Expert Systems With Applications 54 (2016) 88–94

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Spatio-temporal filtering of thermal video sequences for heart rate estimation

Kian Hamedani[a], Zahra Bahmani[a], Amin Mohammadian[a,b,*]

[a] Research Center of Intelligent Signal Processing (RCISP), Tehran, Iran
[b] Department of Biomedical Engineering, Amirkabir University of Technology, 424 Hafez Ave, Tehran 15875-4413, Iran

ARTICLE INFO

Keywords:
Spatio-temporal filtering
Thermal videos
Heart rate
Non-contact

ABSTRACT

In this paper, a novel method for non-contact measurement of heart rate using thermal imaging was proposed. Thermal videos were recorded from subjects' faces. The measurements are performed on three different areas: the whole face, the upper half of the face and the supraorbital region. A tracker was used to track these regions to make the algorithm invulnerable to the subject's motion. After tracking, the videos were spatially filtered using a full Laplacian pyramid decomposition to increase the signal to noise ratio; next, the video frames were successively temporally filtered using an ideal bandpass filter for extracting the thermal variations caused by blood circulation. Finally, the heart rate was calculated by using two methods including zero crossing and Fast Fourier Transform. For evaluating the results, the complement of absolute normalized difference (CAND) index was used which was introduced by Pavlidis. This index was 99.42% in the best case and 92.472% in average for 22 subjects. These results showed a growth in CAND index in comparison with previous work. Zerocrossing outperformed FFT because of the nonstationary nature of thermal signals. Another benefit of our method is that, the videos are taken from the face unlike most of the studies that take it from the neck and Carotid. Neck and carotid are less accessible than faces. Finally, the optimum ROI for estimating the heart rate from face was identified.

1. Introduction

In various fields of science, there is a growing demand to measure the parameters and application of human physiology. In medical applications, human physiological values are used for monitoring patients both in hospitals and homes, caring for elderly, and rehabilitation (Garbey, Sun, Merla, & Pavlidis, 2007; Watanabe et al., 2005). In the other fields such as psychology (Bailón, Mainardi, Orini, Sörnmo, & Laguna, 2010; Hjortskov et al., 2004; Jamieson, Nock, & Mendes, 2012; Ogorevc, Podlesek, Geršak, & Drnovšek, 2011; Shalev et al., 1998), sport (Mesleh, Skopin, Baglikov, & Quteishat, 2012; Pušnik & Čuk, 2014) and automotive industry (Fördős, Bosznai, Kovács, Benyó, & Benyó, 2007; Singh, Conjeti, & Banerjee, 2013), and even in the field of human–machine interface (Piechulla, Mayser, Gehrke, & König, 2003), physiological parameters of human are widely used.

Measuring the cardiac pulse rate is one of the important human physiological parameters. So many contact-based methods have been proposed to estimate the subject's heart rate. Among them Electro-Cardio-Graphy (ECG) is the most famous one (Garbey et al., 2007). ECG requires at least three electrodes to be attached to the subject's skin (Garbey et al., 2007) which can cause discomfort for the subject. Another contact-based approach which is so popular is piezoelectric transducer. This method is based on measuring the pulse through the mechanical effects caused by blood flow in the vessels (Bourlai, Buddharaju, Pavlidis, & Bass, 2009). The drawback of this method is its sensitivity to subject's motion (Bourlai et al., 2009).

Photoplethysmography (PPG) is another contact-based method which depends on the optical properties of the desired skin area, the idea behind PPG is this: Near-infrared light is emitted on the skin some of this light is absorbed, and some other is reflected, cardiac pulse rate corresponds to the backscattered light (Garbey et al., 2007). Doppler ultrasound (Reymond, Merenda, Perren, Rüfenacht, & Stergiopulos, 2009) and Laser Doppler sensing (Poh, McDuff, & Picard, 2011) are more advanced than aforementioned methods which are used in measuring the cardiovascular pulse. All the methods discussed so far are contact-based or require subject's cooperation, but in some cases the subject cannot cooperate or in some psycho-physiological applications

* Corresponding author at: Research Center of Intelligent Signal Processing, No. 12, Bisheh Dd.End, Motahari- Sohrevari Crossroad, Tehran, Iran. Tel.: +982183857000; fax: +982183857100.
E-mail addresses: Hamedani.kian64@gmail.com (K. Hamedani), zabahmani@gmail.com (Z. Bahmani), a.mohammadian@aut.ac.ir, amohammadian@elenoon.ir (A. Mohammadian).

http://dx.doi.org/10.1016/j.eswa.2016.01.022
0957-4174/

(Garbey et al., 2007) imposing the subject to cooperate and attaching electrodes and sensors to his or her skin can affect the emotions and the results of psycho-physiological experiment will be ruined. Besides, in order to apply measurements in a wide range of applications, the existing contact methods with the known limitations would seem inadequate in some cases (Kranjec, Beguš, Geršak, & Drnovšek, 2014).

The most important advantage which is brought by using non-contact methods is about the people like infants which attaching electrode can cause some irritation effect on their skin. By using non-contact methods, these harmful effects can be avoided. In these cases, we have to find a way to measure the cardiac pulse with the least interference. Non-contact methods have somehow made this demand possible, one of the first non-contact methods was proposed by Mikhelson, Bakhtiari, Elmer, and Sahakian (2011), a Radar Vital Signs Monitor (RVSM) which uses an active radar detector to detect the waves backscattered from the chest and other areas of the subject containing information about the heart and respiration cycles. The problem of using this method is the side effect on subject's health caused by the active sensors energy focused on the body (Shalev et al., 1998). By using passive Infrared (IR) detectors, the drawback of active detectors mentioned in the last paragraph can be avoided since they collect the energy emitted from subject's skin and nothing is attached to the subject's body (Pavlidis et al., 2007). Blood perfusion in the skin can cause certain periodical thermal changes that can be detected by mid-wave IR cameras (Yang, Liu, Turner, & Wu, 2008).

Pavlidis and his colleagues were the first people that proposed a method for measuring the heart rate from thermal imaging (Garbey et al., 2007). In their method heart rate corresponds to the dominant frequency estimated from averaging the power spectra of the pixels in the desired region of interest (ROI) which can be carotid or superficial vessels (Sun, Garbey, Merla, & Pavlidis, 2005; Fei & Pavlidis, 2010). We have evaluated our proposed method by the results reported in Garbey et al. (2007).

Some other studies have used the visible videos in order to estimate the heart rate (Poh, McDuff, & Picard, 2010; Poh et al., 2011; Xu, Sun, & Rohde, 2014; Yu, Kwan, Lim, Wong, & Raveendran, 2013; Yu, Raveendran, & Lim, 2014 , 2015). In another study (Wu et al., 2012) a method was introduced for revealing and displaying temporal changes in human face due to blood perfusion in non-IR (visible) videos, these temporal changes are invisible by naked eyes, this method was successful in revealing variations of the face color caused by blood circulation in ordinary visible videos so we thought that applying this method on IR videos can also reveal thermal-temporal changes caused by blood circulation on the face. Although in Wu et al. (2012) only a method was described for displaying the color variations of skin but it is possible to extract the heart rate which corresponds to the color variations displayed. In this work, it is the first time that spatio-temporal filtering (STF) has been proposed for estimating heart rate from thermal video sequences. We have also found the optimum ROI in face for estimating heart rate from thermal videos.

This paper is divided into 5 sections. In Section 2 we introduce the setup which we have used to record data including thermal videos and ground-truth (GT) data. Spatio-temporal filtering for heart rate estimation is also introduced for extracting heart rate, and we say how it is possible to estimate the heart rate from the color amplified thermal videos. We also mention to the importance of region of interest (ROI) selection and tracking. In this section, an algorithm is also explained for vessel segmentation in thermal videos. In section three, the tracker which we used was introduced and the results of heart rate estimation are displayed. Section 4 discusses our results and the results of previous study (Garbey et al., 2007).

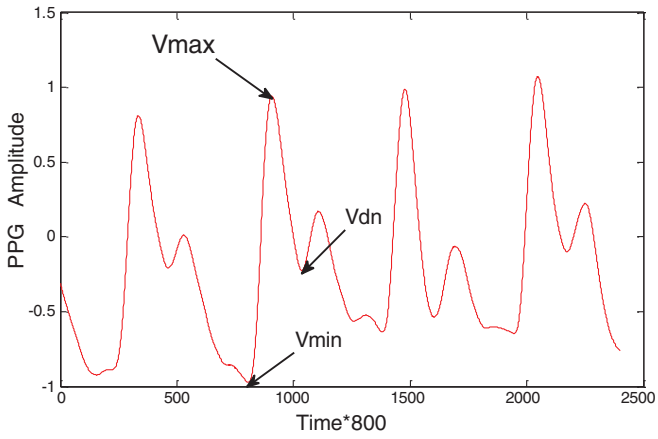

Fig. 1. PPG signal sample recorded from subjects.

2. Method and models

IR band is typically subdivided into three bands: reflective (0.75–3 μm), thermal (3–14 μm) and extreme (14,300 μm) and reflective and thermal bands are again subdivided into four different bands, Near Wave IR (NWIR: 0.75–2.4 μm), Short Wave IR (SWIR: 0.9–2.4 μm), Middle Wave IR (MWIR: 3–5 μm), Long Wave IR (LWIR: 8–14 μm).

The human skin is a good emitter of thermal energy, regardless of the skin color the emissivity factor of skin is 0.98 (Togawa, 1989) and it is supposed to be constant during the test. Object temperature and emissivity are two main factors that determine the object's IR wave radiation (Chekmenev, Farag, Miller, Essock, & Bhatnagar, 2009).

The thermal camera which we used is a MWIR camera with less than 0.01 °C NETD[1] in 30 °C, a 16-bit extended dynamic range with a 320×256 resolution. This camera is of cooling focal plane array (FPA) type, its focal length is 60 mm, and the angle of view is $9.2° \times 7.3°$. The video frames were captured in 25 frames per second rate. The subjects were between 22 and 55 years old (all males). The subjects were seated at 2 m distance frontal to the camera and they stayed relaxed and motionless during the test. The temperature of the room was kept constant about 25 °C. To be sure about having enough data about 15 min of video was recorded for every person but the analysis was performed on a segment with five minutes of length. A portable PPG recording system (Kashef) was used to record the baseline PPG signal. The sampling rate of this system was 800 samples per second. Fig. 1 shows a segment of a PPG signal 3 s; it has 2400 samples. Ground truth PPG signal has been recorded simultaneously with video recording.

2.1. Spatio-temporal filtering (revealing slight variations in videos)

There are so many signals which are beyond of our eye sensitivity but contain a lot of valuable information. One signal which can be mentioned is color variations of face due to blood circulation which can be used to estimate the heart rate (Wu et al., 2012). Eulerian video magnification method has somehow solved this problem. The main idea behind this method is to consider the color values at any pixel as a time series and then applying a temporal filter to the signal in the desired frequency band in order to extract the variations. After the filtering step, the output of the temporal filter is amplified by a factor and then is added to the original frame to make the variations visible (Wu et al., 2012). Fig. 2 shows one example of applying this method on a video,

[1] Noise equivalent temperature difference.

Fig. 2. Magnifying color variations due to blood circulation the first row contains raw frames, the second row contains color magnified frames.

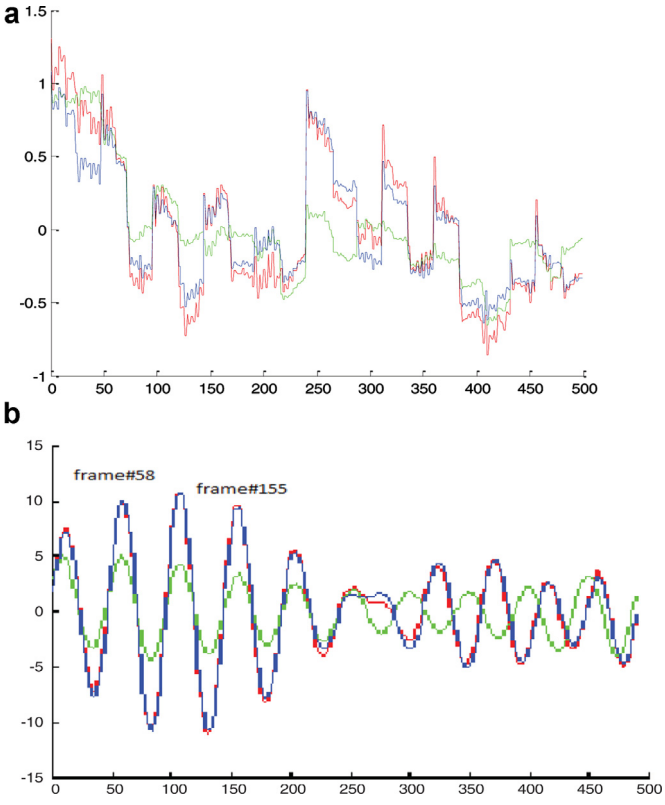

Fig. 3. (A) The average of RGB channels of unmagnified video, and (B) the signals extracted from magnified video.

temporal filtering is performed in a bandwidth including the human pulse rate. Since the circulation of blood in the face is periodic the red channel of face frames becomes bright and dark periodically as it can be seen in the second row of Fig. 2. Fig. 3A shows the average signals of RGB channels of video of Fig. 2 before magnification and Fig. 3B shows those signals of magnified videos. As can be seen in Fig. 3B, there are some peaks in the signal which are related to the moments which the blood pressure in the face is at its peak. These peaks were simultaneous with the moments which the r channel of magnified video had the brightest frames.

2.1.1. Spatial filtering

Before temporal filtering, there is one pre-processing step which is called spatial filtering and its aim is to increase the signal to noise ratio. For computational efficiency the frames of video are spatially low-pass filtered and down-sampled (Wu et al., 2012) by a full Laplacian pyramid (Burt & Adelson, 1983). Laplacian pyramids are useful for reducing the image correlation which is a combination of predictive and transform methods feature. Weighting function is convolved with the image in order to obtain the predicted pixel values. $I_0(i,j)$ is the original image and $I_1(i,j)$ is obtained by applying a lowpass filter on the original image. By subtracting $I_0(i,j)$ from $I_1(i,j)$, the prediction error is achieved:

$$e_0(ij) = I_0(ij) - I_1(ij) \tag{1}$$

In order to generate a Laplacian pyramid, a sequence of error images, $e_0, e_1, \ldots, e_N$ is considered as a Laplacian pyramid, where each error image is the difference between two levels of Gaussian pyramids (Burt & Adelson, 1983).

$$e_l = I_l - EXPAND(I_{l+1}) \quad 1 \le l < N \tag{2}$$

The EXPAND function is defined to expand an $(M+1)$-by-$(N+1)$ array in to a $(2M+1)$-by-$(2N+1)$. This expansion is achieved through interpolating new node values between given values (Burt & Adelson, 1983).

2.1.2. Temporal filtering

After spatial filtering, temporal processing is performed on spatial bands in this way: the values of a pixel are considered as a time series then a bandpass filter in the desired frequency band for example .8–1.6 Hz (48–96 beats per minute) for magnifying the heart pulse is exerted. The extracted bandpass filtered signal is then magnified by factor a, and then it is again added to the original frame to get the final output.

The main factors which have to be adjusted by the user for having a good spatio-temporal filtering result are: (1) α, which is the magnification factor, and (2) l, which is the level of spatial filtering. Increasing a more than a specified value will make this method vulnerable to noise (Wu et al., 2012). In our approach, we extract the signal directly from the videos. We would think that this would increase the accuracy, because in other methods at first video frames are converted to a signal of the pixels average; then, the heart rate is extracted from that signal. However, in our method the signal related to the thermal variation of skin is extracted by using a temporal filter to capture the temporal variations between frames. Fig. 4 depicts the block diagram of STF method for extracting the heart rate.

2.2. Extracting the heart rate

We mentioned that by using spatio-temporal filtering, it is possible to make the slight variations visible, but only observing these variations does not provide any useful information for us. Extracting extra information about the heart rate can be useful as a result of observing the variation of thermal imprints. One step in color magnification is temporal filtering. The output of these temporal filters corresponds to the variations in thermal imprints of the ROI. By applying the temporal filter to the thermal video and extracting its output, it is possible to have a signal related to PPG. A five-minute segment of recorded videos has been selected for analysis.

After that the signal was extracted for estimating the heart rate we used two approaches: (1) using Fast Fourier Transform (FFT) for measuring the predominant frequency, and (2) using zero crossing method for calculating the moments that signal passes zero. The results are stated in Table 1. In order to evaluate the results and compare them with other methods we used complement of the absolute normalized difference CAND index which is introduced in (Garbey et al., 2007):

$$\mathrm{CAND} = 1 - \frac{|\mathrm{GT} - \mathrm{TI}|}{\mathrm{GT}} \tag{3}$$

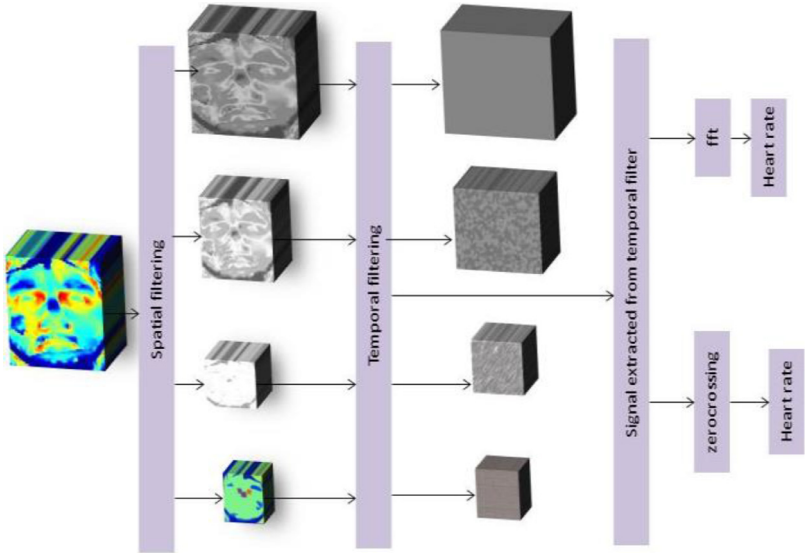

Fig. 4. The block diagram of spatio-temporal filtering algorithm for extracting heart rate.

Table 1
Results of ground truth and proposed method (BPM).

Subject#	GT	Face		Upper half of face		Supraorbital	
		STF-FFT	STF-zerocross	STF-FFT	STF-zerocross	STF-FFT	STF-zerocross
1	80.75	67.69	74.02	62.01	69.26	61.95	69.98
2	73.20	82.64	72.09	82.85	71.95	73.07	71.96
3	70.56	61.02	70.99	60.68	69.60	60.85	70.99
4	82.11	54.80	65.88	54.77	68.92	50.60	71.96
5	57.11	73.70	65.45	73.69	63.79	57.25	63.89
6	79.72	54.67	70.78	55.10	74.80	50.68	69.98
7	72.52	65.64	69.98	66.80	69.00	58.37	68.64
8	92.64	67.12	73.09	64.94	71.76	67.12	68.83
9	68.97	58.96	69.01	62.86	68.87	59.80	66.99
10	78.66	63.03	74 .03	64.20	71.16	55.74	75.75
11	83.99	54.92	72.90	58.27	66.94	54.91	71.96
12	73.46	58.82	69.93	59.73	71.96	64.23	73.04
13	66.00	65.46	67.93	75.89	70.94	65.74	69.95
14	69.84	61.30	77.11	73.85	79.00	86.34	77.12
15	75.30	61.09	72.98	61.02	73.03	68.44	76.92
16	86.94	74.26	78.02	57.79	79.10	57.82	77.06
17	81.28	64.16	71.02	67.94	67.95	57.38	77.83
18	65.28	58.49	66.94	56.81	67.91	58.33	70.85
19	71.54	55.75	74.07	55.76	73.94	55.71	74.97
20	83.47	74.94	77.03	75.04	76.36	72.92	75.04
21	63.10	58.47	66.85	58.35	69.88	58.35	79.03
22	77.55	71.72	74.58	53.64	69.1	70.06	73.75
Average CAND		82.86%	92.46%	80.83%	91.04%	81.77%	91.25%

GT stands for Ground Truth and TI for Thermal Imaging it means the results achieved by thermal imaging. Three ROI's results of heart rate estimation using these ROI's are outlined in Table 1. The results are stated in beats per minute (BPM).

2.3. Region of interest selection and tracking

In Garbey et al. (2007) Pavlidis has mainly focused on Carotid vessel in the neck, since he believes that thermal imprints are more obvious in the Carotid region. The Carotid thermal imprints are said to be more obvious than other areas but considering the Carotid as the main ROI, has its own drawbacks. In some cultures, recording from the neck is not possible because of religious belief and occurs more frequently with women. Usually, in order to perform the procedure, the subject is asked to show his or her neck; thus, their cooperation is required. However, in some conditions when the experiment has to be done with the least possible attention of the subject, this request is not advisable. Thus, we considered it better to consider other parts of the face as the ROI which require less or no cooperation. Some areas like the whole face, the upper half of the face, and the supra-orbital area are good nominees for this purpose. The two last regions are marked as 2 and 3 in Fig. 5. Selecting those can provide information about respiration rate.

2.4. Vessels localization

We implemented a method which was proposed in Buddharaju, Pavlidis, Tsiamyrtzis, and Bazakos (2007) for segmenting the vessels in thermal videos for face recognition. This algorithm is stated in detail: to localize the superficial vasculature, morphological operations are applied. The edges of vessels in thermal videos are not

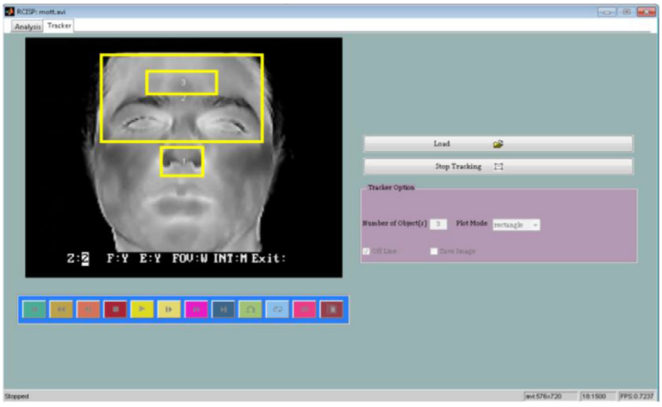

Fig. 5. The graphic user interface which was made for the tracker and ROI's are selected.

as strong; but this problem is resolved using anisotropic diffusion. Object boundaries are enhanced using anisotropic diffusion filter. The anisotropic diffusion filter is mathematically formulated as below:

$$\frac{\partial I(\bar{x},t)}{\partial t} = \nabla(c(\bar{x},t)\nabla I(\bar{x},t)) \tag{4}$$

I is the thermal image, t is time and $\bar{x}$ is the spatial dimension $c(\bar{x},t)$ is diffusion function. The anisotropic diffusion filter is also defined for discrete images.

$$\begin{aligned} I_{t+1}(x,y) = I_t + \tfrac{1}{4}*[&c_{N,t}(x,y)\nabla I_{N,t}(x,y) + c_{S,t}(x,y)\nabla I_{S,t}(x,y) \\ &+ c_{E,t}(x,y)\nabla I_{E,t}(x,y) + c_{W,t}(x,y)\nabla I_{W,t}(x,y)] \end{aligned} \tag{5}$$

One diffusion coefficient is defined for every direction. The diffusion coefficient along the south direction is defined with the following formula:

$$c_{S,t}(x,y) = \exp\left(-\frac{\nabla I^2_{N,t(x,y)}}{k^2}\right) \tag{6}$$

where $I_{N,t} = I_t(x,y+1) - I_t(x,y)$

After the anisotropic filter is implemented on thermal images, edges of the blood vessels are more prominent. To extract the blood vessels, morphological operators are applied on the diffused image. A top hat segmentation method which is the combination of erosion and dilation is applied on diffused images to segment the vessels.

3. Results

3.1. Tracking the ROI

In order to have accurate heart rate estimation, it is necessary for the subjects to be motionless. We ask them to stand motionless, but there are still some slight motions. The tracker which we used is trained by P-N learning method (Kalal, Mikolajczyk, & Matas, 2012).

3.2. Heart rate estimation

The average CAND index of STF-FFT method is around 81% but the average CAND index of STF-zerocrossing method is around 92%. Fig. 6 is a signal extracted from thermal videos using spatio- temporal filtering; its length is about 3 s, and it shows thermal imprints variation in time, which is analogous to the PPG.

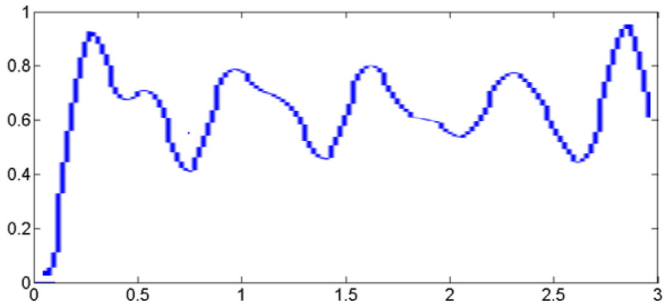

Fig. 6. The signal extracted from the output of temporal filter.

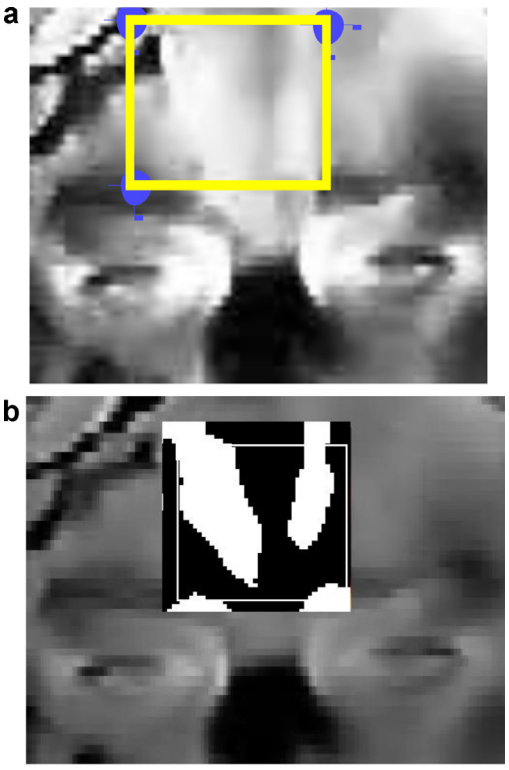

Fig. 7. (A) Thermal imprints of supraorbital region, and (B) vessels existing in supraorbital region extracted.

3.3. Vessels extraction

After implementing the algorithm stated in 2.4 on some thermal videos, the results in Fig. 7 were achieved. As it can be seen in Fig. 7B, there are two main vessels in the supraorbital region. We will state the importance of finding the vessels in the supraorbital region in the following section.

4. Discussion

In Garbey et al. (2007), Pavlidis implemented this experiment on 34 subjects and in his experiment the ROI was Carotid region in the neck which is a better area for this purpose. He achieved 85.6985% of overall CAND index on his subjects. In our experiment, although we did not use Carotid region but we have achieved results which are quite better than the results in Garbey et al. (2007), especially the results of zero crossing method.

As it was mentioned our experiment outperforms the experiment in Garbey et al. (2007), although we don't have access to Carotid region which is a better area for extracting the heart rate using thermal imaging. There are a few reasons for this statement.

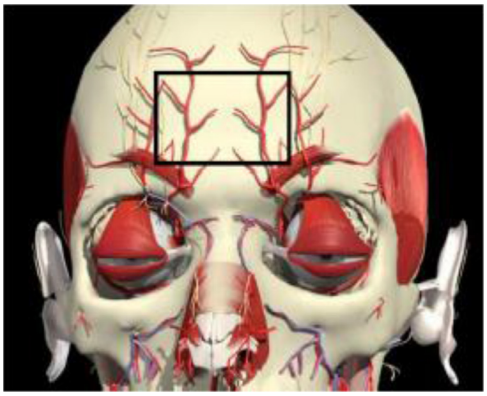

Fig. 8. The pattern of vessels existing in supraorbital region (Moxham, Kirsh, Berkovitz, Alusi, & Cheeseman, 2002).

First, before extracting the heart rate signal from thermal video, we exert Laplacian pyramid spatial filters on videos to increase the signal to noise ratio; the level of spatial filtering depends on the size of the ROI, which for the face area is maximum and for the supra-orbital is minimum, but in Garbey et al. (2007) no such preprocessing is performed on the samples. Second, in our method, the temporal variations which are extracted by temporal filtering then are amplified by a coefficient a which helps to magnify the thermal variations but since in Garbey et al. (2007) no such magnification is done, some thermal variations in skin might not be detected.

Besides the aforementioned benefits of our method, there is also another advantage in our method, as we know when the level of stress increases the PPG envelope varies. In our next work we will investigate this effect on the signals which we have extracted from TI.

4.1. Finding the optimum ROI in face for heart rate estimation

When we take a look at the results, we see that there is not a significant difference between the results achieved from those three ROI's, and they are relatively similar to each other. From this data, we conclude that there is one part in these three regions which dominates the heart rate and has the main role in producing the heart rate. If we know which part has the main role, it is possible to limit the analysis to that ROI. With this limitation, less time is spent, and the processing will be faster without considerable difference in results. By considering the theory of sets, it is possible to say that the part which is in common between all ROI's is the main part in heart rate. The part which is in common between all these ROI's is supra-orbital which is the smallest part. The supraorbital vessels are the main source of blood circulation in the face so it is not necessary to process the whole face to calculate the heart rate, and it is possible to extract the heart rate from thermal videos by just processing the supraorbital region rather than all of the face. The supra-orbital region is an area where arterial blood flow is quite near the skin surface where the thermal variations are stronger than other areas.

As it can be seen in the area of yellow rectangle in Fig. 7, the supraorbital region is brighter than the other parts of the face due to the fact that blood circulating in the vessels in the supra-orbital region makes this part warmer than the other parts of the face. As seen in Fig. 8, there are two main supraorbital vessels in the middle of the forehead, and these vessels have several branches.

The estimated location of blood vessels by this method corresponds with those shown in Fig. 8. So our assumption about the supraorbital region being the dominant part in producing heart rate seems to be proved. Hence, it is absolutely suitable for us to limit the ROI to the supraorbital region in estimating the heart rate from thermal videos according to the aforementioned reasons.

4.2. Nonstationary nature of signals

The question that comes to mind, at this point, is why estimating the heart rate using zerocrossing yields better results than FFT. It is mentioned in Nhan and Chau (2009) that the thermal signals extracted from thermal videos are nonstationary about 50% of the time mostly due to time-varying mean. This nonstationary nature of thermal signals suggests that it is better to analyze thermal signals using nonfrequency domain methods like zerocrossing. These nonfrequency domains include time-frequency methods like wavelet. We tested zerocrossing to determine how it works in analyzing thermal signals extracted from thermal videos. The CAND index showed that the performance of zerocrossing is much better than frequency methods like FFT in counting the number of heart beats. The better performance of zerocrossing is due to the nonstationary nature of signals which methods like FFT fail to analyze.

4.3. Faster dynamics of thermal videos

In Gault and Farag (2013) it was mentioned that visible videos are capable of performing the noncontact heart rate measurements in time periods of 2 s, since thermal videos can capture faster dynamics it is possible to measure the heart rate in time periods even less than 2 s.

5. Conclusion

In this work, it is the first time that spatio-temporal filtering (STF) has been proposed for estimating heart rate from thermal video sequences. We have also found the optimum ROI in face for estimating heart rate from thermal videos. We demonstrated in this paper that with the STF method (spatio-temporal filtering of thermal videos) it is possible to extract thermal 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 thermal variations of the ROI, Two methods were proposed to estimate the heart rate from this signal: (1) STF-FFT, and (2) STF-zerocrossing, and the results of zerocrossing outperformed the results of FFT. The results of our experiment demonstrated a growth in the average CAND index in comparison with the results reported in Garbey et al. (2007). This growth is due to several reasons. First, in our experiment there is one preprocessing step as the spatial filtering to increase the signal to noise ratio, a step not performed in Garbey et al. (2007). Second, in our approach, the thermal variations are extracted directly from the video frames which would be more effective than extracting the variations from the signal, which is achieved by averaging the frame in Garbey et al. (2007). Third, in our approach, it is possible to magnify the thermal variations by a magnification factor which makes the thermal variations visible. In addition to benefits achieved by this method and a growth in the CAND index, extracting a signal which corresponds to the subject's PPG is another benefit of our approach. To date, such a signal has not been introduced in other studies (Garbey et al., 2007). By analyzing this signal, it is possible to extract more features. For extracting heart rate, we found the supraorbital region to be the best ROI. This region also demonstrated that nonstationary signals as well as non-frequency domain methods, like zerocrossing, yielded better results in analyzing signals and estimating heart rate. The Strength of proposed Noncontact method is that this method, using spatial filtering to increase signal to noise ratio, increasing the CAND index, faster dynamics of thermal videos in comparison with

visible videos, extracting a signal which corresponds to PPG and finding the optimal ROI in the face for heart rate measurement. This system could be considered as an expert system which would be potentially replaced with a doctor in the future if further developed. And its weaknesses is that spatial filtering step increases the run time, high prices of thermal cameras, and vulnerability to the motion.

References

Bailón, R., Mainardi, L., Orini, M., Sörnmo, L., & Laguna, P. (2010). Analysis of heart rate variability during exercise stress testing using respiratory information. *Biomedical Signal Processing and Control, 5*, 299–310.

Bourlai, T., Buddharaju, P., Pavlidis, I., & Bass, B. (2009). On enhancing cardiac pulse measurements through thermal imaging. In *Proceedings of the 9th international conference on information technology and applications in biomedicine, ITAB 2009* (pp. 1–4). IEEE.

Buddharaju, P., Pavlidis, I. T., Tsiamyrtzis, P., & Bazakos, M. (2007). Physiology-based face recognition in the thermal infrared spectrum. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 29*, 613–626.

Burt, P., & Adelson, E. (1983). The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications, 31*, 532–540.

Chekmenev, S. Y., Farag, A. A., Miller, W. M., Essock, E. A., & Bhatnagar, A. (2009). Multiresolution approach for noncontact measurements of arterial pulse using thermal imaging. *Augmented vision perception in infrared* (pp. 87–112). Springer.

Fei, J., & Pavlidis, I. (2010). Thermistor at a distance: Unobtrusive measurement of breathing. *IEEE Transactions on Biomedical Engineering, 57*, 988–998.

Fördős, G., Bosznai, I., Kovács, L., Benyó, B., & Benyó, Z. (2007). Sensor-net for monitoring vital parameters of vehicle drivers. *ACTA Polytechnica Hungarica, 4*, 25–36.

Garbey, M., Sun, N., Merla, A., & Pavlidis, I. (2007). Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. *IEEE Transactions on Biomedical Engineering, 54*, 1418–1426.

Gault, T. R., & Farag, A. A. (2013). A fully automatic method to extract the heart rate from thermal video. In *Proceedings of 2013 IEEE conference on computer vision and pattern recognition workshops (CVPRW)* (pp. 336–341). IEEE.

Hjortskov, N., Rissén, D., Blangsted, A. K., Fallentin, N., Lundberg, U., & Søgaard, K. (2004). The effect of mental stress on heart rate variability and blood pressure during computer work. *European Journal Of Applied Physiology, 92*, 84–89.

Jamieson, J. P., Nock, M. K., & Mendes, W. B. (2012). Mind over matter: Reappraising arousal improves cardiovascular and cognitive responses to stress. *Journal of Experimental Psychology: General, 141*, 417.

Kalal, Z., Mikolajczyk, K., & Matas, J. (2012). Tracking-learning-detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 34*, 1409–1422.

Kranjec, J., Beguš, S., Geršak, G., & Drnovšek, J. (2014). Non-contact heart rate and heart rate variability measurements: A review. *Biomedical Signal Processing and Control, 13*, 102–112.

Mesleh, A., Skopin, D., Baglikov, S., & Quteishat, A. (2012). Heart rate extraction from vowel speech signals. *Journal of Computer Science and Technology, 27*, 1243–1251.

Mikhelson, I. V., Bakhtiari, S., Elmer, T. W., & Sahakian, A. V. (2011). Remote sensing of heart rate and patterns of respiration on a stationary subject using 94-GHz millimeter-wave interferometry. *IEEE Transactions on Biomedical Engineering, 58*, 1671–1677.

Moxham, B. J., Kirsh, C., Berkovitz, B., Alusi, G., & Cheeseman, T. (2002). Interactive head & neck (CD-ROM). *Primal pictures*. https://www.primalpictures.com/

Nhan, B., & Chau, T. (2009). Infrared thermal imaging as a physiological access pathway: A study of the baseline characteristics of facial skin temperatures. *Physiological Measurement, 30*, N23.

Ogorevc, J., Podlesek, A., Geršak, G., & Drnovšek, J. (2011). The effect of mental stress on psychophysiological parameters. In *Proceedings of 2011 IEEE international workshop on medical measurements and applications (MeMeA)* (pp. 294–299). IEEE.

Pavlidis, I., Dowdall, J., Sun, N., Puri, C., Fei, J., & Garbey, M. (2007). Interacting with human physiology. *Computer Vision and Image Understanding, 108*, 150–170.

Piechulla, W., Mayser, C., Gehrke, H., & König, W. (2003). Reducing drivers' mental workload by means of an adaptive man–machine interface. *Transportation Research Part F: Traffic Psychology and Behaviour, 6*, 233–248.

Poh, M.-Z., McDuff, D. J., & Picard, R. W. (2010). Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express, 18*, 10762–10774.

Poh, M.-Z., McDuff, D. J., & Picard, R. W. (2011). Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Transactions on Biomedical Engineering, 58*, 7–11.

Pušnik, I., & Čuk, I. (2014). Thermal imaging of hands during simple gymnastics elements on the wooden bar with and without use of magnesium carbonate. *Science of Gymnastics, 67*, 67–72.

Reymond, P., Merenda, F., Perren, F., Rüfenacht, D., & Stergiopulos, N. (2009). Validation of a one-dimensional model of the systemic arterial tree. *American Journal of Physiology-Heart and Circulatory Physiology, 297*, H208–H222.

Shalev, A. Y., Sahar, T., Freedman, S., Peri, T., Glick, N., Brandes, D., et al. (1998). A prospective study of heart rate response following trauma and the subsequent development of posttraumatic stress disorder. *Archives of General Psychiatry, 55*, 553–560.

Singh, R. R., Conjeti, S., & Banerjee, R. (2013). A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals. *Biomedical Signal Processing and Control, 8*, 740–754.

Sun, N., Garbey, M., Merla, A., & Pavlidis, I. (2005). Imaging the cardiovascular pulse. In *Proceedings of IEEE computer* society conference on computer vision and pattern recognition, CVPR 2005*: Vol. 2* (pp. 416–421). IEEE.

Togawa, T. (1989). Non-contact skin emissivity: Measurement from reflectance using step change in ambient radiation temperature. *Clinical Physics and Physiological Measurement, 10*, 39.

Watanabe, K., Watanabe, T., Watanabe, H., Ando, H., Ishikawa, T., & Kobayashi, K. (2005). Noninvasive measurement of heartbeat, respiration, snoring and body movements of a subject in bed via a pneumatic method. *IEEE Transactions on Biomedical Engineering, 52*, 2100–2107.

Wu, H.-Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., & Freeman, W. (2012). Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics, 31*, 65.

Xu, S., Sun, L., & Rohde, G. K. (2014). Robust efficient estimation of heart rate pulse from video. *Biomedical Optics Express, 5*, 1124–1135.

Yang, M., Liu, Q., Turner, T., & Wu, Y. (2008). Vital sign estimation from passive thermal video. In *Proceedings of* IEEE conference on computer vision and pattern recognition, CVPR 2008 (pp. 1–8). IEEE.

Yu, Y.-P., Kwan, B.-H., Lim, C.-L., Wong, S.-L., & Raveendran, P. (2013). Video-based heart rate measurement using short-time Fourier transform. In *Proceedings of* 2013 international symposium on intelligent signal processing and communications systems, ISPACS (pp. 704–707). IEEE.

Yu, Y.-P., Raveendran, P., & Lim, C.-L. (2014). Heart rate estimation from facial images using filter bank. In *Proceedings of* 2014 6th international symposium on communications, control and signal processing, ISCCSP (pp. 69–72). IEEE.

Yu, Y.-P., Raveendran, P., & Lim, C.-L. (2015). Dynamic heart rate measurements from video sequences. *Biomedical Optics Express, 6*, 2466–2480.