An Unobtrusive Measurement Method for Assessing Physiological Response in Physical Human–Robot Interaction

Blaž Jakopin, Matjaž Mihelj, *Member, IEEE*, and Marko Munih, *Member, IEEE*

Abstract—The objective of this work was to develop and validate a novel unobtrusive method for measuring person's physiological response with a low-cost integrated sensory system for use in a physical control task. Two different sensory handles were designed (cylindrical and hemispherical shape) and used in a physical human–robot control task. Twenty-three participants underwent a measurement session with both handles, performing four different tasks for each handle. Two basic task conditions were permuted: physical load (high/low) and task dynamics (high/low). Electrocardiogram, photoplethysmogram, electrodermal activity, and peripheral skin temperature signals were recorded by sensory handles and a reference high-accuracy biosignal amplifier to determine the raw signal correlation between the measurement systems. Additionally, several standardized physiological parameters were calculated and discussed for both systems. Results of raw signal correlation showed a high correlation between the reference measurement system and the sensory handles. Pearson's correlation coefficients were above 0.8 for most of the physiological signals in all task conditions. Some effect of physical load and high task dynamics was registered. In terms of signal quality, the hemispherical design outperformed the cylindrical design. Correlation results show that the proposed system correlates well with the reference system for all tasks. In terms of optimal design for signal quality and comfort, hemispherical handle shape is more appropriate. Unobtrusive nature and short setup time of such a method deems it appropriate for home use, monitoring, and research.

Index Terms—Affective engineering, haptic interaction, human–machine interaction, physiological sensors, unobtrusive sensors.

I. INTRODUCTION

IN THE past few decades, quantitative measurement of persons's affective state and emotion reactivity has become a major topic of research for a broad area of applications, ranging from rehabilitation robotics in human–robot interaction to user experience testing in human–computer interaction. Affective state is a representation of subject's psychophysiological state along two principal dimensions: valence and arousal [1]. This representation could provide valuable information on how a person is motivated, entertained, engaged, and satisfied with a certain task or product.

In human–robot interaction applications, motor rehabilitation has progressed from simple passive mechanisms for physical exercise with physical therapist present at all times to semiautonomous robotic interfaces that can offer personalized therapy for each patient [2], [3] or objectively asses the patient's limb mobility and biomechanical parameters [4]. These are often combined with virtual environments to make the exercise more engaging and to increase the persons's motivation [5], [6]. This type of motor rehabilitation is especially effective with poststroke patients, where motivation and engagement in physical exercise is inherently low. In an effort to improve the success of rehabilitation, several attempts have been made to incorporate a quantitative measure of persons's affective state, that can, together with their motor performance parameters, provide different control strategies of the physical interaction task by adjusting the task difficulty [7]–[9].

Previous research has shown that estimation of person's psychological state is possible by measuring physiological parameters, such as heart rate (HR), skin conductance, peripheral skin temperature, and respiration [10], [11], and adapting the parameters of the physical control task based on the measured physiological response [12]. Similar research was performed by Badesa *et al.* [13], [14], using multisensory data to adaptively change the complexity of the virtual reality task. Enhanced human–robot interaction experience is reported by Guerrero *et al.* [15] using a biocybernetic closed-loop controller to adapt the robot assistance. Similarly, Morales *et al.* [16] presented a new concept of a multimodal assistive robotic system to address the growing trend of patient-tailored assistance. In an effort to improve the healthcare services in the future, Swangnetr and Kaber [17] proposed an algorithm for emotional classification in patient–robot cooperation, using physiological data acquired using standard bioamplifiers. A thorough overview of the affective engineering field has recently been done by Balters and Steinert [18], where applications and physiological measurement methods were discussed. Most of the aforementioned research was based on physiological data analysis, acquired by high-tech and expensive medical equipment. Although this type of equipment is noninvasive, it is usually not user-friendly and does not enable fast setup. Previous research by Dijkers *et al.*

[19] found that many therapists may stop using devices if the setup takes more than 5 min. Fast setup time and ease of use, together with the desired device accuracy, should therefore be considered.

A possible solution for an unobtrusive measurement of person's physiological response can be realized by placing sensors at the grasping point off the haptic robot tip. This way, no additional setup is needed. The person can interface with the virtual environment through robot interaction and have the physiological parameters simultaneously measured by the integrated sensors at the handle. This approach could enable researchers for faster experiments and open the door to home-based measurement solutions, not requiring additional personnel for task setup.

Such thinking is well aligned with notable increase in low-cost, wearable, and unobtrusive sensory systems for healthcare in order to tackle the problem of aging population, ubiquity of chronic diseases, or simply the trend of healthier lifestyle [20], [21]. These sensors often measure some of the person's physiological signals and combine them with biomechanics. After a comprehensive review of commercial low-cost measurement devices, the authors determined that a custom-embedded solution was necessary, since none of the present off-the-shelf solutions can be applicable. Heuer *et al.* [22] measured physiological signals through direct contact electrodes and sensors embedded in the steering wheel of a vehicle. Similarly, physiological sensors were integrated in a smart wheelchair for an unobtrusive measurement during patient use [23]. In order to assess the person's psychological state during physical rehabilitation task without applying additional sensors and electrodes, a simple combination of grasping and measurement is a natural choice.

Our focus is on the measurement of physiological parameters, namely HR and its variability [using electrocardiography and photoplethysmography (PPG)], electrodermal activity (EDA), and peripheral skin temperature. These parameters are mostly linked with physical activity and emotional arousal [24] through the sympathetic part of the autonomous nervous system and have been intensely used in the literature, showing significant changes with regard to different cognitive tasks [25]. Physiological responses can also be influenced by physical activity. In the past, efforts have been made to separate cognitive and physical response to determine which physiological measure is most reliable in motor rehabilitation, where physical workload is also present [26].

This paper starts with the development of an embedded low-cost sensory system and its application in two types of handle configurations having cylindrical and hemispherical shapes. These are placed at the human arm to haptic robot interaction point. The signals from this low-cost embedded system are processed in a same way as the normative data that were gathered from the established commercial system. The goal and motivation for this work is the comparison of signals carrying physiological features. This paper is focused on the reliability of signals and presence of noise and artifacts in the case of a low-cost unobtrusive system versus an established commercial system.
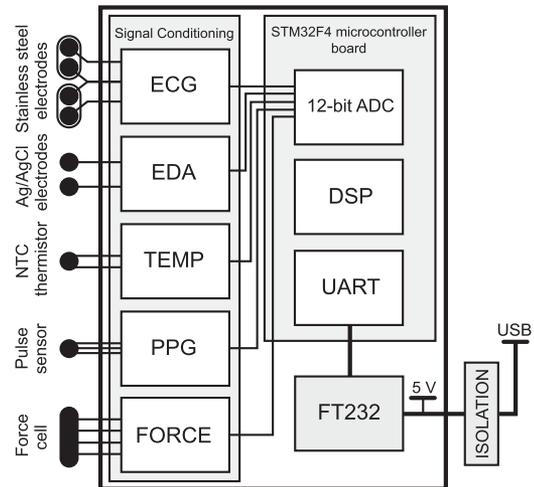

Fig. 1.   Overview of the system architecture. Separate analog front ends are locally regulated to ensure stable operating voltage. Isolation for this prototype is provided externally, and power is supplied by the USB port.

II. SYSTEM DESIGN

A. Sensor Architecture

A full sensory system architecture is illustrated in Fig. 1. Various analog conditioning circuits are used to acquire person's physiological signals.

The single-lead electrocardiogram (ECG) monitor circuit was designed as proposed by Richard and Chan [27] for measuring HR and heart rate variability (HRV). For the purpose of unobtrusive measurement, dry stainless steel electrodes also commonly found in ECG monitors for home use were used.

For measuring EDA, a circuit as proposed by Poh *et al.* [28] was designed, using a nonlinear feedback automatic bias control with low-power operational amplifiers (TLC274 by Texas Instruments), also utilizing dry Ag/AgCl electrodes.

A conditioning circuit for peripheral skin temperature measurement was designed using a negative temperature coefficient (NTC) glass thermistor (62S3KF354G by Betatherm), connected in a Wheatstone bridge and amplified using an instrument amplifier (AD8223 by Analog Devices).

Since ECG monitors require bimanual measurement [29], we upgraded the system with a second HR measurement using PPG. An off-the-shelf PPG sensor (Pulse Sensor by World Famous Electronics) was used in order to enable for unimanual measurement and to test the appropriateness of the PPG signal to serve as a standalone sensor for measuring HR and HRV. Such a setup could provide the possibility for fusing ECG and PPG signals for a more robust HR estimation in the future. Next possible upside of adding the PPG sensor can be the additional clinical parameters that can be extracted using both ECG and PPG data (e.g., pulse transit time [30]). All signal conditioning is integrated on the PPG sensor printed circuit board that, with the adequate gain, already produces an analog waveform for digital conversion.

Additionally, a force cell conditioning circuit was designed to enable the use of grasping force measurement. A low-power

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

JAKOPIN *et al.*: UNOBTRUSIVE MEASUREMENT METHOD FOR ASSESSING PHYSIOLOGICAL RESPONSE IN PHYSICAL HUMAN–ROBOT 3

Fig. 2. Handle designs with integrated physiological sensors. (a) Cylindrical shape (c-handle). (b) Hemispherical shape (s-handle).
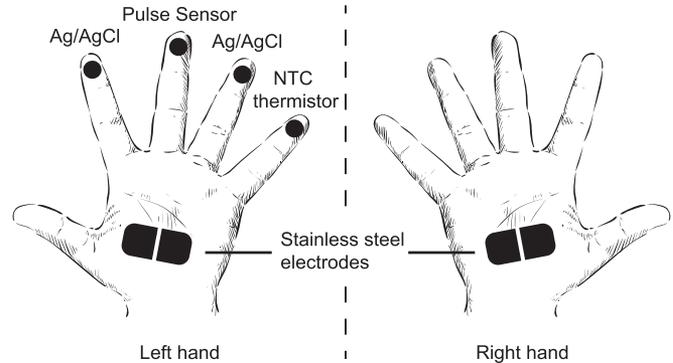

Fig. 3. Predicted measurement locations for both left and right palms at time of interaction with the robot. Left (active) hand is interacting with the robot and is simultaneously being measured for physiological response at the handle. Right (passive) hand is resting on the static handle, populated only by the ECG electrodes.

high-accuracy instrumentation amplifier with a precision reference and differential input amplification (INA125 by Texas Instruments) was used for load cell signal amplification. All conditioning circuits had local voltage references to produce a stable and noise-free analog voltage.

A microcontroller board (STM32F4 Discovery by ST Microelectronics) was used for analog-to-digital (A–D) conversion, signal processing, and communication. Power is drawn directly from the universal serial bus (USB) port on the local computer and then locally regulated on the board. A–D conversion was made by an integrated 12-bit analog-to-digital converter (ADC). With Nyquist frequency set at $f_{\text{Nyquist}} = 100$ Hz, real-time operation is possible, and the most relevant bandwidths of physiological signals are covered in accordance with [31]. QRS complex of the ECG has the highest frequency content of all measured signals and is located in the range of 10 Hz. We oversampled the analog signals with a sampling frequency $f_{\text{sampling}} = 4^n \cdot f_{\text{Nyquist}}$, where $n = 4$, to increase the resolution to 16 bits [32]. After the data have been processed, it is communicated via universal asynchronous receiver/transmitter (UART), through UART to USB data transfer interface (FT232 by Future Technology Devices International), and finally passed to the local computer through the use of external isolation (USB-to-USB isolator by Baaske Medical). Isolation is needed for safety and for reduction of measurement interference.

B. Handle Design

To enable an inconspicuous and unobtrusive measurement of physiological parameters, sensors have to be placed at the point of the haptic interaction between the robot and the human; thus, the electrodes and sensors were integrated in the robot handle. Two mountable robot handles with different shapes were developed for this purpose: one cylindrically shaped (c-handle) and one hemispherically shaped (s-handle) handle, as presented in Fig. 2.

After a preliminary study of comfort in the rehabilitation task within the laboratory staff, s-handle showed much better results regarding comfort than the c-handle. However, the cylindrical shape is more universal and can be easily used in different orientations, for different rehabilitation tasks, enabling grasping force measurements.

Electrodes and sensors were integrated to the handles as proposed in the literature [18] and in a way to be as intuitive and

inconspicuous as possible. The person should be able to grasp the handle with no special care, not needing to focus on the sensors and electrodes during task. Both handles were designed for the left hand to simulate reduced motor capabilities when experimenting with healthy subjects.

Predicted measurement locations are illustrated in Fig. 3. EDA measurement is made through Ag/AgCl electrodes positioned in such a way that distal phalanges of the second and fourth fingers cover the entire surface of the electrode. PPG measurement is made by covering the pulse sensor with the distal phalanges of the third finger. Peripheral skin temperature is measured at the distal phalanges of the fifth finger. Finally, stainless steel electrodes for measuring ECG are appropriately positioned on the handle, enabling contact with proximal/thenar palmar surface.

A second (static) handle was added to the system to ensure that the bimanual measurement of the ECG was possible, and additionally to provide a bimanual frame of reference for the person [33], during the physical control task. The second handle, which was designed for the right hand, is hemispherically shaped for best ergonomic fit and includes only the stainless steel electrodes.

C-handle was designed and 3-D printed in the Laboratory of robotics, while s-handle was hand made for best ergonomic fit out of expanded polystyrene.

III. MATERIALS AND METHODS

A. Hardware

Haptic Master robot (Moog FCS, The Netherlands) was used as the haptic robot interface for the motor rehabilitation task. The end-effector of the robot is equipped with a three-axis force sensor. The robot itself enables force-controlled movement in three degrees of freedom (DOFs). Handles were mounted to the force sensor at the robot end-effector (see Fig. 4).

A 2.4 m × 1.7 m screen was used to display the virtual rehabilitation task. Subjects sat at a distance of 3.2 m in front of the screen, with the robot positioned at their left-hand side.
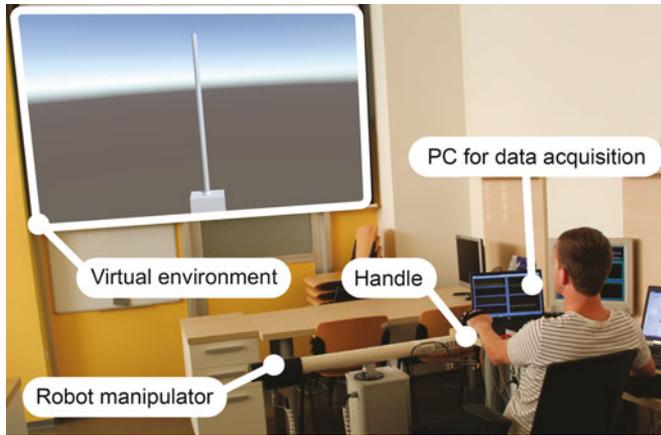
Fig. 4.   Subject performing a physical control task by balancing the inverted pendulum in a virtual environment.

A g.USBamp (g.tec Medical Engineering GmbH, Austria) amplifier was used as a "gold standard" reference physiological measurement system. ECG was recorded using four disposable pregelled electrodes, with one electrode at the left part of the chest, one at the right part of the chest, one at the right part of the abdomen, and a ground electrode applied to the upper right part of the back. EDA was measured with a g.GSR sensor, placed on the distal phalanges of the second and fourth fingers on the nonactive hand, to enable the same measurement conditions as on the active hand. Peripheral skin temperature was measured with a g.TEMP sensor, attached to the distal phalanges of the fifth finger. PPG was measured with a g.PULSE sensor, attached to the distal phalanges of the third finger. All signals were sampled at 100 Hz to enable the same conditions as with the proposed measurement system.

B. Participants

A total of 23 students and staff members (age 26.8 ± 6.9) of the University of Ljubljana participated in the study. Each participant performed a single measurement session, consisting of eight measurement blocks in total, four with each handle. Eighteen participants were males, and five were females. All participants were healthy with no physical or cognitive defects. One female person recording was discarded due to severe signal corruption.

C. Physical Control Task

A mathematical model of the inverted pendulum was implemented in a virtual environment as the physical control task [26]. The inverted pendulum is an inherently unstable system without control, so the participants had to balance the pendulum by applying a virtual force to the cart (robot end-effector). The amount of force needed and the pendulum dynamics were adjusted through the mathematical model by changing the parameters such as gravity, friction, damping, mass, and pole length. If participants did not manage to stabilize the pole, the pole was reset to the vertical position after the cart was brought back to

the middle of the screen. All robot movements and haptics were limited only to the x-axis (horizontal axis) of the robot to enable a 2-D rehabilitation task and to simplify the experiment.

In order to create different physical conditions in the physical control task, mass of the cart was changed between two values (mass at high physical load was set to five times higher in comparison to low physical load). Different dynamic conditions were created by changing the length of the pole between two values (high dynamics were set at one-third of the length of the pole at low dynamics). With high physical load, the robot produced larger reaction forces at the end-effector for the person to move the cart. With high dynamics, the pendulum reacted much faster to the cart movement, also falling much faster. In this condition, participants had to react faster in order to balance the pole. This way, two conditions were permuted to create four different tasks (a similar design, but for different purpose, was used in [26]):

> *Task 1 (T1):* low physical load and low dynamics;
> *Task 2 (T2):* low physical load and high dynamics;
> *Task 3 (T3):* high physical load and low dynamics;
> *Task 4 (T4):* high physical load and high dynamics.

D. Experiment Protocol

The experiment was conducted in a quiet environment of the laboratory, where participants were not disturbed by random noise and other stimuli. Only experiment supervisor and one participant were present in the room during the experiment.

Participants were seated in front of the robot, g.tec sensors were applied to their passive hand and the chest, and the experiment protocol was explained. Support for the active arm was provided by an armchair. The passive arm was resting on a soft comfortable material at approximately the same height as the active arm. After seating, subjects were encouraged to practice T1 only once, for up to 2 min to get the basic knowledge of the inverted pendulum task to provide for a steady and predictable physiological response. In addition, the short practice session allowed the experiment supervisor to comment and correct the palm–handle point of interaction for every subject in order to minimize the effect of motion artifacts during actual tasks.

After the initial practice session, subjects performed four blocks of measurement trials separately for both handle options. Each block consisted of a 3-min rest period and was followed by a 3-min task period. The rest period served as a baseline for physiological measurements. Tasks were selected randomly. After the first four blocks, the handle was replaced. During replacement, participants were allowed to rest briefly with reference sensors attached. Following the replacement, all four blocks of measurements were repeated.

E. Performance and Biomechanical Measures

Performance was evaluated with parameters, such as success rate, mechanical work, and mean frequency of position signals. Success rate is determined by counting how many times the pendulum has fallen during the task. Subjects with better balancing will have a lower count than others. Also, under the task

condition with low dynamics, subjects should, in general, have a lower count than at the high dynamic condition. Total mechanical work W_{total} is calculated as the sum of all work increments, calculated as a dot product of position s and force F signals from the robot

$$W = \int_c \vec{F} \cdot d\vec{s} \rightarrow W_{\text{total}} = \sum F \cdot s \cdot \cos\theta = \sum |F \cdot s|. \tag{1}$$

Since the task was designed for single degree of motion (horizontal movements), we are calculating forces and work using data only from horizontal movements. Total work should increase through the tasks, especially in tasks with greater physical load. Mean frequency of position signals was calculated to confirm the effect of the high dynamic condition and should increase when the high dynamic condition is active. Mean frequency f_{mean} was calculated using Welch's power spectral density estimate (as proposed in [34])

$$f_{\text{mean}} = \frac{\sum pxx \cdot f}{\sum pxx} \tag{2}$$

where pxx is the power spectral density, and f is the frequency vector.

F. Physiological Parameters

Physiological recordings were obtained from both handle and reference measurement systems for every subject. After the experiment, signals were processed offline for 3-min periods of both baseline and task, from which several standardized parameters were extracted for each period.

From ECG recordings, the *mean HR* was calculated as a mean value of time differences between consecutive RR peaks in the QRS complex. Additionally, two standardized HRV parameters [35], [36] were also extracted: the *standard deviation of successive NN intervals (SDNN)* and *the square root of the mean squared differences of successive NN intervals (RMSSD)*. The ECG processing algorithm is the same for both the bimanual measurement and the reference ECG recording. Since ECG signals are taken at various locations, morphology of the signals and timing is not similar enough to provide for raw correlation processing. For determining the correlation of the HRV information between systems, an *HRV signal* was produced by cubic spline interpolation of consecutive HR values.

Same parameters (*mean HR, SDNN*, and *RMSSD*) and same *HR signal* were extracted from the first derivative of the PPG signals.

EDA recording can be decomposed into two separate components: a low-frequency (tonic) component and a higher frequency (phasic) component. The tonic component describes the overall skin conductance over a longer period of time and is obtained by filtering the raw signal with a low-pass filter with a cutoff frequency of 0.05 Hz [37]. Similarly, the phasic component was obtained by high-pass filtering of the raw signal with a cutoff frequency of 0.05 Hz to observe the higher frequency fluctuations of skin conductance that are modulated on top of the slower tonic component. From the tonic component, the mean *skin conductance level (SCL)* parameter is extracted.

Skin conductance responses (SCRs) are a quantitative measure of skin conductance fluctuations in a period of time, classified as an amplitude increase of 0.05 μS and with a peak occurring within less than 5 s after the beginning of the increase. *SCR frequency* was calculated for the duration of baseline and task periods.

Final skin temperature was calculated as an average peripheral skin temperature of the last 2 s of each task period.

G. Motion Artifacts

An extensive study to quantify the effects of motion artifacts on signal quality was devised. First, sources of motion that influence the electrodes and sensors at the handle are determined. These are mostly the horizontal movements of the task, handle grasping, and the passive rotation of the handle. As we have designed our experiment with a redundancy of biomechanical sensors, we can model the artifacts arising from three-DOF sources of motion using the horizontal interaction force from the robot end-effector force cell, grasping force from the handle force cell, and angular velocity from the inertial sensor that was attached to the posterior side of the hand.

As proposed by Sweeney *et al.* [38], different Quality of Signal (QOS) metrics were produced for all three sources of motion. The proposed method utilizes a 1-s moving window that scans across the signal and compares all the sample values to a threshold value. Based on how many sample values are over the threshold, the algorithm will return QOS value between 0 and 1. Value 1 represents maximum motion artifact around that sample. Three new signals were produced using this method: QOS_{F_x} (horizontal force component contribution), $\text{QOS}_{F_{gr}}$ (grasping force component contribution), and QOS_{Ω_y} (angular velocity component contribution). Threshold values were set as suggested in the literature [38]. Threshold for horizontal force (THR_{F_x}) and angular velocity (THR_{Ω_y}) was calculated using three standard deviations of F_x and Ω_y from an average user. Grasping force threshold ($\text{THR}_{F_{gr}}$) was empirically set based on its influence on the signal quality. The threshold values were $\text{THR}_{F_x} = 2$ N, $\text{THR}_{F_{gr}} = 6$ N, and $\text{THR}_{\Omega_y} = 0.1$ rad/s. QOS samples are summed to produce a scalar representation of motion artifact presence for a specific session (QOSS—*quality of signal surface*) and normalized to the entire period of the session to produce a relative measure. QOSS $= 100\%$ would mean that the quantity in question was over the threshold for the entire duration of the session.

Together with the QOS signals that quantify the effect of motion, binary algorithm detection signals (DET) were produced, which quantify the algorithm detection capability and functioning. From ECG and PPG signals, we calculate HR values through a threshold algorithm that uses minimal amplitude and minimal peak distance thresholds. If consecutive HR values are within the average population, the signal will be set on 0; in other case, when detection is failing, it will return 1. For EDA and temperature signals, that are located in the lower frequency ranges, we examine the signals for fast rates of transition that could result only from motion artifacts. A moving window of ten samples and a threshold value of 0.25 μS is implemented for

EDA algorithm detection. For temperature, a 50-sample window with a threshold value of 0.03 K is used. Difference between maximum and minimum value of the window is calculated at every sample.

The positive detection percentage (PDP) parameter is calculated for every subject to produce a mean PDP value for all physiological signals. The number of events (*NOE*) parameter is calculated to describe the frequency of negative detection occurrences as the number of transitions from 0 to 1 in the algorithm detection signal.

In addition, QOS overlap (QOSO) was calculated between algorithm detection signals (DET) and QOS signals as a ratio between the surface of QOS included in (DET) signal and the entire QOS surface to determine how much of the QOS signal is contained within the period of the failed detection

$$\text{QOSO}[\%] = \frac{\sum_{i=1}^{n} \text{QOS}(i) \cdot \text{DET}(i)}{\sum_{i=1}^{n} \text{QOS}(i)} \cdot 100. \tag{3}$$

H. Data Analysis

Performance and biomechanical measures for both handles were compared between four tasks. One-way repeated measures analysis of variance (ANOVA) was used to determine the statistical significance of differences between tasks. The purpose of this step was to show how different task conditions changed during the experiment.

In the second step, similarities of physiological recordings for both handles were tested against the g.tec recordings using Pearson's correlation coefficients (PCCs). For ECG and PPG recordings, the calculated *HR signal* was used for analysis, since raw signals would not produce meaningful results. For EDA and temperature analysis, raw signals were used. Raw signals were filtered prior to analysis to reduce the effect of motion artefacts and noise: for EDA recordings, a low-pass Butterworth filter with a cutoff frequency of 3 Hz was used, while for the temperature, the cutoff frequency was set at 1 Hz. After filtering, signals were split into baseline and task segments, and PCCs were calculated. Additionally, significance of differences between task and baseline PCCs was determined using paired *t*-tests to analyze the effect of task on signal correlation.

For the final step of data analysis, we compared the calculated physiological parameters in task conditions to the same parameters in baseline conditions. Absolute values of parameters were used for comparison of baseline and task conditions using paired *t*-tests. Relative values for each parameter were calculated to show the similarity of the calculated parameters between the handles and the reference system. Relative values were calculated either by subtracting the baseline values from the task values or by additionally dividing the values with the baseline values to find the percentage of the difference. The purpose of this step was to show that by calculating physiological parameters, significant change in baseline-task parameters can be detected by both the proposed and reference systems for the same conditions.

All signals were processed by custom algorithms written in MATLAB (The MathWorks, Inc.). Sigma Plot (Systat Software, Inc.) was used for statistical analysis.
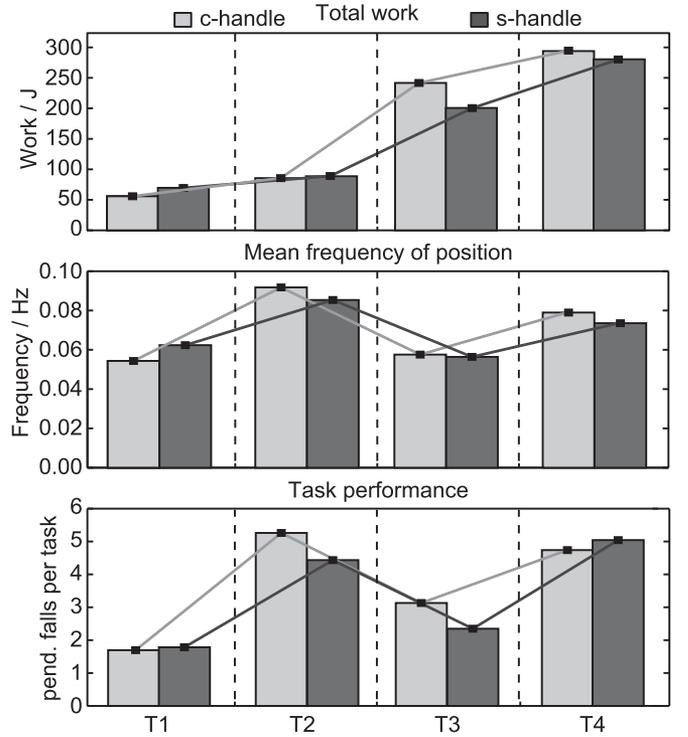

Fig. 5. Total work (top), mean frequency of position (middle), and task performance (bottom). Gray lines are connecting the mean values of both c-handle and s-handle measurements to better illustrate the effect of task on task performance and biomechanical parameters.

The threshold for statistical significance was set at $p = 0.05$. Statistical significance of difference was calculated by either one-way repeated measures ANOVA, followed by the Tukey posthoc test, or by paired t-test. The Kolmogorov–Smirnov test was used to test for normality. Whenever the normality test failed, ANOVA on ranks or signed rank test were used.

IV. RESULTS

A. Performance and Biomechanical Measures

For total work performed by subjects [see Fig. 5(top)], a significant difference was found ($p < 0.001$) among the tasks for both handles, and pairwise comparisons found significant difference ($p < 0.05$) between all low and high physical load conditions (T1–T3, T1–T4, T2–T3, and T2–T4) for both handles.

Mean frequency of the position signal [see Fig. 5(middle)] has shown a significant difference ($p < 0.001$) among the tasks for both handles, and pairwise comparisons revealed significant difference ($p < 0.05$) between low and high dynamic conditions (T1–T2, T3–T2, and T3–T4) for both handles. Additionally, significant difference ($p < 0.05$) was identified between T1 and T4 for c-handle, but not for s-handle. There was a significant difference found for task performance ($p < 0.001$) among the tasks for both handles, and pairwise comparisons found significant difference ($p < 0.05$) between high and low dynamic conditions (T1–T2 and T2–T3) for both handles. Additionally,
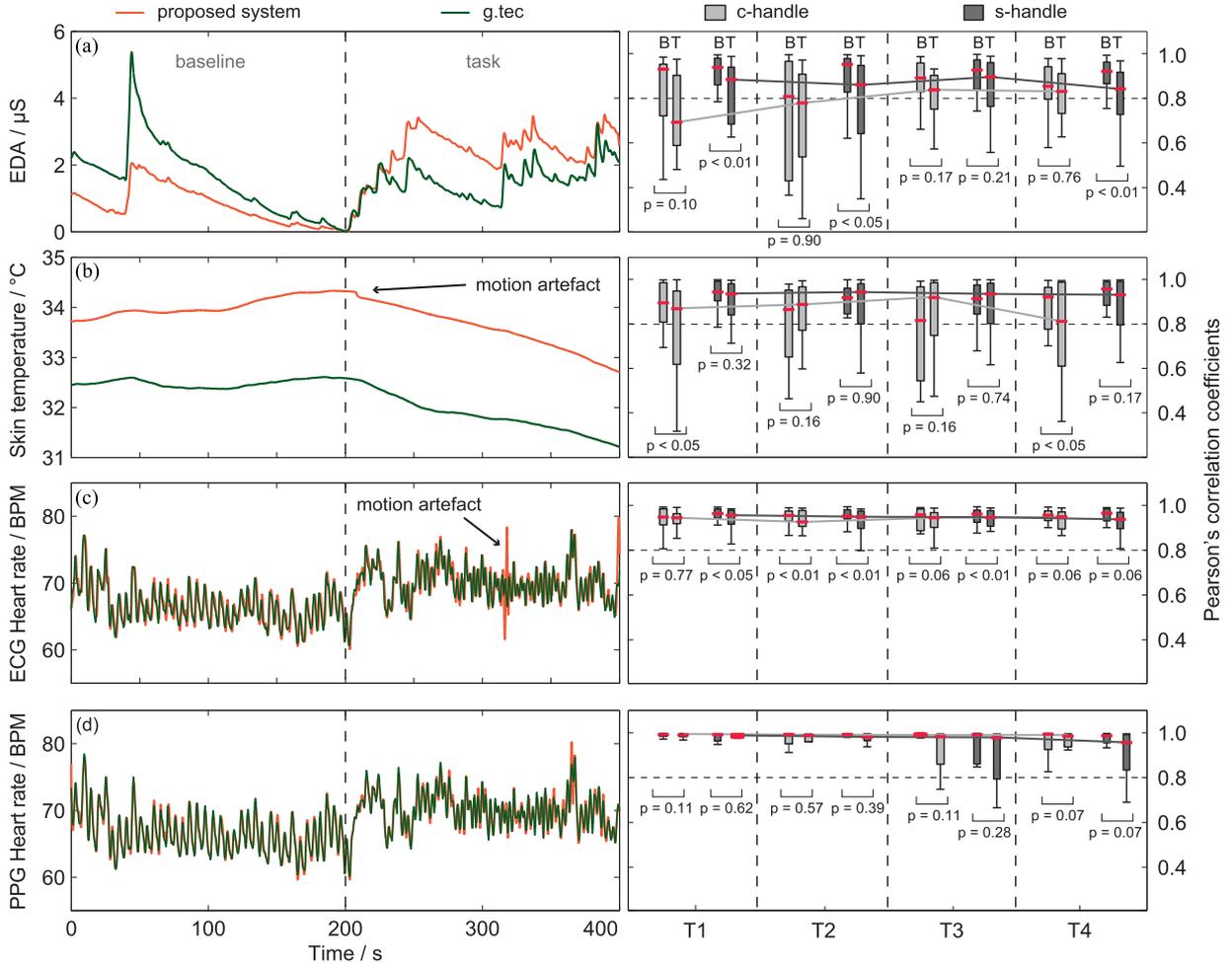
Fig. 6. Raw physiological waveforms (left) and PCCs for all tasks are represented as box plots (right). (a) EDA, (b) peripheral skin temperature, (c) ECG HR signal, and (d) PPG HR signal. Gray lines are connecting the median values of PCCs in tasks to better illustrate the correlation trend between different tasks. B and T are marking baseline and task period for each of the four tasks, respectively.

statistically significant difference ($p < 0.05$) was found between T3 and T4 for s-handle but not for c-handle. Mean values of task performance are given in the bottom graph of Fig. 5.

B. Signal Similarity

Results for signal similarity can be seen in Fig. 6. PCCs were calculated between g.tec and handle signals for both handles. For EDA signals [see Fig. 6(a)], median values of PCCs in baseline $\overline{r_b}$ are higher ($0.81 \leq \overline{r_b} \leq 0.93$ for c-handle and $0.92 \leq \overline{r_b} \leq 0.95$ for s-handle) than in task $\overline{r_t}$ periods ($0.69 \leq \overline{r_t} \leq 0.84$ for c-handle and $0.84 \leq \overline{r_t} \leq 0.9$ for s-handle). A significant baseline-task correlation reduction ($p < 0.05$) was found for T1, T2, and T4 of s-handle results; others were not significant.

For skin temperature [see Fig. 6(b)], median values of PCCs showed a reduction in correlation for T1 ($\overline{r_b} = 0.89$, $\overline{r_t} = 0.87$ for c-handle and $\overline{r_b} = 0.94$, $\overline{r_t} = 0.93$ for s-handle) and T4 ($\overline{r_b} = 0.92$, $\overline{r_t} = 0.81$ for c-handle and $\overline{r_b} = 0.96$, $\overline{r_t} = 0.93$ for s-handle). Increase in baseline-task correlation was found for T2 ($\overline{r_b} = 0.87$, $\overline{r_t} = 0.89$ for c-handle and $\overline{r_b} = 0.92$, $\overline{r_t} = 0.94$ for s-handle) and T3 ($\overline{r_b} = 0.82$, $\overline{r_t} = 0.92$ for c-handle and

$\overline{r_b} = 0.91$, $\overline{r_t} = 0.94$ for s-handle). A significant difference of baseline-task correlation ($p < 0.05$) was found for T1 and T4 of c-handle results; others were not significant.

For ECG_{HR} extracted signals [see Fig. 6(c)], median values of PCCs showed high correlation in both baseline and task periods ($0.95 \leq \overline{r_b} \leq 0.96$ for both handles, $0.93 \leq \overline{r_t} \leq 0.96$ for both handles). For both handles, a significant difference of baseline-task correlation ($p < 0.01$) was found for T2 ($\overline{r_b} = 0.95$ and $\overline{r_t} = 0.93$ for c-handle, and $\overline{r_b} = 0.95$ and $\overline{r_t} = 0.94$ for s-handle).

For PPG_{HR} extracted signals [see Fig. 6(d)], median values of PCCs showed high correlation in T1, T2, and T3 ($\overline{r_b} > 0.99$ and $\overline{r_t} > 0.98$ for all three tasks and both handles). A significant difference of baseline-task correlation ($p < 0.05$) was found for both handles in T4 ($\overline{r_b} = 0.99$ and $\overline{r_t} = 0.98$ for c-handle, and $\overline{r_b} = 0.99$ and $\overline{r_t} = 0.96$ for s-handle).

Fig. 7 presents differences between s-handle and c-handle correlation values in all conditions, showing correlation both in the baseline and in the task. S-handle showed better correlation for most of the signals through all four tasks, except for PPG_{HR} curve that showed higher correlation for c-handle for all tasks.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

8   IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS
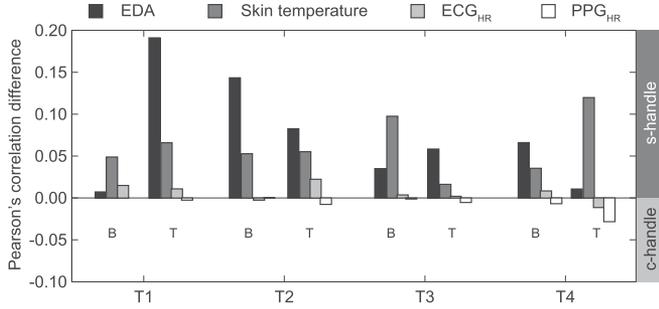

Fig. 7. Difference in PCCs for all physiological signals between s-handle and c-handle. Difference is calculated as $\overline{r}_{s-\text{handle}} - \overline{r}_{c-\text{handle}}$, where a positive value presents better s-handle correlation, and a negative value presents better c-handle correlation.

C. Baseline–Task Parameters

Mean relative values of the calculated physiological parameters are reported in Table III for c-handle and Table IV for s-handle. Values are reported for recordings of both the proposed system (handles) and the reference system (g.tec).

Significant differences in parameter values between the baseline and the task are bolded and additionally marked with asterisks. Results show no significant difference in mean HR_{ECG} and mean HR_{PPG} parameters for both integrated systems and the reference system. For HRV parameters, a significant difference was observed in all four tasks for SDNN_{ECG} and SDNN_{PPG} for both integrated and g.tec measurement systems. For $\text{RMSSD}_{\text{ECG}}$, a significant difference was observed for T2 and T4 for both systems, and $\text{RMSSD}_{\text{PPG}}$ was found significant for T1 and T4 for both systems. *Mean SCL* showed a significant difference in T3 and T4 for c-handle, but not for g.tec. *SCR frequency* was found significant for all four tasks for both c-handle and g.tec, except for T3, where the g.tec parameter is not found significant. *Final skin temperature* difference was found significant across all four tasks.

For s-handle, a significant difference was observed in all four tasks for SDNN_{ECG} and $\text{RMSSD}_{\text{ECG}}$. For SDNN_{PPG}, a significant difference was observed for all four tasks for both s-handle and g.tec, except for s-handle in T3. $\text{RMSSD}_{\text{PPG}}$ was found significant only for g.tec in T1 and T2. *Mean SCL* showed a significant difference in T1 and T2 for s-handle, but not for g.tec. *SCR frequency* was found significant for all four tasks for both s-handle and g.tec, except for T3, where the g.tec parameter is not significant. *Final temperature* difference was found significant, when high physical load was present (i.e., in T3 and T4).

D. Motion Artifacts

Mean values of QOSS are given in Table I. Largest artifact occurrence is found for $\text{QOSS}_{W_y} = 3.4\%$ at T2, followed by $\text{QOSS}_{F_x} = 1.58\%$ for T4. Largest NOE is found for EDA signals at T4, where the detection is failing with an average rate of 1.2 events per minute. High values of PDP are found for both EDA and temperature detection algorithms, where both values are over 99% positive detection within the entire session. A significant decrease is found only for PDP_{ppg} at tasks T3 and T4, where the value lowers to 90.1% and 91.7%.

TABLE I
MEAN VALUES OF QOSS, NOE, AND PDP CALCULATED FOR ALL FOUR TASKS

PARAMETER	TASK 1	TASK 2	TASK 3	TASK 4
QOSS_{F_x} [%]	0.00	0.00	0.06	1.58
$\text{QOSS}_{F_{gr}}$ [%]	0.00	0.36	0.01	0.3
QOSS_{W_y} [%]	0.49	3.4	0.76	1.23
NOE_{ppg} [ev/min]	0.37	0.4	0.60	0.58
NOE_{ecg} [ev/min]	0.43	0.55	0.53	0.52
NOE_{eda} [ev/min]	0.42	0.82	0.49	1.2
NOE_{temp} [ev/min]	0.42	0.73	0.43	0.55
PDP_{ppg} [%]	98.2	95.9	90.1	91.7
PDP_{ecg} [%]	98.1	97.2	96.4	96.6
PDP_{eda} [%]	99.6	99.7	99.4	99.1
PDP_{temp} [%]	99.4	99.6	99.7	99.9

TABLE II
MEAN VALUES OF QOSO BETWEEN ALGORITHM DETECTION SIGNALS (DET) AND QOS FOR T4

	DET_{ppg}	DET_{ecg}	DET_{eda}	DET_{temp}
QOSO_{F_x} [%]	48.6	15.3	4.7	4.29
$\text{QOSO}_{F_{gr}}$ [%]	20.3	16.2	8.7	5.98
QOSO_{W_y} [%]	24.6	11.8	4.2	3.92

Mean values of QOS and algorithm detection overlap are found in Table II, where DET_{ppg} appears to be most influenced by QOSO_{F_x}, with a mean overlap value of 48.6%. A similar effect on DET_{ECG} is shown by $\text{QOSO}_{F_x} = 15.3\%$ and $\text{QOSO}_{F_{gr}} = 16.2\%$. Largest DET_{eda} contribution is by $\text{QOSO}_{F_{gr}} = 8.7\%$.

V. DISCUSSION

A. Signal Similarity

Task conditions were changed to test for signal correlation between two integrated systems (handles) and reference system (g.tec). For EDA, median PCCs experienced a decrease, while transiting from the baseline to the task. Most of the values were above 0.8, except for the first two tasks in c-handle EDA measurement. Although higher values of correlation were calculated for s-handle in the first two tasks, both handles demonstrated a statistically significant drop in tasks for both low and high dynamics in the low physical load conditions (T1 and T2). This implies a noticeable effect of motion artifacts on EDA measurement, induced by more jerkier movements in conditions with low physical load. In high physical load conditions, no significant reduction in correlation was noted for both handles in low dynamics condition (T3), but s-handle showed a significant reduction of correlation in the high dynamic condition. This is to be expected as the high dynamic conditions induce higher frequency of arm and hand motion, which in terms produce more noticeable motion artifacts. Ag/AgCl surface electrodes can easily detect unwanted noise created by motion and signals can get corrupted by a minimal loss of the skin–electrode contact.

It should also be noted that comparison of EDA recordings by themselves is not a simple task, since same-site recording is not

TABLE III
MEAN RELATIVE VALUES OF CALCULATED PARAMETERS FOR C-HANDLE AND THE REFERENCE SYSTEM (G.TEC)

	TASK 1		TASK 2		TASK 3		TASK 4	
	C-HANDLE	G.TEC	C-HANDLE	G.TEC	C-HANDLE	G.TEC	C-HANDLE	G.TEC
mean HR_{ECG} [BPM]	-1.2 ± 2.3	-1.2 ± 2.2	-1 ± 3.2	-1 ± 3.2	-0.4 ± 2.4	-0.4 ± 2.3	-0.4 ± 2.7	-0.4 ± 2.7
$SDNN_{ECG}$ [%]	$-13 \pm 16^{**}$	$-13 \pm 17^{**}$	$-20 \pm 17^{**}$	$-22 \pm 16^{**}$	$-15 \pm 16^{*}$	$-15 \pm 16^{**}$	$-23 \pm 15^{**}$	$-25 \pm 15^{**}$
$RMSSD_{ECG}$ [%]	$-7.8 \pm 15^{*}$	-6.6 ± 14	$-10 \pm 19^{*}$	$-10 \pm 19^{*}$	-7.3 ± 2.9	-7.4 ± 19	$-9.4 \pm 22^{*}$	$-15 \pm 16^{**}$
mean HR_{PPG} [BPM]	-1.2 ± 2.3	-1.1 ± 2.2	-1.7 ± 2.7	-1.7 ± 2.7	-0.4 ± 2.5	-0.3 ± 2.4	-0.8 ± 3.3	-0.5 ± 2.9
$SDNN_{PPG}$ [%]	$-14 \pm 16^{**}$	$-14 \pm 18^{*}$	$-17 \pm 21^{**}$	$-19 \pm 17^{**}$	$-17 \pm 19^{*}$	$-17 \pm 15^{*}$	$-23 \pm 16^{**}$	$-26 \pm 16^{**}$
$RMSSD_{PPG}$ [%]	$-6.2 \pm 16^{*}$	$-7.3 \pm 15^{*}$	0.2 ± 29	-5.1 ± 17	-1.5 ± 34	$-12 \pm 19^{*}$	$-16 \pm 16^{**}$	$-17 \pm 18^{**}$
Mean SCL [μS]	0.6 ± 1.4	0.3 ± 1.7	0.9 ± 2.0	0.7 ± 1.6	$0.8 \pm 1.4^{*}$	-0.5 ± 3.0	$0.5 \pm 1.1^{*}$	0.5 ± 3.0
SCR frequency [%]	$96 \pm 136^{**}$	$68 \pm 91^{**}$	$119 \pm 178^{**}$	$40 \pm 53^{**}$	$126 \pm 241^{**}$	35 ± 103	$58 \pm 69^{**}$	$36 \pm 43^{**}$
Final temperature [K]	$-0.2 \pm 0.3^{*}$	$-0.3 \pm 0.3^{**}$	$-0.2 \pm 0.4^{*}$	$-0.3 \pm 0.3^{**}$	$-0.2 \pm 0.3^{**}$	$-0.3 \pm 0.3^{**}$	$-0.3 \pm 0.4^{*}$	$-0.3 \pm 0.4^{**}$

Statistically significant differences from baseline values are marked with asterisks: * for $p < 0.05$ and ** for $p < 0.01$.

possible due to crosstalk, repeated measures are not reliable due to habituation, and bilateral recording sites will also produce different responses [39].

B. Physiological Parameters

Temporal changes of skin temperature are slower than those of other physiological signals, so most of the higher frequency motion artifacts can be filtered away prior to processing. Thermistor time constant also plays an important role in temporal response, since a sensor with a larger time constant will not show higher frequency motion artifacts. Skin temperature generally increases as the first-order system step response from the moment of contact to reach the final value in baseline. With this in mind, an appropriate thermistor with lower time constant has been used to reduce the rise time. Once the final value is reached, skin temperature generally fluctuates around that value. When a subject is presented with stimuli (task), a significant drop in the skin temperature is induced as a result of increased perspiration. This response appears to improve the skin temperature correlation from the baseline to the task for the second and third tasks. A significant drop in correlation was registered in T1 and T4 for c-handle, which can be a consequence to an unergonomic handle shape and/or nonuniversal sensor placement.

For ECG measurement, high dynamics and low physical load condition showed most significant effect on signal correlation for both handles. This implies that ECG signal correlation is less affected in tasks with higher physical load, providing a firm grip on surface stainless steel electrodes. Since the overall correlation reduction in all tasks is minimal, ECG bimanual measurement is considered a viable option for HR measurement in physical control tasks. A different effect is observed in PPG measurement, where correlation is significantly reduced in high physical load conditions. This is due to PPG sensor measuring slight changes in light reflectance, which in static conditions translates to the pressure wave propagating through the arteries. When different external pressure is applied to the skin–sensor contact, this PPG measurement is distorted. Calculation of PCCs shows that physical load is the main cause of interference, since both T1 and T2 showed no major correlation reduction. Another aspect can also be the handle design and sensor integration.

While fingers can be isolated from movement in low physical load conditions, they have to take an active part in balancing and stabilizing the handle in high physical conditions.

Median values of PCCs mostly reside above the 0.8 line (see Fig. 6), except in c-handle EDA measurements for T1 and T2, which is promising, comparing these results to other literature [28], [40]. S-handle design outperformed the C-handle in all tasks and for all signals, except for the PPG measurements, where the cylindrically shaped handle showed higher correlation for both baseline and task measurements (see Fig. 7). Correlation results show that the proposed system correlates well with the reference system in both cases; in terms of the optimal design for signal quality and comfort, hemispherical shape is more appropriate.

In total, nine parameters were calculated for both handles (see Tables III and IV), most of which showed a significant change in the task at specific conditions, except for mean HR_{ECG} and mean HR_{PPG}. Mean relative values of the HR were very similar between the proposed systems and g.tec for all tasks; thus, we can assume that the HR parameter can be used in both high dynamics and high physical load conditions. No deviation between PPG and ECG measurement was noted for the HR parameter, which implies that this parameter can be determined accurately by both measurement methods for all task conditions.

Values of HRV time-domain parameters tend to decrease in high arousal situations. For SDNN parameter, a significant decrease was found for all task conditions for both handles. While comparing mean relative values of $SDNN_{PPG}$ between both the proposed and reference systems, results show no major error; thus, we concluded that bimanual ECG measurement performs well in terms of measuring SDNN parameter, and also that this parameter showed meaningful information regarding different workload in the task. Results also showed a significant decrease in the $SDNN_{PPG}$ parameter for all four tasks for c-handle, and mean relative values were also comparable between the proposed and reference systems. However, s-handle results in the third task failed to show significance in comparison to the reference system, and also larger error in mean relative values is evident for T4. This is similar to what the raw signal correlation data have shown. The trend in mean relative values is still very similar to that of the reference system, so we concluded that

TABLE IV

MEAN RELATIVE VALUES OF CALCULATED PARAMETERS FOR S-HANDLE AND THE REFERENCE SYSTEM (G.TEC)

	TASK 1		TASK 2		TASK 3		TASK 4	
	S-HANDLE	G.TEC	S-HANDLE	G.TEC	S-HANDLE	G.TEC	S-HANDLE	G.TEC
mean HR_{ECG} [BPM]	-0.1 ± 2.7	-0.1 ± 2.7	$0.7 \pm 0.7\%$	$0.9 \pm 0.9\%$	$0.1 \pm 0.1\%$	$0.2 \pm 0.2\%$	$0.1 \pm 0.1\%$	$0.1 \pm 0.1\%$
$SDNN_{ECG}$ [%]	$\mathbf{-14 \pm 17}^{**}$	$\mathbf{-14 \pm 17}^{**}$	$\mathbf{-21 \pm 21}^{**}$	$\mathbf{-21 \pm 22}^{**}$	$\mathbf{-14 \pm 20}^{*}$	$\mathbf{-13 \pm 20}^{*}$	$\mathbf{-12 \pm 18}^{*}$	$\mathbf{-13 \pm 19}^{*}$
$RMSSD_{ECG}$ [%]	$\mathbf{-10 \pm 16}^{**}$	$\mathbf{-10 \pm 16}^{**}$	$\mathbf{-15 \pm 18}^{**}$	$\mathbf{-17 \pm 18}^{**}$	$\mathbf{-10 \pm 21}^{*}$	$\mathbf{-12 \pm 17}^{*}$	$\mathbf{-6.6 \pm 14}^{*}$	$\mathbf{-8.1 \pm 16}^{*}$
mean HR_{PPG} [BPM]	-0.3 ± 2.7	-0.4 ± 2.7	$0.3 \pm 0.3\%$	$0.5 \pm 0.5\%$	0.1 ± 3	$0.2 \pm 0.2\%$	$0.1 \pm 0.1\%$	$0.2 \pm 0.2\%$
$SDNN_{PPG}$ [%]	$\mathbf{-12 \pm 17}^{*}$	$\mathbf{-14 \pm 17}^{**}$	$\mathbf{-17 \pm 21}^{**}$	$\mathbf{-20 \pm 21}^{**}$	-8 ± 29	$\mathbf{-10 \pm 21}^{*}$	$\mathbf{-12 \pm 20}^{*}$	$\mathbf{-17 \pm 17}^{**}$
$RMSSD_{PPG}$ [%]	-7.1 ± 19	$\mathbf{-11 \pm 15}^{*}$	-10 ± 17	$\mathbf{-17 \pm 21}^{*}$	1.9 ± 53	-13 ± 18	2.8 ± 43	-7.1 ± 21
Mean SCL [μS]	$\mathbf{0.5 \pm 1}^{*}$	$0.1 \pm 0.1\%$	$\mathbf{1.4 \pm 1.7}^{**}$	$0.5 \pm 0.5\%$	$0.1 \pm 0.1\%$	-0.1 ± 18	0.8 ± 2	$0.1 \pm 0.1\%$
SCR frequency [%]	$\mathbf{149 \pm 225}^{**}$	$\mathbf{42 \pm 63}^{**}$	$\mathbf{176 \pm 272}^{**}$	$\mathbf{51 \pm 90}^{**}$	$\mathbf{80 \pm 174}^{**}$	35 ± 69	$\mathbf{86 \pm 123}^{**}$	$\mathbf{59 \pm 106}^{**}$
Final temperature [K]	-0.1 ± 0.5	$\mathbf{-0.2 \pm 0.4}^{*}$	-0.1 ± 0.4	-0.2 ± 0.4	$\mathbf{-0.3 \pm 0.5}^{**}$	$\mathbf{-0.3 \pm 0.4}^{**}$	$\mathbf{-0.3 \pm 0.4}^{**}$	$\mathbf{-0.2 \pm 0.3}^{**}$

Statistically significant differences from baseline values are bolded and marked with asterisks: * for $p < 0.05$ and ** for $p < 0.01$.

with careful sensor design and placement, the SDNN parameter can also be determined via the PPG measurement method.

The RMSSD parameter is also frequently used as the time-domain method for assessing HRV. $RMSSD_{ECG}$ was accurate for both proposed systems, with only larger error noted in T4 for c-handle measurement. $RMSSD_{PPG}$ parameter showed poor performance and larger error in comparison to that of the reference system. The only exceptions were T1 and T4 in c-handle measurement and T1 in s-handle measurement. The $RMSSD_{PPG}$ parameter can thus be reliable only in tasks with low dynamics and low physical load, while $RMSSD_{ECG}$ can be used reliably for all task conditions.

EDA parameters have to be examined with caution, since bilateral measurements can differ in dc levels (tonic component) and conductivity fluctuations (phasic component). As found in previous studies, *Mean SCL* is closely linked to physical workload and in some cases cognitive workload. Interesting effect was found for *Mean SCL* parameter, which showed significant increase only for recording of the proposed system, but not for the reference system. For c-handle, *Mean SCL* was significant in both conditions of high physical load; for s-handle, it was significant for both conditions of low physical load. This can be due to the fact that the reference system was applied to the hand not involved in the physical control task and, hence, did not experience such perspiration throughout the task in comparison to the active hand.

SCR frequency has shown a significant increase from the baseline for all tasks and both handles. The trend of mean relative values between the proposed and reference systems is correlating well for c-handle measurement; however, large deviations in mean relative values can be seen from the results. This can be due to bilateral measurement as mentioned before, or due to the differences between the measurement systems. The reference system uses a preamplifier with an integrated low-pass filter, which could be the reason for the lower SCR detection rate in high arousal situations. The final temperature parameter showed significant reduction in almost all tasks except in T1 and T2 of s-handle measurement, which was expected due to the low physical load condition. In all tasks, low error of mean relative values was found between both the proposed system and the reference system.

C. Motion Artifacts

The highest value of SQSS was calculated for the angular velocity contribution in T2, which is aligned with the experiment protocol, where T2 was supposed to induce fast and jerky movements. It is interesting to point out that horizontal force contribution ($SQSS_{F_x}$) is not raised to a higher level. This could be due to the fact that the virtual mass was minimal in T1 and T2, and thus, it did not produce high enough force to pass the threshold. As expected, the artifact occurrence is increased in T4 for both horizontal force and angular velocity contribution.

The NOE parameters are used to show the mean failing rate of the algorithm and are mostly all below 1 event per minute. The lowest values of NOE are found for T1, which is expected due to the low dynamics and virtual mass.

Similarly, the PDP parameters are showing the percentage of positive detection within the entire session. It appears that although the failing rate of the algorithms is similar across all quantities, the PDP_{eda} and PDP_{temp} are having higher percentages of positive detection within the session. This implies that artifacts are influencing the detection algorithm for a shorter period of time. The lowest value of PDP is found for PPG algorithm detection, where results align with the raw correlation results from previous section.

To try and quantify different contributions of motion for a specific physiological signal, QOSO was calculated for T4 [see (3)]. Results of the different QOSO contributions are seen from Table II. Since PPG signals are most prone to motion artifact disturbance, it is normal to expect higher values of overlap. From the results, horizontal force is the largest contributor to the PPG algorithm fail rate, followed by the angular velocity and finally grasping force. An interesting result is shown for the EDA algorithm fail rate, where the largest contribution is calculated for the grasping force. The smaller values of QOSO are expected for more robust signals (e.g., EDA and temperature), where algorithm is not failing that often.

VI. CONCLUSION

We have presented a novel physiological measurement system for use in a physical rehabilitation task. Human–robot studies, based on physiological data analysis, are currently done mostly

by high-tech and expensive medical equipment with long setup times. To achieve a user-friendly and unobtrusive measurement with fast setup time, sensors were integrated into two different handle designs: a cylindrical and hemispherical shape. To best of our knowledge, this is the first case of an integrated physiological measurement system attached to a robot end-effector for an unobtrusive human–robot interaction measurement.

Experiment included a haptic robot manipulator with a measurement handle attached to its end-effector and a virtual task. By changing the model of the virtual environment, we manipulated task dynamics and physical workload to assess the change in signal integrity. Correlation with the reference system attached to the static hand showed different effects of task conditions on correlation values. For low physical load, most of the noticeable effects were seen only in EDA measurement. PPG measurement was mostly influenced by the high physical load condition. ECG measurement was found to be very robust in all four conditions. The temperature measurement has shown some effect in high physical load and high dynamics conditions. Additionally, most of the calculated physiological parameters have demonstrated robustness to handle-embedded sensory measurement.

To quantify the effect of motion, we devised a comprehensive study of motion artifacts and their effect on the signal quality. Artifacts were modeled as three-DOF sources of motion using the horizontal interaction force from the robot end-effector force cell, grasping force from the handle force cell, and angular velocity from the inertial sensor that was attached to the posterior side of the hand. QOS metrics were calculated and postprocessed to acquire parameters that described the influence of all three motion contributions. Results of raw signal correlation were aligned with the findings of motion artifact processing.

Importantly, results have also proven that hemispherical handle performed better for all physiological measurements than the cylindrical handle, except for PPG measurement. The results speak for reliability of such a measurement approach, which can be applied to all areas of human–machine interaction. Furthermore, these methods can be used in designing home-based rehabilitation and monitoring solutions.

Future work can be done by optimizing the handle design for higher comfort and universal use (sensory placement to cover wide range of hand sizes) and to develop adaptive algorithms for robust monitoring of physiological signals based on motion analysis.

REFERENCES

[1] E. Harmon-Jones, P. A. Gable, and T. F. Price, "Does negative affect always narrow and positive affect always broaden the mind?" *Current Directions Psychol. Sci.*, vol. 22, no. 4, pp. 301–307, 2013.

[2] H. I. Krebs *et al.*, "Robot-aided neurorehabilitation: A robot for wrist rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 3, pp. 327–335, Sep. 2007.

[3] P. Maciejasz, J. Eschweiler, K. Gerlach-Hahn, A. Jansen-Troy, and S. Leonhardt, "A survey on robotic devices for upper limb rehabilitation," *J. Neuroeng. Rehabil.*, vol. 11, 2014, Art. no. 3.

[4] Z. Zhang, Q. Fang, and X. Gu, "Objective assessment of upper limb mobility for post-stroke rehabilitation," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 859–868, Apr. 2016.

[5] F. C. Huang, R. B. Gillespie, and A. D. Kuo, "Human adaptation to interaction forces in visuo-motor coordination," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 3, pp. 390–397, Sep. 2006.

[6] N. Hocine, A. Gouaich, S. A. Cerri, D. Mottet, J. Froger, and I. Laffont, "Adaptation in serious games for upper-limb rehabilitation: an approach to improve training outcomes," *User Modeling User-Adapted Interact.*, vol. 25, no. 1, pp. 65–98, 2015.

[7] D. Novak, M. Mihelj, J. Ziherl, A. Olenšek, and M. Munih, "Psychophysiological measurements in a biocooperative feedback loop for upper extremity rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 4, pp. 400–410, Aug. 2011.

[8] R. Morales, F. J. Badesa, J. Rodriguez, N. García-Aracil, J. M. Azorin, and C. Pérez-Vidal, "A platform for researching on multimodal robot-assisted rehabilitation therapies," in *Proc. 4th IEEE RAS EMBS Int. Conf. Biomed. Robot. Biomechatron.*, 2012, pp. 1398–1403.

[9] K. Knaepen *et al.*, "Psychophysiological response to cognitive workload during symmetrical, asymmetrical and dual-task walking," *Human Movement Sci.*, vol. 40, pp. 248–263, 2015.

[10] K. H. Kim, S. Bang, and S. Kim, "Emotion recognition system using short-term monitoring of physiological signals," *Med. Biol. Eng. Comput.*, vol. 42, no. 3, pp. 419–427, 2004.

[11] A. Hariharan, P. Adam, and M. Thomas, "Blended emotion detection for decision support," *IEEE Trans. Human-Mach. Syst.*, vol. 45, no. 4, pp. 510–517, Aug. 2015.

[12] D. Novak *et al.*, "Psychophysiological responses to robotic rehabilitation tasks in stroke," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 4, pp. 351–361, Aug. 2010.

[13] F. J. Badesa, R. Morales, N. Garcia-Aracil, J. M. Sabater, C. Perez-Vidal, and E. Fernandez, "Multimodal interfaces to improve therapeutic outcomes in robot-assisted rehabilitation," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 1152–1158, Nov. 2012.

[14] F. J. Badesa, R. Morales, N. Garcia-Aracil, J. M. Sabater, A. Casals, and L. Zollo, "Auto-adaptive robot-aided therapy using machine learning techniques," *Comput. Methods Programs Biomed.*, vol. 116, no. 2, pp. 123–130, 2014.

[15] C. R. Guerrero, J. C. F. Marinero, J. P. Turiel, and V. Muñoz, "Using human state aware robots to enhance physical human–robot interaction in a cooperative scenario," *Comput. Methods Programs Biomed.*, vol. 112, no. 2, pp. 250–259, 2013.

[16] R. Morales *et al.*, "Patient-tailored assistance: A new concept of assistive robotic device that adapts to individual users," *IEEE Robot. Autom. Mag.*, vol. 21, no. 3, pp. 123–133, Sep. 2014.

[17] M. Swangnetr and D. B. Kaber, "Emotional state classification in patient–robot interaction using wavelet analysis and statistics-based feature selection," *IEEE Trans. Human-Mach. Syst.*, vol. 43, no. 1, pp. 63–75, Jan. 2013.

[18] S. Balters and M. Steinert, "Capturing emotion reactivity through physiology measurement as a foundation for affective engineering in engineering design science and engineering practices," *J. Intell. Manuf.*, pp. 1–23, 2015.

[19] M. P. Dijkers, P. C. deBear, R. F. Erlandson, K. Kristy, D. M. Geer, and A. Nichols, "Patient and staff acceptance of robotic technology in occupational therapy: A pilot study," *J. Rehabil. Res. Develop.*, vol. 28, no. 2, pp. 33–44, 1991.

[20] D. Giansanti, S. Morelli, G. Maccioni, and M. Brocco, "Design, construction and validation of a portable care system for the daily telerehabilitataion of gait," *Comput. Methods Programs Biomed.*, vol. 112, no. 1, pp. 146–155, 2013.

[21] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," *J. Neuroeng. Rehabil.*, vol. 9, 2012, Art. no. 21.

[22] S. Heuer, B. Chamadiya, A. Gharbi, C. Kunze, and M. Wagner, "Unobtrusive in-vehicle biosignal instrumentation for advanced driver assistance and active safety," in *Proc. IEEE EMBS Conf. Biomed. Eng. Sci.*, 2010, pp. 252–256.

[23] O. Postolache, V. Viegas, J. D. Pereira, D. Vinhas, P. S. Girão, and G. Postolache, "Toward developing a smart wheelchair for user physiological stress and physical activity monitoring," in *Proc. IEEE Int. Symp. Med. Meas. Appl.*, 2014, pp. 1–6.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

12   IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS

[24] J. Kim and E. André, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.

[25] M. Khezri, M. Firoozabadi, and A. R. Sharafat, "Reliable emotion recognition system based on dynamic adaptive fusion of forehead biopotentials and physiological signals," *Comput. Methods Programs Biomed.*, vol. 122, no. 2, pp. 149–164, 2015.

[26] D. Novak, M. Mihelj, and M. Munih, "Psychophysiological responses to different levels of cognitive and physical workload in haptic interaction," *Robotica*, vol. 29, no. 03, pp. 367–374, 2011.

[27] E. Richard and A. D. Chan, "Design of a gel-less two-electrode ecg monitor," in *Proc. IEEE Int. Workshop Med. Meas. Appl.*, 2010, pp. 92–96.

[28] M.-Z. Poh, N. C. Swenson, and R. W. Picard, "A wearable sensor for unobtrusive, long-term assessment of electrodermal activity," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 5, pp. 1243–1252, May 2010.

[29] N. Burns, "Cardiovascular physiology," Sch. Med., Trinity College, Dublin, Ireland, 2013.

[30] E. Gilab, J. M. Vergaracb, and P. Lagunaab, "Pulse transit time variability versus heart rate variability during decreases in the amplitude fluctuations of photoplethysmography signal," *Int. J. Bioelectromagn.*, vol. 12, no. 3, pp. 95–101, 2010.

[31] N. V. Thakor, J. G. Webster, and W. J. Tompkins, "Estimation of QRS complex power spectra for design of a QRS filter," *IEEE Trans. Biomed. Eng.*, vol. BME-31, no. 11, pp. 702–706, Nov. 1984.

[32] R. Stewart and E. Pfann, "Oversampling and sigma-delta strategies for data conversion," *Electron. Commun. Eng. J.*, vol. 10, no. 1, pp. 37–47, 1998.

[33] S. Ullrich, T. Knott, Y. C. Law, O. Grottke, and T. Kuhlen, "Influence of the bimanual frame of reference with haptics for unimanual interaction tasks in virtual environments," in *Proc. IEEE Symp. 3D User Interfaces*, 2011, pp. 39–46.

[34] G. Benbir, S. Ozekmekci, S. Oguz, G. Kenangil, S. Ertan, and E. Akalan, "Quantitative analysis of reduced arm swing frequency in essential tremor," *Eur. Neurol.*, vol. 63, no. 5, pp. 302–306, 2010.

[35] B. M. Appelhans and L. J. Luecken, "Heart rate variability as an index of regulated emotional responding," *Rev. Gen. Psychol.*, vol. 10, no. 3, pp. 229–240, 2006.

[36] Task Force of the European Society of Cardiology *et al.*, "Heart rate variability standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, pp. 354–381, 1996.

[37] W. Boucsein, *Electrodermal Activity*. New York, NY, USA: Springer, 2012.

[38] K. Sweeney, S. McLoone, and T. Ward, "A simple bio-signals quality measure for in-home monitoring," *Biomed. Eng.*, vol. 1, 2010.

[39] D. C. Fowles, M. J. Christie, R. Edelberg, W. W. Grings, D. T. Lykken, and P. H. Venables, "Publication recommendations for electrodermal measurements," *Psychophysiology*, vol. 18, no. 3, pp. 232–239, 1981.

[40] M. van Dooren *et al.*, "Emotional sweating across the body: Comparing 16 different skin conductance measurement locations," *Physiol. Behavior*, vol. 106, no. 2, pp. 298–304, 2012.

Blaž Jakopin received the B.S. degree in electrical engineering in 2014 from the University of Ljubljana, Ljubljana, Slovenia, where he is currently working toward the Ph.D. degree in the Laboratory of Robotics, Department of Measurement and Robotics, Faculty of Electrical Engineering.

Matjaž Mihelj (M'10) received the Graduate, M.Sc., and D.Sc. degrees in electrical engineering from the University of Ljubljana, Ljubljana, Slovenia, in 1996, 1999, and 2002, respectively.

Since 2014, he has been a full Professor with the Faculty of Electrical Engineering, University of Ljubljana.

Marko Munih (M'88) received the Ph.D. degree in electrical engineering from the University of Ljubljana (UL), Ljubljana, Slovenia.

He is currently a Full Professor and the Head of the Laboratory of Robotics, UL, where he was the Head of the Department for Measurements and Process Control, Faculty of Electrical Engineering, from 2004 to 2006.