Human Detection Using Biological Signals in Camera Images with Privacy Aware

Toshihiro Kitajima[1,2(✉)], Edwardo Arata Y. Murakami[1], Shunsuke Yoshimoto[2], Yoshihiro Kuroda[2], and Osamu Oshiro[2]

[1] Samsung R&D Institute Japan, ET Center ET-2 Lab, Minoh Semba Center Bldg. 2-1-11 Semba Nishi, Minoh City, Osaka, Japan
{t.kitajima,arata}@samsung.com
[2] Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka, Japan
{yoshimoto,ykuroda,oshiro}@bpe.es.osaka-u.ac.jp,
http://www.samsung-srj.co.jp
http://oshiro.bpe.es.osaka-u.ac.jp

Abstract. In the Internet of Things (IoT), in which people sensing occurs everywhere, cameras are being increasingly used because they provide inexpensive and effective sensing devices. However, while cameras can provide significant amounts of information, much of that information is personal and there are significant concerns that individual privacy could be compromised. Furthermore, since home appliances are increasingly being connected to the Internet via the IoT, it has become possible for user images to leak out unintentionally. With these concerns in mind, we propose a human detection method that protects user privacy by using intentionally blurred images. In this method, the presence of a human being is determined by dividing an image into several regions and then calculating the heart rate detected in each region. In our performance evaluation, the proposed method showed dominant performance results when compared with an OpenCV-based face detection method, and was confirmed to be an effective method for detecting human beings in both normal and blurred images.

Keywords: Human detection · Heart rate · Privacy · Camera images · IoT

1 Introduction

In the current Internet of Things (IoT) era, vehicles such as trains, airplanes, and automobiles, along with appliances such as televisions and refrigerators, are increasingly being connected to the Internet, which they automatically send to and receive from a wide variety of data. Additionally, various sensors are being attached to consumer electronics and the data they collect are increasingly being sent from and received by products in order to improve the services they provide.

A.M. Madureira et al. (eds.), *Intelligent Systems Design and Applications*, Advances in Intelligent Systems and Computing 557,
DOI 10.1007/978-3-319-53480-0_18

Since users can be expected to select home appliance services that support their conditions, it would be helpful if those services could sense the state of the humans present in their residences.

Meanwhile, as the cost of cameras has declined drastically in recent years, their usage as human sensing devices has increased. However, even though the amount of information that can be gleaned from camera imagery is quite large, which makes cameras very effective as sensing devices, personal information such as faces, clothes, etc., is also acquired, and there are concerns that personal privacy will be compromised. Furthermore, since home appliances are increasingly being connected to the Internet via the IoT, the possibility exists that user image data may leak out unintentionally (Fig. 1). Since such concerns have already convinced numerous consumers not to purchase Internet-connected or camera-equipped products, the development of a sensor device that safeguards the user's privacy is becoming increasingly important.

There are currently a number of devices, such as pyroelectric [1] and temperature sensors [2], that can perform human detection without the use of camera images. These devices can detect the presence of a human within their sensing range, but they cannot detect his or her direction and position. Additionally, while thermopile array sensors [3] can be used to measure the presence of a human, their resolution is very rough, which makes it difficult for them to determine exact locations. Furthermore, individual thermopile sensors are expensive, and their cost increases significantly when they are used as part of a thermopile array.

With these points in mind, this study targets the development of a method whereby a camera can be used to measure the presence, direction, and position of a human being while simultaneously protecting that individual's privacy. Some methods that aim to protect the privacy of camera image data have already been reported. These include software methods to blur human faces after an image is acquired [4,5], a method for replacing the region of the human face with a computer graphics (CG) generated image [6], and a method that requires a framework authentication key to view the images [7]. However, in all of those methods, there is a possibility that the original camera image could be hacked, so it is believed that privacy would be better protected if the image could be made more secure at the acquirement stage.

There are a number of methods for detecting humans using a camera image. One is the face detection method using Haar features and AdaBoost [8], and another is a human detection method using histogram of oriented gradients (HOG) features [9]. These methods are effective in clear images, but they do not work in images that lack sharpness, because contrast and edges characteristics cannot be extracted. The method used in this study applies a different concept that does not consider image appearance features. Instead, our method performs human detection using biological signals obtained from those images. More specifically, before images are captured with the camera, either a frosted glass is placed in front of the camera lens or the focal point of the lens is shifted until the image is sufficiently blurred to protect the user's privacy. Acquired

images are then divided into several regions and fast Fourier transform (FFT) is used to determine whether something is present in a region that has a value corresponding to the heart rate of a human being. In order to verify the effectiveness of this proposed method, a performance evaluation in which it is compared with an existing human detection method is carried out.

The heart rate-based human detection method is described in Sect. 2. A performance comparison in which our method is compared with an existing human detection method using both blurred and normal images is carried out in Sect. 3. A summary is provided and future works are described in Sect. 4.

Fig. 1. Leakage of user privacy via Internet and camera.

2 Method

Currently used human detection methods draw on significant amounts of appearance information, and thus require clear images. However, when it is necessary to protect the user's privacy, clear images are contraindicated, so we propose a method that does not require their use. Reports on a method of measuring heart rates from camera images [10,11] have been published previously. Since this method detects a face region and measures the heart rate in that area, we determined that it would be possible to use this heart rate measurement method to detect the presence of human beings. Specifically, the whole camera image is divided into regions, after which heart rate measurements are taken for each region. If a measured value that corresponds to a human heart rate is detected in a specific region, it is possible that a human face or skin is present in that region.

2.1 Principle of Heart Rate Measurement

In this subsection, the principles of obtaining heart rate measurements from human skin are explained. When used in the context of human blood circulation, pulse refers to the pulsation of blood arteries, which contain large amounts of oxygenated hemoglobin. Figure 2 [13] shows the absorption spectrum of human biological tissue material. Here, it can be seen that oxygenated

hemoglobin absorbs significant amounts of light in the wavelengths from 500 to 580 nm. Figure 3 [14] shows the corresponding relationship between wavelength and color. In this figure, it can be seen that the wavelengths from 500 to 580 nm correspond to green and yellow. More specifically, yellow is from 560 to 580 nm and green is from 490 to 560 nm. Thus, the green wavelength range is wider than the yellow range, and more green light is absorbed into the oxygenated hemoglobin. Commercially available finger-type pulse wave [15] and wristwatch-type [16] sensors measure heart rates by utilizing the property of the color green that is absorbed by the oxygenated hemoglobin. Figure 4 [17] shows the heart rate measurement principle that uses green light. The apparatus is provided with a green light-emitting diode (LED) and a photo-detector. Human skin is irradiated with green light by the LED, and the photo-detector receives the reflected light. As the volume of the blood vessel changes due to the pulse, the value received by the photo-detector varies as a waveform, as can be seen in the right graph of Fig. 4. The heart rate is detected from this waveform.

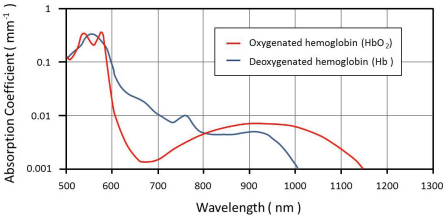

Fig. 2. Absorption spectrum of main substances in living tissue [13].

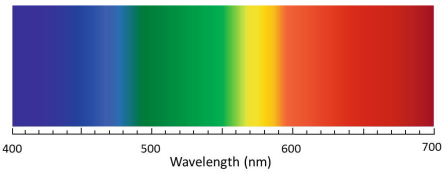

Fig. 3. Wavelength and color of the corresponding relationship [14].

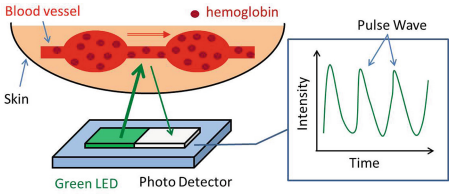

Fig. 4. Pulse wave sensor principle [17].

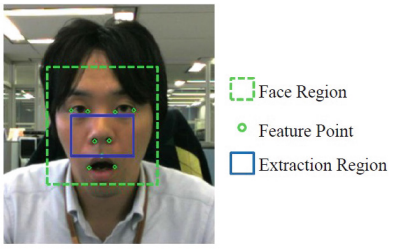

Fig. 5. Face detection and selected extraction region.

When this mechanism is applied to photographs, faces are detected using a commercially available face detection library, as shown in Fig. 5. Eye and mouth movements, as well as the area between the eyes and mouth, are acquired from the face feature points. Since the light in the room and sunlight contains the color green, a red-green-blue (RGB) camera extracts green (G) light as the

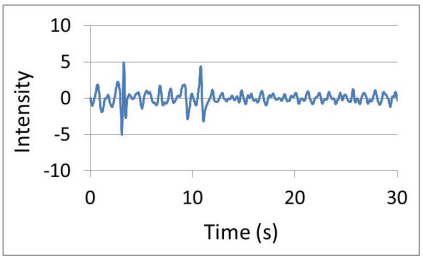

Fig. 6. Heart rate signal extracted from the face region time series data.

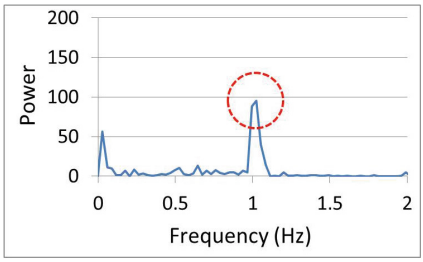

Fig. 7. FFT calculation of heart rate signal showing the detected peak frequency.

reflected wave. Figure 6 shows time series G data in this area. These data include a pulse component, which is equal to the heart rate. Figure 7 shows the result of the data frequency analysis contained in Fig. 6. When the peak surrounded by the red dotted line is detected, it is possible to calculate the human heart rate.

2.2 Human Detection Method

First, before images are captured with the camera, frosted glass is placed in front of the lens or the lens focal point is shifted to reduce image sharpness until human faces cannot be identified. The image in Fig. 8 was obtained by a normal camera, but by shifting the focal point of the lens, the image shown in Fig. 9 was acquired. Next, the entire image is divided into several regions, as shown in Fig. 10. Then, by using the heart rate measuring method discussed in [10], it is possible to calculate the heart rate in each region.

Regions A and B in Fig. 11 show background and human face areas, respectively. Figures 12 and 13 show 30 s of data for the average green intensity values in each region obtained via an RGB camera. Since Fig. 12 shows a background region, there are no substantial changes in green intensity values. In contrast, Fig. 13 shows a human face region, and it can be seen that waveforms indicating a human pulse have been obtained. Next, FFT is performed as input data for these green intensity values. The window function is the rectangular window. The time range window shows the data from current time extending back to the past. Furthermore, the time window moves successively, while overlapping the range processed immediately before. Figures 14 and 15 show the FFT result in each region of the background and human areas. Based on the power spectrum results, it is possible to determine whether data of these regions match the pulse of human being. Since measured values that do not correspond to human heart rates are outside the search range of this method, a heart rate range of 50 to 120 beats per minute (bpm) is set. Strenuous exercise or illness is excluded because the daily life of a human at home is the target of this method. The light gray areas in Figs. 14 and 15 shows areas where the heart rate is less than 50 bpm, and thus outside the scope of the search. Furthermore, when a detected power spectrum is weak, it is unlikely to be a human pulse. Therefore, a threshold value was set, and all areas outside the power spectrum of a human pulse

were excluded. The frequency of the maximum value is selected if multiple power readings exceeding heart rate threshold power are detected. In Figs. 14 and 15, the red dotted line represents the threshold value. Based on this criterion, there is no power value exceeding the threshold in Fig. 14. This indicates that there is no human presence in that area. In Fig. 15, the detected 0.95 Hz frequency power exceeds the threshold. $60 \times 0.95 = 57$, thus indicating a heart rate of 57, so it can be determined that there is a human in this area.

Figure 16 shows the results of this process for all regions. Regions where the power does not exceed the threshold value show a heart rate value of zero. In contrast, the human face region in Fig. 16 shows a heart rate of 57 bpm, so it can be concluded that there is a human in this region. By performing such processing, it is possible to detect the presence of a human being.

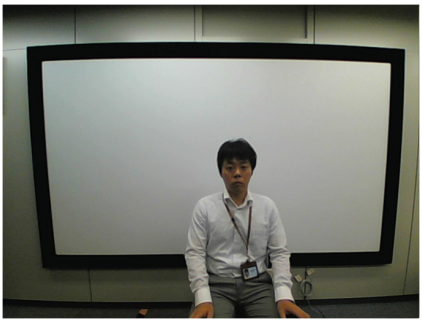

Fig. 8. Normal camera image.

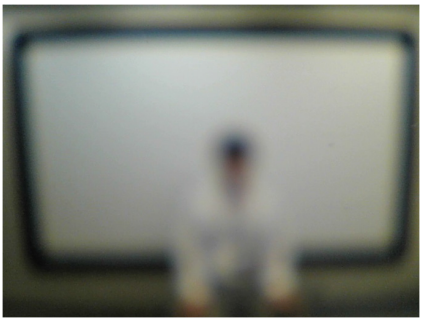

Fig. 9. Out-of-focus image.

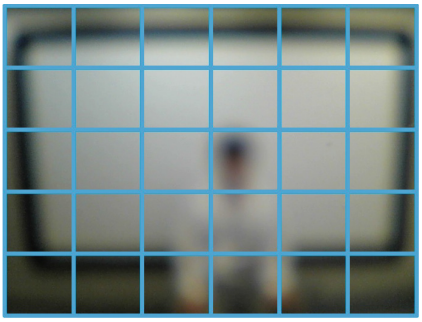

Fig. 10. Divided image to several small regions.

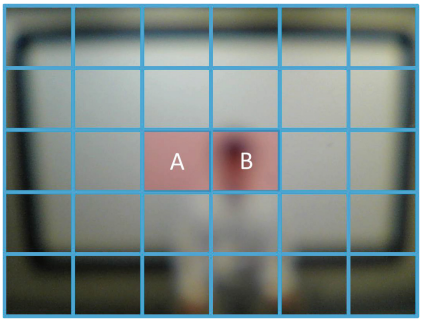

Fig. 11. Background region (A) and human region (B).

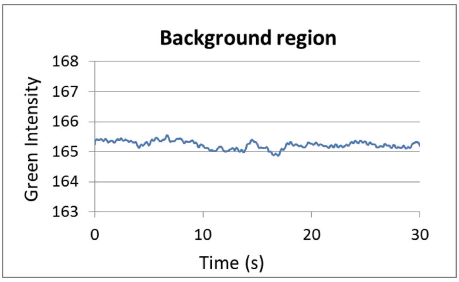

Fig. 12. Green light intensity average of pixels inside the background region (A).

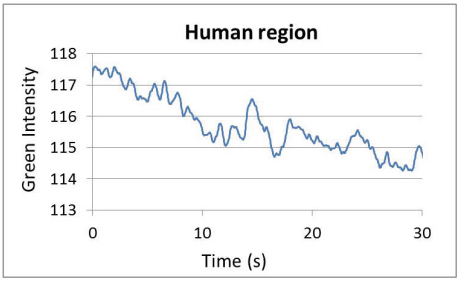

Fig. 13. Green light intensity average of pixels inside the human region (B).

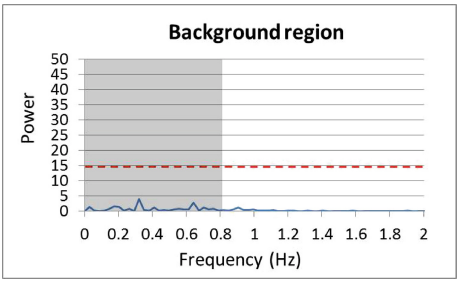

Fig. 14. FFT calculation of the signal in the background region (A). The red dotted line represents threshold values.

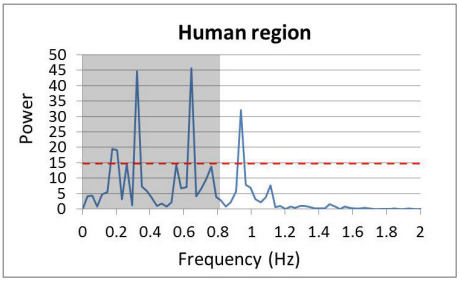

Fig. 15. FFT calculation of the signal in the human region (B). The light gray area is outside the scope of the search.

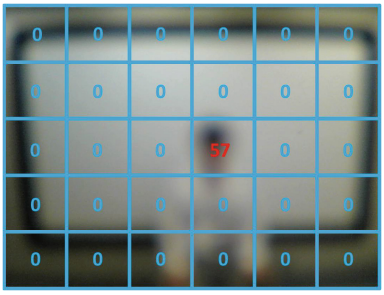

Fig. 16. Heart rate detection results for each region.

3 Performance

3.1 Experiment Environment

To ascertain the utility of our method, performance evaluation of its human detection capability was conducted. In these experiments, a web camera [12] type that is commonly used for video chats, as shown in Fig. 18, was employed. The evaluation using this camera was performed at 30 fps and VGA resolution. The camera specifications are shown in Fig. 17.

Type	RGB
Resolution (pixels)	640 ×480
Frame rate (fps)	30
Gradation (bits)	8 (1 ch)

Fig. 17. RGB camera specifications.

Fig. 18. RGB camera.

Since this report is focused on detecting the heart rate of human beings in stationary states, experiments involving the moving state were not conducted. Subjects sat 1 m from the camera, and stayed in a rest position. The proposed method was compared with an existing face detection method that was implemented in OpenCV Ver. 2.4.3. The minimum search window size of OpenCV is 40 × 40 (pixels), and the scaling factor of the search window was set to 1.2. For our proposed method, the divided region was also 40 × 40 (pixels). The FFT time window of the proposed method was set at 30 s and recording was performed for one minute and 30 s. Since the detection rate of the first 30 s was excluded from the evaluation, the FFT result for the first 30 s was not output. 1800 images remaining one minute were defined as evaluation image. The OpenCV method was also used to evaluate the same 1800 images. Five test subjects participated in this experiment.

The following detection criteria were used. The reference area, which represents the region in which there is human presence, was set manually for all images. If the output area of our human detection method overlaps the reference area, it is determined that human detection is successful. If the areas are not duplicates of each other, a false detection is recorded.

3.2 Experiment Results

While the comparative evaluation of our proposed method and OpenCV was performed using intentionally blurred images, the proposed method is also capable of human detection using normal images. Therefore, an additional evaluation was performed using normal images.

Intentionally Blurred Images. Three types of intentionally blurred images were used. The first were out-of-focus images, as shown in Fig. 9. The second and third were images taken by mounting optical filters in front of the lens.

Fig. 19. Image taken with cyclone filter.

Fig. 20. Image taken with radial filter.

Figures 19 and 20 respectively show photos taken with the cyclone [18] and radial filters [19] used in our experiment.

The receiver operating characteristic (ROC) curve was used in the comparison of the experimental results. The ROC curve shows a representation of detection error rates (false positives) on the horizontal axis and the detection rate (true positives) on the vertical axis. It is possible to create different curves by varying the detector threshold. As the curve is close to the upper left corner, it can be said that the detection performance is good. 27000 images were used as the evaluation images (1800 [images] $\times$ three [filters] $\times$ five [subjects] = 27000 [images]).

One example of the human detection results generated by the proposed method is shown in Fig. 21. The sole detection region is displayed in the green rectangles. The numbers in the rectangle reflect the detected heart rate.

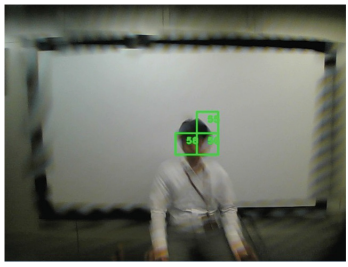

Fig. 21. Human detection result image generated by our proposed method (cyclone filter image). The detection region is displayed in the green squares. The numbers in the squares reflect the detected heart rates.

As shown in Fig. 22, the results from our proposed method are closer to the upper left than those produced by OpenCV. This indicates that the proposed method provides better performance. When the false detection rate was 2.7%, the detection rate was, respectively: proposed method 99.5%, OpenCV 4.1%,

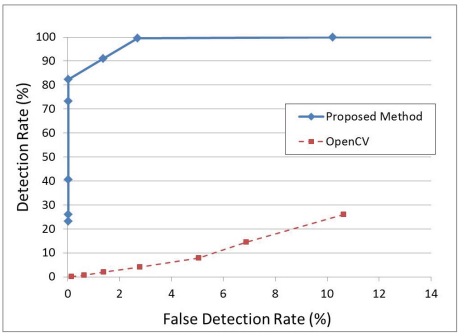

Fig. 22. ROC curves of human detection results with intentionally blurred images. The closer the curve is to the upper left corner, the better the detection performance level is.

which indicates a significant difference. As a whole, OpenCV had low detection rates due to the lack of contrast and edge necessary for facial feature detection. In contrast, the proposed method showed both a high rate of detection and a low rate of false positives. Based on these results, we can state that our proposed method provides an effective human detection method in circumstances where image privacy protection is an important consideration.

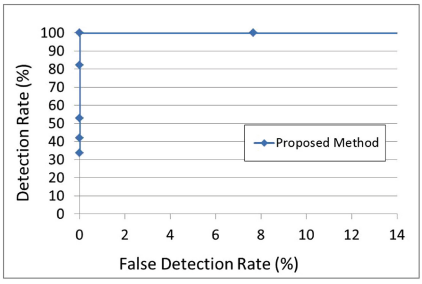

Fig. 23. ROC curve result of the proposed method using normal images.

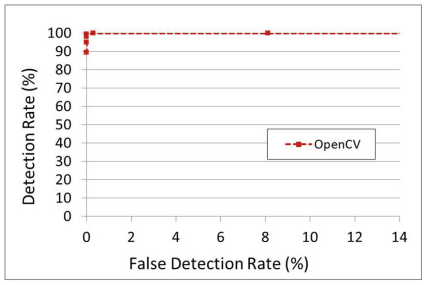

Fig. 24. ROC curve result of OpenCV face detection method using normal images.

Normal Image. Next, a comparison evaluation of the proposed method and OpenCV was performed using normal images, such as that shown in Fig. 8. A total of 9000 images were used in this stage of the experiment (1800 [images] $\times$ five [subjects] = 9000 [images]). The experimental results comparison used the ROC curve in the same manner as discussed above. Overall, the proposed method and OpenCV showed similar results, and it was difficult to comprehend the overlapping curves. Accordingly, the results were shown separately in Figs. 23 and 24 using the ROC curve.

As can be seen in Fig. 24, OpenCV is very close to the upper left, which means it can detect humans almost 100% of the time, with few false positives. However, the proposed method is also very close to the upper left, showing that it too can detect humans almost 100% of the time, with few false positives. From this result, it can be said that the proposed method provides detection performance levels that equal the performance of the existing method, even when normal images are used.

4 Conclusion

Herein, we proposed a novel human detection method that makes use of intentionally blurred images to prevent a subject from being identified, thereby protecting individual privacy. In our method, collected images are divided into regions and the presence or absence of a human being is determined by calculating human heart rates detected in each region. In a performance evaluation using intentionally blurred images, our proposed method showed a higher level of performance than an existing OpenCV-based face detection method. As a result, the proposed method was confirmed to be an effective technique for detecting humans using intentionally blurred images. Additionally, in a separate performance evaluation using sharp normal images, our proposed method showed a level of performance equivalent to the existing OpenCV-based method. This result confirmed that the proposed method can be effectively used with normal images as well. In its current state of development, the proposed method has only considered humans in a rest position. However, since performing detection on moving subjects is often necessary, it will be also be necessary to consider how the inter-frame difference method can be integrated into our proposed method in order to facilitate detection of moving humans. Furthermore, this paper limited detection to a single person. However, if more than one person present, it will eventually be necessary to count people via use of an algorithm for grouping the detection area. Accordingly, it is clear that there is still a need for improvements to our proposed method, and these areas will be the subjects of our future work.

References

1. muRata Infrared Sensors. http://www.murata.com/en-global/products/sensor/infrared. Accessed 2 Aug 2016
2. Thermopile Sensors. http://www.heimannsensor.com/products.php. Accessed 2 Aug 2016
3. Thermopile Array Sensors. http://www.omron.co.jp/ecb/products/sensor/special/humandetectingsensor/. Accessed 2 Aug 2016
4. Dufaux, F., Ebrahimi, T.: Scrambling for video surveillance with privacy. In: Proceedings of IEEE Computer Vision and Pattern Recognition Workshop on Privacy Research in Vision (2006)
5. Yu, X., Babaguchi, N.: Privacy preserving: hiding a face in a face. In: Proceedings of the 8th Asian Conference on Computer Vision, pp. 651–661 (2007)

6. Mitsugami, I., Mukunoki, M., Kawanishi, Y., Hattori, H., Minoh, M.: Privacy-protected camera for the sensing web. In: Information Processing and Management of Uncertainty in Knowledge-Based Systems (2010)
7. Zhang, W., Cheung, S.S., Chen, M.: Hiding privacy information in video surveillance system. In: Proceedings of International Conference on Image Processing (2005)
8. Viola, P., Jones, M.J.: Rapid object detection using a boosted cascade of simple features. In: IEEE CVPR, pp. 511–518 (2001)
9. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE CVPR, pp. 886–893 (2005)
10. Kitajima, T., Choi, S., Murakami, E.A.Y.: Heart rate estimation based on camera image. In: 14th International Conference on Intelligent Systems Design and Applications (2014)
11. Kitajima, T., Choi, S., Murakami, E.A.Y., Yoshimoto, S., Oshiro, O.: Contactless measurement of heart rate using single channel camera images. In: 12th International Conference on Ubiquitous Healthcare (2015)
12. Web Camera Buffalo BSW20KM11BK. http://buffalo.jp/product/multimedia/web-camera/bsw20km11bk/. Accessed 2 Aug 2016
13. Absorption spectrum of main substances in living tissue. http://www.hamamatsu.com/jp/ja/technology/innovation/trs/index.html. Accessed 2 Aug 2016
14. Wavelength, color of the corresponding relationship. http://www.ieij.or.jp/what/yougo.html. Accessed 2 Aug 2016
15. iHealth Wireless Pulse Oximeter. http://www.ihealthlabs.com/health-and-fitness-products-wireless-wireless-pulse-oximeter_80.htm. Accessed 2 Aug 2016
16. Mio ALPHA Heart Rate Sport Watch. http://www.mioglobal.com/product_info.htm. Accessed 2 Aug 2016
17. Pulse wave sensor principle. http://www.rohm.com/web/global/pulse-wave-sensor. Accessed 2 Aug 2016
18. Kenko ZS-CYCLONE. http://www.kenko-tokina.co.jp/discontinued/filters/ef_around/4961607349403.html. Accessed 2 Aug 2016
19. Kenko ZS-RADIAL. http://www.kenko-tokina.co.jp/discontinued/filters/ef_around/4961607349397.html. Accessed 2 Aug 2016