2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)
November 3 - 5, 2015, Seoul, Korea

Reliable Stress Measurement using Face Temperature Variation with a Thermal Camera in Human-Robot Interaction

Mihaela Sorostinean   François Ferland   Adriana Tapus

Abstract— **For companion robots to properly interact with people, they should be able to gain awareness of their effect on the person's state and learn to behave accordingly. Advances in physiological monitoring sensors, such as infrared cameras, may enable robots to get an insight on a person's emotional state, which has a strong correlation with the physiological parameters. Performing this monitoring in a contact-free, user transparent manner is an important requirement. In this paper, we used a thermal imaging sensor mounted on the Meka humanoid robot, to remotely measure temperature variations on the face of the person while interacting with it. The scenario was a card-based quiz with the robot, which was aimed to induce a stressful state and 16 participants took part in the experiment. Several conditions with variations in the distance, height, and gaze of the robot were tested. A statistically significant variation in temperature in the nasal area of the face was measured while the robot was standing in the personal space of the participant. The first preliminary results encourage us to believe that temperature variations across various areas of people's faces can hint companion robots to adjust their behavior in human interactions.**

I. Introduction

While it is still not known the extent to which emotions are physiological or cognitive processes, it is widely agreed that emotional states have an impact on the physiological parameters of a person [1], [2], [3]. Thus, physiological sensing has become an important tool for researchers interested in affective computing [4], [5]. Embodying a robot with physiological measurement capabilities can offer it a continuous method to get an insight on a person's emotional state and behave accordingly.

Recent advances in infrared techniques have offered researchers the possibility to directly assess aspects of emotional experiences through the measurement of skin temperature. The autonomic nervous system (ANS) is an important part of the human emotional response. Fluctuations in the blood flow beneath the skin are controlled by the sympathetic nervous system, which is a part of the ANS, and result in temperature changes at the skin surface, due to either vasoconstriction, which leads to cooling, or vasodilatation, which leads to warming [6]. Furthermore, emotions usually lead to activation of the facial muscles, which can be associated with changes in the blood flow on various regions of the face and therefore with regional changes in temperature.

Hence, it is plausible that thermal imaging can be used to detect patterns of temperature variations across the human face that can be associated with certain emotions. [7] puts together a review of 23 experimental procedures that employ functional thermal infrared imaging for investigations of affective nature. A short overview on changes observed in the temperature of different regions of the face in studies focused on evaluating the level of stress induced by mental workload tasks is also provided. Generally, a decrease in the temperature of the nose and maxillary area has been observed, while other regions of the face, such as the forehead, have shown both increase and decrease in temperature.

Physiological sensors have been previously used to monitor the state of a person [8]. More recently, the authors in [9] present a machine learning approach meant to detect stress in people wearing a physiological sensing device. Participants were asked to perform both stressful and neutral tasks while the proposed technique used continuous physiological measurements to classify their state into "stressful" and "non-stressful". However, a disadvantage of their approach is the necessity of wearing the device, which may be considered intrusive.

In a study related to human-computer interaction, a thermal infrared system was used to measure in a non-invasive manner signs of frustration. Based on experiments conducted with 12 participants, it was observed that users' stress is correlated with increased blood flow in the frontal vessels of the forehead region [10]. Furthermore, an increase in the cutaneous forehead temperature has also been observed in [11], while participants were being interrogated for a mock crime scenario.

Moreover, Ioannou et al. [12] used a thermal infrared imaging system to evaluate the effect of proximity and gaze in a study on human-human interaction. Considering two interpersonal distances, intimate and social, and two gaze conditions, direct and averted, the authors have found strong correlations between these and the temperature in various areas of the face such as the forehead, chin, cheeks, nose, maxillary, and periorbital region. More specifically, an increase in the temperature of these areas has been detected when the experimenter moved in the intimate space of the participant.

Considering the results obtained in these previous works, we propose to use a thermal imaging sensor to investigate the changes that might occur in the physiological state of a person during a human-robot interaction scenario. We believe that while performing a cognitively demanding task, in this case a card game based on geography knowledge, changes in temperature of different regions of the face could be observed. Furthermore, we believe that the presence of the robot could have an impact on the evolution of the

The authors are at the Computer Vision and Robotics Lab, Computer Science and System Engineering Department, ENSTA-ParisTech, Palaiseau, France `firstname.lastname@ensta-paristech.fr`

978-1-4799-6885-5/15/$31.00

person's physiological state. Thus, while varying between the appearance of the robot (Tall or Short), its behavior (Direct or Averted Gaze), and the distance at which the interaction takes place (Social or Personal space), a higher level of change in temperature can be expected when the robot is found in the dominant role, for instance as a Tall robot in the Personal space of the participant.

This paper is structured as follows. Section II describes the robotic platform and the sensors used in our work. Section III presents the experimental design put in place to validate the approach and the hypotheses. Section IV describes the system architecture and the methodology used for the experiment. Section V gives the results from the experiment, and finally Section VI concludes the paper and offers a perspective on future work.

II. Robot and Sensors Test-beds

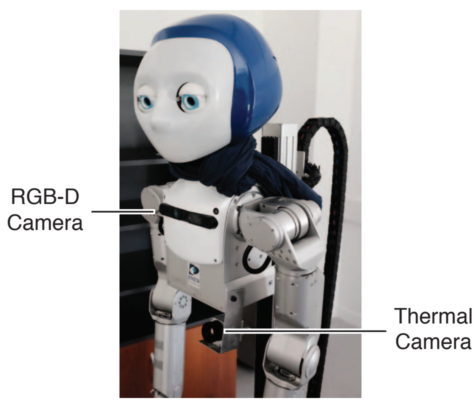


Fig. 1. The Meka humanoid robot with the RGB-D and thermal cameras used in this experiment.

The experiments presented in this work have been conducted with the Meka M-1 humanoid robot, shown in Figure 1. Designed for a wide range of expressive postures, it is a platform particularly well suited for researchers interested in human-robot interaction and social robotics. The robot features compliant force control throughout its body, and an omnidirectional base with a prismatic lift. The head is a 7 Degrees-of-Freedom (DOF) robotic active vision head with high resolution FireWire cameras in each eye, integrated DSP controllers, and zero backlash Harmonic Drive gearheads in the neck. Also, an ASUS Xtion Live Pro RGB-D camera is mounted in its chest, level with the shoulder actuators. It provides a 640x480 RGB image at up to 30 Hz with a matched depth image in the range of 0.5 m to 3.5 m. In order to achieve thermal measurement capabilities, an USB-powered Optris PI 640 thermal camera was mounted on the robot. It is located 25 cm below, 7 cm to the right and two cm behind of the RGB-D camera on a tilt-adjustable mount. At 46 mm wide, 56 mm tall and 90 mm long with its lens, the Optris PI 640 thermal camera has an optical resolution of 640x480 pixels and a spectral range from 7.5 to 13 µm, being able to measure temperatures ranging from $-20°$C to $900°$C at a frame rate of 32 Hz.

III. Experimental design

A. Experimental Conditions

In order to investigate the impact on temperature changes across people's faces according to the stress induced by the robot's behavior, participants were exposed to three conditions in a 2x2x2 between participants study. The conditions in the experiment were:

- **Height:** The height of the robot could vary between **Tall** (1.59 m) and **Short** (1.17 m), both measured at the eye level of the robot;
- **Gaze:** The gaze of the robot was either **Direct**, where the robot was looking directly at the person at all times, or **Averted**, where the robot alternated between looking away and at the card in the person's hand when detected;
- **Distance:** The robot was in the **Personal** space (0.6-0.7 m) or **Social** space (1.2-1.3 m) of the person, measured from the front of the person's face to the front of the chest of the robot.

B. Hypotheses

We formulated two hypotheses for this study:

- **H1:** The rise in temperature in the nose area will be greater when the robot is Tall, has a Direct gaze, and/or enters the Personal space of the person;
- **H2:** People are more likely to rate the interaction as stressful when the robot is Tall, has a Direct gaze, and/or enters their Personal space.

With these hypotheses, the goal is to verify if it is possible with the Meka robot to elicit stress in a participant, and if a variation of the temperature, possibly caused by the stress, can be measured with the thermal camera.

C. Scenario and Experimental Setup

Fig. 2. The experiment setup, showing the robot in its Short configuration, at a Social distance, and with a Direct gaze.

For the sake of this experiment, we designed a card game. The game consisted of the robot asking general geography questions (e.g., "Can you show me the card with the flag of Finland?" or "I would like to see the card with the map of Canada"). The total number of questions asked was 10. Questions came in 5 categories (Flags, Animals, Scenery, Capitals, and Map Outlines) and two difficulty levels.

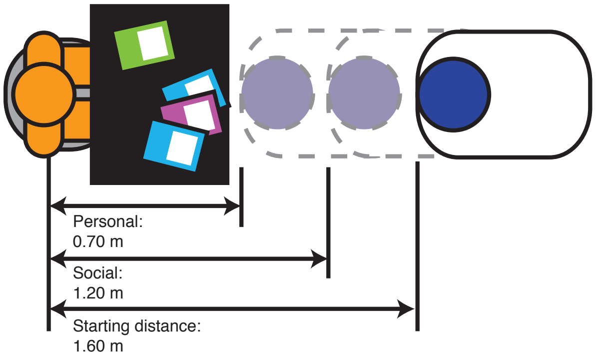


Fig. 3. A diagram of the different interaction distance with the participant in orange and the Meka in white and blue.

Prior to the experiment, participants were asked to refrain from consuming vasoactive substances (e.g., nicotine, caffeine, alcohol) for at least two hours before their interaction with the robot in order to obtain reliable measurements. After entering the experimentation room, the participant sat at a computer located near the entrance and signed an informed consent form. Then, the participant answered a short questionnaire regarding his/her demographic profile (i.e., gender, age, etc.). Additionally, the participants were asked to fill a BigFive Personality Traits questionnaire [13]. Next, the experimenter directed the participant to the experimentation area and introduced both the robot and the game. The participant was asked to sit on a chair with adjustable height, which the experimenter used to ensure that the participant's face was visible in both the frame of the RGB-D camera and the frame of the thermal camera for the length of the experiment. The experimenter then explained that the participant would play a cognitive game with the robot based on general geography knowledge. The game involved 40 cards distributed on a small table in front of the participant. The dimensions of the table were 50 cm in depth, 75 cm in width and 75 cm in height. The table was positioned between the person and the robot and did not interfere in the field of view of the thermal camera. After the general rules of the game were explained, the experimenter reminded the participant to look as much as possible at the robot, and avoid occluding his or her face with the cards when showing them to the robot. The experimenter then left the experimentation room, and the participant was lead to believe to be alone with the robot. However, a second, hidden experimenter monitored the experiment and could interrupt the interaction if any problem occurred.

The interaction started with the robot located at 80 cm from the edge of the table, or approximately 160 cm away from the participant's face, depending on his or her posture. This corresponded to the Social distance according to the proxemics as defined by Hall in [14]. The robot gave a 30 seconds speech welcoming the person, introducing itself and repeating the rules of the game. After the introduction, the robot started moving closer to the person, stopping at one of two distances selected for the interaction: Personal or Social. Figure 2 shows the experimental setup for the Social distance, and Figure 3 presents the three distances used during the experiment.

The participant had only one try per question, but the task was open-ended so there was no global time limit. The participant answered the question by showing a card with the answer on one side and a QR-code on the other side. An example of a card used in the game can be seen in Fig. 4.

Fig. 4. Example of a card used in the quiz.

The card had to be held with the QR-code side toward the robot. After decoding the card, the robot answered either "The answer is correct" or "The answer is wrong" and went to the next question. If the participant did not give an answer in less than 15 seconds, the robot repeated the question in case it was forgotten or not understood correctly. Additionally, the robot informed the participant of its performance after question 3 and 6 by saying a sentence such as "You answered correctly 2 out of 6 questions". For each participant, the set of questions was randomized with an equal number of questions in each category and difficulty level so as to avoid the choice of questions and ordering having an impact on the stress of the person. After the last question, the robot thanked the participant and instructed him or her to fill a questionnaire on their perception of the interaction. Each answer was on a 7-point Likert scale ranging from "Not at all" to "A lot". For this paper, the relevant questions were:

- Q1. Was the robot stressful?
- Q2. Did the distance at which it was standing made you feel intimidated by the robot?
- Q3. How did you feel when the robot started moving towards you? (stressed, scared, surprised)

IV. System Architecture

Figure 6 illustrates the main modules in the architecture that were developed. During the interaction with the robot, the goal is to evaluate the evolution in temperature of the person's face. Therefore, the face represents the region of interest (ROI) in the thermal image. Detection and tracking of the person's face in the images taken from the RGB-D sensor is achieved with the Viola-Jones algorithm [15], a method relying on Haar feature-based cascade classifiers. The OpenCV implementation of the algorithm as well as its pre-trained classifiers were used. In order to improve the speed of the detection algorithm, face detection is first performed in a local window corresponding to the previously detected face, as it is assumed that the person would not move significantly between two image frames. Nevertheless, if a face cannot be detected in the local window, the detection will be carried in the entire image. Moreover, since the detection algorithm is computationally expensive and to further increase the overall detection speed, detection was performed only once every five frames. To implement this, a buffer of twelve frames is maintained and interpolation is applied between successful face detections. The length of the buffer was empirically selected as a compromise between latency for real-time temperature monitoring and the capability to reconstruct long sequences of missing face detections. This approach is useful in maintaining a constant frame rate, 15 fps, for the detected face and therefore a constant sampling rate for the temperature values. After a face is detected in the image frame, its corresponding depth is extracted in order to obtain the real-world coordinates of the person's face. The depth is computed as the average value of the face rectangle in the corresponding registered depth frame. These coordinates are further used for both human detection when controlling the mobile base and the gaze of the robot as well as for face mapping between the RGB and thermal camera.

In order to compute the geometrical relationship between the two cameras, stereo calibration was performed using an appropriately characterized chessboard pattern of black and white squares. As the thermal camera only captures temperature, it cannot distinguish the chessboard pattern on a standard single sheet of paper. To be visible to both the thermal and RGB-D camera, the pattern was constructed in two parts. First, squares holes were cut in a white piece of cardboard paper. The cardboard was heated under a incandescent lamp for a few minutes, and then joined with a black sheet of paper that was covered with a thin layer of cold water. Therefore, depending on the camera, the white cardboard appeared as both white and warm, and the visible parts of the wet sheet as black and cold. Figure 5 shows the chessboard pattern simultaneously visible in both the RGB and thermal image.

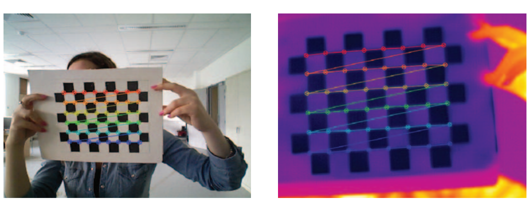


Fig. 5. Camera calibration process.

Two calibration settings were used depending on the relative height of the person and the robot to avoid missing parts of the face in the thermal image. Thereafter, the rotation matrix and translation vector between the two cameras resulting from the calibration process were used to map the face detected in the RGB image to the thermal image, obtaining the ROI. The face considered for further analysis represents the center 60% width and the full height of the mapped face. The temperature of the nasal and peri-nasal region was computed as the average value over a rectangle representing the middle 30% of the width and 25% of the height of the detected face.

In order to enable the robot to autonomously position itself at different distances towards the person, a control system for the mobile base was developed. The control system was built on a hierarchy of behaviors, where the teleoperation behavior controlled through a joystick by the experimenter had the highest priority to allow him to intervene in case the circumstances required it. The second, lower-priority behavior was used to move the robot to the Personal or Social distance from the person. Proportional velocity control using the current distance to the target was applied until the desired distance was reached and the person's face was aligned with the front of the robot.

Capability to decode and localize QR-Codes printed on the cards used for the geography knowledge game was also implemented. Decoding was performed with the ZBar library[1] on the color frames from the RGB-D camera. Moreover, the real-world coordinates of the detected code were computed in the same manner as the ones for the detected face. In case the card was detected outside of the depth range of the RGB-D camera, which can happen if the robot is too close from the person, a fixed depth of 0.50 m was used. As these coordinates were used to control the gaze of the robot, this depth was empirically determined to appear as if the robot was correctly looking at the card from the point of view of the participant.

V. Experimental Results

16 people (14 male, 2 female, average age of 27.5 years) accepted to participate in the experiment. To measure variations in the temperature of the participants, temperature data from the beginning of the first question to the end of the conclusion speech of the robot was selected. To remove outliers such as a card briefly covering the person's face (but not enough to prevent face detection) or false face detection, two techniques were applied. First, temperature readings over and under certain thresholds were discarded. These thresholds were manually selected for each participant to avoid discarding actual data, as face temperature can vary from person to person. For instance, for a face temperature varying between 30°C to 32°C, threshold were set to 25°C and 35°C, eliminating most false measurements. Then, to remove short but steep variations leading to these false measurements, the data was re-sampled with linear interpolation at a rate of 10 Hz to replace empty sequences, and a moving average filter with a window of 20 samples (2 seconds) was applied. The difference between immediate samples was then calculated, and any sample presenting an absolute difference of more than 0.1°C per second was discarded. Finally, least-square regression was applied to fit a linear model on the remaining sparse data. Figure 7 shows a typical filtered dataset from an experiment with the result of the linear regression.

We performed a 2x2x2 analysis of variance (ANOVA) on the temperature change rate as estimated by the linear

[1] https://github.com/ZBar/ZBar

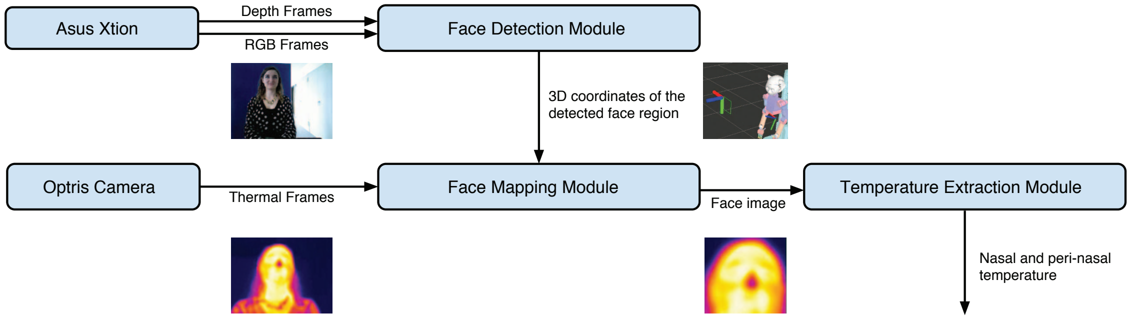


Fig. 6. Communication structure between the main modules.

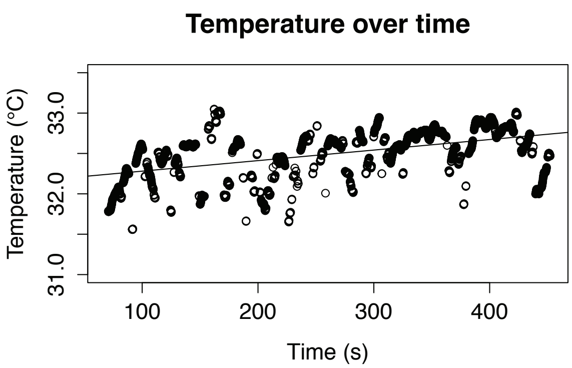


Fig. 7. A typical dataset of filtered temperature readings. Note that the data starts at 68 s, which is after the introduction speech and robot motion. The linear regression was fitted with $r^2 = 0.2324$, $p < 2 \times 10{-}16$, resulting in a rise 0.013 °C/s.

model. As can be seen from results shown in Table I, there was a statistically significant main effect of the Distance of the robot, F(1,8) = 7.214, p = 0.0277. Furthermore, a statistically significant interaction between the Distance and Gaze of the robot was also observed, F(1, 8) = 5.963, p = 0.0404. However, the results also suggest that neither the Gaze nor Height of the robot by themselves were sufficient to provoke a temperature change (p = 0.1501 and p = 0.2277, respectively).

Figure 8(a) shows the temperature variation depending on Distance. A temperature rise can be observed when the robot is in the Personal space of the participant (mean variation: 4.04×10^{-3} °C/s, σ^2=3.53×10^{-3}). Moreover, the temperature shows a slight decrease when the robot is in the Social space instead (mean variation: -1.13×10^{-4} °C/s, σ^2=4.26×10^{-3}). These results suggest that hypothesis H1 is validated with respect to the distance of the interaction.

To evaluate if the robot's behavior and pose had an impact on stress, we analyzed the results from our questionnaire. Table II shows the results from a 2x2x2 ANOVA on the response to Q1. A statistically significant main effect was again observed for Distance, F(1,8) = 10.714, p = 0.0113. Figure 8(c) shows the responses to Q1. It can be observed that participants were more likely to rate the robot as stressful ($\overline{x}$ = 4.13, σ^2 = 1.13) when it entered their Personal space.

TABLE I
2x2x2 ANOVA ON TEMPERATURE VARIATION

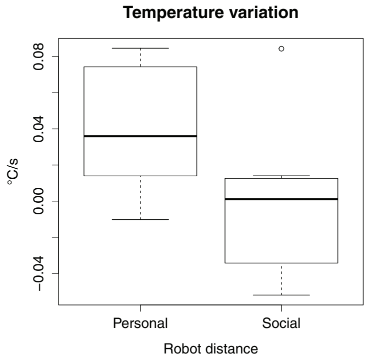
Factor	Df	Mean Sq	F	p
Height	1	1.629e-05	1.707	0.2277
Gaze	1	2.419e-05	2.534	0.1501
Distance	1	6.888e-05	7.214	**0.0277**
Height:Gaze	1	1.033e-05	1.082	0.3286
Height:Distance	1	2.440e-06	0.255	0.6270
Gaze:Distance	1	5.693e-05	5.963	**0.0404**
Height:Gaze:Distance	1	2.796e-05	2.928	0.1254
Residuals	8	7.638e-05	9.550e-06	

This suggests that hypothesis H2 is also supported for the distance of the robot. No additional effect or interaction was found regarding the factors of the experiment. However, a 2x2x2 ANOVA on the answer to the part of Q3 covering surprise reveals that Height has a statistically significant main effect with F(1,8)=7.143, p=0.0282, while Distance or Gaze have none. This suggests that the Meka robot in its Tall configuration surprises more with its motion than in the Short configuration. Figure 8(b) shows the distribution of the answers.

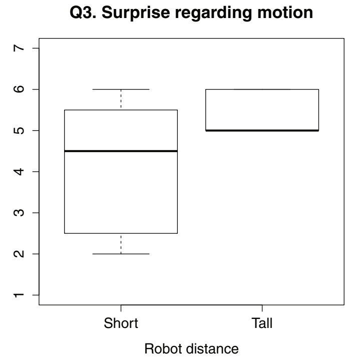
TABLE II
2x2x2 ANOVA ON ANSWERS TO Q1

Factor	Df	Mean Sq	F	p
Height	1	0.063	0.048	0.8327
Distance	1	14.062	10.714	**0.0113**
Gaze	1	3.063	2.333	0.1651
Height:Distance	1	0.563	0.429	0.5311
Height:Gaze	1	0.063	0.048	0.8327
Dist:Gaze	1	0.562	0.429	0.5311
Height:Distance:Gaze	1	1.563	1.190	0.3070
Residuals	8	10.500	1.312	

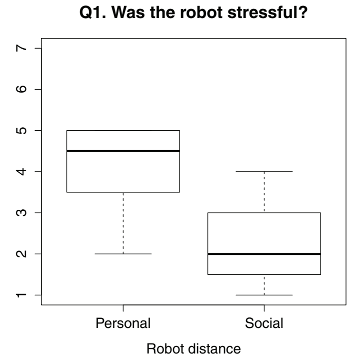
To investigate if participants with whom we observed a temperature rise were more likely to rate the robot as stressful, we performed a one-factor ANOVA on the answers to Q1. While the result shows a marginally statistical significance with F(1,14)=4.253, p=0.0582, we cannot conclude that there is a direct link between rating the robot as stressful and a rise in face temperature at that moment. However, some people might rate their level of stress differently than others. For instance, a single factor ANOVA reveals that participants rated as extroverted by the BigFive personality questionnaire (8 out of 16) were also more likely to rate the motion of the



(a) Temperature variation according to the distance of the robot.



(b) Answers to Q3 according to the distance of the robot.



(c) Answers to Q1 according to the distance of the robot.

robot as surprising, with F(1,14)=10.07, p=0.007.

The surprise caused by the motion of the robot might also explain the decrease in face temperature observed in the Social distance condition. The surprise could have caused a brief, initial rise in temperature at the beginning of the experiment, followed by a decrease to a more normal level when the robot stops outside of the personal space of the participant. A future within-subjects study aimed at measuring an increase in stress or surprise would give a more complete explanation.

VI. Conclusion

This paper presents the design of a thermal imaging-based system mounted on an humanoid robot performing contact-free measurement of temperature variations across people's faces in a stressful interaction. A statistically significant variation of temperature in the nasal and peri-nasal region of the face was observed to be greater when the robot was standing in the personal space of the participant. Results also show that there is a statistically significant interaction between the distance of the robot and the direction of its gaze on this temperature variation. While only a marginally significant correlation could be found between the level of stress rated by the people in the post experiment questionnaire and the increase in temperature, we believe that a more thorough analysis based on personality dimensions of the participants can help us get an insight on what exactly caused temperature variations in the personal, direct condition.

From the results of this work, we posit that thermal imaging sensors can be successfully employed in embodying robots with physiological sensing capabilities in order to allow them to become aware of their effect on people, learn about their preferences, and build a reactive behavior. The presented approach could be applied online, with parameters such as temperature thresholds automatically estimated at the beginning of an interaction with a person. From this perspective, it would be interesting to investigate in future work if a reactive change in the behavior of the robot meant to reduce the stressing effect on the person could reveal a decrease in temperature of the participant to its baseline.

VII. Acknowledgements

We would like to thank the ERASMUS+ program that funded the internship of Mihaela Sorostinean. This work was funded and done in the context of the EU Horizon2020 ENRICHME project, Grant Agreement No: 643691.

References

[1] K. H. Kim, S. Bang, and S. Kim, "Emotion recognition system using short-term monitoring of physiological signals," *Medical and biological engineering and computing*, vol. 42, no. 3, pp. 419–427, 2004.

[2] J. T. Cacioppo and L. G. Tassinary, "Inferring psychological significance from physiological signals." *American Psychologist*, vol. 45, no. 1, p. 16, 1990.

[3] D. Keltner and P. Ekman, "The psychophysiology of emotion," in *Handbook of Emotions*, M. Lewis and J. Haviland-Jones, Eds., 2000, pp. 236–249.

[4] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 10, pp. 1175–1191, 2001.

[5] J. Healey, "Physiological sensing of emotion," *The Oxford Handbook of Affective Computing*, p. 204, 2014.

[6] P. Philippot, R. S. Feldman, *et al.*, *The regulation of emotion*. Psychology Press, 2004.

[7] S. Ioannou, V. Gallese, and A. Merla, "Thermal infrared imaging in psychophysiology: Potentialities and limits," *Psychophysiology*, vol. 51, no. 10, pp. 951–963, 2014.

[8] C. L. Lisetti and F. Nasoz, "Using noninvasive wearable computers to recognize human emotions from physiological signals," *EURASIP journal on applied signal processing*, vol. 2004, pp. 1672–1687, 2004.

[9] V. Sandulescu, S. Andrews, D. Ellis, N. Bellotto, and O. M. Mozos, "Stress detection using wearable physiological sensors," in *Artificial Computation in Biology and Medicine*. Springer, 2015, pp. 526–532.

[10] C. Puri, L. Olson, I. Pavlidis, J. Levine, and J. Starren, "Stresscam: non-contact measurement of users' emotional states through thermal imaging," in *CHI'05 extended abstracts on Human factors in computing systems*. ACM, 2005, pp. 1725–1728.

[11] Z. Zhu, P. Tsiamyrtzis, and I. Pavlidis, "Forehead thermal signature extraction in lie detection," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE, 2007, pp. 243–246.

[12] S. Ioannou, P. Morris, H. Mercer, M. Baker, V. Gallese, and V. Reddy, "Proximity and gaze influences facial temperature: a thermal infrared imaging study," *Frontiers in psychology*, vol. 5, 2014.

[13] L. R. Goldberg, "An alternative" description of personality": the big-five factor structure." *Journal of personality and social psychology*, vol. 59, no. 6, p. 1216, 1990.

[14] E. T. Hall, *The Hidden Dimension*. Doubleday & Co, 1966.

[15] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, no. 2, pp. 137–154, 2004.