Volume 4, Issue 4, April 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Simpler Technique for Extraction of Health Information using Video Imaging

Xavier Gonsalves, M. M. Sardeshmukh
Department of Electronics and Telecommunication
Sinhgad Academy of Engineering, Pune University, India

Abstract—The flow of blood below the skin has certain parameters like blood volume which changes over time. These conversely change the parameters of light spectra reflected from the skin. This light spectra can reveal health parameters like heart rate (HR) and many others. A few new processing methods have been developed in recent years for using this imaging technique called photoplethysmographic imaging in extraction of HR. However, these techniques are complex and their implementation on conventional hardware in a real-time manner looks difficult. In this paper, we will discuss a simpler technique for the extraction of HR and will suggest improvements that can be applied to it.

Keywords— Health parameters, photoplethysmography (PPG), photoplethysmographic imaging, heart rate (HR), face detection, peak counting.

I. INTRODUCTION

Various parameters such as heart rate (HR), respiration rate (RR) and the saturation of oxyhemoglobin in blood are important for diagnosis of a person's health. The conventional methods for acquiring this information serve as reference standards for various purposes. For example, the cardiac electrocardiogram (ECG) using chest electrodes is a well-established technique for observing the electric activity of the heart. Pulse oximetry is also used to monitor oxyhemoglobin saturation and HR for clinical diagnosis and monitoring. Various techniques have been developed after painstaking research and development over many years, some of which are mentioned in [1].

Photoplethysmographic imaging is simple and can be implemented using LEDs or color cameras which record videos of a subject. Monitoring changes in the blood volume characteristics makes phoplethysmographic imaging fit for extraction of health parameters. The light information of the skin reflectance can help in extraction of health parameters. When the heart contracts, blood is pumped to various parts of the body through the arteries. It causes a wave of blood to flow through the various skin tissue. When the capillaries are filled with blood, they block the passage of light through them and thus more light is reflected from the skin. The change in the volume of blood in the directly affects the amount of reflected light. This light can be captured using a conventional camera and a time series signal can be generated from it. Pulse oximetry uses this method to determine the oxygen saturation in arteries which is based on photoplethysmography (PPG). The change in attenuation of light from the body tissue dictates the oxygen saturation in the blood. A detailed description of PPG can be found in the work by Sahindrakar et al. [2] and in [3].

Poh et al. [4] proposed an algorithm based on ICA analysis after their work at the MIT in 2010. A spatial average of all pixels of the region of interest (ROI) which is the face was calculated of each frame for all three channels giving the R, G, and B traces. Their technique employed a concept called independent component analysis (ICA) [5]. The normalized signals were decomposed into three independent components using either joint approximate diagonalization of eigenmatrices (JADE) algorithm developed by Cardoso [6] or FastICA which is based on Hyvärinen's work [7]. After this a Fourier transform was applied on one of the source signals to obtain a power spectrum. The pulse or heart rate (HR) is the highest power frequency in the range of 0.8 – 4 Hz. This work was carried forward by the same team [8]. They incorporated additional post processing steps on the segregated independent component. A new technique of in beat intervals (IBIs) was employed to calculate the HR. In the Lomb periodogram of the time series signal, a high frequency (HF) component related to breathing helps in estimating the respiratory rate (RR) of a subject [9]. In their study the RR was calculated using the center frequency of the HF peak f_{HF} in the PSD which actually equates to $60/f_{HF}$. In our paper however, the focus will be on the HR and not on the RR. A method proposed by Wei et al. [10] employs a similar methodology as that of the IBI series. The raw traces of RGB channels of a face ROI are first used to construct a Lalpacian eigenmap (LE) which reduces the three signals into one by means of dimensionality reduction i.e. it reduces three dimensional data to one dimension. After post processing steps similar to [8], the HR is found out from the IBI series. A comparison of how LE fares when compared with other techniques has been done in [10]. Also, Sahindrakar et al. [2] developed a technique which uses the face as the ROI and tracks it in each frame using the Pyramidal LK method [2], [11] and [12]. The methodology here is to first track and then combine the RGB traces in an additive manner and generate two component signals i.e. 'R – G', 'G – B' along with 'R + G – 2B' and '– 2R + G + B'. In the algorithm these vectors are used to extract the HR by means of data processing and Fourier analysis.

In this paper, we present a simpler processing technique which will help photoplethysmographic imaging in extracting the heart rate (HR). We applied this approach to face videos of many subjects to extract HR measurements and compared them with a blood volume pulse (BVP) oximeter which was used as a reference. Simultaneously, we will be proposing a few directions for future work in this area.

II. METHODOLOGY

There are many techniques which can be used for physiological parameter extraction from PPG signals [1]. But a few techniques developed in recent years give results that agree with those of conventional methods. In our previous work [13], we have investigated many algorithms and a few are mentioned in the previous section. We found that a mixture of simple steps of various techniques suits our need of implementation in C. In this section we have a look at our methodology for extraction of HR.

A. Experimental Setup

For the purpose of parameter extraction, the video imaging of a certain part of the skin is necessary. The region of interest (ROI) that was chosen for experimentation was the face. The face ROI is mostly used for reflection type PPG. The subject has to sit still in front of the web camera. The video of the desired ROI is recorded in color with a frame rate of 15 frames per second (fps) and is saved in AVI format on a computer. For our project, we have recorded our videos in AVI format using a Logitech C310 webcam with 640x480 resolution at 15 fps. The remaining processing part can be done by a program written in MATLAB (The Mathworks Inc.) or using the Open Computer Vision (OpenCV) library for an implementation in C language. But for the purpose of this study OpenCV was used. While capturing the videos the pulse of the subject was also captured using a pulse oximeter which was used as a reference. The pulse oximeter we used is a Contec CMS 50D pulse oximeter. A basic block diagram of the algorithmic (software) side of the system is shown in Figure 1 and a description of its various parts is given below.

B. Image Capture Construct

OpenCV library helps in image capture from a video or from a live webcam. OpenCV gives us a device capture construct. Changing the parameters of this construct allows us to read from a video or from a live webcam.

The videos for our project were read by the device at each run of the algorithm. Accessing the pixel data of the image can be done on a pixel by pixel scale. OpenCV allows us to use certain data types for this purpose [11].

C. Face Detection and Tracking

The face detector used in this project is a pre-trained cascade classifier with *Haar-like features* and is available in the OpenCV library in the form of a class. This classifier is based on the work by Viola and Jones [14] and Lienhart and Maydt [15]. OpenCV allows us to use various classifiers which are basically databases of features for matching with objects of various sizes in the image.

Fine tuning of the face detector parameters can lead to a faster detector. From our experiments, we have noticed that the *scaleFactor* when close to 1.1 gives almost accurate detection of faces. But this reduces the execution speed of the code. We have thus worked on this aspect and feel that a *scaleFactor* of 1.3 and *minNeighbors* being 3 gives quiet good results when considering close to real-time implementations. More on these parameters can be found in [11].

D. Pre-processing

Once the face was detected in an image, the R, G and B information was extracted from each frame. The average value of the R, G and B vector over the entire ROI was then calculated. Thirty seconds worth of image frames were used to create a data trace or window which is used for further processing. Each of the traces was 5-point symmetric moving

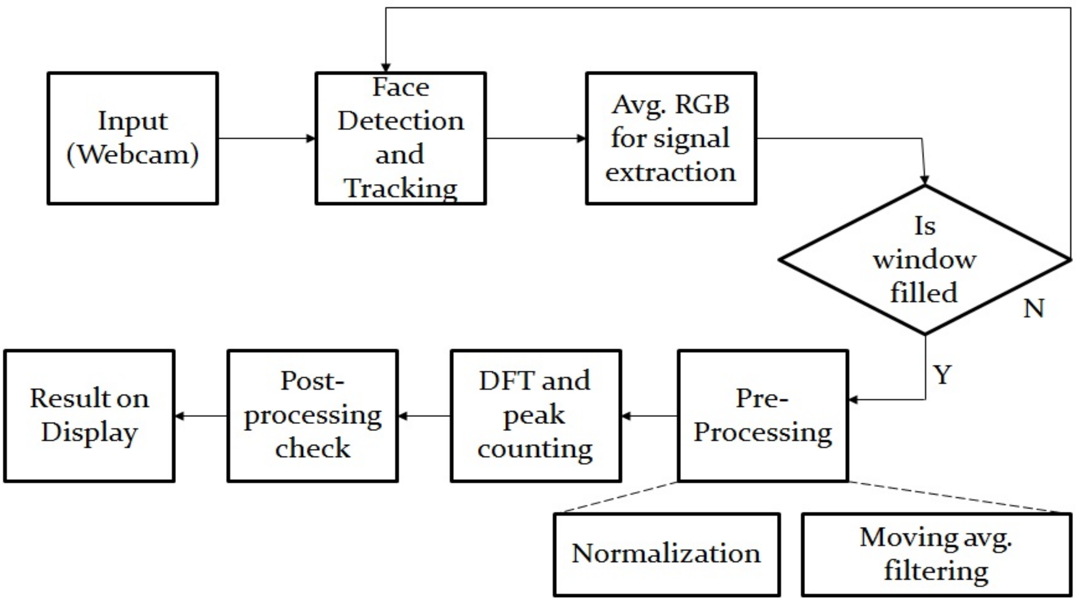


Fig. 1. Block diagram of the software side of the system

average filtered. These traces were then normalized. From these three time series signals a signal was generated which was the signal of 'R-G'. It was then raised to the fourth power. This signal is then normalized over its data samples.

E. Discrete Fourier Transform and further processing

The need of the discrete Fourier transform (DFT) of the time series signal was to find out the frequency with the highest power content in the 30 seconds window. The standard formula of Fourier transform was used for our calculations which is:

$$X_k = \sum_{n=0}^{N} x_n e^{-i2\pi kn/N} \quad (1)$$

where, $k = 0, \ldots., N-1$. This computation is difficult to compute easily in C which was our language used to write the algorithm so, we used Euler's form to simplify it as:

$$X_k = \sum_{n=0}^{N} x_n \cos(\frac{2\pi kn}{N}) - i \sum_{n=0}^{N} x_n \sin(\frac{2\pi kn}{N}) \quad (2)$$

This equation was used to separate the real and imaginary parts, putting them in separate arrays. We later used them to calculate the power spectrum. We scanned our band of interest i.e. 0.8 - 4 Hz to find the component with the highest power among all the other frequencies. This component gives us our heart beat in Hertz which we converted to beats per minute (bpm) which gave us our heart beat in the range 48 - 240 bpm.

Also alongside this we also calculated the number of peaks that were present in this 30 seconds window. We doubled this number which gave us a very close estimate of the HR in bpm. This HR extracted using the peak counting algorithm was used for post processing check which determined the final heart rate at that instant. In our algorithm, the window of 30 seconds incremented by one second on every new run of the algorithm. Thus what is received at the output is a stream of figures displaying the heart rate and this stream updates every one second. At each pass the HR cannot change by more than 12 bpm. If this is not true then the current peak in the DFT is cancelled out and the frequency with the next highest power component is taken as the HR frequency. Along with this there is an added restriction that the HR must be within 30 beats of the current peak counting heart rate.

III. RESULTS

TABLE 1 HEART RATE MEASUREMENTS

Participant No.	**Oximeter HR (bpm)**	**Calculated HR (bpm)**	**Error (bpm)**	**Percentage error (%)**
1	89	88	-1	1.12
2	93	92	-1	1.07
3	75	68	-7	9.33
4	82	76	-4	4.87
5	107	100	-7	6.54

Section II. A., elaborated the use of the face ROI for capture of the PPG signal. An example of the extraction of the photoplethysmographic signal is shown in Figure 2. In Fig. 2(a) there is a depiction of the face detector which is elaborated in section II. C. The area of this rectangle encapsulating the face is shrunk to an approximate value to cover only the skin pixels. Fig. 2(b) shows how the face pixels were split into their RGB components. Fig. 2(c) shows the trace of the first 10 seconds of the 30 seconds window of the green trace from which we extracted the 'R-G' signal. This plot shows the raw trace of the green signal which clearly has the photoplethysmographic information needed for processing.

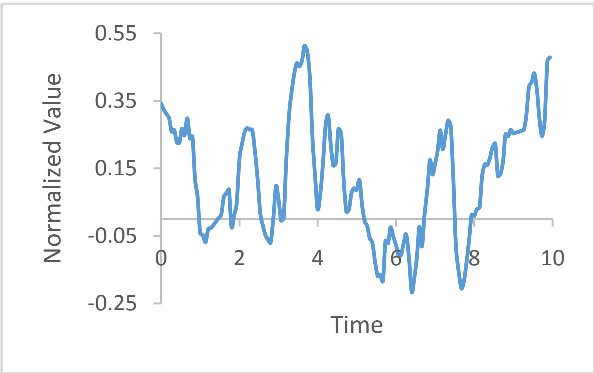
As explained in section II. E., we have used discrete Fourier transform for finding the maximum power component in the range of 0.8 - 4 Hz. A depiction of the Fourier transform of one of our windows is shown in Fig. 2(d). We programmatically suppressed the out of band frequencies. We noticed that the highest component is more or less in the range of the reference HR measured using a pulse oximeter. If we use the maximum power frequency (f_{max}), we can find the heart rate (HR) as:

$$HR_{(bpm)} = 60 * f_{max\,(Hz)} \quad (3)$$

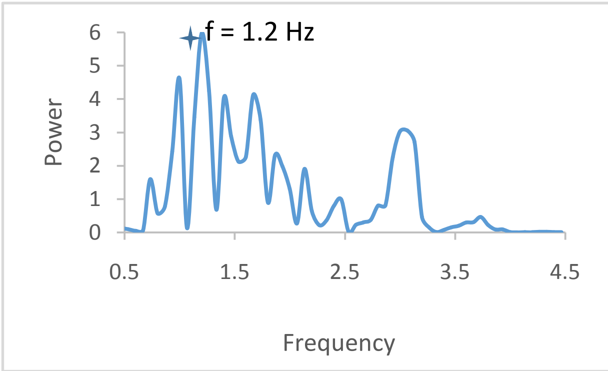
For our experiments we recorded videos of 20 participants (3 videos per participant) and each video was 2 minutes long. Of these 3 videos one video was recorded after the participant did some level of exercise whereas, the other two were recorded while at rest. Table 1 shows our findings on five of the videos of subjects we tested. For these readings we recorded each subject's video of 30 seconds and the system calculated the HR over this 30 seconds window. In this tabulation we have compared the calculated HR with that acquired from a pulse oximeter. We have also calculated the percentage error for each participant and have noticed that it is quiet low. We did so for each video of our database and found out that the absolute average difference was 5.225 bpm.

(a)

(b)



(c)



(d)

Fig. 2. System outputs (a) Face detector working on a video. (b) Face pixels split into R, G and B components. (c) Normalized raw green trace of 10 sec length. (d) Power spectrum with highest power at frequency close to that of the HR

IV. DISCUSSION

Studies have shown that light reflectance from the skin is influenced by various physiological differences at different depths of the tissue [16]. For ease of measurement and comfort of the participant three ROIs are generally used which include the earlobe [17], the face [2], [4], [8] or any one fingertip [18]. But in our study we kept our ROI as the face because it is comfortable to sit in front of the camera for a certain amount of time and also this technique can be used while the subject is at work on a computer.

The main point about our algorithm is that it is simple. This due to the fact that it relies heavily on the 'R-G' component of the extracted RGB traces. We found that just the G trace which is said to contain most of the PPG signal actually does not give good results. While experimenting with other vector pairs based on the study in [2] we found that in fact 'R-G' vector gives good results when compared with the other vectors of this study. When compared with the other algorithms like ICA and LE, our algorithm is simple and does not need any complex processing. The complex algorithms do take a lot of time per run and may not be optimal in a real-time sense. These algorithms can be sought after when a very accurate result of the HR is needed but when a more real-time implementation is required, our algorithm can be considered. This because our output occasionally wanders off but most of the time it remains within the range of the pulse oximeter reading. But, one thing should not be overseen which is that the output updates every one second and so some wander can occur. Also, our algorithm doesn't give good results for the videos which were taken after exercise. The reason for this is that our initial output reading is displayed after an analysis of 30 seconds of frame samples and in this time window the HR of the participant drops by a very large value which is outside our guard condition of 12 beats change and thus the wrong result.

The algorithm discussed above is implemented on video recordings of subjects and not in real-time. It is important that the method used for detecting the face and finding the HR fulfills certain time bounds for implementation in real-time. The face detection algorithm takes more time to implement when accuracy is high, which may cause skipping of frames [2]. Improvements in the classifier can help reduce the time of detection and can make the algorithm fit for real-time implementations.

The other two aspects that are still left are lighting and motion artifacts. For our experiments we have used lighting conditions that are similar to a room illuminated by natural sunlight. Our readings totally relied on this natural source. Yet, there is future scope for comparison with measurements done using a spot light which mimic the real-life situation of a subject working at a regular work desk. Also, our experiments were carried out on subjects who were sitting still in front of the webcam with little or no movement. A test of the motion-robustness similar to the work done by the teams in [4] and [2] can certainly be deemed as future scope for this work. This will ascertain whether or not our algorithm is fit for real-life scenarios.

V. Conclusion

In this paper, we discussed about photoplethysmographic imaging and stated that according to our findings the ‘R-G’ vector is useful for extracting health information from PPG. Thus, giving a simpler technique for HR extraction. We also elaborated future scope for this technique.

PPG and its related methods, specifically the one we discussed in this paper have the capability of being portable and noncontact in nature as it relies on light information which can be easily captured by consumer grade cameras. This makes the technique viable for regular health monitoring without the need of a medical expert.

References

[1] H. Liu, Y. Wang and L. Wang, “A Review of Non-Contact, Low-Cost Physiological Information Measurement based on Photoplethysmographic Imaging,” in *Proc. 34th Annu. Int. Conf. IEEE Eng. in Med. and Biology Society*, IEEE, 2012, pp. 2088-2091.

[2] P. Sahindrakar, G. de Haan and I. Kirenko, "Improving Motion Robustness of Contact-less Monitoring of Heart Rate Using Video Analysis," thesis of the Dept. of Math. and Comp. Science, Eindhoven Univ. of Tech., Eindhoven, Netherlands, Aug. 2011.

[3] Wikipedia, The Free Encyclopedia (2014, April 7). *Photoplethysmogram* [Online]. Available: http://en.wikipedia.org/wiki/Photoplethysmogram

[4] M.-Z. Poh, D. J. McDuff and R. W. Picard, “Noncontact, automated cardiac pulse measurements using video imaging and blind source separation.” *Optics Express*, vol. 18, no. 10, pp. 10762-10774, 2010.

[5] P. Comon, “Independent component analysis, a new concept?” *Signal Processing*, vol. 36, no. 3, pp. 287–314, 1994.

[6] J.-F. Cardoso, “High-order contrasts for independent component analysis,” *Neural Computation*, vol. 11, no. 1, pp. 157–192, 1999.

[7] A. Hyvärinen, “Fast and Robust Fixed-Point Algorithms for Independent Component Analysis,” *IEEE Trans. Neural Networks*, vol. 10, no. 3, pp. 626-634, 1999.

[8] M.-Z. Poh, D. J. McDuff and R. W. Picard, “Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7-11, 2011.

[9] T. E. Brown, L. A. Beightol, J. Koh and D. L. Eckberg, “Important influence of respiration on human RR interval power spectra is largely ignored,” *J. Appl. Physiol.*, vol. 75, no. 5, pp. 2310–2317, 1993.

[10] L. Wei, Y. Tian, Y. Wang, T. Ebrahimi and T. Huang, “Automatic Webcam-Based Human Heart Rate Measurements Using Laplacian Eigenmap,” in *Proc. 11th Asian Conf. Computer Vision*, 2013, vol. 7725, pp. 281-292.

[11] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*, 2nd ed. Sebastopol, CA: O'Reilly, 2008.

[12] J.-Y. Bouguet, "Pyramidal Implementation of the Affine Lucas Kanade Feature Tracker Description of the algorithm," Microprocessor Research Labs, Intel Corp., Santa Clara, CA, 5, 2001.

[13] X. Gonsalves and M. M. Sardeshmukh, “A Review of Processing Techniques for Non-invasive Photoplethysmographic Extraction of Health Parameters,” *Int. J. Adv. Research Elect. Electron. Instrum. Eng., IJAREEIE* [Online], vol. 3, no. 2, pp. 7155–7162, 2014.

[14] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proc. 2001 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, 2001, vol. 1, pp. 511.

[15] R. Lienhart, and J. Maydt, “An extended set of Haar-like features for rapid object detection,” in *Proc. of 2002 Int. IEEE Conf. Image Processing*, 2002, vol. 1, pp. 900.

[16] P. Shi, V. A. Peris, A. Echiadis, J. Zheng, Y. Zhu, P. Y. S. Cheang and S. Hu, "Non-contact reflection photoplethysmography towards effective human physiological monitoring," *J. Med. Biol. Eng.*, vol. 30, no. 3, pp. 161-167, 2010.

[17] M.-Z. Poh, N. C. Swenson and R. W. Picard, “Motion-tolerant magnetic earring sensor and wireless earpiece for wearable photoplethysmography,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 786-794, 2010.

[18] C. G. Scully, J. Lee, J. Meyer, A. M. Gorbach, D. G.-Fraser, Y. Mendelson, and K. H. Chon, “Physiological Parameter Monitoring from Optical Recordings With a Mobile Phone,” *IEEE Trans. Biomed. Eng.*, vol. 59, no. 2, pp. 303-306, 2012.

