2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
October 1-4, 2023, Oahu, Hawaii, USA

2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC) | 979-8-3503-3702-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/SMC53992.2023.10394331

Database for Human Emotion Estimation through Physiological Data in Industrial Human-Robot Collaboration

1st Justin R. Namba
Department of Computer Science
Rochester Institute of Technology
Henrietta, United States
jrn1325@rit.edu

2nd Karthik Subramanian
Dept. Elect. & Microelectronic Eng.
Rochester Institute of Technology
Henrietta, United States
kxs8997@rit.edu

3rd Celal Savur
Dept. Elect. & Microelectronic Eng.
Rochester Institute of Technology
Henrietta, United States
cs1323@rit.edu

4th Ferat Sahin
Dept. Elect. & Microelectronic Eng.
Rochester Institute of Technology
Henrietta, United States
feseee@rit.edu

***Abstract*—We introduce three new multi-modal data sets. They contain physiological and/or emotional information about human interactions with robotic arms in proximity to completing a task in an industrial setting. The data sets provide data from human subjects engaged in the assistive task of assembling a PVC joint pipe with robots. These data streams were collected to analyze and improve the comfort and safety of humans collaborating with robots in proximity in an industrial setting. These data sets can appeal to researchers studying human-robot collaboration, robot adaptation, and affective computing. Our data is stored in various formats, including images and human-readable Comma-Separated Values (CSV) or JavaScript Object Notation (JSON) files.**

***Index Terms*—human-robot collaboration, human-robot interaction**

I. Introduction

The study of collaborative processes between humans and robots, known as human-robot collaboration, involves human emulation and complementary approaches. The former aims to develop robots that mimic human-like abilities, while the latter recognizes the asymmetric capabilities between humans and robots and uses the strengths of robots to compensate for the weaknesses of human agents. Within industrial settings, complementary human collaboration is used in four ways: coexistence, sequential collaboration, cooperation, and responsive collaboration. The first two are the most common forms of collaboration, while the latter two involve simultaneous and real-time interaction between human and robot agents. However, there is a risk of collisions, which can significantly reduce productivity and injure human agents. Measures such as sensor-based detection of humans and limiting the speed and force of robot movements can be taken to ensure safe collaboration. Overall, human-robot collaboration in industrial settings can improve productivity and safety but requires careful consideration of the specific application and level of collaboration between human and robot agents. The goal of this paper is to present new databases for human arousal estimation. The novelty of these databases is that the data they contain derive from human-robot experiments in industrial settings. We summarize our main contributions as follows:

- **HEEP-HRI database** We present a database of human subject responses and features extracted from their physiological data during a human-robot interaction task.
- **HEEP-HRC database** We present a database of human subject responses and features extracted from their physiological data during a human-robot collaboration task.
- **AVE database** We present another database of human subject responses and images that capture their responses while interacting with a robotic arm in proximity.

For many years, researchers have primarily explored human-robot collaboration within controlled environments. These spaces are carefully regulated, encompassing factors such as conversation, position, temperature, luminosity, and actions. However, most of these research endeavors have had limited or no direct physical interactions between human and robot agents. As humans and robots increasingly interact in closer proximity in our daily lives, conducting further research and

This material is based upon work supported by the National Science Foundation under Award No. DGE-2125362. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

gathering data is crucial to ensure seamless and safe cooperation between the two.

Our research focuses explicitly on human-robot cooperation within an industrial setting. This environment allows human and robot agents to work closely together on tasks simultaneously. However, this proximity also introduces a higher risk of collisions, which can lead to physical damage and decreased productivity. To mitigate the occurrence of such collisions, we performed three human-robot cooperation experiments that emulated an industrial setting. Using the Circumplex model-based approach, we collected physiological and/or subjective data from human agents during these experiments to estimate the quality of cooperation between humans and robot agents. Figure 1 depicts our machine learning framework. With the features extracted and data estimated, three comprehensive databases are presented in this paper.

II. Related Work

The estimation of arousal and valence has been the focus of many datasets, such as SAL [1], SEMAINE [3], MAHNOB-HCI [5], Belfast naturalistic 2 [6], DEAP [4], RECOLA [7], SEWA [9], VAM-faces [2], AFEW-VA [8], Aff-Wild [10] and AffectNet [11] databases. The sensitive Artificial Listener (SAL) database stores the emotional states of a human conversing with a robot. The experiment consisted of four subjects recorded conversing with a robot who pretended to be a human agent. The research aimed to identify the signs of a failing conversation between humans and robots. The data collected includes facial and non-verbal signals such as pitch, spectral characteristics, timing, arousal, and valence. Sustained Emotionally colored Machine-human Interaction using Nonverbal Expression (SEMAINE) database also stores the emotional states of humans [3]. The experiment consisted of 20 human subjects that mimicked the SAL experiment, so instead of human-robot conversation, two human subjects communicated without seeing each other, while one of the human agents mimicked the SAL robot [1]. The human subjects were in separate rooms, but they could see each other using teleprompters, giving them the illusion that they were looking directly at each other's eyes. They communicated using a set of speakers. The goal was to analyze the audio and visual models to train a robot to understand human social signals/cues. Multi-Modal Affective Database for Affect Recognition and Implicit Tagging (MAHNOB-HCI) database contains 20 videos of 30 human subjects of various cultural backgrounds watching movie clips. All the human subjects wore sensors to capture physiological signals such as eye gaze, ECG, GSR, respiration amplitude, and skin temperature. Belfast Naturalistic Two consists of recordings of human subjects' faces and torsos while performing a series of 5 emotion laboratory-based tasks (frustration, disgust, fear, surprise, and amusement). The tasks were either passive or active. For example, during the frustration task, a human subject had to move a copper ring attached to a thin rod between the index finger and thumb along a zigzag bare copper wire without letting the ring touch the wire. Only self-report of emotions was documented. [5]

Database for Emotion Analysis Using Physiological Signals (DEAP) is a database of physiological signals (galvanic skin response (GSR), blood volume pressure, respiration pattern, skin temperature, electromyography (EMG), electrooculography (EOG), and electroencephalography (EEG)) of 32 human subjects watching music videos. With wearable sensors attached to them, the human subjects sat on a chair and watched music videos. At the end of each trial, the human subjects reported their emotional levels of arousal, valence, liking, and dominance [4]. REmote COLlaborative and Affective (RECOLA) database contains multi-modal data, including audio, video, and physiological data (electrocardiogram (ECG), electrodermal activity (EDA)), to estimate the arousal and valence of 40 human subjects performing a survival task. In this experiment, the human subjects collaborated with each other to rank 15 items, in order of importance, for the survival of a group of people experiencing a disaster scenario like a plane crash [7]. The SEWA database contains features extracted from audio and video recordings of 398 human subjects of an emotional experiment. The task was for the human subjects to watch advertisements videos and afterward discuss them with each other through video chat [9]. Vera am Mittag (VAM)-faces database contains the emotional responses extracted from the audio and video recordings from the guests in a German TV talk show, 'Vera Am Mittag.' The number of guests ranged from 2 to 5 people [2]. AFEW_VA database contains a collection of 600 video clips extracted from movies. Each frame was annotated with an arousal and valence value that ranges from -10 to 10 [8]. The Aff-Wild Database consists of 298 videos of 200 human subjects reacting spontaneously to various stimuli in the video clips [10]. AffectNet is a database of facial images extracted by querying emotion-related keywords in six different languages from three major search engines: Google, Bing, and Yahoo [11]. A general comparison of these data sets can be seen in Table I.

III. Experimental Design

We design two human-robot interaction (HRI) experiments and one human-robot collaboration (HRC) experiment to accomplish one task: assembling a PVC pipe joint. The first experiment is a human emotion estimation through HRI (HEEP-HRI) physiological signals. The agents are a human, sitting on a chair, and Sawyer, a robotic arm. The process is recorded using a stand-alone camera. The second experiment is a human emotion estimation through physiological signals in HRC (HEEP-HRC). The agents are a human in constant motion, a Sawyer robot, and UR10, a robotic arm that can adapt its speed based on its proximity to a human subject. In both of these experiments, wearable sensors are attached to the human agents to record physiological data. The third experiment is an arousal and valence estimation (AVE) through images. It is the same as the first experiment without the use of sensors to collect human agents' physiological signals. All the human agents have zero to some experience collaborating or interacting in proximity with robots.

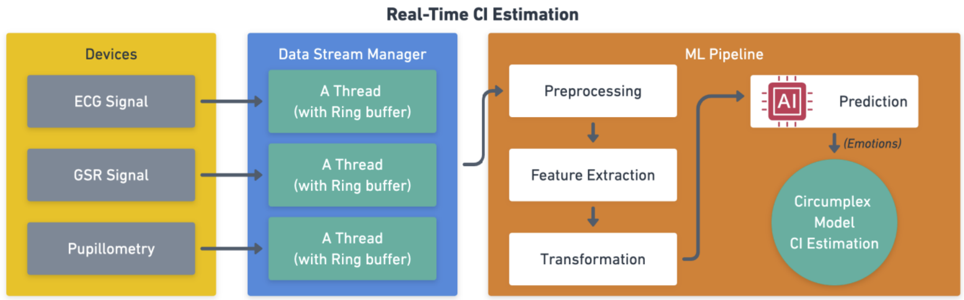

Fig. 1. Circumplex Model (CI) Estimation. [12]

TABLE I
AVAILABLE DATABASES FOR AROUSAL VALENCE ESTIMATION.

Database	Number of subjects	Amount of data	Elicitation	Environment
SAL	4	24 videos	Human-machine interface	Controlled
SEMAINE	20	959 videos	Human-machine interface	Controlled
MAHNOB-HCI	27	20 videos	Watching videos	Controlled
Belfast naturalistic	125	298 videos	Talk-show	Controlled
DEAP	32	40 videos	Watching videos	Controlled
RECOLA	46	46 videos	Online conversations	Controlled
SEWA	398	538 videos	Watching videos – Dyadic conversations	Controlled
VAM-faces	20	1867 images	Talk-show	Controlled
AFEW-VA	240	600 videos	Watching videos	Wild
Aff-Wild	200	298 videos	Watching videos	Wild
AffectNet	450,000	1,450,000 images	Human-machine interaction	Wild
HEEP-HRI	12	72 videos	Human-machine interaction	Industrial setting
HEEP-HRC	40	400 videos	Human-machine collaboration	Industrial setting

A. Human Emotion Estimation Through Physiological Signals in HRI (HEEP-HRI)

The setup of the first experiment consists of 12 human subjects, the Sawyer robot, and a board of 24 partially assembled PVC pipe joints, as shown in Figure 3. The assembly of each PVC pipe joint represents an iteration. It is the period in which the robot picks up a piece, follows a random trajectory, waits for the human to attach the last piece to the PVC pipe joint, and drops the object into a box. Picking, waiting, and dropping represent the three states the robot can be in. Moreover, wearable sensors are attached to all human subjects to record physiological data such as galvanic skin response, heart rate, and pupil dilation to predict their comfort. At the end of each iteration, the human subjects use a computer tablet to enter their emotions related to their comfort level with the robot. This subjective data consists of six values: comfort, safety, surprise, anxiety, boredom, and calmness. Figure 2 shows the customized app that human subjects use to report their feelings with a finger tap on the bars.

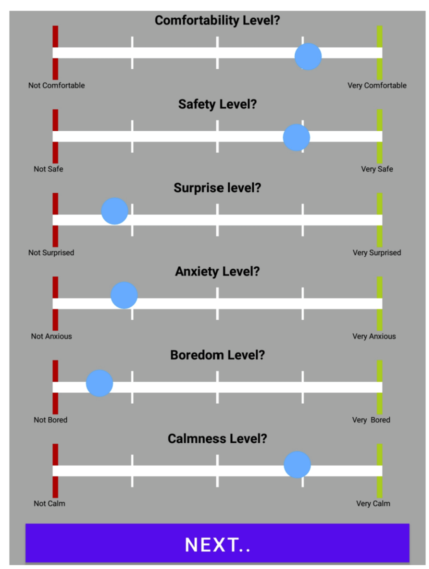

Fig. 2. Customized android app that human subjects use to articulate their feelings during the experiments. [12]

B. Human Emotion Estimation Through Physiological Signals in HRC (HEEP-HRC)

In the second experiment, we want to analyze the comfort of a human agent in motion, collaborating with a robot instead of a stationary human agent interacting with the robot in the first experiment. The setup of the second experiment is shown in Figure 4. It consists of 40 human subjects, two robotic arms called Sawyer and UR10, and a three-piece PVC pipe joint disassembled across three stations labeled A, B, and C. Station D is where the assembled object was placed. The wearable sensors used in the HEEP-HRI experiment are used in this one

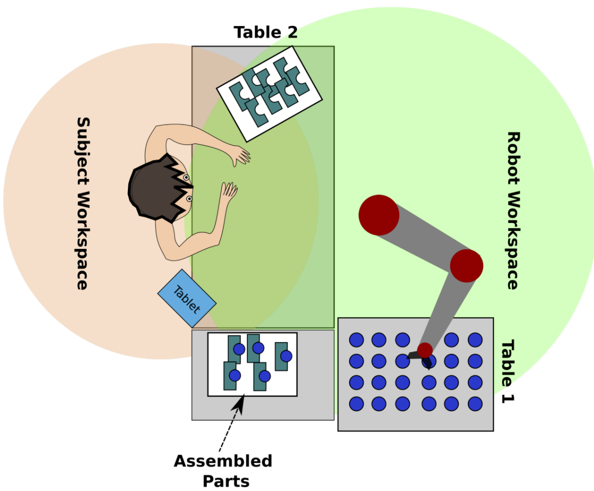

Fig. 3. Illustration of the HEEP-HRI experiment. [12]

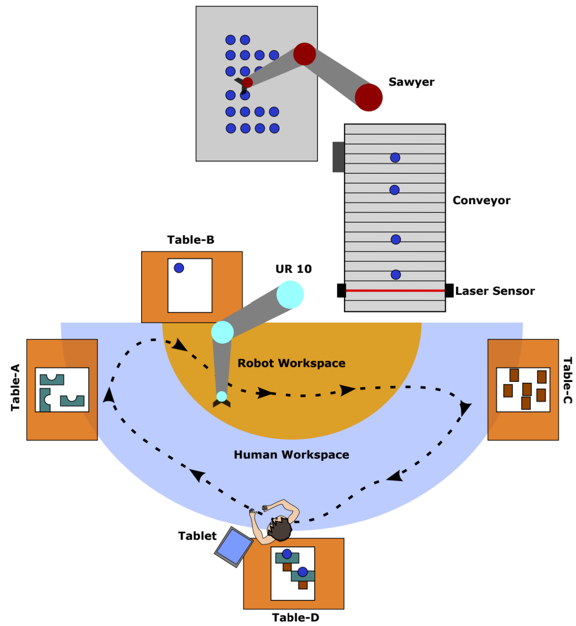

Fig. 4. Illustration of the HEEP-HRC experiment [12].

to collect the human subjects' physiological data. UR10 robot is a hand robot with device sensors attached to detect human proximity. UR10 has three modes: normal, slow, and stop. The robot automatically goes into different modes depending on its distance from the human agent. Each human subject assembled the object for a total of 9 iterations. During an iteration, human subjects pick up a piece on station A, walk to station B to get the second piece placed by the UR10 robot that got it from the Sawyer robot through the conveyor belt, and advance to station C to pick up the last part, and finally proceed to station D to drop the assembled object. At the end of each iteration, the emotions of the human subjects are recorded to evaluate their comfort working with the UR10 robot, just like in the HEEP-HRI experiment. Here the information is stored in Extensible Data Format (XDF).

C. Arousal and Valence Estimation (AVE) through Images Experiment

In the third experiment, we modify the design to represent a more realistic industrial setting by removing the wearable sensors from the human subjects. In such an environment, it would be inefficient and uncomfortable for human agents to always wear sensors while accomplishing their daily tasks. The setup of the AVE experiment was the same as the HEEP-HRI experiment, minus the sensors. During each iteration, the human subject and the Sawyer interactions are visually recorded using a camera later to deduce the human subject's comfort from facial data. Moreover, at the end of each iteration, the human subjects use a computer tablet to enter their emotions related to their comfort level with the robot. This subjective data consists of three values: arousal, valence, and comfort (AVC). They were derived from the arousal and valence model. Arousal refers to an emotion's intensity (low/calm or high/excited), and valence represents pleasantness (good or bad). The data collected was stored in rosbag files.

IV. Data Extraction

This section presents the procedures by which the data streams were extracted and transformed. Data is stored in a relational database for easy analysis and export for the three experiments described above. These data sets can be used for any work within human-robot collaboration or affective computing.

A. HEEP-HRI Database

In this experiment, the emotional and physiological data is stored in rosbag, a file format used to store Robotic Operating System (ROS) messages. The number of rosbag files is equivalent to the number of human subjects. That means a rosbag file contains all the information across the 12 trials in which each human subject assembled the PVC pipe joint throughout 24 iterations. We extract the human subject's physiological data and store it in a relational database. A sample of this data can be seen in Table II. It shows a portion of the features calculated and the response values of the third trial of the 7th human subject.

B. HEEP-HRC Database

This experiment stores emotional and physiological data in Extensible Data Format (XDF). It is a binary file format that stores blocks of big columnar data. The number of XDF files represents the number of human subjects. That means an XDF file contains all the information across the nine files in which each human subject assembled the PVC pipe joint throughout 24 iterations. In the XDF files, data is organized into streams. The number of streams varies across the different trials. A stream may contain physiological data (ECG, GSR)

TABLE II

A SAMPLE OF THE HEEP_HRI EXPERIMENT SUBJECTIVE DATA FOR TRIAL 3 OF SUBJECT 7.

Subject	Trial	Comfort	Surprise	Anxiety	Calmness	Boredom	ECG Rate	GSR Tonic Mean	Pupil Mean
P07	3	0.7609	0.8194	0.8997	0.8729	0.3645	102.7451	1.2089	38.9250
P07	3	1.7140	0.3227	0.6405	1.2910	0.3361	94.7723	-0.5452	38.2106
P07	3	1.2525	1.3027	0.8829	0.6973	0.3629	98.0785	-1.0582	36.7075
P07	3	1.6923	0.2575	0.3311	1.1973	0.6656	100.6050	-0.8697	36.8158
P07	3	1.3261	1.1472	0.8227	0.8445	0.6589	102.8115	-0.5167	36.7621
P07	3	1.7107	0.2993	0.3462	1.3211	0.7274	97.1081	-0.3947	38.2680
P07	3	1.2843	0.6472	0.6221	1.1923	0.7525	98.6855	-0.3935	37.8310
P07	3	1.3094	0.8829	0.6739	1.1605	0.7107	94.7036	-0.2085	39.4558
P07	3	1.8612	0.3495	0.3227	1.3662	0.7826	65.7169	0.0903	26.1024

TABLE III

A SAMPLE OF THE HEEP_HRC EXPERIMENT SUBJECTIVE DATA FOR TRIAL 1 OF SUBJECT 1.

Subject	Trial	Target	Surprise	Anxiety	Calmness	Boredom	ECG Rate	GSR Tonic Mean	Pupil Mean
P00	1	1.8946	0.2107	0.2040	1.7926	0.2842	107.0138	-0.4829	33.0026
P00	1	2.0184	-0.0167	0.1538	1.9448	0.1087	109.7169	-0.0402	33.9900
P00	1	1.8863	0.1355	0.1421	1.9632	0.1438	104.0610	-0.2912	33.9063
P00	1	2.0050	0.0987	0.0936	1.9565	0.1204	106.4450	-0.3073	34.4783
P00	1	1.9699	0.0753	0.1070	1.9582	0.0702	109.2336	0.1801	35.2770
P00	1	1.8127	0.0819	0.3211	1.8612	0.1472	103.5419	0.3311	35.4646
P00	1	1.5920	0.0719	0.1505	1.6104	0.0635	103.5126	1.1284	36.3801
P00	1	1.8930	0.0702	0.0067	1.9983	0.0819	105.6975	0.0828	36.0547
P00	1	1.5786	0.1355	0.4231	1.5903	0.0836	108.9285	0.5511	35.5241

or information about the robotic arm (states). We only extract the human physiological data for each stream and store it in CSV files. These files merge into a data frame and eventually export into a portable relational database for easy export and future updates. A sample of our database is shown in Table III. It shows some of the features calculated from the sensors data and the self-reported feelings of the human subjects after completing each iteration during a trial.

C. AVE Database

All the data collected in this experiment is stored in a rosbag. We extract image frames for each iteration, representing the time the robot picks up to when it drops the object (hence we can capture all the human responses during a random trajectory). We choose frames 10 seconds before and after the time the human subjects enter their responses because we believe those frames contain the information that explains an AVC value. For each iteration, we store the image frames in folders. Each image folder name is created by concatenating a human subject/participant id, trial id, and iteration id; it is structured as follows: **Person ID_Trial ID_Iteration ID**. Table IV shows a sample of our data from the AVC database. The first column is the image folder, whose name communicates the human subject identification, trial number, and iteration number.

Figure 5 depicts the image folder structure for each human subject. For example, P1 is a folder for the first human subject. This folder generally contains a set of 9 folders representing the number of trials accomplished. Nested within these trial folders are a maximum of 24 iteration folders, illustrating the number of times a human subject assisted the Sawyer robot in assembling a PVC pipe joint. Again, some trial folders may contain fewer than 24 iterations folders because human subjects forgot to enter their AVC values. The selected images capture the human subject's responses within the iteration folders.

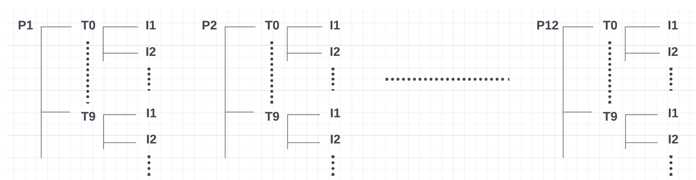

Fig. 5. AVE image folder structure.

The last three columns contain the human subject's arousal, valence, and comfort response. Arousal describes the excitement a stimulus provides to a human subject, and valence represents the pleasantness of a human subject. Both values are used to generate the comfort value.

TABLE IV

A SAMPLE DATABASE OF THE AVE EXPERIMENT. THE FOLDER CONTAINS THE IMAGE FRAMES THAT CAPTURE A HUMAN'S AROUSAL, COMFORT, AND VALENCE VALUES ASSOCIATED WITH EACH ITERATION.

Folder	Arousal	Comfort	Valence
P01_T0_I1	1.115819216	1.299373031	0.9025423527
P01_T1_I1	0.4957627058	1.266457677	1.300847411
P01_T1_I2	1.073446274	1.669278979	0.8022598624
P01_T1_I3	0.651129961	2.02664566	1.338983059
P01_T2_I1	0.6652542353	1.167711616	1.24011302
P01_T2_I2	0.5847457647	1.443573713	1.265536666
P01_T2_I3	0.5593220592	1.379310369	1.265536666
P01_T2_I4	0.5960451961	1.373040795	1.326271176
P01_T2_I5	0.6581920981	1.238244534	1.262711883

D. Sensors used in Experiments for Data Collection

Three device sensors were attached to the human subjects to collect their physiological data, such as galvanic skin response (GSR), heart rate, and pupil dilation. Shimmer3 was used to record GSR, a metric that uses sweat secreted on fingers to identify how a human responds emotionally to an external stimulus. An electrocardiogram (ECG) was used to record the electric activities of a human subject's heart. An eye-tracking device was used to measure the dilation of the pupils.

1) GSR: Galvanic skin response, also known as skin conductance, is a technique used to measure the skin's electrical conductivity to determine how humans respond to an external stimulus. We use a Shimmer3 GSR+ monitor to measure the changes in sweat glands secreted. Changes in skin conductance are automatic and happen fast, without the human being aware or in control. This makes GSR helpful and trustworthy because the information recorded is objective in the terms that human agents do not need to describe or hide their emotional state. Moreover, setting up the GSR device monitor is easy and straightforward. It needs to be connected to a computer that will receive the signals. Two sensors on the GSR device monitor need to be attached to the index and middle fingers of the non-dominant hand of a human agent. The sensor sampling rate was set to 128 Hertz (Hz). Regarding GSR data analysis, we measure the number of GSR peaks over a period. A peak is a response to an emotional stimulus. The greater the number of peaks, the greater the arousal during that period.

2) ECG: An electrocardiogram is a device used to measure the heart's electrical signals to monitor its health. We use Zephyr BioHarness, a chest-worn monitor, to record and transmit the physiological data of a human subject, such as heart rate, breathing rate, and a 3-axis accelerometer. The 3-axis accelerometer measures the activity level of human subjects and their posture. These physiological data are sent wirelessly to a lab computer.

3) Pupil Core labs: Pupil Core Labs is a wearable eye-tracking headset with three cameras. Two of the cameras are used to record eye movements, and the last one points to the subject's perspective. We focus on measuring pupil dilation, which occurs when the center of the iris's size increases to let more light in. It is a physiological response to an external stimulus. In our experiments, the dilation of the pupils means that the human subject is stressed.

V. Descriptive Statistics

Table V shows the features extracted and their descriptions from the wearable sensors used during the HEEP_HRI and HEEP_HRC experiments. We calculate mean HR, RR, SDNN, RMSSD, and pNN50 from the ECG. Mean HR is the average heart rate of the human subject. Mean RR is the average time interval between two consecutive heartbeats. SDNN is the standard deviation of the time interval of two consecutive heartbeats. RMSS is the root-mean-squared of the difference of two consecutive time intervals of two consecutive heartbeats. pNN50 is the percentage of successive RR/IBI intervals that differ by more than 50 ms. From GSR, we derive tonic mean,

TABLE V
Features extracted from ECG, GSR, and Pupillometry signals.

Type	Metric	Unit
ECG	Mean HR	bpm/min
	Mean RR	ms
	SDNN	ms
	RMSSD	ms
	pNN50	%
GSR	Tonic Mean	Micro-siemens
	Tonic Std	Micro-siemens
	Phasic Mean	Micro-siemens
	Phasic Std	Micro-siemens
	Onset Rate	onset/sec
	Peak Amp. Mean	Micro-siemens
	Rise Time Mean	ms
	Recovery Time Mean	ms
Pupil	Pupil Mean	pixel
	Pupil Std.	pixel

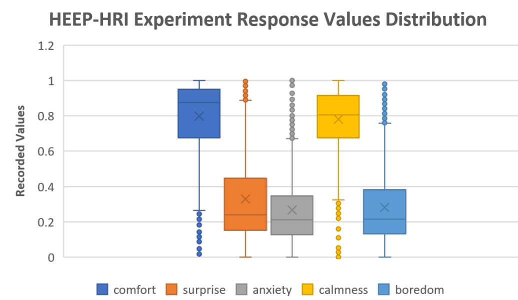

Fig. 6. Distribution of Subjects Responses in HEEP-HRI experiment.

tonic std, phasic mean, phasic std, onset rate, peak amplitude mean, rise time mean, and recovery mean. Tonic mean and tonic sdt (signal detection) are the average and standard deviation of the tonic component of GSR signals. Phasic mean and phasic std are the average and standard deviation of the phasic component of GSR signals—the onset rate of the change in response to a stimulus. Peak amplitude mean, rise time mean, and recovery time mean are the averages of the peak amplitude, rise time, and recovery time, respectively. Lastly, from Pupil labs, we calculate pupil mean and standard deviation, which is the average and amount of variability of the pupil signals, respectively.

A. HEEP-HRI Data Analysis

Figure 6 displays the five responses conveyed by human subjects in the HEEP-HRI experiment. As one can see, it is notable that none of the responses have a normal distribution. The majority of human subjects had values ranging from 0.1 to 0.5 for surprise, anxiety, and boredom, indicating a lack of these emotions. On the other hand, responses for comfort and calmness ranged between 0.7 and 0.9, indicating that most subjects were at ease while interacting with the robotic arm. However, all the boxplots have outliers, indicating a high level of skewness.

B. HEEP-HRC Data Analysis

Figure 7 presents boxplots summarizing the response values of all human subjects in the HEEP-HRC experiment. Most

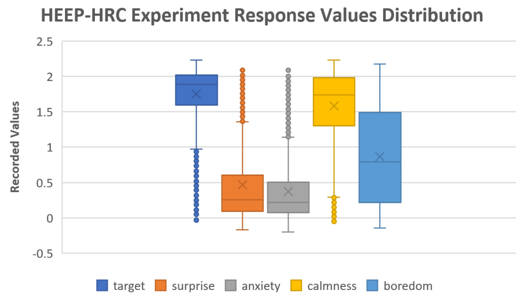

Fig. 7. Distribution of Subjects Responses in HEEP_HRC experiment.

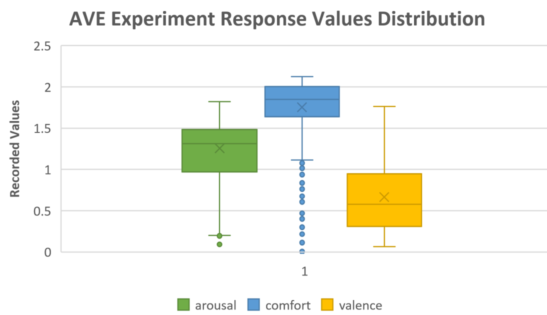

Fig. 8. Distribution of arousal, valence, and comfort values in AVE experiment.

subjects' target (comfort) values were within the range of 1.6 to 2.0, while their calmness levels ranged from 1.4 to 2.0. Meanwhile, their surprise and anxiety levels ranged from 0.1 to 0.6 and 0.1 to 0.5, respectively. However, their boredom levels were more variable and ranged from 0.3 to 1.5. Comfort and calmness had similar left-skewed distributions, indicating that human subjects generally felt calm and comfortable. Conversely, anxiety and surprise had highly right-skewed distributions, suggesting that most subjects did not feel anxious or surprised during the experiment. Boredom had a relatively uniform distribution with slight right-skewness, indicating that an almost equal number of subjects either felt bored or did not.

C. AVE Data Analysis

The distribution of the AVC values for the AVE experiment can be seen in the box plots in Figure 8. The majority have an arousal value between 1 and 1.5, a comfort value between 1.6 and 2.0, and a valence value between 0.3 and 1.8. We detect a lot of outliers ranging from 0.0 to approximately 1.2. Arousal had a distribution that was skewed to the left, which means that a lot more human subjects were excited to interact with the robot in proximity. Very few of them felt calm. Valence did not have a normal distribution and was skewed to the right, denoting that most human subjects did not enjoy interacting with the robotic arm Sawyer. Comfort had an approximately normal distribution skewed to the left, indicating that most of the human subjects felt comfortable working with Sawyer.

VI. Accessing Data

Readers can access our databases through the following link: https://mabl.rit.edu/datasets

VII. Conclusion

Most existing databases for human arousal estimation are limited in two ways: they are designed for data extracted from experiments in controlled settings, and the human subjects' reactions are typically not spontaneous. To overcome these challenges, we present three new databases that include features of subjective responses extracted from human emotion estimation information through physiological data experiments in an industrial setting. Unlike most existing databases, our databases have data from experiments that human subjects interact or collaborate with robots in an industrial environment.

References

[1] Douglas-Cowie, Ellen & Cowie, Roddy & Cox, Cate & Amir, Noam & Heylen, Dirk. (2008). The Sensitive Artificial Listener: an induction technique for generating emotionally colored conversation. Mathematics of Computation - Math. Comput.

[2] Grimm, Michael & Kroschel, Kristian & Narayanan, Shrikanth. "The Vera am Mittag German audio-visual emotional speech database," 2008 IEEE International Conference on Multimedia and Expo, Hannover, Germany, 2008, pp. 865-868, doi: 10.1109/ICME.2008.4607572.

[3] Mckeown, Gary & Valstar, Michel & Cowie, Roddy & Pantic, Maja. (2010). The SEMAINE corpus of emotionally colored character interactions. 2010 IEEE International Conference on Multimedia and Expo, ICME 2010. 1079-1084. 10.1109/ICME.2010.5583006.

[4] Koelstra, Sander & Mühl, Christian & Soleymani, Mohammad & Lee, Jong-Seok & Yazdani, Ashkan & Ebrahimi, Touradj & Pun, Thierry & Nijholt, Anton & Patras, Ioannis. (2011). DEAP: A Database for Emotion Analysis Using Physiological Signals. IEEE Transactions on Affective Computing. 3. 18-31. 10.1109/T-AFFC.2011.15.

[5] Soleymani, Mohammad & Lichtenauer, Jeroen & Pun, Thierry & Pantic, Maja. (2012). A Multi-Modal Affective Database for Affect Recognition and Implicit Tagging. Affective Computing, IEEE Transactions on. 3. 1 - 1. 10.1109/T-AFFC.2011.25.

[6] Sneddon, Ian & McRorie, Margaret & Mckeown, Gary & Hanratty, Jennifer. (2013). The Belfast Induced Natural Emotion Database. Affective Computing, IEEE Transactions on. 3. 32-41. 10.1109/T-AFFC.2011.26.

[7] Ringeval, Fabien & Sonderegger, Andreas & Sauer, Jürgen & Lalanne, Denis. (2013). Introducing the RECOLA multi-modal corpus of remote collaborative and affective interactions. 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, FG 2013. 1-8. 10.1109/FG.2013.6553805.

[8] Kossaifi, Jean & Tzimiropoulos, Georgios & Todorovic, Sinisa & Pantic, Maja. (2017). AFEW-VA Database for Valence and Arousal Estimation In-The-Wild. Image and Vision Computing. 65. 10.1016/j.imavis.2017.02.001.

[9] Kossaifi, Jean, Robert Walecki, Yannis Panagakis, Jie Shen, Maximilian Schmitt, Fabien Ringeval, Jing Han, Vedhas Pandit, Björn Schuller, Kam Star, Elnar Hajiyev and Maja Pantic. "SEWA DB: A Rich Database for Audio-Visual Emotion and Sentiment Research in the Wild." IEEE Transactions on Pattern Analysis and Machine Intelligence 43 (2019): 1022-1040.

[10] Kollias, Dimitrios & Tzirakis, Panagiotis & Nicolaou, Mihalis & Papaioannou, Athanasios & Zhao, Guoying & Schuller, Björn & Kotsia, Irene & Zafeiriou, Stefanos. (2019). Deep Affect Prediction in-the-Wild: Aff-Wild Database and Challenge, Deep Architectures, and Beyond. International Journal of Computer Vision. 127. 10.1007/s11263-019-01158-4.

[11] Mollahosseini, Ali & Hasani, Behzad & Mahoor, Mohammad. (2017). AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. IEEE Transactions on Affective Computing. PP. 10.1109/TAFFC.2017.2740923.

[12] Savur, Celal (2022). A Physiological Computing System to Improve Human-Robot Collaboration by Using Human Comfort Index" (2022). Thesis. Rochester Institute of Technology. Accessed from https://scholarworks.rit.edu/theses/11298