Original Paper

Perfusion

Perfusion
1–7

DOI: 10.1177/0267659114564206
prf.sagepub.com
SAGE

A temporal video-processing method to improve heart rate estimation

M DaneshiKohan[1] and AM NasrAbadi[2]

Abstract

Background: The remote monitoring of vital signs has grown up in recent decades. One of the most frequent examinations in healthcare monitoring is cardiac pulse measurement. Over the past few years, the application of image processing has increased in the field of medical sciences. In this study, a spatio-temporal video processing, namely Eulerian video magnification, was used to develop equipment that can extract minute details of substances. Firstly, the video is input and then spatial decomposition applied on that, followed by temporal filtering to the frames. Generally, the goal of this research was to reveal in videos informative signals that are difficult to see with the naked eye.

Methods: The videos from faces and hands of 32 subjects were captured simultaneously. In addition, to evaluating the results as ground truth, the photoplethysmograph (PPG) signals from PPG sensors that attach to the fingers of subjects were obtained. Subsequently, they were analyzed by applying an Eulerian process for the videos of the hands and face and the signals were extracted from the videos. The heart rate from signals based on 3 methods, including Fast Fourier Transform (FFT), zero crossing and peak detection was extracted.

Results: The results were compared with the heart rate that we extracted from PPG signals by the complement of absolute normalized difference (CAND) index. Suitable results based on the mean CAND of 32 subjects were obtained for the Eulerian algorithm by peak detection for the face videos (92.15%). The heart rate variability for the obtained signals (face, hand and PPG) was compared as it is considered as a proper parameter for medical applications. The Bland-Altman index was also measured for combined graphical and statistical interpretation of the two measurement techniques.

Conclusion: The results revealed that there is a significant improvement in extracting the heart rate by the Eulerian method with peak detection. This technique can work out in real time to extract the heart rate in a contactless way.

Keywords

heart rate; Eulerian video magnification; temporal filtering; spatial decomposition; Bland-Altman index; CAND index

Introduction

Over the past few decades, the field of image processing has gained many interests. Today, it has applications in various fields, such as satellite communication, experimental physics, chemistry, agricultural and medical sciences. Over the past few years, the application of image processing has increased in the field of medical sciences. It has aided scientists to develop equipment that can extract minute details of substances and assists them to identify the problem precisely.[1-3] Generally, contact-free measurement of a cardiac pulse based on the analysis of a video is a new technique in medicine.[4-7]

The human visual system has limited spatio-temporal sensitivity, but many signals that fall below this capacity can be informative. For example, the color of human skin varies slightly with blood circulation. This variation, while invisible to the naked eye, can be exploited to extract pulse rate.[1, 8-10] Hence, efficient tools are needed in order to analyze these unapproached signals, usually containing significant information. Motion magnification is a technique that acts like a microscope for visual motion. It can amplify subtle motions in a video sequence,

[1]Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
[2]Department of Biomedical Engineering, Shahed University, Tehran, Iran

Corresponding author:
Marzieh Daneshi Kohan
Department of Biomedical Engineering
Science and Research Branch
Islamic Azad University
Tehran, 14189
Iran.
Email: marzieh.daneshikohan2015@gmail.com

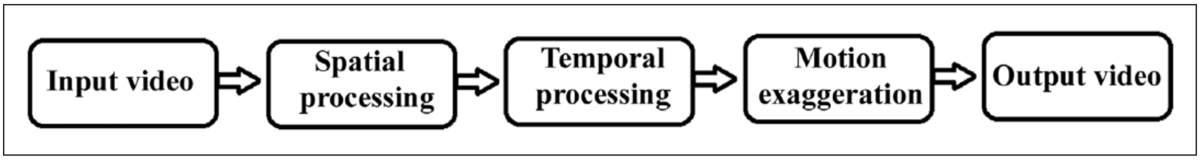

Figure 1. Overview of the Eulerian video magnification framework. The system first decomposes the input video sequence into different spatial frequency bands and applies the same temporal filter to all bands. The filtered spatial bands are then amplified by a given factor, alpha, added back to the original signal and collapsed to generate the output video.

allowing for visualization of deformations that would otherwise be invisible.[11] The method called Eulerian video magnification automatically selects and then amplifies a band of temporal frequencies that includes plausible human heart rates (Figure 1). The amplification reveals the variation of redness as blood flows through the face. For this application, temporal filtering needs to be applied to lower spatial frequencies (spatial pooling) to allow such a subtle input signal to rise above the camera sensor and quantization noise.

The idea of performing physiological measurements on the face was first postulated by Pavlidis and associates [12] and later demonstrated through the analysis of facial thermal videos.[13,14] Previous attempts have been made to unveil imperceptible motion in videos.[11] In addition, temporal processing has been used previously to extract invisible signals and to smooth motions. Poh et al.[15] extracted a heart rate from a video of a face, based on the temporal variation of the skin color, which is normally invisible to the human eye. Based on this idea, Poh et al. tried to apply independent component analysis (ICA) on video images of the human face to extract the underlying blood volume pressure (BVP) for cardiac pulse rate measurement.[16] Similarly, Lewandowska obtained the heart rate directly from a webcam with the particular component analysis (PCA.)[2] According to the Beer-Lambert law, reflected light intensity traveling through facial tissue varies non-linearly with distance. Neither ICA nor PCA could extract the pure BVP from the collected data as both of them are based on linear hypotheses. Lan Wei et al. proposed a novel, webcam-based method to measure human heart rate, using the Laplacian Eigen Map.[17] In this study, a combination of spatial and temporal processing of videos was used to amplify subtle variations. Based on the proposed method, some videos were captured, processed with color amplification and analyzed with PPG signal as ground truth.

Materials and methods

Experimental setup

A basic webcam embedded in a laptop and a camera to record videos of the face and hand was employed. All videos were recorded in color RGB (high-resolution red, green and blue) at 30 frames per second (fps) with a pixel resolution of 640×480 and saved in the audio video interleave (AVI) format on the laptop. Generally, 32 participants (6 males, 26 females) with an age of 10-75 years were enrolled for the experimental measurements. The experiments were conducted indoors and with a varying amount of ambient sunlight entering through windows as the only source of illumination. The subject sat at a table in front of the window and with the laptop at a distance of approximately 400 cm from the webcam and, for recording videos of the hand, we used a camera at a distance of approximately 50 cm from the left hand. During the experiment, participants were asked to keep still and face the webcam while their video was recorded for 15 minutes. Videos of the face and hands were recorded simultaneously; their PPG signals were also recorded using a biofeedback system at a sampling rate of 256 Hz attached to the third finger of the left hand. For generating variable heart rate, music with two different rhythms was played and a headphone was placed on their head and, for the first 5 minutes of the protocol anything was played for the neutral baseline state. After that, they heard relaxing music for five minutes and, for the last five minutes, they heard fast music.

Eulerian method

The approach proposed in this study for motion magnification is emboldened by the Eulerian perspective, where properties of fluid, such as pressure and velocity, evolve over time. In this case, the variation of pixel values over time was amplified and studied in a spatially multi-scale manner. In the Eulerian approach to motion magnification, motion is not explicitly estimated, but rather exaggerated by amplifying temporal color changes at fixed positions. The video sequence is first decomposed into multiple spatial frequency bands. These bands might be magnified differently on the grounds that either they might contain spatial frequencies for which the linear approximation used in our motion magnification does not hold or they might exhibit different signal-to-noise ratios. In the latter case, the amplification is reduced for these bands to suppress artifacts. When the goal of spatial processing is simply to enhance temporal signal-to-noise ratio by pooling multiple pixels, then the video frames are spatially low-pass filtered and down-sampled for computational efficiency. In the general case, a full Laplacian pyramid is computed. Then, temporal processing is performed on each spatial band. The time series is considered corresponding to the value of a pixel in a frequency band and a band-pass filter is applied to extract the frequency bands of interest. The temporal processing is identical for all spatial levels and for all pixels within each level. After that, the extracted band-passed signal is multiplied by a magnification factor that can be specified by the user and may be assuaged automatically, depending on certain parameters. Then, the magnified signal is added to the original and collapses the spatial pyramid to obtain the final output.[1,18]

Obtaining the heart rate from videos

All the videos of the face and hand and physiological recordings were analyzed off-line, written in MATLAB 2013 (MathWorks Inc., Natick, MA, USA).[2,19,20] The open source computer vision (Open CV) (Matlab, MathWorks Inc.) face detection algorithm which is a cascade of boosted classifier was used. The algorithm returned the x and y-coordinates along with the height and width that define a box around the face. The center 60% width and full height of the box as the region of interest (ROI) for our subsequent calculations was selected. Therefore, the videos were cropped as the input of the Eulerian algorithm. After that, a color amplification was applied to the videos of 10 seconds by 10 seconds and extracted signals of the heart rate from it at the same gorge. Thus, there are 10-second signals as output. The videos were spatially filtered using a full Laplacian pyramid decomposition to increase the signal-to-noise ratio. Next, the video frames were temporarily filtered successively, using an ideal band-pass filter with a range of 12 Hz for extracting the variations caused by blood circulation. At this procedure, the signals were analyzed to extract the heart rate. Three common methods, namely FFT, zero crossing and peak detection were employed to extract the heart rate from three signals; face- and hand-estimated signals (with Eulerian method (EVM)) and PPG signals. Therefore, a set of 90 numbers of the heart rate for each person for more analysis (15*60/10) was used.

Statistics

The Bland-Altman index, an index for combined graphical and statistical interpretation of the two measurement techniques, the differences between estimates from the EVM and the PPG sensors, was plotted versus the averages of both systems. The mean and standard deviation (SD) of the differences, the mean of the absolute differences and the 95% limits of agreement (±1.96 SD) were calculated. The root mean squared error (RMSE), Pearson's correlation coefficients and the corresponding p-values were calculated for the estimated heart rate from the EVM and the PPG by three different methods (FFT, zero crossing and peak detection).

Results

CAND index

For evaluating the results, the complement of absolute normalized difference (CAND) index that was introduced by Pavlidis[12] was used.

$$\text{CAND}=1-\frac{|\text{GT-EVM}|}{\text{GT}} \tag{1}$$

GT stands for Ground Truth and EVM for Eulerian method. The results are stated in beats per minute (BPM). These indices for comparing with ground truth were 82.80% and 82.89%, on average, for the face videos of 32 subjects with the FFT and zero crossing methods, respectively, and 86.71% and 87.73%, on average, for the hand videos of 32 subjects with the FFT and zero crossing methods, respectively. Further pre-processing, such as removing singular points from the signals (here we define "singular points" as the points which are greater than (or less than) 10 times the average) was applied. After the removal, the average of two adjacent points to interpolate the vacant position were used. The peak detection technique was employed for extracting the heart rate. The results of the CAND index were 92.15% and 92.13%, on average, for the face and hand videos of the 32 subjects, respectively. These results showed a growth in the CAND index in comparison with previous work.[12] Figure 2 shows an example of applying this method on a video and a sample of the signal extracted from 10 seconds of a video.

In addition, the CAND index for the face and hand signals was used to compare with themselves. It was 85.73%, 83.87% and 89.44%, on average, for the face and hand videos of 32 subjects based on zero crossing, FFT and peak detection, respectively.

Statistic evaluations

Bland-Altman plots were used for combined graphical and statistical interpretation of the two measurement techniques. The differences between estimates from the Eulerian method and the PPG sensor were plotted verses the averages of both systems. The mean and SD of the differences, mean of the absolute differences and 95% limits of agreement (±1.96 SD) were calculated.

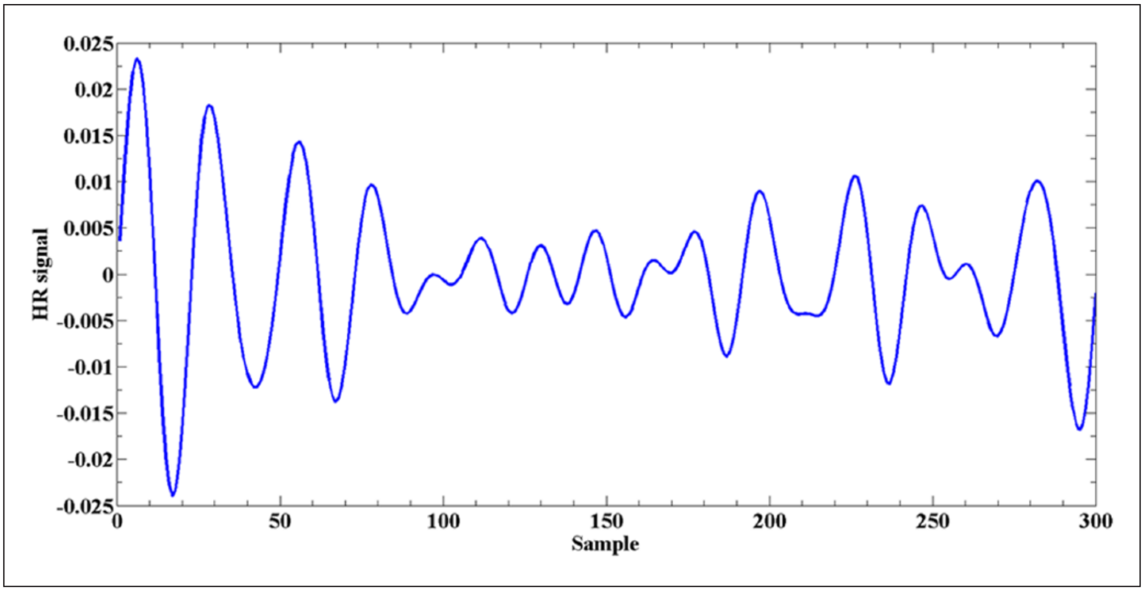

Figure 2. A sample signal of heart rate based on applying the Eulerian method on a subject face video in 10 seconds.

The RMSE, Pearson's correlation coefficients and the corresponding p-values were calculated for the estimated heart rate from the Eulerian method and the PPG sensor. To illustrate the effect of peak detection, we first evaluated the accuracy of heart rate measurements obtained directly from FFT and zero crossing and then compared them with the peak detection technique. When the agreement between 2880 pairs of measurements from 32 participants was tested by Bland-Altman analysis for FFT, the mean bias d was 0.09 bpm with 95% limits of agreement −9.26 to 9.26 bpm (Figure 3a) and, for zero crossing, the mean bias d was 0.095 bpm with 95% limits of agreement −12.78 to 12.78 bpm (Figure 3b). Using the proposed method with peak detection to recover the heart rate reduced the error; d was −0.04 bpm with 95% limits of agreement −3.22 to 3.22 bpm. The RMSE was reduced from nearly 9.13 and 14.65 (obtained from FFT and zero crossing, respectively) to 3.23 bpm and the correlation coefficient r increased from 0.77 to 0.89 ($p<0.001$ for both).

Scatter plot

For more analysis, scatter plots for comparing two measurements were used. For example, the measurements of heart rate between a webcam and a reference sensor was compared. In Figure 4 these plots, on average, for 32 subjects are displayed.

Discussions

In the proposed method, a standard video has been taken as the input and magnified to amplify the small motion in it which was invisible to human eye. This method processes pixels at specific positions in a video where it has low frequencies and amplifies them to see small changes in the video. Our method exhibited a strong degree of linear correlation to a standard PPG pulse measurement method. Furthermore, work improvements in Eulerian color amplification by using peak detection to extract heart rate better were investigated. The method of how to take information from these small videos changes can be applied to different scenarios, such as security, medical, sport or weather purposes.

We demonstrated that, using the STF method (spatio-temporal filtering of videos) makes it possible to extract the variations caused by blood circulation in the face. After extracting these variations, it is possible to estimate the heart rate. From the output of temporal filters, a signal is extracted which corresponds to the variations of the ROI. Three methods were proposed to estimate the heart rate from this signal: STF-FFT, STF-zero crossing, and STF-peak detection, in which the results of peak detection outperformed the results of FFT and zero crossing. The results of our experiment demonstrated a growth in average CAND index in comparison with the reported results.[13] The statistics outcomes of the 32 donators are listed in Table 1. Furthermore, for statistical analysis, the agreement between 2880 pairs of measurements from 32 participants was tested by Bland-Altman analysis; the mean and SD of the differences, the mean of the absolute differences and 95% limits of agreement (±1.96 SD) were calculated. The RMSE, Pearson's correlation coefficients and the corresponding p-values were calculated for the estimated heart rate from the Eulerian method and a reference.

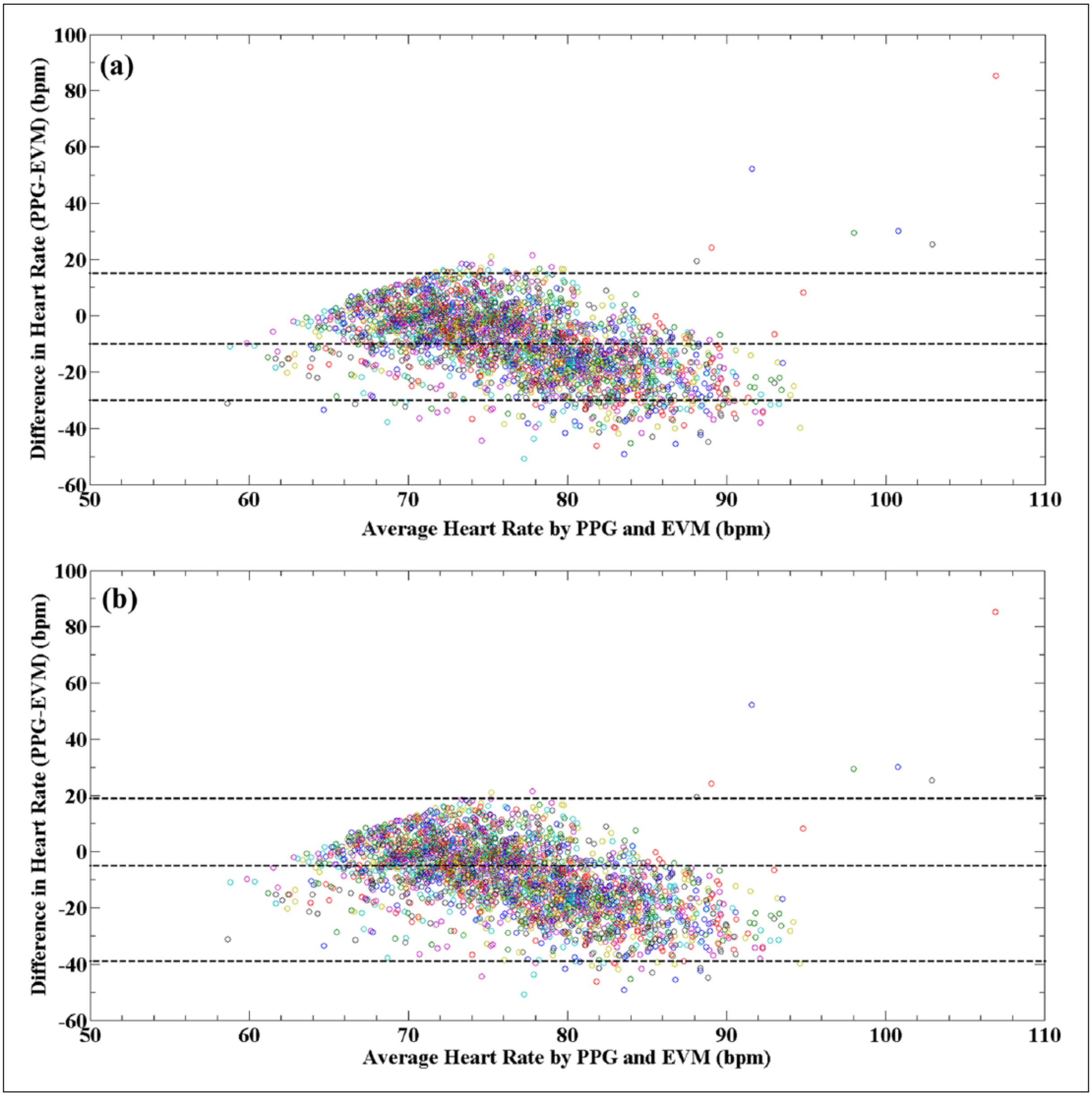

Figure 3. A sample Bland-Altman plot to compare the Eulerian method with (a) the FFT technique and (b) peak detection.

As the method is contact-free, passive and automated (face detection), it opens the way for sustained physiological measurements in the most transparent manner. Almost all the conventional methods require contact and, hence, they compromise the subject's comfort and mobility, especially in long observational periods. Moreover, measurements by these methods are strongly affected by movement artifacts, with no easy way to counteract them. Initially, our method may find applications in sleep studies as well as sport training. In all these cases, long observations are required and intrusive sensing is undesirable since it interferes with the subject's function. Therefore, a contact-free, highly automated, pulse-measurement method will bring considerable value for us.

Conclusions

In this study, a novel methodology for recovering the cardiac pulse rate from video recordings of the human face was developed and demonstrated an implementation, using a simple webcam with ambient daylight providing illumination. To our knowledge, this is the first demonstration of a low-cost method for non-contact heart rate measurements that is automated and motion-tolerant. Given the low cost and widespread availability

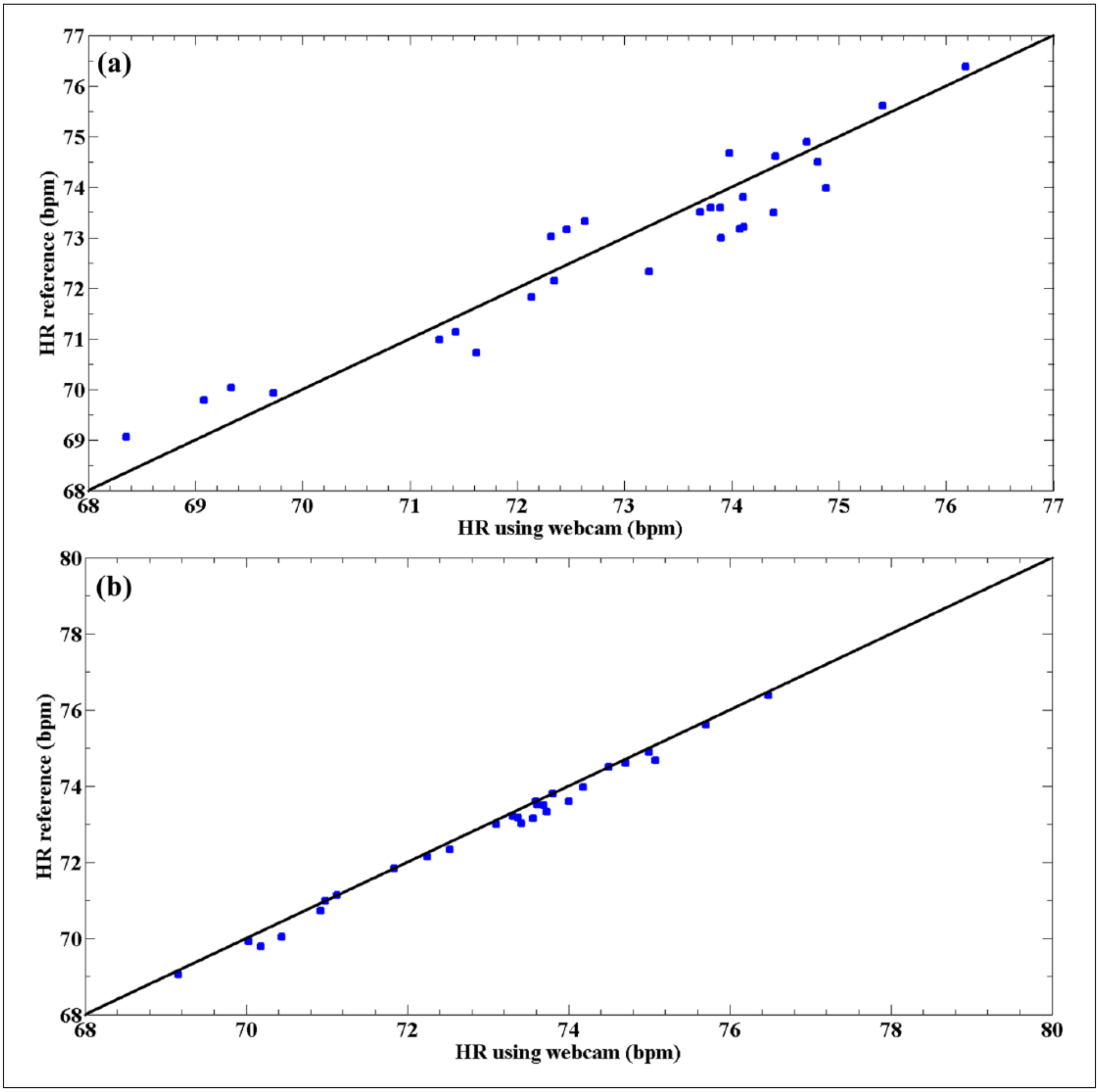

Figure 4. Scatter plot for mean heart rate of 32 subjects; (a) the Eulerian method with FFT, (b) the Eulerian method with peak detection.

Table 1. The statistics results of 32 participants which were reported by Bland-Altman analysis.

Statistics	FFT	Zero crossing	Peak detection
Number of measurement pairs	2880	2880	2880
Mean bias (bpm)	0.09	0.095	−0.04
Mean absolute bias (bpm)	10.23	12.07	5.44
SD of bias	6.25	17.35	1.66
RMSE	9.13	14.65	3.23
Correlation coefficient	0.77	0.75	0.89

FFT: Fast Fourier Transform; bpm: beats per minute; SD: standard deviation; RMSE: root mean squared error.

of webcams, this technology is promising for extending and improving access to medical care. In the future it is recommended to investigate the system using wavelet and other video decomposition, with different filter structures to see improvements in the system.

Declaration of Conflicting Interest

The authors declare that there is no conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

1. Wu H-Y, Rubinstein M, Shih E, Guttag JV, Durand F, Freeman WT. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans Graph* 2012; 31: 65.
2. Lewandowska M, Ruminski J, Kocejko T, Nowak J. Measuring pulse rate with a webcam—a non-contact method for evaluating cardiac activity. *Computer Science and Information Systems (FedCSIS), 2011 Federated Conference on; 2011: IEEE; 2011*. p. 405–410.
3. Wu H-Y, Rubinstein M, Shih E, Guttag J, Durand F, Freeman W. Eulerian Video Magnification for Revealing Subtle Changes in the World. *ACM Trans. Graph. Proceedings SIGGRAPH* 2012; 31.
4. Abdi M, Karimi A. A computational electrical analogy model to evaluate the effect of internal carotid artery stenosis on Circle of Willis efferent arteries pressure. *J Biomater Tissue Eng* 2014; 4: 749–754.
5. Karimi A, Navidbakhsh M, Yamada H, Razaghi R. A nonlinear finite element simulation of balloon expandable stent for assessment of plaque vulnerability inside a stenotic artery. *Med Biol Eng Comput* 2014; 52: 589–599.
6. Karimi A, Navidbakhsh M, Faghihi S. A comparative study on plaque vulnerability using constitutive equations. *Perfusion* 2014; 29: 178–183.
7. Sarod LC, Mandaogade NN. Review on video motion magnification. *International Journal of Advanced Research in Computer Science and Management Studies* 2014; 2: 480–484.
8. Abdi M, Karimi A, Navidbakhsh M, Hassani K, Faghihi S. Modeling of cerebral aneurysm using equivalent electrical circuit (Lumped Model). *Perfusion* 2014; 29: 142–152.
9. Abdi M, Karimi A, Navidbakhsh M, Pirzad Jahromi G, Hassani K. A lumped parameter mathematical model to analyze the effects of tachycardia and bradycardia on the cardiovascular system. *Int J Numer Model* 2014. In press. DOI: 10.1002/jnm.2010.
10. Sahindrakar P. (Thesis) *Improving motion robustness of contact-less monitoring of heart rate using video analysis*. Department of Mathematics and Computer Science, University of Technology, Eindhoven, The Netherlands, 2011.
11. Liu C, Torralba A, Freeman WT, Durand F, Adelson EH. Motion magnification. ACM Transactions on Graphics (TOG) 2005; *ACM*: 2005. p. 519–526.
12. Pavlidis I, Dowdall J, Sun N, Puri C, Fei J, Garbey M. Interacting with human physiology. *Computer Vision Image Understanding* 2007; 108: 150–170.
13. Garbey M, Sun N, Merla A, Pavlidis I. Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. *Biomed Eng, IEEE Transactions on* 2007; 54: 1418–1426.
14. Fei J, Pavlidis I. Thermistor at a distance: unobtrusive measurement of breathing. *Biomed Eng, IEEE Transactions* 2010; 57: 988–998.
15. Poh M-Z, McDuff DJ, Picard RW. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Exp* 2010; 18: 10762–10774.
16. Poh M-Z, McDuff DJ, Picard RW. Advancements in noncontact, multiparameter physiological measurements using a webcam. *Biomed Eng, IEEE Transactions* 2011; 58: 7–11.
17. Wei L, Tian Y, Wang Y, Ebrahimi T, Huang T. Automatic webcam-based human heart rate measurements using laplacian eigenmap. *Computer Vision–ACCV 2012*: Springer; 2013: 281–292.
18. Sarode LC, Mandaogade N. Review on video motion magnification. *Int J Computer* 2014; 2: 89–95.
19. Kwon S, Kim H, Park KS. Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone. *Conf Proc IEEE Eng Med Biol Soc* 2012; 2012: 2174–2177.
20. Karimi A, Navidbakhsh M, Razaghi R. Plaque and arterial vulnerability investigation in a three-layer atherosclerotic human coronary artery using computational fluid-structure interaction method. *J Appl Phys* 2014; 116: 064701–064710.