Extracting Remote Photoplethysmogram Signal From Endoscopy Videos For Vessel and Capillary Density Recognition

Seyyed Hamed Fouladi1, Ilangko Balasingham1,2, Kimmo Kansanen1 and Tor Audun Ramstad1

Abstract—In this paper, we propose a new feature for finding lesions in gastrointestinal tissues. Polyps or cancerous parts have different capillary pattern compared with normal parts. There are polyps which have higher density of vessel or capillary pattern. This feature leads us to extract remote photoplethysmogram signal from different parts of videos from gastrointestinal tissue. Due to the fact that hemoglobin absorbs more light than surrounding tissues, more changes are expected to be observed in the parts with higher density of vessels and capillaries. In the experimental results, rPPG signals is extracted from colonoscopy and endoscopy videos. This feature is used to distinguish between normal and abnormal tissues. It is shown that power of rPPG signal can be used to find lesion areas.

I. INTRODUCTION

Colorectal cancer is one of the cancers common for both male and female that is one of the leading causes of cancer death in the world [1], [2]. The early detection of polyps in intestinal tracts is essential to avoid polyps developing to dangerous cancerous stages [3]. Recent clinical studies have shown that inadequate quality of colonoscopy can lead to a significant polyp miss rate. Patients with missed detection of polyps may later be diagnosed with developed cancer with a survival rate of less than 10% [4].

There are various algorithms trying to diagnose abnormalities from intestinal tissue images and videos. The general idea in developing automatic detection system of colorectal polyps and cancers is to have videos or images which contain normal and abnormal tissues as ground truth. Desired features should be extracted for training a classifier which will be able to distinguish between normal and abnormal classes. The performances of the classifiers are dependent on the selection of appropriate features which characterize healthy and unhealthy tracts. Features which are used in the computer-aided detection approaches include shape, texture and color of the tissues. In [5], [6], shape and curvature analysis is considered to find abnormalities. In [7], the spatial structure of local texture is described the Local Binary Pattern. Support Vector Machine (SVM), Neural Network (NN) are the common classifiers exploited for classifying the images.

Most previous works used to evaluate frame or images from video to find the abnormality. In this paper, we present a feature that helps us to recognize the unhealthy tissues. The aim is to compare characteristic of abnormal tract with normal one in time domain. The idea is that the vessel and capillary pattern for polyps and cancerous tracts is different from healthy tissues [8]–[11]. For instance, hyper plastic polyps (HP) and gastric neoplasia (GN) have dense vessel pattern [10]. We try to capture the density of vessels in the video by extracting remote photoplethysmogram (rPPG) signal. The denser pattern of vessels and capillaries leads us to see more color variations over time. These color changes are not observable to eyes. There are many attempts to capture the changes and extract rPPG from camera recordings like [12], [13].

In this paper, we extract rPPG signal using the fact that blood absorbs more light than surrounding tissues. The video is divided into some Region of Interests (ROI). ROIs are tracked over time and rPPG signal is extracted. The powers of extracted signals show difference between abnormal and normal regions. In the experimental results, we use colonoscopy videos to show the differences in different parts of the videos.

II. VESSEL AND CAPILLARY DENSITY OF POLYPS AND CANCEROUS TISSUES

It has been observed that vessel and capillary patterns of polyps or mucosal lesions are different from healthy and normal parts [8]–[11]. Fig. 1 shows a polyp with dense vascular pattern intensity which is considered neoplastic. There are also polyps with normal or weak vessel and capillary pattern. In this paper, the target is to find the polyps with abnormal pattern. The dense capillary pattern in lesion causes stronger rPPG signal due to hemoglobin light absorption. Therefore larger changes are resulted.

III. EXTRACTING RPPG SIGNAL FROM GASTROINTESTINAL VIDEOS

In this section, we aim to extract rPPG signal from videos. Generally, T seconds of a video is selected. The first frame is selected into several part with or without overlap. In Fig. 2, the video is divided into Regions of Interest (ROI) without overlap. For instance, the frame is divided 18 columns and 13 rows which results 234 ROIs. All ROIs are tracked over time. In this paper, we use exhaustive search in 10 pixels in the neighborhood of the last location of the ROI. After ROI tracking, the video is preprocessed before noise and motion artifacts removal. In the preprocessing part, the mean of
ROIs’ pixels are calculated for red, green and blue channels. Therefore, for each ROI, three time-series are available. The signal mean which is related to the color and texture is removed.

A. Noise Reduction

It is essential to remove noise and unwanted signals after tracking ROIs. Generally, a digital camera sensor has red, green, and blue (RGB) color channels. Designing a bandpass filter with passband 0.5 Hz to 4 Hz can remove noises and artifacts. Furthermore, the color difference method in the color channels is exploited to estimate the rPPG signal through a linear mixture of the Red, Green, Blue channels in [13]–[15]. It is supposed that the sources include noise, motion artifacts and desired signal with non-Gaussian distribution. Since a linear mixture of the sources is available at each channel, the probability density function of observation goes to Gaussian distribution due to central limit theorem. Therefore, the mixing matrix and sources can be extracted by maximizing non-Gaussianity of the output components. When the mixing matrix is achieved heart rate frequency and corresponding power are obtained.

B. Finding Heart Rate Frequency and Its Power

The most important part is to obtain the heart rate frequency and its power. Therefore, the power spectrum of the signals extracted form each ROI needs to be estimated. Depending on the assumption on the signal, various spectrum algorithm can be exploited. In [13], [14], fast Fourier transformation (FFT) is used to estimate the signal spectrum. In these kind of algorithms, the number of samples should be selected large enough to get acceptable results. Also, subspace algorithm can be used to resolve low samples problem and gain a reasonable result. In [12], MUltiple SIgnal Classification (MUSIC) algorithm is exploited to overcome low number of samples difficulty in extracting heart rate frequency from human face. In this algorithm it is assumed that the desired signal contains \( p \) number of sinusoid signals which means sparse spectrum for the desired signal. However, since there are many noise and motion artifacts in the video, FFT based spectrum estimation algorithm is used in this paper. We take 10 seconds of a video for heart rate extraction.

After noise and motion artifacts removal, there might be some peaks along with the heart rate component in the desired signal spectrum. The heart rate frequency normally are expected to be observed from 0.7 Hz to 2 Hz. However, patients are more anxious and nervous during endoscopy experiments. Therefore, higher heart rate frequency are expected to be seen. In [16], R-R interval variability was analyzed from electrocardiograms obtained during endoscopy. Also, time and frequency domain analyses of heart rate variability were performed in their work. It has been seen that the heart rates were approximately from 65 bpm to 95 bpm.

Moreover, the availability of heart rate frequency in different part of video should be considered. The heart rate should be observable in all part while a motion artifact or noise might happen just in a ROI.

C. Analyzing Videos

For each ROI, there are three time series related to RGB channels. \( s_{ij}(n) \) is a vector for the ROI of \( i \)th row and \( j \)th column at \( n \)th frame including red channel stream \( s^R_{ij}(n) \), green channel stream \( s^G_{ij}(n) \) and blue channel stream \( s^B_{ij}(n) \) as

\[
s_{ij}(n) = [s^R_{ij}(n) \ s^G_{ij}(n) \ s^B_{ij}(n)]^T
\]

(1)

There is peak in heart rate frequency in the spectrum of the time series from RGB channels \( S^R(f) \), \( S^G(f) \) and \( S^B(f) \). We form the matrix \( X \) with the elements as

\[
X_{ij} = \frac{\int_{-\infty}^{\infty} S^R_{ij}(f)df}{\int_{-\infty}^{\infty} s^R_{ij}(f)df} + \frac{\int_{-\infty}^{\infty} S^G_{ij}(f)df}{\int_{-\infty}^{\infty} s^G_{ij}(f)df} + \frac{\int_{-\infty}^{\infty} S^B_{ij}(f)df}{\int_{-\infty}^{\infty} s^B_{ij}(f)df}
\]

(2)

\[
\frac{\int_{-\infty}^{\infty} S^R_{ij}(f)df}{\int_{-\infty}^{\infty} s^R_{ij}(f)df} + \frac{\int_{-\infty}^{\infty} S^G_{ij}(f)df}{\int_{-\infty}^{\infty} s^G_{ij}(f)df} + \frac{\int_{-\infty}^{\infty} S^B_{ij}(f)df}{\int_{-\infty}^{\infty} s^B_{ij}(f)df}
\]

228
where $X_{ij}$ is an element of $X$ and $\Omega$ is the area that energy of heart rate frequency is distributed which depends on windowing. For rectangular window, we consider it as $\Omega = \{ f | f_{HR} - \frac{T}{2} \leq f \leq f_{HR} + \frac{T}{2} \}$. Conceptually, $X_{ij}$ is the summation of normalized power of heart rate in RGB channels. The powers need to be normalized due to heterogeneous lighting condition.

### D. The Difficulties for Extracting The Feature

The first difficulty is that the illumination is not uniform in the frames that we tried to decrease the impact by defining a criterion in (2). Second, choosing ROI size is a challenge that affects the performance of analysis. We select ROI sizes based on the number of pixels in the height and the width of video. Also, it is needed to choose the ROIs based on the texture and color to avoid tracking dark areas and signal extraction from them. Third, the recording containing the lesion should be long enough in order to have suitable number of samples for extracting rPPG signal. Low number of samples can cause low resolution spectrum.

### IV. Experimental Results

In the experimental results, three colonoscopy videos are analyzed which have different resolutions and frame rates. Two of them contain polyp and one of them is tagged as healthy. The first frame of each video is divided into several ROIs. All ROIs are tracked over time. Neighbor ROIs have 90% overlap in order to see the trend of decrease or increase of power. Then, rPPG signal is extracted for each ROI. Matrix $X$ is obtained which shows heart frequency power concentration in the video.

In the first experiment, a video containing a polyp is analyzed with frame rate 30 fps and its height and width are 576 and 768 respectively. The first frame of the video is represented in Fig. 3. The ROIs sizes selected for this experiment is 100 pixels in width and 100 pixels in height. The distance between the closest corner of neighbor ROI is 10 pixels in width and height which results 90% overlap in both axis. As it can be perceived from Fig. 4, the area around the polyp has higher value.

In the second experiment, a video containing a polyp is analyzed as shown in Fig. 2 with frame rate 30 fps and its height an width are 1080 and 1920 respectively. The ROIs sizes selected for this experiment is 300 pixels in width and 300 pixels in height. Fig. 5 also shows higher values around the lesion.

In the last experiment, a video showing a healthy tissue is analyzed as shown in Fig. 6 with frame rate 25 fps and its height an width are 576 and 768 respectively. The ROIs sizes selected for this experiment is 100 pixels in width and 100 pixels in height. Fig. 7 represents more distributed power rather than videos containing polyp. However, the color and the texture of ROI must be considered in abnormality detection. It means that the proposed feature can be used along with other feature like color and texture to improve
classification performance.

V. CONCLUSION

We proposed a new feature in order to distinguish between normal and abnormal tissues. The feature is the power of rPPG signal in heart rate frequency extracted from gastrointestinal video. A criterion was defined to find denser vessel and capillary pattern in the area or ROI. Since cancerous and polyps have different patterns, this factor is used to discriminate lesions. This feature can be used along with other features such as color and texture of a tissue for classification. However, the performance is affected by non-uniform illumination, noise and motion artifacts. Also, low number of samples decreases heart rate estimation accuracy and the performance.

VI. ACKNOWLEDGMENT

The research work was performed, in part, of the research project Medical sensing, localization and communications using ultra wideband technology (MELODY) project no. 25122502.

REFERENCES