Is a Precise Distortion Estimation Needed for Computer Aided Celiac Disease Diagnosis?

Michael Gadermayr[1], Andreas Uhl[1], and Andreas Vécsei[2]

[1] Department of Computer Sciences, University of Salzburg, Salzburg, Austria
mgadermayr@cosy.sbg.ac.at
[2] St. Anna Children's Hospital, Department of Pediatrics, Medical University Vienna, Vienna, Austria

Abstract. In computer aided celiac disease diagnosis, endoscopes with wide-angle lenses are deployed which induce significant lens distortions. This work investigates an approach to automatize the estimation of the lens distortion, without a previous camera calibration. Knowing the discriminative power of all sensible distortion configurations, the model parameters are estimated. As the achieved parameters are not highly precise, moreover, we investigate the effect of approximative distortion correction on the classification accuracy. Particularly, we identify one simple but especially for certain features highly effective approximative distortion model.

1 Introduction

Celiac disease [9] is an autoimmune disorder which affects the small intestine in genetically predisposed individuals after introduction of gluten containing nutrient. Characteristic for this disease is an inflammatory reaction in the mucosa of the small bowel caused by a dysregulated immune response triggered by ingested gluten proteins. During the course of celiac disease the mucosa loses its absorptive villi and hyperplasia of the enteric crypts occurs leading to a diminished ability to absorb food.

Computer aided celiac disease diagnosis relies on images taken during endoscopy. The deployed cameras are equipped with wide angle lenses, which suffer from a significant amount of barrel type distortion. Especially peripheral image regions are affected. Thereby, the feature extraction as well as the following classification is affected. Distortion correction (DC) techniques are able to rectify the images. However, although the lens distortion can be undone, especially in peripheral regions the images are blurred because of the required interpolation during image stretching [2].

In [2], the authors extensively analyzed computer aided celiac disease diagnosis in combination with lens distortion correction. Various distortion models have been investigated in combination with various interpolation methods and numerous feature extraction techniques. DC in combination with certain features leads to improved classification accuracies. Especially if considering large neighborhoods [3], distortion correction often is advantageous. Interestingly, the

distortion model as well as the interpolation methods do not have a major impact on the classification performance.

In this paper, we focus on two fundamental questions:

- **Automatized Lens Distortion Calibration**
  Usually, for distortion corrected image classification, the distortion parameters must be computed from calibration pattern images (e.g. checkerboard patterns) captured with the respective camera. Having features, which are sensitive to lens-distortion, we try to estimate these parameters in an exhaustive search, maximizing the classification accuracy. Thereby, the accuracies for all sensible combinations of distortion parameters are analyzed.
- **Approximative DC**
  In [2] the authors showed, that the classification accuracy does not decrease in case of simple DC models. In opposite, they argue that the converse might well be so, as the best results are achieved with the simplest model. The analysis of our automated lens distortion estimation shows, that even a slightly wrong calibration does not necessarily lead to significantly worse accuracies. It is important to be aware that although the lens distortions can be rectified precisely, changes in perspective surely cannot be undone. Therefore a highly precise lens-distortion model seems to be an overkill. Another issue which occurs in DC is that the images are distinctly stretched in peripheral regions. Thus the extracted patches are variably blurred (patches in peripheral regions are blurred stronger). Moreover, even a certain patch is variably blurred, which is due to the non-linear distortion correction transform. The facts mentioned so far give rise to the suspicion that an even more simplified model could lead to similar or even enhanced classification accuracies. Such models are introduced and investigated in this work.

The paper is organized as follows: In Sect. 2, the investigated distortion estimation approach is explained. Section 3 introduces three approximative lens distortion correction models. In Sect. 4, experiments are shown and the results are discussed. Section 5 concludes this paper.

1.1 Distortion Correction in Computer Aided Diagnosis

In our analysis, we focus on the quite simple, but effective distortion model introduced in [8]. In this approach, the circular barrel type distortion is modeled by the division model [1]. Having the center of distortion c and the distortion parameter ξ, an undistorted point x_u can be calculated from the distorted point x_d as follows:

$$x_u = c + \frac{(x_d - c)}{||x_d - c||_2} \cdot r_u(||x_d - c||_2) \; . \tag{1}$$

$||x_d - c||_2$ (in the following r_d) is the distance (radius) of the distorted point x_d from the center of distortion c. The function r_u defines for a radius r_d in the distorted image, the new radius in the undistorted image:

$$r_u(r_d) = \frac{r_d}{1 + \xi \cdot r_d^2} \ . \tag{2}$$

In the following, we use $\eta = \frac{1}{\sqrt{-\xi}}$ as distortion parameter which has to be in the interval $]0, \infty[$. A small η corresponds to a strong distortion and vice versa.

1.2 Features for Classification

The following features are utilized in the experiments:

- **Local Binary Patterns** [10] (LBP).
- **Local Ternary Patterns** [12] (LTP).
- **Extended Local Binary Patterns** [7] (ELBP).
- **Rotationally Invariant Local Binary Patterns** [11] (RLBP).
- **Fourier Frequency Bands** [4] (FOURIER):
  A FOURIER feature consists of the mean of a ring of the Fourier power spectrum (the thickness of each ring is 1 pixel).
- **Contrast Feature** [5] (CONTRAST):
  The Haralick feature CONTRAST is calculated from the gray-level-co-occurrence matrix, which is generated for a specific offset lengths and four different orientations (horizontal, vertical and diagonal).

For distortion estimation (see Sect. 2), where distortion sensitive features are required, LBP-like features (LBP, LTP, ELBP and RLBP) are used with 8 neighboring samples and a radius of 4 pixels. CONTRAST is deployed with an offset length of 4 pixels. The FOURIER feature is utilized with a frequency ring reaching from 7 to 8 pixels in the power spectrum. These configurations turned out to profit from a previous distortion correction (i.e. they are sensible to lens distortions). For the evaluation of the investigated approximative DC (see Sect. 3), the finally achieved classification accuracies of LBP-like features with radii reaching from 1 to 4 pixels are averaged, to get more stable results. The same is done with CONTRAST (offset lengths reaching from 1 to 5 pixels) and the FOURIER feature (rings with outer radii from 5 to 9 pixels).

2 Estimating the Lens Distortion Parameters

A quite intesting aspect is, if it is theoretically possible to estimate the distortion parameters from classification rates. Having an image database suffering from barrel-type distortions and taken with the same endoscope, we compute classification accuracies (for certain distortion sensitive features) for all sensible combinations of the distortion parameters. After filtering this 3-dimensional signal, we choose the configuration with the highest achieved accuracy.

As the quite simple division model is used, this is computationally feasible. The parameters η (strength of distortion) as well as $c = (c_x, c_y)'$ (a vector containing the x- and y-coordinate of the center of distortion) must be computed.

In order to evaluate a strategy to estimate the distortion parameters, it definitely is necessary to have a considerable amount of image data. As we have to content with 3 databases (containing 287 images in total), we do not aim in an extended evaluation. However, we would like to find out if there was a strong correlation between the precision of a distortion configuration and the achieved classification performance which different features. So we can make a statement if it potentially makes sense to estimate the distortion in this way and also if a precise distortion estimation is necessary or recommended for an accurate proceeding classification.

2.1 Estimating the Distortion Strength (η)

The first step is, to define a set of all sensible distortion strengths (H) and centers of distortion in x- and y-direction (C_x and C_y). Then for each combination

$$\{(\eta, c_x, c_y) \mid \eta \in H \wedge c_x \in C_x \wedge c_y \in C_y\}\,, \tag{3}$$

the 3-dimensional tensor $R(\eta, c_x, c_y)$ containing classification rates (overall accuracies) for all sensible distortion configurations must be computed. Next for each η, the 2-dimensional signal $R(\eta, C_x, C_y)$ is filtered with a 2-D Gaussian (with $\sigma = 2$) G_{2D} (to suppress noise) and the maximum is calculated. The achieved 1-D signal (in the following $m(\eta)$) comprises the low-pass filtered maximally achieved classification accuracies, separately for each distortion strength. We investigate if there was a correlation between the actual distortion strength η_a evaluated by calibration and the η with the highest overall achieved accuracy $m(\eta)$.

2.2 Estimating the Center of Distortion (c_x and c_y)

In order to get the actual center of distortion, we compute the mean over the first dimension of $R(\{\eta_{a-2}, \eta_{a-1}, \eta_a, \eta_{a+1}, \eta_{a+2}\}, C_x, C_y)$, where a is the index of the actual distortion strength (e.g. η_{a-1} is the next higher distortion strength). This is done to get more stable results. Then this 2-dimensional signal is filtered with a 2-D Gaussian (with $\sigma = 2$) and again the maximum is calculated. The finally achieved center of distortion consists of the indexes c_x and c_y of this maximum.

3 Approximative Lens Distortion Correction

Since the authors in [2] argued, that the simplest distortion model leads to the best classification results, we will investigate if the model can be simplified even more. For celiac disease diagnosis, images patches (128×128 pixels) are extracted from the original images. Consequently, we do not have to rectify the whole image, but only the patches which are significantly smaller. We compare the following strategies to approximatively model barrel-type lens distortions:

- Precise DC using the **Division Model** (DC): This model has been proposed in [1] and has been investigated in [2].

Table 1: The three databases used for our experiments.

	Marsh-0 patches	Marsh-3 patches	Resolution (pixels)
DB 1	58	75	768 × 576
DB 2	57	60	768 × 576
DB 3	9	28	528 × 522

- Approximating an **affine transformation** (DC3): Having the corner points of the non-linearly transformed square patch (during DC), the best fitting parallelogram is computed (in a least squares sense). Finally, the patch is affine transformed according to the computed parameters.
- Approximative **non-uniform scaling** (DC2): Therefore the shear parameters of the affine transformation are discarded and the patch is transformed according to the scale parameters only.
- Approximative **uniform scaling** (DC1): Therefore, the scale parameters are averaged in order to achieve one single scale for uniform scaling.

The precisions of these four approaches decrease from top to bottom. In our experiments these four strategies are evaluated with various features.

4 Experiments

The image test sets used contains images of the *duodenal bulb* taken during duodenoscopies at the St. Anna Children's Hospital using pediatric gastroscopes. In a preprocessing step, texture patches with a fixed size of 128 × 128 pixels were manually extracted. The size turned out to be optimally suited in earlier experiments [6]. In case of distortion correction, the patch position is adjusted according to the distortion function. To generate the ground truth for the texture patches used, the condition of the mucosal areas covered by the images was determined by histological examination from the corresponding regions. Severity of villous atrophy was classified according to the modified Marsh classification scheme [9]. We aim in distinguishing between images of patients with (Marsh-3) and without the disease (Marsh-0). Our experiments are based on three databases as given in Table 1. Each database corresponds to one distinct endoscope with specific distortion parameters. For the analysis on approximative DC, these three databases are merged. Example texture patches are shown in Fig. 1.

For classification, we use the k-nearest-neighbor classifier. This rather weak classifier has been chosen to emphasize on quantifying the discriminative power of the features proposed in this work. To avoid any bias in the results, leave-one-patient-out cross validation is utilized. For the distortion estimation, a set of sensible distortion strengths H and centers of distortion $C = (C_x, C_y)'$ must be chosen. We identified the following sensible $\eta \in H$

$$\begin{aligned} H = \{&40, 42.5, ..., 77.5, 80, 85, 90, 95, 100, 110, ..., \\ &140, 150, 170, 190, 210, 230, 250, 270, 290, M\}_{\times 10} \end{aligned} \tag{4}$$

and the following sensible offsets (in pixels) $c_x \in C_x$ and $c_y \in C_y$

$$C_x = C_y = \{-40, -35, -30, ..., 30, 35, 40\} \,, \tag{5}$$

where $c = (0, 0)'$ represents the actual center of distortion achieved by calibration. The actual η is between 490 and 580, depending on the endoscope (which corresponds to one distinct database).

4.1 Results

Figure 2 shows the discriminative powers (classification accuracy) with varying distortion strengths η for each database and each feature. The curves correspond to vectors $m(\eta)$, defined in Sect. 2.1. These plots show, that most of the chosen features definitely are variant to lens distortions, as the best rates are achieved in case of an $\eta < M$. However, although some features reach their accuracy-peaks near the real η, it is hard to determine a distortion strength from these figures. Quite interestingly, the best discriminative power by tendency is achieved with greater distortion parameters than the actual η (see also the dotted lines in Fig. 2).

Table 2 shows the run of the accuracy with varying centers of distortion c for each database and each feature. Whereas the gray values correspond to the unfiltered mean values (for a better visualization), the centers are computed as explained in Sect. 3 after a 2-D filtering. As with the distortion strengths η, this center of distortion estimation is not precise enough to detect the the centers exactly. The precisions achieved with the different features are quite similar. With none of the features, the center of distortion can be estimated precisely (within 20 pixels) for all databases. Due to the high variances, we do not recommend this center of distortion estimation. Instead, in most cases using the image center is more reliable. In case of our images the distance between the image center and the center of distortion is always below 10 pixels.

The results above raise one question: Why can a wrong calibration lead to improved classification rates? Actually which the exhaustive search, distortion configurations can be found which might be overfitted with respect to the database used. For example, a patch which is originally wrong classified as Marsh-0 (as villi are mistakenly detected), can be stretched and thereby blurred in case of a wrong calibration (e.g. if $\eta << \eta_a$), in order to be classified as Marsh-3. As such

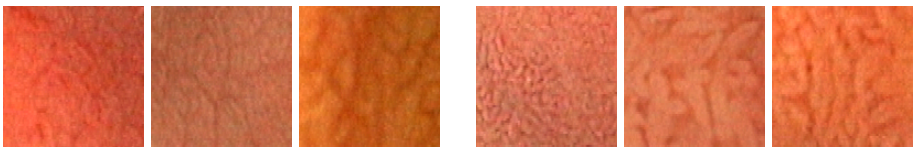


Fig. 1: Example patches of diseased patients (left), showing villous atrophy and healthy patients (right), clearly showing a villous structure.

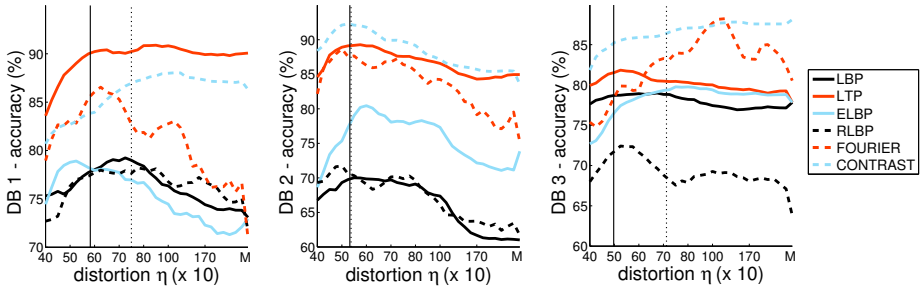
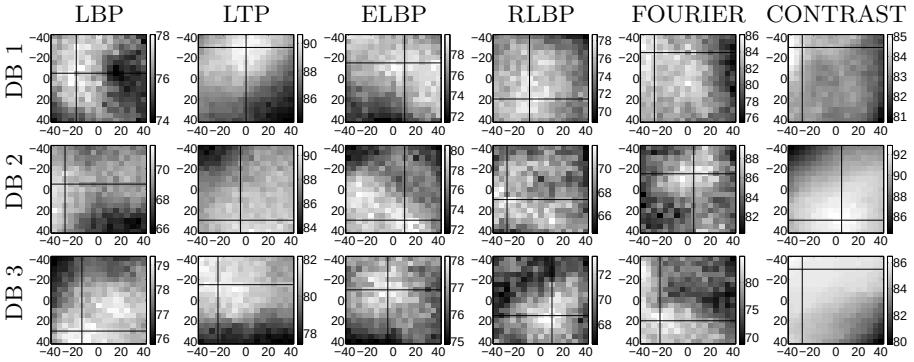


Fig. 2: Maximum classification accuracies for varying distortion strengths η. The solid vertical lines correspond to the (real) η_a acquired by calibration. The dotted lines correspond to the median of the indexes of the highest accuracies (estimated η).

Table 2: Classification accuracies for varying centers of distortion c. The (real) center of distortion is located exactly in the center of the plots. The deviations in x-direction (y-direction) are given on the x-axis (y-axis). The black crosses indicate the computed center of distortion.



overfitting effects commonly can be reduced with larger datasets, we anticipate a more reliable distortion estimation in case of larger image databases.

In Fig. 3 you can see the impact of the introduced approximative distortion models, in comparison to the division model. First of all, the performances of the approach based on distorted images (D), with a slight Gaussian blurring (small σ) in common are highest, but decrease with increasing σ. This is due to the fact that blurring partially compensates the artifacts introduced within distortion correction. In case of concentrating on features with larger neighborhoods (e.g. LBP-like features (LBP, LTP, ELBP) with a radius of four pixels), this curve would be much lower. Especially with the LBP-like features, the simple approximative DC1 approach, which is based on uniform scaling, delivers on average the best rates. We assume, that this is because these texture based features suffer more from inconsistencies (e.g. variable blurring) than from lens

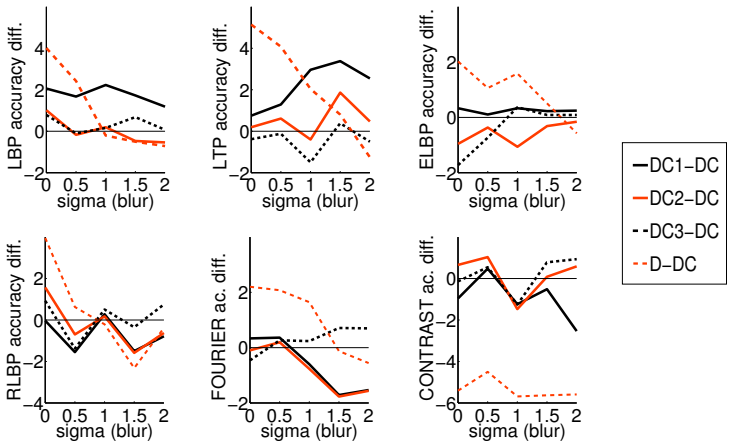


Fig. 3: The accuracy-benefits (and disadvantages) of approximative distortion models in comparison to the division model. The classification rates are given relatively to the approach based on the division model (DC). A positive value indicates a benefit compared to DC.

distortions. For the other features, especially for large σ the more precise DC3 method seems to be slightly advantageously, particularly if compared to DC1 and DC2 (especially with FOURIER features). Considerable decreases on average are only recognized with the FOURIER features and rough approximative methods DC1 and DC2 in combination with a significant blurring.

5 Conclusion

We conclude that a highly precise distortion estimation does not necessarily enhance the classification accuracies even with distortion sensitive features. Thus it is hard to make a good lens-distortion estimation based on the gathered classification rates, particularly in case of a small image databases. However, such an estimation for a proceeding classification might be sufficient, as the classification rates do not necessarily suffer from a (slightly) wrong calibration. Especially the center of distortion can be assumed to be at the image center. The introduced simplified distortion models lead to good accuracies compared to the more precise division model. Particularly for LBP-like features and the simplest distortion approach, enhancements can be observed.

Acknowledgments

This work is partially funded by the Austrian Science Fund (FWF) under Project No. 24366.

References

1. Fitzgibbon, A.W.: Simultaneous linear estimation of multiple view geometry and lens distortion. In: Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR'01). pp. 125–132 (2001)
2. Gadermayr, M., Liedlgruber, M., Uhl, A., Vécsei, A.: Evaluation of different distortion correction methods and interpolation techniques for an automated classification of celiac disease. Computer Methods and Programs in Biomedicine 112(3), 694–712 (Dec 2013)
3. Gadermayr, M., Liedlgruber, M., Uhl, A., Vécsei, A.: Problems in distortion corrected texture classification and the impact of scale and interpolation. In: Proceedings of the International Conference on Image Analysis and Processing (ICIAP'13). pp. 513–522 (Sep 2013)
4. Gadermayr, M., Uhl, A., Vécsei, A.: Barrel-type distortion compensated fourier feature extraction. In: Proceedings of the International Symposium on Visual Computing (ISVC'13). pp. 50–59 (Jul 2013)
5. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Transactionans on Systems, Man, and Cybernetics 3, 610–621 (Nov 1973)
6. Hegenbart, S., Kwitt, R., Liedlgruber, M., Uhl, A., Vécsei, A.: Impact of duodenal image capturing techniques and duodenal regions on the performance of automated diagnosis of celiac disease. In: Proceedings of the International Symposion on Image and Signal Processing and Analysis (ISPA'09). pp. 718–723 (Sep 2009)
7. Liao, S., Zhu, X., Lei, Z., Zhang, L., Li, S.: Learning multi-scale block local binary patterns for face recognition. In: Advances in Biometrics, pp. 828–837 (2007)
8. Melo, R., Barreto, J.P., Falcao, G.: A new solution for camera calibration and real-time image distortion correction in medical endoscopy-initial technical evaluation. IEEE Transactions on Biomedical Engineering 59(3), 634–44 (2012)
9. Oberhuber, G., Granditsch, G., Vogelsang, H.: The histopathology of celiac disease: time for a standardized report scheme for pathologists. European Journal of Gastroenterology and Hepatology 11, 1185–1194 (Nov 1999)
10. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recognition 29(1), 51–59 (Jan 1996)
11. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution Gray-Scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7), 971–987 (2002)
12. Tan, X., Triggs, B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. IEEE Transactions on Image Processing 19(6), 1635–1650 (2010)