Automatic Distortion Correction of Endoscopic Images Captured with Wide-Angle Zoom Lens

Tung-Ying Lee, Tzu-Shan Chang, Chen-Hao Wei, Shang-Hong Lai, Kai-Che Liu, Hurng-Sheng Wu

Abstract—Operation in minimally invasive surgery is more difficult since the surgeons perform operations without haptic feedback or depth perception. Moreover, the field of view perceived by the surgeons through endoscopy is usually quite limited. The goal of this paper is to allow surgeons to see wide-angle images from endoscopy without the drawback of lens distortion. The proposed distortion correction process consists of lens calibration and real-time image warping. The calibration step is to estimate the parameters in the lens distortion model. We propose a fully automatic Hough-entropy-based calibration algorithm, which provides calibration results comparable to the previous manual calibration method. To achieve real-time correction, we use graphics processing unit to warp the image in parallel. In addition, surgeons may adjust the focal length of a lens during the operation. Real-time distortion correction of a zoomable lens is impossible by using traditional calibration methods because the tedious calibration process has to repeat again if focal length is changed. We derive a formula to describe the relationship between the distortion parameter, focal length, and image boundary. Hence, we can estimate the focal length for a zoomable lens from endoscopic images on-line and achieve real-time lens distortion correction.

Index Terms—Endoscopic Image, Optical Distortion, Wide-Angle Distortion Correction, Zoom Lens Calibration

I. INTRODUCTION

MINIMALLY invasive surgery (MIS) is one of the main evolutions of surgical techniques and it provides great benefits to the patient. MIS uses an endoscopic camera and instruments manipulated through small incision on the skin to avoid open surgery. This technique reduces patient scars, infection risk, post-operative morbidity and patient recovery. Although a considerable amount of intervention is still performed with standard open surgery techniques, a large amount of research work is performed every year to develop new instrumentation that replace open surgery procedure by MIS procedure and to make the MIS interventions easier. Indeed, MIS procedures are still more difficult to perform than standard surgery for many reasons. Firstly, surgeons lose depth perception since the camera is monocular and the visual resolution is reduced since the video sensor resolution is still not able to match human eye resolution. Secondly, they lose the haptic feedback since they cannot touch the organs with their hands. Instead, they can obtain haptic feedback with some particular instruments. However, feeling an artery position from blood pulse, for example, is currently not feasible with instruments. Thirdly, the camera field of view is generally quite limited compared to human perception (70° against 160° for human eyes) for almost all abdominal MIS procedures. The main reason is that the wide-angle lens introduces image distortion: straight lines become curved in the distorted image, thus seriously decreasing the hand eye coordination of the surgeon and degrading the capability of size and shape determination.

This small field of view for an endoscopic camera has several drawbacks. When the surgeon introduces the endoscopic camera, a long exploration phase is necessary to localize each organ position and status. And even after this step, the surgeon can never have a broad view of the patient’s organs that allows him to make faster movement. A second drawback is the instrument occultation when they are not in the camera field of view. For some procedures, more than 5 instruments are used at the same time (and hold by assistants) and they cannot be visible at the same time. Hidden instruments should not be moved since they can damage the closest organ. This means that the intervention is slowed down because all the instruments cannot be moved at the same time. Moreover, since the camera holder usually does not manipulate instruments, a very good coordination is necessary among the whole surgical team so that the camera is always focused on what is important during the surgical procedure.

The goal of this paper is to allow surgeons to use an endoscopic wide-angle lens without the drawback of the lens distortion. Since the current cameras are digital, removing distortion can be realized via image processing. In this context, the goal of the paper is to develop a system which can efficiently compensate the wide-angle lens distortion under different zooming conditions. This system is divided into two parts, calibration part and image correction part. In the calibration part, the system captures several images which contain the calibration pattern, and then the distortion parameters of a lens are estimated. These distortion parameters are used to undistort image in the correction part. Although there are existing calibration methods and interpolation methods, there are several challenges to apply them to the practical problem. In practice, current endoscopy provides 60Hz frame rate for visual comfort.
at HD resolution. It is almost impossible to achieve real-time image correction by using only central processing unit (CPU) for the computation due to the large amount of computation involved in the image correction. Secondly, most calibration methods are manual or semi-automatic, and they need some user intervention. Furthermore, traditional methods are still not feasible for a zoomable lens even though the lens can be calibrated before operation. It is because the distortion parameters will be changed when surgeons change the focal length of a lens. In order to make the system practical, a reliable, fast, and automatic calibration is required.

Wide-angle cameras have been widely used in video surveillance, automotive applications, and endoscopic imaging [1]–[12]. A large field of view is obviously needed in these applications. For the application of 3D reconstruction [2], [3] and mosaicing [11], the reconstruction and mosaicing algorithms are based on the undistorted images. Undistorted images are also needed when the system uses several different modalities of images to achieve virtual surgical navigation [4]. The impact of distortion for automatic classification of celiac disease was discussed in [5], [6].

For a lens without zooming, the distortion parameters of an endoscope could be estimated offline. There exist some systems of real-time image distortion correction [4], [7], [9] for endoscopy without zooming. In the system proposed by Shahidi et al. [9], the endoscope was calibrated by using Tsai’s method with some 2D-to-3D point correspondences determined manually. Some systems [8], [10] used dot patterns for calibrating a wide-angle lens, they require careful setting of endoscopy during calibration or assume dot extraction is under controlled environment such that these dots could be organized in a vertical or horizontal fashion. Moreover, the parameter-free distortion correction methods [13], [14] require point correspondences between the input image and calibration pattern, which usually requires user interaction. Thus, an automatic and robust image calibration method is strongly demanded.

A zoomable lens further complicates the problem. The calibration and correction of images acquired with a zoomable lens involves two difficulties. Firstly, it requires a mechanism to calibrate the camera for all possible zoom settings, which either is quite labor intensive [15] or needs special hardware to obtain focal length and radial distortion [16]. These methods are not suitable for calibrating a zoomable endoscope. The second difficulty is to estimate zoom setting of a lens online even if a zoomable lens is fully calibrated.

In this paper, we propose a single-image Hough-based autocalibration for calibrating a wide-angle endoscope with a certain focal length. It is robust against noise and occlusion. Unlike the method by Shahidi et al. [9] that uses small line segments to estimate radial distortion, our method automatically estimates the radial distortion from several distorted straight lines. If the pattern consists of parallel lines, our framework can also utilize the prior knowledge of parallel line to improve the calibration accuracy. We propose a two-stage Hough transform based method for the estimation of the distortion parameters. Furthermore, we extend the proposed system to calibrate the distortion parameters for a zoomable lens.

To describe the relationship between the wide-angle distortion and focal length, we derive a formula that simplifies the process of zoom calibration, and just a few zoom settings need to be calibrated. Hence, we can estimate focal length online and achieve real-time distortion correction. In this paper, we demonstrate the proposed algorithm can achieve real-time correction of wide-angle lens distortion for a zoomable lens.

For the rest of this paper, we present the single-view Hough-based auto-calibration method in Section II. The relationship between the distortion parameter and focal length is derived in Section III. The flow of the whole system is described in Section IV. In Section V, we will the accuracy of the proposed auto-calibration method and the performance of the system.

A preliminary version of this work, given in parts of Section II, was reported in [17]. In this paper, we propose an extended algorithm based on the two-stage Hough transform to take advantage of the parallel-line calibration patterns and develop the calibration and correction framework for a zoomable lens.

II. SINGLE-IMAGE HOUGH-BASED AUTOCALIBRATION

Camera calibration usually uses two kinds of markers, corners and straight lines. A popular calibration toolbox [18] uses corners of chessboards for establishing point correspondences. Although some existing camera calibration methods utilize automatic chessboard detection [19], their performance usually depends on the correctness of chessboard detector, which is easily influenced by illumination. The line-based camera calibration is another popular approach [20]–[23]. In structured environments or straight-line patterns, using line features could be a more suitable option. The projected lines without distortion should be straight; hence, this principle could be used in removing optical distortion [22], [23].

Hough transform [24] is a powerful tool for extracting straight lines. Several previous works have used Hough transform to find the perspective transform [25] or vanishing points [26], [27]. From the distorted images, Habib and Morgan [28] detected clusters of distorted lines and Xu et al. [29] just used Hough transform to find the distortion center given all other distortion parameters.

The most related works to our calibration method are the HTRDC method proposed by Cucchiara et al. [30] and 1-D angular Hough method [31]. These two methods are based on polynomial distortion model and Harris model. The HTRDC method focuses on using a single straight line, and tries to adjust the distortion parameter to maximize the maximal vote assuming the distortion center is known. The 1-D angular Hough method requires less space. This method assumes the orientations of lines are sparsely distributed in the angular Hough space. However, this could be violated due to the perspective effect. A plane with several parallel lines could be served as a pattern for auto-calibration (we denoted it by parallel-line-pattern), but it is a good example that the orientations of lines are not sparsely distributed if the plane of calibration pattern is not parallel to the image plane.

We propose a Hough entropy based method considering all distorted straight lines, and it estimates the distortion...
parameter and the distortion center. The entropy in the Hough space is employed to measure straightness of the undistorted curves. In a naive way, all edge points are undistorted by possible parameters and then the best parameter setting is selected based on the entropy in the Hough space. However, the results will be degraded by multiple lines occurred in the image, because the entropy will be lower when the image is distorted to a smaller size. To overcome this problem, we only count the votes with most possible orientation for each edge point. In addition, the image gradient computed from the distorted image is also distorted. Hence, we derive the undistorted gradient for more compact Hough voting, and it also solves the aforementioned degenerate problem. In the following subsections, we will introduce some preliminaries, including distortion model and Hough transform, and then the Hough entropy method and gradient estimation are presented subsequently. A plane with parallel lines (Parallel Line Pattern) could be served as the calibration pattern in our system. We use the two-stage Hough transform to incorporate the prior of this type of calibration patterns. The first layer Hough map collects votes for straight lines, and the second layer Hough map is used for accumulating pencils of lines. The edge points produced by noise or unrelated curves will be suppressed in the second map. It can make the calibration more accurate.

A. Pin-Hole Camera Model and Division Model

A popular and simple pin-hole camera model contains focal length \( f \) and assumes that camera is located at the origin and faces to negative Z-axis. An object location with respect to the camera coordinate is denoted by \((X, Y, Z)\). The corresponding image point is inversely proportional to the depth \( Z \), i.e., \( x = -fx/Z, y = -fy/Z \). If the camera coordinate and world coordinate are different, there exists a rotation matrix \( R \) and translation vector \( T \) for the coordinate transformation.

Optical distortion can be easily incorporated into the model. Two kinds of optical distortions, radial distortion and tangential distortion were modeled in [32]. The radial distortion at a point is represented by a displacement along the direction of the point to the camera center. The tangential direction is perpendicular to the radial direction. However, low tangential distortion could be compensated by estimating the distortion center [33]. The relationship between the distorted radius of an image point and the undistorted radius is approximated by a low-degree polynomial. Many researchers have found that the higher-order terms could be ignored [34]. In addition, Fitzgibbon [35] suggested a division model which uses a low-order polynomial in the denominator instead of the numerator.

The corresponding undistorted coordinate of an image coordinate \( x_d = (x_d, y_d) \) is defined by \( x_u = (x_u, y_u) \). The division model [35] used to model the optical distortion relates these two coordinates as follows:

\[
r_u = \frac{r_d}{1 + \lambda r_d^2}, \quad r_u = \|x_u - c\|, \quad r_d = \|x_d - c\| \tag{1}
\]

where \( c = (c_x, c_y) \) is the image center. Equation (1) gives the relationship of distorted radius \( r_d \) and undistorted radius \( r_u \). For wide-angle distortion correction by using the division model, we just need to determine the image center \((c_x, c_y)\) and distortion parameter \( \lambda \).

B. Hough Transform

Traditional Hough transform is to transform image features in an image domain to a line parameter space to vote for objects, i.e., lines. Any one line passing through \((x, y)\) can be represented in polar form \((r, \theta)\), and the corresponding line equation is \( r = x \cos \theta + y \sin \theta \), and all possible lines can be collected by \( S_{\text{cont}}(x, y) = \{(r, \theta) | r = x \cos \theta + y \sin \theta \} \). For each edge point \((x, y)\) in the image, all lines passing through \((x, y)\) will receive one vote individually. Finally, the Hough map stores the total number of votes for each line \((r, \theta)\).

In the computer program, the polar coefficients are discretized and a Hough map is a 2D accumulation array used for recording votes. The possible lines passing through \((x, y)\) can be written by a transfer function \( S \):

\[
S(x, y) = \{(x \cos \theta_i + y \sin \theta_i, \theta_i) | \forall \theta_i \} \tag{2}
\]

where \( \theta_i \) is the sampled angle in \([0, \pi]\). The voting process for the edge point \((x, y)\) is denoted by

\[
H(r, \theta) \leftarrow H(r, \theta) + \text{vote}(S, (x, y), \text{inc}) \tag{3}
\]

where \( H \) is an accumulation array and \( \text{inc} \) is the increment (In traditional Hough transform, \( \text{inc} \) is equal to 1).
We propose the Hough entropy method with image gradient estimation for distortion parameter estimation. We found that sometimes the result is degraded because of the votes in the Hough space which are generated from some orientation not similar to the edge direction. Hence, we derive the gradient estimation formula and just vote the most possible orientation for each edge point.

### C. Hough Entropy Method

First, the edge points which could belong to the distorted straight lines are extracted by using Canny edge detector. A typical range of distortion parameters and the possible regions of image center are selected. A uniform sampling is performed in this 3D parameter space. For each candidate parameter triplet \((\lambda, c_x, c_y)\), all edge points are undistorted as follows:

\[
\text{undist}(x_d, y_d; \lambda, c_x, c_y) = \left[ x_d - c_x + \frac{1}{1 + \lambda r_d^2} c_x, y_d - c_y + \frac{1}{1 + \lambda r_d^2} c_y \right]
\]

(4)

Based on these undistorted points, a Hough map \(H(r, \theta)\) is obtained and normalized by dividing the total number of all votes. The entropy of the normalized Hough map is calculated with the following entropy formula.

\[
p(r, \theta) = H(r, \theta)/\sum_{i,j} H(i, j)
\]

(5)

\[
\text{Entropy}(H) = \sum_{v(r, \theta), s.t. p(r, \theta) \neq 0} -p(r, \theta) \ln p(r, \theta)
\]

(6)

Finally, the parameter triplet with the minimum entropy is selected as the best distortion parameter and image center.

### D. Hough Entropy with Gradient Estimation (HEwG)

Fig. 1 illustrates the degenerate problem occurred in Hough entropy method. When the distortion parameter is large, the undistorted image will decrease its size. Hence, the zero counters in Hough map will increase and entropy decrease.

The modified version of Hough entropy method is to use image center and distortion parameter to estimate the orientation of an edge point in undistorted image. Undistorting an image costs too much computation power, because our method examines many candidates in the parameter space. Hence, we derive how to use the image gradient in the division model to simplify the computation in the Hough transform.

For estimating the orientation, we first calculate image gradient on distorted image by using the Sobel masks. Let

\[
\epsilon
\]

be the weight of line \(L\)

\[
\epsilon
\]

obtained and normalized by dividing the total number of all votes. The entropy of the normalized Hough map is calculated with the following entropy formula.

\[
H_{GE}(x, y, g_x, g_y) = \{(r, \theta) | r = x \cos \phi + y \sin \phi \}
\]

(8)

\[
\phi = \cos^{-1} \left( \frac{g_x}{\sqrt{g_x^2 + g_y^2}} \right)
\]

(9)

The best parameter setting is selected by finding the minimal entropy measure.

### E. Incorporation of Hough Framework and Parallel-Line-Patterns (PLPs)

If the calibration pattern is PLP, we employ the two-stage Hough transform to incorporate the prior of this type of calibration patterns. The edge points produced by noise or unrelated curves will be suppressed in the second map. After performing the first Hough transform in Section II-D, the output Hough map \(H_1\) is used as the input to the second Hough transform. The proposed algorithm is to evaluate the entropy of the second Hough map \(H_2\).

The projected version (projected on the image sensor) of a set of 3D parallel lines can be represented by a vanishing point \(V\) in the image space. However, the range of vanishing point is very large, even unbounded. To overcome this problem, we introduce two auxiliary points \(P\) and \(Q\), i.e., \((x_P, y_P)\) and \((x_Q, y_Q)\), in the image for the new representation of the second Hough transform. In practice, we select the left-upper image corner as \(P\) and the image center as \(Q\). Without loss of generality, the vanishing point \(V\) can be represented by the orientations of the two lines, \(PV\) and \(QV\), which are denoted by \((\omega, \rho)\). We use this representation in the second Hough map.
Algorithm 1: Hough Entropy Framework

Input: the initial range of distortion parameter \([\lambda_{\text{min}}, \lambda_{\text{max}}]\),
the initial range of distortion center in \(x\) coordinate \([x_{\text{min}}, x_{\text{max}}]\) and in \(y\) coordinate \([y_{\text{min}}, y_{\text{max}}]\), image \(I\),
sample number \(S\), user-defined tolerance \(\epsilon\)

Output: the best distortion parameter \(\lambda\), the best distortion center \((c_x, c_y)\)

1: Extract the set of edge points \(E\) from \(I\), and the associated gradient \(I_x\) and \(I_y\).
2: for \(l = 1\) to \(20\) do
3:   Sample \(S\) samples over \([\lambda_{\text{min}}, \lambda_{\text{max}}]\)
4:   Sample \(S\) samples over \([x_{\text{min}}, x_{\text{max}}]\)
5:   Sample \(S\) samples over \([y_{\text{min}}, y_{\text{max}}]\)
6:   Generate \(S^3\) possible parameter configurations from previous samples.
7:   for each configuration \((\lambda^l, c_x^l, c_y^l)\) do
8:     if method is PLP extension then
9:       Run Algorithm 3
10:      else
11:         Run Algorithm 2
12:     end if
13:   end for
14: Find the best configuration minimizing Hough entropy \(h_c\).
15: For distortion parameter and distortion center use two nearest samples to update \(\lambda_{\text{max}}, \lambda_{\text{min}}, x_{\text{max}}, x_{\text{min}}, y_{\text{max}}, y_{\text{min}}\),
16: if \((\lambda_{\text{max}} - \lambda_{\text{min}}) < \epsilon\) then
17:   break
18: end if
19: end for

Fig. 4. Hough Entropy Framework

Algorithm 2: Hough Entropy with Gradient Estimation

Input: the configuration \((\lambda^l, c_x^l, c_y^l)\)

Output: the associated Hough map \(H_1\) and Hough entropy

1: Use this configuration to undistort the edge point and calculate undistorted gradient by using (4) and (7).
2: Vote for each undistorted point by using (10) in \(H_1\).
3: Estimate Hough entropy of \(H_1\).

Fig. 5. Hough Entropy with Gradient Estimation

H2. Note that the map is bounded because orientation angles
are limited within 360 (or 180) degrees.

Given the first Hough map \(H_1\), the voting process is given
as follows. A cell in \(H_1\) at position \((r, \theta)\) represents a line
with supporting votes \(H_1(r, \theta)\) as the line \(L\) shown in Fig. 2. All
possible vanishing points of lines parallel to \(L\) are on the line \(L\).
We can vote them by using two-orientation representation.
For each sample \(\omega_i\), we can construct the line \(M\) passing
through \(P\) with orientation \(\omega_i\), \(\rho = xP \cos \omega_i + yP \sin \omega_i\).
The corresponding vanishing point \(V(\omega_i)\) is the intersection
of the line \(L\) and \(M\). The orientation \(\rho_i\) of the line \(QV(\omega_i)\)
can be computed. The location \((\omega_i, \rho_i)\) in \(H_2\) corresponds
to \(H_1(r, \theta)\) votes.

Based on the theory of projective geometry, the intersection
of two lines and the line crossing two points are easily calculated
by cross-product in homogeneous coordinates. Thus, we
define a transfer function \(S_{\text{PLP}}\) with the following equations.

\[
S_{\text{PLP}}(r, \theta) = \{(\omega, \rho) | \rho = \rho(\omega)\} \tag{10}
\]

\[
\rho(\omega) = \tan^{-1}(n_1(\omega)/n_2(\omega)) \tag{11}
\]

\[
(n_1(\omega), n_2(\omega), n_3(\omega)) = (xQ, yQ, 1) \times V(\omega) \tag{12}
\]

\[
V(\omega) = (\cos \omega, \sin \omega, -r(\omega)) \times (\cos \theta, \sin \theta, r) \tag{13}
\]

\[
r(\omega) = xP \cos \omega + yP \sin \omega \tag{14}
\]

Based on the transfer function \(S_{\text{PLP}}\), the line \(L\) will
contribute \(H_1(r, \theta)\) votes at all locations in \(S_{\text{PLP}}\). Finally, we
calculate the entropy of \(H_2\) instead of that of \(H_1\). In Fig. 3,
we depict an example of distortion correction on a chessboard
image and the corresponding \(H_1\) and \(H_2\) maps.

In the cascaded Hough transform [27], several Hough trans-
forms are also applied cascadedly. However, our framework
is different from the traditional cascaded Hough transform.

In the first Hough map, we use radial parameterization for
line, not the slope form in [27]. In the second Hough space,
we introduce two auxiliary points and use two relative angles
instead of the vanishing point. Hence, these two Hough spaces
are limited. We do not need to fold the Hough space as that
used in [27].

F. Optimization

Our camera distortion calibration is an optimization process
to find the best distortion parameters based on the proposed
measures. Our measures are evaluated by entropy of Hough
map, and it is not easy to apply continuous optimization to
solve the problem. The coarse-to-fine search is a possible
strategy [30]. We modified this hierarchical search as follows.
First, we uniformly sample over a bounded parameter space.
The sample with the minimal entropy is selected, and then
we reduce the parameter search space to the neighborhood
of this sample. The next round of sampling, evaluation, and
reduction of parameter search space is repeated until the size
of the parameter search space is less than a user-defined
accuracy. The Hough entropy method with gradient estimation
and the version of incorporation of PLP prior are summarized
in Algorithm 1, 2, and 3 (shown in Fig. 4, 5, and 6).

In our implementation, the typical range of distortion parameter
is given by \([-1/r_{\max}^2, -10^{-7}\]) where \(r_{\max}\) is the distorted
radius of visible regions, the possible range of distortion center
is limited in a \(0.1 \times W\) by \(0.1 \times H\) region which is located in
distortion parameter \(\lambda\) is less than \(\epsilon\), the optimization is finished. In our implementation,
\(\epsilon\) is set to \(-10^{-9}\), and the number of iterations required to meet
this convergence criterion is usually no more than 7.

III. EXTENSION TO A ZOOMABLE LENS

A. Quadratic Prediction of Distortion Parameters

We extend our algorithm to a zoomable lens. Consider the
pin-hole camera model. Assume the focal length is changed to
\(rf\) and the center of camera is stationary. The new projected
point \((x_r, y_r)\) satisfies the following equations, \(x_r = -rf X/Z\)
Algorithm 3: Extension of Parallel-Line Patterns

Input: the configuration \((\lambda_f, c_x, c_y)\)

Output: the associated Hough entropy with incorporation of parallel-line patterns

1: \((H_1, \text{Entropy}(H_1)) = \text{Run Algorithm 2}

2: for each line \((r, \theta)\) in \(H_1\) do

3: Find out all corresponding cells in \(H_2\) by using \(S_{PLP}\), and increase these cells by \(H_1(r, \theta)\)

4: end for

5: Estimate Hough entropy of \(H_2\)

---

**Fig. 6. Extension of Parallel-Line Patterns**

---

**Fig. 7. A zoomable fisheye lens.**

---

and \(y_t = -r f Y/Z\). Most lenses for photography follow this rule.

Inspired by Kannala’s formulation [36], we divide the projection part into radial distortion and perspective projection as shown in Fig. 7. It means the incoming rays to the camera center are refraacted and then the refraacted rays are perspectively projected on the sensor. Originally, the division model is unrelated to focal length. Now we will derive the relationship between the change of distortion parameters and the change of focal length.

Consider focal length \(f\) and \(rf\). Based on our assumption, the image \(J\) with focal length \(rf\) could be approximated by the linearly zoomed version of image \(I\) with focal length \(f\). A 3D point \(X\) is projected to \(x_d^f\) and \(x_d^{rf}\) in the image \(I\) and \(J\), respectively, and these projected points and their undistorted versions can be well approximated with the following equations, \(x_d^f = rx_d^f, x_d^{rf} = rx_d^{rf}\), and \(x_d^{rf} = rr_d^{rf}\).

We obtain the relationship between the change of distortion parameters and the change of focal length. The distortion parameter is inverse-quadratically proportional to focal length.

\[
x_u^{rf} = \frac{x_u^f}{1 + \left(\frac{\lambda_f}{r^2}\right)(r_d^{rf})^2} \Rightarrow \lambda_{rf} = \lambda_f/r^2 \tag{15}
\]

We observe real endoscopic images with different focal lengths. In Fig. 8, we have two kinds of endoscopes E1 and E2 and four focal lengths F1, F2, F3, and F4. We use a circle to fit the boundary of the endoscope in the image. Given distortion parameter of one image in an endoscope, the ratio of two diameters of circles is used to predict the distortion parameter of the other image. We compare the linear prediction model with the derived quadratic prediction model in Eq. (15), and found the quadratic prediction is much better than naive linear prediction. The detailed information is summarized in Table I.

---

**Fig. 8. Real endoscopic images with different focal lengths.**

### Table I

<table>
<thead>
<tr>
<th>Image</th>
<th>(\lambda)</th>
<th>Diameter of Circle</th>
<th>Linear Prediction</th>
<th>Quadratic Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1.F1</td>
<td>-2.9644e-5</td>
<td>210 (px)</td>
<td>-1.8931e-5</td>
<td>-2.8126e-5</td>
</tr>
<tr>
<td>E1.F2</td>
<td>-1.2742e-5</td>
<td>312 (px)</td>
<td>-1.9935e-5</td>
<td>-1.3430e-5</td>
</tr>
<tr>
<td>E2.F3</td>
<td>-7.5446e-6</td>
<td>432 (px)</td>
<td>-3.0577e-6</td>
<td>-5.6195e-6</td>
</tr>
<tr>
<td>E2.F4</td>
<td>-1.6638e-6</td>
<td>~ 712 (px)</td>
<td>-4.1052e-6</td>
<td>-2.2338e-6</td>
</tr>
</tbody>
</table>

---

In the previous examples, we extract the boundary of an endoscope to estimate the focal length. There are two practical issues about extracted boundary, completeness and existence of boundary. First, sometimes extracted boundary of the circle is not complete due to cropping by image sensor. For incomplete boundary of circle, the area of visible region is another good measure instead of diameter of circle and linear prediction can be applied. If the boundary is not available, a simple modification for endoscopes can realize the estimation step. For example, a small marker could be put behind the fisheye lens for online recognition.

The misalignment is another issue. The center of zoomable device and the distortion center of a frontend fisheye lens may not be perfectly aligned. The distorted centers are located in a line at different focal lengths. To overcome this problem, we can linearly interpolate distortion center \((c_x, c_y)\) by using two sets of calibration parameters with two nearest focal lengths.

### IV. THE SYSTEM OF REAL-TIME ZOOMABLE WIDE ANGLE LENS CORRECTION

In this section, we describe the whole distortion correction system. First, the distortion correction system with fixed focal length is introduced. Then, the system is extended to a zoomable lens. Finally, we briefly describe how we achieve real-time image distortion correction by using GPU.

#### A. Fixed Focal Length

The system consists of two parts, calibration and image correction. The calibration could be operated offline. We just need take a single image which contains the calibration pattern, and then perform the proposed Hough-based autocalibration procedure. The distortion parameters are used for correcting lens distortion. The pattern could be a chessboard or parallel lines containing at least 8 lines and at most 20 lines, and the captured image is expected to contain curved lines near the boundary of visible regions.
B. Calibration for A Zoomable Lens

For extending the system to a zoomable lens, two components, focal length estimation component and distortion parameter table, are added into the system as shown in Fig. 9. The input image is binarized by an intensity threshold, and we use the structural analysis [37] for finding the outermost contour on the binary image. First, a raster scan is applied for marking some starting points. Then, the boundary is traced from starting points. An open source implementation for boundary tracing in OpenCV is used in this work. The process can achieve 535 fps at 720-by-540 image size. For images of very high resolutions, regular image downsampling by nearest interpolation can be applied first and the boundary detection is performed on the smaller image to reduce the computational cost. The outlined image boundary can be used to calculate the diameter of the circle and the area of visible regions. The focal length is related to the diameter of the circle for the image boundary as given in Eq. (15).

In order to construct the distortion parameter table, we first calibrate several images with different focal lengths. The distortion parameters, distortion centers, associated diameters, and associated area values are recorded. For an incoming frame of endoscopy, we first perform boundary detection to find the distortion parameter and distortion center with the closest diameter or with two closest areas in the table. If we select parameters with two closest areas, the distortion parameter is linearly interpolated. Elsewise, the ratio of diameter is served as the ratio of the focal lengths and the distortion parameter is calculated by quadratic prediction in Section III.

C. Correction Part

Given distortion parameters, the image has to be corrected in real time. The operation involved in the distortion correction algorithm is parallelizable. It consists of computing the corrected pixel coordinates for all pixels and bilinear interpolation. However, the large amount of computation makes real-time processing on CPU impossible. We use GPU to speed up the computation. First, the corresponding coordinates are computed on GPU in parallel. Secondly, we use texture memory to store distorted image. Because the distorted image is just used for read-only and all threads can efficiently access the texture memory, the bilinear interpolation can also be accomplished in parallel. The traditional lookup table (LUT) implementation is efficient but it requires a large amount of memory. For a zoomable lens, many lookup tables have to be stored for the LUT approach. In this work, we do not store these coordinates in advance. Instead, we compute the corrected image coordinates with GPU on the fly.

V. EXPERIMENTAL RESULTS

A. Quantitative Measure and Related Methods for Calibration

We implement Wang’s line-based method [23], Hough Transform-based method for Radial Lens Distortion Correction (HTRDC) [30], the measure in 1-D angular Hough method (HT1D) [31], CHT measure [27] for experimental comparison. The single image calibration method EasyCam-Calib proposed in [7] needs to extract the corners on the chessboard and cannot handle fronto-parallel configuration. In our dataset, about half of the cases (47.62%) were failed due to noise, specular highlights, low illumination, or orientation of plane of calibration pattern. Hence, we mainly compare our method with line-based methods and Hough-based methods.

We use synthetic datasets and real images captured from fisheye lenses and wide-angle endoscopes in our experiments. Wang’s method requires the user to select lines. In HTRDC, the user needs to select a region of interest that contains a single curve. The performance of these two methods depends on the user selection of lines. Our automatic method is based on Hough transform to aggregate the undistorted edge points and uses the undistorted image gradient to discount many error-prone votes. The results show the accuracy and efficiency of the Hough entropy method is considerably improved by properly using the image gradient information. Even if the lines are carefully selected in the other two methods, our method can provide comparable results.

For evaluating different methods, we select several curved lines in the distorted images. After the correction, the undistorted points located in the same line are fitted with a straight line by using the total least squares. The root mean square of the distance from the undistorted points to the regression line is served as the quantitative measure RMSE. Our Hough entropy method with gradient estimation is denoted by HEwG, and a version without gradient estimation (HEwoG) is served as a baseline method for comparison. The version that incorporates prior of PLPs is denoted by HEPLP. For a fair comparison of different distortion correction methods based on HT1D, HEwoG, CHT, HEwG and HEPLP, all the corresponding measures are based on the division model and optimized in the same hierarchical framework.
B. Synthetic Datasets

We also generate two synthetic datasets, syn1c (ground truth \( \lambda = -5.25e - 6 \)) and syn2c (ground truth \( \lambda = -8.85e - 6 \)). Because HTRDC needs to know the distortion center in advance, we compare the RMSE for different distortion centers. The distortion center is given by \((x + \Delta, y + \Delta)\) where \((x, y)\) is correct distortion center. The RMSE of HEwG and HTRDC are shown in Fig. 11. In our method, we try to use different sizes of Sobel operator to reduce the gradient estimation error. Our experiments show that the proposed method in conjunction with 21 \(\times\) 21 Sobel filters gives the best results.

C. Real Datasets

We have used five different wide-angle lenses to capture the dataset; namely s1c, s2c, s3c, s4c, and s5c. There are 21 images in total, and results for an image in s5c are shown in Fig. 13. All averaged RMSE are shown in Table III and Table II. In Table II, correct distorted centers are given for a fair comparison, because the HTRDC method assumes that the distortion center is given. We compare our method with Wang’s method, HTRDC, HT1D, and CHT. The results show the line selection is very critical to the results of Wang’s and HTRDC methods. However, our method does not require any user interaction, since we use all the edge points. The experimental results show our automatic methods HEwG and HEPLP produced similar results compared to those of Wang17, which requires the user manually select 17 lines on the image. As shown in Fig. 10, the performance of HT1D is degraded if the perspective effect appears in undistorted images, and both the proposed HEwG and HEPLP methods can provide pretty good correction results. The patterns used here are chessboards, and the HEPLP method benefits from the assumption of the parallel-line pattern. In Table III, we compare four methods for determining the radial distortion parameters in the division model. The results show our methods HEwG and HEPLP are very stable and provide the most accurate results when the distortion center is unknown.

d. Efficiency of Calibration and Image Correction

We have tested our distortion correction system on the computer with Intel CPU (i7-860) and Nvidia GPU (C2050). Because the number of edge point will influence the time of calibration, we test on images with different numbers of edge points. The execution time of the proposed calibration algorithm and the image correction procedure for different computing platforms is summarized in Table IV. By combining the GPU and CPU in our implementation of the correction process, the image correction can be accomplished in real time for the undistorted image size of 2847-by-1900.
calibration pattern is chessboard or parallel-line-pattern, the proposed method that incorporated prior will further improve the accuracy. Experimental results show the proposed method provides accurate distortion correction results in comparison with the state-of-the-art distortion correction methods.

The Hough entropy method is used in our distortion correction system to calibrate a wide-angle lens, and then online correct the distorted streaming video. Furthermore, we extend the system to a zoomable lens. We derive the relationship between the focal length and distortion parameter for the division model. Based on this relationship, we sample some different focal lengths (typically 3 or 4) and use quadratic prediction to achieve good approximation.

In this paper, we have shown convincing results and demonstrated it is possible to correct wide-angle optical distortion from images acquired from a zoomable lens in real time. Real-time distortion correction has more applications than offline correction. The accuracy of the calibration and efficiency of image correction in the proposed system are demonstrated on a CPU+GPU computing platform.

The main limitation of our method is that the endoscope boundary must be present in the images. In the future, we would like to incorporate other focal length estimators to relax the limitation and make the system more flexible. The proposed distortion correction algorithm takes advantage of the parallel computing power in GPU to achieve real-time performance. It needs to be modified for other computing platforms with different computational architectures for real-time applications.

VI. CONCLUSION

We proposed a Hough entropy method combined with gradient estimation to correct radial distortion. Our method is automatic and the accuracy is comparable to or better than other line-based methods which require manual selection of lines. The state-of-the-art Hough-transform-based distortion correction method focuses on a region containing a single line. In order to consider all curved lines, we use the entropy in the Hough space as the straightness measure. However, naive extension is not practical because the entropy is easily influenced by the presence of multiple lines. We solve this degenerate problem by using the image gradient estimation. The estimated gradient is derived from the division model and is used to remove the error-prone votes. When the

E. In-vivo Evaluation

We also perform in-vivo evaluation for our system as shown in Fig. 12 and Fig. 14. We provide evaluation by ten endoscopy-related professionals, including surgeons and biomedical engineers. Original video streams and undistorted streams are displayed side by side for subjective comparison. The participants compare them and evaluate the sharpness, lightness, resolution, blurring, and distortion of the video streams before and after applying the proposed distortion correction procedure. Finally, the participants give an overall opinion about the distortion correction. We provide six video clips with different types of camera operations, which contain zoom-in, zoom-out, and camera movements. The results of evaluation are summarized in Table V. The evaluation results show the proposed distortion correction can effectively minimize distortion and preserve the image quality compared to the original videos. Most participants think the correction is helpful for improving the overall image quality. The original videos and corrected videos used in the evaluation are available through the link to http://undistort.erufa.com/

REFERENCES


Fig. 12. Some snapshots of pairs of original (left) and corrected (right) images, including in-vivo evaluation, calibration, and training.
Original  Wang  HTRDC  HEwoG  HEwG  HT1D  HEPLP

Fig. 13. Comparison of distortion correction results on the s5c image by using different methods.

Fig. 14. Several endoscopic images are shown in the first row, and they are captured with different endoscopies and with different focal lengths. The corrected images are shown in the second row.


Tung-Ying Lee (S’08-) received the B.S. and M.S. degrees from the Department of Computer Science, National Tsing Hua University, Taiwan, in 2005 and 2006, respectively.

From 2005 to 2006, he was a member of Multimedia and Knowledge Engineering Laboratory, National Tsing Hua University. He had a research visit to Nara Institute of Science and Technology, Japan, during the summer of 2009. He is currently a Ph.D. candidate in Department of Computer Science at National Tsing Hua University and a member of Computer Vision Laboratory, National Tsing Hua University. His current research interests include medical image restoration, natural image restoration, stereoscopic image processing, computer vision and machine learning.

Tzu-Shan Chang received the B.S. degree from the Department of Information Management, National Chi Nan University, Taiwan, in 2010 and the M.S. degree from the Institute of Information Systems and Applications, National Tsing Hua University, Taiwan, in 2012.

In 2012, she joined ASUSTeK Computer Inc., Taiwan.

Chen-Hao Wei was born in Kaohsiung, Taiwan in 1989. He received the B.S. degree from the Department of Computer Science, National Tsing Hua University, Taiwan, in 2011.

He is currently a M.S candidate in Department of Computer Science at National Tsing Hua University. His research interests include computer vision and image processing.

Shang-Hong Lai (M’95-) received the B.S. and M.S. degrees in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, and the Ph.D. degree in electrical and computer engineering from University of Florida, Gainesville, in 1986, 1988 and 1995, respectively.

He joined Siemens Corporate Research in Princeton, New Jersey, as a member of technical staff in 1995. Since 1999, he returned to Taiwan as a faculty member in the Department of Computer Science, National Tsing Hua University. He is currently a professor in the same department and the director of the Computer and Communication Center in the university. In 2004, he was a visiting scholar with Princeton University. Dr. Lai’s research interests include computer vision, visual computing, pattern recognition, medical imaging, and multimedia signal processing. He has authored more than 200 papers published in the related international journals and conferences. Dr. Lai has been a member of program committee of several international conferences, including CVPR, ICCV, ECCV, ACCV, ICPR, PSIVT and ICME. He has been an associate editor for Journal of Signal Processing Systems since 2010. Moreover, he also served as a guest editor for special issues in Journal of Visual Communication and Image Representation as well as Journal of Signal Processing Systems.

Kai-Che Liu After receiving B.S., M.S., and Ph.D. from National Cheng Kung University in 2001, 2002 and 2005, respectively, he had been working at Industrial Technology Research Institute(ITRI), in the Electronic and Optoelectronics Research Laboratories, first as an engineer and then the assistant manager in the 3D system application division for the research of parallel processing on various processor architectures, and image-processing algorithms for 2D-to-3D conversion, image-based rendering and multi-view imaging. He had also worked on medical image processing and computer aided diagnosis. He has received the Outstanding Research Award four times, the Excellent Technical Paper Award twice and the Technology Promotion Award once in ITRI.

After working in ITRI for several years, he joins the medical image research lab at Asian Institute of TeleSurgery (AITS/IRCAD-Taiwan) as the director in 2010. He has also served as deputy general manager of Ming Shi Co., Ltd., a biomedical company of ShowChwan healthcare system. Moreover, he had been the professor in IRCAD in 2011 and is the faculty in Computational Surgery International Network since 2012. Currently research interesting includes computer assisted surgery, computational surgery, 3D medical system, augmented reality application and surgical robot platform. He was the visiting scholar in CMU and Cornell University.

Hurung-Sheng Wu graduated from National Defense Medical Center, Taiwan in 1982. He had done his clinical research fellow in UCSF USA during 1987–1988. He graduated from Tulane University with Master Degree of Public Health in 2006.

Prof. Wu is the pioneer of robotic surgery in South-East Asia since 2002 and master in NOTES since 2007 in Taiwan. He was the President of the Taiwan Association of Endoscopic Surgery during 2008–2010 and current Dean of Asia Institute Tele-Surgery since 2008.

In addition, he has continued his research leading both clinical and Research & Development in advanced minimally invasive surgery and other aspects of advanced surgical science.