# Contactless Human Physical Signal Measurement for Collaborative Robotic Using Visible-IR Dual Camera

Jiguang Li
*North of England Robotics Innovation Centre*
University of Salford
Salford, UK
J.li56@salford.ac.uk

Wei Yao
*North of England Robotics Innovation Centre*
University of Salford
Salford, UK
J.li56@salford.ac.uk

***Abstract*—Collaborative robotic configurations for monitoring and tracking human beings for safety and efficiency have attracted interest in industrial revolution. The fusion of different types of sensors embedded in collaborative robotic systems significantly improve robotic perception. However, current methods have not deeply explored the capabilities of multi-sensory configurations including visible and thermal sensors. In this paper, we propose a contactless multi-sensor fusion including visible and thermal dual camera for collaborative robots to improve the robotic perception for human safety. Remote photoplethysmography detection and infrared thermal camera were used to measure the heart rate and body temperature.**

***Keywords—Heart rate, Dual camera, Robotics, Thermal camera, Contactless***

## I. Introduction

As we move towards Industry 5.0, the use of robots in industry, as well as our lives, will be increasingly frequent. The utilization of industrial robot sensors in robotics is anticipated to grow rapidly as robotics developments lead to flexible production models [1-2]. Robotic sensor is a device used to evaluate the environment in which the robot is operating and allows the robot to adjust actions based on collected data. A new imperative objective is to design a user-friendly interface to support safe and efficient worker-robot interactions.

Ccollaborative robot (Cobot) for worker-robot collaboration is a promising solution for complex tasks, which can integrate the robots' advantages in strength and accuracy with human ability in intuitive decision-making and adaptability. [3] Monitoring the status of human beings is an important task of such robots.

Physiological signals offer valuable insights into a person's physical and psychological state. Traditionally, these signals are obtained using wearable devices with embedded sensors, however, there are numerous downsides to such devices. Collecting information remotely from users creates opportunities for promoting safer interactive work between humans and the robotics. This non-contact measurement method has emerged as an increasingly noteworthy research area. Visible cameras have been used for contactless information collection [4].

However, existing methods mainly depend on RGB images captured by visible cameras, which are prone to be affected by environmental disturbances, such as poor illumination, fog, and smoke, etc. The usual RGB images sometimes also suffer from privacy issues.

It is important to look beyond their standalone capabilities and explore the synergistic combinations between different types of sensors. In engineering system, the fusion methods are significant because the system could provide capability to systems with different sensors especially beyond that individual system of the sensors. The fusion of different types of sensors embedded in collaborative robotic systems achieves high-quality information and contributes to significantly improve robotic perception. Multi-sensor fusion of data allows integration of data from various sensors makes easier for decision making, planning, executing and control of automation.

Photoplethysmography (PPG) is a simple optical technique for detecting blood volume changes in the microvascular bed of tissues. Remote PPG (rPPG) detection has been used and studied in many fields [5-10]. The advantages of rPPG are that it can be extracted with RGB cameras without the need for additional equipment to detect existing PPGs and it is non-contact.

Heart rate (HR) ranks among the most critical physiological indicators in the human body, significantly indicating an individual's state of physical health. In this work, we will explore the heart rate and temperature information estimation with RGB and infrared camera.

## II. Related Work

Visible and thermal image fusion has been widely used in many applications such as human detection. Shopovska's [11] experiments show that the visibility of pedestrians is noticeably improved especially in dark regions and at night with visible and thermal image fusion. They focus on generating fused images with high visual similarity to regular TrueColor images, while introducing new informative details in pedestrian regions.

Correa et al [12] present a robot that uses an infrared camera to sense humans and an optical camera to acquire a video of the scene. Robust human detection is achieved using thermal and visual information sources that are integrated to detect human-candidate objects, which are further processed in order to verify the presence of humans and their identity using face information in the thermal and visual spectrums. Face detection is used to verify the presence of humans, and face recognition to identify them.

According to Gelfert S, [13] thermal cameras, in combination with an additional sensor, seems to be the most used approach for human detection in regular environments. Results show a detection rate of up to 92% while using a thermal camera to detect humans and an optical camera to acquire the scene.

The visible and infrared dual camara fusion was used for robotic applications. Nikolas Theissen [14] proposes an approach to thermal compensation based on joint power consumption for articulated industrial robots. This work focuses on the modelling, measurement and identification of the change of the kinematic chain of serial articulated industrial robots based on thermo-mechanical deformations due to self-heating. A FLIR SC640 infrared camera is used for thermal sensing.

Coşar et al [15] presents a human re-identification method for service robots using thermal images using entropy-based sampling. In robotic applications, as the number and size of thermal datasets is limited, it is hard to use approaches that require huge amount of training samples. They propose a re-identification system that can work using only a small amount of data.

In recent years, several core rPPG methods have been proposed for extracting the pulse signal from a video including:

Blind source separation (BSS) [6]. It separates temporal RGB traces into uncorrelated or independent signal sources to retrieve the pulse.

CHROM [7], a chrominance-based method, which linearly combines the chrominance signals by assuming a standardized skin color to white balance the images.

PBV[8], which uses the signature of blood volume changes in different wavelengths to explicitly distinguish the pulse-induced color changes from motion noise in RGB measurements. The different absorption spectra of arterial blood and bloodless skin cause the variations to occur along a very specific vector in a normalized RGB-space. The exact vector can be determined for a given light spectrum and for given transfer characteristics of the optical filters in the camera. This algorithm shows much better motion robustness compared to BSS and CHROM methods.

Spatial Subspace Rotation(2SR) [9] which measures the temporal rotation of the spatial subspace of skin pixels for pulse extraction. It is a completely data-driven method to improve motion robustness.

POS (Plane-Orthogonal-to-Skin) [10] defines a plane orthogonal to the skin tone in the temporally normalized RGB space for pulse extraction. POS obtains the overall excellent performance among model-based methods. In this work we use POS algorithm for heart rate estimation.

With the rapid development of deep learning, new technologies have spurred the emergence of many new rPPG methods for heart rate measurement, we will do this in future work.

## III. METHODS

The process flow of our proposed method is shown in Figure 1.

The cameras start to record the video information. Then face detection algorithm is implemented to segment the face from the video. Temperature of the face region is usually different from the ambient environment. Therefore, temperature signal from IR camera can be used to verify the correct face region. Region of interest (ROI), skin segmentation is used to extract the skin region that can be used to calculate the pulse rate. Skin pixel averaging and POS algorithms are implemented for signal extraction. The movement of the head must be considered. Therefore, the active 3D head modelling a motion suppression algorithms are implemented at the same time. The final combined signal is filtered to detect the peak. The heart rate is calculated.

### A. Tools

Two sets of visible (RGB) camera and infrared (IR) camera were used for data acquisition is an RGB camera. The first one is the FLIR ONE dual camera from FLIR Systems,

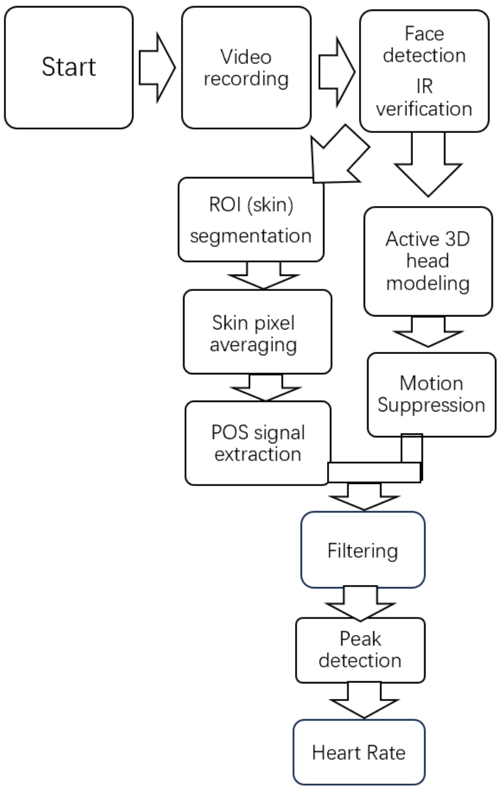


Fig.1 The process flow of our proposed method

Inc. (now Teledyne FLIR). The resolution of the IR camera is 160 × 120. The second one is the designed dual camera with RGB camera and camera with infrared thermopile array from

Fig.2 FLIR ONE dual camera and our designed dual camera with RGB and with infrared thermopile array

Heimann Sensor GmbH. The resolution of the array is 32× 32. These two dual cameras are shown in Figure 2.

The dual camera with FLIR camera was used in this work. The dual camera with Heimann camera will also be explored in industry cobot due to it is more economical for industry application. The dual camera with Heimann camera is mounted on a cobot in n the North of England Robotics Innovation Centre, as shown in Figure 3.

The calculations have been implemented on a laptop with

Fig.3 The dual camera was mounted on the Collaboration robot in the North of England Robotics Innovation Centre

an Intel Core i7 processor (2.70 GHz), NVIDIA® GeForce RTX™ 4050, 6 GB GDDR6, and 16-GB RAM.

### B. *Calculation*

Face detection

First, the face area is detected and captured by the RGB camera. The actual face area used for calculation in this work is a composite of two regions: (a) the cheek area between the eyes and the jaw, (b) the area around the nose, as shown in Figure 4.

We use OpenCV for face detection and rPPG algorithm to get your body signals.

OpenCV already contains many pre-trained classifiers for face, eyes, smiles, etc. In this paper we will be using the face classifier.

Fig.4 The actual face area used for calculation in this work

POS algorithm

POS algorithm is designed targeting certain applications and effects key point is defining two projection axes on the plane that can bound a most likely pulsatile region. Fig. XXX shows the distribution of the pulsatile strength on the plane orthogonal to $(1,1,1)^{T}$ as a function of **z**. Vectors and matrices are denoted as boldface characters, where the column vectors are denoted as **v**. The variable $t$ denotes the time, $T$ denotes the transposition.

Distribution of pulsatile strength on the plane orthogonal to $(1,1,1)^{T}$, The projection plane consists of 360 (discrete) projection axis **z** sampled with 1° difference, where the red/blue color denotes the regions with stronger/weaker pulsatile strength.

Three projection axes are exemplified on the plane: **z1** = $[-2, 1, 1]^{T}$, **z2** = $[1, -2, 1]^{T}$, and **z3** = $[1, 1, -2]^{T}$, which have the pulsatilities −0.64, 0.68, and −0.04. A temporally normalized RGB signal Cn (t) = [Rn (t), Gn (t), Bn (t)], measured from the skin in a video is projected onto **z1, z2 , z3** and obtain **S1 (t), S2 (t), S3 (t)**.

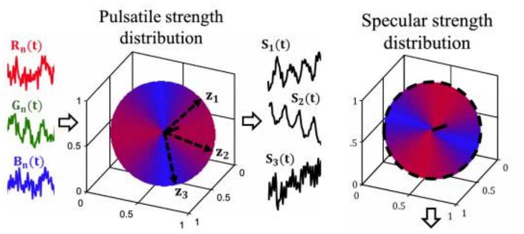


Fig.5. Distribution of pulsatile strength on the plane orthogonal to $(1,1,1)^{T}$, where the pulsatile strength is the absolute pulsatility value. Redesigned from [XXX

POS algorithm computes in the following steps:

1) Spatial averaging
2) Temporal normalization
3) Projection
4) Tuning
5) Overlap-adding

Heart Rate Estimation

In the previous steps we got heart rate signal. Now we estimate heart rate from the signal with peak detection. To refine the signal for peak detection, the signal is usually interpolated using a cubic spline function. The peaks can then be easily identified using a moving window, as they are the maxima within the signal.

Using individual peaks, extracting more information such as HR variability from the inter-beat intervals is possible.

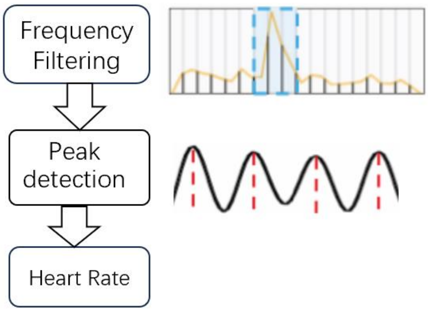


Fig.6 Heart rate estimation after signal extraction and fusion

## IV. Results and Discussion

The images from each camera and the fusion of the images are shown in Figure.7.

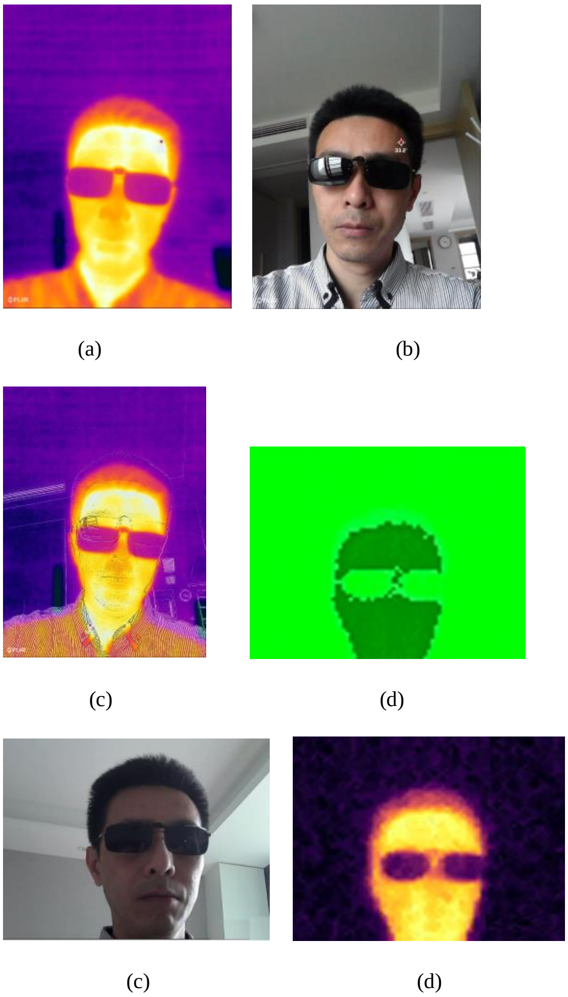


Fig.7 .Images from dual camera (a) Thermal imaging from IR camera of FLIR One,(b)RGB imaging from visible camera of FLIR One (c) the fusion of images from (a) and (b), (d) Thermal matrix from IR array of dual camera with Heimann sensor,(e)RGB imaging from visible camera of of dual camera with Heimann sensor (f) the fusion of images from (d) and (e).

There are several facts could infect the result including skin tone, luminance, movement, etc. Although we already use filter and projection methods to minimize these effects, but we still cannot ignore them. The IR camera helps to identify the skin region on face by measuring the temperature on the face. The face detection by visible camera will make the extraction and calculation of temperature more accurate. Because the calibration of IR is more accurate with the accurate face area.

We did many tests by comparison with apple watch's ECG, and the final error range is +7 bpm. this result is tested in a room with bright and stable light source.

## V. Conclusion and Future work

The goal of the paper was to develop and test visible and infrared fusion system with the idea to facilitate human-robot interaction based on the real-time detection of workers' physical signals during a collaborative task. The heart rate and temperature are successfully measured through this fusion of visible and infrared signals.

With the rapid advancement of research and technology, contactless rPPG methods have found applications in many domains beyond remote HR measurement such as measuring multiple vital signs, affective computing, deepfake detection and so on. Researchers have successfully demonstrated [16] the application of rPPG methods in the field of affective computing, particularly in stress estimation and emotion recognition. It will help to create safe and comfortable environment during the human-robot interaction.

Significant remote photoplethysmography (rPPG) research for estimating of the heart rate has leveraged supervised deep learning for robust signal extraction [17]. We will implement these method to our fusion system in the next step.

The focus of rPPG method research should also continue to be on addressing various influencing factors to improve the performance of rPPG methods to real-world application levels. The fusion of multis-senor makes easier for decision making, planning, executing and control of automation.

A further example is given in article [18] where a system combining facial recognition and audio-lingual affection recognition is proposed. Such approach would not be always applicable in noisy industrial environments, where an implicit emotion recognition system on cobots. It would facilitate tasks and help the robot to connect more with the operator. By using a thermal camera in addition to a depth one, we can more effectively segment human regions and calculate thermal parameters. We will continue to work on the improvement of perception of robotics during the human-robotic interaction for a better human life.

## Acknowledgment

The North of England Robotics Innovation Centre (NERIC) Project has been part-funded by Research England Development (RED). Special thanks to Simon Fong and Nicola Ronan - project managers.

## Funding

The NERIC Project has been part-funded by Research England Development (RED).

## Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

## References


[1] Patil, S., Vasu, V. & Srinadh, K.V.S. Advances and perspectives in collaborative robotics: a review of key technologies and emerging trends. Discov Mechanical Engineering 2, 13 (2023)

[2] Hao Li, Wenfeng Ma, Haoqi Wang, et al,A framework and method for Human-Robot cooperative safe control based on digital twin,Advanced Engineering Informatics, 53,101701,(2022),

[3] Burden, A.G., Caldwell, G.A. & Guertler, M.R. Towards human–robot collaboration in construction: current cobot trends and forecasts. Constr Robot 6, 209–220 (2022).

[4] W. Verkruysse et al., "Remote plethysmographic imaging using ambient light," Opt. Express, vol. 16, no. 26, pp. 21434–21445, Dec.(2008)

[5] M. Lewandowska et al., "Measuring pulse rate with a webcam— A non-contact method for evaluating cardiac activity," in Proc. Federated Conf. Comput. Sci. Inf. Syst., Szczecin, Poland, Sep. 2011, pp. 405–410.

[6] M.-Z. Poh et al., "Advancements in noncontact, multiparameter physiological measurements using a webcam," IEEE Trans. Biomed. Eng., vol. 58, no. 1, pp. 7–11, Jan. 2011.

[7] G. de Haan and V. Jeanne, "Robust pulse rate from chrominance-based rPPG," IEEE Trans. Biomed. Eng., vol. 60, no. 10, pp. 2878–2886, Oct. 2013.

[8] G. de Haan and A. van Leest, "Improved motion robustness of remote PPG by using the blood volume pulse signature," Physiol. Meas., vol. 35, no. 9, pp. 1913–1922, Oct. 2014.

[9] W. Wang et al., "A novel algorithm for remote photoplethysmography: Spatial subspace rotation," IEEE Trans. Biomed. Eng., vol. 63, no. 9, pp. 1974–1984, Sep. 2016.

[10] W. Wang, A. C. den Brinker, S. Stuijk and G. de Haan, "Algorithmic Principles of Remote PPG," in IEEE Transactions on Biomedical Engineering, vol. 64, no. 7, pp. 1479-1491, (2017)

[11] Shopovska, I.; Jovanov, L.; Philips, W. Deep Visible and Thermal Image Fusion for Enhanced Pedestrian Visibility. Sensors 2019, 19, 3727. https://doi.org/10.3390/s19173727

[12] Theissen, Nikolas Alexander. "Articulated industrial robots: An approach to thermal compensation based on joint power consumption." Conference: Laser Metrology and Machine Performance (Lamdamap) XIII, AMRC, Sheffield, UK (2019).

[13] Coşar, S., Bellotto, N. Human Re-Identification with a Robot Thermal Camera Using Entropy-Based Sampling. J Intell Robot Syst 98, 85-102 (2020). https://doi.org/10.1007/s10846-019-01026-w

[14] Gelfert S. A sensor review for human detection with robotic systems in regular and smoky environments. International Journal of Advanced Robotic Systems. 2023;20(3). doi:10.1177/17298806231175238

[15] 64. Correa M, Hermosilla G, Verschae R, et al. Human detection and identification by robots using thermal and visual information in domestic environments. J Intell Robot Syst 2011; 66(1-2): 223-243.

[16] Sabour R.M., Benezeth Y., De Oliveira P., Chappe J., Yang F. Ubfc-phys: A multimodal database for psychophysiological studies of social stress,IEEE Trans. Affect. Comput. (2021)

[17] J. Speth, N. Vance, P. Flynn, A. Czajka, Non-Contrastive Unsupervised Learning of Physiological Signals from Video, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.

[18] Alepis, E., Stathopoulou, I. O., Virvou, M., Tsihrintzis, G. A., & Kabassi, K. (2010, October). Audio-lingual and visual-facial emotion recognition: Towards a bi-modal interaction system. In 2010 22nd IEEE International Conference on Tools with Artificial Intelligence (Vol. 2, pp. 274-281).