

Remote heart rate variability for emotional state monitoring

Yannick Benezeth¹, Peixi Li¹, Richard Macwan¹, Keisuke Nakamura², Randy Gomez², Fan Yang¹

¹ Le2i FRE2005, CNRS, Arts et Métiers, Univ. Bourgogne Franche-Comté, Dijon, France

² Honda Research Institute, Japan, 8-1 Honcho Wako-shi Saitamaken, Japan

Abstract—Several researches have been conducted to recognize emotions using various modalities such as facial expressions, gestures, speech or physiological signals. Among all these modalities, physiological signals are especially interesting because they are mainly controlled by the autonomic nervous system. It has been shown for example that there is an undeniable relationship between emotional state and Heart Rate Variability (HRV). In this paper, we present a methodology to monitor emotional state from physiological signals acquired remotely. The method is based on a remote photoplethysmography (rPPG) algorithm that estimates remote Heart Rate Variability (rHRV) using a simple camera. We first show that the rHRV signal can be estimated with a high accuracy (more than 96% in frequency domain). Then, frequency-feature of rHRV is calculated and we show that there is a strong correlation between the rHRV feature and different emotional states. This observation has been validated on 12 out of 16 volunteers and video-induced emotions which opens the way to contactless monitoring of emotions from physiological signals.

I. INTRODUCTION

In the last 10 years, many modalities have been studied for the analysis of emotions. Facial expression is one of the most studied modality (*e.g.* in [1]), probably because facial expression recognition is quite easy for people. Recently, research on facial expression analysis has shifted its focus from posed facial expression to spontaneous expressions [2]. Facial features have also been used with 3D [3] or thermal sensors [4]. However, emotion recognition is so complicated that even the relation between internal bodily feelings and externally observable expression is still an open research area. As a consequence, another avenue of research in emotion recognition is based on the analysis of physiological signals. To this end, electrocardiogram, photoplethysmogram, skin temperature, electrodermal activity or electromyography have been studied [5], [6], [7]. These technologies have many advantages but require the use of contact sensors.

Among many physiological features, Heart Rate Variability (HRV) is an interesting indicator of autonomic function. The rhythmic beating of the heart at rest was once believed to be regular, however the heart rhythms are now commonly used to recognize emotions, detect stress or more generally changes in Autonomic Nervous System (ANS) [8], [9], [10]. HRV is derived from the electrocardiogram and is a measurement of the beat-to-beat changes in heart rate. The variability in heart rate is regulated by the synergistic action of the two branches of the ANS, namely the sympathetic and

parasympathetic nervous system. The heart rate represents the net effect of the parasympathetic nerves which slow heart rate, and the sympathetic nerves, which accelerate it. These changes are influenced by emotions, stress and physical exercise [11]. Even if the gold standard for HRV measurement is the electrocardiogram, photoplethysmography (PPG) have been successfully used in several studies (*e.g.* in [12]). Photoplethysmography is a technique for detecting microvascular blood volume changes in tissues. The principle of this technology is actually very simple as it only requires a light source and a photodetector. The light source illuminates the tissue and the photodetector measures the small variations in transmitted or reflected light associated with changes in perfusion in the tissue [13].

Interestingly, it has been shown recently [14] that it is possible to recover the cardiovascular pulse wave measuring variations of back-scattered light remotely, using only ambient light and low-cost vision systems. Since this seminal work, there has been rapid growth in the literature pertaining to remote photoplethysmography (rPPG) techniques (*e.g.* [15], [16]). Interested reader may refer to recent reviews on rPPG [13], [17], [18]. The unobtrusive character of rPPG technology significantly increases patient comfort. However, the measured signal has a notably lower Signal to Noise Ratio (SNR), precisely the ratio between the rPPG signal and all possible noises, than the contact measurements. This observation is fundamental because the pulse waveform signal must be very precise to estimate the HRV. HRV requires accurate estimation of time interval between consecutive peaks in the pulse waveform signal and rPPG signals are very sensitive which may lead to poor HRV estimation if false peaks are detected. However, remote measurement of physiological signals correlated with emotional states is highly desirable for realistic and unbiased experiments on emotions. Decoding human underlying emotions in real-time enables improved human-computer interaction on an emotional level.

In this paper, we present a methodology to monitor emotional state from physiological signals acquired remotely. Frequency-feature of HRV estimated using a simple camera is calculated and we show that there is a strong correlation between this feature and different emotional states. Experiments have been conducted using data collected from 16 volunteers and video-induced emotions.

Section 2 presents the data collection; the rPPG algorithm and the rHRV feature are described in section 3 while

experimental results are given in section 4.

II. MATERIAL AND DATA COLLECTION

All experiments of this paper have been conducted using our own dataset. The video frames were obtained with a custom C# application using a USB3 color camera¹ placed at a distance of about 1m from the subject. The camera collects 1024×786 images in 8-bit uncompressed RGB format at 50 frames per second. In addition, two direct DC light sources were used providing a favorable and stable lighting condition. A CMS50E transmissive pulse oximeter was used to obtain the ground truth PPG data. Figure 1 presents two sample images of our dataset.

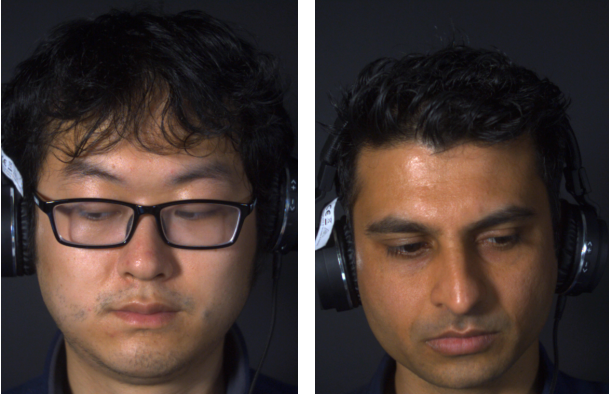


Fig. 1. Two sample images for the experiment from the neutral session.

Sixteen healthy subjects, aged from 23 to 50, participated as volunteers in the experiment. Participants were informed about the protocol and about the purpose of the study. During the experiment, participants were seated in a comfortable chair in a controlled environment while watching videos with emotional content. The affective elicitation was comprised of 2 sessions: after an initial neutral session of 2 minutes, one arousal session of 2 minutes were recorded. For each session, videos from the Internet were shown to the volunteers as video-acoustic stimuli. For the arousal session, selected videos were supposed to elicitate unpleasant emotion of fear or anxiety.

III. HRV FEATURES ESTIMATION FROM VIDEO DATA

The acquired video data are then used to estimate the physiological signals. These signals will be later used for emotional state monitoring. To this end, facial expressions are never used, but we quantify the small variations of light reflected back from the skin due to blood perfusion variations. These pulse waveforms are then used to estimate the HRV.

For each video frame, face detection was first performed using the well-known Viola-Jones face detector [19] implementation provided by the computer vision toolbox of MATLAB. In order to avoid spurious movements of the

detected face, we also use Kanade-Lucas-Tomasi tracking [20] algorithm. Skin detection as formulated by Conaire *et al.* [21] was then performed to select the skin pixels which were then spatially averaged to obtain a triplet of RGB values per frame and concatenated to obtain the RGB temporal trace. The RGB temporal traces are then pre-processed by zero-mean, detrended using smoothness priors approach [22] and band-pass filtered with Butterworth filter (with cut-off frequencies of 0.7 and 3.5Hz). The rPPG signal is then extracted using the chrominance-based method (later called CHROM) [23]. This method applies simple linear combinations of RGB channels and obtains very interesting performance with low computational complexity. Let $y^c(t)$ be the RGB time series obtained after pre-processing, where $c \in \{R, G, B\}$ is the color channel, CHROM method projects RGB values onto two orthogonal chrominance vectors X and Y :

$$\begin{aligned} X(t) &= 3y^R(t) - 2y^G(t), \\ Y(t) &= 1.5y^R(t) + y^G(t) - 1.5y^B(t). \end{aligned} \quad (1)$$

The rPPG signal S is finally calculated with $S(t) = X(t) - \alpha Y(t)$ where $\alpha = \sigma(X)/\sigma(Y)$ and $\sigma(\cdot)$ the standard deviation. Because X and Y are two orthogonal chrominance signals, PPG-induced variations will likely be different in X and Y , while possible motion affects both chrominance signals identically.

Peaks in the rPPG signal were then detected to calculate the remote Heart Rate Variability (rHRV) signal. RHRV signals were resampled to avoid nonuniform sampling before calculating the rHRV periodograms. The same procedure was applied to the PPG signal to obtain the HRV periodograms.

Many features can be estimated from the rHRV temporal signal or the rHRV periodogram. Nardelli *et al.* [24] have presented an interesting comparison of these HRV features in the context of emotion recognition induced by affective sounds. In this paper, we propose to use the $(VLF + LF)/HF$ power ratio with VLF the very low frequency component of rHRV periodogram, *i.e.* below 0.04 Hz, LF the low frequency component ranging between 0.04 and 0.15 Hz and HF the high frequency component comprising frequencies between 0.15 to 0.4 Hz. With respect to the regular LF/HF ratio, this new frequency feature has the advantage of taking into account the VLF component that can also reflect the sympathetic activities.

Similarly than Nardelli *et al.* [24], we experimentally observed that it is necessary to normalize our spectral HRV feature by the feature estimated in the neutral state. More details are given in the next section.

IV. RESULTS

A. Accuracy of Remote HRV Measurements

First, we compare HRV calculated using the camera data to those calculated from the contact sensor. Figures 2 and 3 present a representative example for the qualitative comparison between HRV and rHRV in temporal and frequency

¹EO-23121C from Edmund Optics

domain. It can be observed that the result is very good and the curves match perfectly.

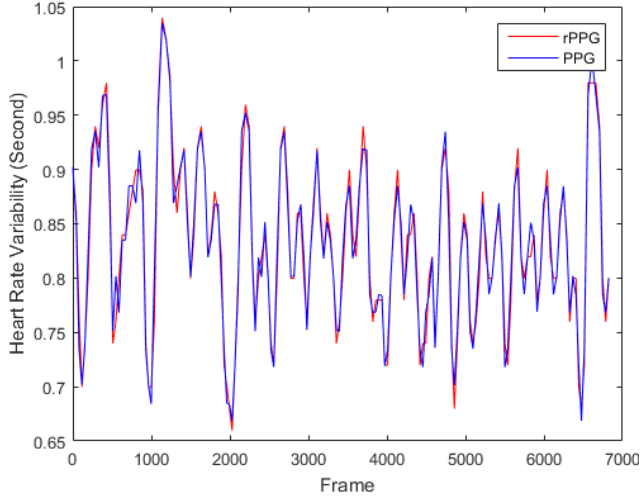


Fig. 2. Comparison between HRV and rHRV temporal signals.

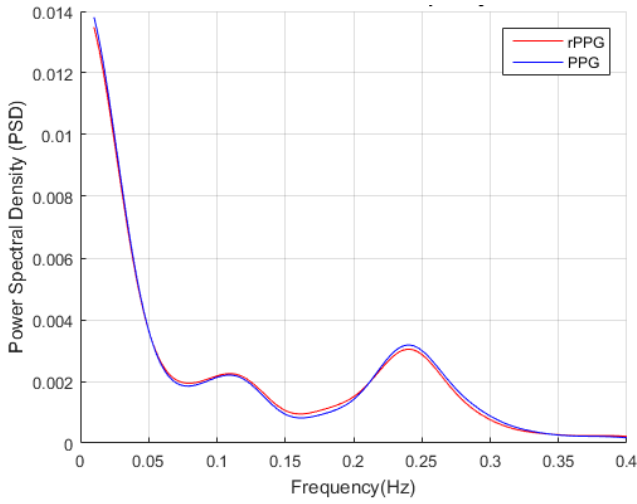


Fig. 3. Comparison between HRV and rHRV periodograms.

Second, we have also quantitatively evaluated the accuracy of rHRV by comparing the integral of energy of VLF , LF and HF components of the HRV and rHRV periodograms with:

$$\begin{aligned} Accuracy(VLF) &= 1 - \frac{|rVLF - VLF|}{VLF} \\ Accuracy(LF) &= 1 - \frac{|rLF - LF|}{LF} \\ Accuracy(HF) &= 1 - \frac{|rHF - HF|}{HF} \end{aligned} \quad (2)$$

where VLF , LF and HF are the integral of the HRV spectrum estimated with the PPG sensor in the corresponding frequency range while $rVLF$, rLF and rHF are obtained using the rHRV periodogram estimated from the video. These metrics have been averaged using the 2 sessions of the 16 videos presented in the previous subsection. We

obtain $Accuracy(VLF) = 98.25\%$, $Accuracy(LF) = 97.59\%$ and $Accuracy(HF) = 96.90\%$. The accuracies estimated in these three spectral ranges are also very satisfactory and show that the camera can accurately capture the HRV signal.

B. Correlation between emotional states and remote HRV

In this paper, we have conducted an experiment to observe the correlation of the $(VLF + LF)/HF$ rHRV feature with two different emotional states, *i.e.* one neutral and one with high arousal elicited by video-acoustic stimuli.

We compare the remotely measured HRV on the two sessions of this experiment. Figure 4 shows $(VLF + LF)/HF$ rHRV feature values in neutral and arousal sessions. Interestingly, 12 participants of the total 16 (75%) have higher values of the feature in the arousal session than in the neutral session. This interesting result validates the methodology and show that the precision of the remote HRV is sufficient to build physiological-based emotional state monitoring using a simple camera.

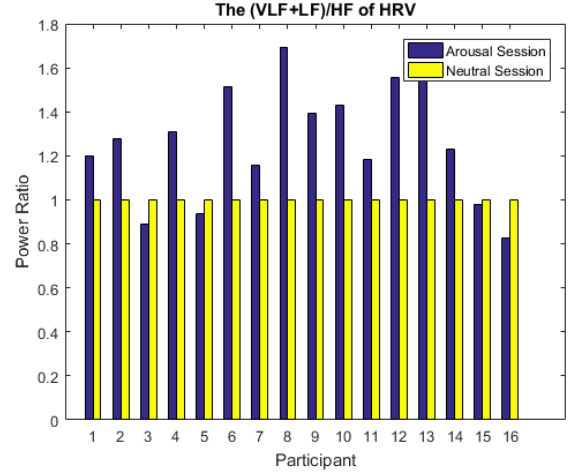


Fig. 4. Normalized $(VLF + LF)/HF$ feature in different emotional states.

In order to remark the importance of normalization procedure in the proposed algorithm, Figure 5 presents non-normalized feature values. It is straightforward to note that a global classifier would fail to recognize the session without normalization. In this case, the obtained classification accuracy would not be significantly different from random guess. Nevertheless, this result remains very interesting and promising because it demonstrates that it is possible to measure changes in the emotional state of a person based on remote physiological data.

Results of 4 participants did not behave as expected. In our opinion, this is not due to the methodology of signal and video processing but rather to the experimental protocol. All the volunteers confirmed that they had been "disturbed" by the arousal-session videos, but the neutral session was probably not satisfactory. Perhaps this was because the participants found the experiment itself stressful. In this case the neutral session may not have represented a truly relaxed state.

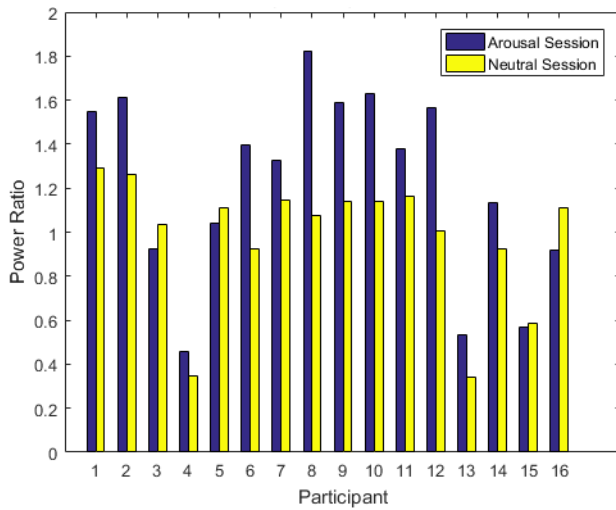


Fig. 5. Non-normalized $(VLF + LF)/HF$ feature in different emotional states.

V. CONCLUSIONS AND FUTURE WORK

Several methods and devices have been proposed for measuring physiological signals and have been used widely in emotion recognition applications. However, physiological sensors usually need to be attached to the human body, which might be intrusive and represents a major hurdle in collecting physiological data in daily life over long term. In this paper, remote HRV frequency-feature is estimated from video data collected with a simple camera. We experimentally show that there is a high agreement between the camera-based HRV and the contact sensor proving that it is possible to accurately capture the HRV signal remotely. Then, using data collected from 16 volunteers and video-induced emotions, we show that there is a strong correlation between the rHRV feature and different emotional states for 12 participants.

Non-contact measurement of physiological signals correlated with emotional states is highly desirable. rHRV is clearly an interesting modality for that matter. However, our work has several limitations that should be addressed in future works. The user-dependence of our feature is very strong and prevent from using an emotion classifier without normalizing the data with a neutral emotional state. Nevertheless, this feature allows to monitor the emotional over time and possibly to detect changes. More work should be conducted on user-independence in order to further open the field of possible applications. The use of multiple features, possibly from other modalities can be investigated to solve this issue. Then, the experimental protocol can also be improved making sure the volunteers are relaxed enough to start the measurements and using more than two emotional states.

REFERENCES

- [1] Tian, Y., Kanade, T. & Cohn, J. F. Facial expression recognition. In *Handbook of face recognition*, 487–519 (Springer, 2011).
- [2] Bartlett, M. S. *et al.* Automatic recognition of facial actions in spontaneous expressions. *Journal of multimedia* **1**, 22–35 (2006).

- [3] Wang, J., Yin, L., Wei, X. & Sun, Y. 3d facial expression recognition based on primitive surface feature distribution. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 2, 1399–1406 (IEEE, 2006).
- [4] Liu, P. & Yin, L. Spontaneous facial expression analysis based on temperature changes and head motions. In *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, vol. 1, 1–6 (IEEE, 2015).
- [5] Nakasone, A., Prendinger, H. & Ishizuka, M. Emotion recognition from electromyography and skin conductance. In *Proc. of the 5th International Workshop on Biosignal Interpretation*, 219–222 (2005).
- [6] Kim, J. & André, E. Emotion recognition based on physiological changes in music listening. *IEEE transactions on pattern analysis and machine intelligence* **30**, 2067–2083 (2008).
- [7] Picard, R. W., Vyzas, E. & Healey, J. Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE transactions on pattern analysis and machine intelligence* **23**, 1175–1191 (2001).
- [8] Vrijkotte, T. G., Van Doornen, L. J. & De Geus, E. J. Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability. *Hypertension* **35**, 880–886 (2000).
- [9] Appelhans, B. M. & Luecken, L. J. Heart rate variability as an index of regulated emotional responding. *Review of general psychology* **10**, 229 (2006).
- [10] Sztajzel, J. *et al.* Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system. *Swiss medical weekly* **134**, 514–522 (2004).
- [11] McCraty, R. Science of the heart: Exploring the role of the heart in human performance [null]. *Boulder Creek, CA: HeartMath Research Center* (2001).
- [12] Lu, G., Yang, F., Taylor, J. & Stein, J. A comparison of photoplethysmography and ecg recording to analyse heart rate variability in healthy subjects. *Journal of medical engineering & technology* **33**, 634–641 (2009).
- [13] Sun, Y. & Thakor, N. Photoplethysmography revisited: from contact to noncontact, from point to imaging. *IEEE Transactions on Biomedical Engineering* **63**, 463–477 (2016).
- [14] Verkruysse, W., Svaasand, L. O. & Nelson, J. S. Remote plethysmographic imaging using ambient light. *Optics express* **16**, 21434–21445 (2008).
- [15] Bobbia, S., Macwan, R., Benezeth, Y., Mansouri, A. & Dubois, J. Un-supervised skin tissue segmentation for remote photoplethysmography. *Pattern Recognition Letters* (2017).
- [16] Macwan, R., Benezeth, Y., Mansouri, A., Nakamura, K. & Gomez, R. Remote photoplethysmography measurement using constrained ica. *IEEE int. conf. on E-Health and Bioengineering* (2017).
- [17] Kranjec, J., Beguš, S., Geršak, G. & Drnovšek, J. Non-contact heart rate and heart rate variability measurements: A review. *Biomedical Signal Processing and Control* (2014).
- [18] McDuff, D. J., Estep, J. R., Piasecki, A. M. & Blackford, E. B. A Survey of Remote Optical Photoplethysmographic Imaging Methods. *int. conf. of the IEEE Engineering in Medicine and Biology Society* (2015).
- [19] Viola, P. & Jones, M. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1, 1–I (IEEE, 2001).
- [20] Lucas, B. D., Kanade, T. *et al.* An iterative image registration technique with an application to stereo vision (1981).
- [21] Conaire, C. O., O'Connor, N. E. & Smeaton, A. F. Detector adaptation by maximising agreement between independent data sources. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, 1–6 (IEEE, 2007).
- [22] Tarvainen, M. P., Ranta-aho, P. O. & Karjalainen, P. A. An advanced detrending method with application to HRV analysis Mika P. Tarvainen, Perttu O. Ranta-aho, and Pasi A. Karjalainen. *IEEE Trans. on Biomedical Engineering* (2002).
- [23] de Haan, G. & Jeanne, V. Robust pulse rate from chrominance-based rppg. *IEEE Transactions on Biomedical Engineering* **60**, 2878–2886 (2013).
- [24] Nardelli, M., Valenza, G., Greco, A., Lanata, A. & Scilingo, E. P. Recognizing emotions induced by affective sounds through heart rate variability. *IEEE Transactions on Affective Computing* **6**, 385–394 (2015).