Identification of dental pain sensation based on cardiorespiratory signals

Teichmann, Daniel; Hallmann, Alexander; Wolfart, Stefan; Teichmann, Maren

Published in:
Biomedizinische Technik

DOI:
10.1515/bmt-2020-0047

Publication date:
2021

Document version:
Final published version

Citation for published version (APA):

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use
This work is brought to you by the University of Southern Denmark. Unless otherwise specified it has been shared according to the terms for self-archiving. If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying this open access version.

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim. Please direct all enquiries to puresupport@bib.sdu.dk
Identification of dental pain sensation based on cardiorespiratory signals

Abstract: The aim of this study is to investigate the feasibility of the detection of brief periods of pain sensation based on cardiorespiratory signals during dental pain triggers. Twenty patients underwent dental treatment and reported their pain events by pressing a push button while ECG, PPG, and thoracic effort signals were simultaneously recorded. Potential pain-indicating features were calculated from the physiological data (sample length of 6 s) and were used for supervised learning of a Random forest pain detector. The best feature combination was determined by Feature forward selection. The best feature combination comprises nine feature groups consisting of four respiratory and five cardiac related groups. The final algorithm achieved a sensitivity of 87% and a specificity of 63% with an AUC of 0.828. Using supervised learning it is possible to train an algorithm to differentiate between short time intervals of pain and no pain solely based on cardiorespiratory signals. An on-site and real-time detection and rating of pain sensations would allow a precise, individuum- and treatment-tailored administration of anesthetics (optimized anesthetic management). Severe periods of pain could be interrupted or avoided by additional anesthetic administration, this would allow more comfortable treatment and thereby, yield a better patient compliance. Additionally, the optimized anesthetic management would be beneficial for medicinal pre-treated patients to reduce the risk of interaction with other drugs (e.g. antidepressants or blood pressure-lowering agents). Also, patients who are not able to communicate could also benefit from an automatic pain indication.

Keywords: machine learning; pain detection; physiological signals.

Introduction

The high occurrence of brief but intense pain sensations during dental treatment is reflected by a high number of patients that suffer from severe fear of dental treatment (about 5–15% of adults in the industrialized nations [1]). For alleviation of such pain sensations, local anesthetics are applied to the mucosa or injected. Here, a minimal dosage is desirable, i.e. drug administration beyond the necessary depth of anaesthesia should be avoided.

An on-site and real-time detection and rating of pain sensations would allow a precise, individuum- and treatment-tailored administration of anesthetics (optimized anesthetic management). Severe periods of pain could be interrupted or avoided by additional anesthetic administration, this would allow more comfortable treatment and, therewith, yield a better patient compliance. Additionally, the optimized anesthetic management would be beneficial for medicinal pre-treated patients to reduce the risk of interaction with other drugs (e.g. antidepressants or blood pressure-lowering agents). Also, patients who are not able to communicate could also benefit from an automatic pain indication.

The response to pain is governed by various changes in metabolism, hormone balance [2], and activation shifts within the autonomous nervous system [3]. The aim of this study is to investigate the feasibility of the detection of brief pain sensations based on cardiorespiratory signals during dental treatment. Twenty patients underwent dental pain triggering and reported their pain events by pressing a push button while cardiorespiratory signals were simultaneously recorded. Potential pain-indicating features were calculated from the physiological data and (together with the pain event annotations) were used for supervised learning of a pain detector. In the Results section the performance of the algorithm is presented, and the best set of physiological features for accurate pain detection is determined.

Various pain measurement strategies that are based on physiological signals have been proposed in the literature and some of them have been translated into commercially available devices. The most established ones are the Analgesia Nociception Index (ANI) [4, 5], the Autonomic Nervous System State Index (ANSSI) [6], the Noxious Stimulation Response Index (NSRI) [7], the Number of Fluctuations of Skin Conductance (NFSC) [8], and the
Surgical Pleth Index (SPI) [9, 10]. The aim of such pain indices is to maintain the anesthesia depth during surgical procedures under general anesthetic. Therefore, all of them have in common that they have been developed for slowly evolving and/or long-lasting pain events and, hence, can only be applied to long time windows of ≥60 s. This makes them unsuitable for dental treatment where mostly short but intense pain sensations occur. Furthermore, the mentioned indices are based on physiological changes that arise during (and, therewith, possibly in interactions with) strong anesthetics (Propofol, Remifentanil, Fentanyl, Sevoflurane) and other drugs (Diazepam, Rocuronium, Cisatracurium) which are seldomly used in ambulant dental treatment.

In a previous study, our group could show that a detection of very short periodontal pain sensations based on Electrocardiography (ECG) and Photoplethysmography (PPG) data is feasible [11]. However, the data used in this previous study were not ideal as the total number of pain samples was quite low and the distribution of pain samples per patient among the study population was very unequally distributed. For the present work, a new clinical study with an improved study protocol has been conducted, with the goal to yield a more equally distributed pain sensation among all patients to allow better generalizing of the pain classifier. The main improvements are: (i) Periodontal probing was solely conducted on the Ramfjord teeth, which are representative for the entire jaws [12]. (ii) Additional to periodontal probing, cold thermal testing of tooth pulp vitality (cold test) [13, 14] has been conducted. This procedure was expected to ensure pain perception during each application and, hence, yield an equally distributed pain sensation among all patients as it was applied to the same number (six) of vital anterior teeth per patient.

In the mentioned previous study, we initially planned to also use respiration data, but such data had to be excluded from the analysis as the respiration sensor turned out to provide very low signal quality with an immense amount of signal loss. Furthermore, the type of ECG lead was not consistent throughout the patients. In the present work, we have eliminated these problems, and the pain detection will base on respiratory features as well.

Materials and Methods

Data collection

Protocol: All evaluations in this work are based on data acquired from a monocentric clinical trial (ethical approval no. EK287/15) conducted at the Department of Prosthodontics and Biomaterials - Center of Implantology, Medical Faculty, RWTH Aachen University (Aachen, Germany). Patients aged ≥20 years who had at all Ramfjord teeth and incisors without crowns and were compos mentis were included in the study as long as none of the following exclusion criteria applied:

- neurologic disease, heart disease, uncontrolled hypertension
- current or past regular abuse of alcohol or drugs
- metallic or electronic implants
- current use of hypnotic, sedative, antidepressive, neuroleptic, or anti-allergic drugs as well as of tranquilizers, analgesics, stimulants, or any drugs affecting heart rate.

Twenty individuals (16 males and 4 females) aged 19–54 years and weighing 63.1–122.9 kg participated in the study. For all participants the same study protocol was applied (see Figure 1).

During the entire study each participant held a push button and was asked to press the button when and for as long as (s)he felt pain. The signal of the push button was recorded simultaneously with the patient’s cardiorespiratory signals. After a baseline phase of 5 min, periodontal probing was performed, i.e. gently inserting a probe (0.2 N pressure-calibrated probe: Kerr-Hawe, Raststatt, Germany) into the sulcus formed by a tooth and its surrounding tissue. Periodontal probing is a common but potentially painful diagnostic method in periodontology. This probing was conducted at six positions around each of the six Ramfjord teeth. Afterwards, a pulp vitality check was conducted for each of the eight incisors (4 upper, 4 lower ones). Here, a cold stimulus (−45 °C ORBIS cold spray, Dental Handels GmbH, Münster, Germany) is used as a provocation test to elicit a short, sharp pain sensation. In dentistry, pulp vitality is usually tested for pulp disease diagnostic.

For the acquisition of physiological data a Somnolab 2 device (Weinmann Emergency Medical Technology GmbH + Co. KG, Hamburg, Germany) was used. The recorded physiological signals were: a Lead-II ECG with a sampling frequency of f_s = 256 Hz, a PPG with f_s = 50 Hz acquired with a transmissive transducer at the left index finger, and a respiratory signal (piezoelectric thoracic effort sensor) with f_s = 30 Hz.

Segmentation and labeling of the time series data

The continuous time series of physiological signals are segmented into smaller time intervals (samples) using a sliding window with a length of 6 s and a 5 s overlap. Each sample is categorized as a ‘pain’ or ‘no pain’ sample. The modulation of physiological signals by pain will take place after the pain sensation, therefore, ‘pain’ intervals should cover this phase. On the other hand there likely will be a certain delay between the beginning of the actual pain sensation and the activation of the push button due to unknown reaction times of individual subjects. To take this into account, the following labeling method was applied for the categorization of samples as ‘pain’ or ‘no pain’ samples: For each indicated pain sensation a pain region was defined which starts with the activation of the push button and ends 11 s after this start point. Each of the 6 s sample intervals that completely lies within such a pain region is labeled as a ‘pain’ sample while each
and a few seconds, segmentation of the data into short time intervals yields a considerably smaller number of ‘pain’ samples compared with those of ‘no pain’ samples. Figure 2 presents the number of collected ‘pain’ and ‘no pain’ samples per person.

Classification approach

In this study, pain detection is achieved by using a binary Random Forest classifier which assigns the periods of the recorded physiological signals to one of the two classes ‘pain’ and ‘no pain’. This classifier has been chosen because it achieved the best results for vital sign based pain detection when compared to other classifiers in our previous study [11]. For this purpose, features are extracted from the physiological signals which enable the discrimination between ‘pain’ and ‘no pain’. These features are time-discrete quantities describing the underlying autonomic mechanisms within the continuous physiological signals. For each input sample a feature vector is calculated. The Features section provides information on the utilized features.

We separated the learning and testing of the classifier into two phases: validation and test phase (according to e.g. [15]). For this purpose, at the very beginning of this work, a test data set was separated by randomly choosing three patients. The rest of the patient data was used as the validation data set. This resulted in 1,991 ‘pain’ and 27,888 ‘no pain’ samples for validation and 358 ‘pain’ and 5,714 ‘no pain’ samples for the final test.

During the validation phase, the mapping is learned empirically from a given data set, i.e. a set of pairs (feature vector, class label) by minimizing a loss function. Following, this mapping is validated. Based on the resulting performance the parameters of the model underlying the classifier are adjusted and training and validation are repeated until the performance is in an acceptable range. During the validation phase the best feature combination was determined by using leave-one-(patient)-out cross validation (LOOCV) and Feature Forward selection (FFS). To account for the unbalanced data set, each LOOCV-step was repeated five times, each time with a number of randomly chosen ‘no pain’ samples equaling the number of ‘pain’ samples (5-fold boot strapping). The resulting algorithm will yield a value (score) for each input sample on whose basis the classification can be made. To further increase the classification performance, this score can be averaged over n consecutive samples (score merging). This yields a reduced classification variance, but increases the effective reaction time by 1 s per merged sample. Since this is a trade-off between prediction accuracy and reaction time, the best value for n was determined as a last step of the validation phase.

During the test phase, the algorithm and its parameters and feature set determined during the validation phase are eventually used for a concluding test with the separate test data set. To do so, the classification algorithm that resulted from the validation phase was trained on the entire validation data and afterwards tested on the independent test set. For this final test, a score threshold above which the classifier assumes input samples as ‘no pain’ intervals has to be defined. In order to conduct unbiased detection without a priori usage of the test data, the optimal operating point for the receiver operating characteristic (ROC) achieved with the validation data set was determined beforehand and then applied to the ROC of the actual test data. The point closest to the point (0, 1) was considered as optimal.

Features

Features were normalized to (i) decrease the probability of providing identification of single patients, and to (ii) address different signal amplitudes among the patients due to different sensor positioning. For this purpose, z-transformation was applied to all features. Since the amount of pain sensations varied among the patients and it was assumed that signal amplitudes are affected by pain sensation, only the standard deviation of the baseline phase was used for the z-transformation.

Frequency spectral bins (FB)

The frequency spectrum of the PPG, ECG, and respiration signal was computed once for each 6 s input sample by fast Fourier transform. This yielded 768 frequency bins with a maximum frequency of 128 Hz for the ECG signal (FBECG), 125 frequency bins with a maximum frequency of 25 Hz for the PPG signal (FBPPG), and 37 frequency bins with a maximum frequency of 15 Hz for the respiration signal (FBRESP), where i specifies the bin no. From each spectrum the first 16 bins were used as features.

Levels of discrete wavelet transform (WL)

The absolute wavelet energies of the first eight levels decomposed by the Discrete wavelet transform [16] for each 6 s input sample were calculated for the ECG (WLECG), PPG (WPPG), and respiration (WLRESP) signal using the Reverse biorthogonal spline wavelet (Matlab name: rbio3.3), Biorthogonal Wavelet (Matlab name: bior3.3), and Fejér-Korovkin wavelet (Matlab name: fkt8), respectively. i denotes the wavelet level no.

Features based on peak values

The average height (H) and the maximum deviation in height (ΔH) of the ECG’s R waves as well as the PPG’s pulse waveform within each 6 s sample interval have been used as features.
Features based on beat-to-beat times

The average pulse beat-to-beat time (BB) as well as the maximum deviation in beat-to-beat times (ΔBB) of the ECG and PPG signal have been used as features.

Features based on the shape of the PPG pulse waveform

The average area under the PPG’s pulse waveform (A) as well as the maximum deviation of such areas (maxΔA) within each 6 s sample interval have been used as features. The average ratio between pulse width and height (W2H) as well as the maximum deviation of such ratio (maxΔW2H) within each 6 s sample interval have been used as features.

Results

Validation

Before performing the FFS, the features described in the Materials and Methods Chapter were summarized into feature groups. The FB and WL features were each reordered by Battacharyya distance (BD) and then combined into feature groups by groups of four. Three further feature groups were formed by those features based on peak values, beat-to-beat times, and the PPG shape, respectively. Figure 3 shows the BD for each feature and Table 1 shows the resulting list of feature groups.

In each step of the FFS, the addition of all feature groups is tested separately and the feature group yielding the highest Area under the curve (AUC) is added to the feature set for the next step. The result of the FFS is presented in Table 2. The highest AUC = 0.72 was achieved in step 9 with a feature set consisting of feature groups 1, 2, 5, 6, 10, 12, 15, 17, 18.

As it can be seen in Table 3, by merging scores of neighboring, overlapping samples (see Materials and Methods), it was possible to push the AUC to higher levels. Although averaging over n = 5 samples yielded the highest AUC, n = 3 was chosen for the final algorithm, since this will yield a faster reaction time while the AUC is still comparable good.

Test

Applying the final pain detection algorithm to the entire test data set yields the ROC curve presented in Figure 4 with an overall AUC of 0.828.

As described in Materials and Methods, the operating point along this curve was chosen based on the ROC of the validation data set (indicated in Figure 4 by a red circle). This way, the pain detection algorithm achieved a sensitivity of 87% and a specificity of 63%. Table 4 presents the performance measures for each of the test patients separately.
Discussion

With the maximum harmonic mean between sensitivity and specificity chosen as the operation point, the final algorithm yielded a sensitivity and specificity of approximately 87 and 63%, respectively. Although this result demonstrates the feasibility of the detection of brief pain sensations based on cardiorespiratory signals, the performance might be too low for clinical use. However, the performance might be adjusted: The working point of the algorithm along the ROC curve could be varied yielding higher values for either sensitivity or specificity. Nevertheless, ultimately, a higher accuracy would be ideal as this would allow high sensitivity and specificity at the same time. The calculation of a time-averaged pain intensity level by accumulation of the detected short pain sensations would possibly be less sensitive to false negatives and allow an higher accuracy. This would give a more general information about the patient’s level of pain sensation during e.g. the last couple of minutes or along the entire procedure instead of information about a short single pain sensation at a specific point in time.

In this study, the Random forest classifier was used for pain detection. This choice was based on the results of our previous study [11], where different classifiers (Artificial neural networks, Support vector machine, K-nearest Neighbors, Random forest) were compared with the Random forest achieving the best performance. However, other classifiers when properly trained might achieve similar or better performances.

Beyond cardiorespiratory signals, there exists a variety of further physiological signals which are easy to record and have been demonstrated to be linked to pain sensation. Although such measures (e.g. electrodermal activity, skin temperature, and electromyography of the face or the trapezius muscle [17]) are relatively easy to measure, the instrumentation necessary for their recording was considered as not as common. Nevertheless, based on the results of the present work detection of brief stomatologic pain based on ECG, PPG, and respiratory effort seems feasible.

Table 2: Result of the feature forward selection process. Best step is underlined.

<table>
<thead>
<tr>
<th>FFS step</th>
<th>Added group</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Group 1</td>
<td>0.625</td>
</tr>
<tr>
<td>2</td>
<td>Group 17</td>
<td>0.667</td>
</tr>
<tr>
<td>3</td>
<td>Group 18</td>
<td>0.684</td>
</tr>
<tr>
<td>4</td>
<td>Group 5</td>
<td>0.693</td>
</tr>
<tr>
<td>5</td>
<td>Group 15</td>
<td>0.702</td>
</tr>
<tr>
<td>6</td>
<td>Group 6</td>
<td>0.712</td>
</tr>
<tr>
<td>7</td>
<td>Group 2</td>
<td>0.713</td>
</tr>
<tr>
<td>8</td>
<td>Group 10</td>
<td>0.714</td>
</tr>
<tr>
<td>9</td>
<td>Group 12</td>
<td>0.716</td>
</tr>
<tr>
<td>10</td>
<td>Group 7</td>
<td>0.713</td>
</tr>
<tr>
<td>11</td>
<td>Group 11</td>
<td>0.714</td>
</tr>
<tr>
<td>12</td>
<td>Group 8</td>
<td>0.715</td>
</tr>
<tr>
<td>13</td>
<td>Group 4</td>
<td>0.712</td>
</tr>
<tr>
<td>14</td>
<td>Group 16</td>
<td>0.708</td>
</tr>
<tr>
<td>15</td>
<td>Group 3</td>
<td>0.711</td>
</tr>
<tr>
<td>16</td>
<td>Group 20</td>
<td>0.705</td>
</tr>
<tr>
<td>17</td>
<td>Group 13</td>
<td>0.702</td>
</tr>
<tr>
<td>18</td>
<td>Group 21</td>
<td>0.697</td>
</tr>
<tr>
<td>19</td>
<td>Group 19</td>
<td>0.691</td>
</tr>
<tr>
<td>20</td>
<td>Group 14</td>
<td>0.692</td>
</tr>
<tr>
<td>21</td>
<td>Group 9</td>
<td>0.687</td>
</tr>
</tbody>
</table>

Table 3: Result of the score merging. # indicates the number of samples which are used for score averaging.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.716</td>
<td>0.727</td>
<td>0.741</td>
<td>0.738</td>
<td>0.742</td>
<td>0.741</td>
<td>0.738</td>
<td>0.735</td>
</tr>
</tbody>
</table>

Table 4: Performance of the final pain detection algorithm applied to each test patient separately and all test patients’ data as a whole (overall).

<table>
<thead>
<tr>
<th>Patient</th>
<th>4</th>
<th>9</th>
<th>16</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.694</td>
<td>0.871</td>
<td>0.876</td>
<td>0.828</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>69%</td>
<td>90%</td>
<td>72%</td>
<td>87%</td>
</tr>
<tr>
<td>Specificity</td>
<td>50%</td>
<td>50%</td>
<td>77%</td>
<td>63%</td>
</tr>
</tbody>
</table>

Figure 4: ROC of the final algorithm applied to the separate test data set. To avoid biasing, the operating point (OptROCval) was determined beforehand based on the validation data set.

Figure 4: ROC of the final algorithm applied to the separate test data set. To avoid biasing, the operating point (OptROCval) was determined beforehand based on the validation data set.
but also challenging. Another ratio for the choice of signals in this study is the ongoing work of the authors’ research group to record such signals by sensor systems which work in an unobtrusive and noncontact way and which are seamlessly integrated into the dental treatment unit [18].

A pain sensation triggers an acute stress response via the sympathetic nervous system by catecholamine production (first stage of the general adaption syndrome) [19]. This will result in increased blood pressure but also tachycardia. Huiku et al. [9] and Paloheimo et al. [6] have demonstrated that this affects the frequency as well as the strength of the heart beat and can be utilized for pain assessment. Therefore, we decided to use metrics that are based on the signal amplitude (H and ΔH), beat-to-beat interval (BB and ΔBB), and pulse waveform (A, maxΔA, W2H, and maxΔW2H) as features. The frequency and wavelet components (FB and WL) of the PPG and ECG have been chosen as features with the idea in mind that those measures might enable the pain detection algorithm to discover further pain-related effects (e.g. muscle tone, blood glucose level, and hypothalamic-pituitary-adrenal axis). Furthermore, there are also pain indicators known from the literature which are based on the fact that the modulation of the heart rate variability by the vagus nerve gets disturbed by pain [20]. But as this is usually observed within a time window of 64 s and analyzed in a frequency range of 0.15–0.5 Hz, it was considered to be too slow to take place within the 6 s interval aimed at in this study.

During the course of the feature selection the addition of respiratory information proved to be of value. Four out of the nine finally selected feature groups are derived from the respiratory signal (see Figure 5). Cardiac information is included by two ECG and three PPG derived feature groups. All selected features are based on signal decomposition (three WL and six FB features).

None of the features derived in the time domain (peak amplitudes, beat-to-beat intervals, and pulse shape) have been chosen by the selection procedure, i.e. their inclusion at later stages of the FFS decreased the AUC. Although these features are certainly linked to pain, such linkage seems to be included by the selected features to the same or a higher content.

In the present study, the AUC achieved in the test phase was much better than the one achieved in the training phase. This might be a coincidence and due to the limited data set. However, we observed a similar result in our previous study and another explanation might be that the data for training of the final algorithm was slightly bigger (one patient) as during the validation phase and that this makes a difference as the total number of patients is low.

The use of cold provocation testing of tooth pulp vitality, which assures the application of the same number of pain triggers per patient, and the restriction of periodontal probing to the Ramfjord teeth turned out to be very useful for achieving a more equally distributed pain sensation among all patients. In comparison with our previous study the new study protocol yielded more than three times the number of indicated pain events (23 button touches) per patient while the coefficient of variation dropped from 16.9 to 5.2%.

Conclusion

Using supervised learning it is possible to train an algorithm (here a Random forest classifier) to differentiate between short time intervals of pain and no pain solely based on cardiorespiratory signals. The final pain detection algorithm in this work achieved a sensitivity of 87% and a specificity of 63% with an AUC of 0.828.

The number of respiratory and cardiac features selected by FFS was almost equal. Features obtained by decomposing the signals into frequency spectral bins or wavelet bands seem to cover the information that is provided by those features that describe pulse shape and beat-to-beat information by metrics derived in the time domain.

Acknowledgments: The authors would like to thank all participating patients and the study nurse Kornelia Schmitz. Research funding: The clinical study and data acquisition was funded by the German Federal Ministry of Economics and Technology (Central Innovation Program SME, ZIM). The dental treatment unit was provided by Ultradent Dental-Medizinische Geräte GmbH & CO. KG, Brunnthal, Germany. Data analysis was partially funded by the European Unions Horizon 2020 research and innovation
programme (ADAS&ME, no. 688900) and the German Research Foundation (DFG, no. TE1174/2-1).

Author contribution: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Conflict of interest: The authors state no conflict of interest.

Publication Ethics: The work described has not been published before and is not under consideration for publication elsewhere.

References