

# Milestone 2: Pulse Rate and Respiratory Rate Extraction

Robert Amelard and Alexander Wong

November 15, 2017

The second milestone from the CHI Project Plan document was “development of algorithm to track pulse rate and respiratory rate via NIR-sensitive camera:”. This report presents the algorithm and experimental results for extracting pulse rate and respiratory rate from a NIR video.

## 1 Experimental Setup

To simulate a hospital bed setting, the imaging system was mounted on the ceiling, with the camera facing directly downward. The distance from the camera to the bed was approximately 2 m. The camera and LED optical paths were designed to image the face and torso region, which are the relevant anatomical locations for extracting pulse and respiration events. 940 nm light was used to illuminate the participant, and the video was recorded at 60 fps based on the results from Milestone 1.

## 2 Algorithm

### 2.1 Face detection

The popular Viola-Jones facial detection algorithm [?] was used to detect the face region of interest (ROI). This method is a fast method used in many applications for detecting frontal faces based on a pre-trained classification and regression tree model. This was implemented as a cascade object detector, where the image is successively downsampled so that the search window’s relative scale is increased, making it robust to changes in scale. Figure 1 demonstrates a graphical example of this process. One particularly useful characteristic about this method is the ability to choose a minimum object window size, which can be determined based on the optical setup. In our case, a  $75 \times 75$  window was empirically determined to be a valid minimum window size for locating faces.

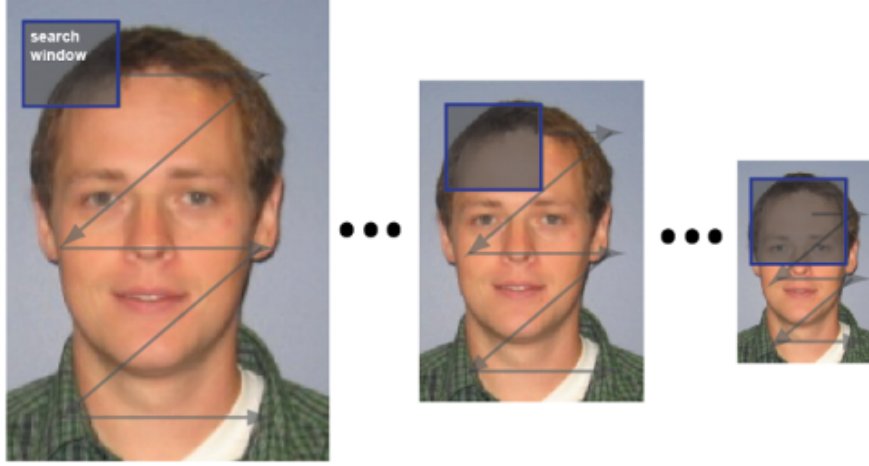


Figure 1: Depiction of a cascade object detector, where the image is successively downsampled until the search window properly encompasses the size of object. From <https://www.mathworks.com/help/vision/ref/vision.cascadeobjectdetector-system-object.html>

## 2.2 Face and Body Segmentation

Using the located face, the regions of the face and body can be determined. The head is important for extracting pulse rate from the facial tissues, and the body is important for extracting the respiratory rate from chest wall expansion and contraction during breathing cycles. Using the facial bounding box as the initial region, the head was defined as the face and neck, and was determined according to the following relationship:

$$\begin{bmatrix} x_h \\ y_h \\ w_h \\ h_h \end{bmatrix} = \begin{bmatrix} x_f \\ y_f - 0.1h_f \\ w_f \\ 1.6h_f \end{bmatrix} \quad (1)$$

where  $(x, y, w, h)$  represents the ROI of the face ( $\cdot_f$ ) and head ( $\cdot_h$ ) based on its top-left coordinate  $(x, y)$  and width and height  $(w, h)$ . Here, the height was extended to incorporate the neck, and was shifted upward to include the entire forehead. Then, the body ROI was defined as  $I \setminus R_h$ , where  $I$  is the image space,  $R_h$  is the head ROI, and  $\setminus$  is the set difference operator.

## 2.3 Absorbance Calculation

In order to extract the blood volume signal, pixel reflectance values were converted to absorbance values:

$$A(x, y, t) = -\log R(x, y, t) \quad (2)$$

where  $R(x, y, t)$  is the reflectance value at pixel coordinate  $(x, y)$  across time, and  $A(x, y, t)$  is the computed absorbance signal. This relationship is derived from the Beer-Lambert law of light attenuation in absorbing media:

$$A = -\log \frac{I}{I_0} \quad (3)$$

where  $I$  and  $I_0$  are the transmitted and incident illumination, respectively.

## 2.4 Pulse Rate and Respiratory Rate Extraction

The pulse rate and respiratory rate were extracting using a physiologically motivated weighted signal fusion approach. Specifically, a signal strength weight was computed for each signal representing the estimated strength of the signal as the underlying physiological (pulse or respiratory) signal. The underlying model uses the physiological prior information of signal quasi-periodicity to estimate signal strength. Normalized spectral entropy was used to model quasi-periodicity:

$$H(z) = \frac{F(z) \cdot \log_2 F(z)}{\log_2 N} \quad (4)$$

where  $N$  is the sample length, and  $F(z)$  is the normalized spectral power of signal  $z$  according to its Fourier transform. Note that, since entropy is a measure of randomness, quasi-periodicity is indicated by small values of entropy (i.e., low randomness). Thus, the weights for each signal were computed as inversely proportional to spectral entropy:

$$w(x, y) = 1 - H(A(x, y)) \quad (5)$$

The pulse rate and respiratory rate signals were then estimated using a weighted fusion of the underlying biophotonic signals:

$$z_{PR} = \frac{\sum_{(x_h, y_h) \in R_h} w(x_h, y_h) A(x_h, y_h)}{\sum_{(x_h, y_h) \in R_h} w(x_h, y_h)} \quad (6)$$

$$z_{RR} = \frac{\sum_{(x_b, y_b) \in R_b} w(x_b, y_b) A(x_b, y_b)}{\sum_{(x_b, y_b) \in R_b} w(x_b, y_b)} \quad (7)$$

where  $R_h$  and  $R_b$  are the regions of the head and body as before. The pulse rate and respiratory rate were then computed as the maximum frequency response of the respective signals.

## 3 Experimental Results

This PR and RR algorithm was evaluated using the setup discussed in Section 1. Figure 2 shows the detected face using the Viola-Jones cascade object detector from the overhead view. The ROI correctly identified the face from face, excluding the neck and upper forehead. This serves as a baseline for extracting PR and RR.

Note how the facial bounding box does not encompass the entire forehead and neck. This is a fundamental characteristic of the face detector, since it was trained to detect faces only. Thus, the bounding box was increased according to Equation 1. Then, the signal strength weights were computed according to Equation 5 in preparation for signal fusion. Figure 3 shows a visual depiction of this process. Brighter areas indicate stronger computed weights. On the left, we can see that the weights are strongest through the neck, cheek and forehead region, indicating that the algorithm identified pulsatility through the carotid and facial arteries. On the right, we can see that the weights are strongest around the chest cavity, indicating that the algorithm identified motion resulting from chest cavity expansion and contraction.

Using these weights, the pulse rate and breathing rate signals were extracted using the signal fusion method from Equation 6 and Equation 7. The pulse rate signal shows strong arterial pulsing that is quasi periodic, resulting in a clean frequency peak at 67 bpm. Similarly, the respiratory signal is slow and quasi-periodic, resulting in a respiratory rate of 22 bpm. These were both consistent with manual observations of the participant’s physiological state.



Figure 2: Frame with automatically detected facial region highlighted.

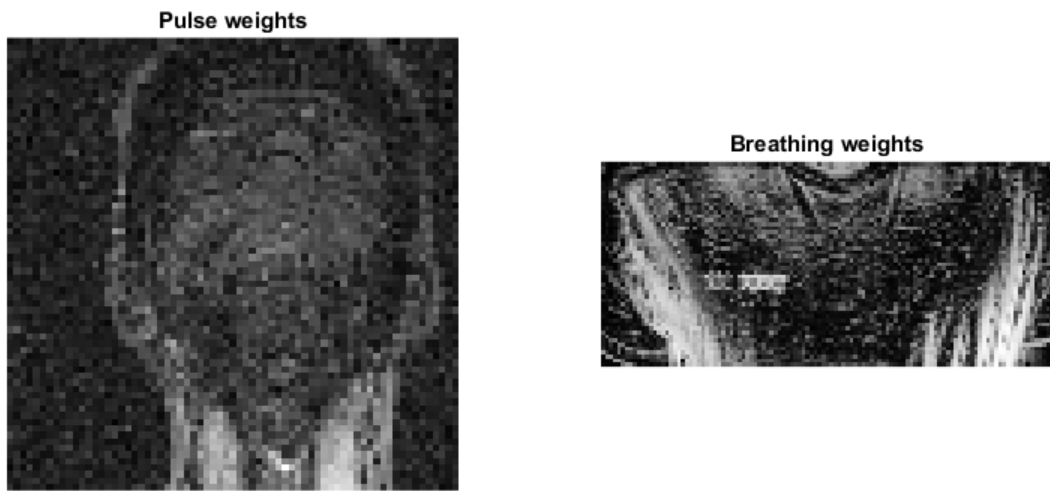


Figure 3: Weights from frame representing the relative strength of signal at each location. The lighter the region, the stronger the weight for extracting the pulse/respiratory rate.

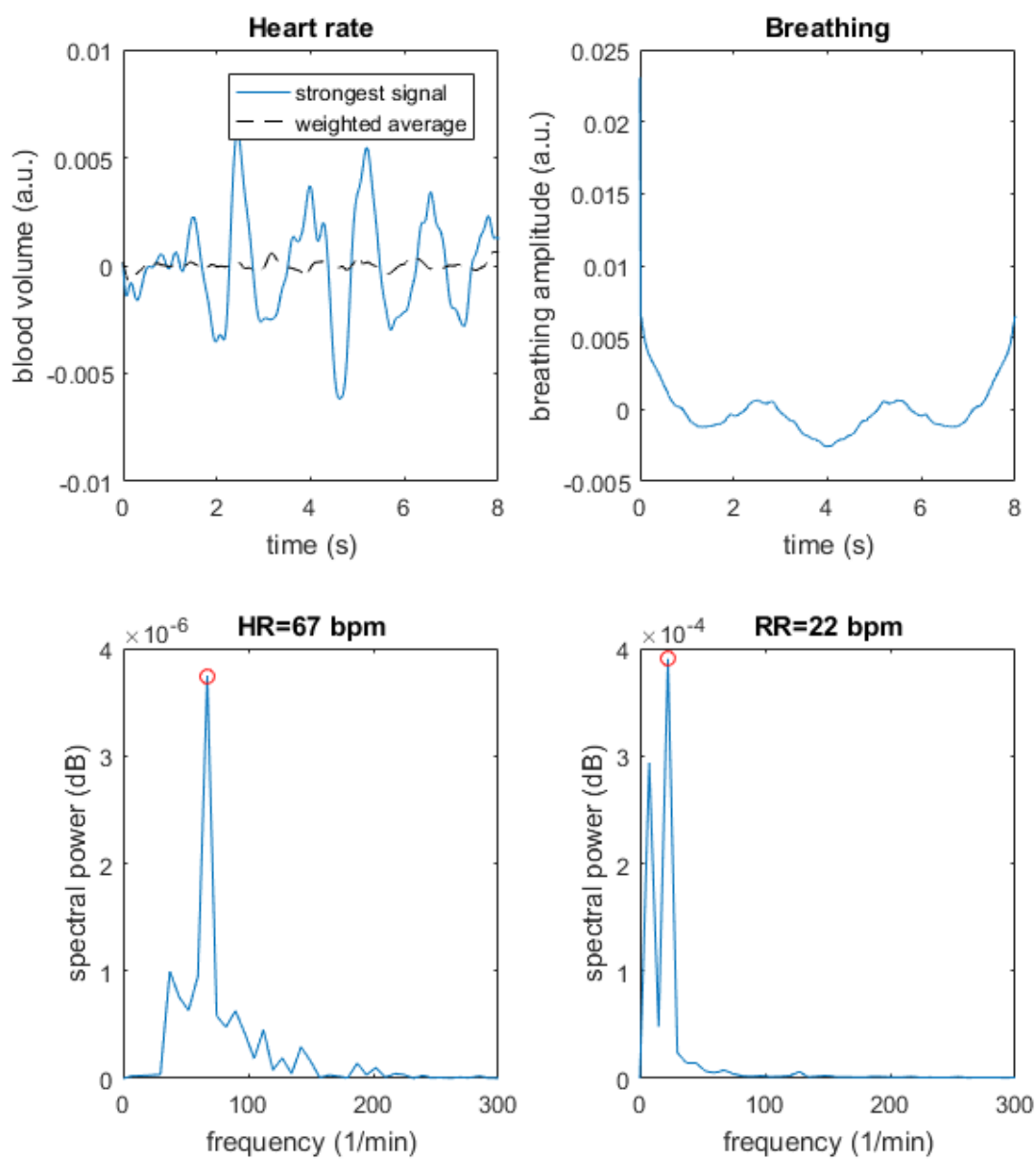


Figure 4: Extracted pulse rate and breathing rate signals (top) and values (bottom).