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).

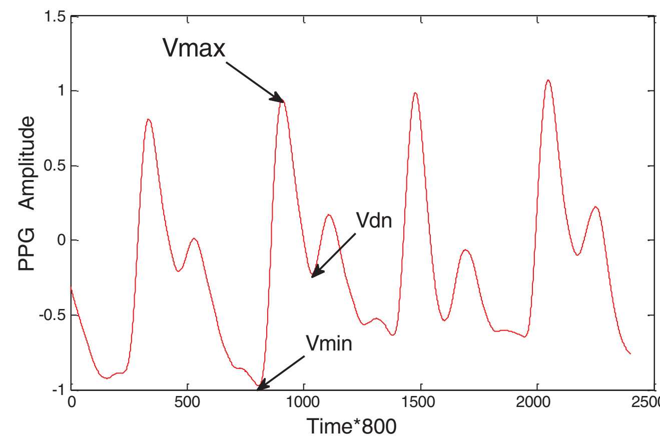

Fig. 1. PPG signal sample recorded from subjects.

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° × 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.