Project Report

Hitesh Kandala

May 6, 2020

Contents

1 Overview

This paper presents a new method of formation of photoplethysmographic (PPG) images with high spatial resolution from video recordings of a living body and is based on the lock-in amplification of every pixel of the recorded video frames. The paper shows that the system is capable to detect the minimal irritations such as scratching of the palm by finger by visualizing dynamic changes in cardiovascular pulse wave.

2 Introduction

Photoplethysmography (PPG) is an optical method to detect a cardiovascular pulse wave travelling through the body. The basic form of PPG technology requires only two optoelectronic components: a light source to illuminate a part of the body and a photodetector to measure small variations in light intensity after light interaction with the illuminated part. The key factors that can affect time-varying component of light intensity are the blood volume, blood-vessel-wall movement, and the orientation of the red blood cells. Generally, this time-varying component provides a signal proportional to changes in skin blood volume but does not produce a quantitative measure.

In this paper, a reference function, required for synchronous detection of cardiovascular pulse waves in PPG images, was formed from a large area of the same images to improve SNR by averaging data from large number of pixels. This technique allows us to obtain PPG images with increased SNR and high spatial resolution which is limited only by the used photosensitive matrix. The main purpose of this paper is to demonstrate visualization of dynamic changes in cardiovascular pulse wave during the cardiac cycle.

3 Methods

3.1 Synchronous Detection

In previous works related to plethysmographic imaging, information about blood pulsations at subject's heart-beat frequency was retrieved after either spectral analysis or narrow band-pass filtering of temporally varying pixels values. After this initial processing only the amplitude data was mapped on the skin image. Amplitude calculation is similar to the envelope demodulation. Synchronous detection, also known as coherent demodulation, while being more complex has several advantages over envelope demodulation. For example, for amplitude modulated signals it provides higher SNR than envelope detection, and, therefore, the amplitude estimate of the oscillations in video image pixels is less susceptible to noise. Synchronous detection can also retrieve the relative phase of blood pulsations, which opens up new possibilities in photoplethysmographic imaging because it would allow visualization of the dynamics of blood pulsations under the skin.

3.2 Experimental set-up

Camera with a global shutter and a detector resolution of 1024 1280 pixels was focused on the skin of the subject palm using a complementary Canon TV-zoom (18-108) lens. The distance

between the camera lens front and the palm was approximately 100 cm. The camera was fixed at a table and focused on the subject's palm refrained from movements during video recording. Duration of recording varied from 30 s to several minutes.

The light source comprising only 2 LEDs with the central wavelength of 530 nm (bandwidth 20 nm) and output power of 30 mW is situated at the distance of 100 cm from the palm. In the experiments presented in this paper, videos were recorded at the frame rate of 20 frames per second with the exposure time for each frame of 8 ms.

3.3 Reference function formation

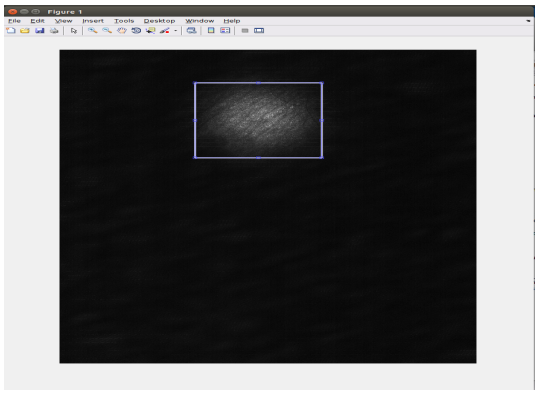


Figure 1: Selected ROI for processing of recorded frames

The first step includes processing of selected ROI of recorded frames to generate a reference function $R_c(\mathrm{t})$, which covers most of the palm image as shown in below figure.

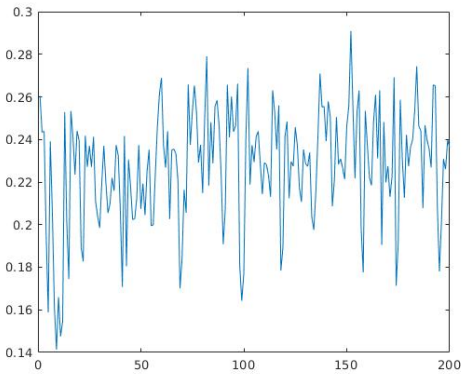
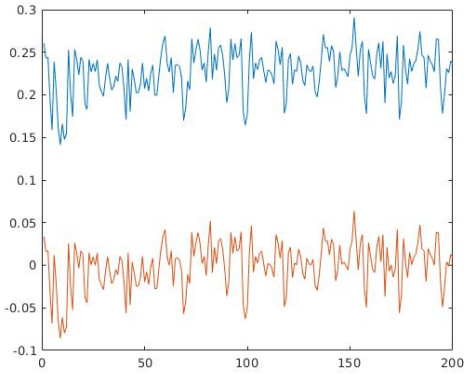
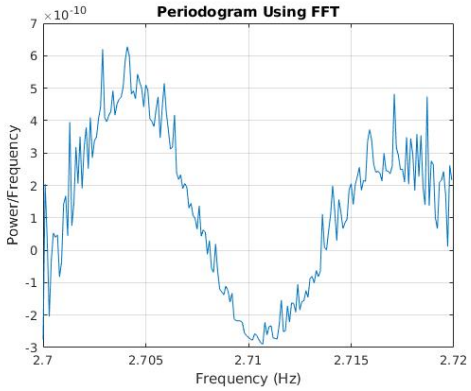


Figure 2: Spatially-averaged pixels value per recorded frame

Then all pixel values within this ROI were spatially averaged resulting in single mean value per each recorded frame. The ratio of mean-pixel-value modulation to the DC level is about of 0.002 which is enough for recognition of cardiac and respiration physiological processes after Fourier analysis.



(a) detrended and original mean-value signal



(b) Power spectrum of the mean-pixel-value time-trace.

Power spectrum of the mean-pixel value of the selected ROI is obtained by applying FFT to the signal after detrending its DC level.

Instabilities of the rates of physiological process during the time of the reflected-images recording result in broadening of the frequencies representative for the cardiac pulsation and breathing. These instabilities originate from both natural variations of the rates and motion artifacts.

After the frequency band corresponded to the heart beats is selected, all other frequencies are truncated, and the inverse Fourier transform is applied exclusively to the truncated spectrum (for frequencies f in the range of C1 $\geq$ f $\geq$ C2). This reconstructs the reference function R_C(t) which represents the heart pulsations. Since the reference function calculated from the inverse Fourier transform has an integer number of periods, its real and imaginary parts are mutually orthogonal. After calculation of inverse Fourier transform it is normalized in such a way that

$$\sum_t Re[R_C(t)]R_C(t) = 1 \tag{1}$$

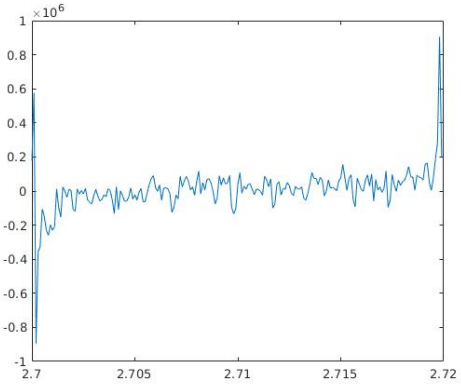


Figure 4: Normalised reference function $R_C(t)$

This reference function is further used for lock-in amplification of the recorded series of images. The normalization allows further evaluation of the mean complex amplitude of the signal synchronized with the heart beats within the chosen frequency range at every pixel of the frame.

The adaptive algorithm also uses the raw time-trace of the spatially averaged pixels value for calculation of the reference function. In contrast with previous algorithm we use only 3-5 periods of the cardiac pulsations for approximate estimation of the heart-beat rate. This estimation is done by applying fast Fourier transform to the initial part of the raw time-trace and selecting the frequency at which the power spectra reaches the maximum within pre-defined frequency band.

3.4 Correlation matrix visualization of blood-volume dynamics

In the second step, the reference function is multiplied with the corresponding recorded frame in such a way that each pixel value of the first frame is multiplied by the same coefficient. Therefore after multiplication, we obtain the series of frames in which pixels have the complex values. Thereafter, a correlation matrix $S_C(x, y)$ is calculated by summarizing complex elements having the same coordinates (x, y) over all the frames in accordance with

$$S_C(x, y) = \sum_t I(x, y, t) R_C(t) \tag{2}$$

Here $I(x, y, t)$ is value of the pixel with coordinates (x, y) of the image frame captured at the moment of t. The correlation matrix $S_C(x, y)$ contains the same number of pixels as any of the initial frames $I(x, y, t)$.

The matrix $S_C(x, y)$ is approximately equal to the cross-correlation function of the reference $R_C(t)$ and the time-varying image of the subject's palm. Since $R_C(t)$ represents the cardiac pulsations, the matrix $S_C(x, y)$ is a lock-in amplification of the pixel values varying in time synchronously with the heart beats. Therefore, the modulus of each pixel value of the correlation matrix $S_C(x, y)$ is proportional to the modulation amplitude of light intensity reflected from the respective point of the subject's skin. Since light-intensity modulation is caused by the blood pulsations, the matrix $S_C(x, y)$ describes spatial distribution of ac-component of the blood-volume pulsations at the heart-beats frequency. In other words, it is PPG image.

Because to its complexity, the matrix $S_C(x, y)$ contains information also about relative phase of blood pulsations in different parts of the observable area as we cannot claim that blood pulsations occur with the same phase everywhere in the whole observable area of the subject's skin. Let us assume, that the reference function, normalized in an observation interval with N samples of mean-pixel-value time-trace is

$$R_C(t) = \frac{2}{N} exp(2\pi i f t), \tag{3}$$

and the pixels values are

$$I(x, y, t) = A(x, y) cos[2\pi i f t + \psi(x, y)] + B(x, y) \tag{4}$$

Here A(x,y) is the amplitude of the pixel value oscillations at the frequency of f, ψ(x,y) is the relative phase of these oscillations, and B(x,y) is the mean pixel value. Therefore, Eq. (2) yields

$$S_C(x,y) = A(x,y)exp[-i\psi(x,y)] \tag{5}$$

Consequently, the real part of $S_C(x,y)$ corresponds to the instant deviation of pixels values I(x,y,t) from their mean values B(x,y) at the moment when the phase of the reference function is equal to zero. The imaginary part of $S_C(x,y)$ in turn describes the instant deviation of pixels values from their mean values at the moment when the phase of $R_C(t)$ is equal to $\pi/2$. Therefore, since the deviation of the pixel value is proportional to the local change in blood volume, we can reconstruct dynamic changes of the blood volume pulsations during the cardiac cycle by calculating a new series of frames $H_C(x,y,t)$ as

$$H_C(x,y,t) = Re[S_C(x,y)]cos\phi(t) + Im[S_C(x,y)]sin\phi(t) \tag{6}$$

Here ϕ(t) = $2\pi f_C$t, where f_C is the mean rate of the heart beats.

Values assigned to the pixels of the frames $H_C(x,y,t)$ can be positive, negative, or zero. Zero level means either absence of the blood-volume pulsations at the heart beats or pulsations which are phase shifted by 90° in respect to the reference function $R_C(t)$. Positive values are pulsations of the blood volume in phase with $R_C(t)$, while negative values are in counter phase. The higher the amplitude of the pulsations, the larger value is assigned to the pixel. The animation shows dynamic changes of the amplitude of the fractional blood volume in the skin during the cardiac cycle and spatial distribution of the relative phase of these pulsations.