Wen Zhang, Hong Hu* and Dongik Shin

Improved XY-axis Calibration Based on Bilinear Interpolation

Abstract: Based on bilinear interpolation, this paper presents an improved calibration method for machine vision systems. An experiment is conducted that involves the construction of a model between the image plane of the camera and the world plane of the object, which provides a relationship of the planes' coordinate systems. An improved distortion correction method is then applied to improve accuracy. The results of the experiment demonstrate the advantages of this method compared to traditional methods.

Keywords: calibration, distortion correction, bilinear interpolation, machine vision

1 Introduction

Although machine vision has become well established, the images obtained are distorted due to the poor structure and installation of cameras. Camera calibration is therefore an important issue in machine vision. Camera calibration is widely considered as the process of estimating the parameters of a pinhole camera model, which approximates the actual camera and then achieves the relationship between image and world points [1]. CCD cameras play an important role in modern machine vision systems because the precision of CCD camera calibration will strongly influence the overall performance of the system [2]. Therefore, to improve the precision of machine vision systems, the precise calibration of the camera must firstly be ensured.

Currently, there are two types of calibration targets used widely in machine vision: the first target is the two-dimensional square matrix [3], where a linear $3 \times 4$ matrix is applied that denotes the perspective transformation; and the second target is chessboard, which applies the non-linear least square method [4] and the deviation pie correction method [5] to correct both parameters of the distortion and the camera itself. Because the second type is applied to compute distortion parameters...
and camera parameters at the same time, the complexity and running time of calculation is increased, both of which researchers are constantly trying to decrease. Reference [6] proposed adaptive distortion correction to improve the calibration speed without loss of the predicted accuracy, while reference [7] only models the distortion parameters without the inner and outer parameters of the camera so that the calibration speed is substantially improved.

Considering that most calibration approaches use detected corners on the chessboard, it is imprecise to correct a certain point located inside the grids because if that point is located exactly at the corners of the chessboard, it can be corrected by corrected coefficients directly. This paper therefore introduces bilinear interpolation in the direction of X-axis and Y-axis respectively to improve the precision of the calibration on these feature points located inside the grids after finishing the camera calibration.

2 Distortion Correction Model

2.1 Imaging Model

The pinhole camera that is the object of this study is shown as Fig. 1. When using a pinhole camera, the geometric mapping from 3D to 2D is called a perspective projection [8]. O-XYZ is the 3D world coordinate system in which the origin O is regarded as the center of calibration board and any feature point P is represented three-dimensionally by X, Y, Z; o-uv is the image pixel coordinate system where the plane (u, v) being perpendicular to the optical axis is just the image plane of the camera (in pixels) and the origin o is in the left upper corner. Given a camera with an optical axis collinear to the Z-axis with all objects located at the plane Z=Z₀, this study aims to find the transformation relationship between the image plane (u, v) and the world plane (X,Y).
2.2 Bilinear Interpolation

In mathematics, linear interpolation is a method to construct new data points with a set of known data points for curve fitting, which applies linear polynomials [9]. If two points are given by the coordinates \((x_0, y_0)\) and \((x_1, y_1)\), the linear interpolation is the straight line between them, which is shown as Fig. 2. For a value \(x\) in the interval \((x_0, x_1)\), the value \(y\) along the straight line is regarded as being given from:

\[
\frac{y - y_0}{x - x_0} = \frac{y_1 - y_0}{x_1 - x_0}.
\]

We will obtain \(y\) by solving (1) with unknown value at \(x\),

\[
y = y_0 + (x - x_0) \frac{y_1 - y_0}{x_1 - x_0},
\]

Which is just the formula for linear interpolation in the interval \((x_0, x_1)\).

The formula can also be understood as a weighted average both in the \(x\) and \(y\) direction. The weights are inversely related to the distance from the end points to the unknown point. That is to say, the closer point has more influence than the farther point. They are therefore supposed to be \((x-x_0)/(x_1-x_0)\) and \((x_1-x)/(x_1-x_0)\), which are normalized distances between the unknown point and each of the end points, and because they sum to 1 it yields the formula for linear interpolation as:

\[
y = y_0 \left(1 - \frac{x - x_0}{x_1 - x_0}\right) + y_1 \left(1 - \frac{x_1 - x}{x_1 - x_0}\right) = y_0 \left(\frac{x_1 - x}{x_1 - x_0}\right) + y_1 \left(\frac{x - x_0}{x_1 - x_0}\right)
\]
Bilinear interpolation is actually an extension of linear interpolation with interpolating functions of two variables (i.e., \(x\) and \(y\)) on a rectilinear 2D grid [10]. The key idea in bilinear interpolation is to perform linear interpolation first in one direction, and then again in the other direction, which is shown as Fig. 3. Although each step is linear in the sampled position, as a whole it is not linear but rather quadratic in the sample location.

\[ f(x, y) \approx \frac{y_2 - y}{y_2 - y_1} f(x, y_1) + \frac{y - y_1}{y_2 - y_1} f(x, y_2) \]

Then proceed with interpolation in the \(y\)-direction to obtain the desired estimate:

\[ f(x, y) \approx \frac{y_2 - y}{y_2 - y_1} \left( \frac{x_2 - x}{x_2 - x_1} f(q_{11}) + \frac{x - x_1}{x_2 - x_1} f(q_{12}) \right) + \frac{y - y_1}{y_2 - y_1} \left( \frac{x_2 - x}{x_2 - x_1} f(q_{21}) + \frac{x - x_1}{x_2 - x_1} f(q_{22}) \right). \]

The same result will be arrived at if the interpolation is done first along the \(y\)-direction and then along the \(x\)-direction.
Similarly, let (5) be understood as a weighted average in x and y directions and the weights are \( t = \frac{x - x_1}{x_2 - x_1} \) and \( s = \frac{y - y_1}{y_2 - y_1} \), so that the bilinear interpolation formula is considered as:

\[
    f(x, y) = (1 - s)(1 - t)f(q_{11}) + tf(q_{21}) + s((1 - t)f(q_{12}) + tf(q_{22})).
\]  

(6)

### 2.3 Distortion Correction Model

As shown in Fig. 4, the whole distortion correction model is based on bilinear interpolation. If we know the world coordinate (X, Y) of a certain point P, we can apply the calibration method to find out its exact image coordinate (u, v). When we get the world coordinates of four vertexes \( P_{11}, P_{12}, P_{21}, P_{22} \), which are (X1,Y1), (X1,Y2), (X2,Y1) (X2,Y2) of the grid where (X,Y) is located, we can apply feature detection to obtain the corresponding image coordinates \( (u_{11},v_{11}), (u_{12},v_{12}), (u_{21},v_{21}), (u_{22},v_{22}) \). On the basis of the bilinear interpolation formula, suppose that u-coordinate and v-coordinate in the image coordinate system are regarded as the unknown function \( u(P) \) and \( v(P) \) respectively, and then the distortion correction model is shown as:

\[
    u(P) = (1 - s)((1 - t)u(P_{11}) + tu(P_{12})) + s((1 - t)u(P_{12}) + tu(P_{22}))
\]

\[
    v(P) = (1 - s)((1 - t)v(P_{11}) + tv(P_{12})) + s((1 - t)v(P_{12}) + tv(P_{22})).
\]  

(7)

That is to say:

\[
    u = (1 - s)((1 - t)u_{11} + tu_{12}) + s((1 - t)u_{21} + tu_{22}),
\]

\[
    v = (1 - s)((1 - t)v_{11} + tv_{12}) + s((1 - t)v_{21} + tv_{22}),
\]  

(8)

Where \( t = \frac{X - X_1}{X_2 - X_1} \) and \( s = \frac{Y - Y_1}{Y_2 - Y_1} \). Therefore, all we need to do is to put world coordinates of \( P_{11}, P_{12}, P_{21}, P_{22} \) into the correction model and then \( (u, v) \) is just the image coordinate for the point p.
Conversely, if the world coordinate \((X, Y)\) of a certain point is known, its exact image coordinate can also be calculated through inverse bilinear interpolation. If the image and world coordinates of four vertexes of the grid where the calculated point location is known, we can also solve the problem from the bilinear interpolation formula. Different from computing the image coordinate \((u, v)\), if we want to solve this equation to find the world coordinate \((X, Y)\) with known \((u, v)\), we first reorganize these equations to be:

\[
\begin{align*}
(u_{11} + u_{22} - u_{12} - u_{21})st + (u_{21} - u_{11})s + (u_{12} - u_{11})t &= u - u_{11} \\
(v_{11} + v_{22} - v_{12} - v_{21})st + (v_{21} - v_{11})s + (v_{12} - v_{11})t &= v - v_{11},
\end{align*}
\]  

(9)

Then the value \(s\) and \(t\) can be solved from this, so the world coordinate of the unknown point \((X, Y)\) will be found as:

\[
\begin{align*}
X &= X_1 + t(X_2 - X_1) \\
Y &= Y_1 + s(Y_2 - Y_1).
\end{align*}
\]  

(10)
3 Distortion Correction Method

3.1 Set the Grid System

Prepare a large chessboard with known dimensions whose vertexes can be represented as a certain set V = \{(X_0, Y_0), (X_1, Y_1), \ldots, (X_n, Y_n)\}, and take a photo of the chessboard using the camera that needs to be calibrated. Apply corner detection to this photograph in order to get the image coordinates corresponding to the vertexes on the chessboard, which can be represented as a set v = \{(u_0, v_0), (u_1, v_1), \ldots, (u_n, v_n)\}. All the calibration is based on these two coordinate sets, and all values to be used to interpolate are from them. The image grid system and the world coordinate system are shown in Fig. 5.

![Fig. 5: Coordinate system](image)

3.2 World-to-Image Coordinate System Calibration

If we know one certain point P=(X, Y) in the world coordinate system, we can find out the exact coordinate (u, v) in the image coordinate system. Fig. 5 shows the principle of CCD camera lens distortion correction by bilinear interpolation where the image coordinate is to be computed based on its known world coordinate.

Thanks to the standard grid in chessboard, it is convenient to determine the exact grid location of P. As shown in Fig. 5(a), we just need to compare X and Y with the points from set V to find out four vertexes P_1 = (X_i, Y_i), P_2 = (X_i, Y_{i+1}), P_3 = (X_{i+1}, Y_i),
P₄ = (Xᵢ₊₁, Yᵢ₊₁), which satisfy P₁, P₂, P₃, P₄ ∈ V and Xᵢ<Xᵢ₊₁, Yᵢ<Yᵢ₊₁. According to the prepared grid system, their corresponding image coordinates can be obtained from set v as (u₁₁, v₁₁), (u₁₂, v₁₂), (u₂₁, v₂₁), (u₂₂, v₂₂). By putting them into the bilinear interpolation formula, the image coordinate (u, v) can be easily solved.

### 3.3 Image-to-World Coordinate System Calibration

Unlike the world-to-image coordinate system calibration, if we only know the image coordinate of a certain point and the prepared grid system, there is no way to find the exact grid location of that point by simply comparing. Because of the lens distortion, shown in Fig. 5(b), which causes the u and v axis not to be straight and perpendicular, it is not possible to compare values in the u and v directions to confirm the four vertexes of the grid for the calculated point. The algorithm therefore calculates the cross product to confirm the location of the grid.

Suppose the grid system in the image plane is as shown as Fig. 6. The main idea of cross product to confirm the location is that we search each grid in the image plane to calculate cross product between vector q₁₁q₂₁ and vector q₁₁q:

\[
q_{11} q_{21} \times q_{11} q = \left| q_{11} q_{21} \right| \sin <q_{11} q_{21}, q_{11} q>.
\]  

(11)

If this cross product is larger than the value 0, vector q₁₁q is located within 180 degrees anti-clockwise of vector q₁₁q₂₁; otherwise it is beyond 180 degree. Once we have calculated the cross products at all four vertexes q₂₁q₂₂×q₂₁q, q₂₂q₁₂×q₂₂q, q₁₂q₁₁×q₁₂q and they are all larger than 0, it is proved that the target point is located inside this grid. We can then get image coordinates of these four vertexes (u₁₁, v₁₁), (u₁₂, v₁₂), (u₂₁, v₂₁), (u₂₂, v₂₂) and their corresponding world coordinates.
Once we have obtained four vertexes, inverse bilinear interpolation is applied to solve the problem. What we need to do is put the image coordinates of these four vertexes into the formula to find \((u, v)\) and finish the calibration.

### 4 Experiments

The current most common calibration method is that based on a planar template, which was proposed by professor Zhang [11]. The method proposed in this paper improves upon that method by introducing a two-step calibration approach by first obtaining a set of photographs taken at different orientations of the planar template and then solving the intrinsic and extrinsic parameter matrix of the camera. An experiment on this improved camera calibration method was conducted in order to prove that it works.

The camera used for calibration was a CCD camera. The model plane we used was a pattern of \(8 \times 8\) squares, which is to say 20 inner corners with each square \(24\text{mm} \times 24\text{mm}\). This prepared chessboard was precisely printed. During this experiment, eight images of this chessboard were first taken in accordance with Zhang's calibration method to obtain the intrinsic and extrinsic parameter matrix and also the distortion parameters. Secondly, we applied the calibration results to 22 points in the image that had been prepared for calibration with world and image coordinates already known, as shown in Fig. 7. Using the known image coordinates of each point, the corresponding world coordinates were calculated by Zhang's method. Suppose \((u_0, v_0)\) is the actual value for the image coordinate and \((u_c,v_c)\) represents the calculated value obtained by parameters of the camera calibration, so the error of each point is calculated by:
The results are shown in Table 1. To prove the increased precision obtained by using the proposed method, we applied bilinear interpolation to the image that had previously been calibrated by Zhang's method. Similarly, we checked the error of each point and the results are shown in Table 1 and Fig. 8.

\[ e = (u_c - u_0)^2 + (v_c - v_0)^2. \] (12)

As can be seen from Table 1, after calibration by the proposed method the average error of these points is clearly smaller than the calibration error incurred using Zhang's method alone. So we can see from the results in this experiment that the method based on bilinear interpolation can increase the accuracy in the transformation between image and world coordinate systems.

<table>
<thead>
<tr>
<th>Calibration procedure</th>
<th>Calibration variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibrated only by Zhang's method</td>
<td>3.480</td>
</tr>
<tr>
<td>Calibrated by both Zhang's method and bilinear interpolation</td>
<td>2.658</td>
</tr>
</tbody>
</table>
5 Conclusion

This paper introduced bilinear interpolation to correct camera distortion and proposed improved algorithms for the world-to-image coordinate system for calibration and the image-to-world coordinate system for calibration. Results from the experiments show that calibration based on bilinear interpolation can be considered as a new calibration strategy for increasing the precision of XY-calibration.

Acknowledgement: The authors gratefully acknowledge the support provided by the National Nature Science Foundation of China under NSFC Grant 51375118, as well as the support from Shenzhen Government Fund JCYJ20160330161701343.

References
