Semantic and Topological Classification of Images in Magnetically Guided Capsule Endoscopy

P. W. Mewes[1,2*], P. Rennert[2*], A. Lj. Juloski[2], A. Lalande[3], E. Angelopoulou[1], R. Kuth[2], J. Hornegger[1]

[1]Pattern Recognition Lab, University of Erlangen-Nuremberg, Erlangen, Germany
[2]Siemens Healthcare, Erlangen, Germany
[3]Groupe Imagerie Médicale, LE2I, Faculté de Médecine, Université de Bourgogne, Dijon, France
[*] These authors contributed equally to this work

ABSTRACT

Magnetically-guided capsule endoscopy (MGCE) is a nascent technology with the goal to allow the steering of a capsule endoscope inside a water filled stomach through an external magnetic field. We developed a classification cascade for MGCE images which groups images in semantic and topological categories. Results can be used in a post-procedure review or as a starting point for algorithms classifying pathologies. The first semantic classification step discards over-/under-exposed images as well as images with a large amount of debris. The second topological classification step groups images with respect to their position in the upper gastrointestinal tract (mouth, esophagus, stomach, duodenum). In the third stage two parallel classifications steps distinguish topologically different regions inside the stomach (cardia, fundus, pylorus, antrum, peristaltic waves) and semantically different viewpoints relative to the stomach wall (close-up, mid-range and panorama view). For image classification, global image features and local texture features were applied and their performance was evaluated. We show that the third classification step can be improved by a bubble and debris segmentation because it limits feature extraction to discriminative areas only. We also investigated the impact of segmenting intestinal folds on the identification of different semantic camera positions. The results of classifications with a support-vector-machine show the significance of color histogram features for the classification of corrupted images (97% accuracy). Features extracted from intestinal fold segmentation lead only to a minor improvement (3%) in discriminating different camera positions.

Keywords: Magnetically-Guided Capsule Endoscopy, Content-Based Image Retrieval, Computer-Aided Diagnosis, Semantic Classification, Upper Gastro-Intestinal Tract

1. INTRODUCTION

Stomach cancer causes nearly 10% of all cancer deaths worldwide, surpassed only by lung cancer. Stomach cancer is particulary frequent in Asia. 43% of the Japanese population suffers from atrophic gastritis,[1] one of the main causes of stomach cancer,[2] and about 20% of the Korean population suffers from gastroesophageal reflux disease.[3] These numbers suggest a high demand for a precise, fast and cost effective way of screening the population for signs and symptoms of cancer, since early detection of cancer increases the survival rate of the patients.[4] The standard procedure for screening and examination of the upper GI tract is the esophagogastroduodenoscopy (EGD).[5] EGD is performed under full consciousness or moderate sedation of the patient and has a relatively low risk of complications (1 out of 1000 cases).[6] However, patient's fears of embarrassment, discomfort, and worries about endoscopic procedures are common.[7,8] These concerns, in conjuction with a lack of adressing patients anxiety, may decrease the willingness of the patient to undergo an EGD procedure.[8,9]
An alternative to EGD is barium x-rays of the upper digestive tract.[10] The so-called barium swallow is used for the examination of the esophagus and can detect esophageal diseases such as achalasia, Schatzki's ring, gastroesophageal reflux disease or cancer of the esophagus.[11–14] Barium swallow is also deployed in the examination of

Philip Mewes: philip.mewes@informatik.uni-erlangen.de, Peter Rennert: P.Rennert@cs.ucl.ac.uk

the stomach and duodenum where it can help in the diagnosis of hiatal hernia, gastric carcinoma and polyps.[15–17] In comparison, EGD provides more in-depth information and outperforms barium x-rays in terms of sensitivity for the detection of esophagitis, gastritis, or small ulcers.[18,19] Today modern barium computed tomography (CT) and magnetic resonance imaging (MRI) using barium or other contrast agents may compete with EGD.[20,21] However, for the upper gastrointestinal tract CT and MRI techniques are costly, time consuming, may cause allergic reaction caused by the contrast agent and do not allow biopsy. A drawback for all x-ray techniques is the radiation dose the patient is exposed to.

Capsule Endoscopes (CE) were introduced in 2000 by Given Imaging*. CE is aimed primarily for small bowel examination which is hard to reach with other types of endoscopy. CE has also been applied to the examination of the colon and esophagus.[22–24] In CE the patient swallows a capsule which then acquires images while travelling through the gastric tract by the natural intestinal peristaltic motion. It is used in case of obscure bleedings or whenever a disease (e.g. Crohn's disease or peptic ulcers) is suspected in the small intestine.

Classical CE is not indicated for stomach examination because the capsule can not be controlled, which makes it impossible to reliably cover the entire surface of the stomach. Magnetically guided capsule endoscopes (MGCE) were introduced in 2010 by Siemens Healthcare† and Olympus Medical Systems Corporation‡ making it feasible to control a capsule endoscope in a water-filled stomach. Two clinical studies show the feasibility of gastric exploration with MGCE.[25,26]

Computer-aided diagnosis for CE and MGCE was previously presented in.[27–35] In[27–29] different descriptors for the task of blood and ulcera detection and topographic segmentation of the GI tract are investigated. Topographic segmentation is addressed in[30] and[31] for the purpose of a more efficient and faster review. In[32] and[33] the issue of eliminating redundant frames, as well as those with intestinal juices, is addressed. In[36] a set of color and texture based features for the detection of intestinal bleedings is presented. These methods are not directly applicable to MGCE. The duodenum, when compared to the stomach, exhibits different pathologies and imaging conditions, such as texture and distance to objects of interest. Camera viewpoints and distance to objects of interest in CE is limited. Furthermore CE exams are conducted in an empty and narrow gastrointestinal (GI) tract, while MGCE exams are conducted in a liquid-distended stomach.

This paper presents a three-stage classification cascade for sorting MGCE images in different semantic and topological categories. Furthermore the third classification stage is subdivided in two parallel classification tasks. A categorized image can ease a post-procedure review process or be used as a starting point for algorithms classifying pathologies inside the upper GI tract such as in Mewes et al.[34,35]

2. THE ANATOMY OF THE STOMACH

The stomach is part of the upper digestive system, connecting to the esophagus at its upper end and the duodenum at its lower end. The stomach itself can be further divided into three major parts. The upper part, called fundus, the middle part, called stomach body, and the lower part called pyloric antrum or simply antrum. Between the esophagus and the fundus is a sphincter call cardia and at the lower end between the stomach and the duodenum is a second sphincter, called pyloric sphincter (pylorus). The stomach wall is covered by gastric folds (or gastric rugae) to increase the stomach surface area for food. The folds allow the stomach to expand, when food enters, without increasing the pressure in the stomach walls.

3. METHOD

Our classification cascade is divided in three stages (see Fig. 1). In the first step (stage #1) severely corrupted images (over-exposure, under-exposure, too many floating particles and debris) are excluded. In a second step (stage #2) images are classified according to the topology of the upper GI tract (mouth, esophagus, stomach and duodenum). In a third step images of the stomach are further classified: Stage (#3b) topologically classifies images according to the position inside the stomach (cardia, fundus, pylorus, antrum, peristaltic contractions)

*Given Imaging, Yoqneam, Israel
†Siemens Healthcare Sector, Erlangen, Germany
‡Olympus Medical Systems Corporation, Tokyo, Japan

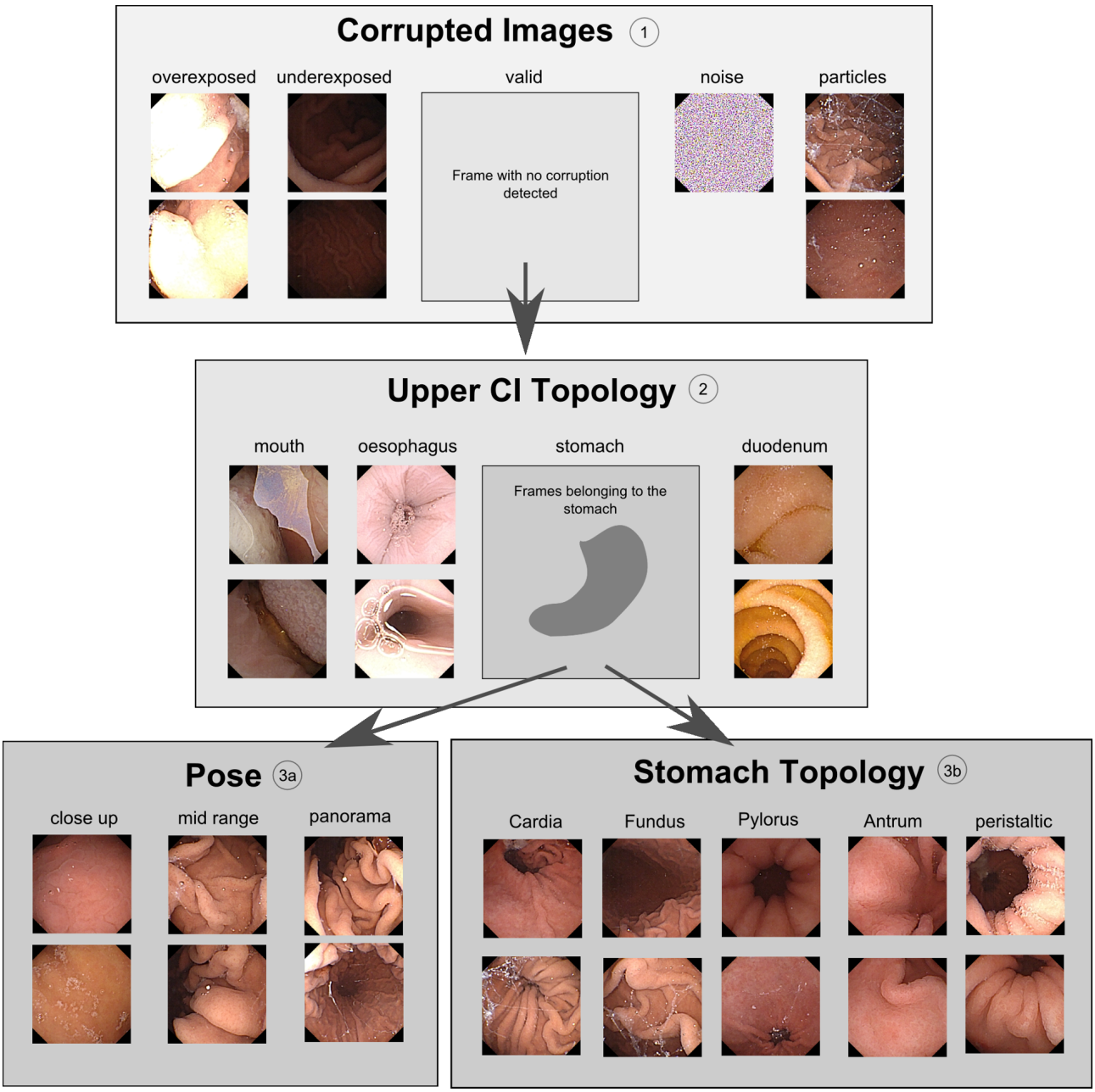

Figure 1: Classification cascade showing the different stages and pipeline of classification, as well as class-specific example images from each stage.

while stage #3a groups images according to the camera viewpoint relative to the stomach wall (close-up, mid-range view, panorama view). The following features were extracted from each image: mean, variance, skewness, kurtosis, energy and entropy for each histogram of an HSV color channel (HSV, 18 features); line-likeliness and directionality of the features as presented by Tamura[37] (LLD, 10 features)(cf. section 3.4); linear binary pattern (LBP, 18 features); stomach fold features (SFF, 201 features)(cf. sections 3.2-3.3). Table 1 shows which features are used at each stage.

3.1 Bubble and Debris Detection

Bubbles and debris appear often in MGCE images. An image region with bubbles usually contains no medically relevant information and corrupts global features, such as color or LBP histograms (Fig. 1).[31] Areas with bubbles have a different texture than the stomach mucosa, the esophagus, the duodenum or the mouth. While the latter regions mostly appear as smooth and intensity-constant regions, areas with bubbles and debris exhibit sharp local changes in contrast caused by transitions from bubble specularities to bubble/mucosa edges and bubble shadows. Based on these observations, we developed the following algorithm (Algorithm 1) for bubble and debris detection:

- Steps 1-2: To distinguish between texture related to image regions with bubbles and texture related to image regions stemming from the rest of the upper GI tract, the response image for the LBP variance operator (8-neighborhood/radius 1)[38] is computed (step 1, Fig 2b). This operator is highly discriminative regarding local changes in contrast and invariant against shifts in gray scale. The resulting variance image is thresholded with an empirically determined threshold (threshold used for this paper is 100) leading to a binary pixel map (step 2, Fig 2c). The input image I and the binary pixel map have the dimension of 431 pixels in width and 432 pixel in height. We denote these dimensions as h and v for the rest of this paper.
- Steps 3-5: Three morphologic operations are performed. Step 3 reduces noise with a small opening operator (Fig 2d). Step 4 connects gaps and openings between debris/bubble segments with a closing operator (Fig 2e). Step 5 removes transitions between stomach folds and flat mucosa that were mistakenly detected as bubbles (Fig 2f). This occurs only for few stomach folds which exhibit a bright illumination close to the camera and a significant dark shadow at the transition to normal mucosa. This high-contrast edge region may be above the empirical threshold applied to the LBP variance image (Step 2) and appears as a thin line on the binary image. Such lines are removed with a small opening element (Fig 2f).
- Steps 6-7: The previously segmented region is limited to small bubble segments by a larger opening operator. The result is subtracted from the original segmented regions of step 5. The remainder of this operation is the small bubble segments, which were previously deleted (step 5, Fig 2g).The binary image is subsequently dilated with an element of the same size (step 7, Fig 2h).
- Step 8: The union of the result of steps 5 and 7 represents the actual bubble area (Fig 2i).

The morphological operations performed in step 6-7 lead to a more conservatively segmented area of bubbles and debris. For a less sensitive version of the algorithm steps 6-7 may be omitted and the segmented area of step 5 is directly dilated with a structuring disk element of size 8 (Algorithm 2 step 6, (Fig 2g)). This latter version may misclassify prominent edges of stomach folds as bubbles, while performing more reliably on the segmentation of edges of bubbles. We refer to the two versions of the Algorithms 1 and 2 as *low* and *high* sensitivity versions. Bubbles detection is applied to stages 2, 3a and 3b of the classification pipeline. Images in classification stage 1 may contain elements other than bubbles and upper GI mucosa such as transmission noise or saturated areas due to overexposure. This stage is therefore not preprocessed with the bubble segmentation algorithm.

3.2 Fold Segmentation

Stomach folds are distinctive visual features in MGCE images that may convey significant information for stages #3a and #3b of the classification cascade. They are oriented along the stomach sphincters from the cardia (part of the stomach attached to the esophagus) towards the antrum to the pyloric sphincter (see section 2 for anatomical description). Thus, the first task in this feature extraction process is to segment the folds in the image. The segmentation focuses on the transitions between the actual fold and the surrounding flat stomach mucosa (Fig.1 stage #3a and stage #3b). These particular regions appear as thin dark-shadowed regions. Algorithm 3 combines results of a Sobel edge detector together with local brightness information to extract these regions based on the following steps:

- Steps 1-2: The gray scale image is convolved with the Sobel kernels. In order to discard weak edges, the edge magnitude image is thresholded with an empirically determined threshold (here 201) leading to a binary image.

Algorithm 1: Bubble and debris detection. Morphologic operations are defined as follows: dilation ($\oplus$), closing ($\bullet$), opening ($\circ$) and disk($\cdot$) denotes a morphologic structuring disk element with a given radius. Algorithm 2: Bubble and debris detection with less sensitivity

Algorithm 1:

Require: I {Input Image, grayscale}

1: $I_{\mathrm{VAR}} \leftarrow \mathrm{VAR}_{(8,1)}(I)$
2: $I_{\mathrm{VAR}_\tau} \leftarrow I_{\mathrm{VAR}} > \tau$
3: $I_{bubbles} \leftarrow \mathrm{disk}(2) \circ (I_{\mathrm{VAR}_\tau})$
4: $I_{bubbles} \leftarrow \mathrm{disk}(11) \bullet (I_{bubbles})$
5: $I_{bubbles} \leftarrow \mathrm{disk}(3) \circ (I_{bubbles})$
6: $I_{smallB} \leftarrow I_{bubbles} - \mathrm{disk}(6) \circ (I_{bubbles})$
7: $I_{smallB} \leftarrow \mathrm{disk}(8) \oplus (I_{smallB})$
8: $I_{bubbles} \leftarrow I_{bubbles} \cup I_{smallB}$

Algorithm 2:

Require: I {Input Image, grayscale}

1: $I_{\mathrm{VAR}} \leftarrow \mathrm{VAR}_{(8,1)}(I)$
2: $I_{\mathrm{VAR}_\tau} \leftarrow I_{\mathrm{VAR}} > \tau$
3: $I_{bubbles} \leftarrow \mathrm{disk}(2) \circ (I_{\mathrm{VAR}_\tau})$
4: $I_{bubbles} \leftarrow \mathrm{disk}(11) \bullet (I_{bubbles})$
5: $I_{bubbles} \leftarrow \mathrm{disk}(3) \circ (I_{bubbles})$
6: $I_{bubbles} \leftarrow \mathrm{disk}(8) \oplus (I_{bubbles})$

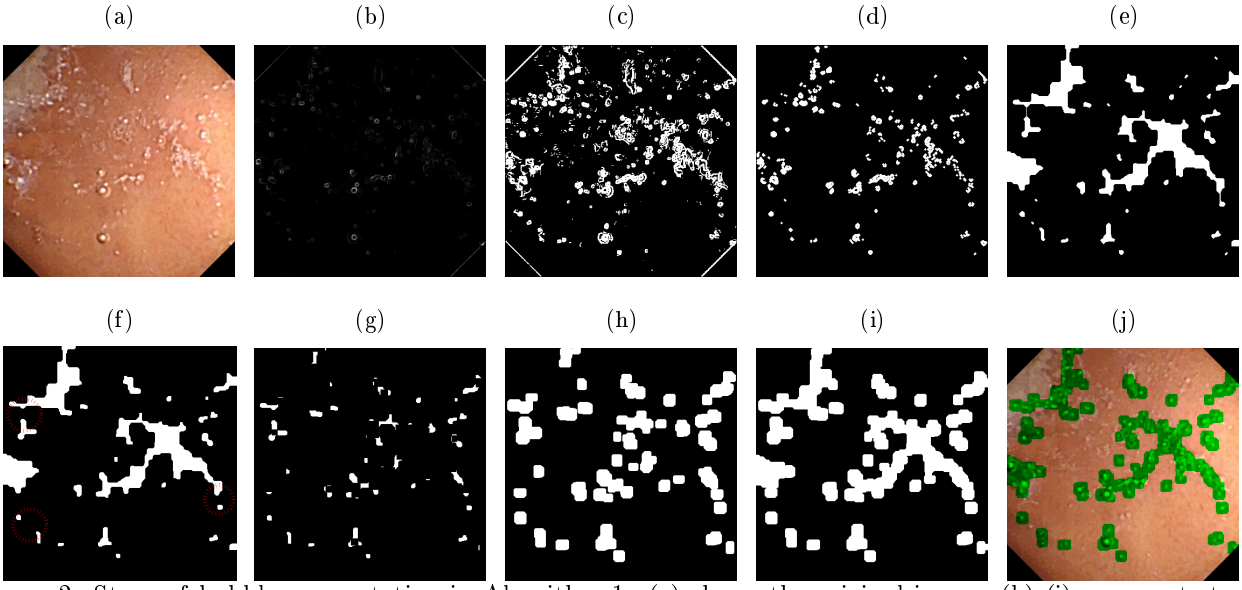

Figure 2: Steps of bubble segmentation in Algorithm 1: (a) shows the original image, (b)-(i) represent step 1 to step 8 of Algorithm 1, where (b) is the response image from the LBP variance operator, images (c)-(i) are binary images, red circles in (f) point out differences between (e) and (f), (j) shows the original image with the superimposed segmented bubble region in green

- Step 3 detects dark pixels in the original image. We compute for each pixel at location (x, y) the mean intensity value at a 15×15 neighborhood around it, $\mathrm{mean}(\mathcal{W}_{x,y}) = \sum_{p \in \mathcal{W}} I_g(p)$ where $I_g(p)$ is the gray-scale value at a pixel p. A pixel is considered dim and thus assigned a value 1 if its value is smaller that the neighborhood average.
- Step 4 computes the intersection of the strong edge mask and the dim pixel mask (i.e. the result of step 2 and 3).
- Steps 5-6: From the intersected binary map 8-connected groups of pixels are computed (Step 5). The size of a pixel group considered for further processing can be limited to a certain size using the threshold ρ. Both versions (with and without threshold) are considered for the fold focus computation, while for the computation of the fold coverage the threshold is not applied (both methods are described in section 3.3). In step 6 the intersected binary mask is element-wise multiplied with the original horizontal and vertical Sobel derivatives ($I_{Sobel(h,v)}$) to obtain the horizontal and vertical derivatives maps for each pixel in the

fold-transition region.

For the sake of simplicity we will refer to the extracted edges of stomach folds simply as stomach folds.

Algorithm 3: Fold segmentation algorithm. card(·) denotes the number of pixels within a 8-connected region

Require: I_g {Input Image, grayscale}
1: $I_{Sobel} \leftarrow I_g * Sobel$
2: $I_{Sobel_\tau} \leftarrow |I_{Sobel}| > \tau$
3: $I_{min} \leftarrow I_{(x,y)} > \text{mean}(\mathcal{W}(x,y))$
4: $I_{intersect} \leftarrow I_{Sobel\tau} \cap I_{min}$
5: $I_{foldMap} \leftarrow \text{card}(\text{8-connected regions}(I_{intersect})) > \rho$
6: $M_{h,v} \leftarrow I_{Sobel(h,v)} \times I_{foldMap}$

3.3 Fold Features (Fold Coverage and Line Focus)

Fold Coverage An indicator of the presence of folds is the percentage of pixels that are set to 1 in the intersected binary map $I_{foldMap}$ (Step 5 in Algorithm 3). This feature solely contains the quantitative presence of folds in the image without information about spatial distribution and orientation. However, the ratio between pixels belonging to folds and pixels not belonging to folds is considered as a feature since quantitative differences in the amount of folds are an indicator for the capsule position relative to the stomach wall. An image captured close to the stomach wall may exhibit fewer folds than an image capturing the complete fundus with many gastric folds.

Fold Focus This feature aims to describe the orientation of folds on the stomach mucosa. From this description, information may be derived about the capsule position relative to the stomach wall (stage #3a) and about the topologic position of the capsule inside the stomach (stage #3b). For instance, if many stomach folds tend to meet in one spot in the image one can assume that the capsule image contains the pylorus or cardia, while a parallel arrangement of the folds, with no obvious vanishing point of the folds, may represent a situation where the capsule is in a perpendicular position to the stomach wall. Stomach folds can be detected as prominent edges or lines in an image. Figure 3 illustrates such segmented lines and their different orientation for two topologically different images.

Stomach fold segmentation is described in Algorithm 3 and leads to a binary pixel map $I_{foldMap}$ that is set to 1 for a pixel considered as part of a stomach fold and 0 if not (Step 5). $I_{foldMap}$ has the same dimensions as the image I_g. We refer to a pixel on this map with the binary value of 1 as $p_i = (p_x, p_y)$ with $i = 1...n$ and n the total number of these pixels. For each of these pixels the horizontal and vertical derivatives of the convolution with the Sobel operator is computed (Step 1 and 6 of Algorithm 3). We denote these maps containing the derivatives as $M_h(x,y)$ and $M_v(x,y)$ for the horizontal(h) and vertical(v) derivatives respectively. The gradient direction for each pixel p_i is denoted as

$$\theta_{p_i} = arctan\left(\frac{M_h(p_x,p_y)}{M_v(p_x,p_y)}\right) \tag{1}$$

Subsequently for each pixel p_i a line $l(x)_i$ going through this pixel is defined as follows:

$$l(x)_i \quad = \quad -\frac{1}{\theta_{p_i}} \cdot (x - p_x) + p_y \tag{2}$$

For the n lines $l(x)_i$ the possible number of line-line intersections is denoted $n_{is} = (\frac{n!}{2!(n-2)!})$ assuming that none of the lines is parallel to each other. Each intersection point is denoted is_j with the coordinates (x,y) and $j = 1...n_{is}$. The function

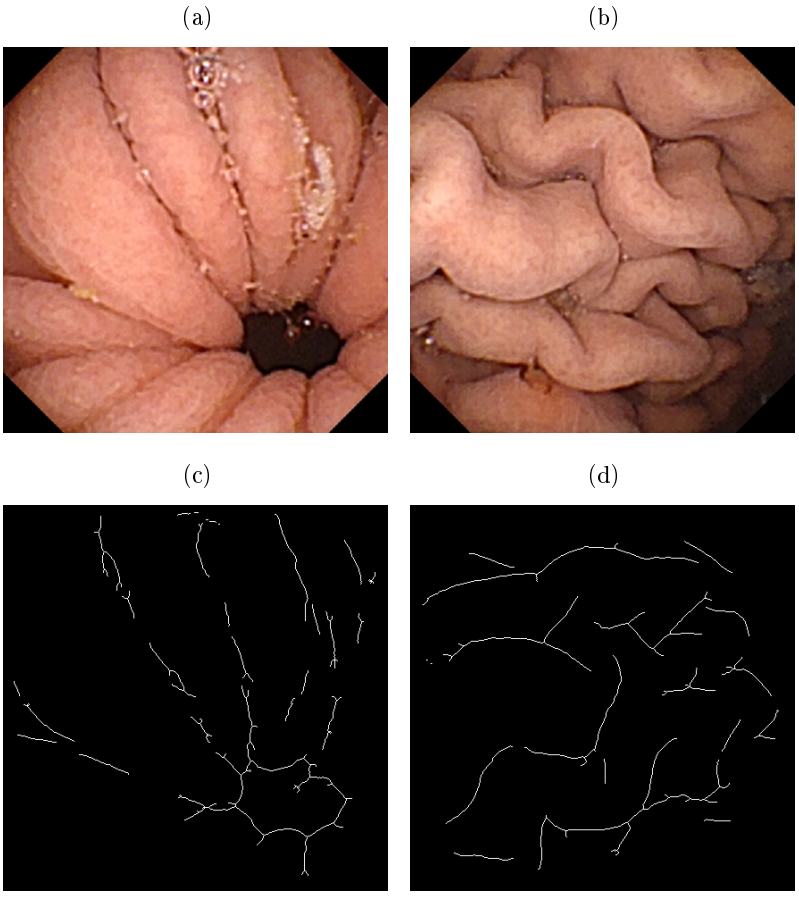

Figure 3: Example of different fold orientations detected by Algorithm 3. (a),(b) Original stomach images: View of the cardia (a) and stomach wall (b). (c),(d) folds extracted from the original images (a) and (b).

$$S(j) = N(l(x)_i \cap l(x)_k) \qquad \forall \quad l(x)_i \nparallel l(x)_k \tag{3}$$

returns the coordinates (x, y) of the line-line intersection point $is_j(x, y)$ for all j unique combinations of two intersecting lines $l(x)_i$ and $l(x)_k$. $N(\cdot)$ rounds the value of x and y to the nearest integer so that $is_j(x, y) \in \mathbb{N}$. The spatial and quantitative distribution of line-line intersection points is translated to a cumulative map cm of the same dimensions (h, v) as the original gray-scale image I_g. Each map pixel $cm_{(x,y)}$ is defined as

$$cm_{(x,y)} = \sum_j \delta(j) \tag{4}$$

with δ an indicator function s.t.

$$\delta = \begin{cases} 1 & \text{if } S(j) = (x, y) \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Thus, cm contains information about the quantitative accumulation of line-line intersections in the image. The mapping to cm is comparable to a parameter space such as the Hough space for circle detection.[39] For each pixel

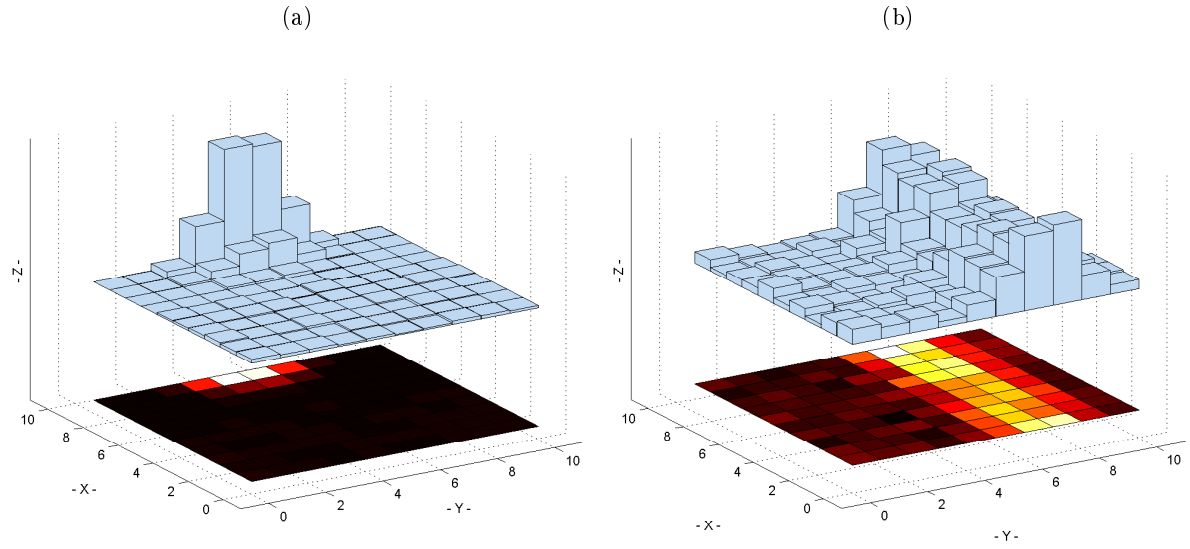

Figure 4: Artificially generated example histograms showing the quantitative occurrences (-Z-) for the accumulated line-line intersections map $cm_{(x,y)}$(-X-) vs. the intensity value of I_g (-Y-). A projected view of bin intensities is shown on the bottom of the image for better visibility. (a) shows a distribution which occurs when many lines meet at few points (high number of occurrences of line intersections in a few bins) The corresponding gray-scale intensities are distributed around the mean of all possible intensities. (b) shows a distribution which occurs when lines meet at many points (uniform number of occurrences of line intersections distributed over all bins). The corresponding gray-scale intensities are distributed around the upper third of the intensity spectrum. The upper third of the intensity spectrum contains dark pixel values. Therefore it can be assumed that the lines meet in dimmer (lower intensity) illuminated background pixels of the image

being part of a stomach fold according to $I_{foldMap}$ a parameter is entered into the parameter space. For cm this parameter is the sum of all lines (Eq. 4) defined for all pixels p_i (Eq. 2) going through a map pixel $cm_{(x,y)}$.

Since cm and I_g have the same size they can both be represented in a combined 2D histogram showing the quantitative occurrences of intersections for a position in cm vs. the intensity value of I_g. The intensity information of the original gray scale image conveys information about the distance of the spatial fold arrangement with respect to the capsule. The more dimly illuminated background results in low intensity values, while the well illuminated foreground results in high intensity values. Figure 4 shows two example histograms illustrating the way information from the previously described method is translated and combined in one 2D histogram. The number of histogram bins on both axes is fixed to 10. Each bin is used as a feature and therefore leads to 100 features per 2D histogram. Since $I_{foldMap}$ can be computed in two ways (step 5 of Algorithm 3) the total number of features computed in section 3.3 is 201 (200 fold focus features + 1 fold coverage feature).

3.4 Tamura Features

Directionality: Implemented as described by Tamura[37] in chap. IV-c. Tamura's original directionality feature is represented as the modes of the histogram of angles (directions) of gradients in the image. Here, instead of representing the feature with the modes, the angle histogram is computed as proposed by Tamura but is represented by six histogram statistics (mean, variance, skewness, entropy, energy, kurtosis).

Line-likeness: Implemented as described by Tamura[37] in chap. IV-d. For the computation of the image co-occurrence matrix the angles 0, $\frac{\pi}{4}$, $\frac{\pi}{2}$ and $\frac{3\pi}{4}$ with the discrete distance of 3 pixels are taken into account. This results in a four-element line-likeness feature. The use of other distances did not have a significant impact

Table 1: Performance of classification stages given in terms of the average accuracy and sensitivity of all classes belonging to a stage. The impact of bubble and debris segmentation (bds) is shown for the stages 2, 3a and 3b.

stage # (# of classes/ # of Images)	Cascade stage	Features (Num. of Features)	bds (Sensitivity High/Low)	Average accuracy	Average sensitivity
1 (2/349)	Corruption	HSV (18)	No	0.97	0.92
2 (4/408)	Upper GI Topology	HSV, LBP (36)	No	0.94	0.90
2 (4/408)	Upper GI Topology	HSV, LBP (36)	Yes (High)	0.94	0.89
3a (3/363)	Camera Pose	HSV, LBP, SFF (237)	No	0.89	0.84
3a (3/363)	Camera Pose	HSV, LBP, SFF (237)	Yes (High)	0.92	0.88
3b (5/361)	Stomach Topology	HSV, LBP, SFF, LLD (247)	No	0.75	0.44
3b (5/361)	Stomach Topology	HSV, LBP, SFF, LLD (247)	Yes (Low)	0.82	0.48

on the classification results and are, therefore, not computed. In order to make the feature rotation invariant, the concatenated line-likeness of all four directions is sorted (circular shift) such that it starts with the least significant number. The total number of Tamura features is therefore ten.

4. RESULTS

Different sets of features described in section 3 were extracted from a dataset of 1481 images. Sets were chosen based on the quality of the described classification results and based on the discriminative potential of individual features. Table 1 summarizes the number of images used for each of the four stages, as well as the exacted composition of previously described features and their total number. To evaluate the discriminative potential of the complete feature-set of each stage classification was performed with support vector machines using radial kernels in a 10-fold leave-one-out crossvalidation. The results of these experiments at each stage are shown in Table 1. The impact of debris and bubble detection on the classification stages 2, 3a and 3b is shown for each of these stages. The best result regarding the low and high sensitivity version of the debris and bubble segmentation algorithm is shown. The average accuracy and average sensitivity is computed for all stages.

5. DISCUSSION

The experimental results show that semantic and topological classification of images stemming from MGCE achieve an accuracy of over 80% for all classification stages. This also holds for the true positive rates of all stages except the topological classification stage 3b. However, images from the cardia, fundus or the pylorus region are difficult to distinguish even for human observers. Deformation of the stomach shape due to peristaltic waves may cause anatomical patterns that can easily be confused with other anatomic regions. Features based on the presence and orientation of stomach folds in an image could improve the results but are strongly influenced by the camera pose. Stage 3a classifies the camera pose with an average accuracy of 92% and an average sensitivity of 88%. Therefore, for future work one should consider shifting the classification stage 3a ahead of the classification stage 3b and use the obtained knowledge as an additional information to solve 3b. However, taking into account the difficulties a human observer has with this classification task there is no guarantee for superior performance of stomach topology classification.

Regarding the bubble and debris segmentation it is notable that this method can improve the classification result for the two classification stages (3a and 3b) inside the stomach, but not significantly for the general topological classification of the upper GI. In these stages (3a and 3b) the classification task becomes more severe. Color characteristics as an obvious discriminative feature for the general topological classification of the upper GI (stage 2) are less substantial in 3a and 3b. Features considering texture and the global fold feature become more important. These features are strongly affected by bubbles and debris which explains the improvement generated by the bubble and debris detector. The use of fold features could improve the sensitivity of the classification result for stage 3b by 2% and by 3% for stage 3a. For stomach topology classification, the bubble segmentation with low sensitivity gave better results than the one with high sensitivity.

6. NEW OR BREAKTHROUGH TO BE PRESENTED

This paper proposes a classification cascade for MGCE images. The presented classification results suggest that techniques from passive capsule endoscopy such as[31] can be applied successfully for some tasks. However, for stomach topology classification they are inadequate for the second and third classification stages. Therefore, we developed an algorithm for fold segmentation and an approach for exploiting this information. We also showed that the detection and exclusion of bubble areas enhanced the classification result especially for stages 2 and 3 of the classification cascade. Our evaluation also showed that different stages of the cascade and different features need different bubble and debris segmentation settings.

7. ACKNOWLEDGMENTS

The authors would like to thank J.F. Rey, I. Pangtay (Institut Arnault Tzanck, Saint-Laurent-du-Var, France), H. Ogata, N. Hosoe, T. Hibi (Keio University School of Medicine, Tokyo, Japan), K. Ohtsuka, N. Ogata, S. Kudo (Showa University Northern Yokohama Hospital, Yokohama, Japan), K. Ikeda, H. Aihara and H. Tajiri (The Jikei University School of Medicine, Tokyo, Japan) for conducting the clinical study and their helpful support. All clinical images have been provided by Dr. Rey and all copyrights of the images are reserved.

REFERENCES

[1] Kimura, K., "Gastritis and gastric cancer. Asia.," *Gastroenterology Clinics of North America* **29**(3), 609 (2000).
[2] Jemal, A., Siegel, R., Xu, J., and Ward, E., "Cancer statistics," *A Cancer Journal for Clinicians* **60**(5), 277 (2010).
[3] Lee, S., Leeb, K. J., Kima, S. J., and Chob, S. W., "Prevalence and risk factors for overlaps between gastroesophageal reflux disease, dyspepsia, and irritable bowel syndrome: a population-based study," *Digestion* **79**(3), 196–201 (2009).
[4] World Health Organisation, "Fact Sheet: Cancer (№297)." online (October 2011).
[5] Ritchie Jr, W., Rhodes, R., and Biester, T., "Work loads and practice patterns of general surgeons in the united states, 1995–1997: a report from the american board of surgery," *Annals of surgery* **230**(4), 533 (1999).
[6] Yusuf, T. and Bhutani, M., "Esophagogastroduodenoscopy," (2011).
[7] Charalampos, M., Nikolaos, D., Marianna, Z., Vassiliki, P., Christos, G., Christos, F., Athanasios, D., Dimitrios, V., Eleni, V., Iannis, Z., et al., "Panic attack during elective gastrointestinal endoscopy," *Gastroenterology Research and Practice* **2011** (2011).
[8] Hackett, M., Lane, M., and McCarthy, D., "Upper gastrointestinal endoscopy: Are preparatory interventions effective?," *Gastrointestinal endoscopy* **48**(4), 341–347 (1998).
[9] Brody, D., "The patient's role in clinical decision-making," *Annals of Internal Medicine* **93**(5), 718 (1980).
[10] Levine, M., Rubesin, S., and Laufer, I., "Barium esophagography: a study for all seasons," *Clinical Gastroenterology and Hepatology* **6**(1), 11–25 (2008).
[11] Pohl, D., Tutuian, R., et al., "Achalasia: an overview of diagnosis and treatment," *Journal of Gastrointestinal and Liver Diseases* **16**(3), 297 (2007).
[12] Groskreutz, J. and Kim, C., "Schatzki's ring: long-term results following dilation," *Gastrointestinal endoscopy* **36**(5), 479–481 (1990).
[13] Irwin, R., French, C., Curley, F., Zawacki, J., and Bennett, F., "Chronic cough due to gastroesophageal reflux. clinical, diagnostic, and pathogenetic aspects.," *Chest* **104**(5), 1511 (1993).
[14] De Oliveira, J., Birgisson, S., Doinoff, C., Einstein, D., Herts, B., Davros, W., Obuchowski, N., Koehler, R., Richter, J., and Baker, M., "Timed barium swallow: a simple technique for evaluating esophageal emptying in patients with achalasia," *American Journal of Roentgenology* **169**(2), 473 (1997).
[15] Kahrilas, P., Kim, H., and Pandolfino, J., "Approaches to the diagnosis and grading of hiatal hernia," *Best Practice & Research Clinical Gastroenterology* **22**(4), 601–616 (2008).
[16] Tunaci, M., "Carcinoma of stomach and duodenum: radiologic diagnosis and staging," *European journal of radiology* **42**(3), 181–192 (2002).
[17] Gordon, R., Laufer, I., and Kressel, H., "Gastric polyps on routine double-contrast examination of the stomach.," *Radiology* **134**(1), 27 (1980).
[18] Dooley, C., Larson, A., Stace, N., Renner, I., Valenzuela, J., Eliasoph, J., Colletti, P., Halls, J., and Weiner, J., "Double-contrast barium meal and upper gastrointestinal endoscopy," *Annals of internal medicine* **101**(4), 538 (1984).
[19] Davidescu, D., "The barium mean examination versus endoscopy," *The Journal of Preventive Medicine* **13**(1-2), 73–80 (2005).
[20] Fornasa, F., Benassuti, C., and Benazzato, L., "Role of magnetic resonance enterography in differentiating between fibrotic and active inflammatory small bowel stenosis in patients with crohn's disease," *Journal of Clinical Imaging Science* **1**(1), 35 (2011).

[21] Siddiki, H., Fidler, J., Fletcher, J., Burton, S., Huprich, J., Hough, D., Johnson, C., Bruining, D., Loftus Jr, E., Sandborn, W., et al., "Prospective comparison of state-of-the-art mr enterography and ct enterography in small-bowel crohn's disease," *American Journal of Roentgenology* **193**(1), 113 (2009).
[22] Iddan, G., Meron, G., Glukhovsky, A., and Swain, P., "Wireless capsule endoscopy," *Nature* **405**, 417 (2000).
[23] Swain, P., Iddan, G., Meron, G., and Glukhovsky, A., "Wireless capsule endoscopy of the small bowel: development, testing, and first human trials," in [*Proceedings of SPIE*], **4158**, 19 (2001).
[24] Iddan, G. and Swain, P., "History and development of capsule endoscopy," *Gastrointestinal Endoscopy Clinics of North America* **14**(1) (2004).
[25] Rey, J., Ogata, H., Hosoe, N., Ohtsuka, K., Ogata, N. Ikeda, K., Aihara, H., Pangtay, I., Hibi, T., Kudo, S., and Tajiri, H., "Feasibility of stomach exploration with a guided capsule endoscope," *Endoscopy* **42**(7), 541–545 (2010).
[26] Rey, J., Ogata, H., Hosoe, N., Ohtsuka, K. a. N., Ogata, N. Ikeda, K., Aihara, H., Pangtay, I., Hibi, T., Kudo, S., and Tajiri, H., "First blinded non-randomized comparative study of gastric examination with a magnetically guided capsule endoscope (mgce) and standard videoendoscope.," *Gastrointestinal Endoscopy* (2011).
[27] Cunha, J., Coimbra, M., Campos, P., and Soares, J., "Automated topographic segmentation and transit time estimation in endoscopic capsule exams," *IEEE T Med Imaging* **27**(1), 19–27 (2007).
[28] Szczypiski, P., Sriram, P., Sriram, R., and Reddy, D., "Model of deformable rings for aiding the wireless capsule endoscopy video interpretation and reporting," *Comp Imag Vis* **32**, 167–172 (2006).
[29] Szczypinski, P. and Klepaczko, A., "Selecting texture discriminative descriptors of capsule endoscopy images," in [*Proceedings of 6th International Symposium on Image and Signal Processing and Analysis, 2009. ISPA*], 701–706, IEEE (2009).
[30] Berens, J., Mackiewicz, M., and Bell, D., "Stomach, intestine, and colon tissue discriminators for wireless capsule endoscopy images," in [*Proceedings of SPIE*], Fitzpatrick, J. and Reinhardt, J., eds., **5747**, 283 (2005).
[31] Mackiewicz, M., Berens, J., and Fisher, M., "Wireless capsule endoscopy color video segmentation," *IEEE Transactions on Medical Imaging* **27**(12), 1769–1781 (2008).
[32] Karargyris, A. and et al., "A video-frame based registration using segmentation and graph connectivity for Wireless Capsule Endoscopy," in [*Life Science Systems and Applications Workshop LiSSA 2009. IEEE/NIH*], 74–79, IEEE (2009).
[33] Vilarino, F., Spyridonos, P., Pujol, O., Vitria, J., Radeva, P., and De Iorio, F., "Automatic detection of intestinal juices in wireless capsule video endoscopy," in [*ICPR 2006*], Tang, Y., Wang, S., Lorette, G., Yeung, D., and Yan, H., eds., **4**, 719–722, IEEE (2006).
[34] Mewes, P. W., Neumann, D., A. Lj. Juloski, Angelopoulou, E., and Hornegger, J., "On-the-fly Detection of Images with Gastritis Aspects in Magnetically-Guided Capsule Endoscopy," in [*SPIE Medical Imaging*], (2010).
[35] Mewes, P. W., Neumann, D., Licegevic, O., J. Simon, J., Juloski, A. Lj., and Angelopoulou, E., "Automatic Region-of-Interest Segmentation and Pathology Detection in Magnetically-Guided Capsule Endoscopy," in [*Conference on Medical Image Computing and Computer Assisted Intervention*], **6893**, 137–144, Springer (2011). [to appear].
[36] Mackiewicz, M., Fisher, M., and Jamieson, C., "Bleeding detection in wireless capsule endoscopy using adaptive colour histogram model and support vector classification," in [*Proceedings of SPIE*], Joseph, M. and Josien, P., eds., **6914**, 69140R (2008).
[37] Tamura, H., Mori, S., and Yamawaki, T., "Textural features corresponding to visual perception," *IEEE Transactions on Systems, Man and Cybernetics* **8**(6), 460–473 (1978).
[38] Ojala, T., Pietikäinen, M., and Mäenpää, T., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**, 971–987 (2002).
[39] Hart, P., "How the hough transform was invented [dsp history]," *Signal Processing Magazine, IEEE* **26**(6), 18–22 (2009).