

©Springer Verlag. The copyright for this contribution is held by Springer Verlag. The original publication is available at www.springerlink.com.

Distortion Adaptive Image Classification - an Alternative to Barrel-Type Distortion Correction

Michael Gadermayr¹, Andreas Uhl¹, and Andreas Vécsei²

¹ Department of Computer Sciences, University of Salzburg, Austria
{mgadermayr, uhl}@cosy.sbg.ac.at

² St. Anna Children's Hospital, Endoscopy Unit, Vienna, Austria

Abstract. The endoscopes, utilized in computer aided celiac disease diagnosis are equipped with wide-angle lenses, which introduce significant barrel-type distortions. Previous work is on rectifying the distortions prior to the feature extraction. However, due to new arising inadequacies this often even decreases the classification accuracy. The idea of this paper is based on the fact, that there is a correspondence between the position of a patch in the image and the amount (and orientation) of lens distortions. Therefore, in order to classify an image patch, only a reduced training set of similarly distorted patches is considered. We show that in most cases with the new approach higher classification rates can be achieved compared to traditional distortion corrected and uncorrected image classification.

Key words: Endoscopy, classification, celiac disease.

1 Introduction

Celiac disease is an autoimmune disorder that affects the small bowel in genetically predisposed individuals of all age groups after introduction of gluten containing food. Characteristic for this disease is an inflammatory reaction in the mucosa of the small intestine caused by a dysregulated immune response triggered by ingested gluten proteins of certain cereals, especially against gliadine. During the course of the disease the mucosa loses its absorptive villi and hyperplasia of the enteric crypts occurs leading to a diminished ability to absorb nutrients.

Computer aided celiac disease diagnosis relies on images taken during endoscopy. The employed cameras are equipped with wide angle lenses, which suffer from a significant amount of barrel-type distortions. Whereas the distortion in central image pixels can be neglected, peripheral regions are highly distorted. Thereby, the feature extraction as well as the following classification is compromised. Based on camera calibration, distortion correction (DC) techniques are able to rectify the images. However, although the barrel-type distortion can be undone, especially in peripheral regions there remains a lack of information, as the DC method stretches the image. The lack of information has to be compensated using an interpolation technique.

In recent studies [1, 2], the impact of barrel type distortions and distortion correction on the classification rate of celiac disease endoscopy images has been investigated.

In [1], the authors have shown that patches in peripheral regions, which are more affected by the distortions are more likely to be misclassified. Furthermore, the higher the distortion difference between a patch and its nearest-neighbor patch the more likely a patch gets misclassified. With distortion correction, the classification rate on average even suffers. In [3], different distortion correction techniques have been investigated. The computer aided celiac disease diagnosis [1–3] is based on 128×128 pixel patches, which are manually extracted from reliable regions in the original images (with e.g. 768×576 pixels). A computer-aided detection of sensible patches is a separate problem definition.

In this paper, the focus is not on distortion correction, but on increasing the classification performance (i.e. the overall classification accuracy), by considering the position of the respective image patches. As the position of a patch correlates with the orientation and strength of the barrel-type distortions, the classifier is enabled to adaptively handle variably distorted patches. In the following, our new approach is called Distortion-Adaptive-Classification (DAC). Actually, DAC is not limited to barrel-type distortions. It can also be applied in case of other systematic image degradations in combination with patch-based image classification, which has become an important field of research (e.g. see [4]).

The paper is organized as follows: In Sect. 2, our new approach is introduced and compared with distortion corrected and uncorrected classification. In Sect. 3, experiments are shown and the results are discussed. Section 4 concludes this paper.

2 Distortion-Adaptive-Classification (DAC): Classification Based on Patch Position Information

2.1 Barrel-Type Distortions

Fig. 1a shows a checkerboard pattern, captured with an endoscope used in celiac disease diagnosis. Especially in peripheral regions, significant distortions can be recognized. It can be seen, that e.g. the scale (i.e. the size of the squares) in these regions considerably decreases.

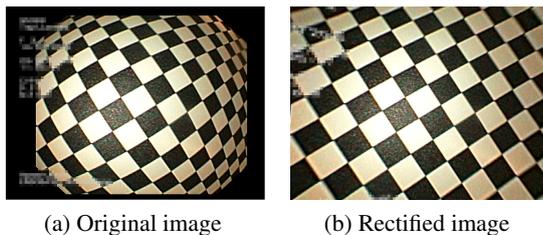


Fig. 1: Distorted and undistorted image of a planar checkerboard pattern.

A general model for barrel-type distortions is given by

$$x_u(x_d) = \hat{x}_c + \frac{(x_d - \hat{x}_c)}{\|x_d - \hat{x}_c\|_2} \cdot r_u(\|x_d - \hat{x}_c\|_2), \quad (1)$$

where x_u is an undistorted point, x_d is a distorted point and \hat{x}_c is the center of distortion. The function r_u , which maps a distance from the center of distortion in the distorted image ($\|x_d - \hat{x}_c\|_2$) to a distance in the undistorted image, is modeled differently. For comparisons in this work, the distortion correction approach introduced in [5], which turned out to be appropriate for our purpose, is utilized. In this approach, r_u is modeled by the division model.

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

Other well known approaches are based on polynomial or parameter free models. Figure 1b shows an undistorted image corresponding to the distorted image shown in Fig. 1a. As can be seen, the geometrical properties are effectively rectified.

2.2 New Approach – an Alternative to Distortion Correction

In [1], the authors concluded, that there is a correlation between the misclassification probability and the radial distance (see (5)) to the nearest neighbor patch, in case of classifying barrel-type distorted celiac disease images. That means, if the patch with the smallest feature distance (i.e. the nearest neighbor patch) has a large radial distance, the patch is more likely to be misclassified. Although distortion correction techniques are able to rectify the geometrical image properties quite well (see Fig. 1), problems within computer aided diagnosis cannot be fully eliminated, as new difficulties arise within the distortion correction. One major problem is that the image has to be distinctly stretched, especially in peripheral regions. This leads to (variably) blurred images.

The basic idea of our new approach is based on the fact, that there is a correspondence between the position of a patch in the image and the amount as well as the orientation of the lens distortions. Therefore, in order to classify a certain image patch, an individually reduced training set of similarly distorted patches is generated. Patches with a high geometrical distance to the current patch show another level of distortions, and therefore are not included in the training set.

For the geometrical distance between two patches, we consider their center points p_1 and p_2 in the original (distorted) image. We identified the following potentially sensible metrics:

– Euclidean Distance:

The quite simple and commonly known Euclidean distance is given by:

$$d_E = \|p_1 - p_2\|_2. \quad (3)$$

– Distortion Level Euclidean Distance:

The Euclidean distance does not incorporate the strengths of distortions, depending on the patch position (peripheral regions vs. center regions). This metric utilizes undistortion to get higher distances in peripheral regions which correspond to higher distortions.

$$d_{DLE} = \|x_u(p_1) - x_u(p_2)\|_2. \quad (4)$$

– **Radial Distance:**

Patches with a similar distance to the center of distortion are similarly distorted as far as its strength is concerned (see (1): r_u only depends on the distance to the center of distortion \hat{x}_c). The idea of this metric is, to totally ignore the distortion orientation and to focus on the strength only.

$$d_{RAD} = \text{abs}(\|p_1 - \hat{x}_c\|_2 - \|p_2 - \hat{x}_c\|_2) . \quad (5)$$

– **Distortion Distance:**

This metric is based on the geometrical shape of the patch in case of undistortion. Similarly warped patches have small distances and vice versa. For simplicity, we only consider the diagonal lengths which are good indicators for distortions. Experiments with other indicators (e.g. using the parameters of an approximated affine matrix) did not lead to significantly different results. The diagonals $diag_i$ are achieved by adding different offsets to the patch center point p and undistorting these corner points (s denotes the side length of a patch (in our case 128)).

$$d_{DD}(p) = |diag_1(p_1) - diag_1(p_2)| + |diag_2(p_1) - diag_2(p_2)| . \quad (6)$$

$$diag_1(p) = \|x_u(p + (\frac{-s}{2}, \frac{-s}{2})) - x_u(p + (\frac{+s}{2}, \frac{+s}{2}))\|_2 . \quad (7)$$

$$diag_2(p) = \|x_u(p + (\frac{+s}{2}, \frac{-s}{2})) - x_u(p + (\frac{-s}{2}, \frac{+s}{2}))\|_2 . \quad (8)$$

For DAC, one distance metric and one threshold must be chosen. The training set for a certain patch consists of all patches with a distance below this threshold. All other patches are simply ignored during classification. We have defined 20 sensible thresholds (e.g. for the Euclidean metric between 80 pixels and infinity (infinity means, DAC is disabled)) for each distance metric. For each feature, the best threshold has been evaluated during exhaustive search.

The classification of images suffering from barrel-type distortions is only one application scenario for DAC. Actually, the new approach is potentially sensible in each case of classifying images with systematic image degradations. With such degradations, we refer to the problem of having variable (average) image properties over the original image where the patches are extracted from. For example, within endoscopy a source of light mounted on top of the endoscope could cause a slightly varying exposure over the image. Features being not invariant to illumination, might suffer from such slight differences, and could profit from DAC. A crucial issue is the identification of an adequate distance metric. The metrics defined in (3) - (6) have been identified to be potentially useful in case of barrel-type distortions. However, especially the Euclidean distance (3) is expected to be good choice in general.

2.3 Features for Classification

In order to investigate the effects of the proposed approach we utilize a couple of texture and shape features. Instead of evaluating the best configuration for each feature, we are computing classification rates with various configurations in order to make a general statement on the effect of our new approach.

The following feature extraction methods are investigated:

- **Local binary patterns** [6] (LBP):
LBP is used with varying radii (r) and a varying number of neighbors (s). The feature vector’s dimensionality is 2^s .
- **Extended local binary patterns** [7] (ELBP):
ELBP, which is an edge based LBP derivative, is used with varying radii and a varying number of neighbors.
- **Local ternary patterns** [8] (LTP):
LTP is used with varying radii and a varying number of neighboring samples and a fixed threshold ($\Theta = 3$), which turned out to be a good choice for various configurations. Although the authors proposed a coding scheme to get only $2 \cdot 2^s$ dimension, to get a very high dimensional feature for our experiments, we implemented LTP with 3^s dimensions.
- **Contrast** [9] (CONTRAST):
The feature vector consists of the Haralick contrast feature [9] calculated for different offsets $(0, r)^T$, $(r, 0)^T$, $(r, r)^T$, $(r, -r)^T$ (i.e. the feature has 4 dimensions). Varying ranges r are investigated.
- **Fourier power spectrum** (FOURIER):
After computing the Fourier power spectrum, rings with varying radii and a thickness of 2 pixels are extracted and the means of these values is calculated. For our experimental usage, we only consider the discriminative power of single rings. (i.e. the dimensionality of a single feature is 1).
- **Shape Curvature Histogram** [10] (SCH):
SCH is a shape feature, especially developed for celiac disease diagnosis. A histogram contains the occurrences of the contour curvature values. In our experiments, we consider various histogram bin numbers (the bin number corresponds with the dimensionality of the feature).

3 Experiments

3.1 Experimental Setup

The image test set used contains images of the *duodenal bulb* taken during duodenoscopies at the St. Anna Children’s Hospital using pediatric gastroscopes (with resolution 768×576 pixels). In a preprocessing step, texture patches with a fixed size of 128×128 pixels were extracted in a manual fashion. The size turned out to be optimally suited in earlier experiments on automated celiac disease diagnosis [11]. Before features are extracted, all patches are converted into gray value images. Using additional color information, no significant improvements are achieved. In case of distortion correction, the patch positions are 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 of biopsies from the corresponding regions. Severity of villous atrophy was classified according to the modified Marsh classification in [12].

Although it is possible to distinguish between the different stages of the disease, we only aim in distinguishing between images of patients with (Marsh 3A-3C) and without the disease (Marsh 0). We decided for this strategy, as the two classes case is more

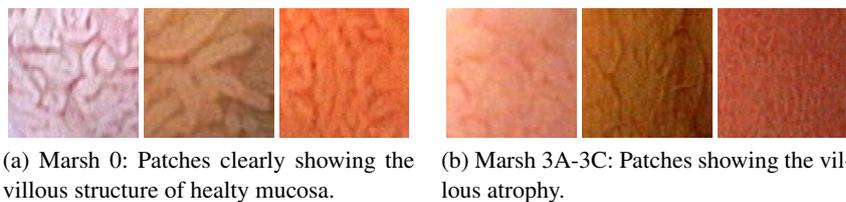


Fig. 2: Example patches of patients without (a) and with the disease (b).

relevant in practice. Our experiments are based on a database containing 135 (Marsh 0) and 115 (Marsh 3A-3C) images, respectively. Example texture patches are shown in Fig. 2. To study the effect of our approach on the classification rate (accuracy), we use leave-one-patient-out cross validation combined with the nearest-neighbor classifier. This rather weak classifier is chosen, as its results can be easily interpreted (see [1]). In an extended future work, the impact of various classifiers will be investigated.

3.2 Results

In Fig. 3, an overview about our results is given. The bars indicate the average improvements with a certain feature and a certain distance metric in comparison to the approach being not based on DAC or distortion correction. The higher (white-colored) top of the bars represent the results, achieved if the threshold is optimized separately for each feature and each distance metric. The lower bars (colored in shades of gray) indicate the mean above the best rate and its 4 neighbors (2 neighbors with smaller and 2 with higher threshold) as far as the distance-threshold is concerned. In order to avoid over-fitting, we will primary focus on these lower values. The quite simple Euclidean distance metric seems to be the best choice, as for each feature it delivers the best of at least highly competitive results (if considering either the higher or the lower rate). Although the differences are quite small, we decided to choose the Euclidean metric for a more detailed comparison.

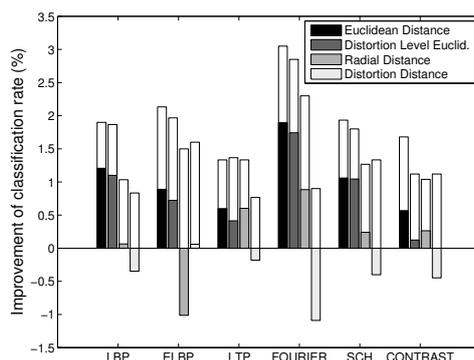


Fig. 3: Overview about the features and the distance metrics.

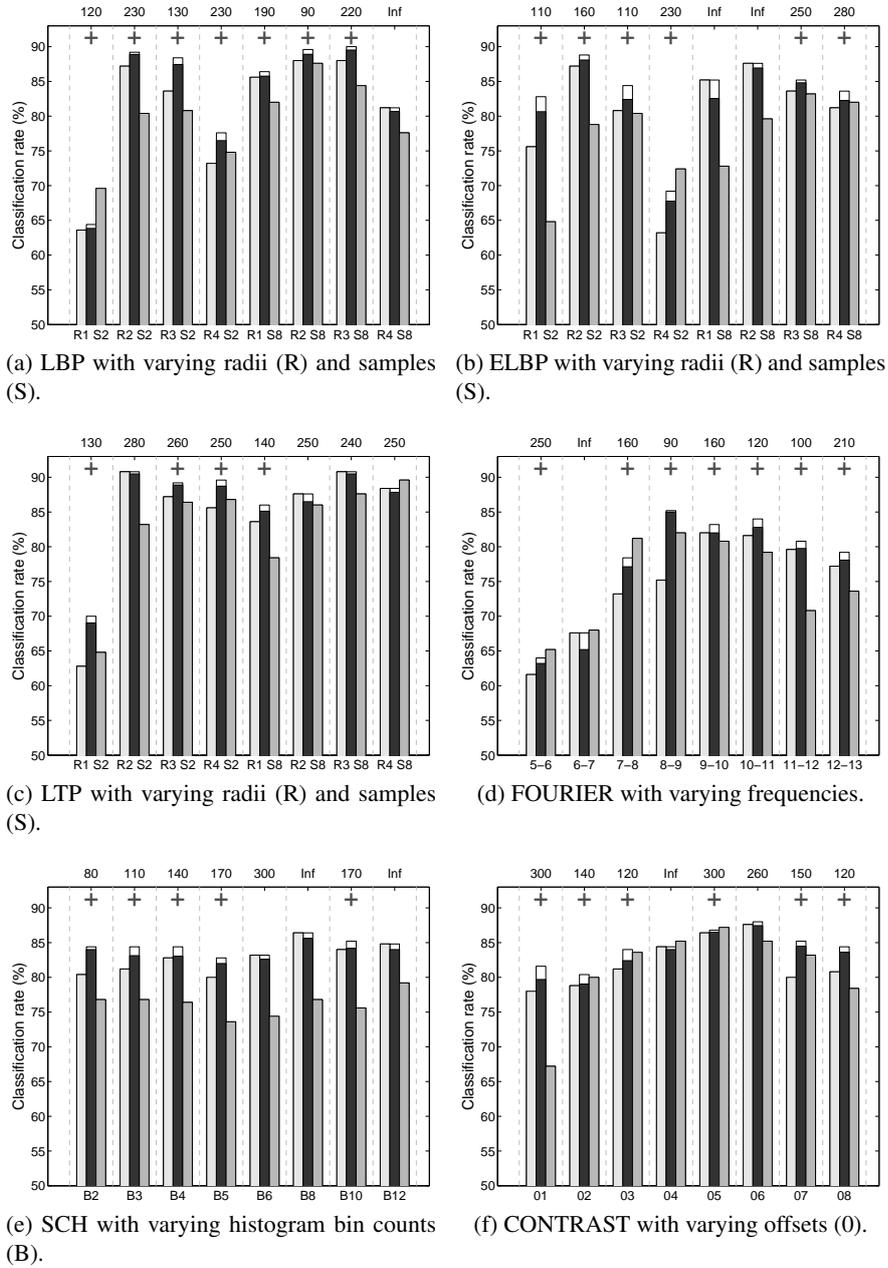


Fig. 4: For each feature and each configuration, the center bar shows the classification rates achieved with our new DAC approach in comparison to the traditional approach without (left bar) and with distortion correction (right bar).

In Fig. 4, the overall classification accuracy results achieved with varying configurations for each feature are visualized. For comparability, the left of the three bars shows the classification rate achieved without distortion correction (noDC) and the right bar shows the rates achieved with distortion correction (DC). The center bar shows 2 different rates achieved with our new DAC approach. The higher (white-colored) top of the bar indicates the best rate achieved with the Euclidean distance metric and varying thresholds. The threshold which corresponds to this rate is shown on top of each column. The lower (dark-colored) bar indicates the mean above the best rate and its 4 neighbors (2 neighbors with smaller threshold and 2 with higher threshold) as far as the distance-threshold is concerned. As in the overview, we will focus on the lower value. A large higher value can be achieved in case of a deceptive search-space. With the averaged lower value, this problem is circumvented. An indicator for the deceptiveness of the respective search-space is given by the difference between the higher and the lower rate. In our plots, a “+” indicates features which profit from our new approach (i.e. the lower rate is greater than (>) the rate achieved with traditional classification).

With LBP and ELBP, especially with the smaller sample size 2, in each case an improvement to noDC can be observed. With 8 samples, only with specific configurations, DAC turns out to be advantageous. The best overall classification rate can be improved with the location consideration. Considering LTP, our new approach seems to be slightly less competitive (especially in combination with samples size 8). The best overall classification rate cannot be outreached. With FOURIER, in most cases an increase of the discrimination power can be achieved and the best overall classification rate can be improved. Although DC also improves the results (in case of low frequencies), our approach obtains the best overall rates. SCH profits from our approach in case of a small histogram bin size (from B2 to B5). The best overall rate cannot be improved. CONTRAST slightly profits in most cases. However, the best overall rate cannot be improved, considering the lower classification value of DAC.

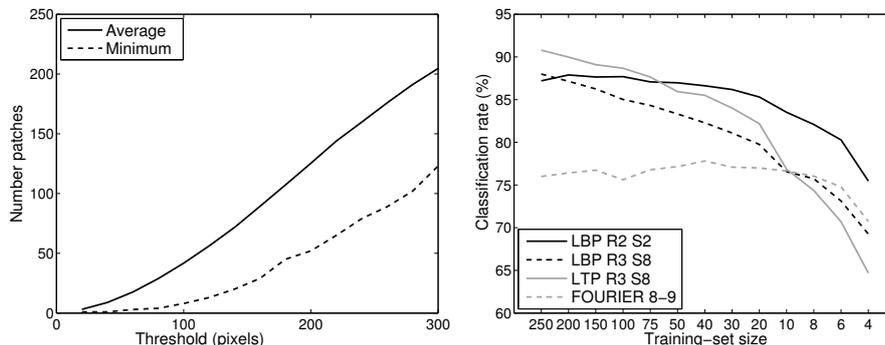
3.3 Discussion

For all features, considering specific configurations, our approach can improve the discriminative performance, however, with some configurations no benefit can be achieved.

With LBP-like features, by tendency with a lower number of samples, the improvement with our method is higher. A similar behavior can be observed with the histogram bin count and the SCH feature. Seemingly, high dimensional features suffer from DAC. Especially with the very high dimensional LTP feature in combination with 8 samples, mostly an increase cannot be achieved. In opposite the 1-dimensional FOURIER features on average lead to the highest benefit of DAC (see Fig. 3).

High dimensional features might suffer because with DAC the training-dataset size is decreased (depending on the chosen threshold). In Fig. 5a, the average and the minimum considered patch counts are given for variable thresholds. Whereas the average count for e.g. a distance threshold of 100 pixels is quite acceptable, the minimum count (which typically corresponds with peripheral patches), is very low.

Figure 5b shows for some example features the average decrease of the classification rate, in case of reducing the training-set randomly (in opposite to DAC where the count is reduced intelligently). Therefore, we averaged the classification rates achieved



(a) This plot shows the average and the minimum considered patch count, dependent on the chosen threshold (the Euclidean distance is used).

(b) Classification rates (y-axis) achieved with example features and shrinking randomly chosen training-sets (the sizes are given on the x-axis).

Fig. 5: The decrease of the training-set size (a) with DAC and the impact in case of a random (instead of an intelligent) decrease (b).

with 30 randomly chosen image sets for each training-set size. We notice that especially, the high dimensional features (LTP and also LBP R3 S8) highly suffer from a reduced training-set. In opposite, with the low-dimensional FOURIER and LBP R2 S2, the classification rate suffers less distinct. As expected, a randomly decreased training-set size on average leads to a loss of discriminative power. This confirms that the increases of the classification rates with DAC are definitely not due to advantages of small training-set sizes. On the contrary, having a larger overall training-set, the patch counts (Fig 5a) could be increased and therefore we anticipate an improvement with DAC even with high dimensional features.

Another quite interesting aspect is to separately investigate features operating in a smaller and others operating in a larger neighborhood. Distortion correction, by tendency profits from larger neighborhood sizes (see LBP, LTP or ELBP with R4 (compared to R1) and FOURIER with small frequencies (compared to large frequencies)). Larger neighborhoods (i.e. lower image frequencies are considered) are more affected by barrel-type distortions on the one hand and less affected by interpolation on the other hand, which is a problem within DC. With DAC, this effect is at least less distinct which is obvious as it is not based on distortion correction in combination with interpolation.

To put it in a nutshell, with DAC and our dataset the classification rate of 24 out of 27 features (88.9 %) with up to 4 dimensions and 29 out of 36 features (80.6 %) with up to 16 dimensions can be robustly improved.

4 Conclusion

We introduced a Distortion-Adaptive-Classification approach, which considers the position of a patch (which corresponds to the distortion level) in the original image for choosing patches for classification. For most features with a traceable dimensionality,

the achieved classification rates are highly competitive. Whereas the traditional classification of barrel-type distortion corrected images leads to benefits only in few cases (especially if lower image frequencies are considered), with DAC in most cases an increased discriminative power can be observed. With very high dimensional features (especially LTP with 8 neighboring samples), we do not achieve systematically improvements. We anticipate an even more advantageous behavior (even for high dimensional features) in the case of a larger image database.

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

References

- Liedlgruber, M., Uhl, A., Vécsei, A.: Statistical analysis of the impact of distortion (correction) on an automated classification of celiac disease. In: Proc. of the 17th Intern. Conf. on Digital Signal Processing (DSP'11), Corfu, Greece (2011)
- Gschwandtner, M., Liedlgruber, M., Uhl, A., Vécsei, A.: Experimental study on the impact of endoscope distortion correction on computer-assisted celiac disease diagnosis. In: Proc. of the 10th Intern. Conf. on Information Technology and Applications in Biomedicine (ITAB'10). (2010)
- Gschwandtner, M., Hämmerle-Uhl, J., Höller, Y., Liedlgruber, M., Uhl, A., Vécsei, A.: Improved endoscope distortion correction does not necessarily enhance mucosa-classification based medical decision support systems. In: Proc. of the Intern. Workshop on Multimedia Signal Processing (MMSp'12). (2012) 158–163
- Häfner, M., Liedlgruber, M., Puespoek, A., Thomas, S., Schoefl, R., Wrba, F., Gangl, A., Uhl, A.: Computer-assisted pit pattern analysis of colonic lesions. *Gastrointestinal Endoscopy* **63** (2006) AB247 – AB247
- 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 Trans Biomed Eng.* **59** (2012) 634–44
- Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distribution. *Pattern Recognition* **29** (1996) 51–59
- Liao, S., Zhu, X., Lei, Z., Zhang, L., Li, S.: Learning multi-scale block local binary patterns for face recognition. In: *Advances in Biometrics*. (2007) 828–837
- Tan, X., Triggs, B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. In: *Analysis and Modelling of Faces and Gestures*. Volume 4778. (2007) 168–182
- Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. *IEEE Trans. on Systems, Man, and Cybernetics* **3** (1973) 610–621
- Gadermayr, M., Liedlgruber, M., Uhl, A.: Shape curvature histogram: A shape feature for celiac disease diagnosis. Technical Report 2012-06, Dept. of Computer Sciences, University of Salzburg, Austria, <http://www.cosy.sbg.ac.at/research/tr.html> (2012)
- 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: Proc. of the 6th Intern. Symp. on Image and Signal Proc. and Analysis (ISPA '09). (2009) 718–723
- 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** (1999) 1185–1194