Thermal Camera Based Physiological Monitoring with an Assistive Robot

S. Coşar[1], Z. Yan[1,2], F. Zhao[1,3], T. Lambrou[1], S. Yue[1], N. Bellotto[1]

***Abstract*— This paper presents a physiological monitoring system for assistive robots using a thermal camera. It is based on the detection of subtle changes in temperature observed on different parts of the face. First, we segment and estimate these face regions on thermal images. Then, by applying Fourier analysis on temperature data, we estimate respiration and heartbeat rate. This physiological monitoring system has been integrated in an assistive robot for elderly people at home, as part of the ENRICHME project. Its performance has been evaluated on a new thermal dataset for physiological monitoring, which is made publicly available for research purposes.**

I. INTRODUCTION

Considering the ageing population and increasing life-expectancy of people in the world, well-being and health monitoring is an important topic in various research fields. For assistive technologies, several approaches, from wearable to contact-free [9], [4], have been proposed for monitoring physiological parameters. Although wearable devices may have advantages over contract-free devices in terms of accuracy, they are often impractical, especially for elderly people with cognitive impairments, who may forget or refuse to wear them. Since many contact-free approaches require to stand in front of a fixed camera or sensor, it is hard to apply them in real-world scenarios for domestic use.

This paper proposes a contact-free physiological monitoring system integrated in a mobile cognitive assistive robot. This is part of the ENRICHME* project, which provides adaptive and cognitive stimulation for the elderly with mild cognitive impairments. the system detects subtle changes in temperature on thermal images, acquired by a thermal camera on the top of the robot (Fig. 1). The software developed for this application estimates three important physiological parameters, i.e. temperature, respiration and heartbeat rate.

The contributions of the paper are threefold:

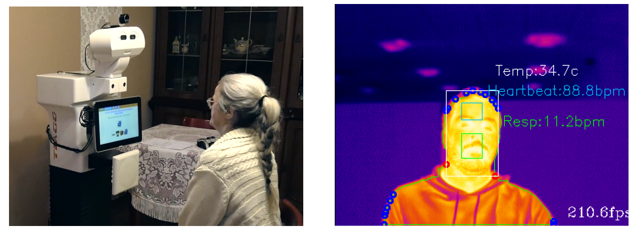
- A practical estimation method of face temperature, respiration rate and heartbeat rate combining face pose and region-based FFT analysis of thermal images;
- A complete and ready-to-use software pipeline for physiological monitoring with a mobile robot;
- A new publicly available thermal dataset for physiological monitoring.

*This work was supported by the EU H2020 projects ENRICHME (No. 643691 - http://www.enrichme.eu) and STEP2DYNA (No. 691154)

[1]Lincoln Centre for Autonomous Systems (L-CAS), School of Computer Science, University of Lincoln, United Kingdom. {scosar,tlambrou,syue,nbellotto}@lincoln.ac.uk

[2]LE2I-CNRS, University of Technology of Belfort-Montbéliard (UTBM), France. zhi.yan@utbm.fr

[3]Department of Computer Science, Liverpool John Moores University, United Kingdom. f.zhao@ljmu.ac.uk



(a) (b)

Fig. 1. The ENRICHME system integrates an assistive robot for the elderly (a) to monitor their physiological state with a thermal camera (b).

The remainder of the paper is as follows. Section II describes some related work on contact-free physiological monitoring. Section III explains how the physiological parameters are estimated. The experimental results with a new public dataset are presented in Section IV. Finally, we conclude this paper in Section V discussing achievements and current limitations.

II. RELATED WORK

Among the growing number of assistive technologies, those for physiological monitoring have become very popular. In this section, we omit wearable devices and focus on contact-free approaches relevant to this paper.

Most of the existing solutions utilize color images acquired from RGB cameras. Based on the fact that the human skin color varies slightly with the blood circulation, the authors in [13] amplify a band of temporal frequencies that include plausible human heart rates, revealing the variation of redness as blood flows through the face. Similarly, in [2], subtle head oscillations that accompany the cardiac cycle are exploited. The cyclical movement of blood from the heart to the head via the abdominal aorta and the carotid arteries causes the head to move in a periodic motion. The authors track feature points on a person's head and use principal component analysis (PCA) to find a periodic signal caused by the pulse. In [11], the respiration rate is estimated by detecting the chest movements using subsequent subtractions.

The previous approaches mostly rely on color information and therefore are affected by shadows and illumination changes. Thermal images can be used to reduce these issues. In [4], the heart rate is estimated by applying a Fast Fourier Transform (FFT) on the temperature signal, extracted from a line-based region along the major superficial vessel. The authors in [1] monitor the respiration rate by observing temperature changes on the nose region. They automatically detect the latter by finding the warmest and coldest points

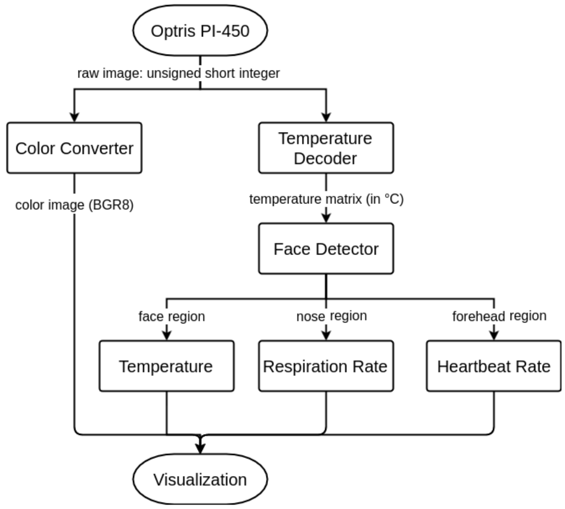


Fig. 2. The flow diagram of the physiological monitoring system.

on the face. Similarly, in [3], the authors exploit a pan-tilt thermal camera system to track the mouth-nose region and estimate the respiration rate by looking at temperature variations. The region is manually initialized by the user and then tracked using optical-flow. These approaches assume a semi-static setup with a steady user in front of the camera. Such assumption, however, is impractical for mobile robots in real-world application, in particular for monitoring elderly people at home. New solutions are needed to estimate physiological parameters on thermal images despite face or camera movements.

III. PHYSIOLOGICAL MONITORING

The system described next extracts face temperature, respiratory rate and heartbeat rate using a thermal camera (Optris PI-450). The main challenges include face pose estimation in thermal images and biometric feature extraction from temperature information. Fig. 2 shows the process steps.

A. Face Temperature

The raw thermal image (height = 288 pixels, width = 382 pixels, frame rate = 27 Hz) is initially converted to a colour image (BGR8) and a corresponding temperature matrix having the same size as the image. Then, we use Algorithm 1 to find the face region within the temperature matrix. Here, the thresholds t_{min} and t_{max} are manually chosen, assuming the body temperature of the user falls between 30°C and 40°C. After this, a morphological closing operation is performed to improve the face contour. Another threshold c_{min} is used on the latter to filter faces that are too small and difficult to process. Then, based on the approach in [12], the convexity of the contour is analysed to detect the actual face region. A visual representation of the process is shown in Fig. 3. For simplicity, and considering only personal robot applications, the current system assumes there is only one face in the image.

Algorithm 1: Face detection in temperature matrix.

Convert temperature matrix to binary image

$$binarypixel = \begin{cases} 255 & \text{if } t \in (t_{min}, t_{max}), \\ 0 & otherwise; \end{cases}$$

Apply morphological closing operation;
Find biggest contour c among all areas;
if $c > c_{min}$ **then**
    Analyse contour convexities to extract face area;
    Estimate nose and forehead locations using facial landmark ratios;
end

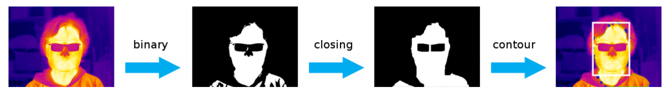
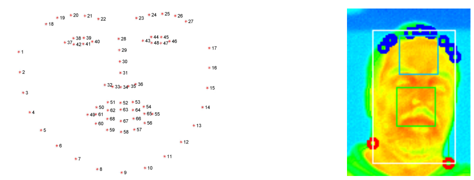


Fig. 3. Face detection in thermal images.



(a) (b)

Fig. 4. (a) the 68 facial landmark points detected using the face pose estimation in [7] and (b) forehead and nose region detection.

The final face temperature is calculated as follows:

$$t_{avg} = \frac{\sum_{i=0}^{n} t_i}{n}, \; t_i \in face\ region, \tag{1}$$

averaging all the n temperature values in the face region.

B. Respiration and Heartbeat Rate

Respiratory and heartbeat rate can be obtained with a Fourier analysis of the temperature data on the nose and the forehead regions, respectively [4], [1]. These regions are extracted with the help of a facial pose estimation algorithm [7], here applied to thermal images. The algorithm detects 68 face landmark points (Fig. 4-a) using the method proposed by [6] and trained on the iBUG 300-W dataset [10]. After detecting the landmark points, we delimit a forehead region of 30x30 pixel above and between eyebrow points. Similarly, we take the 30x30 pixel area around the nose points for the respective region (Fig. 4-b). Our implementation pipeline is similar to a webcam-based pulse detector[1], which uses the green channel of RGB images, instead of thermal images, to estimate the heartbeat rate.

The implementation of our respiration rate estimator is explained by Algorithm 2. A FIFO contains the time series data to be processed. The pre-defined number of samples (n_r) must be appropriately set in order to have enough data for the calculation. In our case, we assume that at least 10 second of data is required to evaluate the respiratory rate with our thermal camera at 27 Hz, and therefore $n_r = 270$.

[1] https://github.com/thearn/webcam-pulse-detector

Algorithm 2: Respiration rate estimation using FFT.

Set a FIFO data buffer with size b_r;
Fill the average temperatures t_i of the nose region (time series data) into the buffer;
Set a minimum number of samples n_r;
if $n_r \leq b_r$ **then**
  Generate evenly spaced time intervals m_i;
  Use one-dimensional linear interpolation to approximate t_i that matches m_i;
  Use Hamming window to minimize the signal side lobe in frequency domain;
  Normalize the temperature signal;
  FFT of the normalized temperature signal;
  Get the respiration rate according to the spectrum amplitude peak;
end

Because of the real-time acquisition, the temperature data could be unevenly spaced in time. Therefore, before applying FFT, we re-sample the evenly spaced t_i values at m_i intervals using a linear interpolation of the original data. We then use a Hamming window [5], [8] to minimize the signal side lobe (unwanted radiation) and improve the quality of the incoming data, resulting in a larger but smoother frequency peak.

$$\tilde{t_i}[n] = w_H[n] \cdot t_i[n] \tag{2}$$

where $w_H[n] = 0.54 - 0.46 \cdot cos(2\pi n/(N-1))$ is the Hamming window function and N is the length of the temperature data.

After the Hamming window, we take the FFT of the signal in the buffer and calculate the respiration rate corresponding to the peak amplitude in the frequency spectrum:

$$f_r = \arg\max_k(\tilde{T_i}[k]) \tag{3}$$

where $\tilde{T_i}[k] = \mathscr{F}\{\tilde{t_i}[n]\}$ is the FFT of the signal. To reduce the effect of noisy observations in thermal images, we also apply a moving average on the estimated respiration rate using a temporal window of b_r samples.

The estimation of the heartbeat rate is similar to the respiration one, but it is based on the temperature of the forehead region rather than the nose region. The buffer size (b_h) and a the number of samples (n_h) are also different.

IV. EXPERIMENTS

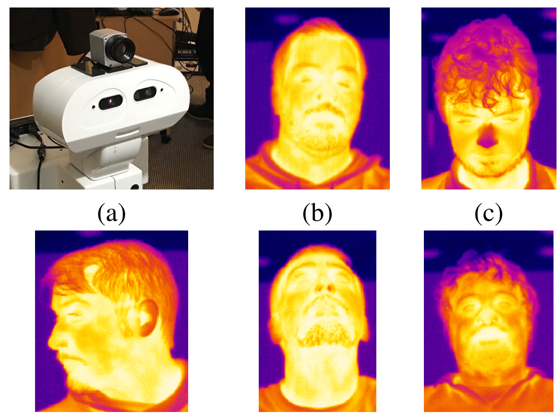
Our software for physiological monitoring is encapsulated into a Robot Operating System[2] (ROS) package, which can be easily installed in any ROS-compatible robot system. This package includes launch and configuration files with pre-defined parameters for physiological monitoring, as shown in Table I. The software is publicly available, together with our new thermal dataset[3]. This section presents the dataset and some preliminary results from the experimental evaluation of our system implemented on the ENRICHME assistive robot.

[2]http://www.ros.org
[3]https://lcas.lincoln.ac.uk/wp/research/data-sets-software/lcas-thermal-physiological-monitoring-dataset/

Name	Description	Default Value	Unit
$temp\_thr\_min$	Lower value for thresholding thermal images (t_{min})	30	°C
$temp\_thr\_max$	Higher value for thresholding thermal images (t_{max})	40	°C
$contour\_area\_min$	Minimum value for the area of the face contour (c_{min})	30	pixels
$resp\_buffer\_min$	Minimum number of samples (buffer) required for respiration rate estimation (n_r)	100	samples
$resp\_buffer\_max$	Size of the buffer for respiration rate (b_r)	1000	samples
$heart\_buffer\_min$	Minimum number of samples (buffer) required for heartbeat rate estimation (n_h)	10	samples
$heart\_buffer\_max$	Size of the buffer for heartbeat rate (b_h)	250	samples

TABLE I

PARAMETERS OF THE SOFTWARE FOR PHYSIOLOGICAL MONITORING AND THEIR DEFAULT VALUES.



(a) (b) (c) (d) (e) (f)

Fig. 5. (a) Optris PI-450 camera, mounted on the top of the robot, to record the thermal dataset. (b-f) Some thermal images from the dataset.

A. Thermal Dataset

We created a thermal-based physiological monitoring dataset of *rosbag* files for the ROS environment. The dataset contains 5 people sitting in front of the robot, observed by an Optris PI-450 thermal imager mounted on the top of the robot, at 1.3 m from the floor (Fig. 5-a). The distance between the robot and the user's face was approximately 1.5 m. Each file contains thermal images of the user's upper body, with a resolution of 382×288. We have recorded two-minute data for each person, with thermal images taken at 27 Hz. The participants were asked to remain still during the first minute, and then to move the head up and down, forward and backward, turning right and left. Each head position was kept for 10 seconds, in total collecting approximately 3,000 continuous images. Some examples are shown in Fig. 5-(b-f).

The ground truth of the temperature is the average value of the forehead region, manually segmented on the thermal image. Each participant wore a fingertip pulse oximeter (Acurio AS-302-L) during the recording, which provided the ground truth for the heartbeat rate. Finally, the ground truth of the respiration rate was obtained by manually annotating the time instants when the person was inhaling, and calculating the breath cycle times from these annotations.

STD of Temperature [°C]		
	Still	Moving
Person 1	0.14	0.19
Person 2	0.07	0.31
Person 3	0.06	0.33
Person 4	0.07	0.22
Person 5	0.09	0.16
Overall	0.09	0.25

TABLE II

STD OF THE TEMPERATURE ESTIMATED BY OUR SYSTEM.

	Still			Moving		
	Mean	STD	RMSE	Mean	STD	RMSE
Temp. [°C]	0.86	0.06	0.86	0.78	0.48	0.88
Resp. [bpm]	3.72	0.78	3.81	5.87	2.18	6.20
Heartbeat [bpm]	29.68	15.76	33.18	18.96	12.15	22.51

TABLE III

ESTIMATION ERRORS OF OUR PHYSIOLOGICAL MONITORING SYSTEM.

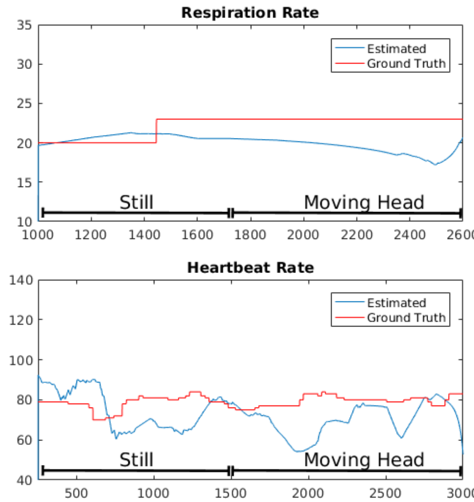


Fig. 6. Examples of respiration/heartbeat rate estimation and ground truth.

B. Physiological Monitoring Results

We ran our system on the collected dataset and estimated face temperature, heartbeat rate and respiration rate. Table II shows the standard deviation (STD) of the estimated temperatures when the persons were still and when they moved their heads. It can be observed that, despite a small increase, the temperature variations in both cases are within acceptable limits. Also, thanks to the known face pose, the head movements seem to influence the temperature estimation very little.

Fig. 6 shows some examples of estimated respiration and heartbeat rate, together with the respective ground truth. Table III, instead, reports the mean and STD of the absolute error for each parameter. Even in this case, we can see that our approach can estimate the face temperature with good approximation (average error $< 1°C$), for both the still and moving cases. However, the results for the respiration rate are less satisfactory, in particular when the person is moving, although the relatively small STD suggests the estimation may be affected by a systematic error, which could be reduced by an opportune calibration. Finally, for the heartbeat, we can see that the estimation is affected by considerable errors, independently of whether the participants were moving or not. This is not completely unexpected though, as the heartbeat rate estimation depends on the forehead position, which is very difficult to localize robustly (more than the nose), and the thermal frequency analysis can be very sensitive to small noises on this region.

V. CONCLUSIONS

This paper presented a contact-free physiological monitoring system for assistive robot applications. The system utilizes a thermal camera, mounted on a robot, to detect subtle changes in temperature of different face regions. Using Fourier analysis, we estimate also respiration and heartbeat rate. The proposed system has been evaluated on a new thermal dataset, made publicly available for the research community. Some preliminary experiments show that the proposed system has the potential to become an effective physiological monitoring solution for assistive robots in unconstrained settings, dealing with small movements of the person or of the robot itself. Although the estimation performance are not always satisfactory, the system has been successfully implemented and used in ENRICHME to determine approximate levels (i.e. high, normal, low) of physiological parameters. In our future work, we will exploit the RGB-D camera of the robot for face tracking to facilitate the localization of nose and forehead regions, improving the estimation of respiration and heartbeat rate, respectively.

REFERENCES

[1] F. Q. AL-Khalidi, R. Saatchi, D. Burke, and H. Elphick. Tracking human face features in thermal images for respiration monitoring. In *AICCSA 2010*, pages 1–6, May 2010.

[2] G. Balakrishnan, F. Durand, and J. Guttag. Detecting pulse from head motions in video. In *CVPR*, June 2013.

[3] R. Chauvin, M. Hamel, S. Brire, F. Ferland, F. Grondin, D. Ltourneau, M. Tousignant, and F. Michaud. Contact-free respiration rate monitoring using a pan-tilt thermal camera for stationary bike telerehabilitation sessions. *IEEE Systems Journal*, 10(3):1046–1055, Sept 2016.

[4] M. Garbey, N. Sun, A. Merla, and I. Pavlidis. Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. *IEEE Trans. on Biomedical Engineering*, 54(8):1418–1426, Aug 2007.

[5] F. J. Harris. On the use of windows for harmonic analysis with the discrete fourier transform. *Proc. of the IEEE*, 66(1):51–83, Jan 1978.

[6] V. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees. In *CVPR*, 2014.

[7] D. E. King. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10:1755–1758, 2009.

[8] A. Nuttall. Some windows with very good sidelobe behavior. *IEEE Trans. on Acoustics, Speech, and Signal Proc.*, 29(1):84–91, Feb 1981.

[9] A. Pantelopoulos and N. G. Bourbakis. A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, 40(1):1–12, Jan 2010.

[10] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 faces in-the-wild challenge: database and results. *Image and Vision Computing*, 47:3 – 18, 2016.

[11] K. S. Tan, R. Saatchi, H. Elphick, and D. Burke. Real-time vision based respiration monitoring system. In *CSNDSP*, pages 770–774, 2010.

[12] W. K. Wong, J. H. Hui, J. B. M. Desa, N. I. N. B. Ishak, A. B. Sulaiman, and Y. B. M. Nor. Face detection in thermal imaging using head curve geometry. In *ICIP*, pages 881–884, Oct 2012.

[13] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 31(4), 2012.