Preliminary version. To appear in *Proceedings of the 26th International Symposium on Computer-Based Medical Systems (CBMS'13), June 20-22, 2013, Porto, Portugal*

Detection of Circular Content Area in Endoscopic Videos

Bernd Münzer, Klaus Schoeffmann, Laszlo Böszörmenyi
Institute of Information Technology, Klagenfurt University, Austria
{bernd,ks,laszlo}@itec.aau.at

Abstract

The actual content of endoscopic videos is typically limited to a circular area in the image center. This area has a dynamic position and size and is surrounded by a dark, but noisy border. In this paper we present a novel algorithm that (1) classifies which frames of an endoscopic video feature the circular content area and (2) determines its exact position and size, if present. This information is very useful for improving the performance of subsequent analysis techniques. It can also be used for more efficient video encoding and economic printing of still images in findings and reports. The evaluation shows that the proposed method is very accurate, robust and efficient in terms of runtime.

1 Introduction

Endoscopy is a minimally invasive technique for screenings and operations in various regions of the human body. A small camera is inserted into the body and produces a video stream that is displayed on a screen. In recent years, it has become common to record the video streams for documentation purposes and for retrospective analysis.

The actual content of endoscopic videos is typically limited to a circular area in the center of the image due to inherent camera characteristics. The pixels outside the circle are very dark and contain no useful information. Unfortunately, they do not form a perfectly homogeneous black area but are subject to intensive noise. Thus, a certain amount of bandwidth has to be wasted to encode them if the video is recorded. When content-based analysis is carried out, the irrelevant border pixels should be omitted because they would severely bias analysis results. If still images are printed on findings and reports the black border wastes a lot of ink.

To address these problems, we propose a new robust algorithm to exactly determine the parameters of the featured circle (centre coordinates and radius). We can use this information to concentrate further content based analysis exactly on the relevant pixels inside the circle. Furthermore, the border pixels can be superimposed with a homogenous black overlay that can be encoded more efficiently. This mask could also have any other color. For example, it would make sense to use a white overlay mask for images in printed findings to save ink. Moreover, video summarization or browsing interfaces can be designed in an optimized way that minimizes wasted screen space by only showing the relevant content area.

2 Related work

Research in the field of endoscopic image/video analysis is mainly focused on the classification of gastrointestinal polyps, lesions or tumors in the context of CAD (computer aided diagnosis) [3] and real-time analysis of laparoscopic video streams for robotic endoscope and instrument guidance as well as for augmented reality [7].

To the best of our knowledge, the idea of determining the exact parameters of the circle has not been addressed yet. However, many methods could benefit from this information by exactly narrowing down analysis to the relevant area. The naive alternative is to ignore the problem and analyze the whole image, but in that case especially global image features are biased by the irrelevant border pixels. Moreover, analysis algorithms have to work on a larger area than necessary. Other rudimental approaches like ignoring all dark pixels below a threshold (e.g., [6]), cropping the image at empirical boundaries so that it only contains the rectangular center of the circular content area [1] or using a predefined static mask for the content area [8] are not fully satisfying because they either analyze parts of the irrelevant border regions or omit parts of the content area which can severely bias analysis results. Only by using a dedicated circle detection algorithm content based analysis can be limited exactly to the relevant area of the image.

A possible alternative to our approach could be a general-domain circle detector like the well-established generalized Hough transform [2]. It is even able to find multiple circles at arbitrary positions in the image. However, this generic capability (which is not required for our purpose) is quite expensive in terms of processing speed due to its brute-force nature. Moreover, our experiments revealed that its accuracy is not high enough and the circle parameters can only be estimated approximately. On the contrary, our approach exploits domain specific knowledge about the possible position and size of the circle and thus is much faster and more accurate.

3 Circle detection algorithm

Our proposed algorithm consists of a number of steps that are performed for each frame of the video. It is not sufficient to detect the circle for one frame and use the result for the entire video because the circle size and position are not standardized and usually vary over time. Moreover, some endoscopes offer a zoom function that can be used to progressively magnify the video image until the content area covers the entire image. Therefore, the circle detection algorithm also has to detect if the frame has a circular content area at all. Examples for circle frames and zoomed frames are depicted in Figure 1.

First of all, we examine the average intensity and variance of small border samples to find out if the frame is likely to have a dark border at all. If this is not the case, it is classified as zoomed frame and the actual circle detection is not executed. Note that this pre-processing step only filters out obvious zoomed frames but does not guarantee that the frame must contain a circular content area.

For frames with a dark border, an edge image is obtained with the Canny edge detector. The first edge pixel from the left and from the right in n rows are considered as edge points. In an optimal case all edge points lie on the contour of the circle. However, the edge image is not perfectly reliable because it can have gaps due to dark areas at the periphery of the circle in the original image. Furthermore, heavy border noise can cause edges to be detected outside the circle. If the frame actually is a zoomed frame that slipped through the black border detection, the points occur at random positions. Examples for typical edge point distributions are depicted in Figure 1.

Each possible combination of three edge points is used to calculate a circle candidate by means of basic geometry. Each of the candidates is checked for plausibility. Here we exploit the domain knowledge about the valid position and radius that are roughly known. Figure 2 depicts three different circle candidates for the second frame of Figure 1. The first two circles are based on two valid edge points and one outlier while the third circle is plausible. Valid edge points are highlighted with a white, outliers with a red circle.

To detect outliers among the edge points we calculate a confidence value $c(e) = cp(e)/cc(e)$ where $cp(e)$ is the number of plausible circle candidates based on e and $cc(e)$ is the total number of circle candidates based on e. If $c(e)$ is considerably smaller than the median confidence of all edge points or smaller than an absolute threshold, e is regarded as invalid. All circle candidates that are based on one or more invalid edge points are discarded.

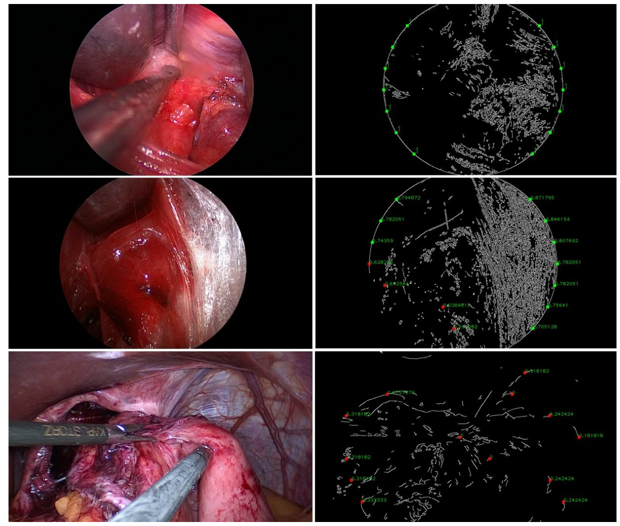

Figure 1. Edgepoints and confidence values

In case of a zoomed frame, the random distribution of edge points leads to a high number of unplausible circle candidates and further to very low confidence values for the edge points rendering them invalid. If the ratio of remaining plausible circle candidates to the original total number of circle candidates is below another threshold, the frame is classified as zoomed frame. Figure 1 depicts the results of the edge point filtering for three different frames. For the first example frame no edge points have to be filtered because all of them are on the circle and therefore have a plausibiltiy of 1. In the figure, red points represent invalidated edge points while green points represent valid edge points. The number beside the points denotes the respective plausibility.

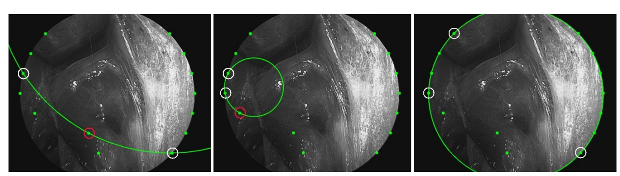

Figure 2. Examples for circle candidates

The remaining circle candidates are compared to the edge image and the candidate with the highest match is selected as final circle. Finally, a temporal filtering is conducted on the sequence of frame classifications and circle parameters to correct outliers. This step assures that each sequence of equally classified frames has a minimum length (e.g., 100 frames) and occasional outliers in terms of position or size are smoothed. A more detailed description of the algorithm can be found in [4].

4 Evaluation

We evaluated our algorithm with five hours of laparoscopic surgery videos. This testset includes 271,161 circle frames and 168,447 zoomed frames. In terms of classification (circle frame (positive) against zoomed frame (negative)), the algorithm achieved a precision of 99.9992%, a sensitivity of 97.7%, a specificity of 99.9988% and an accuracy of 98.57%. The detailed results are summarized in Table 1. A closer investigation revealed that the few misclassifications were caused by very dark, irrelevant frames. That means that for relevant frames, the accuracy is 100%.

Table 1. Circle detection accuracy

		Classification	
Ground truth		Circle	Zoomed
Circle	271,161	264,863	6,298
Zoomed	168,447	2	168,445

In terms of circle parameter accuracy, the average deviation from the ground truth is 0.16% for the x coordinate, 0.44% for the y coordinate and 0.25% for the radius. For a full HD video (1920x1080) this corresponds to an average error of 1-2 pixels which is neglectable. Not a single detected circle had a deviation of more than 2% from the ground truth so we can state that our algorithm is highly accurate both in terms of classification and determination of circle parameters.

To ensure the robustness of the algorithm in terms of avoiding false positives we also tested it without black border detection and outlier correction which led to the same result. In a further analysis run on the TRECVID 2012 general video data set [5] consisting of 8,263 videos with a total length of 200 hours no false positives were incorporated into the final result.

In terms of runtime the algorithm was able to process 161 fps (frames per second) on average when executed on an off-the-shelf Intel Xeon 2.4 GHz processor. Most of the processing time is consumed by the pre-processing operations (edge detection, blur filtering and down-scaling) that could be reused by further analysis techniques. The runtime of the actual detection algorithm is almost negligible (about 2%). Further evaluation details can be found in [4].

5 Conclusions and future work

In this paper we have presented a robust and fast algorithm for differentiating between frames of endoscopic videos with the typical circular content area and zoomed frames. In the former case, the exact parameters of the circle are determined. This information is an important input for further content based analysis algorithms to narrow down analysis to the relevant parts of the image. Moreover it can be used for improving encoding efficiency, for economic printing of findings in terms of ink consumption and for optimizing content visualization in summaries etc.

The evaluation shows that our algorithm is highly accurate and reliable, especially in terms of avoiding false positives. Furthermore, it can also be executed in realtime scenarios. In offline applications it can be used as efficient preprocessing step, especially if the required pre-processing operations (scaling, blur filtering, edge detection) have to be conducted anyway and can be reused.

Future work includes an evaluation of the actual impact on coding efficiency. Moreover, the proposed method will be used as a pre-processing step for further analysis algorithms that we are currently working on. We also plan to evaluate to what extent the consideration of the actual content area improves the performance of these techniques. Furthermore, it should be investigated and quantified to what extent printing ink can be saved if the border region is ignored for printed findings.

References

[1] L. Alexandre, N. Nobre, and J. Casteleiro. Color and position versus texture features for endoscopic polyp detection. In *International Conference on BioMedical Engineering and Informatics, 2008. BMEI 2008*, volume 2, pages 38–42, May 2008.

[2] J. Illingworth and J. Kittler. A survey of the hough transform. *Computer Vision, Graphics, and Image Processing*, 44(1):87–116, Oct. 1988.

[3] M. Liedlgruber and A. Uhl. Computer-aided decision support systems for endoscopy in the gastrointestinal tract: A review. *Biomedical Engineering, IEEE Reviews in*, 4:73–88, 2011.

[4] B. Münzer, K. Schoeffmann, and L. Böszörmenyi. Detection of circular content area in endoscopic videos for efficient encoding and improved content analysis. Technical Report TR/ITEC/12/2.03, Institute of Information Technology (ITEC), Klagenfurt University, Austria, 2012.

[5] A. F. Smeaton, P. Over, and W. Kraaij. Evaluation campaigns and TRECVid. In *MIR '06: Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval*, pages 321–330, New York, NY, USA, 2006. ACM Press.

[6] S. R. Stanek, W. Tavanapong, J. Wong, J. H. Oh, and P. C. De Groen. Automatic real-time detection of endoscopic procedures using temporal features. *Computer Methods and Programs in Biomedicine*, 108(2):524–535, Nov. 2012.

[7] S. Voros, J.-A. Long, and P. Cinquin. Automatic detection of instruments in laparoscopic images: A first step towards high-level command of robotic endoscopic holders. *The International Journal of Robotics Research*, 26(11-12):1173–1190, Nov. 2007.

[8] R. Yao, Y. Wu, W. Yang, X. Lin, S. Chen, and S. Zhang. Specular reflection detection on gastroscopic images. In *2010 4th International Conference on Bioinformatics and Biomedical Engineering (iCBBE)*, pages 1–4, June 2010.