

Visualisierung kontinuierlicher, multimodaler Schmerz Scores am Beispiel akustischer Signale

Masterarbeit

Franz Anders
HTWK Leipzig

Januar 2017

Abstract

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen der medizinischen Schrei-Forschung	2
2.1	Schmerz Scores	2
2.2	Schmerz-Schrei aus medizinischer Sicht	4
2.3	Physio-Akustische Modellierung des Weinens	5
3	Grundlagen der Signalverarbeitung	8
3.1	Statistische Merkmale	9
3.2	Fehlersignale	11
3.3	Korrelation	12
3.4	Faltung	14
3.5	Diskrete Fourier-Transformation	15
3.6	Filter	15
3.7	akustische Modellierung der menschlichen Stimme	15
3.8	Feststellung von Periodizität in Signalen	15
3.8.1	Zero-Crossing-Rate	15
3.8.2	Methoden des Frequenzbereiches	15
3.8.3	Autokorrelation	15
3.8.4	Cepstrum	15
4	Zusammenfassung	16
	Appendices	19

Abbildungsverzeichnis

2.1	Veranschaulichung des Grundvokabulars	6
3.1	Ein zeit-kontinuierliches Signal (A) und ein zeit-diskretes Signal (B)	8
3.2	Ein nicht-periodisches Signal (A) und ein periodisches Signal (B)	9
3.3	Statistische Werte eines Signals über das Intervall [50,200]	10
3.4	Komponentenweise Addition und Multiplikation zweier Signale	11
3.5	Berechnung des MSE, RMSE und SNR eines von Rauschen gestörten Signals	12
3.6	Correlation der Signale x und y	13
3.7	Beispiel einer Running Correlation	14
3.8	Beispiel für die Faltung	15
.1	Boxplot-Auswertung über Sensitivity, Specificity und Accuracy der beiden VAD-Modelle	21

1 Einleitung

2 Grundlagen der medizinischen Schrei-Forschung

2.1 Schmerz Scores

Bei erwachsenen Menschen wird der Schmerzgrad typischerweise durch eine Selbsteinschätzung des Patienten unter der Leitung gezielter Fragen des Arztes vorgenommen. Bei Kindern unter 3 Jahren ist diese Selbsteinschätzung nicht möglich. Schmerz drückt sich in Veränderungen des psychologischen, körperlichen und biochemischen Verhaltens des Säuglings aus. Die für den Arzt am leichtesten feststellbaren Verhaltensänderungen sind von außen wahrnehmbare Merkmale, wie zum Beispiel ein Verkrampfen des Gesichtsausdrucks, erhöhte Körperbewegungen oder lang anhaltendes Weinen. Um eine weitestgehend objektive Schmerzfeststellung zu ermöglichen, wurden sogenannte *Pain-Scores* entwickelt, die durch ein Punktesystem den gesamten Schmerzgrad des Babies quantifizieren.[14] Es existieren *eindimensionale* Pain-Scores, die den Schmerz nur Aufgrund der Beobachtung eines Merkmals beurteilen, so wie beispielsweise die reine Beurteilung des Gesichtsausdrucks. *Mehrdimensionale* (auch *multimodale*) Pain-Scores beziehen mehrere Faktoren in das Scoring mit ein.[1]. Tabelle 2.1 zeigt das Scoring-System „Neonatal Infant Pain Scale“ (NIPS) als Beispiel für eine multimodale Pain-Score. Der Säugling wird anhand der aufgeführten Kategorien bewertet und alle vergebenen Punkte aufsummiert. Ein insgesamt Wert von > 3 zeigt Schmerz an, ein Wert von > 4 großen Schmerz.[7]

Tabelle 2.1: NIPS-Scoring

NIPS	0 points	1 point	2 points
Facial Expr.	Relaxed	Contracted	-
Cry	Absent	Mumbling	Vigorous
Breathing	Relaxed	Different than basal	-
Arms	Relaxed	flexed/stretched	-
Legs	Relaxed	flexed/stretched	-
Alertness	Sleeping	uncomfortable	-

In den meisten mehrdimensionalen Scoring-Systeme werden die Schreigeräusche mit einbezogen. Tabelle 2.2 zeigt eine Übersicht über eine ausgewählte Menge an multimodalen Pain-Scores. Alle Pain-Scores sind für Kleinkinder bis 3 Jahren gedacht. In der Übersicht wird nicht wiedergegeben, welche weiteren Merkmale jeweils in das Scoring mit einbezogen werden, oder welche Ingesamtpunktzahlen auf welche Schmerzintensität hinweisen. Es soll an dieser Stelle nur verdeutlicht werden, welche unterschiedlichen Ansätze zur Bewertung des Schreiens aus medizinischer Sicht im Zusammenhang mit Pain-Scores existieren. Folgende Beobachtungen lassen sich aus der Übersicht ziehen:

1. Die zu beobachtenden Eigenschaften des Weinens werden mit subjektiv behafteten Werten charakterisiert. Beispielsweise wird im N-PASS-System ist ein Schmerz-Schrei als „High-pitched or silent-continuous crying“ beschrieben. Es wird nicht fest definiert, was als „crying“ gilt oder welche Tonhöhe als „high-pitched“ ist. Auch die Erstquellen geben keine festen Definitionen.
2. Es gibt verschiedene Ansätze zur Bewertung des Weinens. Bei CRIE ist die Tonhöhe, bei BIIP die Länge und bei COMFORT die Art des Weinens entscheidend.
3. Die Beschreibungen sind kurz und prägnant gehalten, der Arzt hat in keinem der Modelle auf mehr als drei Parameter des Schreiens zu achten. Die Begründung liegt darin, dass bei allen Modellen a.) das Schreien nur eines von mehreren Faktoren ist, und b.) Die Schmerzbestimmung in einem vorgegebenen Zeitrahmen durchführbare sein muss.

System	P.	Description
FLACC[19]	0	No cry (awake or asleep)
	1	Moans or whimpers; occasional complaint
	2	Crying steadily, screams or sobs, frequent complaints
N-PASS[15]	-2	No cry with painful stimul
	-1	Moans or cries minimally with painful stimuli
	0	Appropriate Crying
	1	Irritable or Crying at Intervals. Consolable
	2	High-pitched or silent-continuous crying. Not consolable
BIIP[6]	0	No Crying
	1	Crying <2 minutes
	2	Crying >2 minutes
	3	Shrill Crying >2 minutes
CRIES[3]	0	If no cry or cry which is not high pitched
	1	If cry high pitched but baby is easily consoled
	2	If cry is high pitched and baby is inconsolable
COVERS[10]	0	No Cry
	1	High-Pitched or visibly crying
	2	Inconsolable or difficult to soothe
PAT[8]	0	No Cry
	1	Cry
DAN[4]	0	Moans Briefly
	1	Intermittent Crying
	2	Long-Lasting Crying, Continuous howl
COMFORT[11]	0	No crying

	1	Sobbing or gasping
	2	Moaning
	3	Crying
	4	Screaming
<hr/>		
MBPS[13]	0	Laughing or giggling
	1	Not Crying
	2	Moaning quiet vocalizing gentle or whimpering cry
	3	Full lunged cry or sobbing
	4	Full lunged cry more than baseline cry

Tabelle 2.2: Übersicht über Pain-Scores

2.2 Schmerz-Schrei aus medizinischer Sicht

Die Frage ist: Woher kommen diese unterschiedlichen Bewertungen des Weinens in Tabelle 2.2? Gibt es eine Pain-Score, die aus wissenschaftlicher Sicht „recht hat“? Dieser Fragestellung unterliegen unterliegen zwei grundlegendere Fragen: 1.) Ist es überhaupt möglich, anhand der akustischen Eigenschaften den Grund für den Schrei abzuleiten, also beispielsweise Hunger, Einsamkeit oder Schmerz? Anders formuliert: Gibt es überhaupt so etwas wie einen Schmerz-Schrei? 2.) Ist es möglich, anhand der akustischen Eigenschaften den Schweregrad des Unwohlseins abzuleiten (also beispielsweise den Grad des Schrei-Versursachenden Schmerzes)?

Die Annahme, dass es möglich ist, aus dem Schreien den Grund abzuleiten, wird als „Cry-Types Hypothesis“ bezeichnet. Die berühmtesten Befürworter dieser Hypothese ist eine skandinavische Forschungsgruppe, auch bezeichnet als „Scandinavian Cry-Group“, die diese Idee in dem Buch „Infant Crying: Theoretical and Research Perspectives“ [2] publik machte. Die Annahme ist, dass die verschiedenen Ursachen *Hunger, Freude, Schmerz, Geburt und Anderes* klare Unterschiede hinsichtlich ihrer akustischen Merkmale aufweisen, welche an einem Spektogramm ablesbar seien. Entsprechende Beispiele werden in dem Buch gegeben. Nur einige Jahre Später zeigte Müller et al [5] in einem Paper, dass bei leichter Veränderung der Bedingungen der Experimente die Unterscheidung nicht möglich ist. Die Gegenhypothese ist, dass Weinen „nichts als undifferenziertes Rauschen“ sei. 50 Jahre später liegt kein anerkannter Beweis für die eine oder andere Hypothese vor. Es gibt nur starke Hinweise dafür, dass die Plötzlichkeit des Eintretens des Schreigrundes hörbar ist. Ein plötzliches Ereignis, wie ein Nadelstich oder ein lautes Geräusch, führen auch zu einem plötzlich beginnenden Schreien. Ein langsam einretendes Ereignis, wie ein langsam immer stärker werdender physischer Schmerz oder langsam eintretender Hunger führen auch zu einem langsam eintretenden Weinens. Da keine Einigung herrscht, wird empfohlen, den Grund aus dem Kontext abzuleiten.[18]

Die Zweite Frage nach der Ableitung der Stärke des Unwohlseins aus den akustischen Eigenschaften des Geschreis wird in der Fachsprache unter dem Begriff *Cry as a graded Signal* subsumiert. Je „stärker“ das Weinen, desto höher das Unwohlsein (*Level of Distress (LoD)*) des Säuglings. Tatsächlich bemessen wird dabei der von dem Beobachter vermutete

Grad des Unwohlsein des Babies, und nicht der tatsächliche Grad, da dieser ohne die Möglichkeit der direkten Befragung des Kindes nie mit absoluter Sicherheit bestimmt werden kann. Dieser vermutete LoD wird entweder durch das subjektive Empfinden der Beobachter oder durch Pain-Scores festgestellt. Ein hohes Level of Distress hat vor allem eine schnelle Reaktion der Aufsichtspersonen zur Beruhigung des Babies zur Folge, womit dem Geschrei eine Art Alarm-Funktion zukommt. Es gibt starke Hinweise darauf, dass das Level of Distress anhand objektiv messbarer Eigenschaften des Audiosignals bestimmt werden kann. So herrscht beispielsweise weitestgehend Einigung darüber, dass ein „lang“ anhaltendes Geschrei auf einen hohen Level of Distress hinweist. Insofern aus dem Kontext des Schreiens Schmerz als wahrscheinlichste Ursache eingegrenzt werden kann, kann aus einem hohen Level of Distress ein hoher Schmerz abgeleitet werden. [18] und [17]

Es herrscht wiederum keine Einigung darüber, welche akustischen Eigenschaften im Detail ein hohes Level of Distress anzeigen. Carlo V Bellieni et al [4] haben festgestellt, dass bei sehr hohem Schmerz in Bezug auf die DAN-Scala (siehe Tabelle 2.2) die Tonhöhe des Geschreis steigt. Qiaobing Xie et al [17] haben festgestellt, dass häufiges und „verzerrtes“ Schreien (ohne feststellbares Grundfrequenz, da der Ton stimmlos erzeugt wird) auf einen hohen Level of Distress hinweist.[18] Diese Uneinigkeit hat wahrscheinlich zu den verschiedenen Bewertungen in den Pain-Scores geführt. 2.2.

2.3 Physio-Akustische Modellierung des Weinens

Das Ziel dieses Kapitels ist die Schaffung eines einheitlichen Vokabulares, auf den sich bezogen wird, um das Schreien eines Babys zu beschreiben. Die hier vorgestellten Begriffe stammen sowohl aus dem Buch „A Physioacoustic Model of the Infant Cry“ H Golub und M Corwin [9] als auch aus dem Paper „Rhythmic organization of the Sound of Infant Cry“ von Zeskind et al.[16]

Die Lautäußerung eines Neugeborenen, umgangssprachlich auch als „Weinen“ oder „Schreien“ bezeichnet, lässt sich im allgemeinen beschreiben als das „rhythmische Wiederholen eines beim ausatmen erzeugten Geräusches, einer kurzen Pause, einem Einatmungs-Geräusch, einer zweiten Pause, und dem erneuten Beginnen des Ausatmungs-Geräusches.“[21].

Das Vokabular, welches insbesondere von H Golub und M Corwin geschaffen wurde, ist sehr umfassend. An dieser Stelle wird eine Auswahl grundlegender Begrifflichkeiten vorgestellt, die in dieser Arbeit gebraucht werden. Sie werden in Abbildung 2.1 veranschaulicht.

Expiration beschreibt den Klang, der bei einem einzelnen, ununterbrochenem Ausatmen mit Aktivierung der Stimmbänder durch das Baby erzeugt wird. [16]. Der von Golub et al [9] verwendete Begriff **Cry-Unit** wird in dieser Arbeit synonym verwendet. Umgangssprachlich ist handelt es sich um einen einzelnen, ununterbrochenen *Schrei*.

Inspiration beschreibt den Klang, der beim Einatmen durch das Baby erzeugt wird.

Burst beschreibt die Einheit von einer Expiration und der darauf folgenden Inspiration. Das heisst, dass die zeitliche Dauer eines Bursts sowohl das Expiration-Geräusch, das Inspiration-Geräusch als auch die beiden Pausen zwischen diesen Geräuschen umfasst.

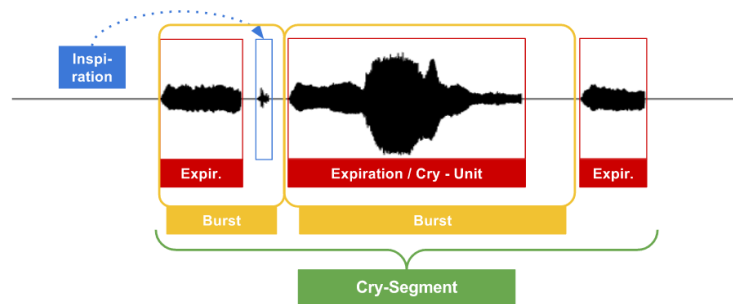


Abbildung 2.1: Veranschaulichung des Grundvokabulars

Praktisch ergibt sich das Problem, dass vor allem bei stärkerem Hintergrundrauschen die Inspiration-Geräusche häufig weder hörbar noch auf dem Spektrogramm erkennbar sind. Daher wird die Zeitdauer eines Bursts oder Cry-Unit vom Beginn einer Expiration bis zum Beginn der darauf folgenden Expiration definiert und somit allein von den Expirations auf die Bursts geschlossen. Implizit wird somit eine Inspiration zwischen zwei Expirations angenommen.

Cry die insgesamt klangliche Antwort zu einem spezifischen Stimulus. Eine Gruppe mehrerer Cry-Units.[9] In dieser Arbeit wird ein *Cry* als **Cry-Segment** bezeichnet, um Verwechslungen zu vermeiden.

Weiterhin wurden von H Golub und M Corwin [9] Cry-Units in eine der folgenden drei Kategorien eingeführt:

Phonation beschreibt eine Cry-Unit mit einer „vollen Vibration der Stimmbänder“ mit einer Grundfrequenz zwischen 250 und 700 Hz. Entspricht umgangssprachlich einem Weinen mit einem „klaren, hörbaren Ton“.

Hyper-Phonation beschreibt eine Cry-Unit mit einer „falsetto-artigem Vibration der Stimmbänder“ mit einer Grundfrequenz zwischen 1000 und 2000 Hz. Entspricht umgangssprachlich einem Weinen mit einem „sehr hohen, aber klaren, hörbaren Ton“.

Dysphonation beschreibt eine Cry-Unit ohne klar feststellbare Tonhöhe, produziert durch Turbulenzen an den Stimmbändern. Entspricht umgangssprachlichen dem „Brüllen oder Krächzen“.

Eine Cry-Unit gehört dabei mindestens einer dieser Kategorien an, kann aber auch in seinem zeitlichen Verlauf die Kategorie wechseln. H Golub und M Corwin [9] stellen weiterhin eine Reihe an charakteristischen Eigenschaften vor, die in Bezug auf ein Cry-Segment berechnet werden.

Latency-Period beschreibt die Dauer zwischen dem zufügen eines Schmerz-Stimulus und dem beginn des ersten Cry-Bursts des Segmentes

Duration beschreibt die insgesamt Zeitdauer des Cry-Segmentes. Es wird keine genaue Definition gegeben, wodurch Beginn und Ende definiert werden. Das Segment endet dort, wo es „scheint, aufzuhören“.

Maximum-Pitch beschreibt die höchste festgetellte Grundfrequenz des Segmentes.

... und viele weitere, die in [9] nachgelesen werden können, aus Platzgründen an dieser Stelle jedoch nicht vollständig genannt werden.

3 Grundlagen der Signalverarbeitung

Ein *Signal* ist eine Funktion eines Parameters mit numerischen Wertebereich. Die Abbildung zwischen Definitions- und Wertebereich kann, aber muss nicht durch eine Formel definiert sein. So fällt $f(x) = \sin(x)$ genauso unter die Definition eines Signals wie eine Folge numerischer Werte, die durch die Aufnahme eines Messgerätes entstanden sind. Weiterhin kommt dem Wertebereich eine gewisse Bedeutung zu, wie *Zeit* oder *Ort*. Ein typisches Beispiel für ein Signal ist die Spannung, die abhängig von der Zeit von einem Mikrofon erzeugt wird. Da in dieser Arbeit nur Signale von Bedeutung sind, deren Wertebereich sich auf die Zeit bezieht, konzentrieren sich alle folgenden Bereiche auf diesen Bereich. Im Zusammenhang mit Signalen wird der Definitionsbereich auch als *unabhängiger Parameter* und der Wertebereich auch als *abhängiger Parameter* bezeichnet. [20, S. 11-12] [12, S. 22-23]

Bei einem zeit-kontinuierlichen Signal $x(t)$ ist der Wertebereich kontinuierlich, wie in Formel 3.1 definiert. Bei einem zeit-diskreten Signal $x[n]$ ist der Wertebereich diskret, wie in Formel 3.2 definiert. So beschreibt beispielsweise $x[17] = s$ den Wert zur Zeit $n = 17$. „Zeit“ hat in diesem Kontext keine Einheit. Ein Wert wird auch als *Sample* oder *Amplitude* bezeichnet. $x[17]$ meint somit das 17. Sample des Signals. Abbildung 3.1 zeigt Beispiele für ein zeit-kontinuierliches und ein zeit-diskretes Signal. [12, S. 22 - 23]

$$x(t) = s, t \in \mathbb{R} \quad (3.1)$$

$$x[n] = s, n \in \mathbb{Z} \quad (3.2)$$

Zeit-diskrete Signale werden häufig dadurch gewonnen, dass ein zeit-kontinuierliches Signal in regelmäßigen Intervallen abgetastet wird. Dieser Prozess wird als *Sampling* bezeichnet und durch Formel 3.3 definiert. Der Parameter T_s wird als *Sampling-Interval* bezeichnet. Das Reziproke $\frac{1}{T_s} = f_s$ heißt *Sampling-Rate* und wird in der Einheit $\frac{1}{s} = \text{Hz}$. Eine Sampling-Rate von $f_s = 44\,100 \text{ Hz}$ bedeutete beispielsweise, dass ein Signal 44100 mal pro Sekunde

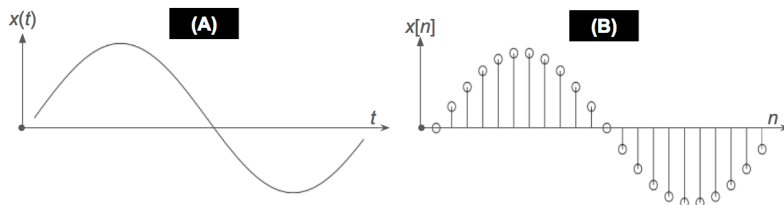


Abbildung 3.1: Ein zeit-kontinuierliches Signal (A) und ein zeit-diskretes Signal (B)

abgetastet wurde.[12, S. 24]

$$x[n] = s(n \cdot T_s), -\infty < n < \infty \quad (3.3)$$

Da in dieser Arbeit nur zeit-diskrete Signale von Interesse sind, werden ab diesem Punkt die Definitionen für zeit-kontinuierliche Signale ausgelassen. Der *Support* ist das kleinst mögliche Zeitintervall, der alle Samples enthält, die nicht den Wert 0 haben, wie Formel 3.4 definiert. Die *Dauer* eines Signales ist die Länge des Supportes nach Formel . Das Signal $x[n] = \cos(n)$, $0 \leq n \leq 3$ hat beispielsweise den Support $[0, 3] = \{0, 1, 2, 3\}$ und die Dauer 4. Ein *unendliches Signal* hat einen unendlichen langen Support, das heißt es gilt $\text{Duration}(x) = \infty$. Ein *endliches Signal* hat einen endlichen Support, das heißt $\text{Duration}(x) \neq \infty$. [12, S. 24]

$$\begin{aligned} \text{Sup}(x) &= [\text{sup}_s, \text{sup}_e], \text{sup}_s, \text{sup}_e \in \mathbb{Z} \\ , x[\text{sup}_s] &\neq 0 \wedge x[\text{sup}_e] \neq 0 \wedge \forall n \notin [\text{sup}_s, \text{sup}_e] : x[n] = 0 \end{aligned} \quad (3.4)$$

$$\text{Duration}(x) = \text{sup}_e - \text{sup}_s + 1 \quad (3.5)$$

Ein Signal gilt als *periodisch*, wenn Formel 3.6 erfüllt ist. Der Parameter N wird als Periode von x bezeichnet. Wenn ein Signal mit N periodisch ist, dann ist es auch mit $2N, 3N, \dots$ periodisch. Die Grundfrequenz N_0 ist das kleinste N , für das Formel 3.6 erfüllt ist. Abbildung 3.2 zeigt ein Beispiel für ein nicht-periodisches und ein periodisches Signal. [12, S. 24]

$$\exists N : \forall n \in \text{Sup} : x[n + N] = x[n] \rightarrow \text{Periodisch}(x, N) = \text{true} \quad (3.6)$$

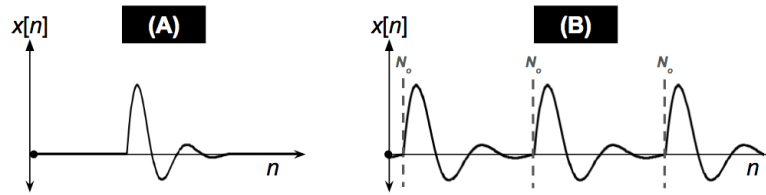


Abbildung 3.2: Ein nicht-periodisches Signal (A) und ein periodisches Signal (B)

3.1 Statistische Merkmale

Im folgenden wird ein Überblick über die häufig verwendete Signaleigenschaften gegeben. Abbildung 3.3 visualisiert die Erläuterungen.

1. Der **Maximalwert** / **Minimalwert** beschreibt den höchsten / niedrigsten in x enthaltenen Wert nach den Formel $\max(x)$ und $\min(x)$.

2. Der **Durchschnittswert** / **Average Value** beschreibt den durchschnittlichen Wert aller Samples von x nach Formel 3.7. Dieser Durchschnittswert wird über dem Intervall $[n_1, n_2]$ berechnet.

$$\text{AVG}(x) = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} x[n] \quad (3.7)$$

3. Der **Mean Squared Value** (MSV) beschreibt den quadrierten Durchschnittswert über eine bestimmtes Intervall nach Formel 3.8. Er wird auch als *durchschnittliche Energie* oder *average Power* bezeichnet.

$$\text{MSV}(x) = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} x[n]^2 \quad (3.8)$$

4. Das **Root Mean Square** (RMS) ist die Wurzel des Mean Squared Value nach Formel 3.9. Der RMS findet häufiger Anwendung als der MSV, da er besser ins Verhältnis zu den Werten des Signals gesetzt werden kann. Er wird im Deutschen auch als **Effektivwert** oder **Durchschnittsleistung** bezeichnet. Da die deutschen Begriffe in einigen Quellen jedoch auch für den MSV verwendet werden, wird an dieser Stelle nur mit den englischen Begriffen gearbeitet.

$$\text{RMS}(x) = \sqrt{\frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} x[n]^2} \quad (3.9)$$

5. Die **Energie** / **Energy** bezeichnet die „Stärke“ eines Signals über einen bestimmten Intervall nach Formel 3.10. Sie entspricht dem MSV-Wert multipliziert der Länge des Intervalls. [12, S. 27-28]

$$E(x) = \sum_{n=n_1}^{n_2} x[n]^2 \quad (3.10)$$

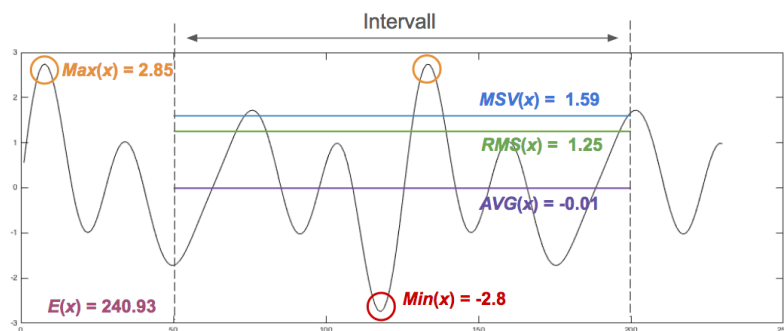


Abbildung 3.3: Statistische Werte eines Signals über das Intervall $[50, 200]$

Die Addition und Multiplikation wird bei Signalen komponentenweise durchgeführt, das heißt $x_1[n] + x_2[n] = y[n]$ und $x_1[n] \cdot x_2[n] = y[n]$. Abbildung 3.4 visualisiert diese Operationen.

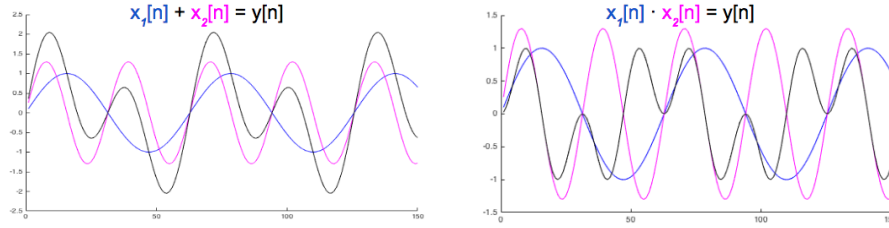


Abbildung 3.4: Komponentenweise Addition und Multiplikation zweier Signale

3.2 Fehlersignale

Die Addition wird unter anderem für die Modellierung des Einflusses von Störungen benötigt. Angenommen, ein Signal x wird übertragen, auf dem Übertragungsweg jedoch durch ein anderes Störsignal wie z.B. Rauschen e überlagert. Dieses Störsignal wird in diesem Zusammenhang auch als „Fehler-Signal bezeichnet. Das resultierende Signal x' wird nach Formel 3.11 berechnet. Kennt man sowohl das Eingangssignal x als auch das Ausgangssignal x' , kann das Störsignal e nach Formel 3.12 berechnet werden.

$$x'[n] = x[n] + e[n] \quad (3.11)$$

$$e[n] = x'[n] - x[n] \quad (3.12)$$

Errechnet man nun den den MSV- oder RMS-Wert des Störsignales e , gibt das Ergebnis einen Eindruck über die „Stärke“ des Fehler-Signals. Der MSE-Wert des Fehlers wird in diesem Zusammenhang auch als *Mean Squared Error* (*MSE*) und der RMS-Wert als *Root Mean Squared Error* (*RMSE*) oder einfach als *Fehler* oder *Error* bezeichnet. Formel 3.13 und 3.14 definieren die Berechnungen des MSE und RMSE. Der RMSE hat im Gegensatz zum MSE den Vorteil, dass er besser ins Verhältnis zu den Werten des Fehlersignals gestetzt werden kann. Ein $\text{RMSE} = 0$ heisst, dass $x = x'$ und somit kein Störsignal vorliegt. Ein $\text{RMSE} = \text{RMS}(x)$ heisst, dass Eingangs- und Störsignal den selben Effektivwert und somit die selbe „stärke“ besitzen. Abbildung 3.5 visualisiert die Berechnung des MSE und RMSE. [12, S: 28 - 29]

$$\text{MSE}(x, x') = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} (x[n] - x'[n])^2 \quad (3.13)$$

$$\text{RMSE}(x, x') = \sqrt{\frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} (x[n] - x'[n])^2} \quad (3.14)$$

Eine weitere Betrachtungsweise bezüglich der Stärke des Rauschens auf das Signal ist, das Eingangssignal ins Verhältnis zum Rauschsignal zu setzen. Formel 3.15 gibt die Definition. Ein $\text{SNR}_{\text{rel}}(x, e) = 1$ heißt, dass das Eingangssignal den selben MSV wie das Fehlersignal hat. Meistens ist der MSV des Eingangssignals in der Praxis sehr viel höher als der des Fehler-Signals. Um den Zahlenraum zu begrenzen, wird die Pseudo-Einheit dB verwendet. Formel 3.16 den so berechneten *Signal-Rausch-Abstand* (SNR , englisch Signal-to-Noise-Ratio). Entgegen des MSE weist ein *niedriger* SNR-Wert auf ein *starkes* Rauschen hin, und ein *hoher* SNR auf ein *schwaches* Rauschen! Abbildung 3.5 visualisiert die Berechnung des SNR.

$$\text{SNR}_{\text{rel}}(x, e) = \frac{\text{MSV}(x)}{\text{MSV}(e)} \quad (3.15)$$

$$\text{SNR}(x, e) = 10 \cdot \lg \left(\frac{\text{MSV}(x)}{\text{MSV}(e)} \right) \text{ dB} \quad (3.16)$$

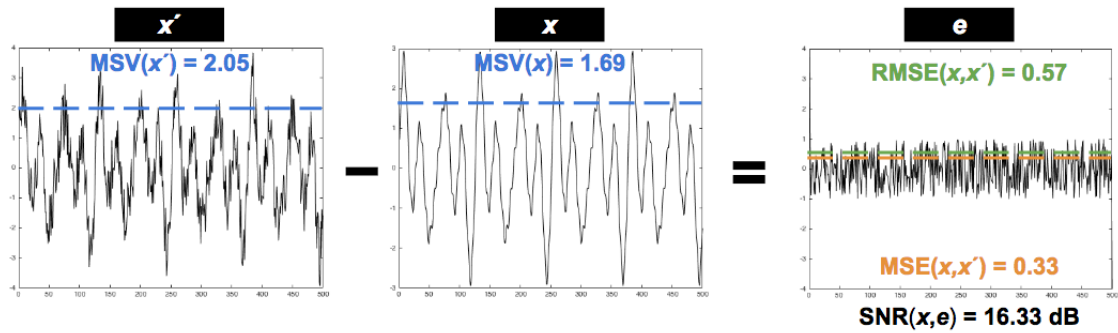


Abbildung 3.5: Berechnung des MSE, RMSE und SNR eines von Rauschen gestörten Signals

3.3 Korrelation

Die *Korrelation* (engl *Correlation*) zweier Signale x_1 und x_2 wird nach Formel 3.17 als die Summe aller Samples des Produktes der beiden Signale über einen bestimmten Intervall $[n_1, n_2]$ definiert. Das Ergebnis ist ein Wert $\in \mathbb{R}$ welches die „Ähnlichkeit der beiden Signale“ kennzeichnet. Ein Positiver Wert weist auf eine *positive Korrelation* hin, ein negativer Wert auf eine *negative Korrelation*, und ein Wert von $\text{Corr}(x_1, x_2) = 0$ auf *keine Korrelation*. Aus der Größe des Wertes kann die Stärke der Korrelation jedoch nicht direkt interpretiert werden. Bei der *normalisierten Korrelation* $\text{Corr}_N(x, y)$ wird daher die Korrelationswert ins Verhältnis zu den Energien der beiden Signale gesetzt, wie in Formel 3.18 definiert. Der Wertebereich der normalisierten Autokorrelation ist $-1 \leq \text{Corr}_N(x, y) \leq +1$. Daraus ergeben sich die in Formel 3.19 definierten Zusammenhänge. Ein Wert von $\text{Corr}_N(x, y) = 1$ wird auch als *perfekte Korrelation* bezeichnet, ein Wert von $\text{Corr}_N(x, y) = -1$ als *anti-perfekte Korrelation* [12, S. 46 - 47] Abbildung 3.6 visualisiert die normalisierte Korrelation

eines Signales x mit den Signalen y_n .

$$\text{Corr}(x, y) = \sum_{n=n_1}^{n_2} x[n] \cdot y[n] \quad (3.17)$$

$$\text{Corr}_N(x, y) = \frac{\text{Corr}(x, y)}{\sqrt{E(x) \cdot E(y)}} \quad (3.18)$$

$$\text{Corr}_N(x, y) = \begin{cases} 1 & \rightarrow x = y \\ -1 & \rightarrow x = -y \end{cases} \quad (3.19)$$

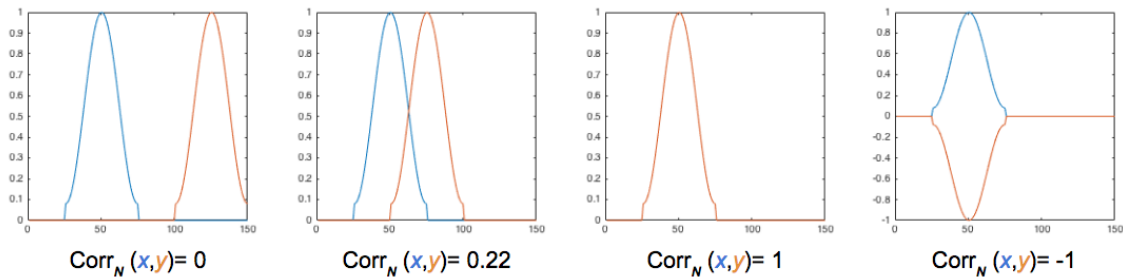


Abbildung 3.6: Correlation der Signale x und y

Die Korrelation und die normalisierte Korrelation werden aufgrund ihrer Eigenschaften verwendet, um ein Signal x in einem Signal y zu detektieren. Häufig ist das Ziel, ein von einem Rauschen e überlagertes Signal $x + e = y$ auf das Vorhandensein des erwarteten Signales x hin zu überprüfen. Wie in Abbildung 3.6 zu sehen ist, ist der Korrelationswert jedoch von der Verzögerung des Signals abhängig. Daher wird in der *Cross – Correlation* das Signal y mit einer verzögerten Varianten des Signals x korreliert, wie in Formel 3.20 definiert. Der parameter k wird als *Lag* bezeichnet und gibt die Verzögerung an.

$$\text{X-Corr}(x, y, k) = \sum_{n=-\infty}^{\infty} x[n - k] \cdot y[n] \quad (3.20)$$

Im Prozess der so genannten *Running Correlation* nutzt man die Cross-Correlation mit den Lags $k = 0 \dots k_{max}$ zur Erstellung des *Korrelationssignals* r , wie in Gleichung 3.21 definiert. Das Signal r gibt Auskunft, zu welchen Verzögerungswerten k die größten Ähnlichkeiten zwischen x und y gefunden wurden.

$$r[k] = \text{X-Corr}(x, y, k) \quad , k = 0, \dots, k_{max} \quad (3.21)$$

Abbildung 3.7 zeigt ein Beispiel für die Erzeugung von r mit der Sliding Correlation. (A) zeigt das zu detektierende Signal x und (B) das Signal y . (C) zeigt das Korrelationssignal r mit den Lags $k = 1, \dots, 1150$ [12, S. 47 - 48]

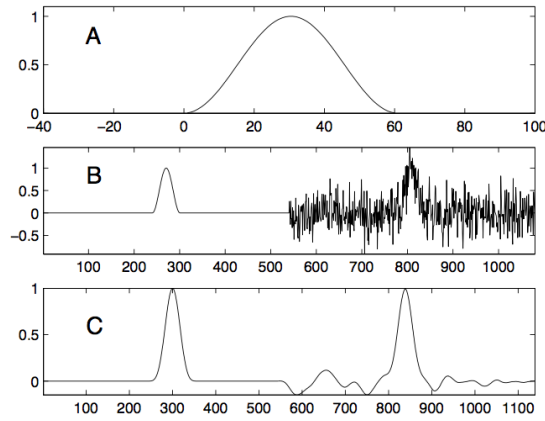


Abbildung 3.7: Beispiel einer Running Correlation

3.4 Faltung

Die *Faltung* (engl. *Convolution*) ist eine der Zentralen Operationen zwischen zwei Signalen, so wie die Addition oder die Multiplikation. Sie wird mit dem Symbol $*$ notiert. Sie wird notiert mit $x * h = y$.

Die Faltung basiert auf der *Faltungs-Summe*, welche die Faltung zunächst Punktweise definiert. Die Gleichung wird in Formel 3.22 abgebildet. In diesem Zusammenhang wird x Eingangs- und y als Ausgangs-Signal bezeichnet. Je nach Anwendungsfall bekommt h den Namen *Faltungs-Kernel*, *Filter-Kernel* oder einfach *Kernel*. [20, S. 107-108]

$$y[n] = x[n] * h[n] = \sum_{i=1}^M h[n] * x[i - n] \quad (3.22)$$

Das Ergebnis einer Berechneten Faltungssumme nach 3.22 ist ein einzelner Wert. Wird die Faltungs-Summe ähnlich der Cross-Correlation für $n = 1 \dots N + M - 1$ durchgeführt, ist das Ergebnis ein Signal. Die tatsächliche Faltung wird in Gleichung 3.23 definiert. x ist ein Signal mit $\text{Support}(x) = [1, N]$ und $\text{Duration}(x) = N$, h ist ein Signal mit $\text{Support}(x) = [1, M]$ und $\text{Duration}(h) = M$ und y ist ein Signal mit $\text{Support}(y) = [1, N + M - 1]$ und $\text{Duration}(y) = N + M - 1$. Das heißt, dass das Eingangssignal um die Länge des Faltungskerns verlängert wird. Abbildung 3.8 zeigt ein Beispiel für die Faltung. [20, S. 115-120]

$$y = x * h = [x[1] * h[1], \dots, x[N + M - 1] * h[N + M - 1]] \quad (3.23)$$

Das neutrale Element der Faltung ist der *Delta-Funktion*, definiert in Gleichung 3.24. Das heißt, dass $x * \delta = x$. Die Faltung ist kommutativ, das heißt $x * h = h * x = y$. [20, S. 107, 113]

$$\delta[n] = \begin{cases} 1 & , n = 0 \\ 0 & , n \neq 0 \end{cases} \quad (3.24)$$

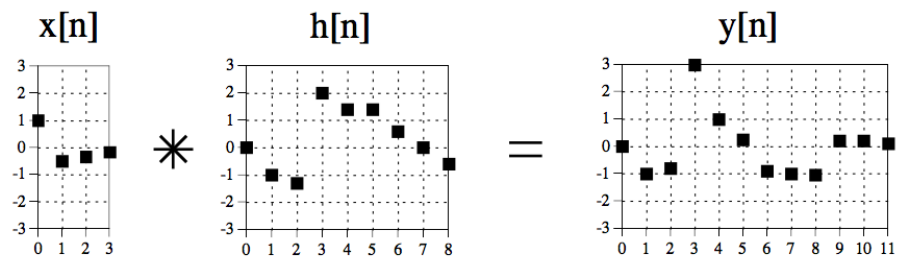


Abbildung 3.8: Beispiel für die Faltung

3.5 Diskrete Fourier-Transformation

3.6 Filter

3.7 akustische Modellierung der menschlichen Stimme

3.8 Feststellung von Periodizität in Signalen

3.8.1 Zero-Crossing-Rate

3.8.2 Methoden des Frequenzbereiches

3.8.3 Autokorrelation

3.8.4 Cepstrum

4 Zusammenfassung

Literaturverzeichnis

- [1] K J S Anand. *Pain in Neonates and Infants*. Elsevier, 2007.
- [2] Zachariah Boukydis Barry Lester. *Infant Crying: Theoretical and Research Perspectives*. Springer, 1985.
- [3] Judy Bildner. *CRIES Instrument Assessment Tool of Pain in Neonates*. City of Hope Pain, 1997. Online unter <http://prc.coh.org/pdf/CRIES.pdf>.
- [4] R Sisto & Giuseppe Buonocore Carlo Bellieni, Franco Bagnoli. Cry features reflect pain intensity in term newborns: An alarm threshold. *Pediatric Research*, 5:142–146, 1. Online unter https://www.researchgate.net/publication/297827342_Cry_features_reflect_pain_intensity_in_term_newborns_An_alarm_threshold.
- [5] H. Hollien & T Murry E Müller. Perceptual responses to infant crying: identification of cry types. *Journal of Child Language*, 1(1):89–95, 1974. Online unter <https://www.cambridge.org/core/journals/journal-of-child-language/article/perceptual-responses-to-infant-crying-identification-of-cry-types/4F0F8088116FCE381851D8D560697A5F>.
- [6] Jan Hamers & Peter Gessler Eva Cignac, Romano Mueller. Pain assessment in the neonate using the Bernese Pain Scale for Neonates. *Early Human Development*, 78(2):125–131, 2004. Online unter <http://www.sciencedirect.com/science/article/pii/S0378378204000337>.
- [7] Health Facts For You. *Using Pediatric Pain Scales Neonatal Infant Pain Scale (NIPS)*, 2014. Online unter <https://www.uwhealth.org/healthfacts/parenting/7711.pdf>.
- [8] Hodgkinson. Neonatal Pain Assessment Tool , 2012. Online unter http://www.rch.org.au/uploadedFiles/Main/Content/rchcpg/hospital_clinical_guideline_index/PAT%20score%20update.pdf.
- [9] Michael J Corwin Howard L Golub. A Physioacoustic Model of the Infant Cry. In *Infant Crying - Theoretical and Research Perspectives*, chapter 3, pages 59 – 82. Plenum, 1985.
- [10] Donna Geiss Laura Wozniak & Charles Hall Ivan Hand, Lawrence Noble. COVERS Neonatal Pain Scale: Development and Validation. *International Journal of Pediatrics*, 2010, 2010. Online unter <https://www.hindawi.com/journals/ijpedi/2010/496719/>.
- [11] Hans M Koot Dick Tibboel Jan Passchier & Hugo Duivenvoorden Monique van Dijk, Josien de Boer. The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. *Pain*, 84(2):367—377, 2000. Online unter

- <http://www.sciencedirect.com/science/article/pii/S0304395999002390>.
- [12] D L Neuhoff. *Signal and Systems I - EECS 206 Laboratory*. The University of Michigan, 2002. Online erhältlich unter: <http://www.eecs.umich.edu/courses/eecs206/archive/spring02/> abgerufen am 11. Januar 2016.
 - [13] Taddio Nulman. A revised measure of acute pain in infants. *J Pain Symptom Manage*, 10:456–463, 1995. Online unter [http://geriatricphysio.yolasite.com/resources/Modified%20Behavioral%20Pain%20Scale%20\(MBPS\)%20in%20infants.pdf](http://geriatricphysio.yolasite.com/resources/Modified%20Behavioral%20Pain%20Scale%20(MBPS)%20in%20infants.pdf).
 - [14] 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>.
 - [15] Steven Creech & Marc Weiss. Patricia Hummel, Mary Puchalski. N-PASS: Neonatal Pain, Agitation and Sedation Scale – Reliability and Validity. *Pediatrics/Neonatology*, 2(6), 2004. Online unter <http://www.anestesiarianimazione.com/2004/06c.asp>.
 - [16] 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>.
 - [17] R Ward & C Laszlo Qiaobing Xie. Automatic Assessment of Infants’ Levels-of-Distress from the Cry Signals. *IEEE Transanctions on Speech and Audio Processing*, 4(4):253–265, 1996. Online unter <http://ieeexplore.ieee.org/document/506929/>.
 - [18] Brian Hopkins & James Green Ronald Barr. *Crying as a Sign, a Symptom, and a Signal*. Mac Keith Press, 2000.
 - [19] J R Shayevitz & Shobha Malviya Sandra Merkel, Terri Voepel-Lewis. The FLACC: A Behavioral Scale for Scoring Postoperative Pain in Young Children. *Pediatric Nursing*, 23(3):293–7, 1996. Online unter https://www.researchgate.net/publication/13998379_The_FLACC_A_Behavioral_Scale_for_Scoring_Postoperative_Pain_in_Young_Children.
 - [20] Steven W. Smith. *The Scientist and Engineer’s Guide to Digital Signal Processing*. California Technical Publishing, 1999. Online erhältlich unter: <http://www.dspguide.com/pdfbook.htm>.
 - [21] P H Wolff. The role of biological rhythms in early psychological development. *Bulletin of the Menninger Clinic*, 31:197–218, 1967.

Appendices

Tabelle .1: Accuracy-Werte der Grenzwertfindung mit REPTree

$S_{Training}$	3 dB				50 dB				50+3 dB			
A_{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%
Freq	82.05%	89.28%	82.71%	84,68%	70.52%	94.37%	55.06%	73,31%	81.75%	91.22%	74.90%	82,62%
Ceps	88.98%	94.72%	92.96%	92,22%	86.83%	94.68%	92.83%	91,45%	88.98%	94.72%	92.96%	92,22%
Corr	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+Freq	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+Ceps	88.98%	94.72%	92.96%	92,22%	86.83%	94.68%	92.83%	91,45%	88.98%	94.72%	92.96%	92,22%
Zeit+Corr	80.45%	73.47%	84.89%	79,60%	49.33%	94.70%	48.66%	64,23%	80.32%	92.35%	88.22%	86,96%
Freq+Ceps	88.98%	94.72%	92.96%	92,22%	70.65%	94.75%	55.06%	73,49%	88.98%	94.72%	92.96%	92,22%
Freq+Corr	82.05%	89.28%	82.71%	84,68%	70.52%	95.60%	95.60%	87,24%	81.75%	94.42%	74.90%	83,69%

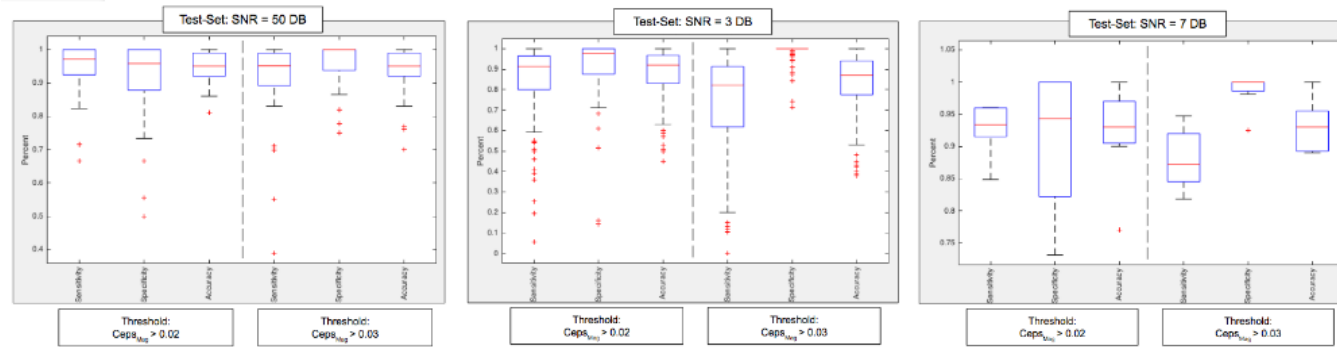


Abbildung .1: Boxplot-Auswertung über Sensitivity, Specificity und Accuracy der beiden VAD-Modelle