



Motion artifact detection in colonoscopy images

Rukiye Nur Kaçmaz^{1,4*}, Bülent Yılmaz^{1,2,3,4}, Mehmet Sait Dündar^{1,4} and Serkan Doğan⁵

Abstract

Computer-aided detection is an integral part of medical image evaluation process because examination of each image takes a long time and generally experts' do not have enough time for the elimination of images with motion artifact (blurred images). Computer-aided detection is required for both increasing accuracy rate and saving experts' time. Large intestine does not have straight structure thus camera of the colonoscopy should be moved continuously to examine inside of the large intestine and this movement causes motion artifact on colonoscopy images. In this study, images were selected from open-source colonoscopy videos and obtained at Kayseri Training and Research Hospital. Totally 100 images were analyzed half of which were clear. Firstly, a modified version of histogram equalization was applied in the pre-processing step to all images in our dataset, and then, used Laplacian, wavelet transform (WT), and discrete cosine transform-based (DCT) approaches to extract features for the discrimination of images with no artifact (clear) and images with motion artifact. The Laplacian-based feature extraction method was used for the first time in the literature on colonoscopy images. The comparison between Laplacian-based features and previously used methods such as WT and DCT has been performed. In the classification phase of our study, support vector machines (SVM), linear discriminant analysis (LDA), and k nearest neighbors (k-NN) were used as the classifiers. The results showed that Laplacian-based features were more successful in the detection of images with motion artifact when compared to popular methods used in the literature. As a result, a combination of features extracted using already existing approaches (WT and DCT) and the Laplacian-based methods reached 85% accuracy levels with SVM classification approach.

Keywords: Image processing, motion-artifact, colonoscopy

Introduction

Large intestine (colon) is the last part of our digestive system. In order to detect any kind of disease in the large intestine, three different methods are employed which are virtual colonoscopy (CT), optical colonoscopy (conventional) and wireless capsule endoscopy. Although each method have advantages and disadvantages, used images in this study were obtained from optical colonoscopy, which is more invasive method when compared to virtual colonoscopy. Optical colonoscopy is the most preferred technique because it has a specific component to remove abnormal tissue during the operation whereas virtual colonoscopy is used only for imaging (1). In addition to these techniques, wireless capsule endoscopy has been launched approximately 20 years ago but this technique produces 60,000 images for each patient, thus it is not an easy task to detect any disease among all the images for an expert. *According to the American Cancer Society*, colorectal cancer prevalence ranks third among all cancer types. New studies show that 97.220 people will suffer from colon cancer in 2018 (1). During the operation, experts either record video or take an image, in order not to miss any abnormalities. Due to the folded structure of colon, disease detection process may not be completed during the operation. Experts may need to make interpretations from videos or images that are obtained during the operation. However, colonoscopy images are exposed to artifacts due to the camera movement, which is called the motion artifact. If an image includes motion artifact, interpretation of

¹Department of Electrical and Computer Engineering, Graduate School of Engineering and Natural Sciences, Abdullah Gül University, Kayseri, Turkey

²Department of Electrical-Electronics Engineering, School of Engineering, Abdullah Gül University, Kayseri, Turkey

³Department of Bioengineering, Graduate School of Engineering and Natural Sciences, Abdullah Gül University, Kayseri, Turkey

⁴Biomedical Instrumentation and Signal Analysis Laboratory (BISA), School of Engineering, Abdullah Gül University, Kayseri, Turkey

⁵Training and Research Hospital, Gastroenterology Clinic, Kayseri, Turkey

*Corresponding author: R. N. Kaçmaz
E-mail: rukiyenurkacmaz@gmail.com

DOI: 10.2478/eibtj-2018-0022

© 2018 Authors. This work was licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 License.

the image and disease detection is not an easy task for experts, thus; separation images with motion artifact and images with no artifact (clear) was tried automatically. Experts have to select among all the videos and images that are clear and after this process is completed they have to decide on the disease type. Deciding clear images visually has two main drawbacks: Firstly, this process is time-consuming for the experts in a highly busy daily schedule, and secondly, since the decisions are subjective and dependent on the experts' performance, human factors such as fatigue, sleepiness or lack of experience may negatively affect the accuracy of clear images and disease detection. Although the conventional colonoscopy images/videos, wireless capsule endoscopy have been discussed, wireless capsule endoscopy (WCE) has a similar problem. In addition to the colonoscopy systems, in recent years, WCE has been developed for human-independent computer-aided screening (1). In this method, the camera is placed in a vitamin-sized capsule. As this capsule travels along the digestive tract, the recorder connected to the patient's waist captures thousands of images. Although this procedure seems more advantageous for the patient, it is not very efficient in terms of examining about 60,000 images. The examination and processing of these images are highly time-consuming, it has not yet become a preferred method (2). Automatic clear image detection is required for both colonoscopy and WCE.

The fast movement of the camera affects many different image types, not only medical images. The detection and elimination of images with motion artifact have been a topic of research in digital images. For example, Seyed et al. conducted a study to eliminate blurred images using de-convolution conjugate gradient method for digital images (3). Iman et al. proposed the Wiener filtering method to make images deblur and eliminate noise on medical images such as colon, brain, and lung (4). Norman et al. proposed another quality measurement study but they used digital image power spectrum to measure the quality of an image (5). Hanghang et al. supposed that Harr wavelet transform was an efficient method to detect a blurred image. That study also gave information about the blurredness level of the image (6). Arnold et al. conducted a similar study by using discrete wavelet transform (DWT) coefficients as features obtained from colonoscopy videos. They considered only the luminance channel of an image (7). Another wavelet-based study was conducted by Atika et al. to restore digital blurry images (8). Dongen et al. aimed at automatic detection of informative frames for early detection of oesophageal cancer. They used the color histogram and the discrete cosine transform (DCT) coefficients as the features to be employed in the classification (9). Another similar study was performed by Xavier et al. who studied the detection of images with motion artifact using DCT features (10).

The aim of our study was to propose Laplacian-based methods for the detection of frames with motion artifact in the colonoscopy videos. The comparison of Laplacian-based methods with the previous approaches such as the wavelet transform (WT) and DCT was also performed. In this study, these tech-

niques were used to extract features from the images and aimed at automatically classifying images as blurry and clear.

Materials and Method

Colonoscopy Images

In this study, 100 colonoscopy images were extracted from the videos obtained from <https://www.gastrointestinalatlas.com/index.html>, (open-source) and Kayseri Training and Research Hospital (KTRH). The images of this database were acquired using white light conventional colonoscopy system. The open-source images were extracted from different videos. In order to improve the diversity in the image set, images were selected from different patients acquired from different angles. Image set included 17 images that were acquired from KTRH. While 10 images were labelled as clear, 7 images were labelled as motion-artifact. Each group contained 50 images. Six sample images from the image database were shown (Fig. 1a - 1f).

Pre-processing

The size of the images coming from different sources was different. While the image size of the open source database was 240x352 pixels, Kayseri Training and Research Hospital image size was 576x720 pixels. In order to perform a reliable comparison, all images (both open source and KTRH) were made the same size, thus all images were cropped and new size of all images became 176x156. The cropping process also eliminated the black border that included the date and patient name on the colonoscopic frames. These images were labelled manually based on visual experience by two researchers and a gastroenterologist. After cropping the images, an adaptive histogram equalization (AdaptHistEq)-(pre-processing step) were used to approach to improve the contrast of the images. This approach is different from a regular histogram equalization because AdaptHistEq is applied on tiles, not on whole images. This function is known as contrast-limited-adaptive histogram equalization. Contrast enhancement was performed on each tile (8x8 pixels). Histogram equalization was applied onto each tile, and then, based on desired contrast width-limit, cropping boundary value was obtained in order to crop the histogram. Then, each histogram was redistributed so that it did not exceed this cropping limit value. Finally, the contrast-limited histograms obtained for gray-level mapping were determined. Thus, limited contrast enhancement was performed and noise amplification was prevented (11). It is easy to observe that adapthisteq function revealed the details of the images. The pre-processing step makes the details more explicit and the selection easier. After pre-processing was completed, features were extracted from images using Laplacian operators (12), wavelet transform (WT) (12), and discrete cosine transform (DCT) (12).

Feature Extraction

Image texture is one of the most important characteristics used to describe a region of interest in an image. Texture features are the measures of intensity variations of a surface that determine properties such as smoothness, roughness, and regularity. After

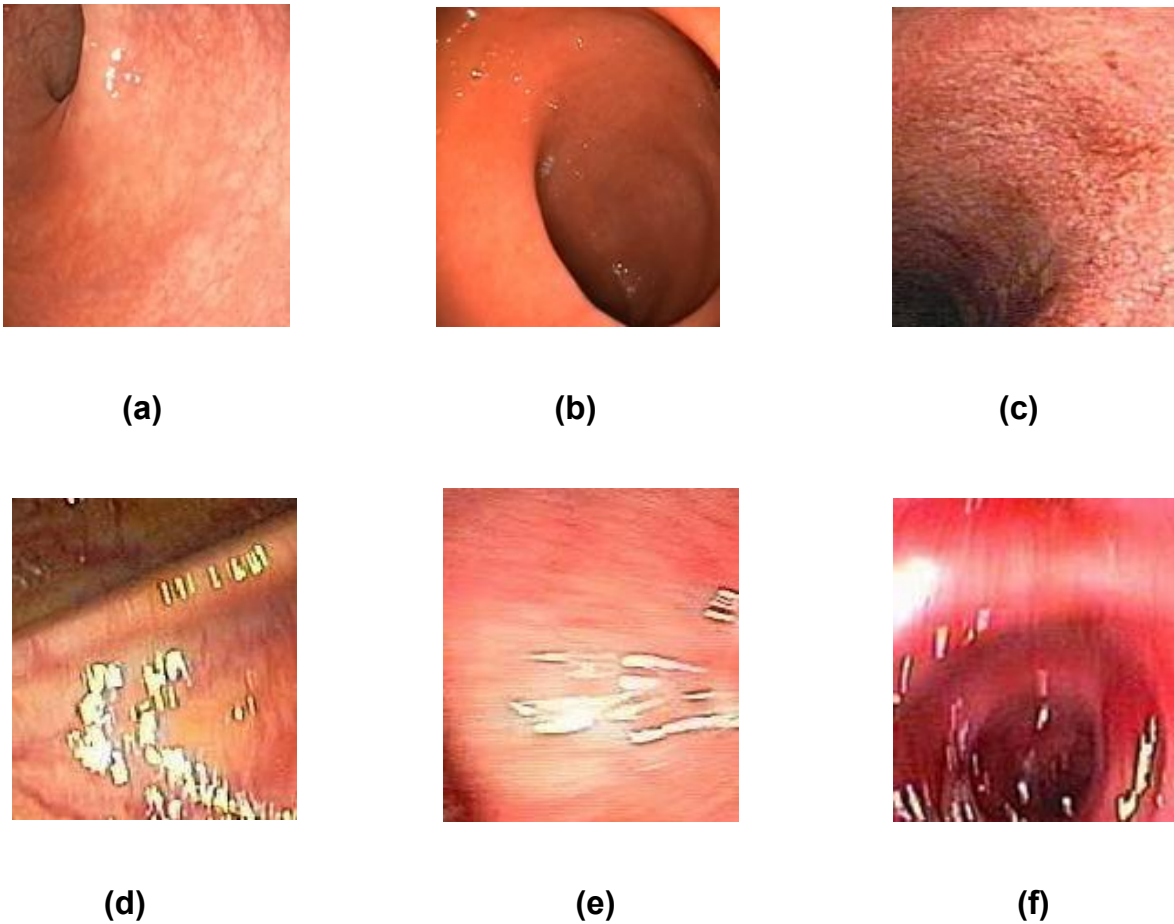


Figure 1. Sample colonoscopy images from our database are shown as clear (a, b, c) and with motion artifact (d, e, f).

pre-processing the next step was to extract image texture features. The features used in this study can be grouped into three families. The first group included DCT-based methods that are energy ratio and reduced energy ratio. The second group included wavelet transform-based methods that were the sum of wavelet coefficients, the variance of wavelet coefficients, and the ratio of wavelet coefficients. The third group consisted of

four Laplacian-based methods; modified Laplacian, the energy of Laplacian, the variance of Laplacian and the diagonal of Laplacian. Laplacian operators have not been used so far for the detection of images with motion artifact. Thus, the proposed method is using Laplacian operators to detect images with motion-artifact. The features were the energy and variance of the Laplacian of the images, the average of pixels obtained using

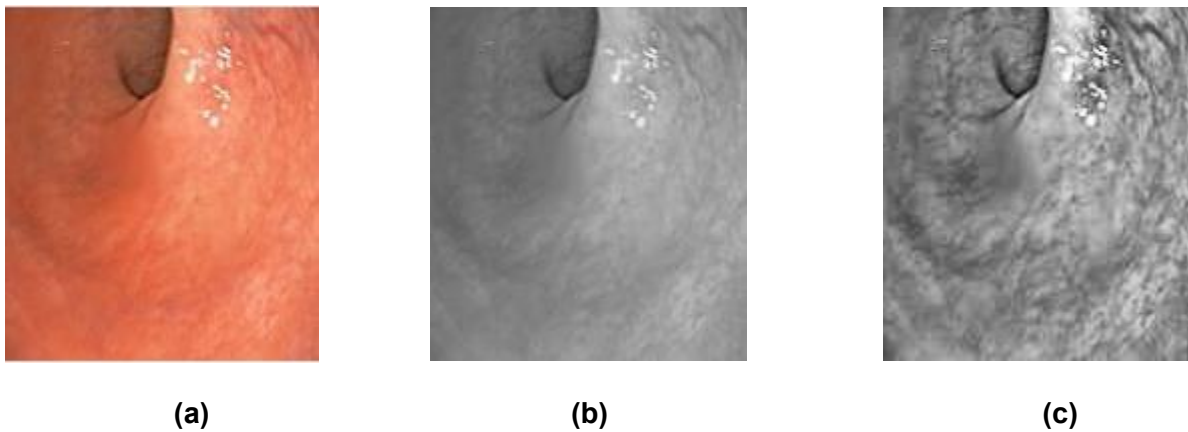


Figure 2. Original colonoscopy image (a), gray-scale representation of the original image (b), and the resultant image of adaptive histogram equalization applied on gray-scale image (c).

diagonal and modified Laplacian operator. Equation 1 and 2 show energy and variance of images respectively. The energy of Laplacian refers to the second derivative of the image (equation 1). The variance of Laplacian is given by equation 2. Modified Laplacian is obtained using equation 3 and is the convolution mask. Diagonal Laplacian is given by equation 4.

$$\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} \Delta I(i,j)^2 \quad [1]$$

$$\phi_{i,j} = \sum_{(i,j) \in \Omega(x,y)} (\Delta I(i,j) - \overline{\Delta I})^2 \quad [2]$$

$$\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} \Delta_m I(i,j) \quad [3]$$

$$\Delta_m I = |I * \mathcal{L}_x| + |I * \mathcal{L}_y| \quad \mathcal{L}_x = \begin{bmatrix} -1 & 2 & -1 \end{bmatrix} \quad \mathcal{L}_y = \mathcal{L}_x^T$$

$$\Delta_m I = |I * \mathcal{L}_x| + |I * \mathcal{L}_y| + |I * \mathcal{L}_{x1}| + |I * \mathcal{L}_{x2}| \quad [4]$$

$$\mathcal{L}_{x1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \mathcal{L}_{x2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Classification

In this study, firstly min-max normalization was applied on the feature set. Extracted features were used in the classification of clear vs. images with motion artifact. In the classification phase, three classification methods have been used such as the support vector machines (SVM), linear discriminant analysis (LDA) and k nearest neighbors (k-NN). SVM, LDA and k-NN were used in this study because they are the most basic and popular classifiers, and have shown to perform well in supervised classification problems (13). Usage of quadratic kernel gives the highest SVM results thus distance (d) was selected as 2. Also, the best accuracy level was obtained when k was selected 10 for the k-NN classifiers. In the classification phase, 10-fold cross-validation (CV) has been preferred. k-fold CV is

a technique to assess predictive models by dividing the original sample into a training set and a test set to evaluate it k times. Each feature family has been used one-by-one and altogether. The discrimination accuracy levels for different feature sets and classification methods have been presented. SVM, LDA and k-NN were used in this study because they are the most basic and popular classifiers, and have shown to perform well in supervised classification problems (13). The criteria used in the performance evaluations are given in equations [5]-[7].

$$Precision = \frac{TP}{TP+FP} \quad [5]$$

$$Recall = \frac{TP}{P} \quad [6]$$

$$f\text{-measure} = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} \quad [7]$$

Results

Obtained results showed that the use of Laplacian-based features in the discrimination of clear images versus images with motion artifact resulted in the best performances for all classifiers when compared to WT and DCT. The classification accuracies were higher than 70% for three classifiers. Laplacian-based features gave 76% accuracy using quadratic SVM, however, the same classifier gave 71% and 66% for WT-based and DCT-based features, respectively. SVM performed the best among all classifiers.

Moreover, in this study three different feature extraction methods were combined and used the same classifiers, SVM and LDA classifiers yielded better accuracy levels, higher than 80%. According to these results, successful blurred image detection is possible using the features obtained from Laplacian-based operators and also a combination of Laplacian, wavelet and discrete cosine transform-based operators. The accuracies, specificities, f-measure scales, sensitivities and area under curve (AUC) obtained for different feature extraction methods and classification approaches (Table 1).

Table 1. Automatic clear vs. blurred image discrimination accuracies, specificities, f-measure scales, sensitivities and AUC for different feature extraction methods and classification approaches.

Performance Measurement	Classification Methods											
	SVM				LDA				k-NN			
	LAP*	WT	DCT	All	LAP*	WAV	DCT	All	LAP*	WAV	DCT	All
Accuracy	%76	%71	%66	%85	%72	%67	%69	%85	%72	%64	%63	%70
Specificity	%72	%86	%62	%84	%76	%72	%62	%82	%82	%78	%70	%78
f-measure	0.76	0.61	0.65	0.85	0.72	0.66	0.68	0.85	0.70	0.61	0.52	0.69
Sensitivity	%80	%56	%70	%86	%68	%62	%76	%88	%62	%50	%56	%62
AUC	0.76	0.76	0.72	0.88	0.72	0.71	0.72	0.88	0.71	0.62	0.66	0.71

* Refers proposed method for discrimination images with motion-artifact from clear images.

Discussion

Both wireless capsule endoscopy and colonoscopy image analysis are still human-dependent systems. The studies reported in the literature demonstrate that many ambiguous images are obtained from colonoscopy and wireless capsule endoscopy. In order to detect any colonic disease accurately, images should be interpretable. Therefore automatic removal of the images with motion-artifact or uninterpretable images would save time for the experts. Some studies were conducted to detect this kind of problem, for example, Pina et al. studied on motion-artifact detection and used edge detection techniques for digital images (14). Levente et al. used blind deconvolution method to detect blurredness. This study focused on the detection of blurred region on digital images (15). In addition to blurred digital image detection, blurred medical image detection approaches were reported in the literature as mentioned in the Introduction section. In those studies, the researchers used DCT and WT for this problem (4)-(7). In this study, using Laplacian-based image features have been proposed for detection of blurry colon images from clear colon images. The results have shown that a combination of features extracted using already existing approaches (WT and DCT) and the Laplacian-based methods reached 85% accuracy levels with SVM classification approach.

In addition to motion artifact there are different noninformative image types like specular reflection and bad contrast in colonoscopy or WCE images. In an extensive future study, images with these artifacts will be included in these image set in order to perform automatic noninformative frame detection.

The number of images included in the dataset here was limited. In a future study, the size of the dataset will be expanded. Several other feature extraction methods will be investigated in the future.

Finally, the computational cost is critical for a real-time noninformative image detection approach. The processing time should be decreased by ten times. For this purpose, GPU-based approaches will be investigated.

Conclusion

In this paper, detection of images with motion-artifact was proposed using Laplacian operator based feature extraction. Apart from this, the features coming from Laplacian operators and previous implemented methods (DCT and WT) were compared and combined for the binary classification problem. The main contribution of this study is that the use of Laplacian-based operators to extract features from colonoscopy images to detect motion artifact is shown to be feasible. It should be noted that combination of Laplacian, DCT and WT based features performs better than individual feature families' accuracies reaching up to 85% by SVM and LDA classifiers.

Acknowledgement

The authors, RNK and MSD, are supported by the Turkish Higher Education Council's 100/2000 Program with a monthly stipend.

Conflict of interest statement

The authors declare that they have no conflict of interest.

References

1. <https://www.cancer.org/cancer/colon-rectal-cancer.html>, Colon Cancer, [online], June, 2018.
2. Tajbakhsh N, Gurudu SR, and Liang J. A classification-enhanced vote accumulation scheme for detecting colonic polyps, *Lect. Notes Comput. Sci.* 2013; 8198:53–62.
3. Mohammad S, Sanipour A, Akhlaghi IA. A new method for eliminating blur caused by the rotational motion of the images, *Peer J Pre Prints* 2015; 3:872-884.
4. Al-qinani IH. Deblurring image and removing noise from medical images for cancerous diseases using a Wiener filter, *Int Res J Eng. Technology* 2017; 4:2354–2365.
5. Nill N.B, Bouzas B. Objective image quality measure derived from digital image power spectra, *Opt Eng.* 1992; 31:813-825.
6. Tong H, Li M, Zhang H, Zhang C. Blur detection for digital images using wavelet transform, 2004 IEEE Int Conf Multimed Expo. 2004; 1:17–20.
7. Arnold M, Ghosh A, Lacey G, Patchett S, Mulcahy H. Indistinct frame detection in colonoscopy videos, *IMVIP 2009 - 2009 Int Mach Vis Image Process. Conf.* 2009:47-52.
8. Jain A, Kanjalkar PM, Kulkarni JV. Estimation of image focus measure and restoration by wavelet, *Proc. - 2011 4th Int Conf Intell Networks Syst. ICINIS 2011*; 73–76.
9. Van Dongen NC, Sommen F, Zinger S, Schoon EJ, De With P. Automatic assessment of informative frames in endoscopic video, *13th International Symposium on Biomedical Imaging 2016*; 1:119–122.
10. Marichal X, Ma WY, Zhang H. Blur determination in the compressed domain using DCT information, *Proc 1999 Int Conf Image Processing 1999*; 2:386–390.
11. Kurt B, Nabiye V. Dijital mamografi görüntülerinin kontrast sınırlı adaptif histogram eşitleme ile iyileştirilmesi, VII. Ulusal Tıp Bilişimi Kongresi 2010; 67–78.
12. Pertuz S, Puig D, Garcia MA. Analysis of focus measure operators for shape- from-focus, *Pattern Recognition* 2013; 46:1415 –1432.
13. Maleki R, Rezaei A, Minaei B. Comparison of classification methods based on the type of attributes and sample size, *J Converging Inf Technology* 2009; 4:94 –102.
14. Marziliano P, Dufaux F, Winkler S, Ebrahimi T. A no-reference perceptual blur metric, *Int Conf Image Process* 2002; 1:57-60.
15. Kovács L, Szirányi T. Focus area extraction by blind deconvolution for defining regions of interest, *IEEE Trans Pattern Anal Machine Intelligence* 2007; 29:1080–1085.