2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Discriminating Stress From Cognitive Load Using Contactless Thermal Imaging Devices

Federica Gioia[1], Maria Antonietta Pascali[2], Alberto Greco[1], Sara Colantonio[2], Enzo Pasquale Scilingo[1]

Abstract— **This study proposes long wave infrared technology as a contactless alternative to wearable devices for stress detection. To this aim, we studied the change in facial thermal distribution of 17 healthy subjects in response to different stressors (Stroop Test, Mental Arithmetic Test). During the experimental sessions the electrodermal activity (EDA) and the facial thermal response were simultaneously recorded from each subject. It is well known from the literature that EDA can be considered a reliable marker for the psychological state variation, therefore we used it as a reference signal to validate the thermal results. Statistical analysis was performed to evaluate significant differences in the thermal features between stress and non-stress conditions, as well as between stress and cognitive load. Our results are in line with the outcomes of previous studies and show significant differences in the temperature trends over time between stress and resting conditions. As a new result, we found that the mean temperature changes of some less studied facial regions, e.g., the right cheek, are able not only to significantly discriminate between resting and stressful conditions, but also allow to recognize the typology of stressors. This outcome not only directs future studies to consider the thermal patterns of less explored facial regions as possible correlates of mental states, but more importantly it suggests that different psychological states could potentially be discriminated in a contactless manner.**

I. INTRODUCTION

There are several sources of stress, such as physical, psychological, emotional, social, etc. A stressful response can be driven by either a positive episode or a negative event. Moreover, the stressful state can be temporary and acute or, if prolonged over time, can become chronic. The physiological and psychological reaction to the different sources of stress is not the same [1], [2], therefore it is important to distinguish the different kinds of stress triggers to better face them.

In clinical practice, stress is assessed through questionnaires and self-reported scales. The subjective nature of these tools may pose several limitations since participants may give socially desirable answers and mask their true psychological state [3] or they might not be aware of their perceived stress level. Furthermore, psychometric tests are usually not stressor-specific.

To overcome this ambiguity, many studies have tried to use peripheral autonomous nervous system (ANS) correlates to objectively assess the subject's psychophysiological state. In fact, the perception of a stressful stimulus from different sources modulates the ANS activity. Particularly, the sympathetic nervous system (SNS) branch dominates over the parasympathetic one, and triggers a cascade of body responses providing a rapid adaptation [4]. The overall effect is an increase in attention and arousal [5], measurable through specific variations of physiological signs such as the increase of heart rate, the blood redistribution to muscles and brain, the deepening of breaths, the pupil dilation, and the increase of sweat glands activity.

The latest researches focused on the development of multiparametric monitoring systems based on wearable devices. Many of these devices are able to monitor the cardiac and respiratory activity, the electrodermal activity (EDA) and the movement [6], [7]. The EDA measures changes in the skin conductance due to the SNS-driven activation of sweat glands in response to psychological stimuli. The EDA signal alone has already been shown to be reliable in the discrimination between stress and non-stress conditions [1] as well as between stress and cognitive load [8]. Moreover, EDA signal has been often used as a reference signal to compare the performance of other signals [9]. However, even though wearable devices are non-invasive, the application of sensors and electrodes on the body can still alter the subject's natural behavior. Furthermore, some of these devices may be uncomfortable, e.g. the EDA systems, to acquire high quality signals, usually requires the use of electrodes attached to the fingers, making activities of daily living difficult.

In addition to wearable devices, a recent challenge is the use of contactless devices. Among these, the thermal camera is an ideal candidate for a multivariate approach through remote sensing, without getting in physical contact with the subject or obstructing his movements. Thermal imaging measures skin temperature, which is determined by the passage of blood in the subcutaneous vessels and by sweat secretion [10]. Both of these phenomena are sympathetically driven [11]. For this reason, thermal imaging is currently taking hold in the psychophysiological field to monitor emotional engagement [12] and to recognize emotions [13]. Previous studies have been able to extract several physiological correlates from thermal signals, such as breathing rate[14], cardiac pulse [15], and cutaneous blood perfusion [16]. Nevertheless, thermal patterns have not been fully understood yet.

Stress elicited by Stroop Task was observed to induce an enhancement of blood flow in the forehead, supraorbital and frontal vessels [17]. So far, the most reliable thermal feature to discriminate between positive and negative valence states is the nose tip change in temperature, which has been successfully used in a real-time classification algorithm to improve social robots interaction with children [12]. How-

[1] Department of Information Engineering & Bioengineering and Robotics Research Centre E. Piaggio, School of Engineering, University of Pisa, Italy.
[2] Institute of Information Science and Technologies, National Council of Research, Pisa, Italy; Pisa, Italy
This work is also partially supported by the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence)
This work is also partially supported by the COST Action The GoodBrother, Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living, CA 19121 https://goodbrother.eu/
Corresponding author email: federica.gioia@phd.unipi.it

ever, to the best of our knowledge, none of the previous studies have distinguished different sources of stress using the thermal signal.

This preliminary study aims to propose thermal imaging technology as a contactless alternative to wearable devices for stress detection. To this aim, we used thermal imaging to discriminate mental stress from cognitive load, investigating, at the same time, which are the most informative facial regions.

We also analyzed EDA as a gold-standard to evaluate the effectiveness of stressful protocols and as a term of comparison for the results obtained through the thermal camera.

II. Materials and Methods

For this study, we recruited 17 healthy volunteers (7 females, mean age = 33, std = 7.25). Each subject signed an informed consent and filled in the Beck Anxiety Inventory (BAI) scale, to measure their level of clinical anxiety [18]. The experiment was approved by the local ethics committee. Subjects suffering from psychiatric or neurological disorders or currently assuming medicines were excluded from the study. Moreover, the subjects were asked to avoid the usage of vasomotor substances (i.e. coffee), moisturizing cream, and make up.

The experiment was comprised of the following sessions: Rest (R1), Stroop test (S), Rest (R2), Mental Arithmetic Task (MA) and Rest (R3) (see Figure 1). All of the resting sessions lasted 5 minutes, while the tasks lasted 3 minutes. The whole protocol was driven by an Android application

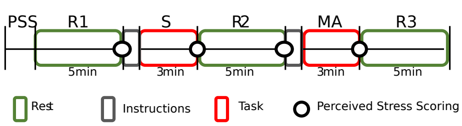

Fig. 1. Experimental Timeline

developed ad hoc for the study and installed on a tablet. The purpose of the App was to minimize the influence of the experimenter on the subjects and to prevent the effect of the speech activity on the EDA signal [19]. Before the start of the experiment, each subject filled in the Perceived Stress Scale (PSS) questionnaire, to assess the subject's sensitivity to stress [20]. The first stressful task was a computerized and paced Stroop Test, which required quick resolution of two incongruous stimuli. The subject had two seconds to press the button corresponding to the tint of the displayed word, that was inconsistent with the meaning. After any mistake or missed answer, a buzzer alerted the subject, and a counter, indicating the number of subsequent right answers, would turn back to zero as a motivational stressor. The second task was a Mental Arithmetic Test (MA) designed to increase the cognitive load of the subject, following the study reported in [2]. Each volunteer had to subtract 7 in series starting from 1022 and type the answer on the tablet. Therefore, the MA required the use of short-term memory and the computation of mathematical operations. In case of any mistake, a popup message indicated that he was required to start over again. During this task, the subject was kept aware of the remaining time with a progressive bar on top of the screen. During both tasks, a clock was ticking in the background to mark the passage of time. Before and after each task, the subjects reported their level of stress choosing a value from 0 (not at all) to 10 (very stressed). Throughout the whole experiment the EDA and the thermal response were recorded from all subjects. The EDA was measured using a shimmer GSR + Unit sensor, with the two electrodes attached to two fingers of the non-dominant hand. The thermal responses were measured using a FLIR A65sc infrared camera with a focal length of 13mm. This camera has spectral range of 7.5–13μm (LWIR), resolution of 640x512 pixel, thermal sensitivity $<0.05°$C and streaming rate 7.5 Hz. In this experiment, a multimodal technology was used, combining thermal and visible information, to obtain a more robust detection and tracking of the facial regions of interest (ROIs). Each participant completed the experiment in a controlled temperature and humidity room, and was always at about 60 cm distance from the camera.

A. Thermal Signal Processing

In this study, we selected 14 facial ROIs, as showed in Figure 2, some of them typically found in the literature and some less common, such as cheeks, nasal septum and chin. Forehead and chin were divided in left, right and central area to achieve a greater degree of spatial detail. The thermal signal was extracted from each ROI as the mean temperature of each ROI over time. The centers of the ROIs were empirically identified through interceptions of straight lines passing through specific landmarks automatically detected by the Yuval Nirkin algorithm [21]. In order to refer the

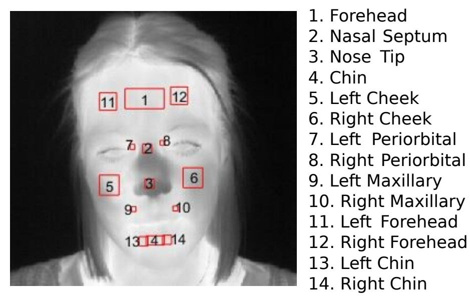

Fig. 2. Facial ROIs

facial landmarks, detected in the RGB image, to the IR coordinate system, all of the paired frames were registered computing the affine-2D transformation matrix matching the fiducial paired points in the two frames. In order to normalize through the subjects the dimension of each ROI, the face of each subject was automatically segmented in the first thermogram to obtain the size of the face; the ROI sizes were set proportionally: for example, the size of each cheek ROI was the 1.5% of the face size. The centers of the ROI were tracked through each experimental phase, using the Matlab PointTracker object based on Kanade-Lucas-Tomasi (KLT), pure translational feature-tracking algorithm [22]. In this way, we were able to follow the movements of the subjects. The thermal signals were smoothed by a moving median filter to remove the high frequency noise before performing the baseline correction. In order to remove the outliers, the normalized thermal signals were considered, and the values exceeding the threshold of 3σ were replaced by the nearest value. Afterwards, we implemented a baseline correction algorithm in order to adjust the strong thermal fluctuation due to the automatic non-uniformity correction (NUC), performed automatically by the camera, which could hide or modify the sought natural physiological response. This correction exploited the derivative of each thermal signal to detect the abrupt changes and update a piecewise constant correction signal. Thus, the obtained correction signal was then subtracted from the original signal, obtaining the corrected final thermal signal. From each thermal signal, we extracted four features: within each non overlapping time window of 20 sec we computed the mean and the standard deviation of both the signal and its derivative.

B. EDA Signal Processing

The EDA signal can be decomposed into a tonic and a phasic component. The former is a slow-varying component, with spectrum below 0.05 Hz, which represents the EDA baseline; the latter is the fast and event related component. Since our protocol administered long-lasting stimuli, we focused on the tonic component, because they can be measured on an ongoing basis over relatively long periods of time [23], and due to its proven stress correlation [1]. Indeed, we are not interested in specific responses to discrete stimuli, instead we intent to monitor the overall alterations in the subject's level of arousal due to the ongoing task over the whole three minute session. Thus, the EDA signal was normalized applying the z-score and decomposed using the cvxEDA algorithm into its tonic and phasic components [24], [25]. From the tonic component, we computed the mean within non-overlapped time windows of 20sec.

C. Statistical Analysis

The EDA and thermal features extracted within the non-overlapped 20-s windows were averaged within each session. For the resting state sessions, we considered only the last three minutes of signal, when subjects were more relaxed. We carried out two levels of analysis as follows:

- *Stress/Non-Stress analysis:* comparison between each task and their preceding resting state (*R1/S*, *R2/MA*);
- *Stress/Cognitive Load analysis:* comparison between the two tasks (*S/MA*).

In both the analyses, we studied the differences in the perceived self-assessed stress scores and the features extracted from the EDA and thermal signals. The comparisons were performed by means of a Wilcoxon signed-rank test with Bonferroni corrections.

III. Experimental Results

BAI questionnaire results showed that none of the recruited subjects presented a pathological state of anxiety (anxiety levels: 13 minimal, 4 moderate, 0 severe); PSS scale showed that the majority of the subjects were incline to moderately feel stress (6 low, 8 moderate, 3 high).

The self-reported stress levels during the experiment showed statistically significant difference between both the MA and S tasks and their preceding resting state (p value<0.01). However, they were not significant in the comparison between MA and S.

As we expected, the EDA showed a significant difference in the mean of its tonic component both in the stress/non-stress analysis and between stress and cognitive load (p value<0.01).

Regarding the thermal features, the mean of the derivative of some specific ROIs discriminated both S and MA from the resting state but not the two stressful tasks (see Table1). Indeed, in agreement with previous studies, the slope of the thermal signal on the nose tip was positive during the resting state and negative during the stressful tasks resulting in a significant p-value (<0.01). Likewise, the nasal septum, the left side of the forehead, and the left maxillary area showed the same trend as the above described nose tip (p value<0.05). Furthermore, the mean of the derivative resulted significant in the comparison between cognitive load and the preceding resting state (p value<0.05) in the following regions: right maxillary, right and left periorbital, chin and right chin.

On the other hand, the mean temperature resulted significant mostly in the comparison between stress and cognitive load (considering the two cheeks and the left forehead, p value <0.05), and in the comparison between cognitive load and preceding resting state (in the right cheek, left side of forehead and nose tip, p value<0.01). Consequently, the mean temperature of the right cheek was able to discriminate all the three experimental conditions.

As concerns the standard deviation of both the thermal signal and its derivative, we did not find any relevant significance.

TABLE I
Results of the statistical comparisons by means of Wilcoxon signed-rank tests - p values

ROI	Feature	R1 / S	R2 / MA	S / MA
Forehead	T mean	-	-	-
	D mean	-	-	-
Nasal Septum	T mean	-	-	-
	D mean	0.0222	0.0055	-
Nose Tip	T mean	-	0.0011	-
	D mean	0.0007	0.0019	-
Chin	T mean	-	-	-
	D mean	-	0.0130	-
L Cheek	T mean	-	-	0.0008
	D mean	-	-	-
R Cheek	T mean	0.0035	0.0022	0.0222
	D mean	-	-	-
L Periorbital	T mean	-	-	-
	D mean	-	0.0064	-
R Periorbital	T mean	-	-	-
	D mean	-	0.0149	-
L Maxil	T mean	-	-	-
	D mean	0.0019	0.0041	-
R Maxil	T mean	-	-	
	D mean	-	0.0130	-
L Forehead	T mean	-	0.0469	0.0222
	D mean	0.0469	0.0416	-
R Forehead	T mean	-	-	-
	D mean	-	-	-
L Chin	T mean	-	-	-
	D mean	-	-	-
R Chin	T mean	-	-	-
	D mean	-	0.0287	-

Note 1: The symbol "-" in the table means p-value>0.05
Note 2: T mean = mean temperature;
D mean = mean of the derivative of the thermal signal.

IV. Discussions and Conclusions

In this study, we assessed the discriminative power between stress and non-stress, and between stress and cognitive load of the thermal signals extracted from different facial regions. We used the EDA signal as a ground-truth for our study to evaluate the effectiveness of our protocol and changes in the psychophysiological state of the subjects. As expected, the EDA tonic activity was able to assess not only a significant difference between resting sessions and tasks, but also between tasks. On the other hand, the analysis of the self-assessed scores of stress perception reinforced the need of an objective physiologically-based method to differentiate the responses to distinct stressors. In fact, such scores did not result significantly different after the two tasks, suggesting that we may perceive them as equally stressful. However, this result is in contradiction with the EDA ones, which indicate a different SNS activity and, consequently, a different of physiological stress level during the two stressful tasks.

To sum up our findings, the analysis of facial thermal imaging showed that: (i) the mean of the derivative signal, is effective to discriminate between stress and non-stress in specific sites (e.g., nose tip or left maxillary); and there are no significant differences between different stressors; (ii) the mean temperature of less-studied regions (i.e. the cheeks and the left forehead) show significant differences between the Stroop color and cognitive load tasks.

In more detail, the nose tip confirmed the results of the current scientific literature, giving us positive feedback on the robustness and repeatability of our analysis. Furthermore, our results suggest that the nasal septum can be considered as an alternative region to the nose tip, as it shows the same dynamics but it is easier to localize and track. It is worthwhile noting that the mean temperature of an unusually considered region as the right cheek can significantly discriminate all the three conditions, candidating itself as the most informative ROI in a stress recognition task. On a speculative level, given the similarity between the right cheek and the EDA tonic statistical results, we could hypothesize a temperature dynamics mostly related to the sweat secretion effect instead of the vascularization one (contrary to what typically happen at the nose tip level).

A limit of this finding could be the non-randomization of the tasks. However, the two tasks were separated by five minutes of resting state in which we expected the subjects to recover and reach the same psychological state as the beginning.

Future works will explore even more facial regions of smaller sizes, allowing a facial tessellation to reach a higher spatial resolution of the findings. In doing so, we will be able to explore also a potential lateralization of the facial thermal patterns. Moreover, we will test the ability of thermal technology to discriminate between further stressors of different nature, physical and psychological, and between different emotional states.

References

[1] A. Greco, G. Valenza, J. Lazaro, J. M. Garzon-Rey, J. Aguilo, C. De-la Camara, R. Bailon, and E. P. Scilingo, "Acute stress state classification based on electrodermal activity modeling," *IEEE Transactions on Affective Computing*, no. 01, pp. 1–1, 2021.

[2] M. Mestanik, A. Mestanikova, Z. Visnovcova, A. Calkovska, and I. Tonhajzerova, "Cardiovascular sympathetic arousal in response to different mental stressors," *Physiological research*, vol. 64, p. S585, 2015.

[3] D. L. Paulhus and D. B. Reid, "Enhancement and denial in socially desirable responding." *Journal of Personality and Social Psychology*, vol. 60, no. 2, p. 307, 1991.

[4] L. D. Godoy, M. T. Rossignoli, P. Delfino-Pereira, N. Garcia-Cairasco, and E. H. de Lima Umeoka, "A comprehensive overview on stress neurobiology: basic concepts and clinical implications," *Frontiers in Behavioral Neuroscience*, vol. 12, p. 127, 2018.

[5] J. G. Betts, P. DeSaix, E. Johnson, J. E. Johnson, O. Korol, D. H. Kruse, B. Poe, J. A. Wise, K. A. Young *et al.*, "Anatomy and physiology," 2014.

[6] Y. S. Can, N. Chalabianloo, D. Ekiz, and C. Ersoy, "Continuous stress detection using wearable sensors in real life: Algorithmic programming contest case study," *Sensors*, vol. 19, no. 8, p. 1849, 2019.

[7] J. Marín-Morales, J. L. Higuera-Trujillo, A. Greco, J. Guixeres, C. Llinares, E. P. Scilingo, M. Alcañiz, and G. Valenza, "Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors," *Scientific reports*, vol. 8, no. 1, pp. 1–15, 2018.

[8] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable eda device," *IEEE Transactions on information technology in biomedicine*, vol. 14, no. 2, pp. 410–417, 2009.

[9] K. Peternel, M. Pogačnik, R. Tavčar, and A. Kos, "A presence-based context-aware chronic stress recognition system," *Sensors*, vol. 12, no. 11, pp. 15 888–15 906, 2012.

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

[11] W. Boucsein, *Electrodermal activity*. Springer Science & Business Media, 2012.

[12] C. Filippini, E. Spadolini, D. Cardone, and A. Merla, "Thermal imaging based affective computing for educational robot," in *Multidisciplinary Digital Publishing Institute Proceedings*, vol. 27, no. 1, 2019, p. 27.

[13] I. A. Cruz-Albarran, J. P. Benitez-Rangel, R. A. Osornio-Rios, and L. A. Morales-Hernandez, "Human emotions detection based on a smart-thermal system of thermographic images," *Infrared Physics & Technology*, vol. 81, pp. 250–261, 2017.

[14] J. N. Murthy, J. Van Jaarsveld, J. Fei, I. Pavlidis, R. I. Harrykissoon, J. F. Lucke, S. Faiz, and R. J. Castriotta, "Thermal infrared imaging: a novel method to monitor airflow during polysomnography," *Sleep*, vol. 32, no. 11, pp. 1521–1527, 2009.

[15] M. Garbey, N. Sun, A. Merla, and I. Pavlidis, "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery," *IEEE transactions on Biomedical Engineering*, vol. 54, no. 8, pp. 1418–1426, 2007.

[16] A. Merla, L. Di Donato, G. Romani, M. Proietti, and F. Salsano, "Comparison of thermal infrared and laser doppler imaging in the assessment of cutaneous tissue perfusion in scleroderma patients and healthy controls," *International Journal of immunopathology and pharmacology*, vol. 21, no. 3, pp. 679–686, 2008.

[17] 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*, 2005, pp. 1725–1728.

[18] A. T. Beck, N. Epstein, G. Brown, and R. A. Steer, "An inventory for measuring clinical anxiety: psychometric properties." *Journal of consulting and clinical psychology*, vol. 56, no. 6, p. 893, 1988.

[19] A. Greco, C. Marzi, A. Lanata, E. P. Scilingo, and N. Vanello, "Combining electrodermal activity and speech analysis towards a more accurate emotion recognition system," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 229–232.

[20] S. Cohen, "Perceived stress in a probability sample of the united states." 1988.

[21] Y. Nirkin, I. Masi, A. T. Tuan, T. Hassner, and G. Medioni, "On face segmentation, face swapping, and face perception," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. IEEE, 2018, pp. 98–105.

[22] B. D. Lucas, T. Kanade *et al.*, "An iterative image registration technique with an application to stereo vision," 1981.

[23] M. E. Dawson, A. M. Schell, and D. L. Filion, "The electrodermal system." 2017.

[24] A. Greco, G. Valenza, A. Lanata, E. P. Scilingo, and L. Citi, "cvxeda: A convex optimization approach to electrodermal activity processing," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 797–804, 2015.

[25] A. Greco, A. Lanata, G. Valenza, E. P. Scilingo, and L. Citi, "Electrodermal activity processing: A convex optimization approach," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2014, pp. 2290–2293.