EEMCS Faculty
Pervasive Systems Group

Device-free monitoring of vital signs to support emotion-aware music systems for people with dementia

Jessica G. J. Vuijk

Committee:
Prof. Dr. P.J.M. Havinga
J. Klein Brinke MSc.
Dr. L.M.A. Braakman - Jansen
Dr. Ir. D. Reidsma

Master Thesis Project

May 2022

UNIVERSITY
OF TWENTE.

Preface and acknowledgements

The project has been composed in consultation with my daily supervisor Jeroen Klein Brinke, based on my own interests and work/interests of Jeroen Klein Brinke at the Pervasive Systems group. I would like to thank him for allowing me to expand my existing skills and develop new ones by working on this project while receiving excellent guidance. In addition, I would like to thank the co-supervisors Nikita Sharma and Christian Wrede, as well as the promoters Prof. Dr. Paul Havinga, Dr. Ir. Dennis Reidsma and Dr. Annemarie Braakman-Jansen for giving me directions where needed, and providing the tools needed to perform this work. Also, the members of the Pervasive Systems group deserve a big thank you. They have shown their interest in my work, listened to my intentions and reflected on the work I performed, providing me with food for thought. Jeanne Parmentier specifically helped me create better quality figures that were also more interpretable. All different perspectives enabled me to improve the quality of this work.

Another thank you goes out to the student that has integrated multiple sensors into the ROPPO chair used in the experiment of this thesis. Moreover, I would like to thank all individuals that have taken a seat in this chair to participate in my pilot study or experiment. Without their contribution, a large portion of this work would not have been possible. Next to that, I am also very thankful for all the respondents to my questionnaires providing me with information essential to construct my experiment.

Moreover, I am grateful for the EEMCS Thesis support group, and especially the sub-group that I have been a part of. Frequent study sessions and setting (realistic) goals together have helped me a lot in staying on track. Suggestions for challenges and structure have served me well.

To my friends and family: I hugely appreciate the support that you have given me. Many of you have been amazing ”rubber ducks”. In listening to my cluttered brainwaves, you have untangled them and inspired alternative solutions to explore. Not only did you help me with the work itself, but the kind and supporting words have also motivated me and aided me to push through.

Specific thanks go to Anita, Jordy, Mayuresh, Eduard and Stijn for helping me address challenges with grammar, overleaf and/or coding.

Summary

The number of individuals diagnosed with dementia is increasing strongly and is expected to continue to do so in the years to come. Accordingly, this also increases the needed amount of healthcare and costs made for dementia. This has caused the World Health Organization to recognise dementia as a threat to society. The quality of care is at risk and new avenues to address this challenge should be explored.

This thesis report builds further on the literature study performed for the course Research Topics, and reports on the project that followed naturally from the observations within Research Topics. It was found that treatments of dementia using music have gained a lot of attention in science given their strong effects on many aspects (e.g. behaviour, character and emotions). The project completed during this thesis is a comparative study of multiple device-free measuring techniques; one in which millimeter wave (mmWave) is used, another where channel state information (CSI) is collected with wireless fidelity (WiFi), and lastly colour intensity tracking using red, green, and blue (RGB) values. This is done to discover the ability of sensors to monitor vital signs. Findings of this work could guide informed decision making in designing an emotion-aware music system to address symptoms of dementia in a cost and care efficient manner. To assess the ability of the sensors to monitor vital signs to support an emotion-aware music system for people with dementia (PWD), heart rate and breathing rate measurements have been collected while subjects were seated in a comfortable chair listening to emotion-inducing music. The emotions include happiness, sadness, anger, and fear. In addition to comparing the ability of sensors to monitor vital signs, this work also includes ethical considerations and privacy risks involved when using the sensors in treatment opportunities for PWD.

It was established that WiFi-CSI would be the best option to further explore in regards to the breathing rate. For the heart rate, mmWave sensing was found to be the most suitable candidate for future research. Additionally, results showed measurement-abilities were minimally affected by differences between participants or states. It is important that this work is continued by including elderly as participants in a similar experiment. Additionally, the sensor data should be combined with self-reported affective state labels to enable emotion-aware music provision. Eventually, this could potentially address symptoms of dementia in a cost-effective and time-efficient fashion, benefiting both the patient and the healthcare sector.

Contents

List of Figures

List of Tables

List of acronyms

bpm	beats per minute
BSS	blind source separation
CSI	channel state information
FFT	fast Fourier transform
FMCW	frequency-modulated continuous-wave
HRV	heart rate variability
MIMO	multiple-input multiple-output
mmWave	millimeter wave
OFDM	orthogonal frequency-division multiplexing
PWD	people with dementia
RAVDESS	Ryerson Audio-Visual Database of emotional Speech and Song
RGB	red, green, and blue
RMSE	root mean square error
RF	radio frequency
WiFi	wireless fidelity

Chapter 1

Introduction

A combination of lifestyle changes and innovations has caused people to grow older, causing a population where the number of people over 60 years old is larger than the number of people under 5 years old [5]. Although the extended lifespan of humans may seem like a positive change, it also comes with its challenges. Dementia is most prominent amongst elderly, and with people living longer, this also means more elderly are present. Accordingly, the portion of people with dementia is growing. In 2015, when the number of patients was around 46.8 million, the costs made globally for care for the disease were estimated to be 818 USD (more than 1% of the global domestic product) [6]. The number of patients is expected to grow tremendously in the upcoming years, possibly even faster than anticipated: While in 2009 the number of patients in 2050 was expected to be 115 million [7], these expectations were intensified in 2015 to 131.5 million in 2050 [6]. With the rise of patients also comes an increase in costs. In the World Alzheimer Report of 2015 [6], the costs made globally for dementia are predicted to reach 2 trillion dollars already in 2030. Besides, a large number of patients puts a considerable burden on the healthcare sector: in the final stages of the disease, patients need personal care in their daily lives. With an increasing number of patients thus also comes an additional need for time and care from caregivers.

If no changes are made, one of the Sustainable Development Goals comes at risk: healthy lives and promotion of well-being for all at all ages [8]. If the need for care for dementia patients is larger than healthcare can provide, this comes at the expense of the quality of life and care of the patients. However, changes can still be made to address the challenge. The global change in dementia asks for cost-effective interventions that decrease the need for, or workload of, caregivers while still maintaining the quality of life of patients. This is something that is encouraged by the World Health Organization recognising dementia as a public health priority [9].

To address the problem that is currently unwrapping itself and will grow in the

Table 1.1: Abstract of the revealed challenges within Research Topics' literature study [4] on music treatments for PWD and emotion detection

Topic	Challenges and observations
Music and dementia	A lack of methodological rigour, need for standardisation. Limited research on long-term effects. Focus on western cultures, limited knowledge of other cultures. No heuristics in what is the best intervention, many intervention design choices and trade-offs. Adaptability to the limitations of persons with dementia in general and the progressive nature of the disease is necessary. Strong results are obtained by using music, but uncertainty about its added benefit compared to other pleasurable events. Research often puts an additional burden on caregivers.
Emotion detection	Limited generalisability due to the specific application design and limited diversity in current studies. Sensor limitations (storage, battery consumption, occasional loss of connection). Computation and responsiveness challenges. Scalability. Understudied (side-)effects of radiation. A lack of standardisation and clear requirements in research. Research is mostly performed in the lab, limited results in natural environments. Adoption (privacy considerations and ethical concerns). Limited certainty of the findings due to the use of small sample sizes in research.

upcoming years, the first action to take is to study current knowledge and existing methodologies on (treating) dementia through systematic literature research. This step has been taken for the course Research Topics. Informed by a brainstorm session, supervisors and the intention to apply pervasive computing to the solution, the main intention of the Research Topics came to be to explore whether a solution could be found in a combination of pervasive systems and music to improve the quality of life of dementia patients, or in other words an emotion-aware music system. The literature search was segmented into three parts [4]: (1) music and the effects on people with dementia, (2) (device-free) sensing of affective states, and (3) (music) recommender systems. As the first two elements are also partially relevant for the project in this thesis report, the work has been summarised in Chapter 3. Additionally, an abstract of the resulting observations and challenges is depicted in Table 1.1. Many of these challenges and observations are relevant to keep in mind when working on this project.

When connecting the challenges of the different topics, multiple research recommendations are found. The main point is ensuring the certainty of the results and increasing generalisability. Future studies should be performed with larger sample sizes, for longer periods, and out of the lab. Diversity (age, gender, culture, etc.) of the sample should be addressed as well.

Some other recurring challenges are more common and persistent; for instance device limitations and the ever-continuing aim to have more efficient and better-performing algorithms.

Despite the challenges, the literature study [4] also showed that the combination of music interventions with pervasive computing in the application domain of dementia has potential, but has (to the author's knowledge) not been explored before. A potential solution to the challenges posed by dementia might be an emotion-aware music system. This system could provide PWD with musical items based on the affective state a person is in to ease symptoms of dementia in a cost-efficient manner, compensating for the limited time of carers while not compensating for the quality of care. The literature study also showed that many methods exist to sense emotions exist, but research on these methods is often performed with small samples and not out of the lab, limiting the generalisability of the results. To narrow down the potential emotion-detection methods and find an approach that fits the application domain, this work explores the ability of three device-free sensing methods - all in the same conditions - to monitor the vital signs. Vital sign values could in turn be used with a model to determine the affective state of a person. The focus is directed toward device-free sensing as it is less intrusive and does not rely on caregivers or PWD to remember to (make the patient) wear sensors. Putting sensors in the environment is especially fitting to the situation as dementia patients in later stages have limited mobility, meaning only a small region of space needs to be monitored and measurements will minimally be affected by motion.

The contribution of this work lies in combining three device-free sensing methods in 1 experiment, which allows for scientifically sound comparison. The specific device-free sensing methods are WiFi-CSI, colour intensity tracking using RGB values and mmWave sensing. Additionally, in contrast to previous studies, this work explores the feasibility to use mmWave sensing from behind instead of the front. This brings additional challenges, such as the heart having a lesser effect on the back, and the signal needs to move through additional fabric of the chair.

The research question that has been constructed for this work is as follows: *How can device-free sensing techniques be used to measure human vital signs to support emotion-aware music for people with dementia?*

Within science, there is ambiguity in the terminology of unobtrusiveness [10] or device-free sensing. In this work, it has been chosen to use the term device-free sensing, with which it is meant that the person that is being monitored is not wearing sensors in-/on-body, but the sensing devices are positioned in the environment instead.

The research question is divided into multiple sub-research questions:

1. What is the relevance of incorporating music in dementia treatments?

2. Which device-free sensing techniques can be utilised for human vital signs monitoring?

3. How are the different device-free sensing techniques affected by the heart rate?

4. How are the different device-free sensing techniques affected by the breathing rate?

5. How does the affective state affect the monitoring of vital signs using device-free sensing techniques?

6. How do interpersonal differences affect the device-free sensing techniques in monitoring the vital signs?

The first 2 questions will be answered by consulting the Research Topics report [4]. The third question will be answered by looking at the difference between a ground truth and the estimations. For the breathing rate, no ground truth is collected. Therefore, it is not possible to determine how far estimations are off from reality. Instead, question 4 will be answered by estimating how likely the output reflects the real respiration rate by comparing it to normal human respiration rates. Question 5 will be answered by separating the processed data per state and comparing the output. Similarly, question 6 will be addressed by splitting the data per participant and then comparing the output for 3 randomly chosen participants.

From the literature, it is known that both heart rate and respiration rate are strong indicators of emotions. Once it is known which device-free sensing techniques are able to measure heart rate and/or breathing rate accurately, the next steps can be taken in designing an emotion-aware music device.

Aside from the (estimated) accuracy, the device-free sensing methods will be compared in other aspects too. This is done to think about the bigger picture later of designing a working pervasive system that is accepted by patients with dementia and adopted by the healthcare sector. The additional elements taken into account in this work include the processing time, storage needed and privacy.

The remainder of this report is organised as follows. In Chapter 2, the different sensing techniques (WiFi-CSI, RGB values and mmWave sensing) are described more elaborately. Then, in Chapter 3, the state of the art is discussed. It is a partial summary of the literature study done during Research Topics [4] and addresses how people with dementia are affected by music and which (device-free) monitoring techniques exist to track vital signs over time. Next, Chapter 4 dives into the wishes and requirements of different stakeholders. The chapter ends with a tabular overview that presents the system requirements using the MosCoW-approach to aid prioritising. Some requirements refer to ethical considerations and privacy risks. These are discussed in Chapter 5. Chapter 8 presents the experiment that has been performed to collect the data that is needed to answer sub-research questions 3 to 6. This experiment was preceded by a questionnaire to collect songs to be used in the experiment and a pilot study to evaluate the experimental setup and data collection plan. This work is presented in Chapter 6 and Chapter 7 respectively. The results of the experiment are presented and discussed in Chapter 9. After interpreting the results, Chapter 10 provides the conclusions. The thesis report is finalised with Chapter 11, suggesting future research avenues and addressing limitations of the current study.

Chapter 2

Background

2.1 WiFi-CSI

WiFi, short for wireless fidelity, is nowadays omnipresent. It is a network that allows for wireless data traffic. One of its biggest advantages is that it is multiple-input multiple-output (MIMO). This means it can meet high demands for data traffic [1]. Data is sent from multiple transmitting antennas to multiple receiving antennas as is illustrated in Figure 2.1.

Additionally, combining MIMO with orthogonal frequency-division multiplexing (OFDM) provides CSI [1]. OFDM is a method where a signal gets decomposed into smaller pieces of signal that are sent simultaneously, but at different frequencies, allowing for more data to be processed at the same time in the same channel. This can be done in the time or frequency domain.

CSI tells one how a packet of information travels from the transmitting node, through the environment, to the receiving node on different carrier frequencies. The travel trajectory (amplitude and phase) is influenced by the environment the signal travels through, such as humans, gestures, and objects. In mathematical form, the channel frequency response can be described as:

$$H(f;t) = \sum_{n}^{N} a_n(t) e^{-j2\pi f \tau_n(t)} \tag{2.1}$$

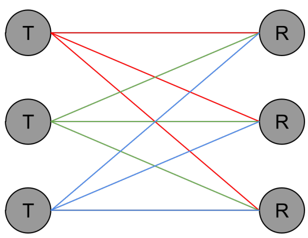

Figure 2.1: Visualisation of MIMO: T = transmitter, R = receiver

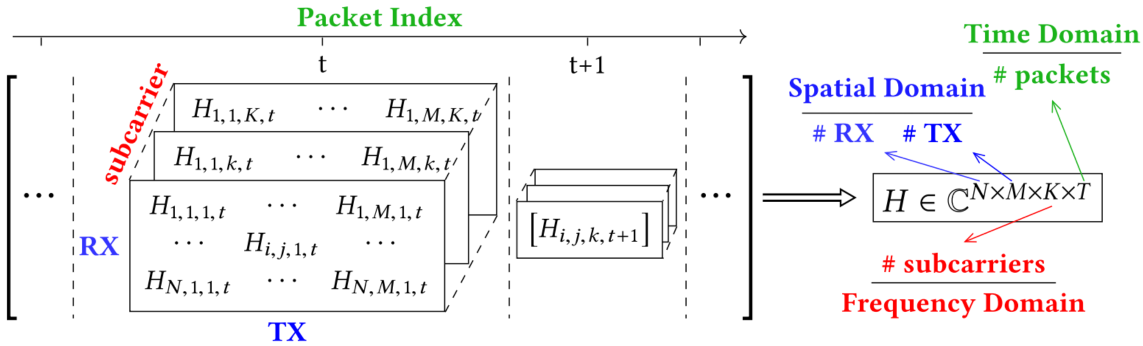

Figure 2.2: An illustration of the 4D CSI tensor: a time series of CSI matrices of MIMO-OFDM channels, by Ma *et al.* [1]. $i = i$th transmitting antenna, $j = j$th receiving antenna, k = subcarrier, t = moment in time

where $a_i(t)$ is the amplitude attenuation factor, $\tau_i(t)$ is the propagation delay, and f is the frequency of the subcarrier [1]. By tracking this information over time, this sensing method can be used for many different applications, including motion detection, activity recognition and fall detection. To be specific, WiFi-CSI allows one to track changes in the time, spatial and frequency domain. When using WiFi-CSI for a MIMO-OFDM channel, the output is a 3D-matrix with M transmitting antennas, N receiving antennas and K subcarriers that represent the change in phase and amplitude of multi-path channels. A fourth dimension is created when tracking this data over time. The shape of the output data is visualised more clearly in a figure by Ma *et al.*, see Figure 2.2.

2.2 RGB

Many methods to estimate the heart rate remotely exist, among which also one that makes use of colour intensity tracking using RGB values. RGB is an abbreviation for red, green, and blue. These are the 3 primary colours in colour coding that, when applying additive colour synthesis, will construct a colour in a frame in the digital space of a visual frame in real space [11]. Wang, Pun and Chanel [2] have performed an exhaustive survey on the topic of remotely detecting the heart rate from frontal face videos. Among the different methods studied in their paper is also colour intensity tracking. Their findings are summarised in the following sentences.

Colour intensity tracking is a method that originates from work by Poh *et al.* [12]. Although the method has been used and adapted by many researchers over time, a standard approach remains: ”Recover the heartbeat signal using blind source separation (BSS) on the temporal changes of face colour” (p. 2) [2]. In other words, this means that colour (RGB) values are being tracked for changes caused by car-

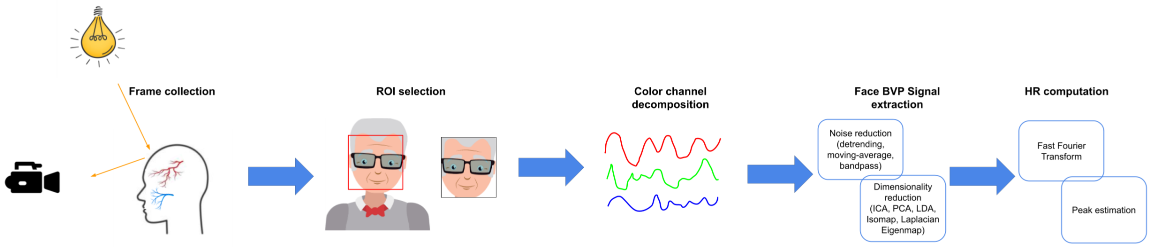

Figure 2.3: Standard procedure in extracting heart rate from RGB values based on work by Wang *et al.* [2]

diac activity. The colour intensities change over time due to the change in blood volume. Compared to other tissue in the face, blood absorbs more light. Each time the heart pumps blood through the veins, this causes a variation in volume, and in turn a change in reflection and absorption of light in the blood [13]. While this is not visible to the human eye, it is something that can be detected by computers. This method is linked to a PPG: photoplethysmogram, which is a method that detects changes in blood volume using changes in light absorption. Each individual channel, red, green or blue, provides different information. Therefore, many different methods to go from RGB values to a heart rate estimation exist. Nonetheless, each channel contains a part of PPG information. For a detailed perspective on multiple approaches of tracking colour intensities to estimate the cardiac rhythm, consult the study by Wang *et al.* [2]. A summary of the standard approach is visualised in Figure 2.3. Here, it can be seen that a region of interest, the face, is selected from the frame recorded by the camera, to which channel decomposition is applied. Next, noise and dimensionality reduction are applied. Lastly, either peak estimation or fast Fourier transform (FFT) is used to determine the heart rate. FFT is a popular method but only works under the assumption that the heart rate will be the highest power of the frequency spectrum (within the normal heart rate frequencies). It is mathematically defined as:

$$F(\omega) = \int_{-\infty}^{\infty} f(x) e^{-i\omega x dx} \tag{2.2}$$

Compared to FFT, peak estimation allows to extract more data: as is in the name, it looks for peaks within the data. This enables the extraction of the time between two consecutive peaks, which is needed to determine heart rate variability features. This is something that cannot be done with FFT as it is only applicable to a time frame, and does not allow for the detection of brief changes in the cardiac rhythm.

Performance of intensity tracking can be affected by many aspects, think about facial hair like a beard, or glasses. But not only do physical aspects matter, but also the quality of the camera and the lighting conditions matter. To eliminate such effects, experiments are often performed in strict lab settings, but these limit the

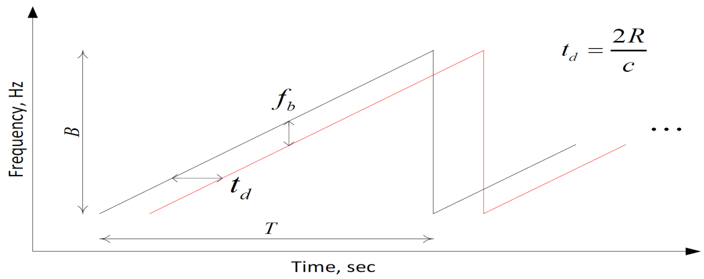

Figure 2.4: Visualisation of FMCW by Texas Instruments [3]

informational gain on how well the results translate to the real world.

2.3 MmWave

Millimeter wave, often abbreviated to mmWave, is a modern sensing technique that, without any contact, can detect an object and provide different characteristics of that object [14]. It does so by transmitting electromagnetic waves, and anything within the signal's path will reflect this wave. The analysis of this reflected signal allows extracting characteristics like velocity, range or angle.

MmWave sensing has multiple characteristics that make it applicable in versatile conditions. For example its operating spectrum is between 30 GHz and 300 GHz, allowing extremely small wavelengths to produce accurate results for small changes (in the order of millimetres) in distances [15]. Moreover, it can travel through different materials (e.g. clothing), but is unaffected by situational circumstances such as light, rain or dust [14]. Additionally, privacy risks are limited as the sensor does not collect identifiable information.

To use mmWave sensing for vital sign monitoring, the chest displacement is the information of interest. For humans, this is $1 - 12$ mm for breathing and $0.1 - 0.5$ mm for the heart rate [3]. Given these minimal displacements in space, mmWave sensing seems especially suitable to track these given its small wavelengths.

A mmWave sensor transmits a frequency-modulated continuous-wave (FMCW) signal and analyses the signal that is reflected by the object within its range [3]. The periodic linearly-increasing frequency chirps (FMCW) can mathematically be depicted as [3]:

$$s(t) = e^{j(2\pi f_c t + \pi \frac{B}{T} t^2)} \tag{2.3}$$

An illustration of what each variable reflects is given in Figure 2.4. However, by the time the signal arrives at the receiver, this is a delayed version of what was originally sent, which transforms the previous formula (Formula 2.3) as follows [3]:

$$r(t) = e^{j(2\pi f_c (t - t_d) + \pi \frac{B}{T} (t - t_d)^2)} \tag{2.4}$$

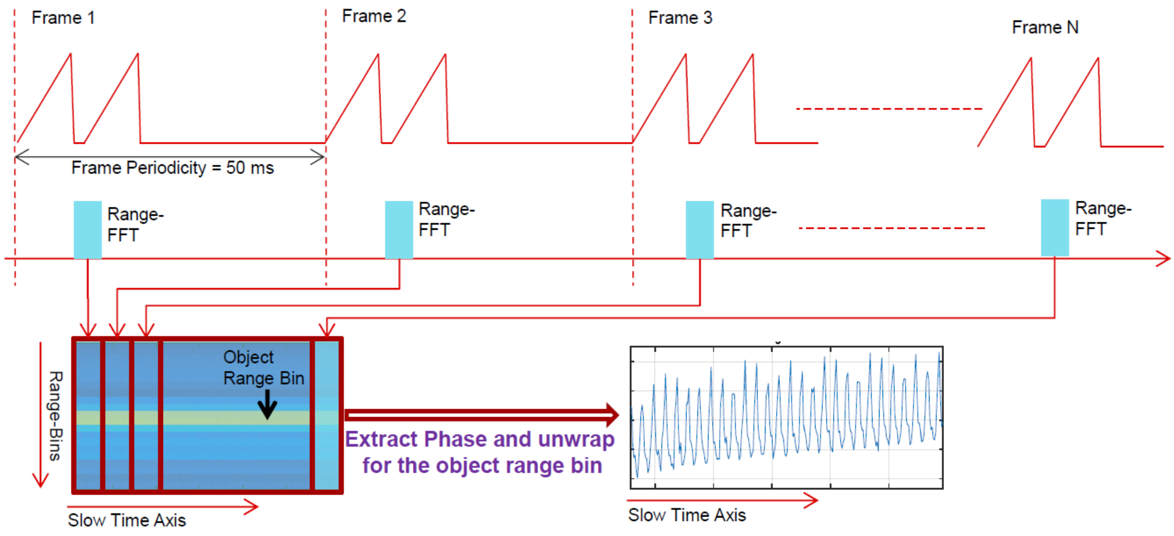

Figure 2.5: FFT extraction of FMWC signal to obtain the phase by Texas Instruments [3]

To extract a beat signal, the transmitted signal and received signal are combined as reflected in the following formula [3]:

$$b(t) = s'(t)r(t) \approx e^{j(4\pi \frac{BR}{cT} t + \frac{4\pi}{\lambda} R)} = e^{j(f_b t + \phi_b)} \tag{2.5}$$

Nonetheless, this does not resolve into a breathing rate or heart rate. To extract this data, the change in phase, caused by chest displacement, with time is needed. This can be calculated using the following formula [3]:

$$\Delta\phi_b = \frac{4\pi}{\lambda} \Delta R \tag{2.6}$$

To obtain the phase, one needs to take the FFT (Formula 2.2) of the beat signal (Formula 2.5). This process is reflected in Figure 2.5. Combining all these elements into a final formula, the equation to extract a heart rate or breathing rate can be described as:

$$x(m, nT_s) = \frac{\lambda}{4\pi} \phi_b(m, nT_s) \tag{2.7}$$

In this formula, m describes the range-bin, n the chirp index, and T_s the time between 2 consecutive measurements.

Chapter 3

State of the Art

3.1 Music and dementia

Music is something that has been studied elaborately in relation to dementia; ranging from determining what good music therapy is to how the brain of PWD responds to music [4]. The fact that it has been researched so intensively already hints at its great potential.

A first benefit of music compared to pharmacological alternatives is that pharmacological treatments often come with side effects and have shown limited effect [16, 17]. Another benefit of music is its positive contribution to people's lives. Music therapy in essence has a positive effect on people in general. It is officially defined as ”the use of music and/or its components (sound, rhythm, melody and harmony) by a qualified music therapist, in individual or group relationships in the context of a formally defined process, with the aim of facilitating and promoting communication, relationships, mobilisation, expression, organisation and other relevant therapeutic goals intended to meet physical, emotional, mental, social and cognitive needs” (p. 294) [18]. People with dementia respond particularly well to music therapy. For PWD, music therapy can help reduce secondary symptoms like a change in personality and inappropriate social behaviour [19].

Another reason to choose music is that people with dementia, even in late stages of the disease, are able to learn, recognise and respond to music [17, 20, 21, 22, 23, 24, 25]. Taking into account that patients over time lose their ability for language comprehension and production, music offers an alternative form of communication [17, 18].

The benefits of using music in treatments range even further than the aforementioned effects. Other advantages shown in the literature are: the creation of positive behaviour such as laughter [26, 27, 28, 29, 30], addressing depression symptoms [31, 32], enhancing emotional relaxation [17, 26, 30, 33] while reducing agitation [16, 26, 29, 34, 35], supporting social interactions [16, 18, 32, 33, 36], reducing behavioural disorders [17, 18, 32] and the stimulation of memories of the patient's own life [18, 37].

Possible explanations were found in the literature for these positive effects [4]. Firstly, the fact that the mechanisms that process music are closely intertwined with the pleasure and reward systems [38, 39]. Secondly, music can unveil repressed feelings with its emotional and evocative elements [16, 40].

Given these extensive, positive and strong effects of music on people with dementia, the incorporation of music in finding a solution to the challenges dementia is currently posing has gained a lot of momentum.

3.2 Device-free monitoring of human vital signs

The solution to explore in the future is an emotion-aware music system. To make the system emotion-aware, many approaches can be used as discovered in the literature study performed for Research Topics [4]. Multiple indicators of emotion were found, among which heart rate and its variability, weight, skin conductance, respiration rhythms, and motion. In this work, only respiration and cardiac activity are part of the scope. In this thesis report, a conscious decision has been made to focus on device-free sensing techniques. Some benefits of device-free monitoring are that it is less intrusive than using on-body sensors, requires no effort by the subject like would be necessary for self-report, and it does not require the subject or caregiver to memorise to wear the sensor. The following paragraphs will discuss existing literature on vital sign monitoring using WiFi-CSI, colour intensity tracking with RGB values, and mmWave sensing. Besides, Appendix A addresses some additional device-free emotion monitoring explorations, including speech and movement.

Firstly, vital sign estimations using WiFi-CSI will be addressed. This has been applied in different settings: while sleeping (e.g.[41, 42, 43]), during driving (e.g. [44, 45], in the healthcare sector (e.g. [46]) or other (e.g. [47, 48, 49]). Additionally, literature is written that combines the information of different studies: [50, 51, 52, 53]. What can be observed based on this set of literature, is that it is a rather new approach: the works are published between 2015 and 2022. This observation is also supported by the fact that PhaseBeat [47], work published in 2020, recognised themselves as the first ones to use WiFi-CSI to monitor both heart rate and respiration at the same time.

Similar to one of the challenges described in the Research Topics report for emotion detection [4], the previously mentioned works also show a lack of standardisation and requirements for reporting. For example, the work by Gu *et al.* [41] reaches an accuracy of 94.215%, while work by Liu *et al.* [42] reports the performance by providing the estimation error in beats per minute for the heart rate. Half of their estimation errors are less than 2 beats per minute (bpm) and over 90% of the estimation errors are less than 4 bpm away from the real value. Their breathing rate estimations have a mean estimation error lower than 0.4 breaths per minute, even over larger distances (5 m - 10 m). In another experiment by Ali *et al.* [54], a median error of less than 1.19 breaths per minute for real-world in-home sleep experiments was obtained. These promising findings encourage the exploration of WiFi-CSI in the situation of use intended for this work (music treatments for PWD).

In addition to the promising results for monitoring an individual, research by Gao *et al.* [49] shows that WiFi-CSI is something that could also be used to monitor the respiration of multiple individuals at the same time. They reached an accuracy of 97.5%. Also for mmWave sensing, it has been shown that multiple persons can be monitored at the same time: in a thorough study on vital sign monitoring using mmWave, Wang *et al.* [55] obtained a median error of 0.19 breaths per minute and 0.92 bpm for the breathing rate and heart rate respectively. Other works [56, 57, 58] support the finding that it is possible to monitor multiple people in parallel.

Based on the literature, it was found that mmWave has more similarities with WiFi-CSI. Firstly, it is a rather novel method as well, with publications appearing mostly appearing between 2014 and now. Moreover, it has been applied in multiple conditions (sleeping: [59], domestic or working environment: [56, 60, 61, 62, 63], healthcare: [64], or other:[55, 65, 66]), but estimating how their results translate to the situation of use studied in this paper (music treatments for PWD) is hard due to the difference in circumstances during testing and lack of standardisation. Still, literature that combines multiple works using mmWave sensing exists: [62, 67]. Work by Suzuki *et al.* [68] even targeted their research specifically at nursing homes, which is expecially relevant for this work. However, they only report specific results for their experiment with youngsters: a correlation of 0.92 with the ECG measurements.

Something that stands out within the research using mmWave sensors is that the setup gets a lot of attention: participants are frequently asked to move as little as possible given the susceptibility of the sensing method to noise, and multiple works explore different distances from the sensor, the angle towards it, and the interference of obstacles in between the sensor and the subject. However, limited research has gone to positioning the sensor behind the person of interest and at a minimal distance. This is something that this thesis report will contribute to science.

Compared to the previous methods, vital sign monitoring using colour intensity tracking is less novel, but still topical with the subset of papers in this chapter appearing between 2014 and 2021. Again, the application domain or target audience varies, but examples are frequently found in the healthcare domain. For example, work by Villarroel *et al.* [69] studies in-hospital patients and obtain a mean absolute error of 2.8 bpm for the heart rate and 2.1 breaths per minute for the majority of valid data using cameras for a longer period of time. Lyra *et al.* [70] also study patients, specifically those in the intensive care unit. The difference between their work and this work is the use of infrared compared to RGB values. Also Tarassenko *et al.* [71] apply their experiment using a camera in the healthcare domain, namely during dialysis. Also Negishi *et al.* [72] study the vital signs of patients, specifically those with seasonal influenza. However, in contrast to the previously mentioned results, they report their performance in correlations: both cardiac activity and respiration estimations had a correlation 0f 0.87 with the ground truth.

An example of work outside healthcare can be found in literature by Aubakir *et al.* [73]. They report the performance of their experiment with healthy subjects in a lab-like setting using a smartphone camera. They obtained an accuracy of 92% for cardiac activity and 94% for respiration. Also Tasli *et al.* [74] have performed work outside the hospital setting. They obtained an average error for all their videos, where participants were allowed to freely move their heads, of 4.2 bpm. Work by Hwang and Lee [75] report on their respiration estimations using RGB values using the average error. This was exceptionally small with an average error of approximately 0.1 breaths per minute.

Previously reported RGB works mostly reflect on their own work. However, work by Rouast *et al.* [76] compares multiple studies specifically monitoring the heart rate using face videos. Based on this, it can be concluded that the heart rate can be obtained using red, green, and blue values, but that performance quality varies among different strategies. The results presented in their paper range from 0.40 root mean square error (RMSE)/bpm to 7.73 RMSE/bpm.

Overall, it can be seen that research is diverse on all three methods, but also hard to compare given the different circumstances and methods of reporting. Many works also specifically ask participants to minimise movement which could mean that the reported performance is not similar to realistic scenarios. Moreover, some works only include heart rate, others only respiration, and some even both at the same time. To address this challenge, this thesis project aims to test the three device-free sensing methods in the same conditions for their ability to monitor respiration and cardiac activity at the same time. The experiment that was done to enable this fair comparison is described in detail in Chapter 8.

Chapter 4

System requirements

To design and develop a system, it is important to list the system requirements and to collect expected barriers. Only by meeting the requirements and addressing expected obstacles, the system stands a chance of user acceptance and meeting the expectations of stakeholders. The stakeholders identified in this work are people with dementia, formal and informal caregivers, society as a whole and the Pervasive Systems group.

In the existing body of literature, multiple requirements concerning unobtrusive monitoring of people with dementia can be found. In a study by Wrede *et al.* [77], the requirements of in-home monitoring of PWD from the perspective of formal and informal caregivers have been collected. Requirements identified in their research include the following: In-home monitoring of PWD should support prevention and proactive measures, prevent information overload, reduce privacy concerns, and reduce ethical concerns. However, as the aim of the system is not to communicate information about the state of a patient to others, but rather to act on the affective state of a patient directly, some of the identified requirements do not apply to this research. These include:

- The prevention of information overload: information communication (e.g. sharing the affective state of the person being monitored with relatives) is not part of the system's goal. The system's goal is to directly address the affective state of a person with dementia. It is beyond the scope of this study to explore the opportunities for information communication. Still, this could be an interesting avenue to explore in case the system's purpose gets expanded to providing caregivers (mental or physical) health status information to address additional challenges posed by dementia.

- Supporting prevention and proactive measures: the requirements based on this theme mentioned by Wrede *et al.* [77] are only partially applicable. The requirements include: voice-based coaching functions, autonomous detection and alerting of emergencies, and recognising patterns and deviations. However, based on the purpose of this system and its pervasive design, it does not include vocal coaching. On the other hand, pattern detection could be used to monitor changes in the vital signs in relation to the affective state to provide smart music recommendations. Still, the information that is gathered will, at this stage, not be used to alarm caregivers. The purpose is to directly address the vital signs. Nonetheless, vital sign status information could be extremely relevant to share with caregivers and should be something to consider in system extensions.

Next to the requirements listed in their work [77], the paper also addresses the expected barriers. These include information overload, privacy concerns and ethical concerns. While the first expected barrier does not apply, the latter 2 will be addressed in more detail in Chapter 5.

Lastly, their work [77] also contains a table that reflects the judgements of both formal and informal caregivers on specific unobtrusive in-home monitoring goals. Among the goals that score high are addressing agitation, apathy, and the positive and negative emotional state. Knowing that music has proven its ability to affect these elements in people with dementia, and that this is very relevant according to caregivers, this emphasises the relevance of this work and its future elaborations where a system is being designed that implements music to treat symptoms of dementia.

The work of Wrede *et al.* [77] is complemented by multiple works of others [78, 79, 80, 81] which have studied the perspectives and opinions of healthy (older) adults towards in-home monitoring of older adults. The general consensus amongst these papers is that the technology will be accepted if the usefulness of monitoring is clear. Once the (older) adults see the benefits of being monitored and sharing this data, such as maintaining independence, enhanced family communication, and detection of changes and acting on those, they have an accepting attitude towards monitoring systems [78, 79, 80]. Undesirable aspects (e.g. high costs, privacy, and usability barriers) should be circumvented or solved by the designers [79, 80].

To specify how the given findings relate to this thesis, a list has been composed based on the MoSCoW-method [82]. This is an approach to prioritise the requirements of the system to be designed for the intended use case.

Table 4.1: MoSCoW-list of requirements for designing an emotion-aware music system

Priority	Requirement
Must have	The system must determine the affective state of PWD
	The system must provide smart music recommendations
	The system must be multi-modal
	The system must be device-free
Should have	Models should have an accuracy above chance on test sets
	The system should have a market-ready design
	The system should be affordable
	The system should support prevention and proactive measures
	The system should reduce privacy concerns
	The system should reduce ethical concerns
	The purpose/usability of the system should be explanatory and outweigh privacy concerns
	The system should circumvent usability barriers
Could have	Models could have an accuracy above 65% on test sets
	The system could be controllable using voice-commands
	Measure the affective state and provide smart music recommendations to a larger audience (e.g. students)
Will not have	The system will not enable data-sharing

Chapter 5

Ethical considerations

Many ethical questions come along with designing a system that affects the emotional state of a person. For example, is it ethical to influence a person's mental state? As mentioned earlier, the intention of designing such a system is to improve the quality of life of a patient in a cost and time efficient manner. However, it is most prominently difficult to define what a good quality of life is for a person with dementia. Additionally, the assumption is made that improved quality of life can be found in improving the affective state, where improving the affective state refers to bringing the person to a more positive emotional state. Nonetheless, research [83] has shown that there are also good things about being in a bad state of mind at times.

In addition, using a system to address symptoms of dementia takes the human-side of care out of the picture. While this is partially the intention as time and availability of health care are limited, one may ask whether this is something we as humans should want. Society, as it is nowadays, is already more individualised compared to a few years ago [84]. Developing systems that take the human even more out of the picture could lead to elderly being more isolated than they at times already might be.

An ethical concern of another order is that the system will be used by subjects that might not have the mental abilities anymore to be aware of what the system does and how it works. In other words, they are not able to make an informed decision to determine whether the system is something that they would want to use. This is something that needs to be determined by formal or informal caregivers.

5.1 Privacy

In addition to the aforementioned ethical considerations, privacy is also a risk that needs to be taken into account. In the previous chapter, privacy has already been touched on. Privacy concerns vary between studies mentioned in Chapter 4: Although Wild *et al.* [78] found that the privacy concerns were lower than hypothesised in their work, work by Boise *et al.* [85] showed that privacy concerns are still prominent. In their work, more than half of the participants raised privacy concerns at the start and these increased during the study. In the study by Berridge and Welte [81], privacy was even the most-frequently mentioned concern. Still, overall it is concluded that the trade-off between privacy and usefulness of monitoring is won by the latter. In the paragraphs following, privacy concerns will be addressed in direct relation to the sensors that are studied in this work.

Using a camera can be perceived as very privacy sensitive. A camera can ”see” everything that happens within its frame. In this work, it specifically looks for a person’s face, which makes it easy to trace the data back to a specific individual. However, research by Chen *et al.* [86] amongst elderly has shown that the acceptance and use of a camera is context-dependent, and part of a trade-off with the benefits that a system might bring.

Work by Singh *et al.* [87] specifically mentions the conscious decision to move from cameras to mmWave sensing to decrease privacy infringements. However, radio frequency (RF)-sensing does not come without risks as mentioned by Yadav *et al.* [88].

Also, when keeping in mind the use case, intended users (people with dementia) might be more hesitant to adopt a system that contains a microphone. The intended users could consider it to be privacy sensitive, even if the only data that is being analysed is acoustics and not semantics. As mentioned in Table 4.1, a future system should reduce privacy concerns. This means that even in the case that future research points out that audio recording would be a reliable and accurate device-free sensing technique, attention should be given to ensure the user that his or her privacy is not at risk. For example in the information brochure added to the product and conversations between the seller and buyer or caregiver and patient.

An important factor to take into account is the difference between perceived privacy risks and actual privacy risks. It could be that people perceive cameras as more privacy-sensitive than WiFi-CSI. This could be caused by the lack of knowledge on the amount of information that could be extracted from WiFi-CSI. It is only fair to have people make use of the system when they make an informed decision. In other words, it is important to make users (or their informal/formal caregivers in this use case) aware of the privacy risks that they are facing when using the system, and what is done to make the system secure. Security of the system could for example be addressed in the method of communication of the system. Methods of communication and their advantages and disadvantages can be found in the Research Topics report [4].

Chapter 6

Questionnaire

6.1 Song collection

From the information collected in State of the Art (Chapter 3), it can be derived that music has the ability to elicit emotions. However, it has not been described which songs elicit specific affective states. As during the experiments it is needed to induce happiness, sadness, anger, and fear using music to study the ability of device-free strategies to monitor changes in the affective state of a subject, it is necessary to collect data on which musical items elicit these emotions. Therefore, a questionnaire (Google Form) has been sent out, where participants are asked to provide the title and singer(s) of 3 songs for each affective state studied in this thesis (happiness, sadness anger, and fear). Specifically, respondents were asked to provide songs that elicit the affective state in contrast to what they would listen to when they are already in a specific state of mind. Additionally, respondents have been asked which genre(s) elicit each affective state for them. The questionnaire has been shared with the Interaction Technology group chats of the University of Twente, the Pervasive Systems group, and with the social circle and family members of the author. This group has been selected as most likely they are comparable to the sample that will be collected for the experiment. The exact list of the questions can be found in Appendix B.

6.2 Results and data analysis

Once the questionnaire was closed, 41 responses had been collected. Interesting observations include that the same song was given for multiple affective states by 1 respondent (e.g. respondent 40 saying that '5 minutes' by Her induces sadness and anger for him or her). This could be an indicator that 1 song can induce multiple emotions at the same time. Equivalently, the same song was suggested for diverse affective states by different respondents (e.g. respondent 26 saying that 'Hold on'

Table 6.1: Descriptives of playlists constructed using the responses of the emotion-inducing song collecting questionnaire

Affective state	URL	# of songs	Percentage
Happiness	https://open.spotify.com/playlist/3PySJEfOEUqjl3hlJWr3Fb?si=9e31bdbba5e64eb6	122	28%
Sadness	https://open.spotify.com/playlist/2AWhMvc4Zo8h2qs3aZ867C?si=626eec80477e4052	121	27.8%
Angry	https://open.spotify.com/playlist/4xyheP9nf3J8WK8WRyUyQL?si=5d9e07f244214d4d	103	23.7%
Fearful	https://open.spotify.com/playlist/2P68Wwpdq3mfeiE8X4ltkY?si=90c82f7247ef4b90	89	20.5%

by Armin van Buuren and Davina Michelle induces sadness, while the same song induces happiness for respondent 2). This could be an indicator that the state that a song induces is personal.

Results might have been influenced by an order bias (the order of affective states was not randomised) as well as recall bias. When looking into the responses, it stands out that some popular songs are mentioned frequently and the majority of songs are modern songs. Hence, although respondents might feel a certain affective state with greater intensity by a different song, they might not have been able to recall such a song.

With these notes and observations in mind, the data analysis was started. The songs provided by the respondents of the questionnaire were put together into 4 playlists on Spotify, one for each affective state. Table 6.1 presents its descriptives. What can be noticed is that although the number of songs per playlist is relatively similar, they do not all contain the exact same number of songs. This is caused by multiple factors: some participants experienced difficulties in recalling songs for specific affective states, especially recalling songs that induce fear was hard. Moreover, not all songs were available on Spotify.

With the playlists ready, the Spotify API was used to extract the following audio features: danceability, energy, loudness, speechiness, instrumentalness, liveness, valence, and tempo for each song per playlist.

Once all data was available, a first inspection of the data took place. An outlier was discovered and removed: Curbi with Triple six. This song was the only song which had a positive loudness, while all others were negative.

The next step has been to produce the descriptives, which has been done using IBM SPSS [89]. Figure C.1 provides the descriptives per playlist.

In addition to the descriptives, the correlations between the features were printed (Table C.2). Strong correlations (> 0.5) were found between danceability and valance as well as energy and loudness. Moderate correlations (between 0.3 and 0.5) were found thrice (danceability, intrumentalness and valence) with loudness. Based on these findings, it was decided to exclude loudness in the remainder of the research.

Lastly, the ANOVA table was printed together with the measures of association, see Table C.3 and Table C.4 respectively. All audio features are significant when taking the threshold for significance at 0.05.

Then, when looking at the eta squared per feature, which depicts the effect size, it is highest for energy and valence. This finding means that the variation in affective states can most strongly be attributed to energy and valence. This observation forms a strong connection with Russel's circumplex model of affection, a continuous approach with a valence axis and arousal axis to model affect [90].

6.3 Song selection

To select the songs the following approach has been taken: Find a playlist on Spotify with songs for each individual affective state (happiness, sadness, anger or fear). Extract the audio features for this playlist. Filter down this list using the descriptives mentioned above until 2 to 4 songs are remaining. Then the author estimated based on intuition which songs would induce the state with greater intensity. The used playlists and their descriptives can be found in Table D.1.

For the songs inducing happiness, this meant that first all songs that are more than 0.1 off from the average energy and valence for happy songs were eliminated. 7 out of 80 songs remained. Then, the same was done for danceability and only 3 songs remained. Of those 3, the principal investigator selected 'Girls just wanna have fun' by Cyndi Lauper and 'You can't hurry love' by The Supremes.

For angry songs, the same start was made by eliminating the songs that are more than 0.1 off from the average energy and valence. This resulted in 33 songs remaining. Doing the same with danceability resulted in 15 songs remaining. Lastly, more than 0.5 diversion from the average speechiness was also used as a criterion to eliminate songs, leaving 4 songs for the author to choose from. 'Nightmare' by Halsey and 'Ready for it' by Taylor Swift were selected.

To select 2 sad songs, eliminating songs with more than 0.1 diversion from the average energy and valence resulted in 19 songs remaining. Doing the same for danceability left 10 songs. Then, trying 0.3 diversion on speechiness did not result in any changes. Lastly, only songs that were less than 10 away in tempo from the average were included. This resulted in 3 songs remaining. From these 'Where's my love' by SYML and 'Moral of the story' by Ashe were selected.

Finally, 2 songs were selected to induce fear by again eliminating songs that were more than 0.1 off on energy and valence. This caused 6 songs to remain. Doing the same for danceability resulted in 4 songs remaining. The principal investigator selected 'Play with fire' by Sam Tinnesz and 'Wires' by The Neighbourhood.

Chapter 7

Pilot study

Multiple approaches exist to elicit emotions, for example using narratives, recalling certain personal events or, as studied in this thesis, music. One of the elements of the pilot study includes validating whether emotions can be elicited using songs and whether the songs selected in the previous chapter have the ability to change the affective state. In case solely listening to a musical item is not enough to induce a specific emotion or cause a change in the affective state, the next alternative to be explored is the combination with the visuals. The first method is closest to the intended design/application of the system studied in this thesis and therefore increases the certainty of the findings translating to the intended situation of use. However, if it is impossible to change the affective state, the ability of the device-free strategies to monitor these changes cannot be tested. Therefore, the inclusion of visuals might be an interesting alternative. In case this is still not sufficient, these elements could be complemented by providing lyrics. The last alternative is to try to elicit these emotions using narratives, and if need be with the addition of visual stimuli. Studying whether solely audio is sufficient to change the affective state of a person or whether additional elements should be present is done using a questionnaire made in PsyToolkit [91, 92]. The questionnaire contains 10 songs (2 per affective state, except for fear which has been complemented with 2 alternative fearful songs given the observed challenge in recalling fearful songs) that are intended to induce different affective states. These songs have been selected as described in Chapter 6. The survey contains 2 trials, where each affective state is present (at least) once per trial and played in randomised order. After each song, participants are asked to answer multiple questions among which also to rate their happiness, sadness, anger and fear on a scale from 1 (not at all) to 5 (extremely). If these values have changed after listening to a song, it has been validated (under the assumption that no circumstances occur that could have changed the state) that the affective states can be induced using music. Additionally, one trial contains solely audio, whereas the other trial has additional visuals by presenting the official music video of that song. If the songs that are presented with visuals obtain significantly higher scores for the self-assessed affective state of the intended state of the song, this would be an indicator that the inclusion of visuals would add to the ability to induce the affective states. If neither trial can induce the states properly, a new survey should be constructed that includes a trial that provides lyrics and a trial that provides lyrics and visuals. Again, if this does not produce significantly better results than solely audio, it can be concluded that music combined with visuals and/or lyrics is not enough to induce the affective states for healthy adults. The last option would be to resort to reading narratives or recalling autobiographic memories. Although this diverts from the musical scope of this study, it does allow to test the device-free measuring techniques on their ability to monitor vital signs while experiencing different affective states.

The last aim of the pilot study is to test the experimental setup as a whole. This includes ensuring that there are enough power sockets and USB ports to connect the sensors and provide them with power. In addition, it should be explored whether the software elements can all run at the same time for the duration of a whole experiment (45 minutes to 1 hour).

7.1 Findings and observations

The pilot study was performed with 5 participants ($f : 2$, $m : 3$). The ages of the participants ranged from 20 to 64 years old ($m_{age} = 38$). The average time taken to complete the survey was 44.4 minutes. This is an indicator that the intended experiment procedure taking an hour is likely close to reality. What needs to be taken into account here is that the pilot study contained 2 more songs than the actual experiment will have. On the other hand, this average of 44.4 minutes does not yet include the time it takes to read the information letter, sign the consent form, give instructions and prepare the sensors (e.g. directing the camera and putting on the wristband). Hence, assuming this balances out, there is also buffer time in case technical difficulties come up.

7.1.1 Song selection and the ability to induce an affective-state

Important to keep in mind for the results of the questionnaire within the pilot study is that the participants are friends and family that have been asked to help. Possibly, this adds a desirability bias to the results. For instance, it could be that if participants expect that a song is supposed to induce anger, that they will also rate this affective state in the question where they are asked to what extent they also feel the different affective states. This is something to keep in mind when interpreting the results.

Moreover, the sample is extremely small ($N = 5$). The songs have been selected based on the survey explained before, with a larger sample size, and this pilot study merely serves as validation of whether the songs are able to induce the intended affective states. In short, given the small sample size of the pilot study, the results can only be interpreted tentatively and intuitively, but are good indicators of whether the songs deduced from the survey are appropriate to be used in the experiment.

A summary of the results of the survey is shown in Table F.1.

It is extremely challenging to conclude whether visuals add to the ability to induce music based on the available quantitative data. For each affective state, one song included visuals and the other did not. However, the mistake that has been made here is that for each participant in the pilot study, the same song per affective state included visuals or not. For a stronger certainty of interpreting the tabular data, it would have been better to alternate the song per emotion that includes visuals. Now, the ability to create a more intense experience of an affective state could also be caused by the song itself. Nevertheless, aside from the quantitative data obtained in the pilot studies, visual observations were made as well. The investigator noticed that even though visuals were presented on a tablet, participants did not always look at them. Some participants looked around, others closed their eyes. One of the participants even verbally expressed afterwards that he felt more capable of completely absorbing the music when he had his eyes closed. Combined with the intended use in mind, it has been decided to not include visuals in the experiment.

In Table F.1, the following can be observed: For the happy songs, the intended state is in line with the associated state as well as the most strongly experienced state. The same holds for the sad songs. However, the happy songs have the ability to induce the intended state with a higher intensity. Nonetheless, all 4 songs are evaluated as appropriate for the experiment. For the angry songs, a larger difference can be observed. Whereas 'Nightmare' is clearly (to the participants) an angry song and is able to induce an angry state in the participants, 'Ready for it' is less strongly associated with anger and also the different affective states lie closely together in intensity. Based on this observation, 'Ready for it' should be replaced. For the last affective state, fearful, 4 songs have been tested, 2 of them based on the survey ('Play with fire' and 'Wires'), and 2 based on the principal investigator's intuition ('Ghost forest' and 'The hidden chamber'). Interestingly, 'Play with fire' most frequently got associated with anger, but still induced fear the strongest. This contrasts with the previously mentioned desirability effect. 'Ghost forest' scored equally on the associated states for fear and sadness. However, again fear was induced the strongest. For 'Wires' both anger and sadness scored equally on the associated state, as well as the experienced intensity of the different states. Hence, this song is strongly discouraged to be used in the experiment. Lastly, 'The hidden chamber'

Table 7.1: Audio features of the songs used in the experiment to induce affective states

Affective state	Artist	Song	Danceability	Energy	Speechiness	Instrumentalness	Liveness	Valence	Tempo
Happy	Cyndi Lauper	Girls just wanna have fun	0.710	0.799	0.0382	0.001	0.349	0.725	120.372
Happy	The Superemes	You can't hurry love	0.617	0.810	0.0798	0.000	0.134	0.650	97.389
Sad	SYML	Where's my love?	0.608	0.385	0.0253	0.019	0.113	0.365	103.940
Sad	Ashe	Moral of the story	0.572	0.406	0.0427	0.000	0.102	0.265	119.812
Angry	Halsey	Nightmare	0.611	0.622	0.0734	0.000	0.359	0.538	146.110
Angry	Drowning pool	Bodies	0.656	0.932	0.0708	0.001	0.144	0.544	130.936
Fearful	John Carpenter	Halloween theme	0.636	0.466	0.0317	0.942	0.156	0.296	136.206
Fearful	John Williams	Jaws theme	0.364	0.268	0.0549	0.886	0.088	0.063	122.540

scored highest on the associated and experienced affective state for sadness, while the intended state was fear. Therefore, this song has also been rejected to be used in the experiment. In short, 1 angry song and 2 new fearful songs should be looked for.

The new fear-inducing alternatives that have been selected are 'Jaws Theme' by John Williams and 'Climbing walls' by Radiohead. 'Bodies' by Drowning Pool was selected to induce anger. Another brief questionnaire was constructed where the newly selected songs were presented. Respondents were asked the same questions as in the pilot study. Responses ($N = 7$) have been collected and are summarised in Table F.2.

Both for the 'Jaws Theme' and 'Bodies' it can be seen that the intended state is also the most frequently associated affective state. Moreover, when averaging the experienced state, it can be observed that the intended state is also the most intensively experienced affective state. Therefore, these 2 songs are accepted as songs to be used in the experiment. However, 'Climbing walls' has been rejected as most people felt sadness most intensively and also associated this state most frequently with the song. In a similar process, 'Halloween Theme' by John Carpenter has been tested and shown its ability to induce fear, see Table F.3

The audio feature values for each of the songs that confirmed its ability to induce the intended affective state are depicted in Table 7.1. These are the songs to be used in the experiment that is described in the next chapter.

7.1.2 Procedure

During the different pilot studies with the participants, multiple challenges occurred. When performing the pilot study with the first participant, the Asus Xtion Pro Live stopped recording mid-experiment. After exploring this issue further, it was found that the way the code was written overloaded the working memory of the laptop. The code has been adjusted to ease the workspace and the issue did not reoccur.

Additionally, the first pilot study was recorded with a smartphone. This recording was also ended mid-experiment, as the recording of the whole experiment took more space than was available on the phone. Therefore, a switch was made to a Panasonic camera with tripod and 64 GB SD card. After each experiment, the data of this camera and the Asus Xtion Pro Live should be transferred to a hard drive with more space to prevent any storage issues.

In the pilot study with the second participant, the ability to record audio using the tablet on which the questionnaire was completed was explored. These recordings stopped after a few minutes without a discovered explicable reason. Additionally, concerns exist because of the distance to the mouth and the effects of a person tapping the screen on the audio recording. Therefore, it was concluded that if the audio element would be included in the experiment, a microphone should be attached to the chair. Moreover, a human error was made and the WiFi-CSI data of the first participant was overwritten by the data of the second participant.

The last 3 sessions took place on the same day. For all 3 participants, something went wrong with the collection of the Empatica heart rate data. As this also went wrong for the second participant, and no visuals cues led the researcher to question the recording during the conduction of the pilot study, it is important to look at an alternative method that helps the researcher make sure data is being recorded. Whereas during the pilot study the researcher made use of the internal memory of the wristband and transferred the data using a USB cable afterwards, the method to be used during the experiments is connecting the wristband to a smartphone via BlueTooth. Using the E4 real-time app, it is possible to stream and store the data. This means that the researcher can check the app to ensure the wristband is collecting data.

For all participants, it was observed that the mmWave sensor was only occasionally and briefly able to collect breathing rate values. So, although the software allows the collection of this data, the setup as it currently is does not allow for consistent measurements. Nonetheless, it will be included in the experiment to see whether the measurement values could be accurate.

Additionally, it was found that the experiment is sensitive to human error: if the filenames of the WiFi-CSI data collection are not changed, new data will overwrite old data and this old data will be lost. Moreover, if one wants to end the Time-of-Flight data collection, the researcher should press 'ctrl+c' and not the stop button of the GUI. Otherwise, the data will not be stored. Similarly, the stream of the Asus Xtion Pro Live is not stored if you press the stop button in the IDE. Instead, one should press 'q' while having the 'OpenCV'-window open. If one forgets to check the box 'save data' in the mmWave GUI, the data collected by the mmWave sensor will also not be stored. Lastly, many steps need to be taken to have all elements running for the experiment, allowing one to easily overlook or forget something. To prevent such issues to occur, a list of steps has been created to follow during each experiment.

Chapter 8

Experiment

8.1 Data acquisition

8.1.1 Study design and sampling procedure

The experimental procedure is depicted in Figure 8.1. Preceding participation in the experiment, participants are asked to read and sign a consent form. After signing, participants are asked to put on an Empatica E4 wristband and take a seat in the chair. Once all sensors are started, the participant gets a sign that he or she can start the questionnaire. The questionnaire has been created using PsyToolkit [91, 92]. This is easy-to-use, free software often used in the psychological domain for experiments. The questionnaire starts with collecting demographical data. This is collected to get a clearer picture of the sample used in this study. Additionally, their initial affective state is collected. Then, the main focus of the experiment starts. Participants are asked to listen to a song while wearing headphones to elicit a specific affective state (happiness, sadness, anger or fear). How the songs have been selected is described in Chapters 6 and 7. After a song is finished, participants

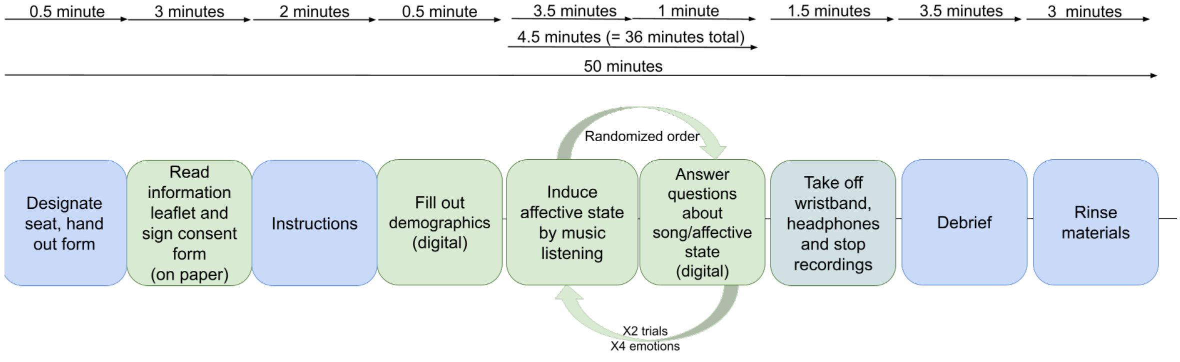

Figure 8.1: Experimental procedure

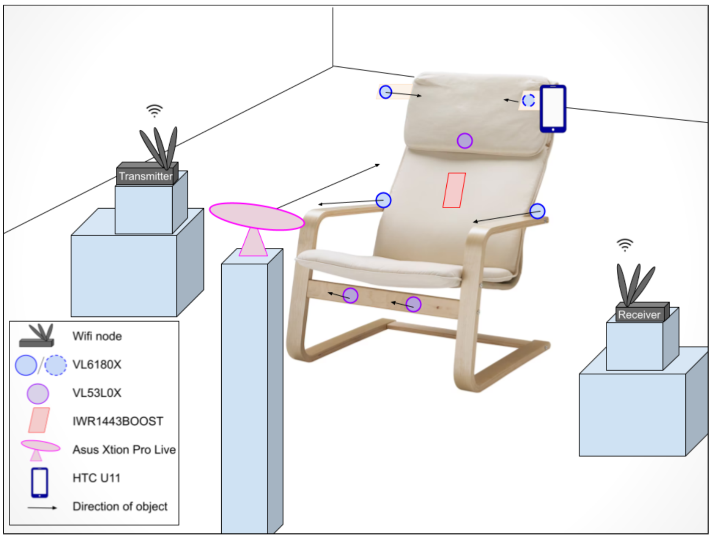

Figure 8.2: Setup of the experiment

are asked multiple questions: whether they are familiar with the song, to rate how much they like the song on a scale from 1 (like it very bad) to 10 (like it very much), to indicate to which affective state they associate the song and lastly they have to indicate to what extent they experience the 4 affective states on a scale from 1 (not at all) to 5 (extremely). This segment is repeated for each of the 4 affective states, and repeated twice to study consistency in the behaviour. Once the questionnaire is completed, all sensors are stopped. The exact list of questions can be found in Appendix E. The data will be collected and stored on the PS drive for 10 years.

The study design got ethically approved by The Ethics Committee of the University of Twente (Behavioral, Management, and Social Sciences), in accordance with the European regulations (IRB Approval Code: 211354).

The participants for this study have been collected using the BMS faculty's Test Subject Pool system SONA. This system is designed to provide (Psychology and Communication Science) bachelor students with experience in empirical research. Hence, in return for their participation, they earn credits relevant to their studies.

Participants were only allowed to partake if they were over 18 years old. Additionally, participants were excluded if they were, at the time of doing the experiment, undergoing, or had been undergoing in the last year, medication or therapy for their mental well-being. A final criterion was that people who had a hearing impairment had something to correct for this, think about hearing aids.

8.1.2 Hardware

With the songs present to induce affective states, there is also a need for hardware to monitor the vital signs. This is done using several pieces of hardware. In Figure

8.2, the location and direction of the hardware that has been used are visualised. In addition to what can be seen in the figure, a camera was directed from above and behind the Asus Xtion Pro Live towards the IKEA Pello chair. The upcoming lines will elaborate on the figure and describe the sensors in more detail.

Firstly, an HTC U11 was used to collect audio data. This audio data could be used in the future for the exploration of extracting respiration from audio files. The HTC U11 contains 4 microphones that together enable 3D audio records. The phone was attached to the chair to the left side of a subject's face.

Secondly, Time-of-Flight sensors have been put on multiple positions of the chair. More specifically, VL6180X sensors have been put at the back of both armrests and on both sides of the head. Additionally, 2 VL53L0X sensors have been put right underneath the seat near the calves and 1 in the backrest. The difference between these sensors is in the distance they are able to measure. The first measures shorter distances (5mm to 200mm) compared to the latter (30mm to 1000mm) [93]. The Time-of-Flight sensors have been connected to a Raspberry Pi 4.

Thirdly, a mmWave sensor (IWR1443BOOST by Texas Instruments[94]) has been put slightly above the centre of the backrest of the chair. The seat padding in the back has been modified to minimise the chances of a person seated in the chair noticing the sensor.

Fourthly, 2 WiFi nodes have been positioned to the sides of the chair, 160 cm apart. They are put on the same horizontal line as the knees, and vertically positioned at chest height.

Fifthly, participants were asked to wear an Empatica E4 wristband to obtain a ground truth for the heart rate.

Lastly, an Asus Xtion Pro Live was positioned in front of the chair to collect the RGB data stream. This camera was at approximately 80 cm distance from the face.

The previously mentioned tools are all used for sensing. However, additional hardware was necessary to complete the experiment. An Apple iPad was used to let participant enter their answers to the questions as well as present them with the songs. Bose Headphones 700 with noise-cancelling were used to present the participants with high-quality audio while not being affected by environmental distractions.

8.1.3 Software

To process the raw data, multiple analytical approaches are used. Figure 8.3 puts the abstracted system architectures included side to side, but the upcoming subsections will describe them in more detail. The data per sensor are combined in the end by combining rows that have matching timestamps. Note that processing methods for the Time-of-Flight sensors and audio recordings are not discussed, as

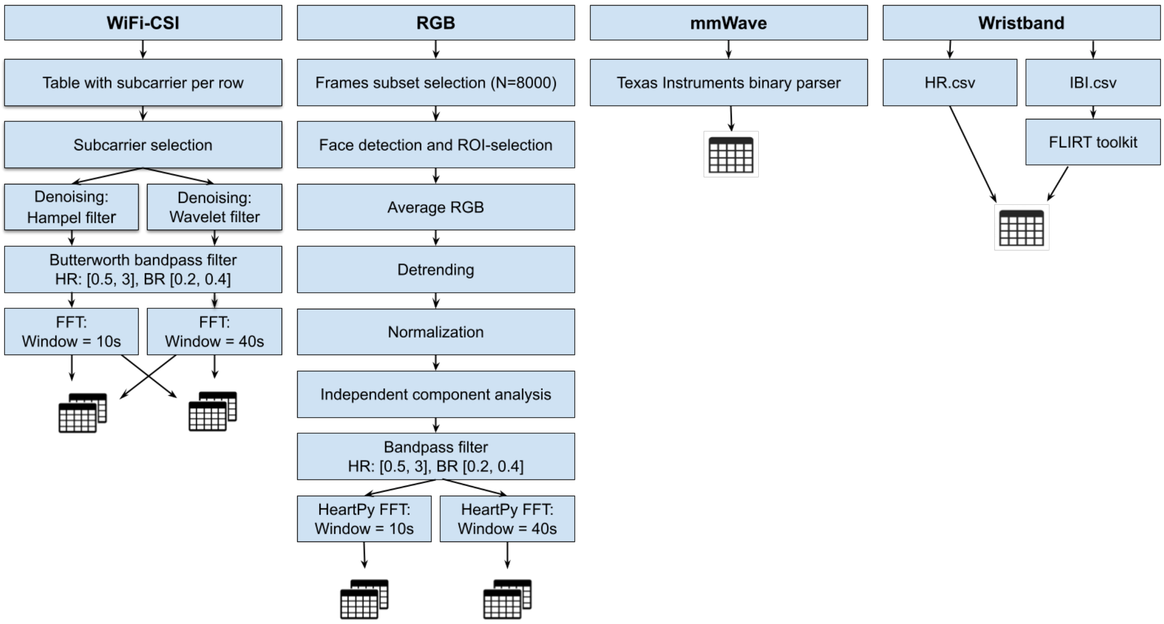

Figure 8.3: Processing methods of the raw data

these elements have been eleminated by time from analysis.

Wristband

The Empatica E4 wristband [95] needs to be connected to the desktop app E4 Manager when not using the app to extract the data from it. By connecting the wristband to the PC using a USB wire, and synchronizing the data in the app, the data gets stored on one's account. By opening the account on the empatica website, it is possible to either visually inspect the data, but in this case also most interestingly to extract the files. By pressing download, the account owner gets multiple files: ACC.csv, BVP.csv, EDA.csv, HR.csv, IBI.csv, tags.csv, TEMP.csv. Each file stores a different element of data: Acceleration, blood-vessel pulse, electric dermal activation, heart rate, inter-beat interval, time tags of self-selected important events and the temperature respectively. In this research, only the HR.csv-file and IBI.csv-file are being used. Nonetheless, the other files have been stored in case they might be interesting for other research.

The heart rate over time can be obtained directly from the HR.csv-file. However, in addition to the heart rate itself, its variability can be indicative of the affective state as well [96] and was therefore part of the processing. heart rate variability (HRV) features are described in detail by Shaffer and Ginsberg [97]. The HRV features have been extracted using the IBI.csv file. The HRV-features are extracted using FLIRT [98]. This is an open-source Python library for wearable data processing. Specifically, the function 'get_hrv_features' [99] was used. It generates features in both the time and frequency domain, as well as non-linear features and statistical features. The output of this function is merged with the heart rate based on matching timestamps and stored in a data frame.

WiFi-CSI

In this work, the following approach has been used to extract the vital signs: The nodes used during the experiment deliver 2 files per participant: A data file which contains the channel-state information, and a text file which contains the timestamps of the data. To process these files, they were opened using Matlab R2021b [100]. Jeroen Klein Brinke provided code (based on [101]), to create a table where each row represents an individual subcarrier and the columns represent the signal-to-noise ratio over time. From here on, the code has been adapted to first extract the variance per subcarrier. The strongest varying subcarrier from the original table is then selected and stored separately to continue working with it. After that, 2 data vectors are generated by applying different denoising filters. One vector is created by using a Hampel denoising filter, while the other is created using a wavelet denoising filter. The Hampel filter just looks for outliers within the 6 surrounding samples (3 in front and 3 behind), and if the sample of interest has a larger difference from the median than 3 standard deviations of that sample, it is replaced with the median [102]. The wavelet denoising filter on the other hand filters based on the knowledge that the data is in the shape of a wave [103].

The remainder of the processing is exactly similar for both vectors. The next step is to apply a Butterworth bandpass filter. To extract a heart rate data vector, the Butterworth filter that has been passed onto the denoised data has a lower cutoff frequency at 0.8 Hz (= 48 bpm) and a higher cutoff frequency at 3 Hz (180 bpm). These values are rather at the exceptional side of the heart rate spectrum but have been chosen as the vital signs are studied while experiencing (possibly intense) emotions. With the same motivation, the low and high cutoff frequencies for the breathing rate are 0.2 Hz and 0.4 Hz respectively, corresponding to 12 and 24 breaths per minute. Once the Butterworth filtering has been extracted, the Fast Fourier Transform is applied to a window of data. From that result, the highest peak is used to determine the frequency of breathing and the heart beating. The window size that has been applied is 10 seconds. This window has been chosen because the ground truth obtained from the Empatica wristband also uses a 10-second window to construct its output. In addition, the same method has been tried with a 40-second window as work by Gu *et al.* [41] mentioned how a 40-second window is needed to obtain accurate results. The values obtained from the different window sizes will be studied for their difference in accuracy later in this report. Combined with the corresponding timestamps, the data is stored in a new excel sheet.

RGB

To extract the heart rate and breathing rate from the colour stream obtained by the Asus Xtion Pro Live, for each participant a numpy-file was created with all frames (60 frames per second) obtained during the experiment.

As these files are extremely large (in the order of GigaBytes), they are processed in segments of 8000 frames. Otherwise, the working space would be overloaded. These segments have an intersection to prevent holes in the data caused by segmented processing. The processing is based on code available on GitHub by David Hass, Spencer Mullinix, Hogan Pope [104, 105]. The first step in the process after reading in frames was to detect if there is a face present in the frame. In the case a face is found, the region of interest (the face) is extracted and the average red, green, and blue values were stored. In case no face is detected, the red, green, and blue values are set to 100.

Next, the data is detrended and normalised. In turn, independent component analysis is done to determine the best component. This component is used to extract the actual blood-vessel pulse. This data is then filtered using a moving average filter as well as a bandpass filter. Similar to the WiFi-CSI processing, the high and low cut-off values are set to 0.8 and 3 for heart rate and 0.2 and 0.4 for breathing rate. The output is then processed by Python Heart Rate Analysis Toolkit. Again, segment-wise processing is used: 10-second windows are applied because this is also what is done in the Empatica wristband processing. Similarly, windows of 40 seconds are applied to compare this with the results of WiFi-CSI. The output of this toolkit is a dictionary. The dictionaries are merged for each segment of 8000 frames and then transformed into a data frame, combined with the timestamps, to be able to store it as a CSV-file.

mmWave

To measure the heart rate using mmWave, a lab from Texas Instruments was used, the vital signs lab of the mmwave_industrial_toolbox_4_9_0 [106]. Not only does this provide a graphical interface that combined with a mmWave sensor enables the monitoring, but it also provides a parser to extract the data from the binary file (main_readGUISavedBinary.m). This file has been adapted such that the output is written to a CSV-file. The process applied by Texas Instruments is reflected in Figure 8.4.

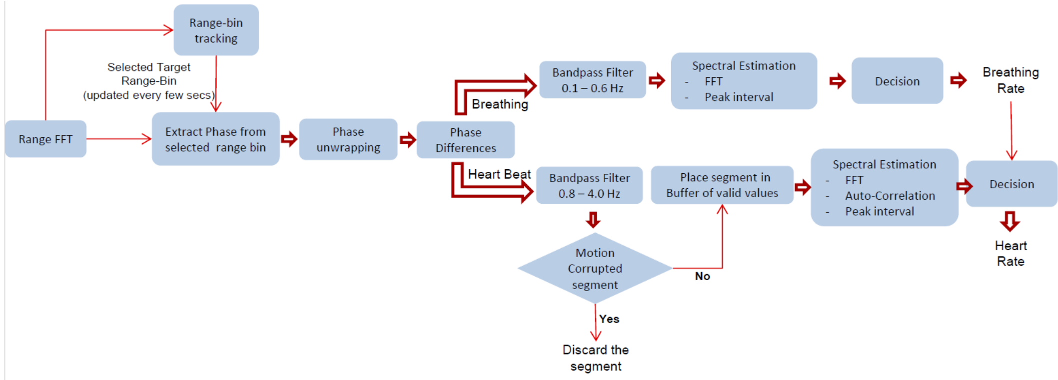

Figure 8.4: Approach by Texas Instruments [3] to extract the heart rate and breathing rate from the mmWave sensor data.

8.2 Methodology

8.2.1 Data analysis

In this experiment, multiple aspects were studied. The main focus was studying which methods of device-free monitoring are suitable to monitor heart rate and respiration. To study how well the sensors in the system architecture function, the (absolute) error between the device-free strategy and a ground truth are calculated for the heart rate. The ground truth was collected by the participant wearing an Empatica E4 wristband. For the respiration rate, no ground truth was collected to prevent drop-outs of participants because of the need to wear an intrusive respiration sensor. In this case, the accuracy is estimated by comparing the estimated values to a normal human respiration rate. This is most commonly between 12 and 15 respirations per minute [107], but has been extended to $12 - 24$ in this work, taking into account that the respiration could increase due to experiencing an affective state intensely.

The data analysis started with a first inspection of the heart rate. The online website of Empatica allows the user to inspect the data visually. An example is given in Figure 8.5. The top/red line is a visualisation of the blood vessel pulse, the middle/purple line visualises the acceleration measured by the wristband and the bottom/orange line is the heart rate extracted from interbeat-intervals. This kind of visualisation allows for an easy first inspection of the data. In this case, a blue line has been added by the author to the figure at a point where the data switches from noisy to normal. This can be explained by the fact that the wristband was put on while it was already sensing and before starting the questionnaire. Another observation is the regular peaks in acceleration: These can be explained by participants sitting back and relaxing while listening to the song and then moving again to an-

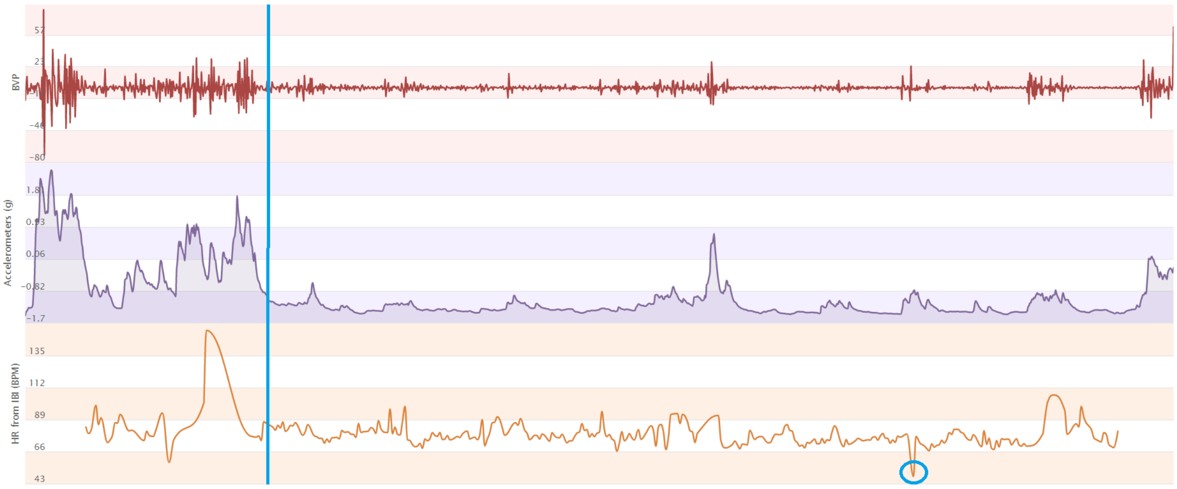

Figure 8.5: Visualisation wristband data

swer the questions on the tablet before/after the songs. Both previously mentioned observations are expected for all participants. The interesting part lies in observing outliers, as for example illustrated in the valley accentuated by the circle in the figure. No extreme acceleration has been detected, but the heart rate is exceptionally low.

After the first inspection, the actual data processing starts. The ground truth heart rate is subtracted from the estimated heart rate. This allows one to see whether the output of the sensing techniques over- or underestimate the heart rate. In addition to the error, the absolute value of the error is also extracted. This allows easy comparison of the error size. A boxplot is generated for each method using RStudio [108] for the estimated heart rates, the errors and absolute errors to make the differences visual. Before the boxplot can be generated, the data has to be restructured. Where in the original dataset each row contains all different values for each timestamp, to generate the boxplot, this data is split into multiple rows, where each sensing method and processing method gets its own row. In addition to the visuals, the data is also extracted in detail in tabular form. To do so, the following R-functions have been used: sum, sd, var, and mean. Also, tapply's summary function is used. Elements of this data have been used to construct the values for the first and third quartile values:

$$mean \pm qt(0.975, df = sum - 1) * sd/sqrt(sum) \tag{8.1}$$

Similar to the heart rate analysis, a boxplot is also generated using RStudio [108] for the estimated breathing rate for each tested method. In this case, the (absolute) error could not be computed due to the lack of a ground truth. Again, a table is constructed with the detailed data of the estimations using the same methods as for the details of the (absolute error in) heart rate estimations.

Aside from studying the vital signs, it is also studied whether the songs used in the questionnaire affected the participants' affective states. This is done by generating boxplots of the output to the question to what extent the participants experienced the affective states induced by the song they just listened to (i.e. the intensity of the affective states). In addition to the intensity, participants were also asked to what extent they like a song and whether they are familiar with them. The effect of these on the experienced intensities of the affective states has been studied by means of the correlation and a t-test using SPSS [89].

Next to the questionnaire, interest is also shown in movement/posture in relation to the affective state of a person, see Appendix A. Since only a portion of the sensors was working during the experiment, the data has been disregarded. Still, observations that might be relevant for future research are reported in Chapter 9.

Chapter 9

Results and Discussion

The study took place in January 2022 $(12th - 28th)$ in a conference room at the University of Twente. In total, 54 participants participated $(m : 9, f : 44, o : 1, m_{age} = 20)$. The sample includes participants from different backgrounds and ethnicities but is predominantly white/European $(N = 44)$. Table G.1 provides an overview per participant and sensing technique which data file was generated successfully (note that raw data may sometimes be available, but the processed data is not).

9.1 Survey

Although 54 subjects participated in the study, the output file of the questionnaire in PsyToolkit [91, 92]. contains more rows. Incomplete rows, caused by technical difficulties have been removed, resulting in complete survey data for 53 participants.

Firstly, Figure 9.1 presents a summary of the intensity of the affective states reported by participants after listening to a song. What can be seen is that for both happy songs, the intended state was induced correctly. When looking at the sad songs, it can be seen that for 'Moral of the story', sadness is only slightly more present. For both anger-inducing songs, the average induced happiness is even higher than for anger (2.70 to 2.23 and 2.37 to 2.21). While during the pilot study, previous iterations of fear-inducing songs frequently also induced anger to a stronger extent, this happens to a lesser extent in the final 2 selected songs ('Jaws Theme' and 'Halloween Theme'). Nonetheless, what can be seen is that for all other than happiness-inducing songs, the intended state is present frequently together with happiness. A possible explanation can be found in the psychological domain for this. Within psychology, a difference is made between mood and emotion. A mood is a longer-lasting state, usually lasting for hours or days without a clear reason, while emotions are short-lasting, usually minutes and can often be assigned a specific cause [109]. This can be translated to this work by looking at the self-assessed initial affective states of the participants. In Figure H.1, the initial state of participants is

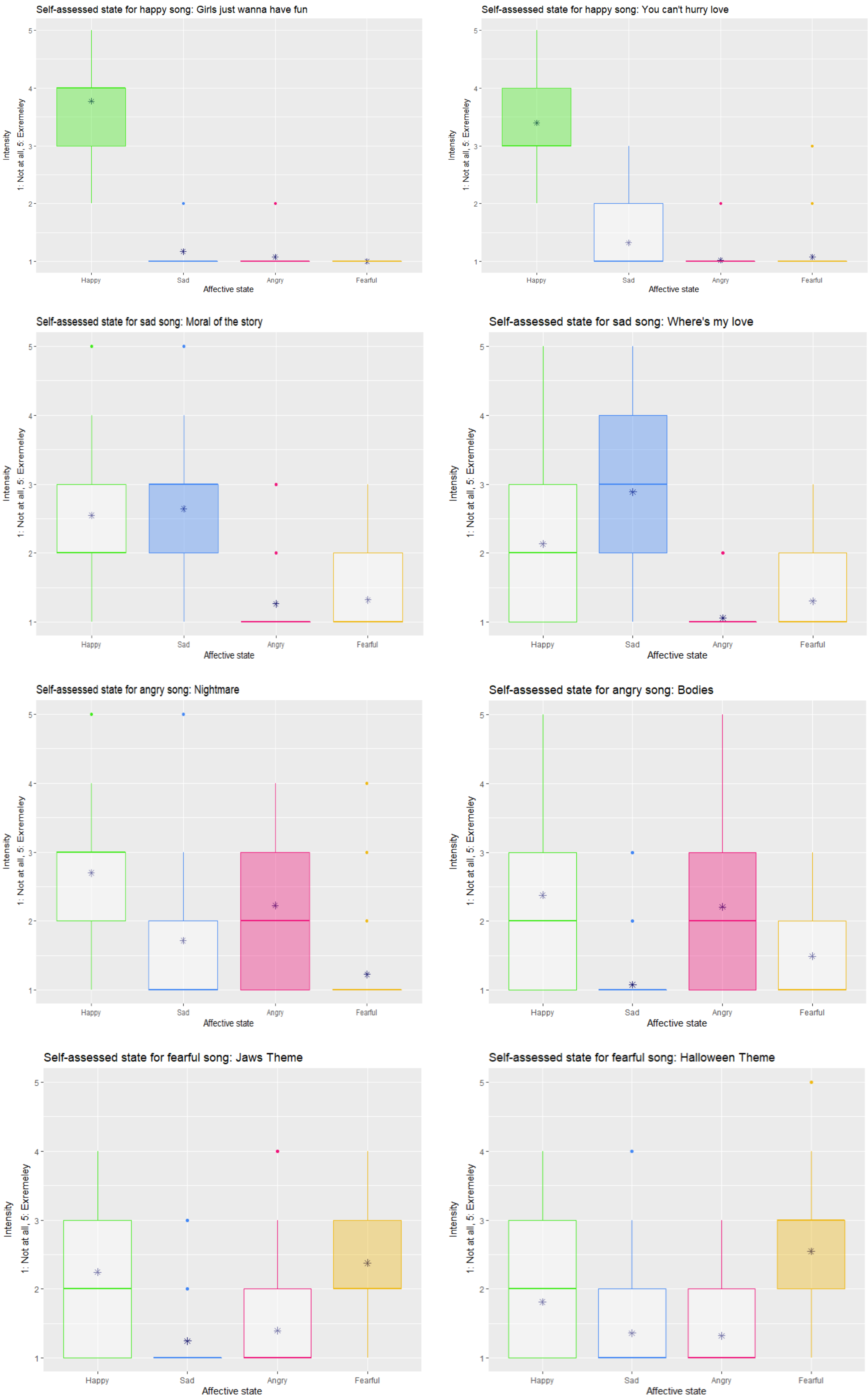

Figure 9.1: Self-reported affective state per song over all participants. The boxplot of the intended state is highlighted and the mean is marked with a star

shown. The most strongly present state in this figure is happiness. One might argue that this initial state might refer to the mood present in the participants, and that what is visible in Figure 9.1 is that the emotion caused by the song is present in combination with the mood a participant is in. Yet, another philosophical question that could be asked is whether what we feel matches the state that our body signals. In other words, do people correctly interpret their own affective states? This closely relates to 3 major theories of emotion: James-Lange Theory, Cannon-Bard Theory and Schachter and Singer's Two-Factor theory [110]. All 3 theories argue about physiological arousal and emotions. While the first states that emotion occurs after the body is aroused, the second theory argues that these happen at the same time. The last theory argues that physiological arousal gets a label first, and then the emotion is experienced. Hence, even within the psychology domain, there is no clear answer to the question on the relation between what our body signals and the label a person gives to the emotion he or she is experiencing. In case that future research finds out that what people think they feel and what they are actually feeling is not in line, self-report of the emotional state cannot be used as a ground truth in research as labels for the emotional state.

Another element of research applied to this survey element within the experiment is the effect of familiarity on the experienced intensity of the intended state. Based on the information present in Table 9.1, this seems not to be the case: Levene's test for equality for variances was not significant for all songs but 1: 'Where's my love'. However, the t-test for equality of means for this song leads to another insignificant value, 0.831, so it cannot be concluded that the 2 groups differ. In short, this means that the intensity of the intended affective state does not differ for listeners that are familiar with the song compared to the ones that are not.

Another effect that is explored in this work is the relation between how much a person likes a song and the intensity of the intended state. This is done by means of Pearson's correlation, resulting only in significant values for the happy songs 'Girls just wanna have fun' and 'You can't hurry love'. When using a scale where a high degree is between ± 0.50 and ± 1.00, a moderate degree is between ± 0.30 and ± 0.49, and a low degree is below ± 0.29, both songs have a correlation of high degree (0.502 and 0.603 resp.) between the intended state and the extent to which the participant liked the song.

9.2 Heart rate measurements

The heart rate estimations for each sensor and method are depicted in Figure 9.2. Outliers are disregarded in this figure to ease visual interpretation. Refer to Figure I.1 to inspect the outliers as well. Along with Figure 9.2, Table 9.2 presents an

Table 9.1: Exploration of the effects of familiarity and appreciation on the intensity of emotion

Emotion	Song	F/U	Levene's test	t-test	Sign r	r
Happy	Girls just wanna have fun	51/2	0.733	-	0.000	0.502*
Happy	You can't hurry love	39/14	0.291	-	0.000	0.603*
Sad	Moral of the story	31/22	0.819	-	0.703	-0.054
Sad	Where's my love	19/34	0.000	0.831	0.740	0.047
Angry	Nightmare	15/38	0.393	-	0.938	-0.110
Angry	Bodies	18/35	0.847	-	0.055	-0.265
Fearful	Jaws theme	25/28	0.442	-	0.345	-0.132
Fearful	Halloween theme	23/30	0.705	-	-0.596	-0.074

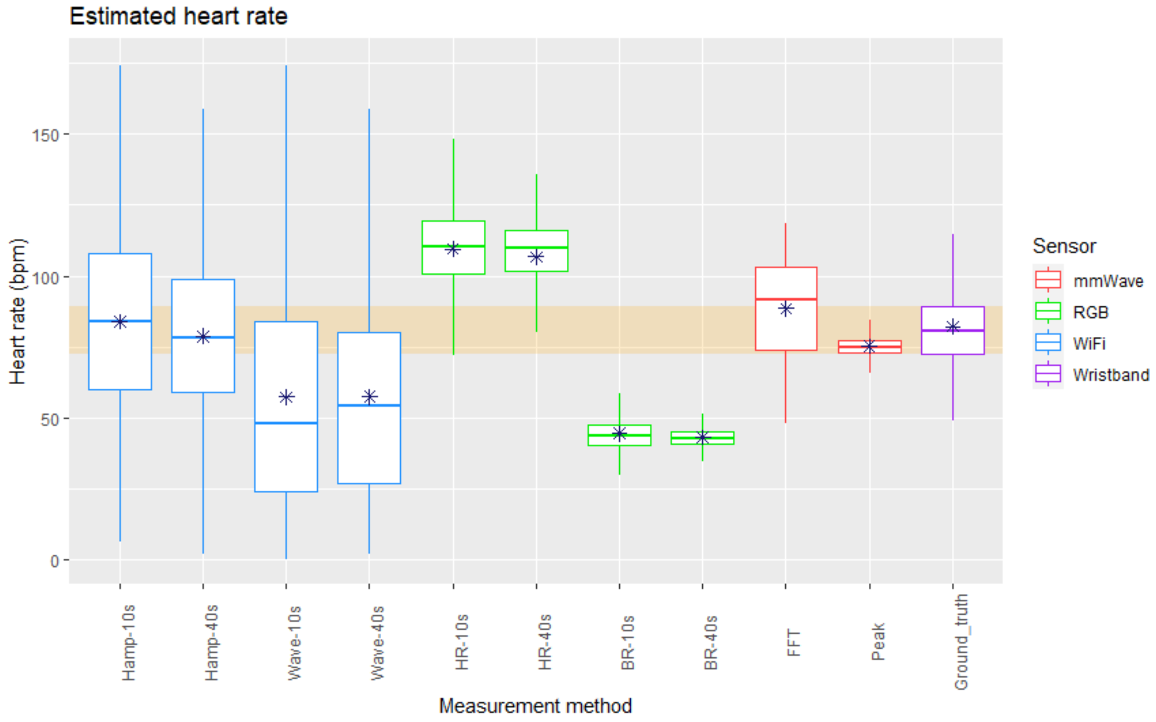

Figure 9.2: Heart rate estimations when disregarding outliers. The interquartile range of the ground truth heart rate is highlighted and the mean is marked with a star

Table 9.2: Summary descriptives of estimated heart rates, abstract of Table I.1

Sensor	Approach	Mean	Median	Variance	Std. Deviation	Min	Max	Range	Interquartile range
Wristband	-	82.32	80.40	214.85	14.66	49.00	187.93	138.93	16.59
WiFi-CSI	Hampel-10s	85.48	84.00	1378.79	37.13	6.00	282.00	276.00	48.00
WiFi-CSI	Hampel-40s	79.13	78.00	1378.13	31.89	2.00	588.00	586.00	40.00
WiFi-CSI	Wavelet-10s	58.47	54.00	1508.17	38.84	0.00	438.00	438.00	60.00
WiFi-CSI	Wavelet-40s	57.65	54.00	1176.27	34.30	2.00	233.00	231.00	53.00
RGB	BR-filter-10s	44.41	43.64	39.48	6.28	28.80	128.57	99.77	7.30
RGB	BR-filter-40s	43.09	42.86	11.55	3.40	32.14	70.59	38.45	4.27
RGB	HR-filter-10s	109.54	110.49	231.71	15.22	30.51	514.29	483.78	19.05
RGB	HR-filter-40s	106.76	109.74	197.92	14.07	31.03	135.89	104.86	14.23
mmWave	Peak	75.21	75.00	13.12	3.62	60.94	89.06	28.12	4.68
mmWave	FFT	88.50	91.41	321.69	17.94	48.05	118.36	70.31	29.29

abstracted summary of the descriptives of the sensing methods.

A first inspection of the heart rate estimations based on Figure 9.2 and Table 9.2 tells that the WiFi methods produce estimations in much larger ranges than the other sensing methods, both over and under regular heart rate ranges. Additionally, a difference can be seen between the different filters for the WiFi sensor: While the Hampel filter's interquartile ranges centralise (60.00 to 108.00 and 59.00 to 99.00) in the same range as the wristband, the wavelet filter's interquartile ranges (24.00 to 84.00 and 27.00 to 80.00) fall below this.

When doing a first inspection of the estimations of the RGB sensor, a large difference can be found between the different methods: While the RGB methods using a heart rate range filter tend to overestimate the heart rate, RGB methods using a breathing-rate range filter always underestimate the heart rate. Moreover, the breathing rate filter estimations are extremely small in range with interquartile ranges 4.27 and 7.30 This is even smaller than that of the ground truth: 16.59. The size of the ranges of the heart rate range filters (interquartile ranges: 14.23 and 19.05) are comparable to that of the ground truth, yet not between the same values.

Lastly, the mmWave sensor produces a clear difference in estimations. While the use of peak estimation results in estimations in an extremely small range (28.12), this is much larger for the use of FFT (70.31). Additionally, it can be seen that FFT tends to overestimate the heart rate ($m = 88.50$), while peak estimation tends to underestimate the heart rate ($m = 75.00$) compared to the ground truth ($m = 82.32$).

Overall, solely based on the heart rate estimations themselves, the mmWave sensor using FFT seems like the best option to choose, but peak estimation also seems relatively good. However, this figure and table only give a first and overall impression of the (ranges of) estimations. It provides limited information on how far off individual estimations are and the tendency to over- or underestimate if the estimations fall within the expected range. To evaluate this, Figure 9.3 was produced, also telling more about the precision of the estimations. Again, this figure

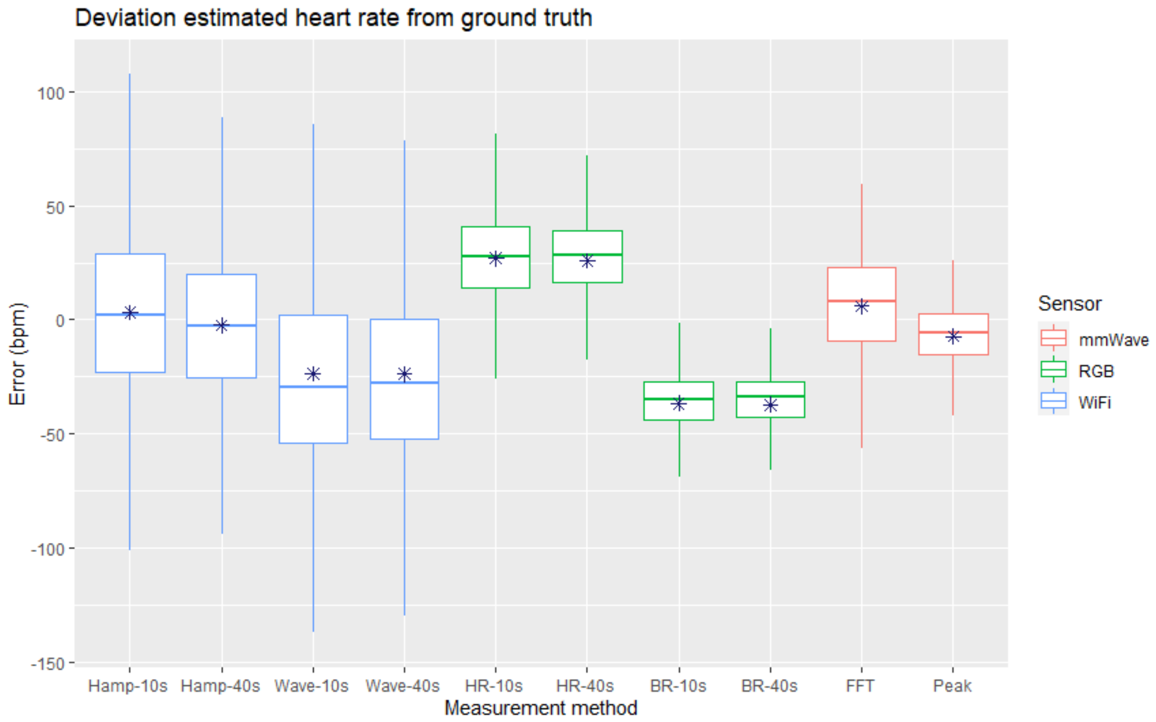

Figure 9.3: Error (bpm) in estimating the heart rate per sensor and processing method when disregarding outliers. The mean is marked with a star

disregards outliers, but a figure including the outliers is available in the Appendix (Figure I.2. Again, the figure is supported by Table 9.3. From this information, it can be concluded that the WiFi sensor again produces values in the widest range (-166.58 to 498.42), meaning precision is low. The precision of the other methods is comparable: all interquartile ranges are between 15.62 and 31.78, while preferably it is near zero. Also, the ranges are comparable if outliers are ignored.

Since the means of the Hampel filter are near zero (4.20 and -2.17), it can be concluded that the Hampel methods do not have a tendency to overestimate or underestimate. Similarly, the mmWave sensing methods also do not strongly tend to one side ($m = -7.27$ and $m = 6.02$). On the other hand, supporting the previous first impression, the RGB methods and the WiFi methods using a Wavelet filter do have a strong tendency to either underestimate or overestimate: For the wavelet filter ($m = -22.81$ and $m = -23.65$) and the breathing rate range filter ($m = -36.74$ and $m = -37.54$) this is underestimating, and for the heart rate range filter ($m = 26.96$ and $m = 25.96$) this is overestimating.

Based on the previous information, again the preference is given to monitoring the heart rate using the mmWave sensor.

Lastly, a figure was produced to visualise the absolute error (so disregarding whether estimations are above or below the actual value). Similar to before, Fig-

Table 9.3: Summary descriptives of error in estimated heart rates, abstract of Table I.2

Sensor	Approach	Mean	Median	Variance	Std. Deviation	Min	Max	Range	Interquartile range
WiFi-CSI	Hampel-10s	4.20	2.38	1552.50	39.40	-127.45	209.82	337.27	52.80
WiFi-CSI	Hampel-40s	-2.17	-2.80	1183.93	20.88	-119.40	498.42	617.82	45.87
WiFi-CSI	Wavelet-10s	-22.81	-29.22	1658.40	40.72	-166.58	351.95	518.53	56.41
WiFi-CSI	Wavelet-40s	-23.65	-27.87	1319.24	36.32	-150.27	160.90	311.17	52.16
RGB	BR-filter-10s	-36.74	-35.09	251.32	15.85	-151.05	46.57	197.62	16.90
RGB	BR-filter-40s	-37.54	-33.74	332.39	18.23	-145.20	2.82	148.02	15.62
RGB	HR-filter-10s	26.96	27.41	430.55	20.75	-100.04	447.82	547.86	26.83
RGB	HR-filter-40s	25.96	28.14	370.78	19.26	-81.16	71.77	152.93	22.70
mmWave	Peak	-7.27	-5.54	226.86	15.06	-112.37	26.00	138.37	17.89
mmWave	FFT	6.02	8.23	519.95	22.80	-120.02	59.26	179.28	31.78

Table 9.4: Summary descriptives of absolute error in estimated heart rates, abstract of Table I.3

Sensor	Approach	Mean	Median	Variance	Std. Deviation	Min	Max	Range	Interquartile range
WiFi-CSI	Hampel-10s	31.30	26.22	590.26	24.30	0.00	209.82	209.82	32.93
WiFi-CSI	Hampel-40s	27.43	23.12	435.99	20.88	0.00	498.42	498.42	29.31
WiFi-CSI	Wavelet-10s	39.28	37.82	636.04	25.22	0.00	351.95	351.95	38.91
WiFi-CSI	Wavelet-40s	36.45	34.73	550.21	23.46	0.00	160.90	160.90	38.22
RGB	BR-filter-10s	36.79	35.10	247.31	15.72	0.03	151.05	151.02	16.91
RGB	BR-filter-40s	37.54	33.74	332.37	18.23	0.43	145.20	144.77	15.62
RGB	HR-filter-10s	29.39	28.12	294.07	17.15	0.00	447.82	447.82	26.63
RGB	HR-filter-40s	28.85	28.92	208.62	14.44	0.00	81.16	81.16	21.28
mmWave	Peak	12.14	9.24	132.35	11.50	0.00	112.37	112.37	11.58
mmWave	FFT	19.35	17.24	181.72	13.48	0.00	120.02	120.02	19.25

ure 9.4 presents the results disregarding outliers, but the alternative can be found in Figure I.3. Additionally, the figure comes along with a statistics summarising table, Table 9.4. Figure 9.4/Table 9.4 reflects the recall of each method. Boxplots that are small in length and near-zero represent a method that has a high recall. Again, WiFi-CSI produces the worst results, having long boxplots meaning low recall. Nonetheless, the recall of the RGB methods is not significantly better. The recall using a breathing rate range filter is lower than that using the heart rate range filter. A similar difference can be found between the Wavelet and Hampel filters. Moreover, it is reflected that a larger window, that of 40 seconds, mostly produces smaller (absolute) errors/higher recall compared to that of a 10-second window using the same analysis method. Based on Figure 9.4 and Table 9.4, peak estimation of the mmWave sensor with an average absolute error of 12.14 bpm is the most favourable device-free sensing method when looking at the recall.

The errors between the ground truth and the estimated heart rate in this research are of a much larger order than the errors obtained in research used for vital signs extraction using WiFi-CSI mentioned previously in this work (Chapter 3). However, a difference between this work and the works mentioned previously is that others

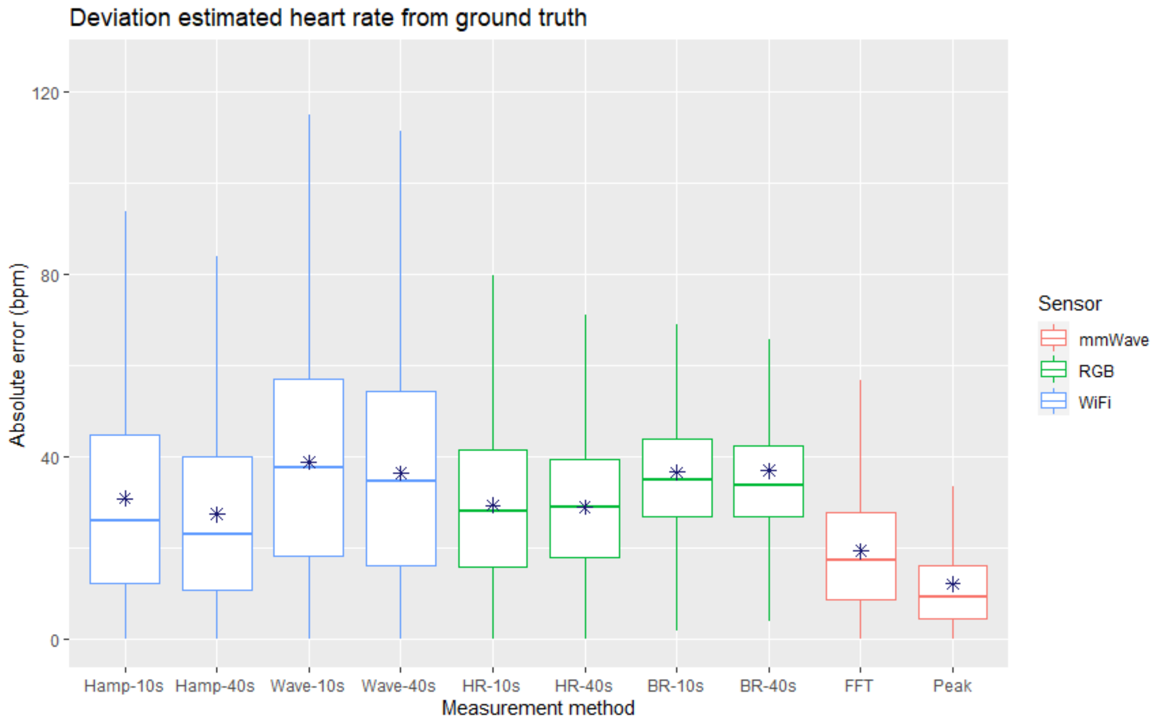

Figure 9.4: Absolute error (bpm) in estimating the heart rate per sensor and processing method when disregarding outliers. The mean is marked with a star

measured the heart rate while the subjects were sleeping, mostly lying still. In this research, participants were awake and listening to music. The music caused some participants to move along, while others actually leaned back and relaxed. Minor movements could have caused noise in the WiFi-CSI and since the physical differences are of an intense small order, minor movements could have a large effect on the results. Nonetheless, when inspecting the camera data combined with the output, it was observed that the estimations were also far off frequently when no movement was observed.

Another explanation for the results could be the manner of analysing the data and estimating the heart rate. In this work, only the Hampel filter and a wavelet filter have been explored, but more exist. Similarly, it could be that the wrong subcarrier has been selected. In this work, a decision was made based on literature to only select the most varying subcarrier. However, in some other works, a choice was made to select multiple subcarriers and average their values and use this output to determine the heart rate.

Compared to the output of WiFi-CSI, mmWave sensing produced much more realistic heart rate estimations. Still, the average absolute error is 12.14 beats per minute off. The author of this work would therefore recommend further exploring this opportunity: In this work, the sensor was put at the back of a person and close to

Table 9.5: Summary descriptives of breathing rates, abstract of Table I.4

Sensor	Approach	Mean	Median	Variance	Std. Deviation	Min	Max	Range	Interquartile range
WiFi-CSI	Hampel-10s	16.75	18.00	32.18	5.67	0.00	66.00	66.00	6.00
WiFi-CSI	Hampel-40s	16.58	17.00	16.27	4.03	2.00	41.00	39.00	6.00
WiFi-CSI	Wavelet-10s	16.08	18.00	34.26	5.85	0.00	66.00	66.00	6.00
WiFi-CSI	Wavelet-40s	16.08	17.00	17.86	4.23	2.00	60.00	58.00	4.00
RGB	BR-filter-10s	10.50	10.56	5.27	2.30	0.00	23.13	23.13	2.83
RGB	BR-filter-40s	10.71	10.47	9.97	3.16	0.00	24.55	24.55	4.30
RGB	HR-filter-10s	14.53	14.07	30.41	5.51	0.00	45.24	45.24	8.47
RGB	HR-filter-40s	13.46	13.40	20.06	4.48	0.00	33.34	33.34	6.36
mmWave	Peak	28.47	28.12	12.03	3.47	14.06	37.50	22.90	4.69
mmWave	FFT	18.63	18.75	44.55	6.67	7.03	33.98	26.95	10.55

the body, while the lab by Texas Instruments was designed for using the sensor in front of the body at a larger distance. Playing around with the position of the sensor (both with respect to the direction towards the body and the distance from the body) may lead to finding a setting that obtains better results.

Just like with WiFi-CSI, the error in estimating the heart rate using colour intensity tracking is at times far off. However, its performance lies between that of WiFi-CSI and mmWave sensing. Since the code was based on real-time heart rate estimations, the error is most likely not caused by participants moving the little amounts that they did. More likely, it could be caused by the distance of the camera: as the distance was further away to be able to position it on a table in front of the face, the region of interest used for sensing, the face, was only a small part of the frame. If this camera would be positioned closer to the face, there were more pixels to average on and the output would likely be more accurate. However, a camera even closer to the face could be experienced as more intrusive. Another risk with using RGB values is its dependence on the lighting conditions. Participants in the experiment were facing the windows, potentially overexposed to the light. On the other hand, underexposure would also cause inaccurate output.

Additionally, many methods of processing this kind of data to extract the heart rate exist. In this work, the order that was chosen was to detrend the averaged RGB values, then apply normalisation, then apply the bandpass filter and lastly use the HeartPy toolkit with FFT as a method to extract the frequency spectrum while also Welch's method was available. Another method to extract the frequency spectrum or another order of processing could possibly generate better results.

9.3 Respiration measurements

No ground truth has been collected in the experiment for the breathing rate. Hence, it is impossible to determine how accurate the measurements are based on how far

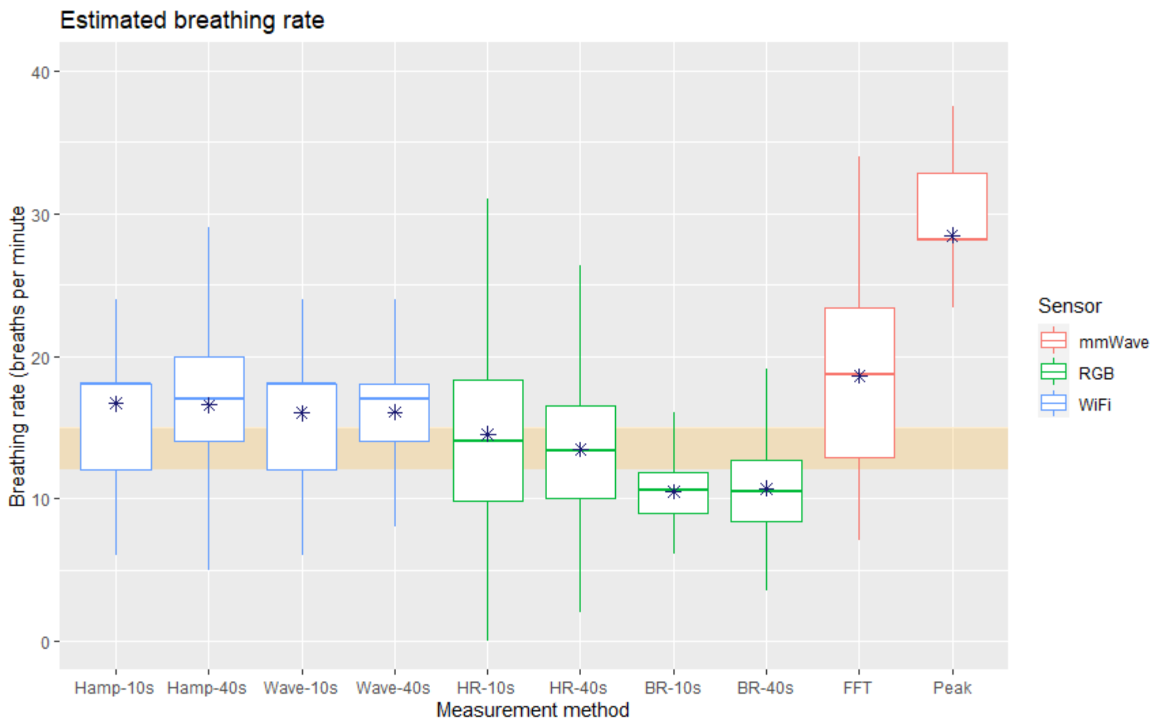

Figure 9.5: Breathing rate estimations when disregarding outliers. The range of a normal human breathing rate is highlighted and the mean is marked with a star

.

away from the ground truth they are. Nonetheless, the estimated breathing rates can be compared to a normal human breathing rate to estimate which device-free sensing method is most likely to be closest to the actual breathing rate. It should be taken into account that the range of values in the estimations is much smaller than it is for the heart rate. This means that a deviation of 1 breath per minute in monitoring respiration is worse compared to 1 beat per minute in the heart rate. In this work, the ranges that have been used to filter out the data that is relevant for breathing based on frequencies are values between 12 and 24 breaths per minute, taking into account that the respiration rate might increase when experiencing different emotions intensely. According to Radboud UMC [107], the most common respiration rate is between 12 to 15 breaths per minute, which will be used to estimate the accuracy of the breathing rate estimations.

Firstly, what stands out in Figure 9.5 is that all WiFi-CSI approaches, independent of filter and window, produce nearly comparable results. They all likely have a tendency to overestimate the breathing rate.

Unexpectedly, the Hampel filter with a 40-second window produces estimations that fall further out of the normal human respiration range compared to the same method with a 10- second window. However, the 10-second window size has higher outliers, see Figure I.4. Another unexpected result is that the heart rate range filter produces on average better estimations ($m = 14,53$ and $m = 13.46$) compared to the breathing rate range filter ($m = 10.50$ and $m = 10.71$). However, the range of estimations is much wider. In other words: recall is likely high for the heart rate range filter, but precision is low, while this is the other way around for the breathing rate range filter. Moreover, the breathing rate range filter likely causes a tendency to underestimate.

In contrast to the heart rate results, the estimations of the breathing rate using the mmWave sensor are likely strongly inaccurate. The estimations fall mostly and strongly out of the normal human respiration rate.

The estimated breathing rate of WiFi-CSI, irrespective of the filter and window size, is mostly within the human range. However, the 40-second window produces smaller outliers. Although research with a ground truth should decide how accurate estimations are, the method seems close to reality, and when off, estimations are most likely overestimated.

The estimations of mmWave sensing regarding breathing are mostly overestimated. On average, the best performing method (FFT) estimates 18.63 breaths per minute. This is strongly above the most common respiration rate mentioned previously ($12 - 15$). Moreover, these values are calculated based on the estimations that the system was able to make. However, most of the time, the Texas Instruments lab was not able to extract a breathing rate. Therefore, this approach of sensing the

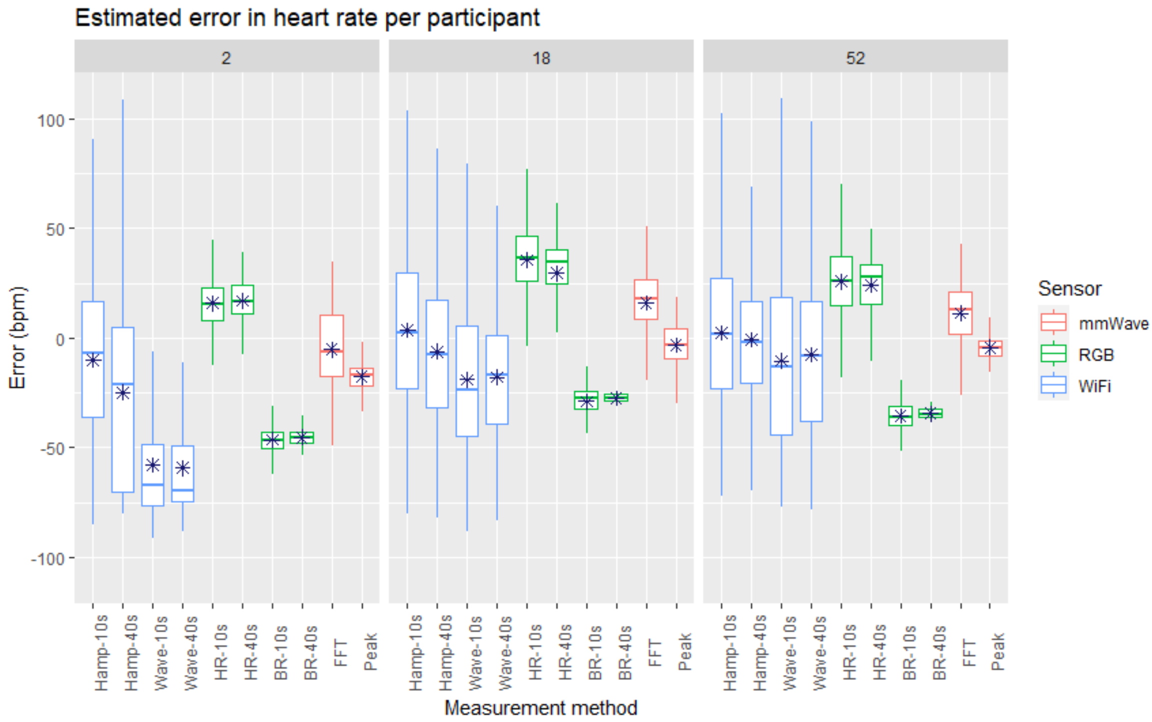

Figure 9.6: Comparison of error in estimating the heart rate over 3 participants using different sensing and processing methods when disregarding outliers. The mean is marked with a star

breathing rate is not recommended.

Although the results for the heart rate do not look promising using colour intensity tracking, they do so for the breathing rate. Using the heart rate range filter and a window size of 40 seconds resulted in estimations mostly within the normal human respiration range. Still, it remains unknown whether these estimations are accurate since no ground truth has been collected.

9.4 Effects of inter-personal differences and states on monitoring

In addition to looking at how well the different device-free monitoring techniques perform in general, it has also been explored whether the sensing techniques are possibly affected by the differences between participants or affective states. To do so, Figure 9.6 and Figure 9.7 have been generated for the heart rate, where the error is shown per processing method per state or participant. From this, it can be derived whether the performance of a sensing method is affected by a state or person. On the other hand, for the breathing rate, this is not possible since there is no ground truth. Instead, figures (Figure 9.8 and Figure 9.9) have been

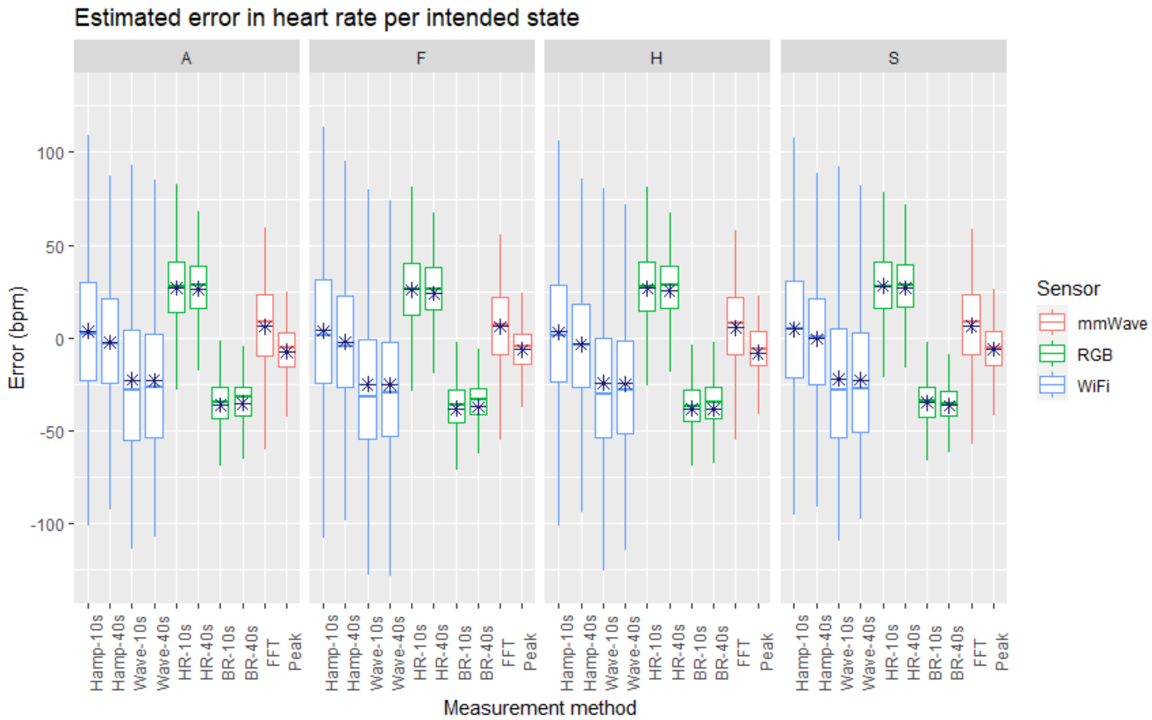

Figure 9.7: Comparison of error in estimating the heart rate per intended affective state using different sensing and processing methods when disregarding outliers. The mean is marked with a star

created that visualise the estimated breathing rate per state or participant. Only under the assumption that the breathing rate estimations are (near) correct, it can be derived whether the performance of the sensors for the breathing rate is affected by differences between states or participants. However, based on Section 9.3, it is likely that this assumption does not apply at all times. Therefore, these figures merely serve as being indicative of differences in the performance of sensing and processing methods between different persons or states.

In Figures 9.6 and 9.8, 3 participants have been randomly selected, because including all participants would cause cluttered visuals. Participant 2 is a 20-year-old female from Europe, participant 18 is a 27-year-old female from Asia and participant 52 is a 20-year-old European that identified as other.

9.4.1 Heart rate comparisons

In Figure 9.6, it can be seen that for the RGB methods with respect to the heart rate, the difference in the performance of methods stays the same over different participants: heart rate range filtered data overestimates while breathing rate range filtered data tends to underestimate. Also for mmWave sensing, FFT always produces results with a wider range and with a higher value than peak estimation. Yet,

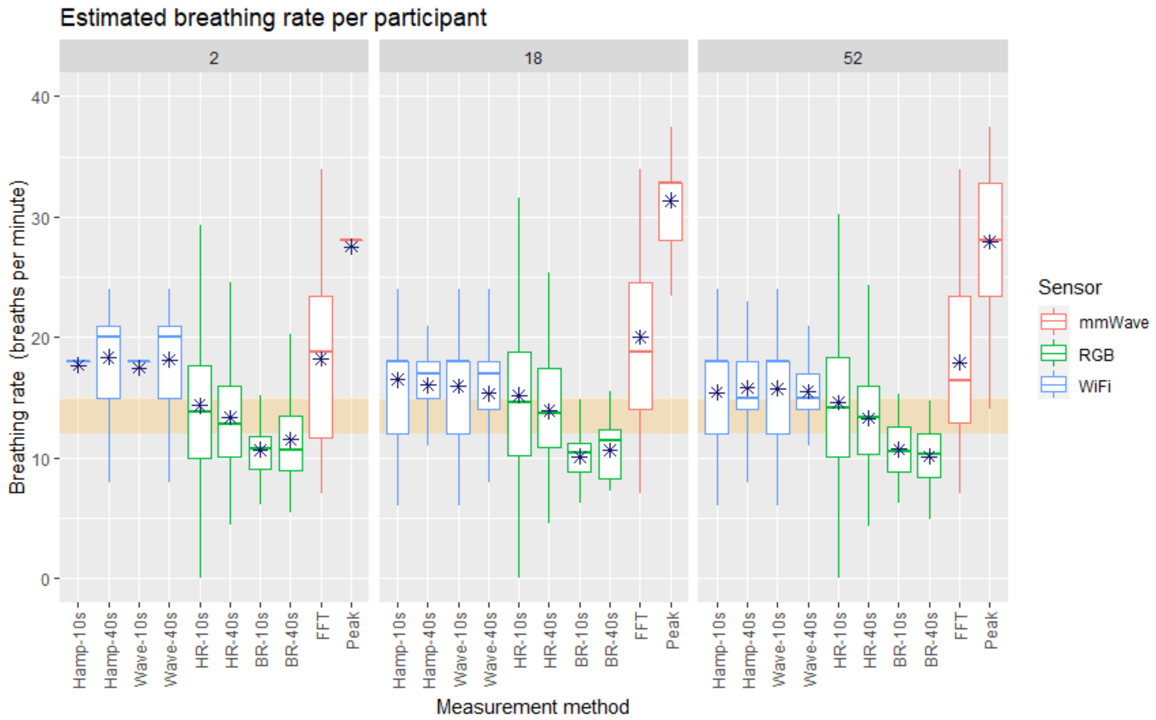

Figure 9.8: Comparison of breathing rate estimations over 3 participants using different sensing and processing methods when disregarding outliers. The mean is marked with a star

for WiFi-CSI, a significant difference can be observed: the Wavelet filter for participant 2 produces more frequently stronger underestimated heart rates. In general, all methods are shifted slightly downwards. Unfortunately, no camera data is available for participant 2 to look into the visual differences between her and the other 2 participants.

In Figure 9.7 it can be seen how the different device-free sensing methods are affected by the intended affective states included in this study. No strong visual differences can be observed, meaning that the error in measuring is slight to not at all affected by the different affective states. Only a minor difference can be found in the ranges of WiFi-CSI, especially for the Hampel filter with a 10-second window and both Wavelet filter windows. These tend to variate at the bottom between -80 and -180 and at the top between 75 and 95. However, as seen in Section 9.2, WiFi-CSI was evaluated as one of the lesser performing methods and would not be recommended for further explorations

9.4.2 Breathing rate comparisons

In contrast to the observations for the heart rate comparisons per participant, Figure 9.8 shows that there are differences between the estimations per participant.

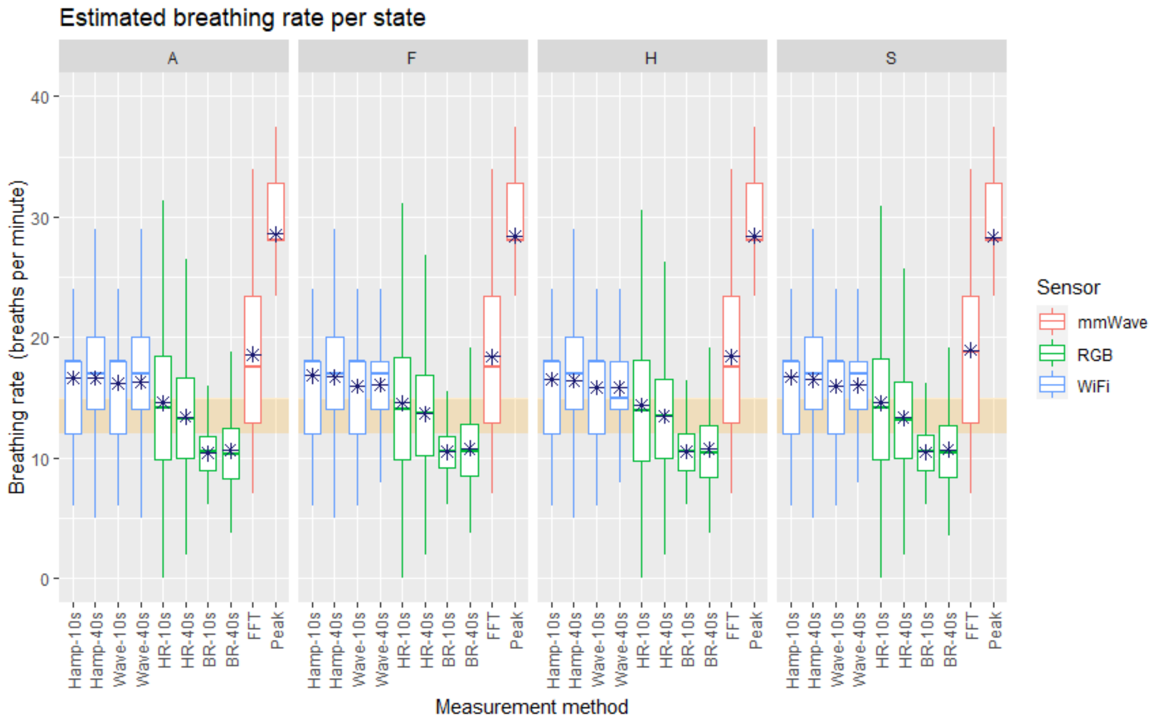

Figure 9.9: Comparison of error in breathing rate estimations per intended affective state using different sensing and processing methods when disregarding outliers. The mean is marked with a star

Again, participant 2 is the one that stands out the most, given that for 3 methods (WiFi-CSI using 10-second windows and peak estimation with the mmWave sensor) only limited measurements were successfully collected. On the other hand, participant 18 and 52 generate nearly comparable results. The biggest difference found between the 2 of them is the difference in range using the mmWave sensor with peak estimation, where participant 18 produces values in a smaller range. When comparing the charts of participants 18 and 52 to that of participant 2, another difference is found in the sense that the WiFi-CSI estimations with a 40-second window are considerably higher for participant 2 than the other 2. It could be that participant 2 has a higher respiration rate in general, and WiFi-CSI was the only one to catch this difference. However, since there is no ground truth, there is no means to validate this. Other arguments to explain this difference have not been found.

When looking at Figure 9.9, it can be seen that the charts for each state are nearly identical. Although the visual differences (e.g. the mean and median shifting slightly) are minimal, this does not mean that the breathing rate cannot be used in an emotion-aware music system. A computer is more sensitive to small differences than the human eye, and these small changes could potentially be enough (in combination with other information such as the heart rate (variability)) to detect a person's affective state. Future research will have to show whether this is also the case using sensors and methods used in this project.

9.4.3 Comparison conclusions

Based on the information from the previous paragraphs, it can be concluded that the performance of sensors and processing methods for estimating the heart rate is minimally affected by inter-personal differences and states. This also holds for the breathing rate per state. However, the ability of monitoring methods appears to differ for the breathing rate per participant: participant 2 generated different output compared to 2 other participants. However, given that the difference is found between 1 European, and another European and Asian, this tends could be indicative of the fact that the difference is not caused by ethnic differences. However, it is impossible to generalise this observation based on the limited number of cases.

9.5 Movement analysis

Although the intention at the start was to include Time-Of-Flight sensors in the experiment and analysis, they have been excluded for further analysis. Multiple sensors broke down prior to and while conducting the experiments, and could not be replaced in time. Nonetheless, some relevant observations were made that are described in the following paragraphs.

A first observation that is made is that the positioning of the sensors should be adaptable if the sensor is intended to be used by more than 1 person. Differences in length for example could cause a non-adjustable position of the sensor to measure something else than intended to. Think for example about the head: putting the sensors vertically at the height of the centre of the headrest could for one person mean the Time-of-Flight sensors would be measuring the top of the head, which has a large range of motion, while for someone else this is at neck height, which has a marginal range of space.

Another observation is that the output that is generated could be affected by the clothes a person is wearing. For example, participants wearing flared jeans nearly always had small Time-of-Flight output, as their pants were flaring out even though their legs might have been stretched more forward. The same holds for sleeves: Tight sleeves versus oversized hoodies will differ in results.

Another point of feedback made by one of the participants is that the sensors should be less observable by the person seated in the chair. The participant indicated hesitance in putting the arm on the rest in fear of breaking things.

9.6 File size and processing time

Aside from the reasonably good respiration estimations, a positive aspect of using WiFi-CSI is the limited amount of storage needed: the files (data and text, for WiFi-CSI data and timestamps, resp.) produced during the experiment are of a size of 40.5 MB per participant. More specifically, the data file is around 38 MB, and the timestamps file 2.5 MB. Moreover, the time spent to iterate over each participant and estimate the heart rate was in the order of minutes.

Similar to the WiFi-CSI, the storage needed for the data generated by the Texas Instruments GUI is small: around 30 MB per participant. Additionally, the time to process the data was again in the order of minutes.

Another disadvantage of the RGB-sensing, next to the minimal privacy, is the storage it needs as well as the processing time. The NumPy-files (frames and timestamps) stored by collecting the RGB stream of the Asus Xtion Pro live produced during the experiment are of a size of $28-40$ GB per participant (the size is strongly affected by the duration). More specifically, the data file is about $25-38$ GB, and the timestamps file 2.5 MB. Also extracting the heart- and breathing rate from this data was in the order of hours per participant.

Chapter 10

Conclusions

What becomes clear from this work is that researching solutions to the problems caused by the rising cases of dementia is urgent: with the increasing number of cases currently and in the upcoming years, costs will rise tremendously. If the healthcare sector does not grow accordingly or find solutions that decrease the time and costs needed to address dementia, this will come at the cost of quality of care and life of PWD. This thesis project aims to contribute to finding alternative solutions: in exploring multiple device-free sensing techniques to monitor the vital signs of a subject, future work has more knowledge on how WiFi-CSI, mmWave sensing, and RGB data streams are affected by respiration and the heart rate during music listening. This in turn enables informed decisions in future design choices of systems that play emotion-aware music to improve the affective state of people with dementia.

The decision to do this research with a future emotion-aware music system has been made, because existing literature has shown the strong effects, especially for people with dementia, of incorporating music in treatments. Benefits include that music allows for an alternative form of communication, it addresses depressive symptoms, it encourages interaction, reduces agitation, and much more.

Table 10.1: Abstract tabular overview of the device-free sensing technologies

Tool	Measuring	Precision	Recall	Processing time	Storage	Privacy
WiFi-CSI	Heart rate	- -	- -	++	++	-
WiFi-CSI	Respiration	-	+	++	++	-
RGB	Heart rate	0	-	- -	- -	- -
RGB	Respiration	+	++	- -	- -	- -
mmWave	Hear rate	0	0	++	++	+
mmWave	Respiration	-	- -	++	++	+

[1] Scale: – – (vey negative) to ++ (very positive).

To monitor the emotional state, many methods exist. For multiple reasons, among which the limited mobility of dementia patients, a focus was given to device-free sensing methods. These are used to monitor the vital signs unobtrusively because respiration and cardiac activity have been mentioned in the literature to be indicative of the mental state. In this report, 3 methods have been studied in-depth: WiFi-CSI, colour intensity tracking using RGB values and mmWave sensing.

An experiment was performed with 54 young adults where their heart rate and respiration rate were monitored using the previously mentioned device-free sensing methods while an affective state was induced by music. In the previous chapters, the results have been reported and discussed in detail. Table 10.1 presents an abstracted overview of how the best method for each different technique performed. From this table, it can be derived that WiFi-CSI in the current setup and with the current processing techniques was not affected enough by the heart rate to generate an output that would be reliable enough in a future emotion-aware system. For colour intensity tracking using RGB values, this was better, but the mmWave sensor performed best on both precision and recall. In combination with the other aspects, including, privacy, processing time and storage, the author evaluates mmWave sensing as the most favourable option in monitoring cardiac activity. On the other hand, WiFi-CSI is recommended as the most suitable candidate for future explorations in monitoring respiration. A trade-off needs to be made between accuracy and other critical elements: processing time, storage and privacy. Since no ground truth was collected, the actual accuracy remains unknown. The estimated precision and recall of the mmWave sensor are the lowest, and that of the RGB values is the highest. However, WiFi-CSI is more favourable based on the other aspects being processing, storage and privacy. Keeping in mind the intended situation of use, WiFi-CSI might be considered the better option here: it also allows the extraction of other data that could be relevant in addressing the increased amount of time and care needed for people with dementia, e.g. fall detection or movement deterioration. Additionally, WiFi-CSI is a relatively new method compared to the use of cameras. Possibly, when more research is done on the analysis methods to monitor the breathing rate, this might result in higher precision and recall of the output.

Based on comparing the output of 3 participants, as well as the comparison of the output per state, it has been concluded that the performance of sensing methods in monitoring the heart rate is minimally affected by inter-personal differences and differences in affective states. A similar observation was done for the breathing rate per state. A larger difference was found for inter-personal differences in monitoring the breathing rate: for one of the participants, three methods (WiFi-CSI with a 10 second window and either filter, as well as mmWave sensing using peak estimation) generated non-usable output.

Overall, it can be concluded that for actual implementation, still a lot of research and experimenting needs to be done. The findings of this work and the identified opportunities should be further explored and tested with people of older age to determine its applicability in the intended use case.

Chapter 11

Limitations and future work

Although the number of participants ($N = 54$) has led to a larger sample size compared to most works studied in Research Topics, which was one of the identified challenges in existing research, the sample is limited to some extent in diversity. To recruit participants, a database of the Faculty of Behavioural, Management and Social Sciences was used. This meant that only students of the Psychology and Communication Sciences programmes could participate. Moreover, these students following are mostly female and white. Nonetheless, a few participants that took part are male and/or from another cultural background. For in-depth studies on how different cultural backgrounds are affected by music, the generated data set in this work would not suffice.

Another limitation within this work with respect to the participants is that they were all young adults. The decision to only include young, healthy students has been made because the aim was to perform a *safe* experiment with an acceptable sample size at a time when COVID numbers were extremely high in the Netherlands. However, the intended users are people with dementia, predominantly elderly. It could be that there is a difference in how well the sensors monitor the heart rate, respiration or movement due to physical differences between elderly and youngsters. Think for example about the ability to identify a face in the RGB-data stream frames. The accuracy is possibly affected by the wrinkles and thinner skin on the face of an elder person.

Not only physical differences between youngsters and elderly are relevant in this study. Another thing to keep in mind is that elderly might be affected differently by music and may process emotion in a different manner. This holds especially for people with dementia, see Section 3.1. This means that even it though would be possible to identify the affective state of a participant in this data set with high accuracy, it would not necessarily mean these results would translate to people with dementia too.

A specific domain of research that studies the effect of ”growing older” is gerontology. The author would like to emphasise the need for research in this field and recommend performing a similar experiment to this work with people with dementia based on the findings. Steps are already being made in this regard, see [111, 112]. The author of this thesis report would encourage researchers to continue work in this direction.

The last limitation of this project is time. Extensive amounts of data have been collected, but not all of it has been processed. For example, the WiFi-CSI data could be consulted for activity recognition. Applying it in a dementia domain, WiFi-CSI has much broader potential. Vital sign detection could be extended to activity recognition, muscle deterioration and fall detection for example.

Similarly, the fixed time frame did not allow to wait for the non-functioning Time-of-Flight sensors to be replaced in time for the experiment. This prevented the opportunity to generate data that could in turn be used to explore the ability of Time-of-Flight to monitor movement/posture as an indicator of emotion.

This last sentence leads to a final point to be made in regards to limited time: the collected data of the sensors has not been combined with the self-reported affective states to research whether the generated data is suitable to produce a model that estimates the subject's emotion. Something that will be essential in designing an emotion-aware music system. Nonetheless, it is more important that this is done in an experiment where data is collected from elderly, and preferably people with dementia, to cover the intended situation the system would be used in. As mentioned before, these people might experience music and emotions in a different fashion.

Bibliography

[1] Y. Ma, G. Zhou, and S. Wang. WiFi Sensing with Channel State Information: A Survey. *ACM Comput. Surv*, 52:36, 2019. doi: 10.1145/3310194.

[2] C. Wang, T. Pun, and G. Chanel. A comparative survey of methods for remote heart rate detection from frontal face videos. *Frontiers in Bioengineering and Biotechnology*, 6(MAY):33, 5 2018. ISSN 22964185. doi: 10.3389/fbioe.2018.00033.

[3] Texas Instruments Incorporated. TI mmWave Labs Vital Signs Measurement (Version 1.4). Technical report, 2017.

[4] J. G. J. Vuijk. Pervasive Emotion Detection and the Effect of Music on People with Dementia: Current Directions and future Challenges. Technical report, University of Twente, Enschede, 9 2021.

[5] WHO. Ageing and health, 2 2018. URL https://www.who.int/news-room/fact-sheets/detail/ageing-and-health.

[6] M. J. Prince, M. M. Guerchet, G. C. Ali, and M. Prina. World Alzheimer Report 2015 - The Global Impact of Dementia - Research Portal, King's College, London. Technical report, Alzheimer's Disease International, London, 8 2015.

[7] C.P. Ferri, R. Sousa, E. Albanese, W.S. Ribeiro, and M. Honyashiki. World Alzheimer Report 2009-Executive Summary. *Alzheimer's Disease International*, pages 1–22, 2009.

[8] United Nations. Goal 3 — Ensure healthy lives and promote well-being for all at all ages, 2021. URL https://sdgs.un.org/goals/goal3.

[9] World Health Organization. *Global action plan on the public health response to dementia 2017–2025*. 2017. ISBN 9789241513487.

[10] N. Sharma, J. Klein Brinke, J. E. W. C. Van Gemert-Pijnen, and L. M. A. Braakman-Jansen. Implementation of Unobtrusive Sensing Systems for Older Adult Care: Scoping Review. *JMIR Aging 2021*, 4(4):1–21, 10 2021. ISSN 25617605. doi: 10.2196/27862.

[11] R. Keim. Understanding Color Models Used in Digital Image Processing, 8 2018. URL https://www.allaboutcircuits.com/technical-articles/understanding-color-models-used-in-digital-image-processing/.

[12] M.-Z. Poh, D. J. McDuff, and R. W. Picard. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 18(10):10762, 5 2010. ISSN 1094-4087. doi: 10.1364/OE.18.010762.

[13] W. Verkruysse, L. O. Svaasand, and J. S. Nelson. Remote plethysmographic imaging using ambient light. *Optics Express*, 16(26):21434, 12 2008. ISSN 1094-4087. doi: 10.1364/OE.16.021434.

[14] Texas Instruments Incorporated. Introduction to mmWave radar sensing: FMCW radars — TI.com Training Series. URL https://training.ti.com/node/1139153.

[15] "Universitetet i Agder". Introduction to mmWave Sensing. URL https://www.uia.no/en/studieplaner/topic/IKT722-G.

[16] Y. Lin, H. Chu, C. Y. Yang, C. H. Chen, S. G. Chen, H. J. Chang, C. J. Hsieh, and K. R. Chou. Effectiveness of group music intervention against agitated behavior in elderly persons with dementia. *International Journal of Geriatric Psychiatry*, 26(7):670–678, 7 2011. doi: 10.1002/gps.2580.

[17] S. Samson, S. Clément, P. Narme, L. Schiaratura, and N. Ehrlé. Efficacy of musical interventions in dementia: Methodological requirements of nonpharmacological trials. *Annals of the New York Academy of Sciences*, 1337(1): 249–255, 3 2015. doi: 10.1111/nyas.12621.

[18] A. Raglio and M. V. Gianelli. Music Therapy for Individuals with Dementia: Areas of Interventions and Research Perspectives. *Current Alzheimer Research*, 6:293–301, 2009. doi: 10.2174/156720509788486617.

[19] H. M. Ridder and D. Aldridge. Individual music therapy with persons with frontotemporal dementia: Singing dialogue. *Nordic Journal of Music Therapy*, 14(2):91–106, 1 2005. doi: 10.1080/08098130509478132.

[20] A. Baird and W. F. Thompson. When music compensates language: a case study of severe aphasia in dementia and the use of music by a spousal caregiver. *Aphasiology*, 33(4):449–465, 4 2019. doi: 10.1080/02687038.2018.1471657.

[21] A. Baird and S. Samson. Memory for music in Alzheimer's disease: Unforgettable? *Neuropsychology Review*, 19(1):85–101, 3 2009. doi: 10.1007/s11065-009-9085-2.

[22] S. Samson, D. Dellacherie, and H. Platel. Emotional power of music in patients with memory disorders: Clinical implications of cognitive neuroscience. In *Annals of the New York Academy of Sciences*, volume 1169, pages 245–255. Blackwell Publishing Inc., 2009. doi: 10.1111/j.1749-6632.2009.04555.x.

[23] T. Takahashi and H. Matsushita. Long-term effects of music therapy on elderly with moderate/severe dementia. *Journal of Music Therapy*, 43(4):317–333, 2006. doi: 10.1093/jmt/43.4.317.

[24] M. El Haj, L. Fasotti, and P. Allain. The involuntary nature of music-evoked autobiographical memories in Alzheimer's disease. *Consciousness and Cognition*, 21(1):238–246, 3 2012. doi: 10.1016/j.concog.2011.12.005.

[25] L. L. Cuddy and J. Duffin. Music, memory, and Alzheimer's disease: Is music recognition spared in dementia, and how can it be assessed? *Medical Hypotheses*, 64(2):229–235, 2005. doi: 10.1016/j.mehy.2004.09.005.

[26] N. Ziv, A. Granot, S. Hai, A. Dassa, and I. Haimov. The effect of background stimulative music on behavior in Alzheimer's patients. *Journal of Music Therapy*, 44(4):329–343, 2007. doi: 10.1093/jmt/44.4.329.

[27] N. Hanson, K. Gfeller, G. Woodworth, E. A. Swanson, and L. Garand. A Comparison of the Effectiveness of Differing Types and Difficulty of Music Activities in Programming for Older Adults with Alzheimer's Disease and Related Disorders. *Journal of Music Therapy*, 33(2):93–123, 1996. doi: 10.1093/jmt/33.2.93.

[28] R. M. Mathews, A. A. Clair, and K. Kosloski. Keeping the beat: Use of rhythmic music during exercise activities for the elderly with dementia. *American Journal of Alzheimer's Disease and other Dementias*, 16(6):377–380, 2001. doi: 10.1177/153331750101600608.

[29] E. Götell, S. Brown, and S. L. Ekman. The influence of caregiver singing and background music on vocally expressed emotions and moods in dementia care. *International Journal of Nursing Studies*, 46(4):422–430, 4 2009. doi: 10.1016/j.ijnurstu.2007.11.001.

[30] A. Schall, J. Haberstroh, and J. Pantel. Time series analysis of individual music therapy in dementia: Effects on communication behavior and emotional well-being. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry*, 28(3):113–122, 9 2015. doi: 10.1024/1662-9647/a000123.

[31] T. Särkämö, S. Laitinen, A. Numminen, M. Kurki, J. K. Johnson, and P. Rantanen. Pattern of emotional benefits induced by regular singing and music listening in dementia. 64(2):439–440, 2016. doi: 10.1111/jgs.13963.

[32] P. Han, M. Kwan, D. Chen, S. Yusoff, H. Chionh, J. Goh, and P. Yap. A controlled naturalistic study on a weekly music therapy and activity program on disruptive and depressive behaviors in dementia. *Dementia and Geriatric Cognitive Disorders*, 30(6):540–546, 1 2011. doi: 10.1159/000321668.

[33] M. van der Vleuten, A. Visser, and L. Meeuwesen. The contribution of intimate live music performances to the quality of life for persons with dementia. *Patient Education and Counseling*, 89(3):484–488, 12 2012. doi: 10.1016/j.pec.2012.05.012.

[34] H. Ragneskog, K. Asplund, M. Kihlgren, and A. Norberg. Individualized music played for agitated patients with dementia: Analysis of video-recorded sessions. *International Journal of Nursing Practice*, 7(3):146–155, 6 2001. doi: 10.1046/j.1440-172X.2001.00254.x.

[35] L. A. Gerdner. Individualized music for dementia: Evolution and application of evidence-based protocol. *World Journal of Psychiatry*, 2(2):26–32, 2012. doi: 10.5498/wjp.v2.i2.26.

[36] B. Lesta and P. Petocz. Familiar Group Singing: Addressing Mood and Social Behaviour of Residents with Dementia Displaying Sundowning — Australian Journal of Music Therapy. *Australian Journal of Music Therapy*, pages 2–17, 1 2006.

[37] L. L. Cuddy, R. Sikka, K. Silveira, S. Bai, and A. Vanstone. Music-evoked autobiographical memories (MEAMs) in alzheimer disease: Evidence for a positivity effect. *Cogent Psychology*, 4(1):1–20, 2017. doi: 10.1080/23311908.2016.1277578.

[38] R. J. Zatorrea and V. N. Salimpoor. From perception to pleasure: Music and its neural substrates. *Proceedings of the National Academy of Sciences of the United States of America*, 110(SUPPL2):10430–10437, 6 2013. doi: 10.1073/pnas.1301228110.

[39] S. Koelsch. Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience*, 15(3):170–180, 3 2014. doi: 10.1038/nrn3666.

[40] J. L. Agustus, C. J. Mahoney, L. E. Downey, R. Omar, M. Cohen, M. J. White, S. K. Scott, L. Mancini, and J. D. Warren. Functional MRI of music emotion

processing in frontotemporal dementia. *Annals of the New York Academy of Sciences*, 1337(1):232–240, 3 2015. doi: 10.1111/nyas.12620.

[41] Y. Gu, X. Zhang, Z. Liu, and F. Ren. WiFi-based real-time breathing and heart rate monitoring during sleep. *2019 IEEE Global Communications Conference, GLOBECOM 2019 - Proceedings*, 12 2019. doi: 10.1109/GLOBECOM38437.2019.9014297.

[42] Jian Liu, Yan Wang, Yingying Chen, Jie Yang, Xu Chen, and Jerry Cheng. Tracking vital signs during sleep leveraging off-the-shelf WiFi. *Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2015-June:267–276, 6 2015. doi: 10.1145/2746285.2746303.

[43] R.S. Thakur. *Vital Sign Monitoring based on Remote PPG and WIFI*. PhD thesis, California State University, Sacramento, 2019.

[44] M. Hussain, A. Akbilek, F. Pfeiffer, and B. Napholz. In-Vehicle Breathing Rate Monitoring Based on WiFi Signals. *2020 50th European Microwave Conference, EuMC 2020*, pages 292–295, 1 2021. doi: 10.23919/EUMC48046.2021.9338026.

[45] T. Zheng, Z. Chen, C. Cai, J. Luo, and X. Zhang. V2iFi: in-Vehicle Vital Sign Monitoring via Compact RF Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(2):1–27, 6 2020. ISSN 2474-9567. doi: 10.1145/3397321.

[46] B. Tan, Q. Chen, K. Chetty, K. Woodbridge, W. Li, and R. Piechocki. Exploiting WiFi Channel State Information for Residential Healthcare Informatics. *IEEE Communications Magazine*, 56(5):130–137, 5 2018. ISSN 15581896. doi: 10.1109/MCOM.2018.1700064.

[47] X. Wang, C. Yang, and S. Mao. On CSI-Based Vital Sign Monitoring Using Commodity WiFi. 2020. doi: 10.1145/3377165.

[48] B. Fang, N. D. Lane, M. Zhang, A. Boran, and F. Kawsar. BodyScan: Enabling Radio-based Sensing on Wearable Devices for Contactless Activity and Vital Sign Monitoring. 2016. doi: 10.1145/2906388.2906411.

[49] Q. Gao, J. Tong, J. Wang, S. Member, Z. Ran, and M. Pan. Device-Free Multi-Person Respiration Monitoring Using WiFi. *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*, 69(11), 2020. doi: 10.1109/TVT.2020.3020180.

[50] Y. He, Y. Chen, Y. Hu, and B. Zeng. WiFi Vision: Sensing, Recognition, and Detection With Commodity MIMO-OFDM WiFi. *IEEE INTERNET OF THINGS JOURNAL*, 7(9), 2020. doi: 10.1109/JIOT.2020.2989426.

[51] X. Wang, X. Wang, and S. Mao. RF Sensing in the Internet of Things: A General Deep Learning Framework; RF Sensing in the Internet of Things: A General Deep Learning Framework. *IEEE Communications Magazine*, 56, 2018. doi: 10.1109/MCOM.2018.1701277.

[52] D. Zhang, Y. Zeng, F. Zhang, and J. Xiong. WiFi CSI-based vital signs monitoring. *Contactless Vital Signs Monitoring*, pages 231–255, 1 2022. doi: 10.1016/B978-0-12-822281-2.00020-2.

[53] D. Wu, D. Zhang, C. Xu, H. Wang, and X. Li. Device-Free WiFi Human Sensing: From Pattern-Based to Model-Based Approaches. *IEEE Communications Magazine*, 55(10):91–97, 10 2017. ISSN 01636804. doi: 10.1109/MCOM.2017.1700143.

[54] K. Ali, M. Alloulah, F. Kawsar, and A. X. Liu. On Goodness of WiFi based Monitoring of Vital Signs in the Wild. *IEEE Transactions on Mobile Computing*, 3 2020. ISSN 15580660. doi: 10.48550/arxiv.2003.09386.

[55] F. Wang, F. Zhang, C. Wu, B. Wang, and K. J. R. Liu. ViMo: Multiperson Vital Sign Monitoring Using Commodity Millimeter-Wave Radio; ViMo: Multiperson Vital Sign Monitoring Using Commodity Millimeter-Wave Radio. *IEEE Internet of Things Journal*, 8(3), 2021. doi: 10.1109/JIOT.2020.3004046.

[56] Z Yang, P. H. Pathak, Y Zeng, X Liran, and P Mohapatra. Monitoring vital signs using millimeter wave. In *Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pages 211–220, New York, NY, USA, 7 2016. ACM. ISBN 9781450341844. doi: 10.1145/2942358.2942381.

[57] A. Ahmad, J. C. Roh, D. Wang, and A. Dubey. Vital signs monitoring of multiple people using a FMCW millimeter-wave sensor. *2018 IEEE Radar Conference, RadarConf 2018*, pages 1450–1455, 6 2018. doi: 10.1109/RADAR.2018.8378778.

[58] S. M.M. Islam, N. Motoyama, S. Pacheco, and V. M. Lubecke. Non-contact vital signs monitoring for multiple subjects using a millimeter wave FMCW automotive radar. *IEEE MTT-S International Microwave Symposium Digest*, 2020-August:783–786, 8 2020. ISSN 0149645X. doi: 10.1109/IMS30576.2020.9223838.

[59] C. Yang. *Contact-free Vital Sign Monitoring Using Phase Difference of Channel State Information*. PhD thesis, Auburn Uniersity, Auburn, 12 2017.

[60] S. Bakhtiari, T. W. Elmer, N. M. Cox, N. Gopalsami, A. C. Raptis, S. Liao, I. Mikhelson, and A. V. Sahakian. Compact millimeter-wave sensor for remote monitoring of vital signs. *IEEE Transactions on Instrumentation and Measurement*, 61(3):830–841, 3 2012. ISSN 00189456. doi: 10.1109/TIM.2011.2171589.

[61] D. T. Petkie, C. Benton, and E. Bryan. Millimeter wave radar for remote measurement of vital signs. *IEEE National Radar Conference - Proceedings*, 2009. ISSN 10975659. doi: 10.1109/RADAR.2009.4977021.

[62] A. Singh, S. U. Rehman, S. Yongchareon, and P. H. J. Chong. Multi-Resident Non-Contact Vital Sign Monitoring Using Radar: A Review. *IEEE Sensors Journal*, 21(4):4061–4084, 2 2021. ISSN 15581748. doi: 10.1109/JSEN.2020.3036039.

[63] M. Alizadeh, G. Shaker, J. C. M. De Almeida, Plinio Pelegrini Morita, and Safeddin Safavi-Naeini. Remote monitoring of human vital signs using mm-Wave FMCW Radar. *IEEE Access*, 7:54958–54968, 2019. ISSN 21693536. doi: 10.1109/ACCESS.2019.2912956.

[64] Y. Wang, W. Wang, M. Zhou, A. Ren, and Z. Tian. Remote Monitoring of Human Vital Signs Based on 77-GHz mm-Wave FMCW Radar. *Sensors*, 20: 1–23, 2020. doi: 10.3390/s20102999.

[65] W. Wang, Z. Jia, C. Xu, G. Luo, D. Zhang, N. An, and Y. Zhang. Feasibility study of practical vital sign detection using millimeter-wave radios. *CCF Transactions on Pervasive Computing and Interaction*, 3(4):436–452, 12 2021. ISSN 25245228. doi: 10.1007/S42486-021-00080-4/FIGURES/17.

[66] S. Churkin and L. Anishchenko. Millimeter-wave radar for vital signs monitoring. *2015 IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems, COMCAS 2015*, 12 2015. doi: 10.1109/COMCAS.2015.7360366.

[67] H. Wang, J. H. Cheng, J. C. Kao, and T. W. Huang. Review on microwave/millimeter-wave systems for vital sign detection. *WiSNet 2014 - Proceedings: 2014 IEEE Topical Conference on Wireless Sensors and Sensor Networks*, pages 19–21, 2014. doi: 10.1109/WISNET.2014.6825499.

[68] S. Suzuki, T. Matsui, M. Kagawa, T. Asao, and K. Kotani. An approach to a non-contact vital sign monitoring using dual-frequency microwave radars for elderly care. *Journal of Biomedical Science and Engineering*, 06(07):704–711, 2013. ISSN 1937-6871. doi: 10.4236/jbise.2013.67086.

[69] M. Villarroel, J. Jorge, C. Pugh, and L. Tarassenko. Non-Contact Vital Sign Monitoring in the Clinic. In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pages 278–285. IEEE, 5 2017. ISBN 978-1-5090-4023-0. doi: 10.1109/FG.2017.43.

[70] S. Lyra, L. Mayer, L. Ou, D. Chen, P. Timms, A. Tay, P. Y. Chan, B. Ganse, S. Leonhardt, and C. H. Antink. A Deep Learning-Based Camera Approach for Vital Sign Monitoring Using Thermography Images for ICU Patients. *Sensors 2021, Vol. 21, Page 1495*, 21(4):1495, 2 2021. ISSN 1424-8220. doi: 10.3390/S21041495.

[71] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. A. Clifton, and C. Pugh. Non-contact video-based vital sign monitoring using ambient light and auto-regressive models. *Physiological Measurement*, 35(5):807, 3 2014. ISSN 0967-3334. doi: 10.1088/0967-3334/35/5/807.

[72] T. Negishi, S. Abe, T. Matsui, H. Liu, M. Kurosawa, T. Kirimoto, and G. Sun. Contactless Vital Signs Measurement System Using RGB-Thermal Image Sensors and Its Clinical Screening Test on Patients with Seasonal Influenza. *Sensors 2020, Vol. 20, Page 2171*, 20(8):2171, 4 2020. ISSN 1424-8220. doi: 10.3390/S20082171.

[73] B. Aubakir, B. Nurimbetov, I. Tursynbek, and H. A. Varol. Vital sign monitoring utilizing Eulerian video magnification and thermography. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2016-October:3527–3530, 10 2016. ISSN 1557170X. doi: 10.1109/EMBC.2016.7591489.

[74] H. E. Tasli, A. Gudi, and M. Den Uyl. Remote PPG based vital sign measurement using adaptive facial regions. *2014 IEEE International Conference on Image Processing, ICIP 2014*, pages 1410–1414, 1 2014. doi: 10.1109/ICIP.2014.7025282.

[75] H.-S. Hwang and E.-C. Lee. Non-Contact Respiration Measurement Method Based on RGB Camera Using 1D Convolutional Neural Networks. *Sensors*, 21(10):3456, 5 2021. ISSN 1424-8220. doi: 10.3390/s21103456.

[76] P. V. Rouast, M. T. P. Adam, R. Chiong, D. Cornforth, and E. Lux. Remote heart rate measurement using low-cost RGB face video: a technical literature review. *Frontiers of Computer Science 2018 12:5*, 12(5):858–872, 9 2018. ISSN 2095-2236. doi: 10.1007/S11704-016-6243-6.

[77] C. Wrede, A. Braakman-Jansen, and L. Van Gemert-Pijnen. Requirements for unobtrusive monitoring to support home-based dementia care: Qualitative study among formal and informal caregivers. *JMIR Aging*, 4(2), 4 2021. ISSN 25617605. doi: 10.2196/26875.

[78] K. Wild, L. Boise, J. Lundell, and A. Foucek. Unobtrusive In-Home Monitoring of Cognitive and Physical Health: Reactions and Perceptions of Older Adults:. *https://doi-org.ezproxy2.utwente.nl/10.1177/0733464807311435*, 27 (2):181–200, 4 2008. doi: 10.1177/0733464807311435.

[79] D. Bradford, J. Freyne, and M. Karunanithi. Sensors on My Bed: The Ups and Downs of In-Home Monitoring. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7910 LNCS:10–18, 2013. doi: 10.1007/978-3-642-39470-6{_}2.

[80] E. Phoebe, I. Hunter, D. Whiddett, C. Lockhart, H. Guesgen, and A. Singh. User Requirements for Technology to Assist Aging in Place: Qualitative Study of Older People and Their Informal Support Networks. *JMIR Mhealth Uhealth 2018;6(6):e10741 https://mhealth.jmir.org/2018/6/e10741*, 6 (6):e10741, 6 2018. doi: 10.2196/10741.

[81] C. Berridge and T. F. Wetle. Why Older Adults and Their Children Disagree About In-Home Surveillance Technology, Sensors, and Tracking. *The Gerontologist*, 60(5):926–934, 7 2020. ISSN 0016-9013. doi: 10.1093/GERONT/GNZ068.

[82] H. van Vliet. Software Engineering: Principles and Practice. 2007.

[83] D.T Kenrick. 7 Good Things About Feeling Bad, 2013. URL https://www.psychologytoday.com/us/blog/sex-murder-and-the-meaning-life/201306/7-good-things-about-feeling-bad.

[84] R. Veenhoven. QUALITY-OF-LIFE IN INDIVIDUALISTIC SOCIETY. *Social Indicators Research*, 48:157–186, 1999.

[85] Boise L, Wild K, Mattek N, Ruhl M, Dodge HH, and Kaye J. Willingness of older adults to share data and privacy concerns after exposure to unobtrusive

in-home monitoring. *Gerontechnology : International Journal on the Fundamental Aspects of Technology to Serve the Ageing Society*, 11(3):428–435, 1 2013. ISSN 1569-1101. doi: 10.4017/GT.2013.11.3.001.00.

[86] A. T. Y. Chen, M. Biglari-Abhari, and K. I. K. Wang. Context is King: Privacy Perceptions of Camera-based Surveillance. *Proceedings of AVSS 2018 - 2018 15th IEEE International Conference on Advanced Video and Signal-Based Surveillance*, 2 2019. doi: 10.1109/AVSS.2018.8639448.

[87] A. Deep Singh, S. Singh Sandha, L. Garcia, and M. Srivastava. RadHAR: Human Activity Recognition from Point Clouds Generated through a Millimeter-wave Radar. *Proceedings of the 3rd ACM Workshop on Millimeter-wave Networks and Sensing Systems - mmNets'19*, 19, 2019. doi: 10.1145/3349624.

[88] S. K. Yadav, S. Sai, A. Gundewar, H. Rathore, K. Tiwari, H. M. Pandey, and M. Mathur. CSITime: Privacy-preserving human activity recognition using WiFi channel state information. *Neural Networks*, 146:11–21, 2 2022. ISSN 0893-6080. doi: 10.1016/J.NEUNET.2021.11.011.

[89] IBM Corp. IBM SPSS Statistics for Windows, 2019.

[90] J. A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 12 1980. doi: 10.1037/H0077714.

[91] G. Stoet. PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods 2010 42:4*, 42(4):1096–1104, 11 2010. ISSN 1554-3528. doi: 10.3758/BRM.42.4.1096.

[92] G. Stoet. PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *https://doi-org.ezproxy2.utwente.nl/10.1177/0098628316677643*, 44(1):24–31, 11 2016. ISSN 00986283. doi: 10.1177/0098628316677643.

[93] Adafruit VL6180X Time of Flight Distance Sensor - 5 to 100mm - STEMMA QT. URL https://www.kiwi-electronics.nl/nl/adafruit-vl6180x-time-of-flight-distance-sensor-5-to-100mm-stemma-qt-2761.

[94] IWR1443BOOST Evaluation board — TI.com. URL https://www.ti.com/tool/IWR1443BOOST.

[95] E4 wristband. URL https://www.empatica.com/en-eu/research/e4/.

[96] R. Rakshit, R. V. Reddy, and P. Deshpande. Emotion detection and recognition using HRV features derived from photoplethysmogram signals. *ERM4CT*

2016 - Proceedings of the 2nd Workshop on Emotion Representations and Modelling for Companion Systems, pages 3–8, 11 2016. doi: 10.1145/3009960.3009962.

[97] F. Shaffer and J. P. Ginsberg. An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, 5:258, 9 2017. ISSN 22962565. doi: 10.3389/FPUBH.2017.00258.

[98] S. Föll, M. Maritsch, F. Spinola, V. Mishra, F. Barata, T. Kowatsch, E. Fleisch, and F. Wortmann. FLIRT: A feature generation toolkit for wearable data. *Computer Methods and Programs in Biomedicine*, 212:106461, 11 2021. ISSN 0169-2607. doi: 10.1016/J.CMPB.2021.106461.

[99] API Reference — FLIRT 0.0.3 documentation. URL https://flirt.readthedocs.io/en/latest/api.html#module-flirt.hrv.

[100] MATLAB. version 9.11 (R2021b), 2021.

[101] Dhalperi. Linux-80211n-csitool-supplementary, 2015. URL https://github.com/dhalperi/linux-80211n-csitool-supplementary.

[102] MATLAB. Outlier removal using Hampel identifier - MATLAB hampel - MathWorks Benelux, . URL https://nl.mathworks.com/help/signal/ref/hampel.html.

[103] MATLAB. Wavelet signal denoising - MATLAB wdenoise - MathWorks Benelux, . URL https://nl.mathworks.com/help/wavelet/ref/wdenoise.html.

[104] D. Hass, S. Mullinix, and H. Pope. Heart Rate Detection Using Remote Photoplethysmography, 2020. URL https://github.com/mullisd1/CV_Heartrate.

[105] P. van Gent, H. Farah, N. van Nes, and B. van Arem. Analysing noisy driver physiology real-time using off-the-shelf sensors: Heart rate analysis software from the taking the fast lane project. *Journal of Open Research Software*, 7 (1):32, 10 2019. ISSN 20499647. doi: 10.5334/JORS.241/METRICS/.

[106] Texas Instruments Incorporated. Vital Signs Lab, 2017.

[107] Radboud UMC. Een normale ademhaling. URL https://www.radboudumc.nl/patientenzorg/behandelingen/revalidatie-bij-neuromusculaire-aandoeningen/ademhalingsoefeningen-bij-neuromusculaire-aandoeningen/een-normale-ademhaling.

[108] RStudio Team. RStudio: Integrated Development for R, 2020.

[109] "Paul Ekman Group". Mood vs Emotion: Differences & Traits. URL https://www.paulekman.com/blog/mood-vs-emotion-difference-between-mood-emotion/.

[110] R. J. Comer, E. Gould, and A. Furnham. *Psychology*. 2013. ISBN 978-1-119-94126-2.

[111] D. S. Nazareth. Emotion Recognition in Dementia: Advancing technology for multimodal analysis of emotion expression in everyday life. In *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. IEEE, 9 2019. ISBN 978-1-7281-3891-6. doi: 10.1109/ACIIW.2019.8925059.

[112] D. S. Nazareth, M.-P. Jansen, K. P. Truong, G.J. Westerhof, and D. Heylen. MEMOA: Introducing the Multi-Modal Emotional Memories of Older Adults Database. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 9 2019. ISBN 978-1-7281-3888-6. doi: 10.1109/ACII.2019.8925462.

[113] E. Kanjo, L. Al-Husain, and A. Chamberlain. Emotions in context: examining pervasive affective sensing systems, applications, and analyses. *Personal and Ubiquitous Computing*, 19(7):1197–1212, 10 2015. doi: 10.1007/s00779-015-0842-3.

[114] R. A. Calvo and S. D'Mello. Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, 1(1):18–37, 1 2010. doi: 10.1109/T-AFFC.2010.1.

[115] S. R. Livingstone and F. A. Russo. The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. *PLOS ONE*, 13(5), 5 2018. ISSN 1932-6203. doi: 10.1371/journal.pone.0196391.

[116] R. Chu. Speech Emotion Recognition with Convolutional Neural Network. URL https://towardsdatascience.com/speech-emotion-recognition-with-convolution-neural-network-1e6bb7130ce3.

[117] B. McFee, C. Raffel, D. Liang, D. P. W. Ellis, M. McVicar, E. Battenberg, and O. Nieto. librosa: Audio and Music Signal Analysis in Python. *PROC. OF THE 14th PYTHON IN SCIENCE CONF. (SCIPY*, 2015:1. URL https://github.com/bmcfee/librosa.

[118] F. Chollet. Keras, 2015. URL https://github.com/keras-team/keras.

[119] J. Esch. A Survey on ambient intelligence in healthcare. *Proceedings of the IEEE*, 101(12):2467–2469, 2013. doi: 10.1109/JPROC.2013.2286654.

[120] H. Zacharatos, C. Gatzoulis, and Y. L. Chrysanthou. Automatic emotion recognition based on body movement analysis: A survey. *IEEE Computer Graphics and Applications*, 34(6):35–45, 11 2014. ISSN 02721716. doi: 10.1109/MCG.2014.106.

Appendix A

Exploration

In order to determine the affective state of a subject, multiple human acts can be consulted. The upcoming paragraphs describe possible methods in more detail.

A.1 Camera

One possible technique to monitor the emotional state using device-free sensors is with the use of a camera. The data used to determine the emotion could either be an image or a video [113]. Facial expressions can be extracted to estimate an emotional state. However, these can be mocked. As video materials is more elaborate, this also allows to extract gestures such as nodding or head shaking.

A.2 Microphone

Another possible method is to add a microphone to the environment. This data could be analysed on its content as well as its acoustics [114]. Content could contain negative or positive terms that could be an indicator of the emotional state of the speaker. In a less direct manner, acoustics like the sounds of crying or vibrations in the voice, which are universal while a language is not, could also be indicative of a specific affective state. As this is less privacy-intrusive, it is explored in more detail in the following paragraphs.

To explore the ability of using acoustic features in speech for emotion classification, the Ryerson Audio-Visual Database of emotional Speech and Song (RAVDESS) database [115] has been used. This is an academically published and validated database consisting of 7356 files. Among these files are both speech files, as well as song files. For speech, the available emotions are calm, happy, sad, angry, fearful, surprise and disgust. For song, this is calm, happy, sad, angry and fearful. The emotions (except for 'neutral) are expressed in 2 different intensities, being normal and strong. To generate this database, 24 (f:12) professional actors have been

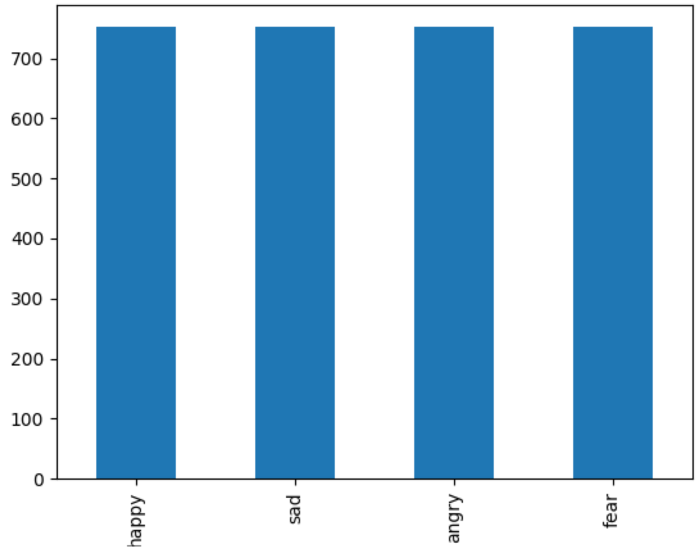

Figure A.1: Distribution of classes used in RAVDESS

asked to vocalise 2 lexically-matched statements: ”Kids are talking by the door” and ”Dogs are sitting by the door” All vocalisations are available in 3 formats: audio-only, video-only and audio-video.

To produce a model fitting the system design in this thesis, the data has been filtered to only include vocalisations that concern happiness, sadness, anger and fear. This means that 3008 files remain, where all emotions are equally distributed (i.e. 752 files per emotion). The equal distribution is visualised in Figure A.1. In order to train the model, a decision has been made to only make use of the audio files. Visual data is not relevant in this case, as the point of focus here is the acoustic features of speech. However, another conscious decision has been made to include both speech and song, as this is deemed relevant in the use case studied in this thesis, where future users of an emotion-aware music box are presented musical items.

From the remaining data files, the mel-frequency cepstral coefficients have been extracted as features for classification. The model that has been applied has been derived from [116]. The code has been developed using python. Librosa 0.8.1 [117] has been used to extract the acoustic features, while Keras 2.6.0 [118] has been used to develop the deep learning model.

Although the trained model shows promising results(see Table A.1 and Figure A.3, A.4 and A.2), and confirms findings that speech can be used to detect emotions [119], it has been decided to not further explore this modality in the remainder of the project. The chances that users speak sufficient amounts to generate enough speech data for accurate emotion classification are low. It is assumed that while

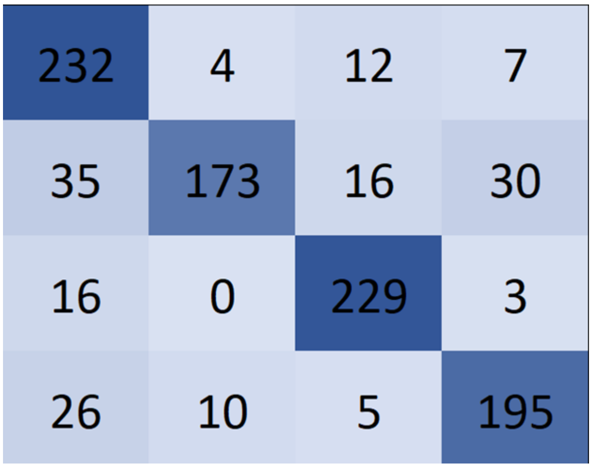

Figure A.2: Confusion matrix of emotion classifier based on the MFCC of speech and singing fragments of RAVDESS.

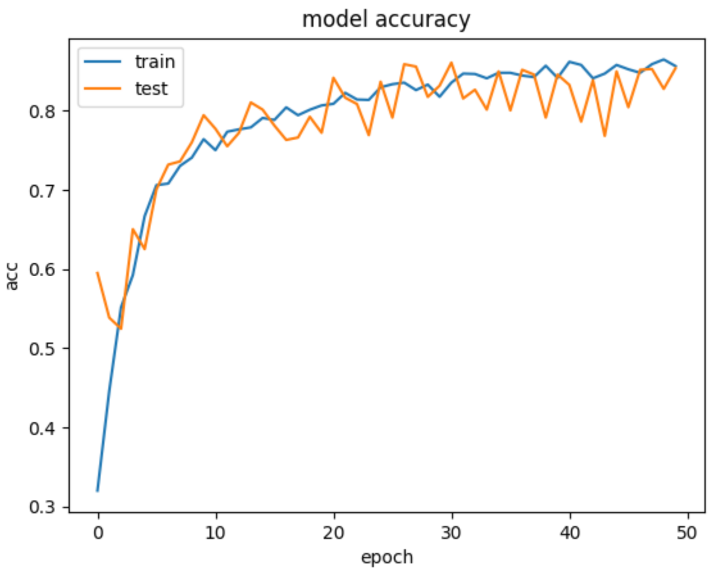

Figure A.3: Accuracy plot of RAVDESS.

Table A.1: Classification report RAVDESS

	Precision	Recall	F1-score	Support
Happy	0.94	0.77	0.85	248
Sad	0.64	0.96	0.77	248
Angry	0.94	0.88	0.91	249
Fearful	0.85	0.63	0.73	248
Accuracy			0.81	993
Macro avg	0.84	0.81	0.81	993
Weighted avg	0.84	0.81	0.81	993

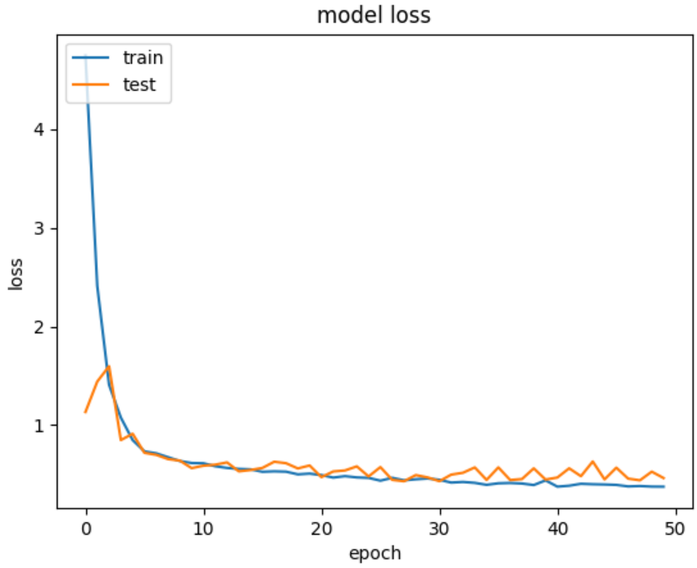

Figure A.4: Loss plot of of RAVDESS.

listening to the music, listeners will speak little to not at all, and possibly only hum. In short, this means that the use of acoustic features is likely a good device-free opportunity for emotion classification, however not in this use case.

As speech is deemed inappropriate for this use case, another opportunity in audio is explored in this thesis: respiration rate.

A.3 Movement

The final indicator of an affective state discussed in this report is movement/posture. Zacharatos, Gatzoulis and Chrysantou [120], present an extensive literature study that presents, amongst others, emerging techniques (in 2018) in using body movement for emotion recognition. Measuring movement can be done in numerous manners, e.g. tracking the activity of a persons body on visual data or using on-body motion sensors. However, as these are likely considered invasive, this work includes Time-of-Flight time series and WiFi-CSI to track human movement.

Appendix B

Song collecting questionnaire

1. Please list your top 3 songs that make you experience happiness.

Please format your answer as follows: Artist(s) - Title
In case you experience difficulties coming up with songs, you could get inspiration from scrolling trough your playlists.
For example on Spotify or YouTube.

Number 1 - Making me feel happy

Number 2 - Making me feel happy

Number 3 - Making me feel happy

----

2. Which music genre(s) make(s) you feel happy?

You can select multiple answer options or add your own if you wish to.

☐ *Blues*
☐ *Classical*
☐ *Country*
☐ *Disco*
☐ *Jazz*
☐ *Metal*
☐ *Pop*
☐ *Hiphop*
☐ *Reggea*
☐ *Rock*

☐ *Other...*, ________________

3. Please list your top 3 songs that make you experience anger.

Please format your answer as follows: Artist(s) - Title
In case you experience difficulties coming up with songs, you could get inspiration from scrolling trough your playlists.
For example on Spotify or YouTube.

Number 1 - Making me feel angry

____

Number 2 - Making me feel angry

____

Number 3 - Making me feel angry

____

4. Which music genre(s) make(s) you feel angry?

You can select multiple answer options or add your own if you wish to.

☐ *Blues*
☐ *Classical*
☐ *Country*
☐ *Disco*
☐ *Jazz*
☐ *Metal*
☐ *Pop*
☐ *Hiphop*
☐ *Reggea*
☐ *Rock*
☐ *Other...*, ________________

5. Please list your top 3 songs that make you experience sadness.

Please format your answer as follows: Artist(s) - Title
In case you experience difficulties coming up with songs, you could get inspiration from scrolling trough your playlists.
For example on Spotify or YouTube.

Number 1 - Making me feel sad

____

Number 2 - Making me feel sad

____

Number 3 - Making me feel sad

____

6. Which music genre(s) make(s) you feel sad?

You can select multiple answer options or add your own if you wish to.

☐ *Blues*
☐ *Classical*
☐ *Country*
☐ *Disco*
☐ *Jazz*
☐ *Metal*
☐ *Pop*
☐ *Hiphop*
☐ *Reggea*
☐ *Rock*
☐ *Other...*, ________________

7. Please list your top 3 songs that make you experience fear.

Please format your answer as follows: Artist(s) - Title
In case you experience difficulties coming up with songs, you could get inspiration from scrolling trough your playlists.
For example on Spotify or YouTube.

Number 1 - Making me feel fearful

____

Number 2 - Making me feel fearful

____

Number 3 - Making me feel fearful

____

8. Which music genre(s) make(s) you feel fearful?

You can select multiple answer options or add your own if you wish to.

☐ *Blues*
☐ *Classical*
☐ *Country*
☐ *Disco*
☐ *Jazz*
☐ *Metal*
☐ *Pop*
☐ *Hiphop*
☐ *Reggea*
☐ *Rock*
☐ *Other. . .* , ________________

Appendix C

Statistical outputs song collecting questionnaire

Table C.1: Descriptives of audio features per affective state

playlist		danceability	energy	loudness	speechiness	instrumentalness	liveness	valence	tempo
Angry	Mean	.590767	.723340	-6.44851	.121486	5.800E-002	.205162	.475207	120.61111
	Std. Deviation	.1510162	.2140581	3.117632	.1235590	1.7800E-001	.1658700	.2186526	30.110991
	Minimum	.2980	.1290	-18.291	.0255	0.0E+000	.0290	.0389	65.043
	Maximum	.9350	.9870	-2.087	.6720	9.0E-001	.8680	.9470	200.039
Fearful	Mean	.505238	.510796	-10.98302	.083957	1.880E-001	.181500	.355017	118.78535
	Std. Deviation	.1864691	.2705863	5.362272	.1019499	3.3410E-001	.1797900	.2642246	25.740741
	Minimum	.0882	.0706	-26.966	.0263	0.0E+000	.0243	.0287	62.952
	Maximum	.8540	.9890	-1.909	.5450	1.0E+000	.9580	.9640	174.111
Happy	Mean	.652943	.713984	-6.47489	.077275	2.600E-002	.191243	.638352	123.99667
	Std. Deviation	.1426383	.1738396	3.013235	.0787177	1.1700E-001	.1596145	.2086380	26.991388
	Minimum	.1510	.1250	-20.414	.0248	0.0E+000	.0486	.0990	74.268
	Maximum	.9350	.9710	-2.098	.5410	9.0E-001	.9650	.9680	194.081
Sad	Mean	.515628	.444551	-9.24993	.047338	2.900E-002	.141379	.330247	113.75031
	Std. Deviation	.1505430	.2245987	3.844143	.0329182	1.1840E-001	.0993109	.1732021	29.606745
	Minimum	.2030	.0581	-19.457	.0243	0.0E+000	.0497	.0393	58.714
	Maximum	.9210	.9720	-2.810	.3070	7.0E-001	.6170	.8010	186.162
Total	Mean	.569805	.599682	-8.16291	.080783	6.800E-002	.178675	.456050	119.27867
	Std. Deviation	.1675393	.2522289	4.265073	.0918332	2.0410E-001	.1531364	.2485118	28.428348
	Minimum	.0882	.0581	-26.966	.0243	0.0E+000	.0243	.0287	58.714
	Maximum	.9350	.9890	-1.909	.6720	1.0E+000	.9650	.9680	200.039

Table C.2: Correlations between audio features

		danceability	energy	loudness	speechiness	instrumentalness	liveness	valence	tempo
danceability	Pearson Correlation	1	.279**	.368**	.254**	-.283**	.056	.573**	.061
	Sig. (2-tailed)		.000	.000	.000	.000	.247	.000	.207
	N	435	435	435	435	435	435	435	435
energy	Pearson Correlation	.279**	1	.800**	.139**	-.257**	.242**	.493**	.128**
	Sig. (2-tailed)	.000		.000	.004	.000	.000	.000	.008
	N	435	435	435	435	435	435	435	435
loudness	Pearson Correlation	.368**	.800**	1	.113*	-.439**	.181**	.410**	.126**
	Sig. (2-tailed)	.000	.000		.018	.000	.000	.000	.008
	N	435	435	435	435	435	435	435	435
speechiness	Pearson Correlation	.254**	.139**	.113*	1	-.100*	.010	.154**	.189**
	Sig. (2-tailed)	.000	.004	.018		.038	.838	.001	.000
	N	435	435	435	435	435	435	435	435
instrumentalness	Pearson Correlation	-.283**	-.257**	-.439**	-.100*	1	-.093	-.270**	-.124**
	Sig. (2-tailed)	.000	.000	.000	.038		.053	.000	.010
	N	435	435	435	435	435	435	435	435
liveness	Pearson Correlation	.056	.242**	.181**	.010	-.093	1	.139**	-.018
	Sig. (2-tailed)	.247	.000	.000	.838	.053		.004	.706
	N	435	435	435	435	435	435	435	435
valence	Pearson Correlation	.573**	.493**	.410**	.154**	-.270**	.139**	1	.143**
	Sig. (2-tailed)	.000	.000	.000	.001	.000	.004		.003
	N	435	435	435	435	435	435	435	435
tempo	Pearson Correlation	.061	.128**	.126**	.189**	-.124**	-.018	.143**	1
	Sig. (2-tailed)	.207	.008	.008	.000	.010	.706	.003	
	N	435	435	435	435	435	435	435	435

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table C.3: Analysis of variance table for audio features

			Sum of Squares	df	Mean Square	F	Sig.
tempo * playlist	Between Groups	(Combined)	6618.278	3	2206.093	2.763	.042
	Within Groups		344127.923	431	798.441		
	Total		350746.201	434			
danceability * playlist	Between Groups	(Combined)	1.615	3	.538	21.952	.000
	Within Groups		10.567	431	.025		
	Total		12.182	434			
energy * playlist	Between Groups	(Combined)	6.784	3	2.261	46.797	.000
	Within Groups		20.827	431	.048		
	Total		27.611	434			
speechiness * playlist	Between Groups	(Combined)	.308	3	.103	13.219	.000
	Within Groups		3.352	431	.008		
	Total		3.660	434			
instrumentalness * playlist	Between Groups	(Combined)	1.690	3	.563	14.805	.000
	Within Groups		16.395	431	.038		
	Total		18.085	434			
liveness * playlist	Between Groups	(Combined)	.261	3	.087	3.775	.011
	Within Groups		9.917	431	.023		
	Total		10.178	434			
valence * playlist	Between Groups	(Combined)	6.916	3	2.305	49.961	.000
	Within Groups		19.887	431	.046		
	Total		26.803	434			

a. The grouping variable playlist is a string, so the test for linearity cannot be computed.

Table C.4: Measures of association of audio features

	Eta	Eta Squared
tempo * playlist	.137	.019
danceability * playlist	.364	.133
energy * playlist	.496	.246
speechiness * playlist	.290	.084
instrumentalness * playlist	.306	.093
liveness * playlist	.160	.026
valence * playlist	.508	.258

Appendix D

Playlists used to select songs for experiment

Table D.1: Descriptives (at the time of use) of the playlists that have been used to select the musical items for the experiment.

Affective state	Playlist	URL	Number of items	Likes
Happy	Have a great day!	https://open.spotify.com/playlist/37i9dQZF1DX7KNKjOK0o75?	80	5.109.365
Sad	Sad songs :(	https://open.spotify.com/playlist/6nxPNnmSE0d5WlplUsa5L3?	159	58.502
Angry	Angry songs :)	https://open.spotify.com/playlist/5bWn8XcgI3hv2vpWjNq8la?	188	9293
Fearful	Scary songs	https://open.spotify.com/playlist/5AM4lgcUAw5sokybXj3ny7?	79	79

Appendix E

Questionnaire experiment

1. How old are you?

______

2. What gender do you identify as?

O Male
O Female
O Prefer not to say
O Prefer to self-describe (please specify), _________________

3. Please enter your occupation (e.g. student psychology, physics professor, retired or unemployed) below.)

____

4. Where were you born?

O North America
O Central America
O South America
O Europe
O Africa
O Asia
O Australia
O Caribbean Islands
O Pacific Islands
O Prefer not to say
O Prefer to self-describe (please specify), _________________

5. Please specify your ethnicity)

O Hispanic or Latino or Spanish Origin of any race

O American Indian or Alaska Native
O Asian
O Black or African American
O Native Hawaiian or Other Pacific Islander
O White
O Two or more races
O Prefer not to say
O Unknown
O Prefer to self-describe (please specify), _________________

6. To what extent do you feel the affective states mentioned below. Please rate on a scale from 1 to 5.

	1. Not at all	2. Slightly	3. Moderately	4. Very much	5. Extremely
Happy	○	○	○	○	○
Sad	○	○	○	○	○
Angry	○	○	○	○	○
Fearful	○	○	○	○	○

The remaining set of questions has been asked for each individual song (where the song order has been randomised):

- *Happy: Cyndi Lauper - Girls just wanna have fun*
- *Sad: Ashe - Moral of the story*
- *Angry: Halsey - Nightmare*
- *Fearful: Jaws - Theme song*
- *Happy: The Supremes- You can't hurry love*
- *Sad: SYML - Where's my love*
- *Angry: Drowning pool - Bodies*
- *Fearful: John Carpenter - Halloween Theme*

7. Please listen to the *complete song* once. Are you familiar with this song?
O Yes
O No

8. Do you like this song? Please rate it on a scale from 1 (very bad) to 10 (very much).

______

9. Please indicate to which affective state you will *associate* the last song?

O Happy

O Sad

O Angry

O Fearful

10. To what extent do you *feel* the affective states mentioned below induced by the last song. Please rate on a scale from 1 to 5.

	1. Not at all	2. Slightly	3. Moderately	4. Very much	5. Extremely
Happy	○	○	○	○	○
Sad	○	○	○	○	○
Angry	○	○	○	○	○
Fearful	○	○	○	○	○

Appendix F

Pilot study results

Table F.1: Summarised results of the ability of songs to induce an affective state (first iteration)

Song	Incl visuals	Intended state	Frequency associated state	Average experienced state
Girls just wanna have fun	Yes	Happy	Happy: 5* Sad: 0 Angry: 0 Fearful: 0	Happy: 4.2* Sad: 1.6 Angry: 1.8 Fearful: 1.6
You can't hurry love	No	Happy	Happy: 5* Sad: 0 Angry: 0 Fearful: 0	Happy: 4.4* Sad: 1.4 Angry: 1.0 Fearful: 1.4
Moral of the story	Yes	Sad	Happy: 1 Sad: 4* Angry: 0 Fearful: 0	Happy: 2.2 Sad: 3.4* Angry: 2.2 Fearful: 2.2
Where's my love (acoustic)	No	Sad	Happy: 1 Sad: 3* Angry: 0 Fearful: 1	Happy: 2.2 Sad: 3.4* Angry: 1.4 Fearful: 2.0
Nightmare	Yes	Angry	Happy: 0 Sad: 0 Angry: 5* Fearful: 0	Happy: 1.6 Sad: 3.2 Angry: 3.8* Fearful: 1.8
Ready for it	No	Angry	Happy: 0 Sad: 1 Angry: 3* Fearful: 1	Happy: 2.4 Sad: 2.2 Angry: 2.6* Fearful: 2.4
Play with fire	Yes	Fearful	Happy: 0 Sad: 1 Angry: 3* Fearful: 1	Happy: 1.6 Sad: 2.4 Angry: 2.6 Fearful: 3*
Ghost forest	Yes	Fearful	Happy: 0 Sad: 2* Angry: 1 Fearful: 2*	Happy: 1.4 Sad: 2.6 Angry: 2.2 Fearful: 3.0*
Wires	No	Fearful	Happy:0 Sad: 2* Angry: 2* Fearful: 1	Happy: 1.8 Sad: 2.8* Angry: 2.8* Fearful: 2.4
The hidden chamber	No	Fearful	Happy: 0 Sad: 4* Angry: 1 Fearful: 0	Happy: 2.2 Sad: 3.0* Angry: 2.0 Fearful: 2.2

Table F.2: Summarised results of the ability of songs to induce an affective state (second iteration)

Song	Intended state	Frequency associated state	Average experienced state
Jaws Theme	Fear	Happy: 1 Sad: 1 Angry: 1 Fearful: 4*	Happy: 1.7 Sad: 1.6 Angry: 1.7 Fearful: 3.1*
Bodies	Anger	Happy: 0 Sad: 0 Angry: 7* Fearful: 0	Happy: 1.7 Sad: 1.7 Angry: 3.4* Fearful: 2.3
Climbing walls	Fear	Happy: 0 Sad: 5* Angry: 1 Fearful: 1	Happy: 1.4 Sad: 3.1* Angry: 2.1 Fearful: 2.0

Table F.3: Summarised results of the ability of songs to induce an affective state (third iteration)

Song	Intended state	Frequency associated state	Average experienced state
Jaws Theme	Fear	Happy: 0 Sad: 1 Angry: 0 Fearful: 6*	Happy: 1.6 Sad: 2.1 Angry: 1.6 Fearful: 3.1*

Appendix G

Overview processed data

Table G.1: Overview of collected and processed data per participant

Participant	Survey	mmWave	WiFi-CSI	RGB	Wristband	HRV features
1	✓	✗	✗	✗	✓	✓
2	✓	✓	✓	✓	✓	✓
3	✓	✓	✓	✓	✓	✓
4	✓	✓	✓	✓	✓	✓
5	✗	N/A	N/A	✓	N/A	N/A
6	✓	✓	✓	✓	✓	✓
7	✓	✓	✓	✓	✓	✗
8	✓	✓	✓	✓	✓	✓
9	✓	✗	✓	✓	✓	✗
10	✓	✓	✓	✓	✓	✓
11	✓	✓	✗	✓	✓	✗
12	✓	✓	✓	✓	✓	✓
13	✓	✓	✓	✓	✓	✓
14	✓	✓	✗	✓	✓	✗
15	✓	✓	✓	✓	✓	✓
16	✓	✓	✓	✓	✓	✓
17	✓	✓	✓	✓	✓	✓
18	✓	✓	✓	✓	✓	✓
19	✓	✓	✓	✓	✓	✓
20	✓	✓	✓	✓	✓	✓
21	✓	✓	✓	✗	✓	✓
22	✓	✓	✓	✓	✓	✓
23	✓	✓	✓	✓	✓	✓
24	✓	✓	✓	✓	✓	✓

Table G.1: Overview of collected and processed data per participant

Participant	Survey	mmWave	WiFi-CSI	RGB	Wristband	HRV features
25	✓	✓	✓	✓	✓	✓
26	✓	✓	✓	✓	✓	✓
27	✓	✓	✓	✓	✓	✓
28	✓	✓	✓	✓	✓	✓
29	✓	✓	✓	✓	✓	✓
30	✓	✓	✓	✓	✓	✓
31	✓	✓	✓	✓	✓	✓
32	✓	✓	✓	✓	✓	✓
33	✓	✓	✓	✓	✓	✓
34	✓	✓	✓	✓	✓	✓
35	✓	✓	✓	✓	✓	✓
36	✓	✓	✓	✓	✓	✓
37	✓	✓	✓	✓	✓	✓
38	✓	✓	✓	✓	✓	✓
39	✓	✓	✓	✓	✓	✓
40	✓	✗	✓	✗	✓	✓
41	✓	✓	✓	✓	✓	✓
42	✓	✓	✓	✓	✓	✓
43	✓	✓	✓	✓	✓	✗
44	✓	✓	✓	✓	✓	✓
45	✓	✓	✓	✗	✓	✗
46	✓	✓	✓	✓	✓	✓
47	✓	✓	✓	✓	✓	✓
48	✓	✓	✓	✓	✓	✓
49	✓	✓	✗	✓	✓	✓
50	✓	✓	✓	✓	✓	✓
51	✓	✓	✓	✗	✓	✓
52	✓	✓	✓	✓	✓	✓
53	✓	✓	✓	✓	✓	✓
54	✓	✓	✓	✓	✓	✓
	53	50	49	49	53	47

Appendix H

Self-reported Initial state in experiment

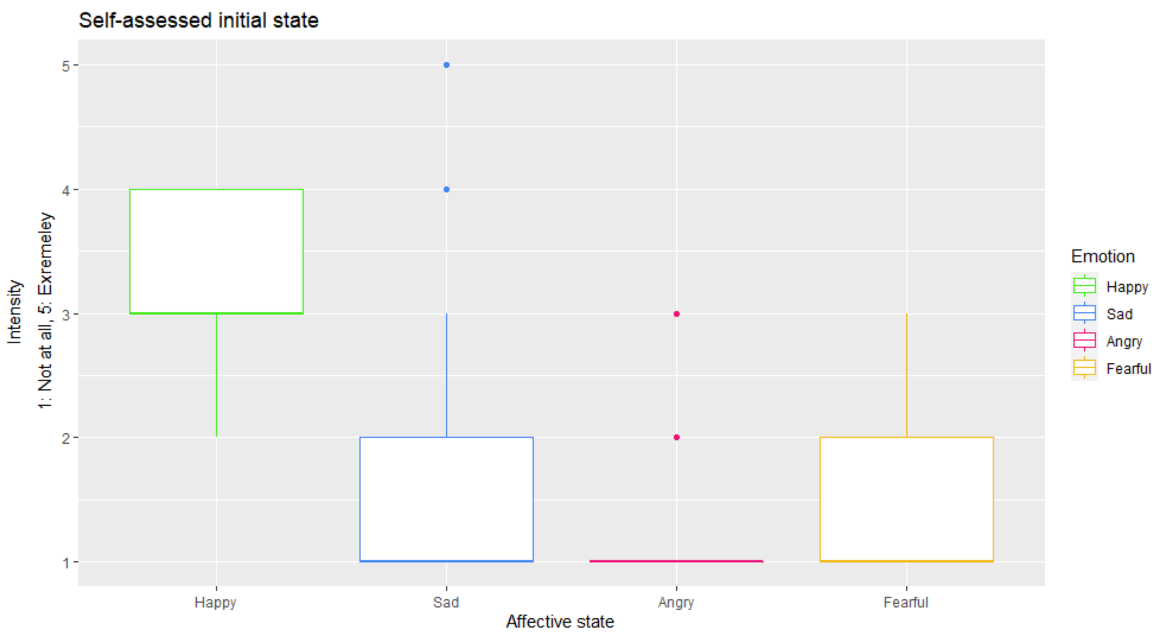

Figure H.1: Summary of the initial state of participants.

Appendix I

Detailed results of vital sign estimations and (absolute) errors.

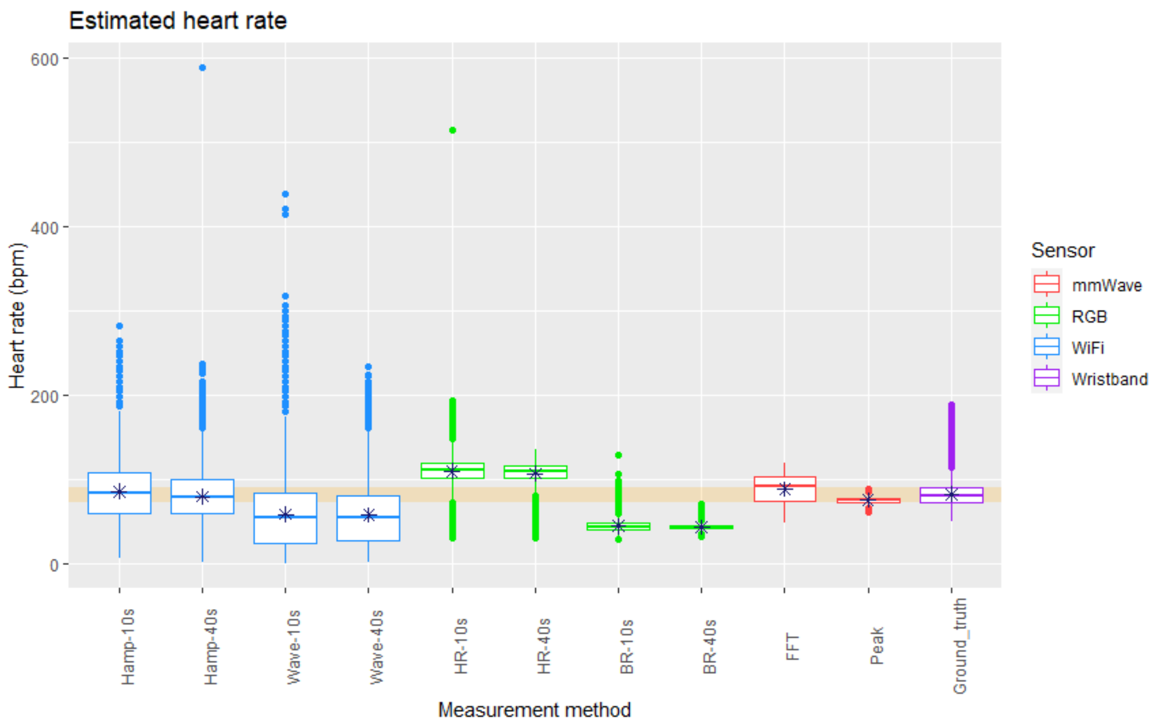

Figure I.1: Heart rate estimations. The interquartile range of the ground truth heart rate is highlighted and the mean is marked with a star

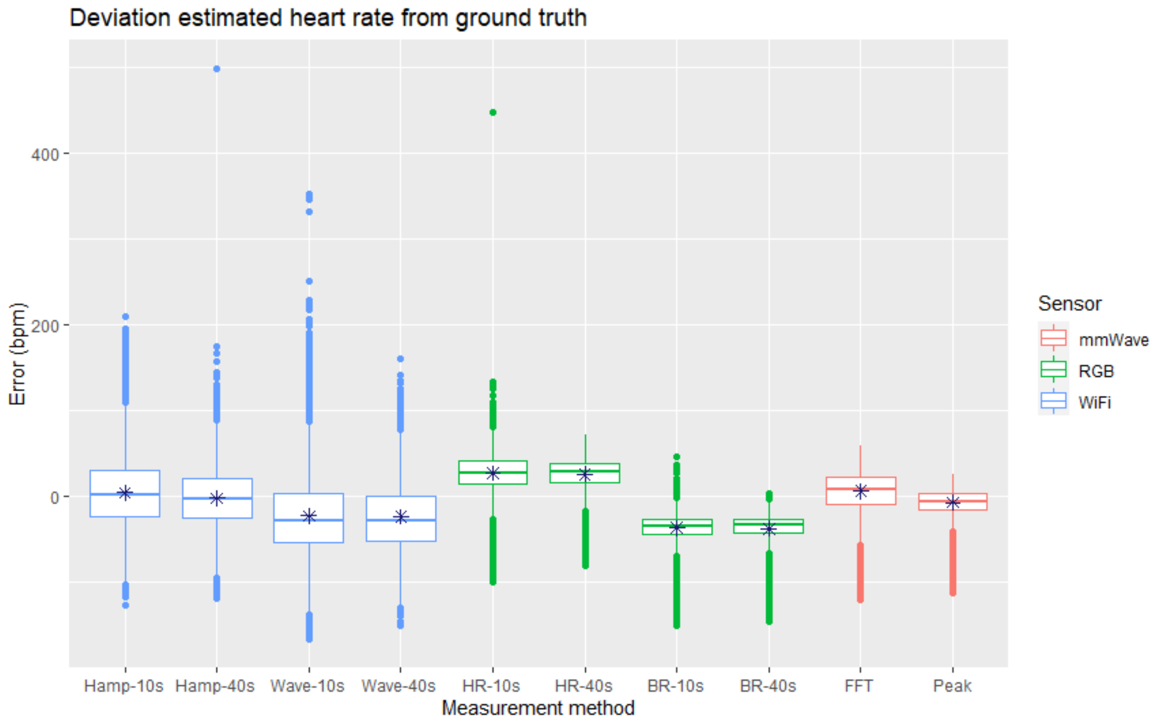

Figure I.2: Error (bpm) in estimating the heart rate per sensor and processing method. The mean is marked with a star

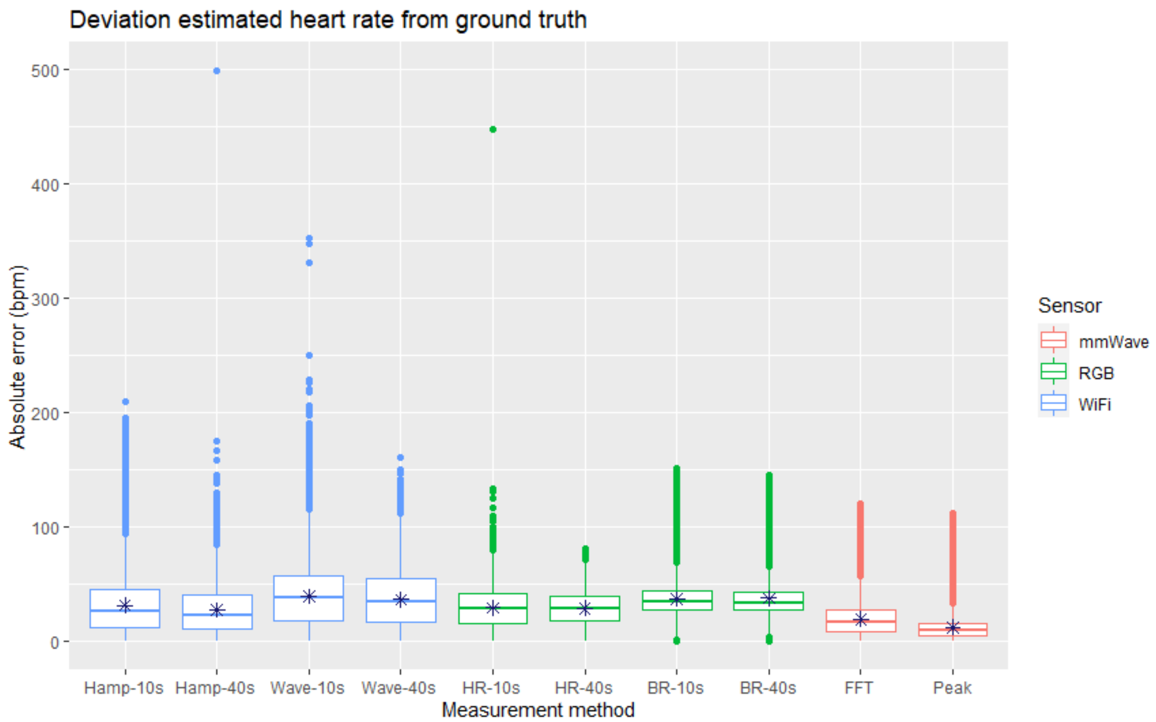

Figure I.3: Absolute error (bpm) in estimating the heart rate per sensor and processing method. The mean is marked with a star

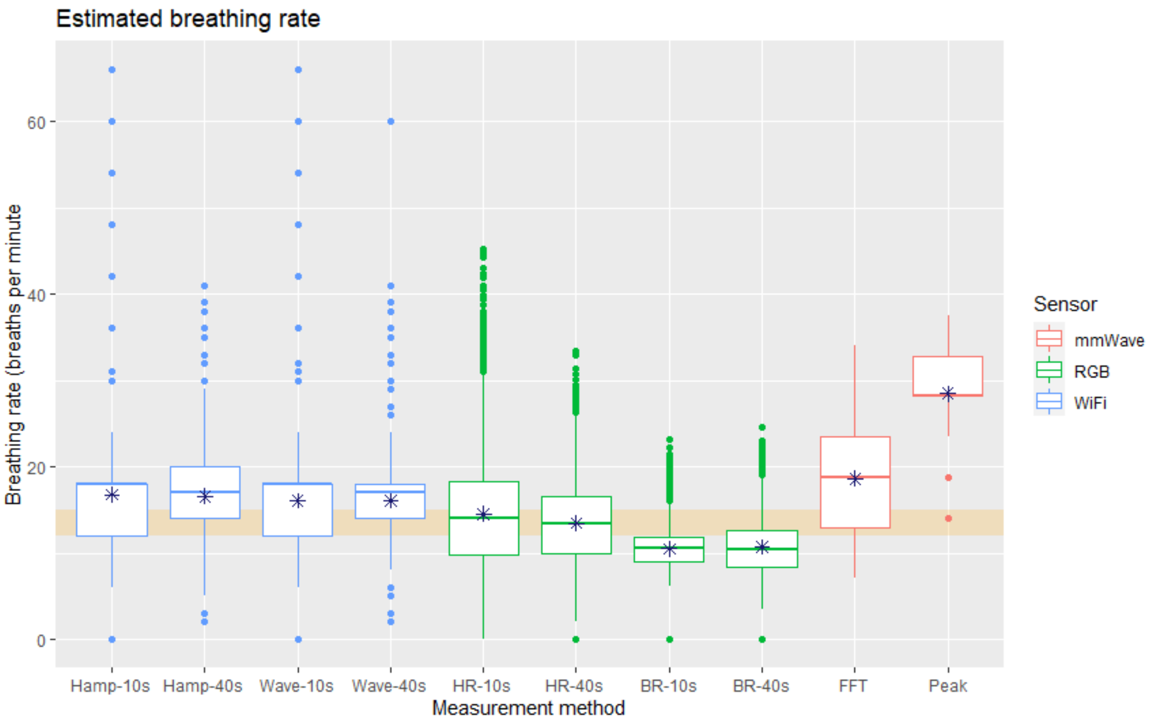

Figure I.4: Breathing rate estimations. The range of a normal human breathing rate is highlighted and the mean is marked with a star

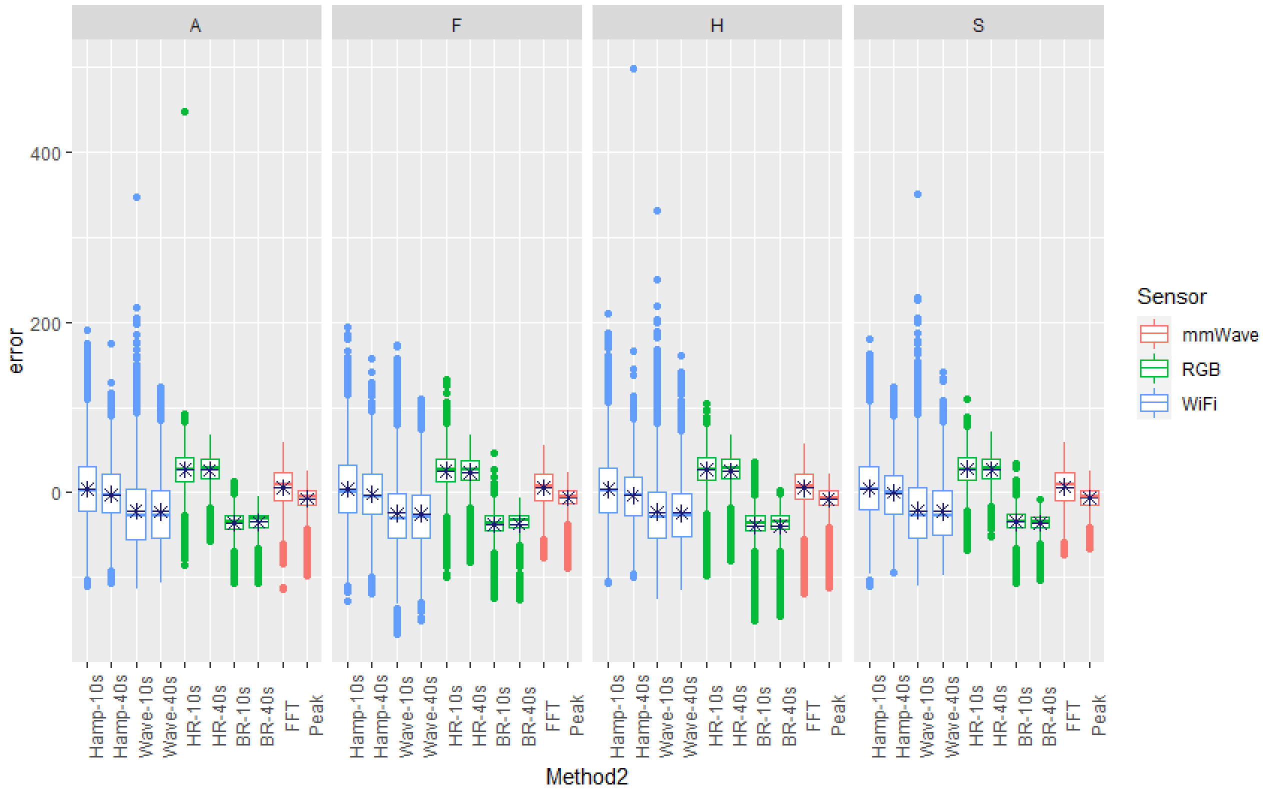

Figure I.5: Comparison of error in estimating the heart rate per intended affective state using different sensing and processing methods. The mean is marked with a star

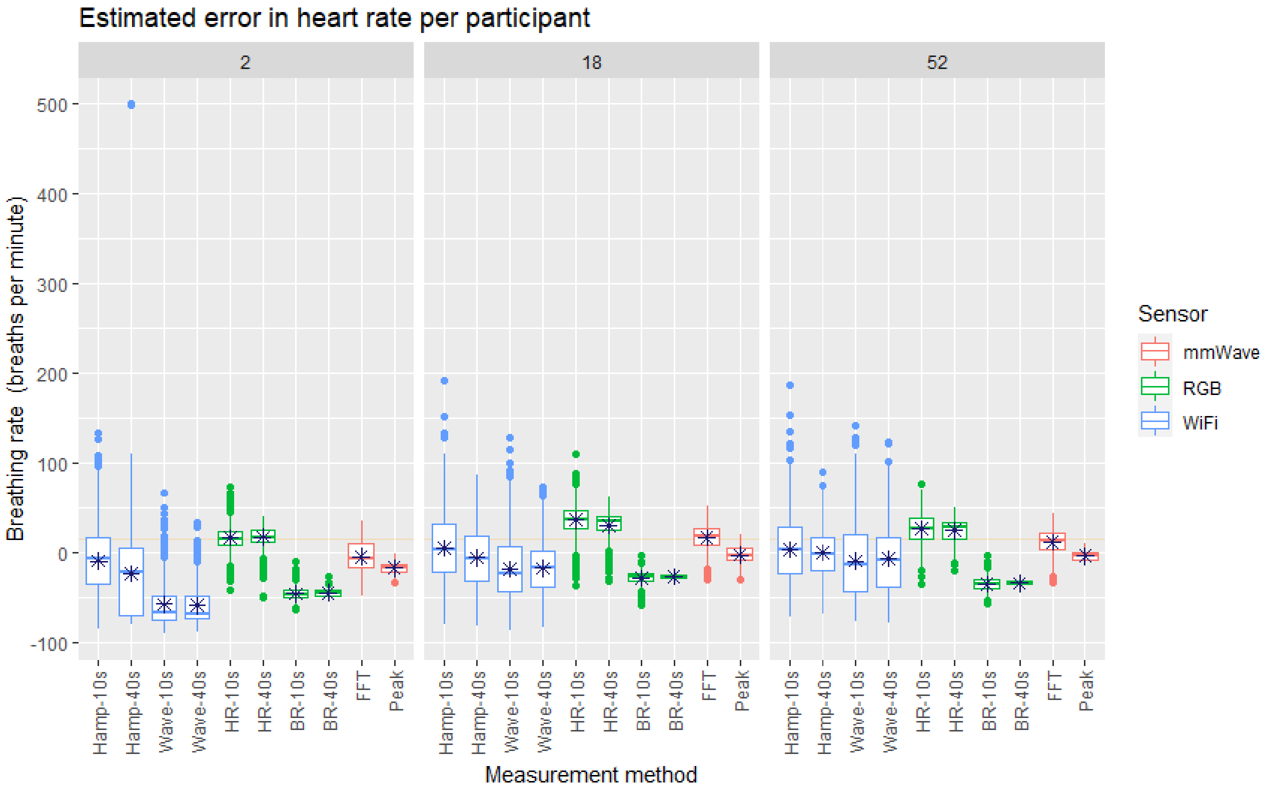

Figure I.6: Comparison of error in estimating the heart rate per participant using different sensing and processing methods. The mean is marked with a star

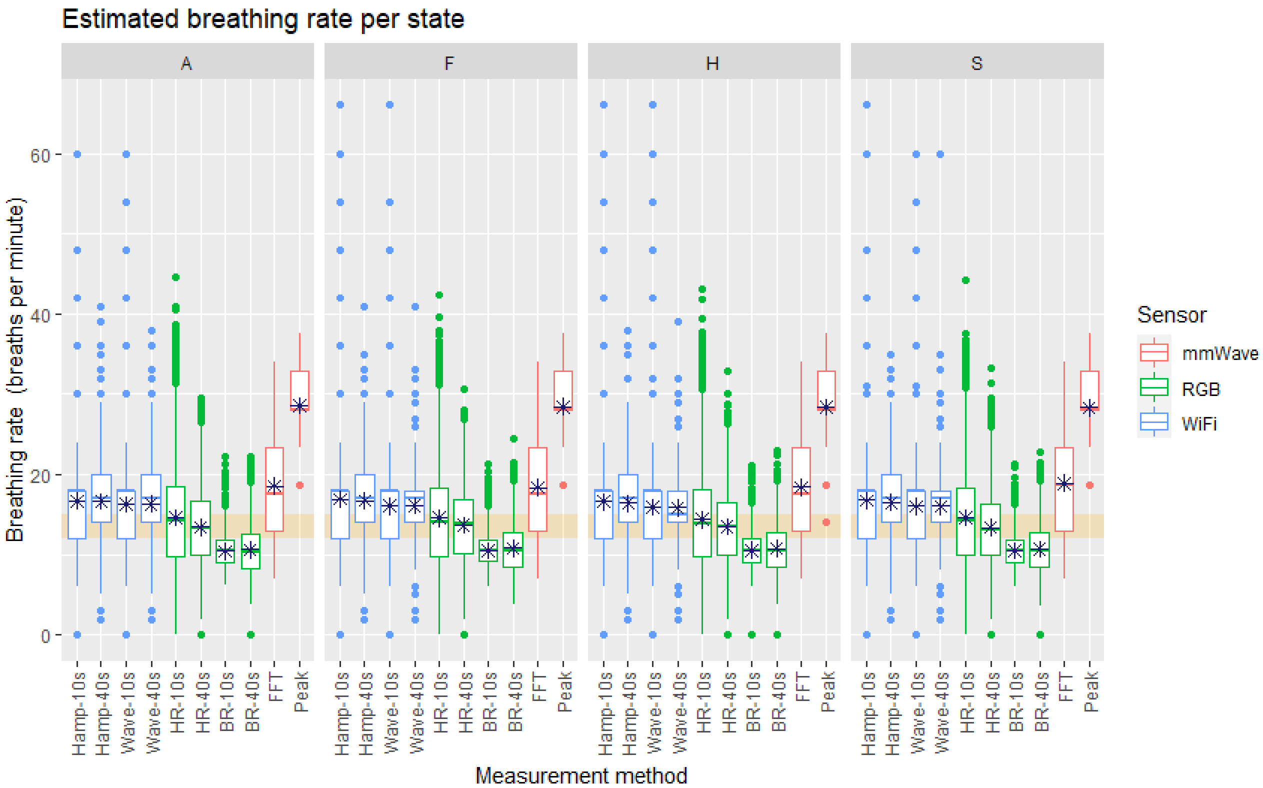

Figure I.7: Comparison of breathing rate estimations per intended affective state using different sensing and processing methods. The mean is marked with a star

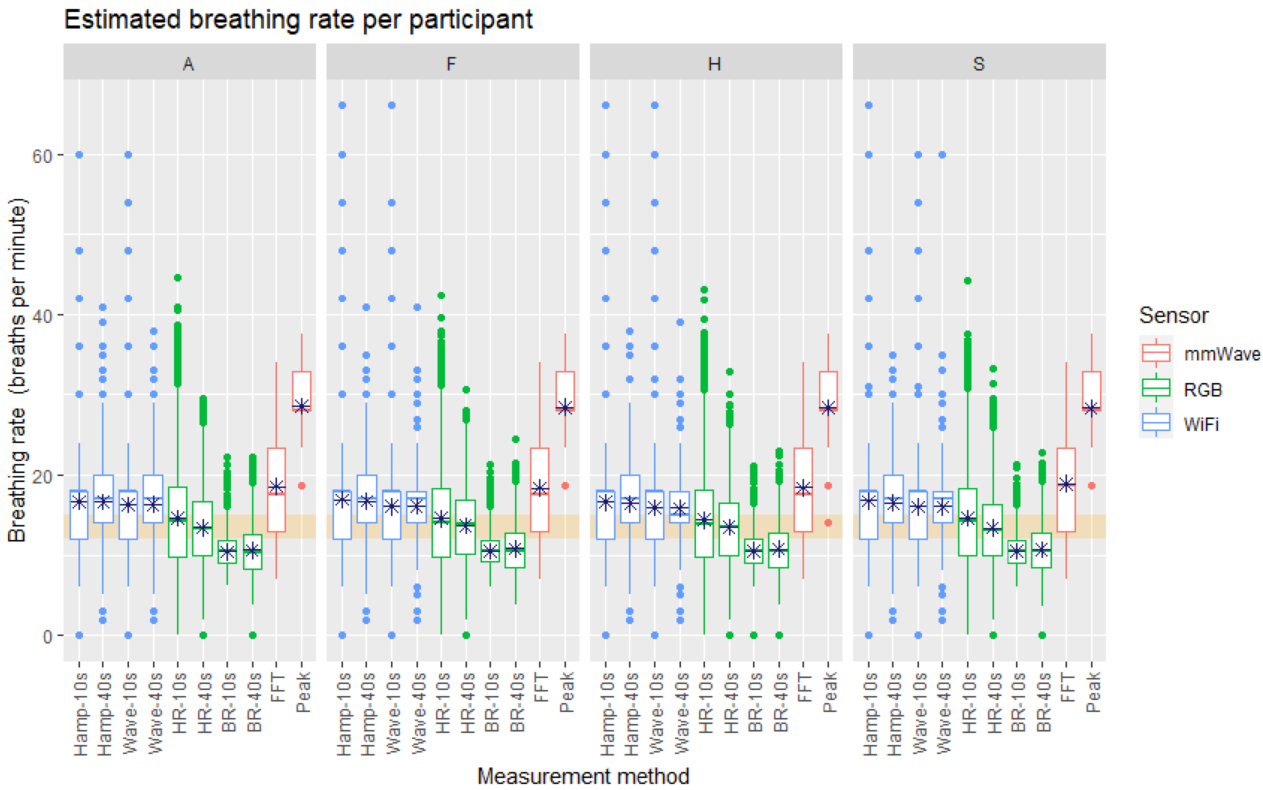

Figure I.8: Comparison of breathing rate estimations per participant using different sensing and processing methods. The mean is marked with a star

Table I.1: Descriptive statistics of heart rate estimations

Sensor	Measuring	Approach	Window	Mean	Lower bound 95% CI	Upper bound 95% CI	1st Quartile	3rd quartile	Median	Variance	Std. Deviation	Min	Max	Range	Interquartile range
Wristband	HR	-	10s	82.32	82.23	82.41	72.30	89.25	80.40	214.85	14.66	49.00	187.93	138.93	16.59
WiFi-CSI	HR	Hampel	10s	85.48	85.21	85.74	60.00	108.00	84.00	1378.79	37.13	6.00	282.00	276.00	48.00
WiFi-CSI	HR	Hampel	40s	79.13	78.91	79.36	59.00	99.00	78.00	1378.13	31.89	2.00	588.00	586.00	40.00
WiFi-CSI	HR	Wavelet	10s	58.47	58.20	58.74	24.00	84.00	54.00	1508.17	38.84	0.00	438.00	438.00	60.00
WiFi-CSI	HR	Wavelet	40s	57.65	57.41	57.90	27.00	80.00	54.00	1176.27	34.30	2.00	233.00	231.00	53.00
RGB	HR	BR-filter	10s	44.41	44.35	44.47	40.22	47.52	43.64	39.48	6.28	28.80	128.57	99.77	7.30
RGB	HR	BR-filter	40s	43.09	43.04	43.14	40.79	45.06	42.86	11.55	3.40	32.14	70.59	38.45	4.27
RGB	HR	HR-filter	10s	109.54	109.43	109.64	100.62	119.67	110.49	231.71	15.22	30.51	514.29	483.78	19.05
RGB	HR	HR-filter	40s	106.76	106.66	106.87	101.62	115.85	109.74	197.92	14.07	31.03	135.89	104.86	14.23
mmWave	HR	Peak	-	75.21	75.19	75.23	72.66	77.34	75.00	13.12	3.62	60.94	89.06	28.12	4.68
mmWave	HR	FFT	-	88.50	88.39	88.61	73.83	103.12	91.41	321.69	17.94	48.05	118.36	70.31	29.29

Table I.2: Descriptive statistics of error in heart rate estimations

Sensor	Measuring	Approach	Window	Mean	Lower bound 95% CI	Upper bound 95% CI	1st Quartile	3rd quartile	Median	Variance	Std. Deviation	Min	Max	Range	Interquartile range
WiFi-CSI	HR	Hampel	10s	4.20	3.92	4.48	-22.93	29.87	2.38	1552.50	39.40	-127.45	209.82	337.27	52.80
WiFi-CSI	HR	Hampel	40s	-2.17	-2.42	-1.92	-25.65	20.22	-2.80	1183.93	20.88	-119.40	498.42	617.82	45.87
WiFi-CSI	HR	Wavelet	10s	-22.81	-23.10	-22.52	-53.88	2.53	-29.22	1658.40	40.72	-166.58	351.95	518.53	56.41
WiFi-CSI	HR	Wavelet	40s	-23.65	-23.91	-23.40	-51.98	0.18	-27.87	1319.24	36.32	-150.27	160.90	311.17	52.16
RGB	HR	BR-filter	10s	-36.74	-36.89	-36.58	-43.87	-26.97	-35.09	251.32	15.85	-151.05	46.57	197.62	16.90
RGB	HR	BR-filter	40s	-37.54	-37.79	-37.28	-42.61	-26.99	-33.74	332.39	18.23	-145.20	2.82	148.02	15.62
RGB	HR	HR-filter	10s	26.96	26.82	27.11	14.08	40.91	27.41	430.55	20.75	-100.04	447.82	547.86	26.83
RGB	HR	HR-filter	40s	25.96	25.81	26.11	16.23	38.93	28.14	370.78	19.26	-81.16	71.77	152.93	22.70
mmWave	HR	Peak	-	-7.27	-7.37	-7.18	-15.02	2.87	-5.54	226.86	15.06	-112.37	26.00	138.37	17.89
mmWave	HR	FFT	-	6.02	5.87	6.16	-9.08	22.70	8.23	519.95	22.80	-120.02	59.26	179.28	31.78

Table I.3: Descriptive statistics of absolute error in heart rate estimations.

Sensor	Measuring	Approach	Window	Mean	Lower bound 95% CI	Upper bound 95% CI	1st Quartile	3rd quartile	Median	Variance	Std. Deviation	Min	Max	Range	Interquartile range
WiFi-CSI	HR	Hampel	10s	31.30	31.13	31.47	12.17	45.10	26.22	590.26	24.30	0.00	209.82	209.82	32.93
WiFi-CSI	HR	Hampel	40s	27.43	27.29	27.58	10.62	39.93	23.12	435.99	20.88	0.00	498.42	498.42	29.31
WiFi-CSI	HR	Wavelet	10s	39.28	39.10	39.45	18.27	57.18	37.82	636.04	25.22	0.00	351.95	351.95	38.91
WiFi-CSI	HR	Wavelet	40s	36.45	36.28	36.61	16.15	54.37	34.73	550.21	23.46	0.00	160.90	160.90	38.22
RGB	HR	BR-filter	10s	36.79	36.64	36.95	26.97	43.88	35.10	247.31	15.72	0.03	151.05	151.02	16.91
RGB	HR	BR-filter	40s	37.54	37.28	37.79	26.99	42.61	33.74	332.37	18.23	0.43	145.20	144.77	15.62
RGB	HR	HR-filter	10s	29.385	29.26	29.51	15.73	41.36	28.12	294.07	17.15	0.00	447.82	447.82	26.63
RGB	HR	HR-filter	40s	28.85	28.81	29.03	18.02	39.30	28.92	208.62	14.44	0.00	81.16	81.16	21.28
mmWave	HR	Peak	-	12.14	12.07	12.21	4.42	16.00	9.24	132.35	11.50	0.00	112.37	112.37	11.58
mmWave	HR	FFT	-	19.35	19.27	19.43	8.61	27.86	17.24	181.72	13.48	0.00	120.02	120.02	19.25

Table I.4: Descriptive statistics of breathing rate estimations.

Sensor	Measuring	Approach	Window	Mean	Lower bound 95% CI	Upper bound 95% CI	1st Quartile	3rd quartile	Median	Variance	Std. Deviation	Min	Max	Range	Interquartile range
WiFi-CSI	BR	Hampel	10s	16.75	16.71	16.79	12.00	18.00	18.00	32.18	5.67	0.00	66.00	66.00	6.00
WiFi-CSI	BR	Hampel	40s	16.58	16.55	16.61	14.00	20.00	17.00	16.27	4.03	2.00	41.00	39.00	6.00
WiFi-CSI	BR	Wavelet	10s	16.08	16.04	16.12	12.00	18.00	18.00	34.26	5.85	0.00	66.00	66.00	6.00
WiFi-CSI	BR	Wavelet	40s	16.08	16.05	16.11	14.00	18.00	17.00	17.86	4.23	2.00	60.00	58.00	4.00
RGB	BR	BR-filter	10s	10.50	10.45	10.55	9.02	11.85	10.56	5.27	2.30	0.00	23.13	23.13	2.83
RGB	BR	BR-filter	40s	10.71	10.66	10.76	8.38	12.68	10.47	9.97	3.16	0.00	24.55	24.55	4.30
RGB	BR	HR-filter	10s	14.53	14.49	14.57	9.82	18.29	14.07	30.41	5.51	0.00	45.24	45.24	8.47
RGB	BR	HR-filter	40s	13.46	13.43	13.50	10.00	16.63	13.40	20.06	4.48	0.00	33.34	33.34	6.36
mmWave	BR	Peak	-	28.47	28.45	28.49	28.12	32.81	28.12	12.03	3.47	14.06	37.50	22.90	4.69
mmWave	BR	FFT	-	7.03	18.59	18.67	12.89	23.44	18.75	44.55	6.67	7.03	33.98	26.95	10.55