Noncontact heart rate measurement using high sensitivity camera in low light environment

Genki Okada[1], Keiichiro Kagawa[2], Shoji Kawahito[2], Norimichi Tsumura[1]
1) Graduate School of Advanced Integration Science, Chiba University, CHIBA, JAPAN
2) Research Institute of Electronics, Shizuoka University, HAMAMATSU, JAPAN

Abstract

In this paper, we propose the method of remote estimation for heart rate (HR) and heart rate variability spectrogram (HRVS) by analyzing hemoglobin concentration from facial RGB images in low light environment. The emotion monitoring has a great potential in areas such as market research, safety measure, medical, and robot systems and so on. Various methods were proposed for the remote emotion estimation. Especially, the methods to estimate emotion by analyzing physiological signals from RGB video are expected to be used practically. However, these studies could not be applied in dark places where monitoring is necessary such as infant's bedrooms, crime prone roads and car drivers. Therefore, the proposed method uses the high sensitivity RGB camera capable of capturing videos at low illuminance. As the result, we could measure HR with over 99% accuracy and estimate HRVS with high accuracy in low light environments of 10 lux that is the same brightness as the twilight where this low light environment is required in monitoring infant's bedrooms, crime prone roads, car drivers.

1. Introduction

The emotion monitoring has a great potential in areas such as market research, safety measure, medical, and robot systems. Among such applications, prevention of potential accidents or crime is important for safe social system. For example, monitoring with a surveillance camera can detect people who are trying to commit a crime or who are excessively excited. Furthermore, accidents due to dangerous driving can be prevented by monitoring the driver of the car and measuring drowsiness and concentration. This technique for emotion monitoring has been studied for a long time. Many researchers have attempted to realize emotion recognition using such as facial expressions, voices and physiological signals. Especially, physiological signals attract the most attention in recent years for emotion recognition. In the study of physiological psychology, it is known that there is strong correlation between the physiological response by the action of the autonomic nervous system and human emotional state. Furthermore, physiological signals are less affected by the social and cultural differences [1]. We can estimate the original emotions that the people were trying to hide or that could not be recognized even in themselves.

Park *et al.* [2] measured physiological signals such as skin temperature, electrodermal activity, photoplethysmography and electrocardiogram of 12 healthy participants before and after they watched the movies that are elicit seven emotions (happiness, sadness, anger, fear, disgust, surprise, and stress) by the electrodes. As a result, they classified seven emotions with around 90% accuracy by selecting useful features for emotion recognition by means of particle swarm optimization to features that were obtained by analyzing the measured physiological signals. In this way, it is possible to classify emotions using the physiological signals. However, this study is not practical because they used special measuring devices such as contact type devices. Moreover, the contact type devices might be uncomfortable for the participants because external factors such as an electrode make participants feel the stress by giving a burden.

Kurita *et al.* [3] and Okada *et al.* [4] realized remote heart rate variability (HRV) measurement system using RGB camera by analyzing hemoglobin concentration from facial color images. They identified if participants were relaxed or stressed or felt 5 emotions by performing frequency analysis on the heart rate variability. In this study, it becomes possible to detect stress without causing unnecessary discomfort to the participants. However, they could not be applied to low brightness images taken in the dark such as inside the car and at night because they were using ordinary camera.

Zhao *et al.* [5] measured heart and respiration rates by applying delay-coordinate transformation and independent component analysis. Since they used the camera that is sensitive to visible and near infrared light, this method can be applied for both day and night conditions. However, their approach required a near infrared light emitting diode.

In this paper, we propose the method to measure a pulse wave by ultra-high sensitivity camera in low illumination environment, which does not require special light sources or contact devices. Our method deals with the hemoglobin concentration obtained by analyzing the facial images captured with the ultra-high sensitive RGB camera capable of captured with high inter-scene dynamic range even at low illuminance. In Section 2, we describe the non-contact pulse wave measurement technique using the hemoglobin pigment separation of the facial image in the prior study [3][4]. In Section 3, we perform the experiments to measure the pulse wave in various illumination environments. In Section 4, the results of pulse wave estimation in each illuminance environment are described. In Section 5, we discuss our results. Finally, in Section 6, we describe the conclusion and future works.

2. Method of Remote Measurement for Pulse Wave

Various methods of pulse wave measurement were proposed using a camera. Kurita *et al.* [3] and Okada *et al.* [4] measured the pulse wave in a non-contact manner by detecting the hemoglobin concentration from the facial image and acquiring the temporal change. This method is based on biological optics and has high credibility. Therefore, in this paper, independent component analysis (ICA) was applied to the RGB pixel values of the facial image to perform skin pigment separation and to extract hemoglobin component images. We treat the change of the average pixel value of the hemoglobin component images as pulse wave.

Figure 1 shows the model of human skin. Human skin is multilayer structure that can be roughly divided into the epidermis, dermis and subcutaneous tissue. In practice, the boundary surface of each layer has an irregular shape. However, we treat the boundary surface as plane shape for simplicity. Human skin contains melanin and hemoglobin pigments. Color tone of human skin is greatly affected by these pigments. Melanin pigments exist in epidermis and hemoglobin pigments exist in dermis. Therefore, melanin and hemoglobin pigments can be regarded as being present in the spatially independent by assuming that epidermis is melanin layer and dermis is hemoglobin layer. Light incident on the human skin is divided into surface reflection light and internal reflection light emitted to the outside of the skin after repeated absorption and scattered inside the skin. While surface reflection light represents the color of the light source such as gloss, internal reflection light represents the color of the skin. In this paper, we take the images without surface reflection light by using polarizing plates placed in front of the camera and the light source orthogonal to each other [6]. When the Modified Lambert-Beer law is assumed to be established with respect to the observation signal that is the reflecting light, the observation signal can be represented by the following equation by logarithmic conversion from image space to density space.

$$\begin{aligned}\boldsymbol{v}^{\log}(x,y) &= -\rho_m(x,y)\boldsymbol{\sigma}_m - \rho_h(x,y)\boldsymbol{\sigma}_h \\ &\quad + p^{\log}(x,y)\boldsymbol{1} + \boldsymbol{e}^{\log}\end{aligned} \qquad (1)$$

where, $\boldsymbol{v}^{\log}$ is the converted observation signal, (x, y) is pixel location, ρ_m and ρ_h is the concentration of melanin and hemoglobin pigment respectively, $\boldsymbol{\sigma}_m$ and $\boldsymbol{\sigma}_h$ is the absorption cross section of melanin and hemoglobin pigment respectively, $p^{\log}$ is the parameter for shading due to the shape of the skin, $\boldsymbol{1}$ is the vector of the strength of the shading and $\boldsymbol{e}^{\log}$ is the bias vector. Hence, we can regard melanin and hemoglobin pigments as independent signals as Figure 2 shows. Therefore, it is possible to obtain the melanin and hemoglobin pigment concentration distribution from RGB values of the facial images.

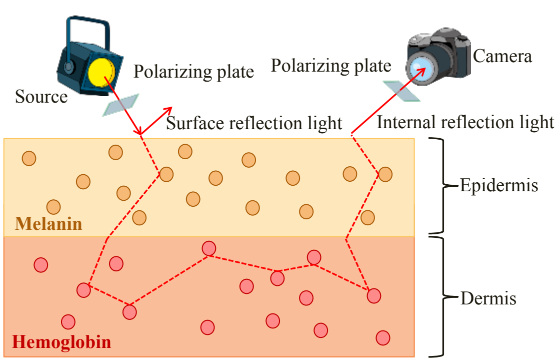


Figure 1 Movement of the light incident on the skin model

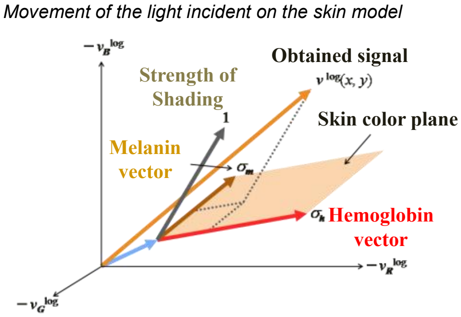
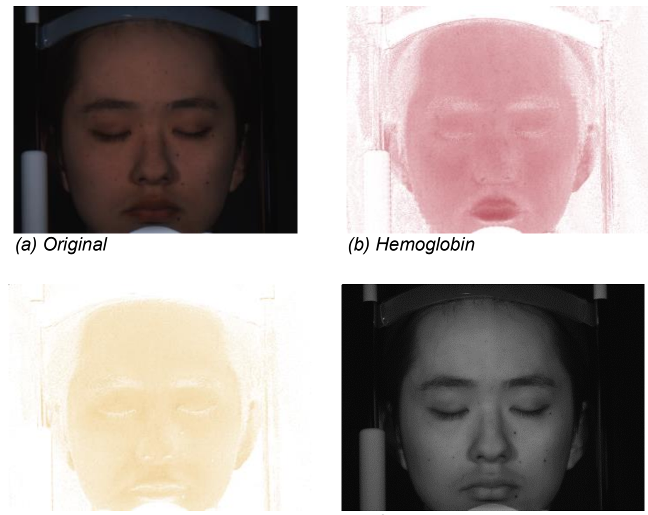


Figure 2 The obtained signals and the three independent signals

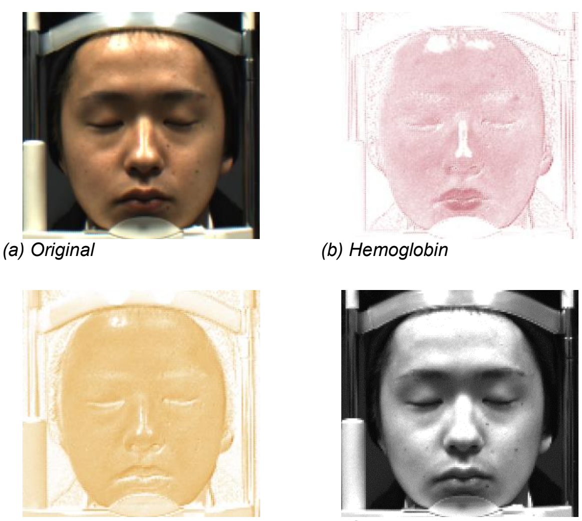
Figure 3 (b), (c) and (d) are the extracted melanin and hemoglobin pigments and shading in whole facial image shown in Figure 3 (a) by independent component analysis. These images are obtained without surface reflection light using polarizing plates. Figure 4 (a) is the facial image taken under fluorescent lights. In the case that facial image contains the surface reflection light, we can also apply skin pigment separation as shown in Figure 4 (b), (c) and (d) using each pigment component color vector estimated from the internal reflection image shown in Figure 3 (a).



(a) Original *(b) Hemoglobin*

(c) Melanin *(d) Shading*

Figure 3 The result of skin pigment separation for internal reflection image using polarizing plates; (a) Original, (b) Hemoglobin, (c) Melanin, (d) Shading



(a) Original *(b) Hemoglobin*

(c) Melanin *(d) Shading*

Figure 4 The result of skin pigment separation for fluorescent lamps image; (a) Original, (b) Hemoglobin, (c) Melanin, (d) Shading

The change of the average pixel values of the hemoglobin component images represents the signal of the blood volume change. However, temporal changes in pixel values acquired from the entire image also include changes due to blinks of eyes and body movements. Therefore, a region excluding the eyes from the face region detected by the Viola Jones method, the forehead (the half width and 25% height from the top of the face region) and the around mouth region (the half width and the half height from the bottom of the face region), was set as the region of interest (ROI).

Figure 5 shows the sets of ROI in the hemoglobin component image and Figure 6 shows the change along the time of the average pixel values in the ROIs in the hemoglobin component images. The peaks of the signal of the blood volume change correspond to the peaks of electro cardiogram waveform called R wave. The intervals between R waves are called RR intervals that are important for heart rate analysis. In order to make it easy to the peak detection, the signal was performed detrending [7]. Subsequently, the detrended signal was multi-band-pass filtered with a Hamming window to reduce noise in the original wave. The multi-band-pass filter was adjusted for each signal by frequency-converting the detrended signal and setting the width of the peak closest to 1 Hz and its second harmonic to cut-off frequency. 1 Hz corresponds to 60 beats-per-minute (bpm), which is a general heart rate in a normal state. The filtered signal was interpolated with three-dimensional spline function at 50Hz to match the sampling frequency of the electrocardiogram measured as the correct value. The RR intervals were calculated by the peak detection with respect to the filtered signal. Figure 6, 7, 8 shows the detrended, filtered signal and RRintervals.

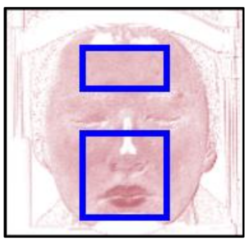


Figure 5 The area of the set ROI

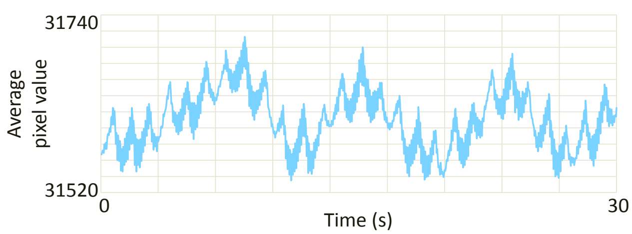


Figure 6 Average pixel values of hemoglobin component images

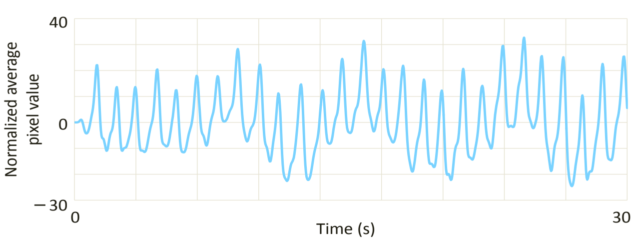


Figure 7 Normalized, detrended and filtered signal

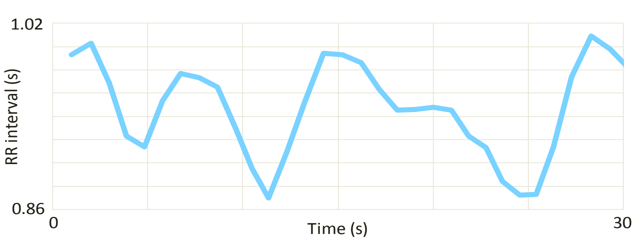


Figure 8 RR intervals

The influence of noises by the camera due to low illuminance increase, and the accuracy of heart rate analysis is increased. In this paper, to remove these noises, the RR intervals were filtered by non-causal of variable threshold algorithm [8].

We calculated heart rate (HR) and HRVS for a comparison between the electrocardiogram and the camera method. HR was calculated as 60/RR, where RR represents the mean of the RR intervals. Heart rate variability (HRV) which is the variations between consecutive heartbeats is modulated by both the sympathetic and parasympathetic branches of the autonomic nervous system. Since the most conspicuous periodic component of HRV is considered to range from 0.15 to 0.4 Hz. In addition to the physiological influence of breathing on HRV, high frequency (HF) component in this range is generally believed to be of parasympathetic origin. Another widely studied component of HRV is the low frequency (LF) component ranging from 0.04 to 0.15 Hz that have been thought to be of both sympathetic and parasympathetic origin. The components of HRV have been found to correlate with such as age, mental and physical stress, and attention [9]. Therefore, highly accurate estimation of HRV spectrum is important. We created heart rate variability spectrogram by calculating the power spectral density (PSD) of the RR intervals for each moving window. The PSD was calculated by the Lomb periodogram. Conventional spectral analysis techniques such as Welch's method require that the input signal be uniformly sampled. If the sampling is not uniform like the RR interval, the signal needs to be resampled or interpolated to a uniform sample rate. However, such processing can add undesirable artifacts to the spectrum, which lead to parsing errors. Since the Lomb periodogram directly processes non-uniform samples, resampling and interpolation are unnecessary and it is very useful for spectral analysis of RR intervals. We used the window of 1 minute and the step size was 1 second. Figure 9 shows an example of HRVS.

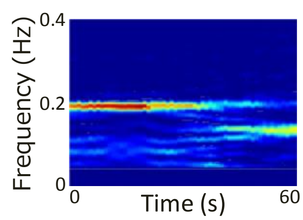


Figure 9 Heart rate variability spectrogram

3. Experiment under various intensities of illuminance

In this section, we discuss the experiments for heart rate and heart rate variability measurements by analyzing the hemoglobin component images from the facial images. The experiments were conducted in the dark room as shown in Figure 10. The 18bit camera [Xviii : ViewPLUS] and the dimming light source were placed away 1 meter from the participants.

We used ultra high sensitivity camera with high inter-scene dynamic range even if the exposure conditions are hardly changed that can capture simultaneously subjects with illuminance of 0.01 lux and 400 lux or more at 30 fps. 0.01 lux is the illuminance of a half moon light or the starlight. 400 lux is the illuminance under fluorescent light. Figure 11 shows images obtained by capturing the participant with illuminance of 10 lux using the 18bit camera. These images were obtained by extracting arbitrary 8 bit data from 18 bit data for visualization. Figure 11 (a) is an image of high bit data of 11 to 18, which is the similar output result as an image obtained by capturing the same scene with a general camera. Figure 11 (b) is an image of medium-bit data of 7 to 14, and it is possible to confirm the participant's face. Figure 11 (c) shows an image of low-bit data from 1 to 8, saturation occurs because the illuminance of the subject is high with respect to the sensitivity of the low bit.

In this study, pulse wave was acquired by the aforementioned method using the values obtained by demosaicing all the data of 1 to 18 bits. The dimming light source can be continuously adjusted the brightness, and we adjusted the brightness manually in accordance with the illuminance of participant's face. The faces were fixed using the chin rest in this paper.

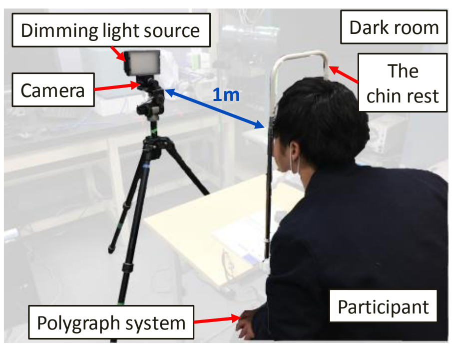
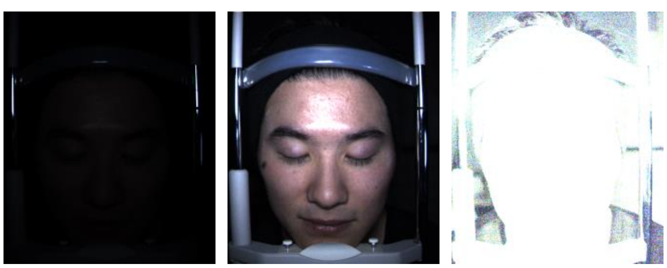


Figure 10 The experiment environment



(a) 11-18bit *(b) 7-14bit* *(c) 1-8bit*

Figure 11 The obtained images from the 18bit camera in 10 lux; (a) 11-18bit, (b) 7-14bit, (c) 1-8bit

Prior to the experiment, the participants were introduced the procedure of the experiments and they had an adaptation time to feel comfortable in the experiment system. We measured with an illuminometer, the brightness of the light source was adjusted so that the illuminance of the faces would be 5, 10, 50, 100, 200, 300 lux. The participants' faces were captured for two minutes at each illuminance.

We measured correct value of heart rate from the electrocardiogram using by the polygraph system [RMT-1000 : NIHON KOHDEN. Inc]. We set the measurement resolution at 50Hz, and applied low-pass filter. Cut-off frequency was set at 15Hz. It is enough to get the peaks of R waves. From these signals, we obtained the RR intervals and calculated heart rate as ground truth data by the method used to hemoglobin component signal.

4. Results

Figure 12 and Table 1 show the accuracy of HR at each illuminance. The accuracy was calculated as shown in the following equation.

$$\text{The accuracy} = 100 - \frac{|E-G|}{G} \times 100 \quad (2)$$

where, E is the estimated HR and G is the ground truth data obtained from the electrocardiogram. In the illuminance environment of 300, 200, 100, 50, 10, 5 lux, the estimation accuracy of the heart rate by the 18bit camera were 99.81, 99.79, 99.80, 99.80 99.49, and 89.15, respectively. By contrast, the accuracy with an 8bit camera were 99.73, 99.78, 99.84, 99.92, and 42.96. Figure 13 shows the noises on the captured images at low illuminance, 10 and 5 lux, respectively.

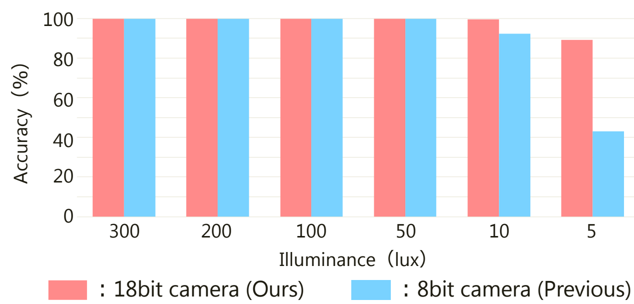
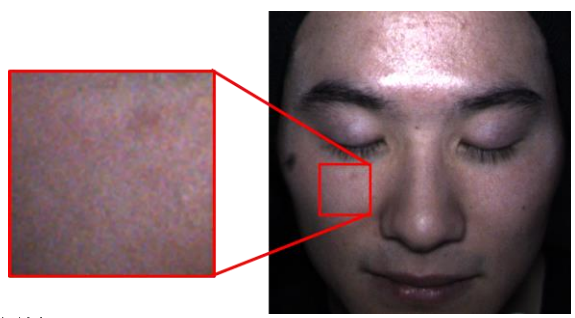


Figure 12 The accuracy of HR at each illuminance

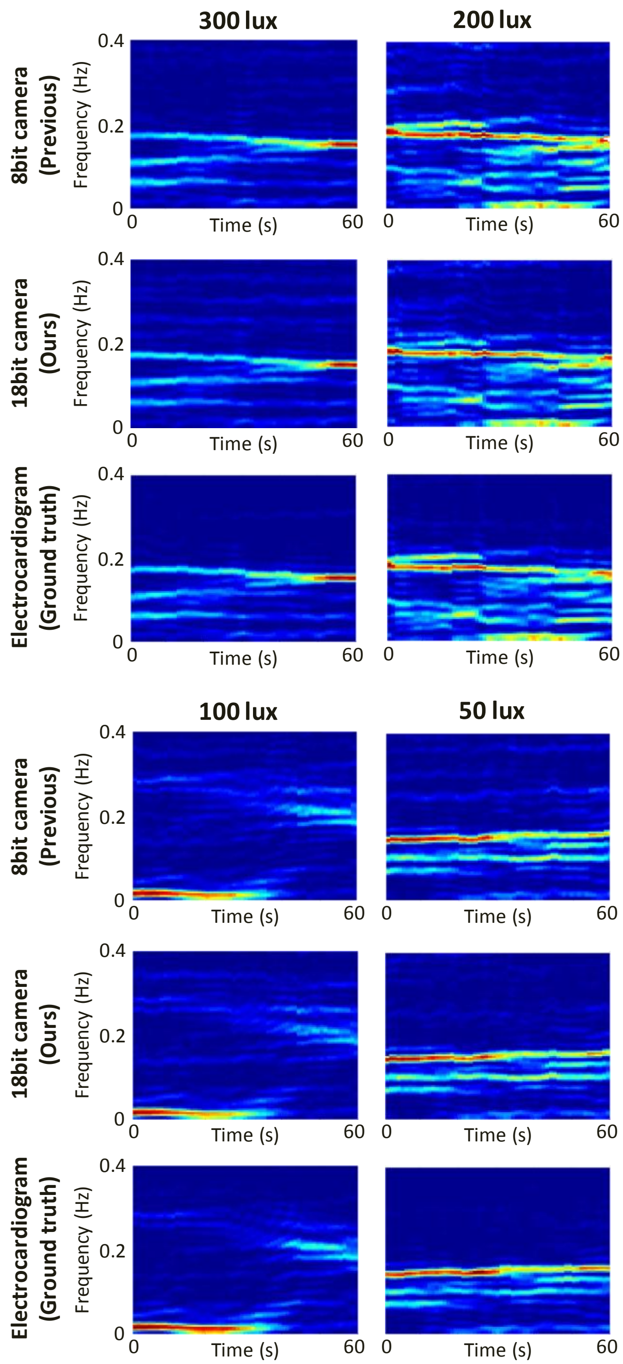


(a) 10 lux

(b) 5 lux

Figure 13 Noises on the captured images at low illuminance

Figure 14 shows the comparison of the HRVS obtained from the previous 8bit camera, the ultra high sensitivity camera and the electrocardiogram at each illuminance. When the illumination environment was between 300 and 50 lux, the results obtained by the cameras were almost the same with the calculated from electrocardiogram. In the case of 10 lux, the noise in estimated result of the 18bit camera increased slightly, but the frequency with the highest power spectral density was similar to the ground truth. On the other hand, the HRVS calculated using the 8bit camera differed from the ground truth. The estimated HRVS by the cameras was different from the created by the electrocardiogram.



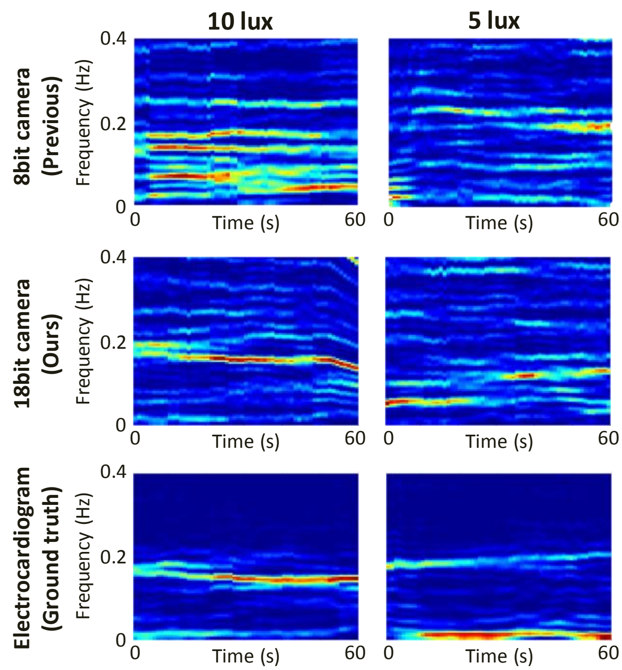


Figure 14 The comparison of the HRVS

5. Discussion

Illuminance of 300, 200, 100, 50, 10 and 5 lux are brightness of general office under fluorescent lamp, family living room, night entrance, under street light, twilight, and residential street, respectively. The heart rate was estimated with nearly 100% accuracy used the 18bit camera in lower illuminance environments. The results show that our method can be applied to general household environments and roads. Heart rate measurement in such a dark place is useful for sudden infant death syndrome, driver monitoring and crime deterrence. The estimation at 5 lux was lower accuracy than the estimated in the brighter illumination environments. This low accuracy seems to be caused by noise due to the camera was dominant as shown in Figure 13. However, this result shows that the accuracy was improved considerably compared with the 8bit camera.

Heart rate variability is modulated by the autonomic nervous system. Spectral analysis of heart rate variability is important since mental state such as stress and drowsiness can be estimated from the function of the autonomic nervous system. The calculated heart rate variance spectrograms were almost perfectly reproduced at 50 lux and more. Even at 10 lux, although the noise increased slightly, the strongest spectra were almost the same as the ground truth. The ratio of HF and LF or the strongest spectrum is widely used for stress detection. Therefore, the result seems to be sufficient accuracy to detect stress. The estimated value when the face illuminance was 5 lux was different from the correct value. This result seems to be affected by the predominant noise that was not removed even by the processing such as spatial averaging, multi-band-pass filter and NC-VT. In order to cope effectively such noise, additional processing should be considered to be required before extracting hemoglobin.

6. Conclusion and Future works

We measured the pulse wave from the hemoglobin concentration obtained by analyzing the RGB image of the face taken with the high sensitivity camera in the low illuminance environment. Heart rate and heart rate variabil ity spectrogram useful for emotion estimation were calculated from the measured pulse wave. The estimation accuracy of heart rate and heart rate variability spectrogram were almost 100% under the illumination more than 10lux.

Our future works are improving of the accuracy at low illuminance of 5 lux or less by eliminating random noise due to the camera and the application to emotion monitoring for drivers. Emotion can be estimated regardless of lighting conditions from low to high illuminance by using the ultra-high sensitivity camera.

Acknowledgements

This work was supported in part by the MEXT/JST COI STREAM program.

References

[1] O. Alaoui-Ismaili, O. Robin, H. Rada, A. Dittmar and E. Vernet-Maury, "Basic emotions evoked by odorants: comparison between autonomic responses and self-evaluation," Physiology and Behavior, vol. 62, pp. 713-720, 1997.

[2] B.-J. Park, E.-H. Jang, S.-H. Kim, C. Huh, and J.-H. Sohn, "Seven emotion recognition by means of particle swarm optimization on physiological signals: Seven emotion recognition," in Proc. 9th IEEE ICNSC, Apr. 2012, pp. 277–282.

[3] Kurita K, Yonezawa T, Kuroshima M and Tsumura N, "Non-Contact Video Based Estimation for Heart Rate Variability Spectrogram using Ambient Light by Extracting Hemoglobin Information," Color and Imaging Conference, Volume 2015, Number 1, October 2015, pp. 207-211

[4] Okada G, Yonezawa T, Kurita K and Tsumura N, " Emotion Monitoring Using Remote Measurement for Physiological Signals by Camera" Color and Imaging Conference, Volume 2016, Number 1, November 2016, pp. 147-151

[5] Zhao F, Li M, Qian Y, Tsien JZ (2013) Remote Measurements of Heart and Respiration Rates for Telemedicine. PLoS ONE 8(10): e71384. doi:10.1371/journal.pone.0071384

[6] Ojima, N. and Minami, T. and Kawai, M. Transmittance measurement of cosmetic layer applied on skin by using processing. : Proceeding of The 3rd Scientific Conference of the Asian Societies of Cosmetic Scientists, vol.114, 1997.

[7] M. P. Tarvainen, P. O. Ranta-aho, and P. A. Karjalainen, "An advanced detrending method with application to hrv analysis," Biomedical Engineering, IEEE Transactions on, vol. 49, no. 2, pp. 172–175, 2002.

[8] J. Vila, F. Palacios, J. Presedo, M. Fernández-Delgado, P. Felix, and S. Barro, "Time-frequency analysis of heart-rate variability," IEEE Eng. Med. Biol. Mag., vol. 16, no. 5, pp. 119–126, Sep./Oct. 1997.

[9] G.G. Berntson, J.T. Bigger Jr., D.L. Eckberg, P. Grossman, P.G. Kaufmann, M. Malik, H.N. Nagaraja, S.W. Porges, J.P. Saul, P.H. Stone, and M.W. Van Der Molen. Heart rate variability: Origins, methods, and interpretive caveats. Psychophysiol, 34:623-648, 1997.

