Respiratory detection using non-contact sensors

Chen Xinqiang¹, Huang Zhicheng¹, Su Lumei¹,²,a

¹School of Electrical Engineering and Automation, Xiamen University of Technology, Xiamen430070, China
²Xiaolong Institute of Engineering and Technology, Second floor, Building A, Longzhou, Industrial Park, Longyan364000, China

Corresponding author: sulumei@163.com

Abstract: Current respiratory detection study are mainly based on contact respiratory detection technology which has the limitation that human body should wear corresponding detection instruments. This paper studies three types of non-contact respiratory detection technologies. (1) Kinect-based respiratory detection method, which uses the depth sensor to obtain the depth value of the chest and abdomen region; (2) Ordinary web camera-based respiratory detection method, which applies optical flow to detect the motion information of the chest and abdomen region; (3) Thermal imager based respiratory detection method, which obtains the temperature change information of the nostril region from thermal video. The advantages and disadvantages of these three non-contact respiratory detection technologies are explored according to the experimental results. Kinect-based respiratory detection method is less affected by environmental interference, but the detection accuracy is lower and the detection range is limited. The detection method based on ordinary web camera is more accurate. However, the environmental illumination requirements are high. The detection method based on the thermal imager is not affected by the ambient light source, but the detection has poor real-time performance.

1. Introduction
Respiratory frequency is an important physiological index, which is closely related to the health of the human body. There are many contact respiratory detection technologies that can detect respiratory frequency. The technologies are mainly divided into two categories. One category is a thermistor-based respiratory frequency detecting device. The other category is a chest-belt type breathing detection device based on pressure sensor [1]. Since contact respiratory detection technologies requires the human body to wear the corresponding detecting instrument, it has certain inconvenience. But non-contact respiration detection technology is simpler. People can check the respiratory rate in their daily lives through smartphones, tablets and so on. Therefore, this paper studies three non-contact respiratory detection technologies. First, based on the detection method of the Kinect camera, the built-in depth sensor is used to obtain the depth value of the chest and abdomen region undulation [2]. Second, ordinary web camera-based respiratory detection method, which utilizes optical flow to detect the motion information of the chest and abdomen region [3]. Third, Thermal imager based respiratory detection method, which obtains the temperature change information of the nostril region from thermal video. Such technologies are lower in cost, more feasible, and can be widely promoted.
2. Non-contact respiratory detection technology

2.1. Non-contact respiratory detection system
The system is mainly divided into three parts. In the first part, three different sensors, Kinect, web camera and thermal imager are used for feature acquisition, respectively, and a data set strongly related to the respiratory signal is obtained. In the second part, since the data sets strongly related to the respiratory signal obtained by the above three detection technologies are time domain information. The time domain data sets need to be converted into frequency domain information by fast Fourier transform (FFT). So that the respiratory signal frequency distribution map can be obtained.

2.2. Three non-contact respiratory frequency detection method

2.2.1. Kinect-based respiratory depth feature extraction
When the human body is breathing, the fluctuation of the chest and abdomen area will cause the depth value of the depth image to change. The depth value is the gray value at each pixel. Let the pixel coordinates of each frame depth image be \((x, y)\). Select a pixel area \((x + w, y + h)\) with width \(w\) and height \(h\) as the region of interest (ROI). Position the chest and abdomen area to the ROI. And matrix summation of the chest and abdomen depth values obtained in each frame. When the sum is divided by the total number of pixels \(w \times h\), the average depth value \(I_D\) can be obtained as

\[
I_D = \frac{\sum_{k=1}^{w \times h} I_k}{w \times h}
\]

(1)

where \(I_k\) is the depth value of the kth pixel point of each frame of the depth image.
Finally, after obtaining the average depth value of each frame of the depth image, the average depth value of successive frames is taken as a data set, and this data set is used as a strong correlation data set of the respiratory signal.

2.2.2. Image-based respiratory motion feature extraction
The undulation of the chest and abdomen area during breathing corresponds to the change in brightness of the pixels on the image. Using the optical flow method to transform the velocity of the chest and abdomen region into the instantaneous velocity of pixel motion. By grading the image sequence, the motion information of the same gray pixel point between adjacent frames is found. Set \(I(x, y, t)\) and \(I(x + \delta x, y + \delta y, z + \delta z, t + \delta t)\) as two image sequences before and after. Then define an energy function.

\[
E(u, v) = \iint \left[ (I_x u + I_y v + I_t)^2 + \alpha^2 (P \nabla \mu P^2 + P \nabla v P^2) \right] dx \, dy
\]

(2)

where \((I_x u + I_y v + I_t)^2\) is gray scale change factor, and \(\alpha^2 (P \nabla \mu P^2 + P \nabla v P^2)\) is smoothing factor. The ideal optical flow field should make the two factors as small as possible. The optical flow value can be obtained by operation.

\[
| \lambda_k - \lambda_{k-1} | < \text{thresh} \tag{3}
\]

where

\[
\lambda = I_x \mu + I_y \nu + I_t 
\]

(4)

2.2.3. Respiratory temperature feature extraction based on thermal imager
When the human body breathes, the hot air exhaled by the nostrils causes the temperature of the nostril area to be greater than the temperature of the surrounding air, thereby causing a change in the thermal characteristics of the nostril area. First, the nostril area is used as the ROI. The temperature of \(\delta \times \delta\) pixel points in the vicinity of the nostril region was taken as a strong correlation data set, and the sum
of the data sets of each frame was calculated. Secondly, when the sum of the temperature of each frame is divided by the number of pixels, the average temperature $I_T$ can be obtained as

$$I_T = \frac{\sum_{k=1}^{\delta \times \delta} I_k}{\delta \times \delta}$$

(5)

where $I_k$ is the temperature value of the kth pixel point of each frame.

Finally, a continuous change in the temperature of the nostril can be obtained, that is, the change law of the respiratory signal can be analyzed.

2.3. **FFT-based respiratory rate detection**

All strong correlation data sets of respiratory signals are obtained by the above three non-contact respiratory signal detection technologies. They can only reflect the law of change in breathing. However, with the fast Fourier transform FFT, the time domain data set can be converted into frequency domain information, so that the human body's respiratory rate can be observed.

2.4. **Prediction of respiratory signals using LSTM**

This paper studies a method for predicting respiratory signals based on LSTM. LSTM unit is mainly composed of three "gates", which are input gate, forget gate and output gate [4]. The input gate is used to update the unit status, the forget gate determines what information is discarded or retained, and the output gate determines the output at the current moment. The formula for each "gate" is defined as follows:

$$z = \tanh(W_z[h_{t-1}, x_t])$$  

(6)

$$i = \sigma(W_i[h_{t-1}, x_t])$$  

(7)

$$f = \sigma(W_f[h_{t-1}, x_t])$$  

(8)

$$o = \sigma(W_o[h_{t-1}, x_t])$$  

(9)

$$c_t = f \cdot c_{t-1} + i \cdot z$$  

(10)

$$h_t = o \cdot \tanh(c_t)$$  

(11)

where $z$, $i$, $f$, $o$, $c_t$, $h_t$ are input value, input gate, forget gate, output gate, new state and output. $W_z$, $W_i$, $W_f$, $W_o$ are corresponding parameter matrices. $\sigma$ is a sigmoid function. $h_{t-1}$, $x_t$, and $c_{t-1}$ are the output of the previous time, the input of the current time, and the state of the previous time.

3. **Experimental results and analysis**

3.1. **Experimental result**

3.1.1. **Kinect-based respiration detection**

After obtaining the average depth value of each frame depth image, the average depth value of successive frames is taken as a data set, and this data set is used as a respiratory signal strong correlation data set. The experiment we designed was to fix the Kinect camera on the desk, and the tester sat about 1m away from the Kinect camera. During the test, the tester should try to stay still to avoid too much interference during the test. When ready, the tester begins normal breathing.

As can be seen in the Figure 1(a), the depth image data obtained by Kinect is not high precision. Therefore, the obtained waveform diagram does not respond well to the respiratory rate.
3.1.2. Respiratory detection based on ordinary web camera

The designed experiment is that the tester and the camera maintain a distance of about 0.5m to test the human body's breathing in the normal breathing mode. Gradient information on the chest and abdomen was collected in the experiment at 512 frames. First, normal breathing is performed, deep breathing is performed after the 180th frame, and normal breathing is resumed after the 300th frame. As shown in Figure 1(b), the amplitude of deep breathing is significantly larger than that of normal breathing, and the waveform is relatively smooth, which reflects the breathing frequency well, and the detection accuracy is high.

3.1.3. Respiratory detection based on thermal imager

The flir thermal imager is used to collect temperature change data in the nostril area, which can be approximated as the change law of human breathing. Align the marked point of the flir thermal imager to the nostril area, and then start to measure the temperature of the nostril area. As shown in Figure 2(a), when the human breathes, the temperature of the gas exchanged in the nostril area changes within the range of 36-40 °C. And the temperature waveform is zigzag, the flir thermal imaging camera has a low sensitivity to reflecting temperature changes in continuous time.

3.2. FFT to obtain the respiratory frequency spectrum

By analyzing the respiratory waveforms obtained by the above three non-contact detection methods, Kinect-based and thermal imager based respiratory detection method have low precision, and the detection precision of ordinary web camera-based respiratory detection method is high. Therefore, the FFT transform is performed by selecting a strong correlation data set of the respiratory signals collected by the ordinary web camera. The respiratory data of 512 frames of normal breathing collected by the ordinary camera is used as input data, and the fast Fourier transform FFT is performed to obtain a respiratory signal frequency distribution map. As shown in Figure 2(b), the human body's respiratory frequency is considered to be 0.05 Hz at this time.
3.3. LSTM-based respiratory signal prediction
First collect a strong correlation data set of respiratory signals through an ordinary web camera and input this data set into the LSTM model. Set the number of iterations to 100 and train the LSTM model. The results of respiratory signal prediction are shown in Figure 3. The blue line represents the observed respiratory waveform, and the red is the predicted respiratory waveform. It can be seen from the figure that the repetition rate of the blue curve and the red curve is high, and the root mean square error (RMSE) is 0.13829, which can meet the basic prediction requirements. Therefore, this method is suitable for predicting respiratory signals.

Figure 3. LSTM predicts respiratory signals

4. Conclusion
This paper studied three non-contact respiratory detection methods. The experimental results indicated that they each have their advantages and disadvantages. (1) Kinect-based non-contact breathing detection methods are not easily interfered by the light source environment. However, the detection rate of this method is low, and the requirements on the equipment are high, making it less practical. (2) The detection accuracy of ordinary web camera-based respiratory detection method is higher, and the hardware cost is lower, which is most suitable for practical application promotion. However, its application has certain limitations, and the brightness of the environment is also required. (3) Thermal imager based respiratory detection method is not affected by human motion and environmental light, and is most suitable for detecting respiratory signals of human body in real time. However, the equipment costs are relatively high, and they are also affected by ambient temperature.

Acknowledgments
This work is financially supported by Natural Science Foundation of department of science and technology of Fujian Province(No.2017J01510), Foundation for Science and technology of Longyan(No.2017LY91), Science Research Foundation for Overseas Scholars of Xiamen (No.XRS[2016]314-05).

References
[1] Xiao X Y, Huang S L, Chen S J, et al. Design of a wearable respiration monitoring system based on pressure sensor. Transducer and Microsystem Technologies, 2016,35(02):126-129.
[2] Li B, Xie D, Duan W J, et al. Kinect-based non-contract type breathing detecting system. Transducer and Microsystem Technologies,2017,36(03):107-109.
[3] Kumar M, Veeraraghavan A, and Sabharwal A. DistancePPG: Robust non-contact vital signs monitoring using a camera [J]. Biomedical optics express, 2015, 6(05): 1565–1588.
[4] Zeng M R, Zheng Z S. Luo S. Human behavior recognition combining two-stream CNN with LSTM. Modern Electronics Technique, 2019, (19):37-40.