Post-print of: A. Kwasniewska, J. Ruminski, M. Szankin and K. Czuszynski, "Remote Estimation of Video-Based Vital Signs in Emotion Invocation Studies," 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, 2018, pp. 4872-4876, doi: 10.1109/EMBC.2018.8513423.

Remote Estimation of Video-Based Vital Signs in Emotion Invocation Studies

A. Kwasniewska, *Student Member IEEE, EMBS*, J. Ruminski, *Member IEEE, EMBS*, M. Szankin, *Member IEEE, and* K. Czuszynski

Abstract— **The goal of this study is to examine the influence of various imitated and video invoked emotions on the vital signs (respiratory and pulse rates). We also perform an analysis of the possibility to extract signals from sequences acquired with cost-effective cameras. The preliminary results show that the respiratory rate allows for better separation of some emotions than the pulse rate, yet this relation highly depends on a subject. The invoked positive emotion resulted in a respiratory rate difference > 1.8bpm, comparing to the average respiration rate of all neutral results (in 89% of observations). Visual facial expression in many cases was insufficient for emotion recognition (in video based experiment only 11.4% of visual responses were classified as an expected emotion).**

I. Introduction

The recent demographic change and global aging in industrialized countries has led to increased necessity of delivering systems that may support elderly people in their daily routines and, thus, foster their autonomy. To address this need, Ambient Assisted Living (AAL) solutions are often considered for remote healthcare applications [1]. AAL focuses on delivering services and concepts that allow for maintaining safety and well-being of individuals, while being beneficial for economy (better management of limited resources) and society (increased overall quality of life) [2].

Remote patient monitoring solutions that utilize image processing algorithms are often focused on investigating the changes within the facial area, as it is a highly sensitive region of the body [3] which allows for acquiring information about wellbeing and state of health. In some researches the analysis of emotions has already been considered for various remote patient monitoring applications, e.g. sentiment analysis [4]. Emotions are used for pain analysis and management by exploiting both spatial and temporal pain information from facial videos [5]. Some diseases, e.g. paralysis can alter the motor skills of the facial muscles [6]. Emotions recognition is also helpful in an indirect evaluation of other health problems, e.g. patients suffering from stroke [7] or neuropsychiatric disorders [8].

This work has been partially supported by NCBiR, FWF, SNSF, ANR and FNR in the framework of the ERA-NET CHIST-ERA II, project eGLASSES – The interactive eyeglasses for mobile, perceptual computing and by Statutory Funds of Electronics, Telecommunications and Informatics Faculty, Gdansk University of Technology.

A. Kwasniewska (alicja.kwasniewska@pg.edu.pl), J. Ruminski (jacek.ruminski@pg.edu.pl), K. Czuszynski (krzczusz@student.pg.gda.pl) are with Gdansk University of Technology, Faculty of Electronics, Telecommunications and Informatics, Deptartment of Biomedical Engineering, Gdansk, Poland

A. Kwasniewska (alicja.kwasniewska@intel.com), M. Szankin (maciej.szankin@intel.com) are with Intel Corp., San Diego, CA, USA

Detection of emotions from visible light images has been well studied and currently achieves a high recognition accuracy [9]. Yet, processing RGB images causes more and more concerns [10], as there is the potential for security threats and privacy violations in the era of always-on devices equipped with visible light cameras. This can be addressed by using thermal images, because instead of showing specific shape and color of an object, they show the temperature distribution, so the risk of revealing individuals identity in the thermal images is much lower [11]. Moreover, emotion recognition from thermal images could make use of bio-signals instead of facial expressions, as in RGB images. Additionally, the dynamic changes of facial temperature patterns can be potentially related to psychological and physiological status of the observed individual. This is potentially very useful because suppressing or masking bio-signals representing emotional response is very hard [12].

The rationality and validity of extracted emotional information are main debates in cognitive affective computing studies. The achieved results are often dictated by the quality and diversity of selected data itself [13]. Yet, only limited databases of thermal, infrared images are available for the affective computing studies. Some examples include NVIE database [14], MAHNOB database [15], and Equinox database [16]. The NVIE database contains visible and thermal infrared images for spontaneous expressions and induces emotions using emotion-stimulating videos. These studies were mainly based on facial-expressions and related thermal patterns observable in an image. In such experiments, authors are typically looking for the best features in single images that could offer the highest emotion recognition accuracy [17][18][19]. However, there is an important research question if physiological signals (e.g. pulse wave, respiration wave) obtained from the visible and thermal sequences can carry useful information about emotions. In [20] pulse and skin conductance were used for multimodal emotion recognition system for evaluating positive, neutral and negative responses resulting in 41.2% accuracy. Video-based emotions invocation studies [21] proved that 2 emotional states can be also detected with accuracy of 71.4% by extracting heart rate variabilities from ECG signals. According to [22], analysis of emotions with thermal imaging can be more versatile than with RGB sensors, as it allows for observing changes that are impossible to communicate by some patients, e.g. infants. The results presented in [23] showed a parallelism between facial temperature distribution of mother and a child in distressing situations.

In this preliminary study, we were mainly interested:

- if physiological signals can be extracted from visible and thermal infrared videos (using cost-effective cameras) for the needs of emotional analysis,

- if parameters of signals (pulse rate or a respiratory rate) may change for various imitated and video-invoked emotions and how they are changing,
- what kind of video-based physiological signals properties could be important in further studies focused on affective-computing,
- what kind of conditions should be met to develop a remote telemedicine service for the evaluation of emotional changes of a supported person.

Therefore, we propose to analyze respiratory rate changes in two scenarios: as an effect of imitating emotions by subjects and as a natural response to visual stimulus. For respiratory rate estimation, we use sequences acquired with low-cost small thermal camera module that can be embedded e.g. in wearable devices [24] or almost transparently added to the already existing home infrastructure. Simultaneously, we record videos using a RGB camera to evaluate visible facial expression changes and try to calculate a pulse rate. By making the use of multimodalities, we want to determine if there is a correlation between vital signs and facial muscles changes in both controlled and uncontrolled emotional states.

The paper is organized as follows: in Section II we demonstrate methodology used for respiratory rate and pulse extraction and emotion recognition. Section III presents the experimental results, further discussed in Section IV. Finally, we conclude the paper in Section V.

II. Methodology

The validation of emotions influence on vital signs was evaluated by performing experiments on the group of 11 healthy volunteers (age: 33.7±11.3) in a testing room at an ambient temperature between 22–25 °C. In our studies, we analyzed respiratory rate calculated by analysis of temperature changes in a nostril area and pulse estimated using imaging photo-plethysmography [25]. Video sequences were recorded using two cameras placed at a distance ~ 0.4-1m from a subject, thermal camera aimed at a face upward at an angle of 15° to make the nostril area more exposed. Thermal images were captured with FLIR® Lepton – a small (<1cm^2), high-dynamic range (14bit) thermal camera, with a resolution of 80x60 pixels and sampling frequency at 9Hz. For the visual spectrum video acquisition and analysis, the Logitech Webcam 9000 Pro camera was used (30 frames per second at 640x480 resolution). Participants were introduced to experiments using an online questionnaire. By filling it they provided information about their age, heart problems and the ease of getting irritated on the scale from 0-10. Information about the goal of the study and the organization of the experiment was described to every participant. They all agreed to participate in experiments as volunteers. After that, they were led through a series of tasks divided into two experiments. The first consisted of imitating 4 emotions (I1 neutral, I2 joy, I3 fear, I4 disgust) 1 min. each, with intervals of 2-min relaxation pause between them, as presented in Fig. 1. The participants were asked to simulate emotions, not only facial expression.

In the second experiment, subjects were presented a series of videos (available at [26]) selected to invoke following emotions: joy (V2 - funny scenes from a gym), disgust (V4 – eating worms), fear (V6 – a dark basement with ghosts), sadness (V8 – dying animals), separated by other clips that were supposed to induce the neutral mood (V1 – an empty road, V3 – an ocean, V5 – a snail race, V7 - clouds). The steps of this procedure are presented in Fig. 2. After watching the videos, participants were asked to name the dominant emotion that was accompanying them during watching each clip. During data acquisition in both experiments the pulse was recorded using the Sanitas SPO25 finger pulse oximeter.

Figure 1. Procedure of the 1st experiment; imitated facial expressions related to given emotions: I1 neutral, I2 joy, I3 fear, I4 disgust

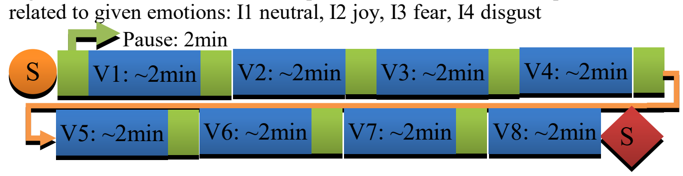

Figure 2. Procedure of the 2nd experiment; video-induced emotions V1 neutral, V2 joy, V3 neutral, V4 disgust, V5 neutral, V6 fear, V7 neutral, V8 sadness

For further analysis, data fragments from the beginning and the end of each recording were used (thermal sequences ~400 first and ~400 last samples; RGB ~500 first and ~500 last samples) to accommodate eventual inertia of emotional response. Using short data segments allowed for reducing possible motion artifacts. In the first step of data analysis intensity of radiation changes were extracted from the manually selected region of interests (ROI) (Fig. 3) around nostril area using the skewness, variance or average operator for all pixel values inside the ROI to achieve the best quality of the extracted signal on the recorded sample as explained in details in [27]. Obtained signals were filtered using moving average filter and the 4th-order high pass Butterworth filter with cutoff frequency of 0.125Hz applied for baseline removal. Respiration rate estimator based on the dominated peak in the frequency spectrum for the autocorrelation function (*eRR_ac*) [27] was then applied to the filtered signal to calculate the respiratory rate. The accuracy of the method was previously verified in [28].

Additionally, for 6 participants the visible light videos were analyzed to obtain the pulse rate. For this, the RGB recordings were converted using the H264 codec, rescaled to 800x600 pixels and transformed to the YUV420P color space

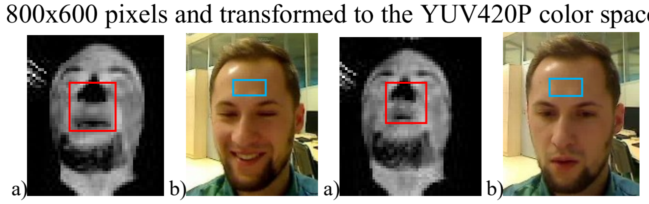

Figure 3. Approximated ROI position for respiratory (a) and pulse (b) rates estimation during imitating joy and fear emotions

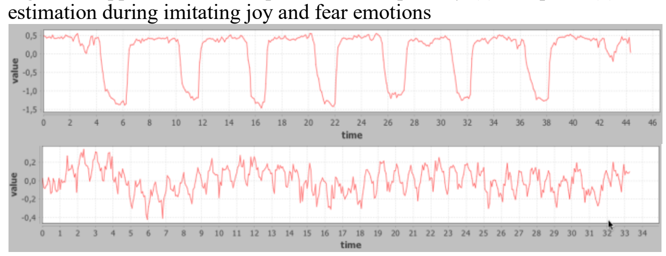

Figure 4. Examples of raw signals extracted from sequences: a) respiration wave (skewness-based data aggregation in each ROI), b) pulse wave (average-based data aggregation in each ROI)

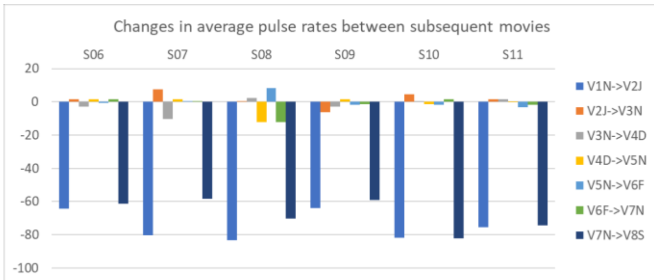

Figure 5. Changes in estimated pulse rate in transition from subsequent video-based emotion invocation studies for subjects S06-S11.

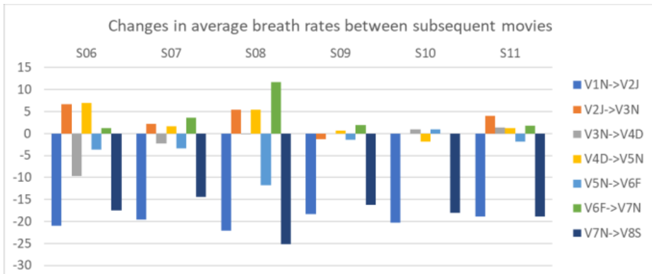

Figure 6. Changes in estimated breath rate in transitions from subsequent video-based emotion invocation studies for subjects S06-S11.

with frame rate reduction to 15.02 frames per second. Then, pixels values inside the manually selected area (on the forehead – Fig. 3) were averaged for each frame producing signals, further analyzed to estimate a pulse rate. Initially, all signals were filtered with the band pass (frequency range between 0.67Hz and 4Hz) Butterworth filter. After that, *ePR_sp* estimator [25] (the frequency value of the dominating peak in the frequency spectrum for the autocorrelation function applied to the filtered signal) was used to determine the pulse rate. The accuracy of the method was previously verified in [29]. Additionally, the level of the pulse rate was checked using readouts from the pulse oximeter. Examples of raw signals extracted from sequences are presented in Fig. 4.

Finally, to compare the relation between the facial expression, vital signs and perceptible emotional response, the RGB videos were analyzed with the Microsoft Emotion Cognitive Service [30]. The use of this API allowed for obtaining the confidence across a set of emotions represented by facial expressions: anger, contempt, disgust, fear, joy (happiness), neutral, sadness, surprise.

III. Results

The respiratory rate evaluated from thermal image sequences and pulse rate calculated from visible light images for both experiments: imitating emotions and invoking emotions by video stimulus are collected in Table I and II, respectively (1^{st} indicates values estimated using 400 for thermal and 500 for RGB samples at the beginning of sequence, analogically 2^{nd} - last 400-500 samples). Label *tp* indicates a technical problem encountered during data collection; *fb* - a face turned away from the camera. Table III presents emotions self-estimated by subjects after watching videos. Using the Microsoft Emotion Cognitive Service, we obtained dominant emotions for each frame in the sequence. Table IV demonstrates two most frequent emotional responses and the percentage of frames in which it was recognized for each sequence. Emotions (Table III, IV) are labelled with abbreviations: anger A, joy J, neutral N, surprise Sr, fear F, disgust D, sad S, unrecognized u. In Fig. 5. and Fig. 6. we plotted differences between estimated vital signs for subsequent video-based emotion invocation studies. At first for each subject average of estimated vital signs for each sequence was calculated (average of 1^{st} and 2^{nd}). Then, differences between average values were presented on the plotted graphs. For Fig. 5, Fig. 6. and Fig. 7. we presented results only for S06-S11, as only for them all tested data (both heart rate and pulse rate) was captured.

TABLE I. Respiratory rate evaluated from thermal images

	I1 N		I2 J		I3 F		I4 D		V1 N		V2 J	
Sub	***1st***	***2nd***	***1st***	***2nd***	***1st***	***1st***	***1st***	***2nd***	***1st***	***2nd***	***1st***	***2nd***
S01	21.60	20.30	18.90	21.60	21.60	21.60	16.20	17.60	21.60	18.90	tp	tp
S02	20.30	20.30	18.00	19.80	18.90	18.90	24.30	25.20	18.90	18.90	28.40	20.30
S03	18.69	22.85	21.60	21.60	18.90	18.90	19.80	18.00	18.90	21.60	17.60	17.60
S04	11.70	10.80	14.72	16.69	16.20	16.20	16.20	14.40	16.20	14.40	17.28	14.85
S05	13.50	13.50	16.20	16.20	12.15	12.15	14.85	13.50	12.15	14.85	tp	15.12
S06	21.60	22.70	21.60	23.10	22.20	22.20	17.30	20.90	22.20	14.30	21.60	20.30
S07	14.40	13.10	14.40	14.40	tp	tp	12.60	14.40	tp	tp	22.90	16.20
S08	18.50	18.50	14.00	15.80	27.00	27.00	24.00	27.70	27.00	18.90	21.60	22.70
S09	14.40	12.60	21.60	21.60	11.20	11.20	14.40	21.60	11.20	11.20	17.60	18.90
S10	18.00	18.00	18.00	19.80	17.60	17.60	19.80	21.60	17.60	17.60	21.60	18.90
S11	21.60	21.60	21.60	19.80	17.60	17.60	21.60	20.30	17.60	17.60	18.90	18.90
	V3 N		**V4 D**		**V5 N**		**V6 F**		**V7 N**		**V8 S**	
S01	18.90	21.60	16.20	17.60	18.90	21.60	17.60	21.60	18.90	20.30	21.60	21.60
S02	18.90	17.60	18.90	17.60	18.90	14.90	23.40	fb	21.60	23.00	20.30	fb
S03	18.90	19.80	20.30	18.90	16.20	18.90	20.30	20.30	18.90	17.60	16.20	20.30
S04	15.12	14.72	14.04	10.80	11.88	14.40	15.12	15.12	12.15	10.80	13.50	14.85
S05	12.15	13.50	13.50	16.20	12.15	13.50	15.75	fb	12.15	13.50	12.15	13.50
S06	16.20	15.40	14.40	30.60	12.90	12.90	18.00	21.60	18.00	14.40	17.10	17.80
S07	12.60	12.60	13.50	16.20	12.60	12.60	14.90	13.50	10.80	10.80	14.40	14.40
S08	14.50	14.50	20.10	19.80	21.60	18.00	27.00	23.60	14.40	12.60	25.20	25.20
S09	16.20	16.20	14.90	14.90	14.90	14.90	14.90	16.20	13.50	14.90	16.20	16.20
S10	18.00	18.00	18.00	18.00	18.00	19.80	18.00	16.20	18.00	18.00	18.00	18.00
S11	12.10	12.10	16.80	15.60	17.60	17.60	19.80	18.00	16.20	18.00	18.00	19.80

TABLE II. Pulse rates evaluated from visible light sequences

	I1 N		I2 J		I3 F		I4 D		V1 N		V2 J	
Sub	***1st***	***2nd***	***1st***	***2nd***	***1st***	***2nd***	***1st***	***2nd***	***1st***	***2nd***	***1st***	***2nd***
S06	59.76	59.76	57.18	59.47	61.73	59.42	59.63	61.68	59.86	57.80	66.73	61.97
S07	57.26	56.58	69.43	72.00	66.88	66.88	66.86	70.20	61.71	59.14	72.70	88.27
S08	84.60	84.98	89.99	91.60	82.80	88.20	88.19	89.99	82.80	81.00	87.27	79.31
S09	59.39	57.59	82.78	84.57	63.00	64.80	64.80	59.87	57.66	57.66	61.20	66.60
S10	75.27	76.87	74.85	75.30	76.80	79.89	80.40	81.00	76.46	79.15	80.49	82.90
S11	73.36	70.64	76.79	70.71	75.80	80.14	78.62	71.41	75.48	72.23	75.00	75.96
	V3 N		**V4 D**		**V5 N**		**V6 F**		**V7 N**		**V8 S**	
S06	61.66	59.09	64.10	59.52	57.85	59.99	61.22	59.49	59.80	59.80	59.91	62.70
S07	57.65	59.45	69.51	63.07	55.86	55.86	57.65	57.65	55.80	59.75	57.60	59.02
S08	90.00	77.40	86.52	81.11	82.90	89.19	75.65	72.05	82.84	81.68	70.20	70.20
S09	61.20	77.40	63.03	63.03	59.44	61.24	63.00	61.20	59.40	61.67	59.45	58.65
S10	77.99	78.93	84.13	81.73	85.37	81.35	80.40	83.72	80.40	80.60	82.85	81.30
S11	73.05	79.00	78.26	77.23	82.33	76.08	75.19	83.19	72.14	80.16	73.65	75.25

TABLE III. Emotions self-estimated by subjects after watching videos in experiment 2

Sub	**V1 N**	**V2 J**	**V3 N**	**V4 D**	**V5 N**	**V6 F**	**V7 N**	**V8 S**
S01	N	J	N	D	N	F	N	S
S02	N	J	N	D	N	F	N	S
S03	N	N	N	D	N	F	N	N
S04	N	J	N	D	N	J	N	S
S05	J	J	N	D	N	F	N	S
S06	N	J	N	D	N	F	N	S
S07	N	J	N	D	N	N	N	N
S08	N	J	J	D	N	F	J	S
S09	N	J	N	D	N	J	N	S
S10	N	J	N	N	N	N	N	S
S11	N	N	N	D	N	F	N	S

TABLE IV. TWO MOST FREQUENT EMOTIONAL RESPONSES AND THE PERCENTAGE OF FRAMES IN WHICH IT WAS RECOGNIZED [%]

Sub	**I1 N**	**I2 J**	**I3 F**	**I4 D**	**V1 N**	**V2J**	**V3 N**	**V4 D**	**V5 N**	**V6 F**	**V7 N**	**V8 S**
S01	N100	J95 N5	Sr82 N18	N49 D30	N100	J74 N26	N100	N71 D15	N100	N99 F1	N100	N100
S02	N99 u1	J82 N10	N80 u20	N92 J5	N22 u78	J72 N5	N49 u51	N77 u23	N72 u28	N36 u46	N89 u11	N22 u78
S03	N100	J93 Sr4	Sr94 N5	A43 N40	N100	N93 J6	N100	N88 J12	N100	N92 J7	N100	N99 Sr1
S04	N96 u4	J92 N4	N87 J10	N44 J35	N96 u4	N82 J16	N89 J5	N71 J24	N95 u4	N91 u5	N85 u12	N93 J3
S05	N99 u1	J90 N10	N97 J1	N96 u4	N93 J3	N64 J30	N99 u1	N93 u6	N97 J2	N49 J24	N89 J7	N82 S10
S06	N100	J80 N20	N93 J5	N53 A39	N100	J92 N7	N100	N78 A16	N100	N100	N100	N100
S07	N100	J99 N1	N87 J13	N92 S7	N100	J89 N11	N100	J59 N40	N100	N98 A1	N100	N100
S08	N100	J95 N5	N100	N100	N96 J4	N65 J35	N100	N61 J39	N100	N71 J27	N100	N96 J3
S09	N70 J25	J64 u32	N76 J24	N50 J47	N100	J58 N13	N100	J64 N32	N56 J42	N80 J20	N100	N98 J2
S10	N97 u3	J23 u77	N33 u59	J61 u26	N31 u68	N38 u53	N100	N68 u31	N48 u51	N42 u57	N50 u50	N45 u55
S11	N64 u35	J23 u76	J26 u63	S7 u86	N37 u62	N14 u85	N55 u45	N25 u74	N40 u53	N70 u29	N95 u5	N87 S12

The relation between estimated vital signs (pulse and respiratory rate) for S06-S11 while inducing emotions with positive (joy) and neutral videos is shown in Fig. 7. Horizontal axis presents respiration rate, vertical axis pulse rate. Red circles correspond to results of measurements took for the latter samples (400-500 last samples), orange for the initial ones (400-500 first samples). Fig. 8. depicts the relation between a pulse and a respiratory rate for subject S07 and S10 in emotion invocation study (joy and neutral videos).

IV. DISCUSSION

In this research, we analyzed the influence of emotional states on respiratory and pulse rates. The conducted experiments covered both controlled and uncontrolled expression of emotions. The vital signs were calculated using estimation methods applied to signals extracted from nostril (thermal images) and forehead (visible light) areas.

The resolution of calculation of vital signs is limited due to the finite frequency resolution. Preprocessed visible light videos were sampled with the frequency of 15Hz. Given the 500 samples after the preprocessing, the frequency quantum is equal 15/500*60~= 1.8bpm (beats per minute). For the respiratory wave: 9/400*60=1.35bpm (breaths per minute). However, in some cases shorter respiratory signals (N~=300) wave were used due to rapid movements of individuals (e.g. for "joy"), who sometimes acted very emotionally, although all volunteers were asked to remain still during experiments. Then the resolution was about 9/300*60=1.8bpm. Therefore, the value of 1.8bpm was used as a safe threshold to indicate the change of the rate between measurements.

In 68.2% of observations (30/44) no difference in respiration rates was observed between the first and last periods of the measured signals for neutral movies. For simulated emotions, the corresponding result was 91% (10/11). Analysis performed for the invoked positive emotion (joy) showed that in 67% of cases (6/9) a difference higher than 1.8 bpm was noted comparing to the average respiration rate of the previous neutral result and 89% (8/9) comparing to

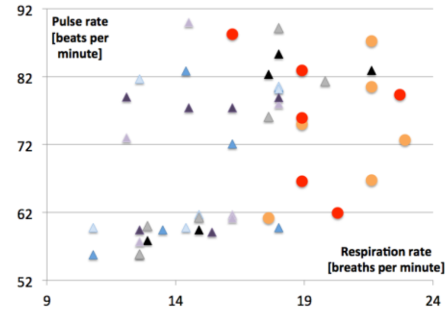

Figure 7. Pulse rate vs. respiratory rate for "V2 joy" stimulation video (circles) and neutral videos (triangles)

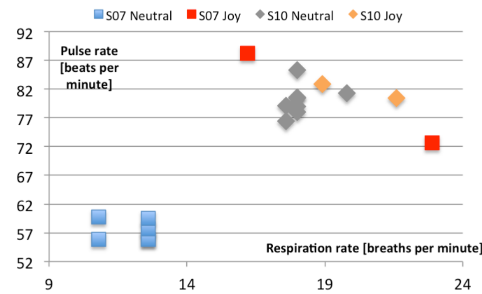

Figure 8. Pulse rate vs. respiratory rate for 2 subjects (S07, S10)

the average respiration rate of all neutral results. As presented in Fig. 5. and Fig. 6. highest differences in both vital signs were observed at the transition from 1st neutral to joy and from last neutral to surprise videos. The exact values in differences depends on the subject and the achieved preliminary results should be further confirmed.

Some image sequences recorded for the 'V2 joy' movie contained high head movements and mouth activities. Most of such situations were compensated using the larger size of the ROI and the skewness operator to aggregate data, as presented in Fig. 3. However, for 3 (out of 22) subsequences it was impossible. In such cases, a nose (a mouth) detection (or tracking) algorithms in thermal imagery should be implemented (e.g. as presented in [31] and [32]). Moreover, in certain cases we observed some irregularities (wide spread of the results, both for different subjects and same subjects tested twice in short time span) in the processed signals for emotion 'Joy'. Such patterns could be used in further work as a possible source of information in classification of emotions.

The results of this preliminary study show that the respiratory rate allows for better separation of some emotions than the pulse rate (Fig. 5). This relation highly depends on a subject. For example, for the subject S07 results for the invoked "neutral" emotions are closely aggregated (Fig. 6) showing similar values of the respiratory and pulse rates. However, for the invoked "joy" emotion the results are clearly separated from the cluster of the "neutral" responses. We discovered that in real situation an observed person can react not only with a changed facial expression but also with a more dynamic movement, like rotating a head or covering the view. Therefore, the designed algorithm for emotional analysis should be also sensitive to such events, which could be even more challenging. Some solutions could be based on a face or gaze detection algorithms and a changed pattern in the signal extracted from the analyzed ROI to recognize the moment of face disappearance from the field of view. It was also observed that visual changes in the facial expression in many cases are insufficient for emotion recognition. For emotion-imitation experiment 39% of samples were labeled with the expected emotion by the Microsoft Emotion API, while for video-stimulus only 11.4% (neutral videos not considered, treated as relaxation pauses). Therefore, we see a need for further development of emotion recognition system based on bio-signals.

V. Conclusion

A series of in-depth analysis was performed in this work to evaluate the possibility of physiological signals extraction from visible and thermal infrared videos for emotional analysis. The preliminary results proved that it is possible to detect changes in vital signs related to emotional responses both in controlled (imitation of emotion) and uncontrolled (video stimulated) scenarios. It was also shown that susceptibility to emotional stimulation using videos could differ between individuals. For future work, we plan to utilize more advanced computer vision algorithms to accommodate the system for possible movements of patients and improve emotion recognition accuracy, e.g. by using deep neural networks. This will also require much more data, which is important for deep networks and for the proper experimental verification of the preliminary results obtained in this study.

References

[1] A. Dohr, R. Modre-Opsrian, M. Drobics, D. Hayn and G. Schreier, “The Internet of Things for Ambient Assisted Living” Seventh International Conference on Information Technology: New Generations, Las Vegas, NV, 2010, pp. 804-809, 2010

[2] J. Morak, M. Schwarz, D. Hayn, G. Schreier, “Feasibility of mHealth and Near Field Communication Technology Based Medication Adherence Monitoring”, Engineering in Medicine and Biology Society (EMBC) 2012 Annual International Conference of the IEEE, pp. 272-275, 2012, ISSN 1557-170X.

[3] P.M. Prendergast, “Anatomy of the Face and Neck” in: Shiffman M., Di Giuseppe A. Cosmetic Surgery. Springer, Berlin, Heidelberg, 2013

[4] A. D. Torres et al., “Patient Facial Emotion Recognition and Sentiment Analysis Using Secure Cloud with Hardware Acceleration” in Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications, Book chapter, Elsevier, 2017

[5] M. Bellantonio et al. “Spatio-temporal Pain Recognition in CNN-Based Super-Resolved Facial Images” Nasrollahi K. et al. (eds) Video Analytics. Face and Facial Expression Recognition and Audience Measurement. FFER 2016, VAAM 2016. Lecture Notes in Computer Science, vol 10165. Springer, Cham, 2017

[6] S. He, J.J. Soraghan, B.F. O'Reilly, “Objective Grading of Facial Paralysis Using Local Binary Patterns in Video Processing” 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, pp. 4805-4808, 2008

[7] R. Yuvaraj, M. Murugappan, M. Norlinah, K. Sundaraj, M. Khairiyah, “Review of Emotion Recognition in Stroke Patients” Dementia and Geriatric Cognitive Disorders; 36:179-196. Karger, 2013

[8] C. Kornreich, P. Philippot, “Dysfunctions of Facial Emotion Recognition in Adult Neuropsychiatric Disorders: Influence on Interpersonal Difficulties” Psychologica belgica 46(1-2), 79-98, 2006

[9] M. S. Hossain, G. Muhammad, “An Emotion Recognition System for Mobile Applications” in IEEE Access, vol. 5, pp. 2281-2287, 2017

[10] M. Abomhara, G. M. Køien, “Security and Privacy in The Internet of Things: Current Status and Open Issues”, International Conference on Privacy and Security in Mobile Systems, Aalborg, pp. 1-8. 2014

[11] S.Z. Nielsen, R. Gade, T.B. Moeslund, H. Skov-Petersen, “Taking the Temperature of Pedestrian Movement”, in: Public Spaces”, Transportation Research Procedia. 2014 Jan 1; 2:660-8.

[12] S. Wioleta, “Using Physiological Signals for Emotion Recognition”, In: Human System Interaction (HSI), 2013 The 6th International Conference on 2013 Jun 6 (pp. 556-561). IEEE.

[13] X. Li, X. Du, Y. Zhang, L. Ying, C. Li, “Research on The Performance Comparing and Building of Affective Computing Database Based on Physiological Parameters”, Journal of Biomedical Engineering, 31(4), pp.782-787, 2014

[14] S. Wang, Z. Liu, S. Lv, Y. Lv, G. Wu, P. Peng, F. Chen, X. Wang, “A Natural Visible and Infrared Facial Expression Database for Expression Recognition and Emotion Inference”, IEEE Transactions on Multimedia, 2010, 682–691

[15] S. Petridis, B. Martinez, M. Pantic, “The Mahnob Laughter Database”, Image and Vision Computing Journal, 2013, 31(2): 186–202

[16] A. Selinger, D. A. Socolinsky, “Appearance-Based Facial Recognition Using Visible and Thermal Imagery: A Comparative Study” Technical report, DTIC Document, 2006

[17] B. Hernández, G. Olague, R. Hammoud, L. Trujillo, E. Romero, “Visual Learning of Texture Descriptors for Facial Expression Recognition in Thermal Imagery”, Computer Vision and Image Understanding, 2007

[18] Y. Yoshitomi, “Facial Expression Recognition for Speaker Using Thermal Image Processing and Speech Recognition System”, in: Proceedings of the 10th WSEAS International Conference on Applied Computer Science. 2010, 182–186

[19] W. Shangfei, H. E. Menghua, G. Zhen, H. E. Shan, J. I. Qiang, “Emotion Recognition from Thermal Infrared Images Using Deep Boltzmann Machine”, Front. Comput. Sci., 2014, 8(4): 609–618

[20] K. Takahashi, “Remarks on Computational Emotion Recognition from Vital Information” Proc. of 6th International Symposium on Image and Signal Processing and Analysis, Salzburg 2009, pp.299-304.

[21] H. W. Guo, Y. S. Huang, C. H. Lin, J. C. Chien, K. Haraikawa, J. S. Shieh, “Heart Rate Variability Signal Features for Emotion Recognition by Using Principal Component Analysis and Support Vectors Machine” 2016 IEEE 16th International Conference on Bioinformatics and Bioengineering (BIBE), Taichung, pp. 274-277

[22] D. Cardone, A. Merla, “The Thermal Dimension of Psychophysiological and Emotional Responses Revealed by Thermal Infrared Imaging” 2014 IEEE International Conference on Image Processing (ICIP), Paris, 2014, pp. 1942-1946.

[23] B. Manini, D. Cardone, S. J. Ebisch, D. Bafunno, T. Aureli, A. Merla, “Mom Feels What Her Child Feels: Thermal Signatures of Vicarious Autonomic Response While Watching Children in a Stressful Situation”, Frontiers in Human Neuroscience, 7 (299), 1-10, 2013

[24] R. McCall, N. Louveton, J. Rumiński, D2.1 The Specification and Overall Requirements of the eGlasses Platform. Technical Report, Univ. of Luxembourg, [http://orbilu.uni.lu/handle/10993/16763], (ISBN: 978-2-87971-125-6), Accessed: Nov. 25, 2017

[25] J. Ruminski, “The Accuracy of Pulse Rate Estimation from The Sequence of Face Images” in Human System Interactions (HSI), 9th International Conference on (pp. 518-524). IEEE, 2016

[26] Videos used in the work for Emotion Invocation Studies https://www.youtube.com/playlist?list=PLFcEURt_ucKjnzRzjmAV45J8ryMlFvmeg, Accessed: Jan. 31st, 2018

[27] J. Ruminski, A. Kwasniewska, “Evaluation of Respiration Rate Using Thermal Imaging in Mobile Conditions” Application of Infrared to Biomedical Sciences, pp. 311-346. Springer Singapore 2017.

[28] J. Ruminski, “Analysis of The Parameters of Respiration Patterns Extracted from Thermal Image Sequences”, in Biocybernetics and Biomedical Engineering. -Vol. 36, iss. 4, s.731-741, 2016.

[29] J. Ruminski, “Reliability of Pulse Measurements in Videoplethysmography”, in Metrology and Measurement Systems, Vol. 23, iss. 3, s.359-371, 2016

[30] Microsoft Emotion Cognitive Service: https://azure.microsoft.com/en-us/services/cognitive-services; Accessed: Jan. 15th, 2018

[31] A. Kwasniewska, J. Ruminski, P. Rad, “Deep Features Class Activation Map for Thermal Face Detection and Tracking”, in Human System Interactions (HSI) 2017 10th International Conference on 2017 Jul 17 (pp. 41-47). IEEE.

[32] A. Kwasniewska, J. Ruminski, K. Czuszynski, M. Szankin, “Real-time Facial Features Detection from Low Resolution Thermal Images with Deep Classification Models”, Journal of Medical Imaging and Health Informatics 2018, in print