



FAKULTÄT  
Informatik, Mathematik und Naturwissenschaften

---

# VISUALISIERUNG KONTINUIERLICHER, MULTIMODALER SCHMERZ SCORES AM BEISPIEL AKUSTISCHER SIGNALE

---

Abschlussarbeit zur Erlangung des akademischen Grades  
MASTER OF SCIENCE

Studiengang MEDIENINFORMATIK

*Autor:*  
Franz ANDERS

*Erstbetreuer HTWK:*  
Prof. Dr. Mario HLAWITSCHKA

*Zweitbetreuer LaBP:*  
Mirco FUCHS

15. März 2017



# Eidesstattliche Versicherung

Ich erkläre hiermit, dass ich diese Masterarbeit selbstständig ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel verfasst habe. Alle den benutzten Quellen wörtlich oder sinngemäß entnommenen Stellen sind als solche einzeln kenntlich gemacht. Diese Arbeit ist bislang keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht worden. Ich bin mir bewusst, dass eine falsche Erklärung rechtliche Folgen haben wird.

Leipzig, 15. März 2017

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Ableitung der Schmerz Score</b>	<b>3</b>
2.1	Segmentierung . . . . .	3
2.2	Extrahierung von Eigenschaften und Ableitung der Schmerzscore . . . . .	6
2.2.1	Extrahierung von Eigenschaften . . . . .	8
2.2.2	Ableitung der Pain-Score . . . . .	11
<b>3</b>	<b>Visualisierung</b>	<b>16</b>
<b>4</b>	<b>Zusammenfassung</b>	<b>17</b>
	<b>Appendices</b>	<b>19</b>
<b>A</b>	<b>Auswertungsdetails VAD</b>	<b>20</b>

# Abbildungsverzeichnis

2.1	Beispiel für eine Segmentierung mit einem Grenzwert von $t_s = 6$ s . . . . .	5
2.2	Beispiel für die Ableitung von Pain Scores für ein Signal nach einer fiktiven Pain Scale ohne Beobachtungszeitraum oder Aktualisierungsintervall . . . .	14
2.3	Beispiel für die Ableitung von Pain Scores für ein Signal nach einer fiktiven Pain Scale mit $t_{act} = 7.5$ s und $t_{obs} = 15$ s . . . . .	15
A.1	Boxplot-Auswertung über Sensitivität, Spezifität und Genauigkeit der beiden besten Modelle . . . . .	21

# Tabellenverzeichnis

2.1	Fiktive Pain Scale . . . . .	13
3.1	Fiktive Pain Scales zur Erläuterung der Visualisierung . . . . .	16
A.1	Klassifikationsgenauigkeiten der durch den REPTree entworfenen Modelle .	20

# 1 Einleitung

Die Schmerzdiagnostik Neugeborener ist eine Aufgabe, die in vielen medizinischen Kontexten notwendig ist: Angefangen von der Überwachung prozeduraler Eingriffe wie Impfungen, über die Therapie von Entzugserscheinungen infolge des Drogenkonsums der Mutter bis hin zu Komplikationen bei der Geburt, welche die Beobachtung der Neugeborenen über mehrere Wochen bis Monate erfordert. Es gibt starke Hinweise darauf, dass insbesondere lang anhaltender Schmerz, welcher nicht korrekt erkannt und dessen Ursache somit nicht beseitigt wird, das Schmerzempfinden des Betroffenen bis in das Erwachsenenalter nachhaltig beeinflussen kann.[3, S. 402]

Babys äußern das Vorhandensein von Schmerz anhand einer Reihe von Merkmalen. Die für den Menschen am leichtesten sichtbaren Zeichen sind beispielsweise das Verkrampfen des Gesichtsausdruckes oder das Weinen. Aber auch physiologische Veränderungen wie der Erhöhung der Herzfrequenz oder des Blutdruckes lassen Rückschlüsse auf den Schmerzgrad zu.[6, S. 440] Auf Basis dieser Indikatoren werden im klinischen Alltag zur Schmerzdiagnostik sogenannte *multimodale Pain Scales* eingesetzt. Eine Pain Scale listet eine Reihe von Merkmalen, wie zum Beispiel den Hautfarbe vor das Weinen, für die eine Korrelation mit Schmerz festgestellt wurde. Für jedes Merkmal werden, je nach Grad der Ausprägung, Punkte vergeben, um schlussendlich den Schmerzgrad als *Pain Score* zu quantifizieren.[3, S. 406]

Forschungsbestrebungen der letzten Jahre haben sich zum Ziel gesetzt, die Schmerzbewertung zu automatisieren. Das Ziel dieser Arbeit ist die Entwicklung eines Konzeptes zur kontinuierlichen Schmerzbewertung bei Neugeborenen in einem multimodalen Verbund. Dabei soll der Schmerzgrad nicht nur abgeleitet, sondern auch visualisiert werden. Der Fokus liegt dabei auf der Analyse des Schmerzindikators *Weinen*, das heisst der Analyse akustischer Signale. Ausgangspunkt aller Entwürfe und Überlegungen sind die über viele Jahre erprobten Pain Scales.

Die Lösung dieser Aufgabenstellung erfordert die Kombination verschiedener Wissenschaftsdisziplin, wie zum Beispiel der medizinischen Schreiforschung oder der akustischen Modellierung der menschlichen Stimme. Kapitel ?? bietet eine Einführung in die wichtigsten Grundlagen. In Kapitel ?? wird ein Überblick über Veröffentlichungen mit ähnlichen Zielstellungen gegeben sowie dass in dieser Arbeit entwickelte Konzept grundlegend erläutert. Die folgenden Kapitel beleuchten die einzelnen Bausteine dieses Konzeptes genauer. In Kapitel ?? werden Methoden zur Feststellung des Vorhandenseins von Weingeräuschen in einem akustischen Signal vorgestellt. Dazu werden klassische Methoden der Voice Activity Detection bezüglich ihrer Leistungsfähigkeit im speziellen Anwendungsfall kindlicher Lautäußerungen eruiert und evaluiert. In Kapitel ?? werden Methoden entworfen, um aus den festgestellten Weingeräuschen den Schmerzgrad abzuleiten. Möglichkeiten der Visualisierung der Pain Score werden abschließend in Kapitel 3 vorgestellt.

Diese Arbeit leistet vor allem den Entwurf eines modular aufgebauten und erweiterbaren Konzeptes, welches die kontinuierliche Feststellung und Visualisierung von Schmerz Scores

---

ermöglicht. Das Modul, welches am ausgiebigsten evaluiert werden konnte, ist die Voice Activity Detection. Da alle darauf folgenden Module ausgiebigere Datenerhebungen in Zusammenarbeit mit medizinischen Fachkräften benötigen, wurden hier begründete Vorschläge zur Aufgabenbewältigung entworfen, welche jedoch nicht evaluiert werden konnten. Das Ziel dieser Arbeit war insbesondere nicht die Erlangung neuer Erkenntnisse über die tatsächliche Schmerzwahrnehmung Neugeborener.



## 2 Ableitung der Schmerz Score

In Kapitel ?? wurden Methoden vorgestellt, durch die einem Audiosignal die Cry-Units erkannt und markiert werden. Diese Cry-Units bilden den Ausgangspunkt, um eine Schmerz Score ableiten zu können.

Wie in Kapitel ?? erläutert wurde, wird eine Schmerz Score typischerweise nicht aus den Informationen einer einzelnen Cry-Unit, sondern aus Gruppen mehrerer Cry-Units innerhalb eines längeren Beobachtungszeitraumes geschlossen. Dafür ist es zunächst notwendig, diejenigen Cry-Units zu identifizieren, die gemeinsam einem schmerzversuchenden Stimulus zugeordnet werden können. Diese Gruppierung von Cry-Units wird in dieser Arbeit als *Segmentierung* bezeichnet und in Kapitel 2.1 vorgestellt. Die so entstandenen Segmente bilden die Basis zur Ableitung von Schmerz Scores, welche in Kapitel ?? diskutiert wird.

### 2.1 Segmentierung

Wie in Kapitel ?? erläutert wurde, wird die Schmerzdiagnose für Post-Operativen Schmerz mit Pain Scales typischerweise umgesetzt, indem das Baby in bestimmten Intervallen besucht und für einen festgelegten Zeitraum beobachtet wird, worauf den Schmerzgrad für einen bestimmten Zeitpunkt diagnostiziert wird. So empfiehlt die Pain Scale PAT, das Baby alle 30 Minuten zu besuchen und für 15 bis 30 Sekunden zu beobachten. Für viele Pain Scales konnten diese Beobachtungsintervalle und Zeiträume nicht in Erfahrung gebracht werden.

Dieses Vorgehen der Schmerzdiagnose nur zu festgelegten Zeitintervallen macht keinen Sinn für ein kontinuierliches System, da es den Vorteil der Kontinuität eliminieren würde. Die Frage ist also, wann eine medizinische Fachkraft den Anfangszeitpunkt einer Schmerzdiagnose festlegen würde, würde es das entsprechende Baby rund um die Uhr beobachten, und nicht nur eine Visite zu festgelegten Zeitpunkten durchführen. Die logisch nächstliegende Antwort ist, die Schmerzdiagnose genau dann zu starten, wenn das Baby anfängt, zu schreien.

Die nächste Frage ist, wann man die Beobachtung wieder beendet und eine Schmerz Score für den beobachteten Zeitraum festlegt. Da einige Pain Scales aus Tabelle ?? Beobachtungszeiträume angeben, kann argumentiert werden, dass die Beobachtung direkt nach dem für die jeweilige Scale angegebenen Zeitraum beendet werden und eine Pain Score abgeleitet werden kann, auch, wenn das Baby weiterhin schreit. Dabei stellt sich die Frage, welche Zeiträume man für die Pain Scales festlegt, die keinen selber keinen vorschreiben. Die zweite Frage ist, ob eine Vorzeitige Beendigung der Beobachtung überhaupt sinnvoll ist, wenn ein automatisierte System zur Verfügung steht und der Beobachter somit „unendlich viel Zeit hat“. Daher wird an dieser Stelle zuerst die grundlegendere Frage gestellt: Was ist der längste sinnvolle Zeitraum, für den eine Pain Score abgeleitet werden kann? Die nächstliegende Antwort ist: Dann, wenn das Kind aufhört, zu schreien.

Auf Basis dieser Argumentation wurde das folgende Vorgehen zur kontinuierlichen Seg-

mentierung entwickelt: Wenn das Baby keine Äußerungen von sich gibt, weil es beispielsweise schläft, wird keine Cry-Unit festgestellt, und somit existiert auch momentan kein offenes Segment. Fängt das Baby an, einen Laut von sich zu geben, also eine Cry-Unit zu produzieren, wird ein neues Segment eröffnet und die Cry-Unit diesem Segment hinzugefügt. Weitere Cry-Units werden so lange diesem Segment hinzugefügt, wie die Dauer der Stille nach einer Cry-Unit einen festgelegten Grenzwert  $t_s$  nicht überschreitet. Ein Cry-Segment wird dann geschlossen, wenn das Baby für einen festgelegten Zeitraum keine Laute mehr von sich gibt, also „aufhört, zu weinen“. Das Endzeitpunkt des Segmentes wird als der Endzeitpunkt der letzten Cry-Unit des Segmentes festgelegt.

Formel 2.1 definiert ein *Cry-Segment*  $[CS]$  als Datentyp. Ein Cry-Segment ist eine Liste von Cry-Units. Alle Cry-Units erfüllen die Nebenbedingung 2.2, das heißt, dass die Distanzen aller benachbarter Cry-Units eines Cry-Segments unterhalb des Grenzwertes  $t_s$  liegen.

$$CS = [cu_0, \dots, cu_n] \quad (2.1)$$

$$\forall cs \in CS : \forall i = 0 \dots \text{length}(cs) - 2 : d(cs[i], cs[i + 1]) < t_s \quad (2.2)$$

Der Start-Zeitpunkt eines Cry-Segments wird nach Formel 2.3 als der Startzeitpunkt der ersten Cry-Unit des Segments definiert. Das Ende eines Segmentes wird definiert als der Endzeitpunkt der letzten Cry-Unit nach Gleichung 2.4.

$$\text{start}(cs) = cs[0].\text{start} \quad (2.3)$$

$$\text{end}(cs) = cs[n].\text{end} \quad (2.4)$$

Algorithmus 1 zeigt einen Pseudocode, wie die Segmentierung nach dem beschriebenen Prinzipien offline durchgeführt wird. Input des Algorithmus ist die Liste aller Cry-Units  $CU_{all} = [cu_0, \dots, cu_m]$ , die durch das Decision-Smoothing nach Algorithmus ?? entstanden ist. Das Ergebnis des Algorithmus ist die Liste  $CS_{all}$ , die alle gefundene Cry-Segmente  $[cs_0, \dots, cs_n]$  enthält. Der Algorithmus eignet sich nicht für eine Online-Segmentierung, da das Ende eines Segmentes erst bei Beginn eines neuen Segmentes festgestellt wird, wobei beliebig viel Zeit zwischen den Beiden Segmenten liegen kann. Wurde beispielsweise  $t_s = 1$  min festgelegt, und die Pause zwischen zwei Segmenten beträgt eine Stunden, so wäre das Ende des ersten Segmentes 59 Minuten zu spät festgestellt worden. Bei einer online durchgeführten Segmentierung empfiehlt es sich, die Dauer der Stille nach jeder neu erkannten Cry-Unit kontinuierlich zu messen und ein Segment sofort zu beenden, wenn ein Stillezeitraum den Grenzwert  $t_s$  überschreitet.

Abbildung 2.1 zeigt die nach dieser Methode durchgeführte Segmentierung anhand eines Beispiels.

Das hier vorgestellte Vorgehen wurde absichtlich möglichst einfach gehalten, damit der Sinn des Parameters  $t_s$  leicht ersichtlich ist und somit von der medizinischen Fachkraft selbständig festgelegt werden kann. Schlussendlich ist eines der Hauptziele dieser Segmentierung, unnötige Berechnungen von Schmerz-Scores in den nachfolgenden Schritten zu vermeiden, so lange keine Cry-Units vorliegen. Das Ende eines Segmentes ist außerdem ein günstiger Zeitpunkt, um die Parameter des Kompressors im Pre-Processing auf Basis des

---

**Algorithm 1** Gruppierung von Cry-Units zu Cry-Segments
 

---

```

1: function SEGMENTCRYUNITS( $CU_{all}, t_s$ )
2:    $CS_{all} \leftarrow []$ 
3:    $cs \leftarrow [CU_{all}[0]]$ 
4:   for  $i = 1, \dots, \text{length}(CU_{all}) - 1$  do
5:      $cu_i \leftarrow CU_{all}[i]$ 
6:      $cu_{i-1} \leftarrow CU_{all}[i - 1]$ 
7:     if  $d(cu_{i-1}, cu_i) < t_{seg-max}$  then
8:        $cs \leftarrow [cs_i, cu_i]$ 
9:     else
10:       $CS_{all} \leftarrow [CS_{all}, cs]$ 
11:       $cs \leftarrow [cu_i]$ 
12:    end if
13:  end for return  $CS_{all}$ 
14: end function
    
```

---

RMS-Wertes des Segmentes zu aktualisieren (siehe Kapitel ??). Trotz der Trivialität dieser laufenden Segmentierung liegt hier ein wichtiger Unterschied im Gegensatz zu vergleichbaren Systemen, wie zum Beispiel das von Cohen et al. [1], bei dem die Entscheidung über Cry/not-Cry für Segmente mit einer festen Fenstergröße von 10 Sekunden vorgenommen wird.

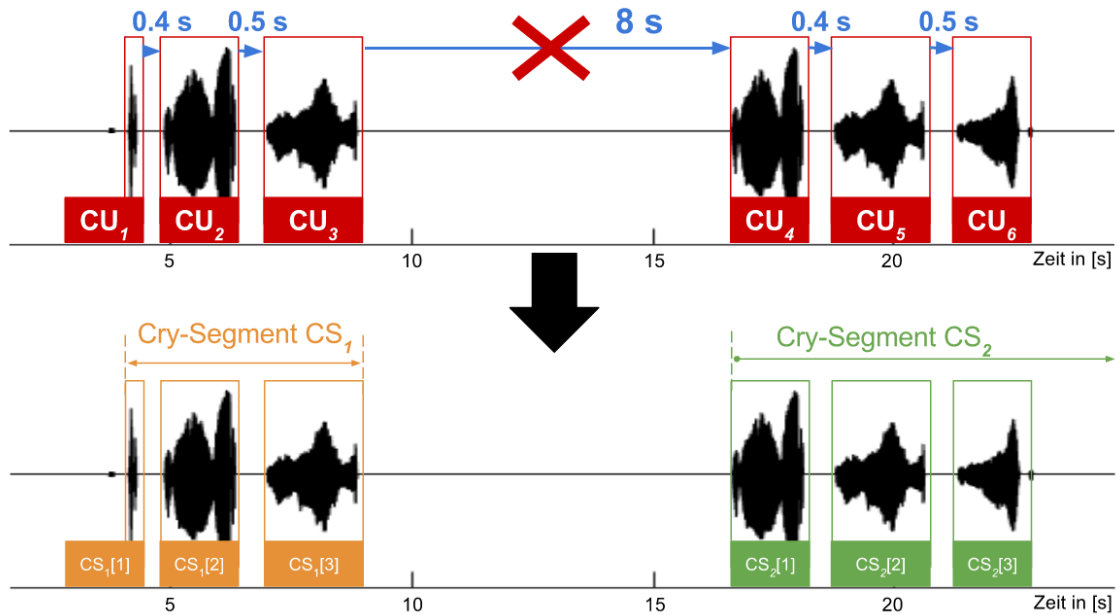


Abbildung 2.1: Beispiel für eine Segmentierung mit einem Grenzwert von  $t_s = 6$  s

## 2.2 Extrahierung von Eigenschaften und Ableitung der Schmerzscore

Das Ergebnis der Segmentierung ist eine Liste an Cry-Segmenten  $cs_0, \dots, c_n$ . Diese Cry-Segmente bilden nun die Basis für die Ableitung der Pain Score. Die medizinische Fachkraft, die das System verwendet, muss dabei zuerst die Wahl treffen, welche Pain Scale verwendet werden soll. Das einfachste denkbare Vorgehen ist die Ableitung genau einer Pain Score aus den globalen Eigenschaften eines Segmentes, wobei diese Ableitung erst vollzogen werden kann, sobald ein Segment abgeschlossen wurde und alle Informationen für dieses Segment vorliegen. Es wird also jedem Segment genau eine Pain Score zugewiesen. Das Vorgehen wird am Beispiel der NIPS aus Tabelle ?? verdeutlicht: Dabei steht die Abwesenheit von Weinen für null Punkte, „mumbling“ (murmeln) für einen Punkt und „vigorous“ (energisch) für zwei Punkte. Bei Abwesenheit von Lautäußerungen, also der Zeitraum zwischen den Segmenten, werden also keine Punkte = null Punkte vergeben. Ein Segment, dessen Qualität insgesamt als „murmelnd“ bewertet wird, erhält einen Punkt, und ein Segment, welches als insgesamt als „energisch“ bewertet wird, zwei Punkte. Das Problem ist offensichtlich: „murmelnd“ und „energisch“ sind subjektiv behaftete Begriffe und lassen sich nicht ohne weiteres aus den Eigenschaften eines Segmentes feststellen.

Es werden zwei verschiedene Lösungs-Strategien für dieses Problem vorgestellt.

### Strategie 1

---

... löst das Problem mit Hilfe von *Regression* (Siehe Kapitel ??):

1. Man erstellt eine Datenbank mit Aufnahmen von kindlichen Lautäußerungen, die man segmentiert.
2. Man errechnet „so viele *objektiv* messabare Eigenschaften wie möglich“ für jedes Segment, wie zum Beispiel die insgesamt Länge, die durchschnittliche Länge der enthaltenen Cry-Units, durchschnittliche Tonhöhe usw.
3. Man bittet medizinische Fachkräfte, für jedes Segment der Datenbank eine Score bezüglich einer Pain Scale zu vergeben. Dadurch erhält man eine gelabelte Test-Datenbank.
4. Man verwendet einen *Regressionsalgorithmus*, um den Zusammenhang zwischen den in Schritt 2 objektiv gemessenen Eigenschaften der Segmente und den in Schritt 3 vergebenen *Scores* herzustellen. An dieser Stelle kann zum Beispiel die in Kapitel ?? beschriebene multiple lineare Regression verwendet werden. Man erhält somit einen Regressor für jede Pain Scale.
5. Möchte man für neue, unbekannte Segmente die Pain Score prognostizieren, nutzt man den entsprechenden Regressor.

---

Das Vorteil dieses Vorgehens ist, dass das Problem der Übersetzung der objektiv messbaren Parameter in die subjektiv behafteten Begriffe überbrückt wird, indem die Regression direkt von den objektiv messbaren Parametern auf die Pain Score durchgeführt wird. Der Nachteil ist, dass eine Testdatenbank für jede Pain Scale aufgebaut werden muss. Wird ein neue Pain Scale eingeführt, muss der Regressor für diese Scale durch erneutes Labeln festgestellt werden. Ein weiterer Effekt der Abbildung des Problems als Regression ist, dass ein Regressor

in einen kontinuierlichen Zahlenraum abbildet. Es sind also Regressionsergebnisse wie zum Beispiel 2.8 denkbar. Diese „bessere Auflösung“ kann als Vorteil betrachtet werden. Ist jedoch eine direkte Übersetzung der Pain Scale inklusive der ganzzahligen Punktzahlen gewünscht, so stellt sich die Frage, ob eine 2.8 auf- oder abzurunden ist.

### Strategie 2

---

... löst das Problem mit Hilfe von Klassifizierung (Siehe Kapitel ??):

1. und 2. entsprechen Strategie 1
3. Man sammelt alle subjektiven Begriffe, die in Pain Scales verwendet werden, wie zum Beispiel „murmelnd“, „energisch“, usw.
4. Man bittet medizinische Fachkräfte, jedes Segment der Datenbank mit denjenigen Begriffen zu labeln, die die jeweilige Person für zutreffend hält.
5. Man Verwendet einen *Klassifizierungsalgorithmus*, um einen Zusammenhang zwischen den in Schritt 2 festgestellten objektiv messbaren Eigenschaften der Segmente und den *subjektiv behafteten Begriffen* zu finden. Man erhält somit einen Klassifikator für jeden Begriff, der binär in *positive = zutreffend* und *negative = nicht zutreffend* klassifiziert.
6. Möchte man für neue, unbekannte Segmente die Pain Score prognostizieren, so wird für jede mögliche Score der Pain Scale überprüft, ob für alle subjektiv beschreibenden Begriffe der entsprechende Klassifikator ein positive prognostiziert. Die Ableitung der Score ist somit ein weiteres Klassifizierungsproblem, wobei eine Score einer Klasse entspricht und genau dann abgeleitet werden kann, wenn alle Voraussetzungen für die Klasse erfüllt sind.

---

Der Vorteil dieser Methode ist, dass auch zum Zeitpunkt der Erstellung der Testdatenbank unbekannte Pain Scales zu einem späteren Zeitpunkt eingebunden werden können, insofern alle in dieser neuen Pain Scale verwendeten subjektiv behafteten Begriffe bereits gelabelt vorliegen, weil sie auch in anderen Pain Scales verwendet wurden. Das Vorgehen erlaubt somit eine gewissen Flexibilität bezüglich zukünftig entwickelter Pain Scales. Der Nachteil dieser Methode ist, dass durch die Umwandlung der eigentlich quantitativ geordneten Score einer Pain Scale in qualitative Klassen aus einem implizit als Regression zu betrachtenden Problem ein Klassifizierungsproblem macht. Dies wirft neue Fragen auf, wie zum Beispiel: Angenommen, bei einer fiktiven Pain Scale wird jede Score mit jeweils drei subjektiv behafteten Begriffen beschrieben, und bei der Klassifizierung eines Segmentes wird festgestellt, dass für jede Punktzahl genau zwei der drei Begriffe erfüllt werden. Welche Score wird dann prognostiziert? Ein anderes Beispiel wird am Beispiel der der NIPS-Score aus Tabelle ?? verdeutlicht: Angenommen, ein Cry-Segment enthält hörbar „starkes“ Schreien, es kann jedoch weder „mumbling (murmelnd)“ noch „vigorous (energisch)“ abgeleitet werden. Demzufolge müsste dieses Segment eine Score von 0 Punkten erhalten, wobei ein Mensch in dieser Situation eventuell „stark“ zu „heftig“ uminterpretieren und 2 Punkte vergeben hätte. Strategie 1 ist weniger anfällig für dieses Problem.

In jedem Fall werden medizinische Fachkräfte benötigt, um das Labeling der Cry-Segmente durchzuführen, was aus Zeitgründen im Rahmen dieser Arbeit nicht möglich ist. Die Aquisierung von Audioaufnahmen von Babys sowie das Labeling der Aufnahmen erfordern nicht nur Zeit, sondern das Fachwissen über das Führen und die Auswertung von Interviews.

### 2.2.1 Extrahierung von Eigenschaften

Im vergangenen Kapitel wurde erläutert, dass die Basis für die Ableitung einer Pain Score für ein Segment die Extraktion von „so vielen Features wie möglich“ ist. In diesem Kapitel wird präzisiert, welche Features gemeint sind. Varallyay [8, S. 16 - 17] schlug vor, drei Kategorien an Features zu betrachten: 1.) Features des Zeitbereichs, 2.) Features der Frequenzbereichs, und 3.) Melodie-bezogene Attribute. Diese Kategorisierung wurde für diese Arbeit übernommen.

In Kapitel ?? wurde beschrieben, welche Features in der medizinischen Schreiforschung typischerweise extrahiert wurden. In Kapitel ?? wurde diskutiert, dass 1.) nicht bewiesen ist, welche Features die „wichtigsten“ sind und 2.) keine Einigung darüber herrscht, wie genau bestimmte Features zu berechnen sind. Basierend auf den in diesem Kapitel vorgestellten Features werden in diesem Kapitel konkrete Berechnungsvorschriften definiert. Welche von diesen Features tatsächlich im Zusammenhang mit Schmerz stehen, lässt sich erst in der anschließenden Nutzung der Features zur Regression oder Klassifizierung der Pain Scales feststellen, welche jedoch im Rahmen dieser Arbeit nicht durchgeführt werden kann.

#### Features des Zeitbereiches

Mit Features des Zeitbereiches sind solche gemeint, die sich allein aus Kenntnis der der Start- und Endzeitpunkte der im Segment enthaltenen Cry Units sowie deren Zeitbereiche gewinnen lassen, wie beispielsweise die durchschnittliche Länge der Cry-Units oder die durchschnittliche Energie der Cry-Units. In diesem Kapitel gilt die Konvention, dass eine Cry-Segment  $cs$  insgesamt  $N$  Cry-Units enthält, die Indexierung wird mit  $0 \dots N - 1$  definiert.

**Segment-Length:** Zeitliche Länge des Segmentes:

$$\text{Segment-Length}(cs) = cs[N - 1].end - cs[0].start \quad (2.5)$$

**Density:** Relativer Anteil der Cry-Units an der Länge des Segmentes („Dichte“)

$$\text{Density}(cs) = \frac{\sum_{i=0}^{N-1} \lambda(cs[i])}{\text{Segment-Length}(cs)} \quad (2.6)$$

**Tempo:** Das Verhältnis zwischen der Dauer des Segmentes und der Anzahl der Cry-Units. Dieses Feature ähnelt dem von LaGasse et al. [4, S. 85] als *Utterances* bezeichneten Feature.

$$\text{Tempo}(cs) = \frac{N}{\text{Segment-Length}(cs)} \quad (2.7)$$

**Statistics of Cry-Units:** Statistische Auswertungen bezüglich der *Länge der Cry-Units*  $\text{stats}_{cu}(cs)$ : Durchschnitt, Median, Minimum, Maximum und Standardabweichung

der Cry-Units. Das  $\text{mean}_{cu}(cs)$ -Feature wird von LaGasse et al. [4, S. 85] und vielen weiteren Schreiforschern als *Mean Duration* bezeichnet.

$$\text{stats}_{cu}(cs) = \begin{cases} \text{mean}_{cu}(cs) = \text{mean}_{i=0 \dots N-1} \{\lambda(cs[i])\} \\ \text{median}_{cu}(cs) = \text{median}_{i=0 \dots N-1} \{\lambda(cs[i])\} \\ \text{min}_{cu}(cs) = \min_{i=0 \dots N-1} \{\lambda(cs[i])\} \\ \text{max}_{cu}(cs) = \max_{i=0 \dots N-1} \{\lambda(cs[i])\} \\ \sigma_{cu}(cs) = \sigma_{i=0 \dots N-1} \{\lambda(cs[i])\} \end{cases} \quad (2.8)$$

**Statistics of Bursts:** <sup>1</sup> Die in Gleichung 2.8 definierten Features können ebenso in Bezug auf die *Längen der Bursts* errechnet werden, in dem in jeder Gleichung  $\lambda(cs[i])$  ersetzt wird durch  $cs[i].start - cs[i-1].start$ . Die Indexierung muss auf  $i = 1, \dots, N-1$  begrenzt werden.

$$\text{stats}_{burst}(cs) = \begin{cases} \text{mean}_{burst}(cs) = \text{mean}_{i=1 \dots N-1} \{cs[i].start - cs[i-1].start\} \\ \text{median}_{burst}(cs) = \text{median}_{i=1 \dots N-1} \{cs[i].start - cs[i-1].start\} \\ \dots \end{cases} \quad (2.9)$$

**Statistics of Pauses:** Nach dem selben Muster werden die statistischen Auswertungen bezüglich der *Längen der Pausen* ermittelt. Eine Pause entspricht in diesem Zusammenhang der Distanz zwischen zwei aufeinander folgenden Cry-Units nach Gleichung ??.

$$\text{stats}_{pause}(cs) = \begin{cases} \text{mean}_{pause}(cs) = \text{mean}_{i=1 \dots N-1} \{d(cs[i-1], cs[i])\} \\ \text{median}_{pause}(cs) = \dots \end{cases} \quad (2.10)$$

**Statistics of Energies:** Zunächst wird die Liste aller in den Cry-Units enthaltenen Signalfenster definiert nach Gleichung 2.11. Eine Cry-Unit hat die Signalfenster  $cu.windows = x_0[], \dots, x_m[]$

$$x_{seg}[] = cs[0].windows[0], \dots, cs[N-1].windows[m] \quad (2.11)$$

Die Liste  $x_{seg}[]$  hat  $R$  Elemente, die Indexierung wird definiert mit  $0, \dots, R-1$ . Gleichung 2.12 definiert die Features bezüglich der MSV-Werte („Lautstärken“) des Segmentes. Der MSV-Wert als Maß des durchschnittlichen Energiegehaltes wurde in Gleichung ?? definiert.

$$\text{stats}_{msv}(cs) = \begin{cases} \text{mean}_{msv}(cs) = \text{mean}_{i=0 \dots R-1} \{MSV(x_{seg}[i])\} \\ \text{median}_{msv}(cs) = \dots \end{cases} \quad (2.12)$$

---

<sup>1</sup>Erläuterung zum Begriff *Burst* in Kapitel ??

Diese statistischen Auswertungen bezüglich der Länge der Cry-Units und Bursts wurden beispielsweise von Zeskind et al. [7] vorgenommen, wenn auch nicht Computer-gestützt. Es ist zu bemerken, dass in der klassischen Schreiforschung zeitliche Features im geringeren Maße in Betracht gezogen wurden als Features des Frequenz-Bereiches. Die einzigen zeitliche Features, die zum Beispiel von Wasz-Hockert et al. [5], Fuller [2] und LaGasse et al. [4] berechneten, sind *die durchschnittliche Länge der Cry-Units* (hier  $\text{mean}_{cu}(cs)$ ) und die *Latenz zwischen Reiz und erster Cry-Unit*, welche nur auf Basis des Audiosignals nicht feststellbar ist. Es spricht jedoch nichts dagegen, die hier vorgestellten Features trotzdem zu erproben. Die anschließende Nutzung der Features zur Regression/Klassifizierung wird Auskunft darüber geben, welchen Beitrag diese Features zur Schmerzdiagnose leisten können.

### Features des Frequenzbereiches und der Melodie

Mit Features des Frequenz-Bereiches sind diejenigen Features gemeint, die sich aus der Short Time Fourier Transformation der Cry-Units gewinnen lassen. Um die Features durch mathematische Formeln definieren zu können, wird zuerst das *Spectrum des Segmentes*  $X_{seg}[\ ]$  nach Formel 2.13 als die Liste aller Frequenz-Bereiche der Signalfenster der Cry-Units des Segmentes definiert. Die Indexierung von  $X_{seg}[\ ]$  läuft, wie bei  $x_{seg}[\ ]$  von  $0, \dots, R-1$ . Nach dem selben Muster wird das *Cepstrum des Segmentes*  $c_{seg}[\ ]$  definiert.

$$X_{seg}[\ ] := \bigvee_{x_i[\ ] \in x_{seg}} : |DFT\{x_i[\ ] \cdot w[\ ]\}| \quad (2.13)$$

Die folgenden Features des Frequenzbereiches lassen sich mit den in dieser Arbeit vorgestellten Methoden berechnen:

**Tensness:** Das Feature, welches in Kapitel ?? als „Ratio2“ beschrieben wurde. Es wurde von Fuller [2] eingeführt und beschreibt die Spannung des Vokaltraktes als Verhältnis der Energien oberhalb von 2000 Hz zu unter 2000 Hz. Wie bei den statistischen Auswertungen der Features des Zeitbereiches kann für das gesamte Segment der Durchschnitt, Median, Maximum, Minimum und Standardabweichung berechnet werden.

$$\text{stats}(Tensness) = \begin{cases} \text{mean}_{Tens}(cs) = \text{mean}_{i=0 \dots R-1} \left\{ \frac{\sum_{k=0}^{2000 \text{ Hz}} X_{sec}[i][k]}{\sum_{j=2000 \text{ Hz}}^{f_s} X_{sec}[i][j]} \right\} \\ \text{median}_{Tens}(cs) = \dots \end{cases} \quad (2.14)$$

**Clarity:** Wie in Kapitel ?? erläutert wurde, lässt eine stark ausgebildete Spitze im oberen Cepstrum-Bereich auf ein stimmhaftes Signal schließen. Ein hoher Anteil stärkerer Cepstrum-Peaks lässt also auf vermehrt phonierte Laute schließen, geringere Cepstrum-Peaks auf dysphoniertere Laute (Siehe Kapitel ??). Dieses durchschnittliche Wert dieses Features trifft eine Aussagen über den Anteil dysphonierter Laute, die Standardabweichung ähnelt dem in Kapitel ?? vorgestellten *Cry-Mode Changes*-Feature.

$$\text{stats}_{clarity}(cs) = \begin{cases} \text{mean}_{Clarity}(cs) = \text{mean}_{i=0 \dots R-1} \left\{ Ceps_{mag}(c_{seg}[i]) \right\} \\ \text{median}_{Clarity}(cs) = \dots \end{cases} \quad (2.15)$$



Alle weiteren Features, die in Kapitel ?? vorgestellt wurden und sich auf den Frequenzbereich beziehen, lassen sich nicht mehr mit den in dieser Arbeit vorgestellten Methoden extrahieren. Entweder beziehen sie sich auf die Lage der Formanten, oder basieren auf der Feststellung der Grundtonhöhe. In dieser Arbeit konnten aus Platzgründen jedoch keine Methoden zur Extraktion dieser Informationen vorgestellt werden. Gleiches gilt für die Feststellung des Melodieverlaufs, welche ebenfalls auf der Feststellung der Grundtonhöhe basiert. Das Muster, nach dem diese Features berechnet werden können, sollte aus den bisher vorgestellten Features ersichtlich sein. So lassen sich beispielsweise die Features bezüglich der Grundtonhöhe nach Formel 2.16 ableiten. Dabei sei  $f_0(X_i[ ])$  eine idealisierte Funktion, welche die Grundtonhöhe  $f_0$  für das Frequenzfenster  $X_i[ ]$  berechnet. Da für die Definition der weiteren Features idealisierte ebenfalls Funktionen angenommen werden müssten, wird die Festlegung weiterer Features an dieser Stelle nicht fortgeführt.

$$\text{stats}_{\text{pitch}}(cs) = \begin{cases} \text{mean}_{\text{Pitch}}(cs) = \text{mean}_{i=0 \dots R-1} \left\{ f_0(X_{\text{seg}}[i]) \right\} \\ \text{median}_{\text{Pitch}}(cs) = \dots \end{cases} \quad (2.16)$$

## Diskussion

Bei allen vorgestellten Features handelt es sich, nach dem Vorbild der in Kapitel ?? vorgestellten Features der klassischen Schreiforschung, um solche, bei denen die Reihenfolge der Cry-Units nicht mit in Betracht gezogen wird. Angenommen, ein Segment besteht aus  $n$  Cry-Units, wobei genau eine Hälfte der Cry-Units kurz und die andere Hälfte der Cry-Units lang ist. Das  $\text{stats}_{cu}(cs)$ -Feature wird bezüglich des Durchschnittes, Minimum, Maximum etc. die selben Werte berechnen, unabhängig davon, ob sich die kurzen Cry-Units allesamt am Beginn des Segmentes, am Ende des Segmentes oder mit den langen Cry-Units durchmischt befinden. Bei der anschließenden Nutzung der Features zu Regression/Klassifizierung wird sich zeigen, wie sehr sich diese Features zur Ableitung von Pain-Scores eignen. Stellt sich heraus, dass sich die Features nicht eignen, ist es eventuell notwendig, neue Features zu definieren, die die Position der Cry-Units mit in Betracht ziehen.

### 2.2.2 Ableitung der Pain-Score

Mathematisch wird die Ableitung einer Pain Score als Funktion  $PS_{\text{Scale}} : cs \mapsto \mathbb{N}$  definiert, welche eine Cry-Segment  $cs$  unter Verwendung der Pain-Scale auf einen Pain-Score  $\mathbb{N}$  abbildet. Der konkrete Inhalt der Funktion ist für die entsprechende Pain Scale mit dem in diesem Kapitel beschriebenen Vorgehen zu ermitteln. Die Ableitung genau einer Pain Score für ein Cry-Segment stellt den einfachsten Fall dar. Dies ist für bestimmte Anwendungsfälle eventuell nicht ausreichend:

1. Die Score kann erst nach der Beendigung eines Segmentes abgeleitet werden, was für einigen Kontexte möglicherweise zu spät ist. Besonders die Schmerzdiagnostik während Schmerzverursachenden Prozeduren kann das häufigere „Aktualisieren“ der Schmerzscores notwendig machen.
2. Falls der Schmerz innerhalb eines Segmentes stark ab- oder zunimmt, ist dieser Verlauf nicht erkennbar. Es würde lediglich der „durchschnittliche Schmerz“ des Segmentes abgeleitet werden.

Das vorgestellte Prinzip wird daher erweitert, indem ein Aktualisierungsintervall  $t_{act}$  und Beobachtungszeiträume  $t_{obs}$  eingeführt wird.

### Aktualisierungsintervall

Die Grundlegende Idee des Aktualisierungsintervalls ist, bei einem momentan offenen Segment in regelmäßigen Abständen die Features abzufragen und direkt die Pain Score abzuleiten, um Zwischenergebnisse zu erhalten. Der am häufigsten umsetzbare Fall ist, ein Aktualisierung nach jeder neu dem Segment hinzugefügten Cry-Unit vorzunehmen. Der am wenigsten häufige Fall ist der bereits genannte, die Aktualisierung erst bei Beendigung eines Segmentes durchzuführen. Eine offensichtliche Variante zur Festlegung von  $t_{act}$  ist die Bestimmung eines zeitlichen Wertes. Ein  $t_{act}$  von beispielsweise 10s würden bedeuten, dass alle 10 Sekunden ein neuer Pain Score für ein Segment berechnet wird. Die Beendigung eines Segmentes würde in jedem Fall eine Ableitung der Pain Score auslösen und einen „erzwungenen Aktualisierungszeitpunkt“ darstellen. Die folgenden Möglichkeiten zur Festlegung von  $t_{act}$  sind denkbar:

- $t_{act}$  als globaler, nicht veränderbarer Wert. Da in der Literatur keine Vorschlag diesbezüglich gefunden werden konnte, müsste ein sinnvoller Wert in Absprache mit medizinischen Fachkräften eruiert werden.
- Man überlässt der medizinischen Fachkraft, die das System überwacht, die Festlegung des Aktualisierungsintervalls. So kann die Person selber bestimmen, wie häufig sie eine Aktualisierung der Pain Score wünscht.
- Die feste Bindung des Aktualisierungsintervalls an die verwendete Pain Scale. Die CRIES-Scale ist beispielsweise für das post-operative Monitoring gedacht und benötigt somit möglicherweise weniger häufige Aktualisierungen als der DAN, welcher zur Schmerzdiagnostik während einer Operation eingesetzt werden kann (siehe Tabelle ??). Da die Pain Scales nicht für die kontinuierliche Schmerzdiagnostik ausgelegt sind, lässt sich aus den für einige Pain Scales eventuell vorgeschriebenen Beobachtungsintervallen kein Aktualisierungsintervall für ein kontinuierliches System argumentieren.

Wenn  $t_{act}$  als zeitlicher Wert definiert wird, kann es passieren, dass eine Aktualisierung in einem offenen Segment durchgeführt wird, während gerade eine neue Cry-Unit markiert wird und noch nicht abgeschlossen wurde. Da die Funktion  $PS_{Scale} : cs \mapsto \mathbb{N}$  nur für Cry-Segmente definiert wurde, die vollständige Cry-Units enthalten, wird diese „halbe Cry-Unit“ nicht mit zur Ableitung der Pain Score verwendet. Der Hintergrund für diese Entscheidung ist, dass bestimmte Funktionen zum Ableiten der Features für das Segment ansonsten fehlerhafte Ergebnisse liefern können. Die Funktion  $PS$  wird folglich auf das Subsegment  $cs_{sub}$  angewandt, welches alle Cry-Units des möglicherweise noch offenen Segmentes  $cs$  beinhaltet, die zum Aktualisierungszeitpunkt  $t$  vollständig beendet wurden. Der Endzeitpunkt des Subsegmentes wird, wie bei herkömmlichen Segmenten, auf den Endzeitpunkt der letzten vollständigen Cry-Unit im Subsegment gelegt.

### Beobachtungszeitraum

Es gibt Eigenschaften, die sich implizit auf den gesamten Zeitraum *Beginn des Segmentes*  $start(cs)$  bis *Aktualisierungszeitpunkt*  $t$  beziehen, wie beispielsweise die *Zeitliche Länge des*

*Segmentes* nach Formel 2.5. Dieser Zeitraum ist gleichzeitig der längst mögliche Zeitraum innerhalb eines Segmentes, der für die Ableitung der Pain Score mit einbezogen werden kann. Es ist jedoch auch möglich, einen kürzere Beobachtungszeitraum  $t_{obs}$  zu wählen. Dies hat zur Folge, dass bei der Ableitung die ersten Cry-Units des Segmentes ausgelassen werden, die außerhalb des Beobachtungszeitraums liegen. So können zeitliche Veränderungen der Pain-Score innerhalb eines Segmentes detaillierter dargestellt werden. Es sind wiederum verschiedene Varianten zur Festlegung von  $t_{obs}$  denkbar:

- Festlegung eines globalen oder eines von der medizinischen Fachkraft frei wählbaren Wertes, so wie bei dem Aktualisierungsintervall  $t_{act}$ .
- Die feste Bindung des Aktualisierungsintervalls an die verwendete Pain Scale. Einige Pain Scale empfehlen bestimmte Beobachtungszeiträume. So wird beispielsweise bei der NIPS-Scale ein Beobachtungszeitraum von einer Minute empfohlen (Siehe Kapitel ??). Es müsste wiederum in Zusammenarbeit mit medizinischen Fachkräften eruiert werden, ob diese, für die manuelle Schmerzdiagnostik vorgesehenen Werte auch für ein automatisiertes System Sinn machen.
- Eine weitere Variante ist,  $t_{obs}$  an den Wert von  $t_{act}$  zu binden. Fall  $t_{obs}$  frei festlegbar sein soll, muss das Personal nicht zwei Werte festlegen. Ein Verhältnis von  $t_{obs} = k \cdot t_{act}$  würde mit  $k = 1$  nicht-überlappende Beobachtungszeiträume und mit  $k = 2$  überlappende Beobachtungszeiträume erzeugen.

Der Beobachtungszeitraum  $t_{obs}$  schränkt somit die Länge des Subsegmentes  $cs_{sub}$  weiter ein, und zwar in diesem Fall bezüglich der Startzeitpunktes. Der Zeitraum innerhalb des Segmentes  $cs$ , der zur Bildung von  $cs_{sub}$  genutzt wird, ist der Zeitraum *Aktualisierungszeitpunkt*  $t - t_{obs}$  bis  $t$ . Es werden nur solche Cry-Units von  $cs$  in  $cs_{sub}$  übernommen, die innerhalb dieses Zeitraumes vollständig begonnen und beendet werden konnten. Der Anfangszeitpunkt des Subsegments ist somit der Anfangszeitpunkt der ersten vollständigen Cry-Unit innerhalb Beobachtungszeitraumes, der Endzeitpunkt des Subsegmentes entspricht dem Ende der zum Aktualisierungszeitpunkt zuletzt vollständig beendeten Cry-Unit.

### Beispiel

---

Die in diesem Kapitel vorgestellten Methoden zur Ableitung von Pain Scores werden anhand eines Beispiels verdeutlicht. Tabelle 2.1 definiert eine fiktive Pain Scale. Es werden sowohl die subjektiv beschreibenden Kriterien angegeben, als auch die objektiv messbaren Kriterien, die mit einer der in Kapitel 2 beschriebenen Strategien gefunden wurden. In diesem Fall hängt die Pain Score von der Durchschnittslänge der Cry-Units des Segmentes  $mean_{cu}(cs)$  ab (siehe Gleichung 2.8).

Tabelle 2.1: Fiktive Pain Scale

Name		0 Punkte	1 Punkte	2 Punkte
„Fiction Scale“	Subj. Krit.:	Kein Weinen	normales Weinen	energisches Weinen
	Obj. Krit.:	Außerhalb Segment	$mean_{cu}(cs) < 1\text{ s}$	$mean_{cu}(cs) \geq 1\text{ s}$

Gleichung 2.17 definiert die Funktion  $PS_{Fiction}$  dieser Pain Scale.

$$PS_{Fiction}(cs) = \begin{cases} 1 & , \text{wenn } mean_{cu}(cs) < 1 \text{ s} \\ 2 & , \text{sonst} \end{cases} \quad (2.17)$$

Abbildung 2.2 zeigt ein Beispielsignal, für das Pain Scores nach dieser Pain Scale abgeleitet werden. In dem Signal werden die stimmhaften Signalbereiche schwarz und das Hintergrundrauschen grau dargestellt. Es sind insgesamt 10 Cry-Units zu erkennen. Die ersten fünf Cry-Units haben jeweils eine Länge von 2.25 s, die letzten fünf Cry-Units eine jeweilige Länge von 0.75 s. Das Signal wurde nach der in Kapitel 2.1 beschriebenen Methode segmentiert mit  $t_s = 5$  s und so alle 10 Cry-Units zu einem Segment zusammengefasst. Das Segment erstreckt sich von Sekunde 10 bis Sekunde 37.5. Für die Bereiche, die außerhalb des Segmentes liegen, wird implizit eine Pain Score von 0 festgelegt. Für das Segment wurde eine durchschnittliche Länge der Cry-Units von  $mean_{cu}(cs) = 1.5$  s gemessen und dem zufolge eine Pain Score von 2 abgeleitet. In diesem Fall wurde ohne Beobachtungs- und Aktualisierungsintervall gearbeitet. Wäre die Analyse also kontinuierlich vorgenommen worden, so wäre nach Feststellung der ersten Cry-Unit das Segment eröffnet, nach Überschreitung der maximal zulässigen Stille von  $t_s = 5$  s nach der 10. Cry-Unit das Segment geschlossen, und daraufhin  $PS_{Fiction}(cs) = 2$  berechnet worden.

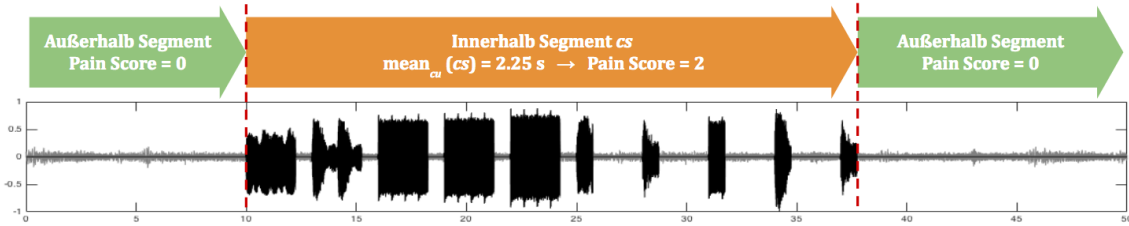


Abbildung 2.2: Beispiel für die Ableitung von Pain Scores für ein Signal nach einer fiktiven Pain Scale ohne Beobachtungszeitraum oder Aktualisierungsintervall

Abbildung 2.3 zeigt die Ableitung der Schmerz Scores, wenn zusätzlich ein Aktualisierungsintervall von 7.5 s und ein Beobachtungszeitraum von 15 s gewählt wird. Nach dem das Segment durch die Cry-Unit an Sekunde 10 eröffnet wurde, werden Aktualisierungen zu den Zeitpunkten  $t = 17.5$  s, 25 s, 32.5 s und 40 s durchgeführt, verdeutlicht durch die kleinen, roten Pfeile in der Abbildung. Wie zu sehen ist, wird bei jeder Aktualisierung innerhalb des Beobachtungszeitraumes ein Subsegment gebildet, für das Subsegment die Features errechnet und die Pain Score abgeleitet. Der Anfangszeitpunkt jedes Subsegmentes ist der Anfang der erste Cry-Unit innerhalb des jeweiligen Beobachtungszeitraumes, und das Ende des Subsegmentes das Ende letzten Cry-Unit im jeweiligen Beobachtungszeitraum. Beispielsweise erstreckt sich das bei der 3. Aktualisierung der Beobachtungszeitraum von 17.5 – 32.5 s, das Subsegment jedoch von 19 – 32 s aufgrund der Lage der Cry-Units. Durch die Verwendung des Beobachtungs- und Aktualisierungsintervalls wird erkennbar, dass in diesem Beispiel der Schmerzgrad innerhalb des Segmentes nach hinten hin abnimmt.

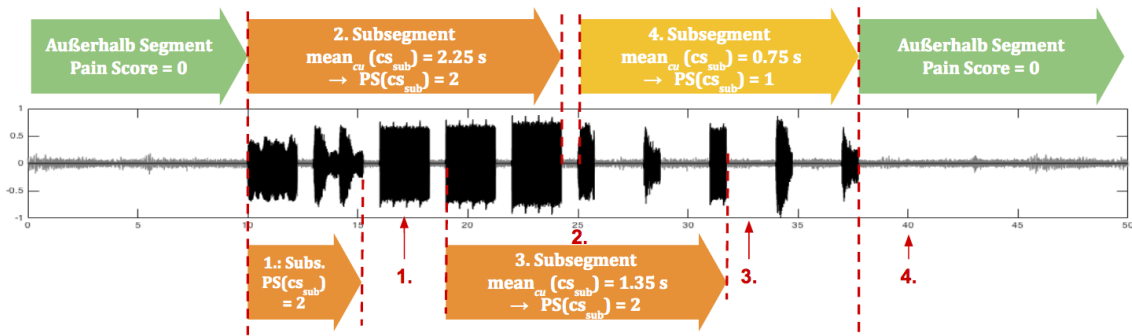


Abbildung 2.3: Beispiel für die Ableitung von Pain Scores für ein Signal nach einer fiktiven Pain Scale mit  $t_{act} = 7.5\text{ s}$  und  $t_{obs} = 15\text{ s}$

### 3 Visualisierung

Zur Erläuterung der Visualisierung werden die in Tabelle 3.1 aufgeführten fiktiven Pain Scales definiert.

Tabelle 3.1: Fiktive Pain Scales zur Erläuterung der Visualisierung

Score	„Length-Scale“		„Max-Scale“	
	Subjektive Krit.	Objektive Krit.	Subjektive Krit.	Objektive Krit.
0 Punkte	kein Weinen	Außerhalb Segment	kein Weinen	
1 Punkt	kurzes Weinen	$S\text{-Length}(cs) \leq 1 \text{ min}$	seichtes Weinen	$0 \text{ s} < \min_{cu}(cs) \leq 1 \text{ s}$
2 Punkte	langes Weinen	$S\text{-Length}(cs) > 1 \text{ min}$	stärkeres Weinen	$1 \text{ s} < \min_{cu}(cs) \leq 2 \text{ s}$
3 Punkte	-	-	energisches Weinen	$\min_{cu}(cs) > 2 \text{ s}$

Gleichung

## 4 Zusammenfassung

# Literaturverzeichnis

- [1] Rami Cohen and Yizhar Lavner. Infant Cry Analysis and Detection. In *27th Convention of Electrical and Electronics Engineers in Israel*. IEEE, 2012. Online unter [https://www.researchgate.net/publication/261116332\\_Infant\\_cry\\_analysis\\_and\\_detection](https://www.researchgate.net/publication/261116332_Infant_cry_analysis_and_detection).
- [2] Barbara Fuller. Acoustic Discrimination of three Cry Types. *Nursing Research*, 40(3), 1991. Online erhältlich unter: [https://www.researchgate.net/publication/21125005\\_Acoustic\\_Discrimination\\_of\\_Three\\_Types\\_of\\_Infant\\_Cries](https://www.researchgate.net/publication/21125005_Acoustic_Discrimination_of_Three_Types_of_Infant_Cries).
- [3] Bonnie Stevens Huda Huijer Abu-Saad, Gerrie Bours and Jan Hamers. Assessment of pain in Neonates. *Seminars in Perinatology*, 2(5):402–416, 1998. Online unter <https://www.ncbi.nlm.nih.gov/pubmed/9820565>.
- [4] A. Rebecca Neal Linda L. LaGasse and Barry M. Lester. Assessment of infant cry: Acoustic cry analysis and parental perception. *Mental retardation and developmental disabilities*, 11(1):83–93, 2005. Online unter <https://www.ncbi.nlm.nih.gov/pubmed/15856439>.
- [5] Katarina Michelsson Ole Wasz-Hockert and John Lind. Twenty-Five Years of Scandinavian Cry Research. In *Infant Crying - Theoretical and Research Perspectives*, chapter 3, pages 59 – 82. Plenung, 1985.
- [6] J L Mathew P J Mathew. Assessment and management of pain in infants. *Postgrad Med J*, 79:438–443, 2003. Online unter <http://pmj.bmj.com/content/79/934/438.full>.
- [7] Susan Parker-Price & Ronald Barr Philip Zeskind. Rythmic organization of the Sound of Infant Cry. *Dev Psychobiol*, 26(6):321–333, 1993. Online unter <https://www.ncbi.nlm.nih.gov/pubmed/8119482>.
- [8] Gyorgy Ivan Varallyay. *Analysis of the Infant Cry with Objective Methods*. PhD thesis, Budapest University of Technology and Economics, 2009. Online erhältlich unter: <https://pdfs.semanticscholar.org/5c38/b368dc71d67cbfab3077a50536b086d8eec.pdf>.



# Appendices

## A Auswertungsdetails VAD

Tabelle A.1: Klassifikationsgenauigkeiten der durch den REPTree entworfenen Modelle

$SNR_{Training}$	3 dB				50 dB				50+3 dB			
$SNR_{Test}$	3 dB	50 dB	7 dB*	Mean	3 dB	50 dB	7 dB*	Mean	3 dB	50 dB	7 dB*	Mean
Zeit	77.81%	79.02%	86.04%	80,96%	49.33%	94.70%	48.66%	64,23%	77.54%	92.47%	84.38%	84,80%
Spektrum	82.05%	89.28%	82.71%	84,68%	70.52%	94.37%	55.06%	73,31%	81.75%	91.22%	74.90%	82,62%
Cepstrum	88.98%	94.72%	92.96%	<b>92,22%</b>	86.83%	94.68%	92.83%	<b>91,45%</b>	88.98%	94.72%	92.96%	<b>92,22%</b>
Autokorr.	80.45%	73.47%	84.89%	79,60%	73.07%	87.14%	77.98%	79,39%	77.90%	84.88%	82.84%	81,87%
Zeit+Spektrum	82.05%	89.28%	82.71%	84,68%	70.52%	94.37%	55.06%	73,31%	81.75%	91.22%	74.90%	82,62%
Zeit+Cepstrum	88.98%	94.72%	92.96%	<b>92,22%</b>	86.83%	94.68%	92.83%	<b>91,45%</b>	88.98%	94.72%	92.96%	<b>92,22%</b>
Zeit+Autokorr	80.45%	73.47%	84.89%	79,60%	49.33%	94.70%	48.66%	64,23%	80.32%	92.35%	88.22%	86,96%
Spektrum+Ceps	88.98%	94.72%	92.96%	<b>92,22%</b>	70.65%	94.75%	55.06%	73,49%	88.98%	94.72%	92.96%	<b>92,22%</b>
Spektrum+Autokorr.	82.05%	89.28%	82.71%	84,68%	70.52%	95.60%	95.60%	87,24%	81.75%	94.42%	74.90%	83,69%

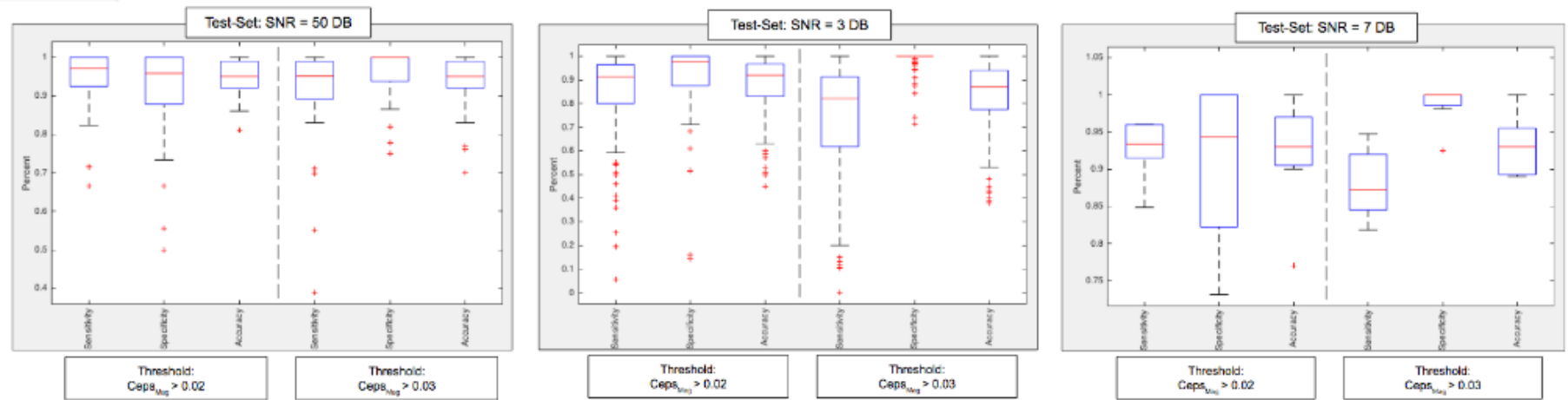


Abbildung A.1: Boxplot-Auswertung über Sensitivität, Spezifität und Genauigkeit der beiden besten Modelle