbf{R_1}(\mathbf{p_i} - \hat{\mathbf{t}})/\mathbf{X'_i} + \mathbf{R_2}(\mathbf{p_i} - \hat{\mathbf{t}})/\mathbf{Y'_i})/\mathbf{2} \tag{11}$$

where $\mathbf{R_1}$ and $\mathbf{R_2}$ are the row vectors of rotation matrix $\hat{\mathbf{R}}$.

For the purpose of comparative analysis, we also developed an algorithm based on the epipolar geometry for the adopted special camera setup. This algorithm is called Simplified Epipolar Geometry Algorithm (SEGA). First, from [4], [3], it is known that the essential matrix $\mathbf{E}$ can be estimated by $\mathbf{E} = \mathbf{TR}$ where $\mathbf{T} = \begin{pmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix}$. Therefore, $\mathbf{E}$ =$\mathbf{TR}$

$$= \begin{pmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & t_y \\ 0 & 0 & -t_x \\ -t_y\cos\theta - t_x\sin\theta & -t_y\sin\theta + t_x\cos\theta & 0 \end{pmatrix}$$

Thus, $\mathbf{e} = (e_1, e_2, e_3)$, can be estimated by:

$$\mathbf{e} = (\mathbf{A^T A})^{-1} \mathbf{A^T b}$$

where $\mathbf{A} = \begin{pmatrix} x_1 & X'_1 z_1 & Y'_1 z_1 \\ x_2 & X'_2 z_2 & Y'_2 z_2 \\ \vdots & \vdots & \vdots \\ x_n & X'_n z_n & Y'_n z_n \end{pmatrix}$

and $\mathbf{b} = (-y_1, -y_2, \cdots, -y_n)^T$.

Based on the above definitions, the steps in the SEGA algorithm are described as follows:

1. The rotation angle $\hat{\theta}$ can be estimated by:

$$\hat{\theta} = \arccos\hat{\cos}\,\theta$$

where

$$\begin{pmatrix} \hat{\sin}\,\theta \\ \hat{\cos}\,\theta \end{pmatrix} = \begin{pmatrix} e_3 & -e_2 \\ -e_2 & -e_3 \end{pmatrix}^{-1} \begin{pmatrix} e_1 \\ 1 \end{pmatrix}$$

2. The 2D rotation matrix $\hat{\mathbf{R}}$ can be estimated using Equation 6.

3. The 2D translation vector $\dot{\mathbf{t}} = (t_x, t_y)^T$ up to a scale factor can be estimated by:

$$\dot{\mathbf{t}} = \hat{\mathbf{R}}^{\mathbf{T}} \begin{pmatrix} -e_3 \\ e_2 \end{pmatrix}$$

4. Use the method below to estimate sets of scale factors $\beta_{x,i}$ and $\beta_{y,i}$:

$$\beta_{\mathbf{x,i}} = \frac{\mathbf{x_i} - \mathbf{z_i}\hat{\mathbf{R}}^{\mathbf{T}}\mathbf{P_i}}{\mathbf{t_x}}$$

$$\beta_{\mathbf{y,i}} = \frac{\mathbf{y_i} - \mathbf{z_i}\hat{\mathbf{R}}^{\mathbf{T}}\mathbf{P_i}}{\mathbf{t_y}}$$

In order to eliminate the effects of noise, the scale factor β can be estimated as:

$$\beta = \frac{Med(\beta_{x,i}) + Med(\beta_{y,i})}{2}$$

where $Med()$ means the median value of a series.

5. The finally calibrated translation vector $\hat{\mathbf{t}}$ is estimated as:

$$\hat{\mathbf{t}} = \beta\dot{\mathbf{t}}$$

6. The depth of each point $\hat{\mathbf{z}}'_\mathbf{i}$ can be estimated using Equation 11. Because the filter is in front of the camera, the depth of each point should be positive. Therefore, a sign check is required on the final calibrated translation vector $\hat{\mathbf{t}}$ and on the depth $\hat{\mathbf{z}}'_\mathbf{i}$ of each point on the filter as follows:

if $\sum_{i=1}^{k} sign(\hat{\mathbf{z}}'_\mathbf{i}) \leq 0$, then $\hat{\mathbf{t}} = -\hat{\mathbf{t}}$, and $\hat{\mathbf{z}}'_\mathbf{i} = -\hat{\mathbf{z}}'_\mathbf{i}$

where $sign(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{otherwise} \end{cases}$ and k is an application dependent integer (normally, $k \geq 3$).

4 Experimental Results

Both SGA and SEGA algorithms were implemented as follows. From a 3D image database, we selected 30 points $\mathbf{p_i} = (\mathbf{x_i}, \mathbf{y_i}, \mathbf{z_i})^\mathbf{T}$ $(i = 1, 2, \cdots, 30)$ on a 3D object described in a reference coordinate frame. These points were subject to controlled rotations of 20, 30, 40, and 50 degrees around a rotation axis $\mathbf{h} = (\mathbf{0}, \mathbf{0}, \mathbf{1})^\mathbf{T}$ with a constant translation vector $\mathbf{t} = (\mathbf{3}, \mathbf{4}, \mathbf{0})^\mathbf{T}$ yielding their correspondence points $\mathbf{p'_i} = (\mathbf{x'_i}, \mathbf{y'_i}, \mathbf{z'_i})^\mathbf{T}$ in camera centred coordinate frame. Finally, correspondence points were projected on the image plane $\mathbf{z'} = \mathbf{1}$ assuming a focal length f of the camera equal to 1. We thus, have one set of 3D object points in the reference coordinate frame and their corresponding 2D perspective image points $\mathbf{P'_i} = (\mathbf{X'_i}, \mathbf{Y'_i})^\mathbf{T}$ $(i = 1, 2, \cdots, 30)$ on the image plane. These are the control sets of points that serve as reference for error estimation.

In order to simulate real world noise contaminated data, Gaussian random noise was then added to the coordinates of object points and their image points and used the proposed SGA and SEGA algorithms to estimate the orientation parameter represented as the 3D rotation matrix $\hat{\mathbf{R}}$, the position parameter represented as the 3D translation vector $\hat{\mathbf{t}}$ and the structural data represented as depth $\hat{\mathbf{z}}'_\mathbf{i}$. We defined the relative calibration error of 3D rotation matrix $\hat{\mathbf{R}}$ as $e_R = ||\hat{\mathbf{R}} - \mathbf{R}||/||\mathbf{R}||$, the relative calibration error of 3D translation vector as $e_t = ||\hat{\mathbf{t}} - \mathbf{t}||/||\mathbf{t}||$, and the relative calibration error of the structure as $e_{z'} = (\hat{\mathbf{z}}'_\mathbf{i} - \mathbf{z}'_\mathbf{i})/|\mathbf{z}'_\mathbf{i}|$. Experimental results are summarised in Tables 1 and 2.

An overall analysis of the tables reveals that SGA algorithm is superior both in accuracy and performance compared with the SEGA algorithm. The main reason is that the SEGA algorithm does not take into consideration the rigid constraints of distance between feature points and does not use angular information. Experiments have shown that the SEGA algorithm is not only sensitive to noise, but also sensitive to rounding errors. Unless accurate enough data can be obtained, satisfactory performance of the algorithm cannot be guaranteed.

On the other hand, the SGA algorithm has still room for further performance improvements when compared with the SEGA method, as the latter makes full use of all points information while the SGA algorithm uses only partial points information, such as those that are far from the fixed point $\mathbf{c}$. Thus, by selectively using fewer feature points, it is possible to vastly reduce the computation time of the algorithm without loss of accuracy and this is critical for real time applications. The tables also show that the accuracy of the algorithms tends to decrease when data are corrupted by heavier noise. This is not a surprising result and, relatively, while the accuracy of the SGA algorithm can still be considered satisfactory, the accuracy of the SEGA algorithm decreases more rapidly.

Rotation angle(deg)	20			
Calibration methods	e_R (%)	e_t (%)	$e_{z'}$ (%)	time (s)
SGA	2.41	9.74	0.93	0.005
SEGA	2.33	10.33	1.51	0.03
Rotation angle(deg)	30			
Calibration methods	e_R (%)	e_t (%)	$e_{z'}$ (%)	time (s)
SGA	1.86	7.54	-1.79	0.01
SEGA	2.94	12.79	-1.98	0.04
Rotation angle(deg)	40			
Calibration methods	e_R (%)	e_t (%)	$e_{z'}$ (%)	time (s)
SGA	1.46	5.96	6.61	0.005
SEGA	3.64	15.49	-0.25	0.03
Rotation angle(deg)	50			
Calibration methods	e_R (%)	e_t (%)	$e_{z'}$ (%)	time (s)
SGA	1.11	4.60	0.71	0.005
SEGA	4.32	18.08	0.73	0.03

Table 1: Relative average calibration errors and calibration time for data corrupted by Gaussian noise with mean zero and deviation 0.03125.

Rotation angle(deg)	20			
Calibration methods	e_R (%)	e_t (%)	$e_{z'}$ (%)	time (s)
SGA	8.17	33.11	-12.35	0.01
SEGA	9.29	34.72	9.32	0.03
Rotation angle(deg)	30			
Calibration methods	e_R (%)	e_t (%)	$e_{z'}$ (%)	time (s)
SGA	5.95	24.12	-0.5	0.01
SEGA	11.25	41.96	-3.01	0.03
Rotation angle(deg)	40			
Calibration methods	e_R (%)	e_t (%)	$e_{z'}$ (%)	time (s)
SGA	4.57	18.61	-98.84	0.005
SEGA	12.91	43.39	103.9	0.04
Rotation angle(deg)	50			
Calibration methods	e_R (%)	e_t (%)	$e_{z'}$ (%)	time (s)
SGA	3.51	14.45	-32.30	0.005
SEGA	13.97	54.63	-48.36	0.03

Table 2: Relative average calibration errors and calibration time for data corrupted by Gaussian noise with mean zero and deviation 0.0625.

5 Conclusions

A special system setup for a real time automatic inspection of filter components has been designed based on the analysis of geometrical properties of image correspondence vectors synthesised into a single coordinate frame. The setup design is likely to render automatic inspection feasible and reliable within given performance constraints. The proposed SGA algorithm based on this special setup provides the closed form solutions to all estimated parameters necessary to decide whether a filter is to be rejected or not. While our approach is substantially different from epipolar geometry, we pointed out that linear algorithms based on epipolar geometry represent a common linear approach to solve 3D-2D and 2D-2D problems. For a comparative study of performance, we also developed another algorithm, SEGA, based on epipolar geometry.

Experimental results have demonstrated that the overall performance of the SGA algorithm is superior to SEGA's. Moreover, it has been shown that the SGA algorithm represents a practical proposition that meets the requirements of the task in terms of real time performance and high reliability. Work on algorithm refinements and its application to on-line inspection of filter components is under way and results will be reported in the near future.

References

[1] T.S. Huang and A.N. Netravali, "Motion and structure from feature correspondence: a review", *Proceedings of the IEEE*, Vol. 82, No. 2, pp. 252-268, February 1994.

[2] A. Mitiche and J.K. Aggarwal, "A computational analysis of time-varying images", *Handbook of Pattern Recognition and Image Processing*, T.Y. Yuong, K.S. Fu, Academic Press, 1986.

[3] J. Weng, T.S. Huang, and N. Ahuja, "Motion and structure from two projective views: algorithms, error analysis, and error estimation", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 5, pp. 451-476, 1989.

[4] R.Y. Tsai and T.S. Huang, "Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, No. 1, pp. 13-27, 1984

[5] T.S. Huang and O.D. Faugeras, "Some properties of the E matrix in two-view motion estimation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 12, pp. 1310-1312, 1990

[6] R.M. Haralick, H.J. Chung-Nan Lee, X. Zhuang, V.G. Vaidya, and M.A. Kim, "Pose estimation from corresponding point data", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 6, pp. 1426-14461, 1989.

[7] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 9, no. 5, pp. 698-700, 1987.

[8] K. Kanatani, "Analysis of 3-D rotation fitting", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 16, no. 5, pp. 543-549, 1994.

[9] S. Linnainmaa, D. Harwood, and L.S. Davis, "Pose determination of a three-dimensional object using triangular pairs", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 10, No. 5, pp. 634-647, 1987.

[10] P.R. Wolf, *Elements of photogrammery*, New York: McGraw-Hill, 1974.

[11] S. Ganapathy, "Decomposition of transformation matrices for robot vision", *Pattern Recognition Letters*, Vol. 2, pp. 401-412, 1989

[12] M. Fischler and R.C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography", *Communications of ACM*, Vol. 24, No. 6, 99. 381-395, 1981.

[13] D. Oberkampf, D. F. DeMenthon, and L. S. Davis, "Iterative pose estimation using coplanar feature points", *Computer Vision and Image Understanding*, Vol. 63, No. 3, pp. 495-511, 1996.

[14] N. Cui, J.J. Weng, and P. Cohem, "Recursive-batch estimation of motion and structure from monocular image sequences", *Computer Vision and Image Understanding*, Vol. 59, No. 2, pp. 154-170, 1994.

[15] S.H. Joseph, "Optimal pose estimation in two and three dimensions", *Computer Vision and Image Understanding*, Vol. 73, No. 2, pp. 215-231, 1999.

[16] H. Araujo, R.L. Carceroni, and C.M. Brown, "A fully projective formulation to improve the accuracy of Lowe's pose-estimation algorithm", *Computer Vision and Image Understanding*, Vol. 70, No. 2, pp. 227-238, 1998.

[17] D. W. Eggert, A. Lorusso, R. B. Fisher, "Estimating 3-D rigid body transformations: a comparison of four major algorithms", *Machine Vision and Applications*, vol. 10, pp. 272-290, 1997.