



Akustisches Interface zur Tremoranalyse

BACHELORARBEIT

aus

Computermusik und Multimedia

eingereicht am

Institut für Elektronische Musik und Akustik

Universität für Musik und darstellende Kunst Graz

Jänner 2014

Studiengang: Elektrotechnik-Toningenieur

Verfasserin: Renate Walter

Betreuer: Univ.Prof. Dr.phil. Gerhard Eckel

Inhaltsverzeichnis

1	Einführung	4
1.1	Projektüberblick	6
1.2	Motivation	7
1.3	Ziel	7
2	Begriffe	8
2.1	Sonifikation	8
2.2	Tremor	10
3	Klinische Diagnoseverfahren	12
3.1	Klassische Diagnose	12
3.2	Accelerometrie	12
4	Probleme der klinischen Diagnose	15
4.1	Probleme der klassischen Diagnose	15
4.2	Probleme der Accelerometrie	15
5	Sonifikation als Analysewerkzeug	16
6	Messumgebung	17
6.1	Signalweg	17
6.2	Equipment	18
6.2.1	Beschleunigungssensor	19
6.2.2	Interface Unit - Subject-Unit	19
6.2.3	Base-Unit	20
6.2.4	Datenerfassung - Power 1401 mkII	21
6.2.5	Software - Spike	22
7	Datenaufbereitung	23
7.1	Automatic Gain Control	24
7.2	Principal Component Analysis (PCA)	24
7.3	Amplitudenmodulation	26
7.4	Frequenzmodulation	27
7.5	Chorus	27
8	Sonifikationen	28
8.1	Erste Umsetzungen	28
8.1.1	Karplus-Strong-Synthese	28
8.1.2	Spectral Shaping	29
8.2	Finale Umsetzungen	31

8.2.1	„VOCODER“	31
8.2.2	„BIRDS“	37
9	Benutzeroberfläche	41
10	Ausblick	44
A	Principal Component Analysis (PCA)	45
A.1	PCA zur Dimensionsreduzierung	45
A.2	PCA zur Bestimmung von rotatorischen und translatorischen Signalanteilen	46
B	Empirical Mode Decomposition (EMD)	48
B.1	Einführung	48
B.2	Algorithmus	49
B.3	Probleme in Hinblick auf die Online-EMD	52

1 Einführung

Das vorliegende Dokument gibt einen umfassenden Überblick über das vom Land Steiermark geförderte Projekt „*Akustisches Interface zur Tremoranalyse*“ aus dem Förderprogramm *Human Technology Interfaces 2011*.

In einer Entwicklungsphase von 14 Monaten wurde in Zusammenarbeit des Instituts für Elektronische Musik und Akustik (IEM) der Kunstuniversität Graz mit der Universitätsklinik für Neurologie der Medizinischen Universität Graz nach Methoden zur Sonifikation von Tremordaten geforscht.

Die Programmierung der Sonifikationen erfolgte größtenteils in *Pure Data (PD)* und *MATLAB* und wurde von den beiden Entwicklern *MMag. David Pirrò (IEM)* und *DI Alexander Wankhammer (IEM)* umgesetzt.

Als Einstieg erfolgt nun ein kurzer Überblick über die im Laufe des Projekts gesammelten Eindrücke.

Eingangs wurde ich mit der Problematik bei der Diagnosefindung vertraut gemacht und konnte als Testpatientin die Erfahrung einer Untersuchung, wie sie in Abschnitt 3 beschrieben wird, nachvollziehen. Dabei konnte ich mir auch ein Bild über das vorhandene Equipment machen, das an der Universitätsklinik für Neurologie zum Einsatz kommt.

Zahlreiche Datensätze mussten gesammelt werden um möglichst viel Vergleichsmaterial zur Verfügung zu haben. Dabei hatte ich die Möglichkeit Untersuchungen beizuwohnen und Beobachtungen zu machen, die sich einerseits auf die Messung selbst konzentrierten, aber auch auf die Situation der Untersuchung selbst, wo mir die Notwendigkeit einer möglichst raschen und sicheren Diagnose bewußt wurde.

Im Laufe des Projekts wurden „Workshops“ mit den teilnehmenden ÄrztInnen durchgeführt, wo die entwickelten Sonifikationen vorgestellt wurden. Bei dieser Gelegenheit wurden die realisierten Klänge gemeinsam evaluiert und erste Testdiagnosen durchgeführt. Es wurde auf die unterschiedlichen Hörerfahrungen eingegangen und Ideen für weitere Entwicklungen gesammelt.

In gemeinsamen Gesprächen mit den beiden Entwicklern wurden mir die Grundlagen der einzelnen Sonifikationen erklärt und mögliche Probleme aufgezeigt.

Zahlreiche Trainingsstunden wurden investiert um bei der End-Evaluierung teilzunehmen. Die verschiedenen Eigenschaften der Sonifikationen wurden dabei hörend erfasst und die Zuordnung der Klangbilder zu den Erkrankungen geübt.

Diese Bachelorarbeit basiert nun auf persönlichen Beobachtungen, geteilten Erfahrungen und Gesprächen mit den Entwicklern und Ärztinnen. Zahlreiche Notizen, Pläne, Unterla-

gen und Informationen aus Handbüchern und wissenschaftlichen Artikeln wurden gesammelt und zur Darstellung des Sachverhalts vereint. Die Informationen zu den technischen Beschreibungen in den Kapiteln 4.2, 7, 8.2, A.2 und B.3, zu denen keine direkte Quelle angegeben werden kann, entstammen aus Gesprächen mit den beiden Entwicklern und gesammelten Notizen.

Am Ende dieser Einleitung möchte ich mich vor allem bei *MMag. David Pirrò* und *DI Alexander Wankhammer* für den bereichernden Erfahrungsaustausch bedanken.

Abschnitt 1 gibt einen Überblick über die teilnehmenden Institutionen und Personen, sowie Motivation und Ziel hinter dem Projekt.

In **Abschnitt 2** werden die Begriffe Sonifikation und Tremor erläutert um tieferen Einblick in die Materie zu erhalten.

Die klassischen Diagnosemethoden, wie sie im klinischen Betrieb angewandt werden, werden im **Abschnitt 3** beschrieben. Er enthält ebenfalls eine Einführung über die Accelerometrie, die die Grundlage zur Datengewinnung aller Sonifikationsmethoden darstellt. **Abschnitt 6** enthält Informationen über das Messequipment, wie es an der Universitätsklinik für Neurologie vorzufinden ist und stellt den Prozess der Datengewinnung dar.

In **Abschnitt 7** erfolgt eine Einführung in die Signalverarbeitung, die der Datenaufbereitung und -modifikation zu Grunde liegt.

Die umgesetzten Sonifikationen, angefangen von ersten Ideen bis zu tatsächlichen Realisierungen werden in **Abschnitt 8** vorgestellt. Neben der Beschreibung der unterschiedlichen Algorithmen erfolgt eine Darstellung der Klangeindrücke welche eine Differenzierung der sonifizierten Krankheitsbilder erlauben.

Abschnitt 9 gibt Einblicke in die grafische Benutzeroberfläche, die zur Bedienung des Sonifikationstools entwickelt wurde.

In **Appendix A und B** sind mathematische Hintergrundinformationen zu finden, auf denen die Realisierungen der einzelnen Sonifikationen basieren.

1.1 Projektüberblick

Projekttitel: *Akustisches Interface zur Tremoranalyse*

Förderprogramm: *Human Technology Interfaces 2011*

Projektstart und -ende: *01.09.2011 - 31.10.2012 (14 Monate)*

Teilnehmende Forschungseinrichtungen:

*Institut für Elektronische Musik und Akustik (IEM),
Universität für Musik und Darstellende Kunst Graz
Universitätsklinik für Neurologie,
Medizinische Universität Graz*

Projektleitung:

*Univ. Prof. Mag. Dr. Robert Höldrich (IEM)
Univ. Prof. Dr. Reinhold Schmidt (Univ.-Klinik für Neurologie)*

Projektmitarbeiter:

*Dr. Petra Katschnig-Winter (Univ.-Klinik für Neurologie)
Dr. Petra Schwingenschuh (Univ.-Klinik für Neurologie)
MMag. David Pirrò (IEM)
DI Dr. Alois Sontacchi (IEM)
DI Alexander Wankhammer (IEM)
Renate Walter (IEM)*

1.2 Motivation

„Physiologischer Tremor zeigt sich bei betroffenen Patienten als rhythmische, unwillkürliche Zitterbewegung eines oder mehrerer Körperabschnitte und ist die am häufigsten auftretende Bewegungsstörung beim Menschen. Da Tremor durch unterschiedlichste neurologische Erkrankungen hervorgerufen werden kann, ist eine richtige und möglichst rasche Diagnosefindung die Grundvoraussetzung für eine erfolgreiche Behandlung.“ [1]

1.3 Ziel

„Der menschliche Gehörsinn hat sich im Laufe der Evolution zu einem Messinstrument entwickelt, dessen Genauigkeit, Adaptionfähigkeit, Fehlertoleranz und hierarchische Strukturierungsfähigkeit jedes technische System in den Schatten stellt. Die gesamte neue Forschungsströmung im Bereich der Sonifikation nutzt dieses Potential der mehrdimensionalen Abbildungsmöglichkeit beim Hörprozess, um neuartige und innovative Werkzeuge zur Analyse und Auswertung komplexer Datenströme zu entwickeln. Vor allem bei der Analyse stark dynamischer Vorgänge - wie sie etwa Tremorbewegungen darstellen - bietet daher das Hinzuziehen der Fähigkeiten des menschlichen Gehörs zur Erkennung von Mustern (hier vorwiegend Rhythmik, Dynamik, Einzelereignisse und Fehlertoleranz) enorme Vorteile.“ [1]

2 Begriffe

2.1 Sonifikation

„Sonifikation ist die nichtsprachliche Darstellung von Daten mittels Klang und Geräuschen.“ [2]

Diese sehr allgemein gehaltene Definition erlaubt nur eine vage Vorstellung über den Begriff der Sonifikation.

Eine detailliertere Beschreibung lautet:

„Jede Technik, welche Daten als Eingabe erhält und im Ergebnis einen Klang liefert, kann genau dann als Sonifikation bezeichnet werden, wenn folgende Bedingungen zutreffen:

- *Der Klang korrespondiert mit den Eigenschaften der Daten.*
- *Die Transformation ist vollständig systematisch. Das bedeutet, dass es eine exakte Definition gibt, wie die Daten (und optionale Interaktionen) zu Klängen führen.*
- *Die Ausführung (Klangerzeugung) ist reproduzierbar: mit gleichen Daten und gleichen Interaktionen müssen strukturell gleiche Klänge entstehen.*
- *Das System kann mit unterschiedlichen Daten, aber auch wiederholt mit den gleichen Daten benutzt werden.“ [3]*

Die Analyse von Daten über das Gehör wird zwar schon erfolgreich eingesetzt, die Selbstverständlichkeit der Anwendung hat sich in vielen Bereichen noch nicht durchgesetzt. Erfolgreiche Beispiele dazu sind:

- der *Geigerzähler* der über Knacksgeräusche Hinweise über die Anzahl auftretender Ionen auf die Elektroden des Geigermessgeräts gibt,
- die *Doppler-Sonographie* erlaubt die Bestimmung von Richtung und Geschwindigkeit des Blutes in Blutgefäßen oder
- die Analyse und Überwachung von Erdbewegungen über *seismographische* Aufzeichnungen.

Das Gehör als Analysewerkzeug besitzt zahlreiche Vorteile gegenüber der visuellen Datenanalyse. Im Folgenden werden einige genannt:

Mehrdimensionalität: Die Fähigkeit des Gehörs gleichzeitig mehrere Dimensionen zu erfassen stellt grafische Interpretationsformen häufig in den Schatten. *„So kann beispielsweise in einem Sprachsignal gleichzeitig der Sprecher, sein Geschlecht, der Wortsinn, der*

Gemütszustand und sogar Heiserkeit oder andere Erkrankungen erkannt werden.“ [3]

Fokussierung: Der Hörsinn ist fähig, sich unter Beisein anderer Schallquellen auf ein Klangbild zu konzentrieren. Als Beispiel sei hier ein Konzertbesuch genannt, wo man ein Orchester als eine einheitliche Klangquelle wahrnehmen kann oder sich auf das Spiel einzelner Instrumente konzentriert und dieses somit in den Fokus seiner Wahrnehmung stellt.

Mustererkennung: Aufgrund der hohen Zeitauflösungsfähigkeit des Gehörs ist es besonders geeignet feinste Muster in Form von Rhythmen aus Klangströmen herauszufiltern und wieder zu erkennen. Die Fähigkeit, geringste Änderungen über die Zeit wahrzunehmen, ermöglicht besonders die Überwachung laufender Prozesse und die Analyse zeitvarianter Daten.

Auditive Gestalten: Durch die Lernfähigkeit des Menschen Klänge an Objekte zu binden, ist es möglich aus einem Klangstrom heraus verschiedene dieser sogenannten *auditiven Gestalten* zu differenzieren und zu benennen.

Diese und weitere Eigenschaften werden genutzt um eine klangliche Darstellung von Datenströmen umzusetzen und eine Interpretation der Ereignisse zuzulassen. (vgl. [3])

Als umfangreiche Informationsquelle zur Sonifikation und ihren Anwendungen sei an dieser Stelle „*The sonification handbook*“ (T. Hermann, A. Hunt, J. G. Neuhoff (Hsg.), Logos Verlag, Berlin Germany, 2011) genannt, das frei zum Download¹ steht.

¹<http://sonification.de/handbook/download/TheSonificationHandbook-HermannHuntNeuhoff-2011.pdf> (zuletzt aufgerufen am 31.10.2013)

2.2 Tremor

In diesem Abschnitt wird ein kurzer Überblick über die untersuchten Tremorerkrankungen gegeben. Die medizinischen Informationen dazu entstammen den „*Leitlinien zur Diagnostik und Therapie in der Neurologie*“ (siehe [4]).

Definition: Tremor ist definiert als unwillkürlich rhythmische Oszillation eines oder mehrerer Körperteile. Zur klinischen Diagnose bezieht man sich unter anderem auf folgende Kriterien:

Aktivierungsbedingung - unter welchen Umständen tritt die Zitterbewegung auf.

- Rest (Ruheposition)
- Posture (Halteposition)
- Action (Bewegung)

Frequenz - mit welcher Frequenz tritt die Zitterbewegung auf.

- niederfrequent: 2 - 4 Hz
- mittelfrequent: 4 - 7 Hz
- hochfrequent: > 7 Hz

Amplitude - wie groß ist die Auslenkung der betroffenen Körperteile.

Dauer der Erkrankung

Familienanamnese - liegt eine vererbungsbedingte Erkrankung vor.

Reaktion auf Alkohol - ist nach Konsum von Alkohol eine Veränderung wahrnehmbar.

Drogenmissbrauch - gibt es äußere Faktoren, die einen Tremor verursachen.

Unter der grossen Anzahl an unterschiedlichen Tremores wurden im vorgestellten Projekt folgende Tremorformen zur Sonifikation herangezogen:

Parkinson: Das wichtigste Kriterium für den Parkinson-Tremor ist die Diagnose des Parkinsonsyndroms. Neben einer verlangsamten Bewegung und Steifigkeit der Muskulatur ist der *einseitige Ruhetremor* ein typisches Zeichen der Parkinson-Erkrankung.

Essentieller Tremor: Hierbei handelt es sich um die zweithäufigste Tremorform. Die Ursachen für den essentiellen Tremor sind oft nicht bekannt. Zumeist ist er erblich bedingt und stellt eine progrediente Erkrankung dar. Halte- und Aktionstremor sind hier vorherrschend, der *Halte-tremor* ist zumeist das dominierende Symptom.

Psychogener Tremor: Er entsteht aufgrund der Kontraktion antagonistischer (gegensätzlicher) Muskeln, kann plötzlich auftreten und spontan wieder verschwinden. Häufig wird er nach besonderen Stresssituationen beobachtet (zB. nach Unfällen, grosser Angst, seelischer Belastung...). Er stellt ein Mischbild zwischen den verschiedenen Tremorformen dar und kann sowohl in *Ruhe- und Halteposition als auch in Aktion* beobachtet werden. Als typisches Merkmal kann man die wechselnde Zitterfrequenz nennen.

3 Klinische Diagnoseverfahren

3.1 Klassische Diagnose

Grundsätzlich erfolgt die klinische Diagnose über die direkte Beobachtung der Patientinnen und Patienten und der betroffenen Körperteile durch geschulte NeurologInnen, häufig unter Zuhilfenahme von Videoanalysemethoden. Außerdem gibt es klinische Kriterien (siehe Abschnitt 2.2) und technische Hilfsmittel, wie die Accelerometrie (siehe unten), die zur Diagnosefindung beitragen. (vgl. [7])

3.2 Accelerometrie

Grundlage für die, in dieser Bachelorarbeit beschriebenen Sonifikationsmethoden, ist die sogenannte *Accelerometrie*, die bereits als Hilfsmittel zur Diagnosestellung im klinischen Betrieb eingesetzt wird. Dabei werden die Zitterbewegungen der betroffenen Körperteile mittels Beschleunigungssensor aufgenommen. Zusätzlich erfolgt über Hautelektroden die Aufnahme eines Elektromyogramms (EMG).

Applikation der Sensoren

Zunächst wird jeweils am zweiten Glied des linken und rechten Zeigefingers ein kleiner und leichter Beschleunigungssensor (Accelerometer) mit Klebeband befestigt. Zur Aufnahme eines oberflächlichen Elektromyogramms werden Elektroden an signifikanten Stellen des Arms befestigt. Für die Sonifikation der Tremordaten werden allerdings nur die Daten der Beschleunigungsaufnehmer verwendet daher wird hier auf die Beschreibung der Aufnahme des Elektromyogramms verzichtet.

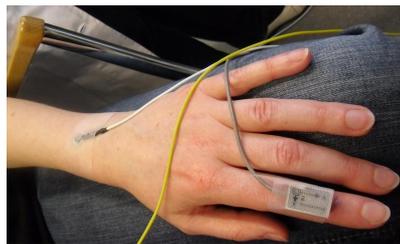
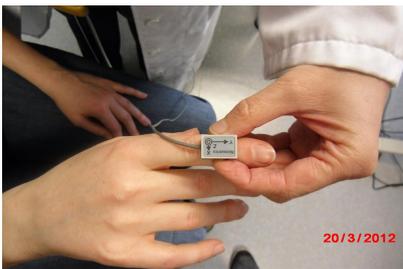


Abbildung 3.1: Befestigung der Sensoren an der Hand

Ablauf der Untersuchung

Grundsätzlich werden die verschiedenen Aktivierungsbedingungen durchlaufen. Dazu gibt es drei unterschiedliche Phasen:

„**Rest**“ - Ruheposition: Die Hände liegen entspannt am Oberschenkel auf.

„**Posture**“ - Halteposition: Die Arme werden ausgestreckt vor dem Körper gehalten.

„**Action**“ - Aktion: Der Zeigefinger wird abwechselnd an die eigene Nase und an einen

fixen Punkt geführt.

Jede Position wird für mindestens 30 Sekunden gehalten bzw. durchgeführt und die Daten, wie unter Abschnitt 6 beschrieben, aufgezeichnet.



Abbildung 3.2: Aktion

Klassische Auswertung der Daten

Zur Bestimmung üblicher Charakteristika der Tremores wird das Leistungsspektrum der aufgezeichneten Daten herangezogen.

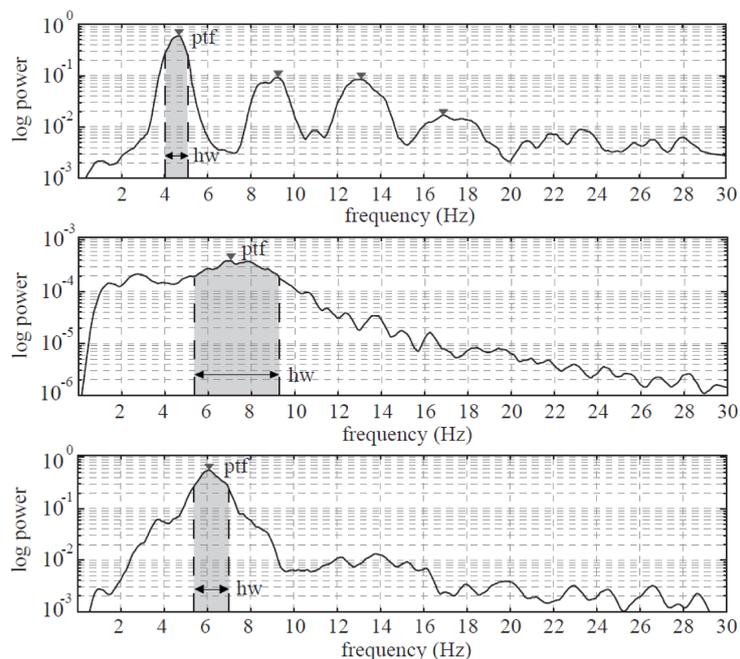


Abbildung 3.3: Leistungsspektren von Parkinson (oben), Essentiellem Tremor (Mitte) und Psychogenem Tremor (unten); ptf...peak tremor frequency; hw...half width [7]

Aus dem Leistungsspektrum werden Parameter ermittelt, die wichtige Hinweise auf den zugrundeliegenden Tremor liefern können.

Die Parameter sind wie folgt:

- Mittenfrequenz des größten „Peaks“ (*peak frequency*).
- Gesamtleistung des Spektrums im Bereich zwischen 1 Hz und 30 Hz.
- Halbwertsbreite - Breite die sich ergibt, wenn der Pegel, ausgehend von der Mittenfrequenz, um je 3dB gefallen ist.
- Leistung innerhalb der Halbwertsbreite (grau markierter Bereich in der Abbildung).

Daraus ergeben sich grundlegende Beobachtungen:

- Parkinson: nahezu harmonisches Spektrum.
- Essentieller Tremor: kein eindeutiger und sehr breiter „Peak“ im unteren Bereich des Spektrums.
- Psychogener Tremor: hauptsächlich ein eindeutiger „Peak“. (vgl. [7])

4 Probleme der klinischen Diagnose

4.1 Probleme der klassischen Diagnose

Die Nachteile der klassischen Diagnose ergeben sich aus den folgenden Gegebenheiten:

- **schlechte zeitliche Auflösungsfähigkeit des Auges**

Das Auge ist nicht in der Lage schnelle Bewegungsänderungen zu verfolgen. Weder Frequenz noch Amplitude der Zitterbewegung können mit bloßem Auge näher bestimmt werden. Das Erkennen von zugrundeliegenden Bewegungsmustern ist somit kaum möglich.

- **hoher Zeitaufwand bei der Auswertung von gesammelten Daten**

Daten aus Videoaufnahmen und der Accelerometrie können erst nach der Untersuchung der betroffenen Person ausgewertet werden. Dies erhöht den Zeitaufwand zur Diagnosefindung erheblich.

- **oft mehrere Sitzungen mit dem Patienten / der Patientin nötig**

Nachdem sich typische Krankheitsbilder erst über einen langen Beobachtungszeitraum herausstellen, sind mehrmalige Untersuchungen mit den TremorpatientInnen notwendig. Zudem kann die Wirkung von eingestellten Medikamenten oft erst nach einem längeren Zeitraum beurteilt werden.

- **schlechte Integration in den klinischen Alltag**

Gerade aus den oben genannten Nachteilen ergeben sich Schwierigkeiten wenn möglichst rasch eine richtige Diagnose gestellt werden soll. (vgl. [7])

4.2 Probleme der Accelerometrie

Die gewonnenen Parameter aus dem Leistungsspektrum enthalten nur schwache Hinweise auf die zugrundeliegende Erkrankung. Zur Gewinnung des Leistungsspektrums wird eine sogenannte *Fast Fourier Transformation (FFT)* durchgeführt, wodurch man Informationen über die, dem Signal innewohnenden Frequenzen und deren Amplitude erhält.

Bei der klinischen Analyse werden sehr große Zeitfenster (> 30 Sekunden) zur Bildung des Spektrums verwendet. Jeglicher zeitlicher Zusammenhang wird somit zerstört. Zudem werden die Leistungsspektren aus den drei Bewegungsrichtungen gemittelt wodurch jede Richtungsinformation verloren geht.

In den Sonifikationen versucht man nun diese wichtigen Indikatoren zu erhalten um die erfolgreiche Analyse zu erleichtern.

5 Sonifikation als Analysewerkzeug

Im Bezug auf das Projekt werden nun wichtige klangliche Eigenschaften vorgestellt, die maßgebend zur Unterscheidbarkeit und Erkennbarkeit der verschiedenen Tremorarten beitragen:

Tonhöhe: Die Frequenz erlaubt die Zuordnung des Gehörten zu einer wahrgenommenen Tonhöhe und somit zur Unterscheidung von tiefen, mittleren und hohen Tönen. Der Frequenzbereich in dem Klangereignisse als Tonhöhe wahrgenommen werden, erstreckt sich für das gesunde (junge) Ohr von 20 Hz bis 20 kHz. Durch Frequenzänderungen kann die Tonhöhe als gleichbleibend, höher oder tiefer eingestuft werden. Die Art und Weise der Frequenzänderung bzw. das Vorhandensein stabiler Frequenzen kann ein typisches Merkmal der zu analysierenden Krankheit darstellen.

Klang: Je nach Anzahl und Verhältnis der sogenannten Harmonischen, also Obertöne zu einem Grundton, wird ein Klang charakterisiert. Anregungsart und Einschwingverhalten spielen ebenfalls eine große Rolle bei der Differenzierung verschiedener Klangquellen. Die spezifischen Klänge ermöglichen es, diese an sogenannte auditive Gestalten zu koppeln, der Klang wird also erkannt und einer bekannten Klangquelle zugeordnet.

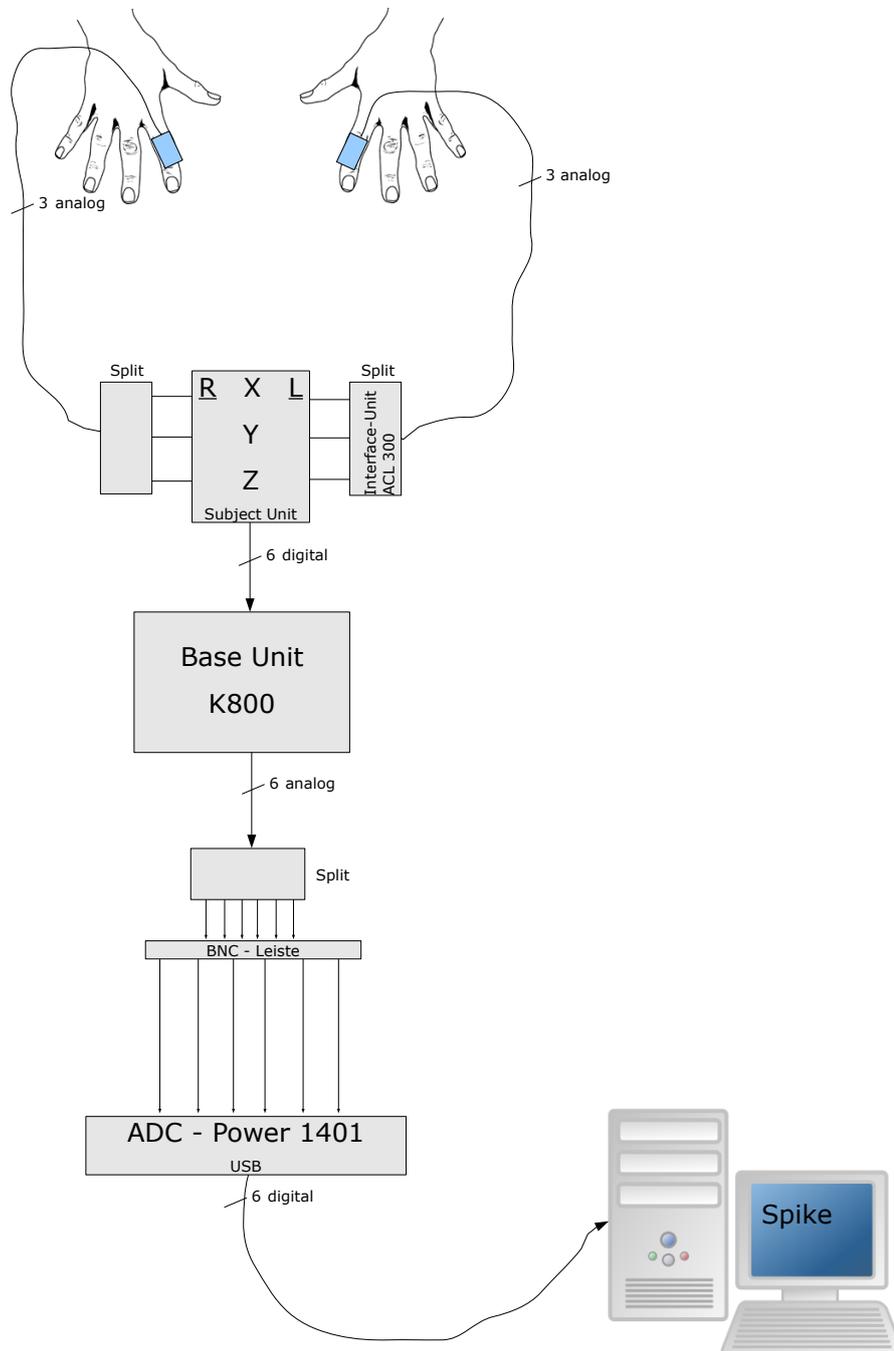
Mustererkennung: Durch die besonders hohe Zeitauflösung des Gehörs (bis hin zu wenigen Mikrosekunden) ist der Mensch in der Lage feinste Muster aus einem Klangstrom zu unterscheiden. Dabei ist es besonders sensibel auf Änderungen, die sich über die Zeit erstrecken. Diese Fähigkeit macht das Gehör zu einem wertvollen Werkzeug zur Überwachung und Analyse zeitvariabler Daten. (vgl. [3])

Lautstärke: Die Lautstärke als variabler Parameter ist schwer zu verfolgen und in der Sonifikation im Allgemeinen weniger gut geeignet. Amplitudenänderungen können jedoch als klangformendes Mittel eingesetzt werden wie dies bei der Amplitudenmodulation (siehe Abschnitt 7.3) der Fall ist.

Räumlichkeit: Zur einfacheren Unterscheidbarkeit verschiedener Signalquellen, können diese (vereinfacht) im Stereobild frei positioniert werden. In den hier präsentierten Sonifikationen werden die Informationen der rechten und linken Hand auf jeweils rechten und linken Kanal des Audioauspielwegs gelegt. Dies hat nun weniger den Sinn eine Räumlichkeit zu schaffen, als vielmehr die gleichzeitige Präsentation die Möglichkeit gibt, Synchronitäten der wiedergegebenen Daten zu erkennen. Daraus können wichtige Schlüsse auf das zugrundeliegende Krankheitsbild gezogen werden.

6 Messumgebung

6.1 Signalweg



6.2 Equipment

Das hier gezeigte Equipment kommt an der Universitätsklinik für Neurologie zur Anwendung und wird zur *Accelerometrie* (siehe Abschnitt 3.2) eingesetzt. Es enthält für die Forschung in diesem Bereich übliche Messgeräte und Anwenderprogramme.



Abbildung 6.1: Messequipment an der Universitätsklinik für Neurologie

Im folgenden Abschnitt werden die einzelnen Komponenten beschrieben:

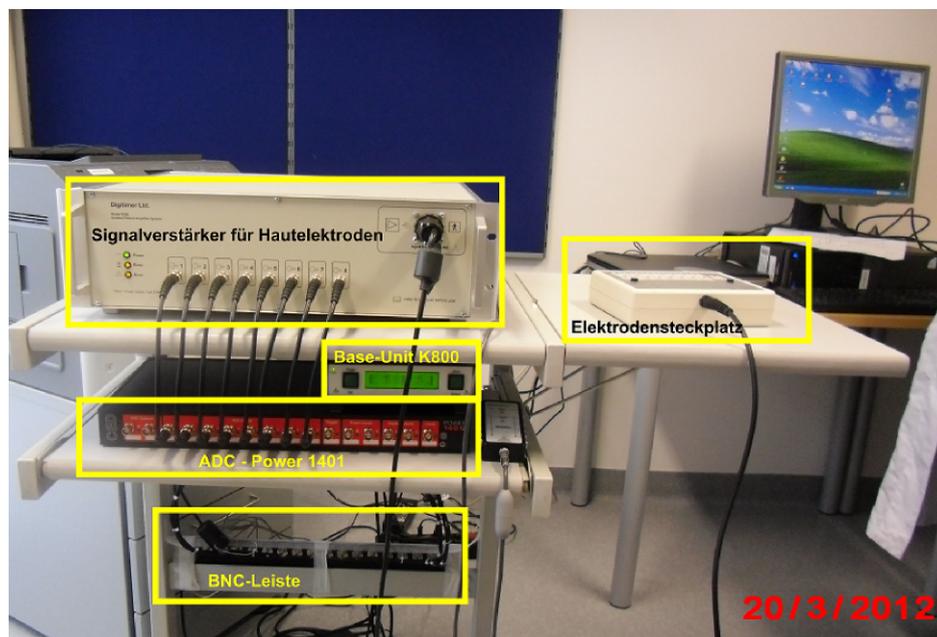


Abbildung 6.2: Beschreibung der Messgeräte

6.2.1 Beschleunigungssensor

Zur Erfassung der Zitterbewegungen der Hände kommt ein kleiner und leichter Präzisions-Beschleunigungsaufnehmer² zum Einsatz, der die Bewegungsabläufe in allen drei Raumrichtungen aufnimmt.

Zur Messung der Beschleunigung wird im Allgemeinen die Wirkung der Beschleunigung auf ein gedämpftes Feder-Masse-System gemessen. Die Wirkung der Beschleunigung ist die Verschiebung gegenüber dem Gehäuse einer, an das Gehäuse elastisch angekoppelten seismischen Masse. Über die Beziehungen $F = m \cdot a$ und $a(t) = \frac{dv(t)}{dt} = \frac{d^2x(t)}{dt^2}$ wird die Beschleunigung ermittelt, ihre Einheit ist $\frac{m}{s^2}$. Beschleunigungssensoren arbeiten üblicherweise nach kapazitiver, elektromagnetischer oder piezoelektrischer Wirkungsweise.

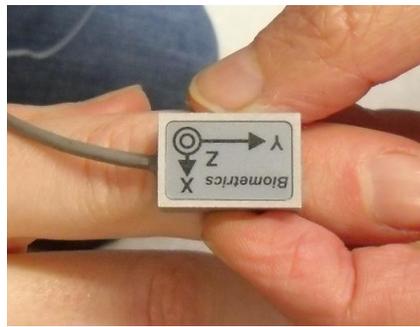


Abbildung 6.3: 3-Achsen Accelerometer

6.2.2 Interface Unit - Subject-Unit

Die Elektronik, die die Versorgung des aktiven Sensors, Daten-Buffer und einstellbare Filter enthält, ist in der ACL300-Interface-Unit untergebracht. Die sechs analogen Signale, die von den Beschleunigungssensoren beider Hände kommen, werden in der Subject-Unit zusammengeführt, digitalisiert und an die Base-Unit weitergeleitet. (vgl. [5])

²<http://www.biometricsltd.com/accelerometer.htm> (zuletzt aufgerufen am 31.10.2013)

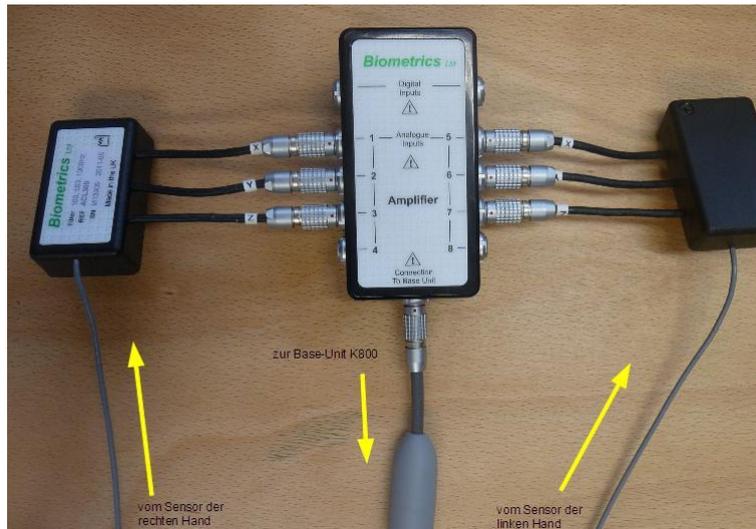


Abbildung 6.4: Subject-Unit

6.2.3 Base-Unit

Die Base-Unit K800³ ist ein Präzisions-Signalverstärker und erlaubt die Zusammenführung von Signalen verschiedener Sensortypen. Hier werden die von der Subject-Unit übertragenen Daten ins Analoge zurück konvertiert. Um den Beschleunigungssensor in eine definierte Ausgangslage (x-,y-,z-Ausrichtung) zu bringen erfolgt hier ein Nullabgleich zu Beginn jeder Sitzung, indem der Sensor auf einer ebenen Platte aufgelegt wird. (vgl. [5])



Abbildung 6.5: Base-Unit

Die sechs analogen Kanäle der Base-Unit werden an einer BNC-Leiste gesplittet und zur Analog-Digital-Wandlung an das Interface „Power 1401“ weitergeführt. Im geplanten Echtzeitsystem sollen die Daten an dieser BNC-Schnittstelle abgegriffen und in ein Audiointerface geführt werden.

³<http://www.biometricsltd.com/k800.htm> (zuletzt aufgerufen am 31.10.2013)

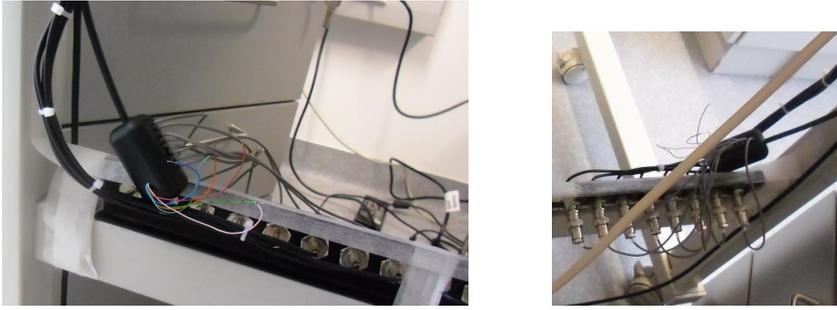


Abbildung 6.6: Split der analogen Kanäle an der BNC-Leiste

6.2.4 Datenerfassung - Power 1401 mkII

Die Datenerfassung erfolgt mit dem Hochleistungs-Interface Power 1401 mkII⁴, das häufig in klinischen Labors zum Einsatz kommt. Es erlaubt schnelle Datenerfassung und -analyse für Multitasking-Echtzeitanwendungen. Das Interface enthält einen Prozessor und On-Board-Memory um schnell Daten zu erfassen, unabhängig von unterschiedlichen Abtastraten und komplexer On-line-Analyse. Der Host-Computer wird somit entlastet und kann andere Rechnertätigkeiten ausführen, wie zB. Datenverarbeitung und weitere -analyse. (vgl. [12])

Die 6 analogen Kanäle, von der BNC-Leiste kommend, werden an der Rückseite in den ADC geleitet. An der Vorderseite kommen die verstärkten Signale der Hautelektroden an. Die Daten werden mit einer Abtastrate von 1.000 Hz und einer Auflösung von 16 bit abgetastet und mittels USB an den PC weitergeleitet. Die digitalisierten Daten der Beschleunigungssensoren und der Hautelektroden werden letztendlich mit der Software *Spike* aufgezeichnet.



Abbildung 6.7: Power 1401 in der Signalkette

⁴<http://www.science-products.com/Products/CatalogC/Interfaces/Power1401/Power1401.html> (zuletzt aufgerufen am 31.10.2013)

6.2.5 Software - Spike

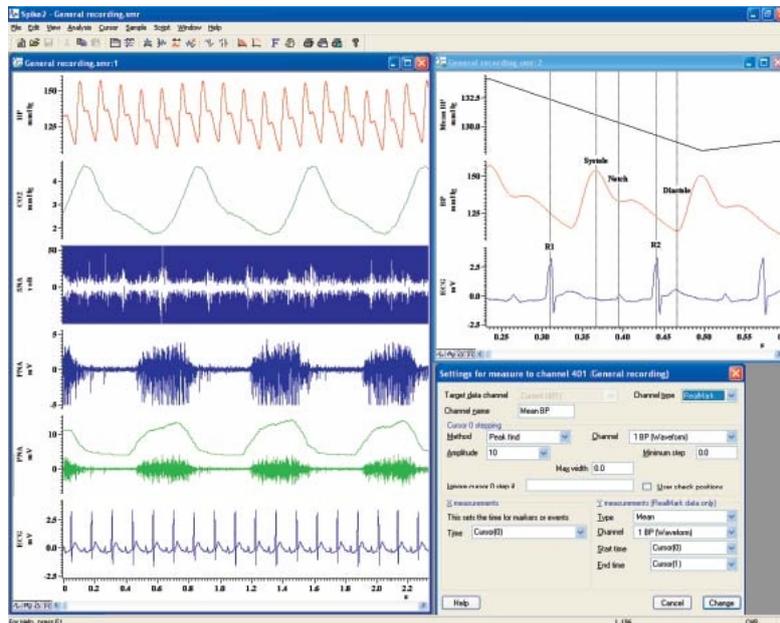


Abbildung 6.8: Datenerfassung mittels Spike [13]

Die Software *Spike*⁵ ist ein unter Forschungseinrichtungen weit verbreitetes Datenerfassungs- und -verarbeitungsprogramm. Es bietet zahlreiche Darstellungs- und Analysemöglichkeiten der aufgezeichneten Daten wie zB. Wellenformdarstellung, setzen von Markern und Triggerpunkten, Sonogrammdarstellung, Ermittlung des Leistungsspektrums, Korrelationsfunktionen uvm. (vgl. [13])

⁵<http://www.science-products.com/Products/CatalogC/Acq&AnaSoftware/Spike2/Spike2.html> (zuletzt aufgerufen am 31.10.2013)

7 Datenaufbereitung

Die Signale, für jede Raumrichtung (x, y, z) und für beide Hände liegen nach der Datenerfassung mit einer Abtastrate von 1.000 Hz und einer Auflösung von 16 bit vor. *Spike* erzeugt daraus eine sogenannte .smr-Datei.

Upsampling und Filterung

Eine MATLAB[®]⁶-Routine erzeugt aus den .smr-Dateien .wav-Dateien, wobei ein Upsampling auf die Audiodatenrate von 44.1 kHz durchgeführt wird. Mittels DC-Filter (Hochpass mit 1 Hz Grenzfrequenz) werden zunächst Gleichanteile gefiltert, ein Tiefpassfilter erster Ordnung mit 70 Hz Grenzfrequenz beschränkt den relevanten Frequenzbereich. Da die Tremorsignale Frequenzen zwischen 2 Hz und 15 Hz aufweisen, ist eine Grenzfrequenz von 70 Hz ausreichend hoch gewählt.

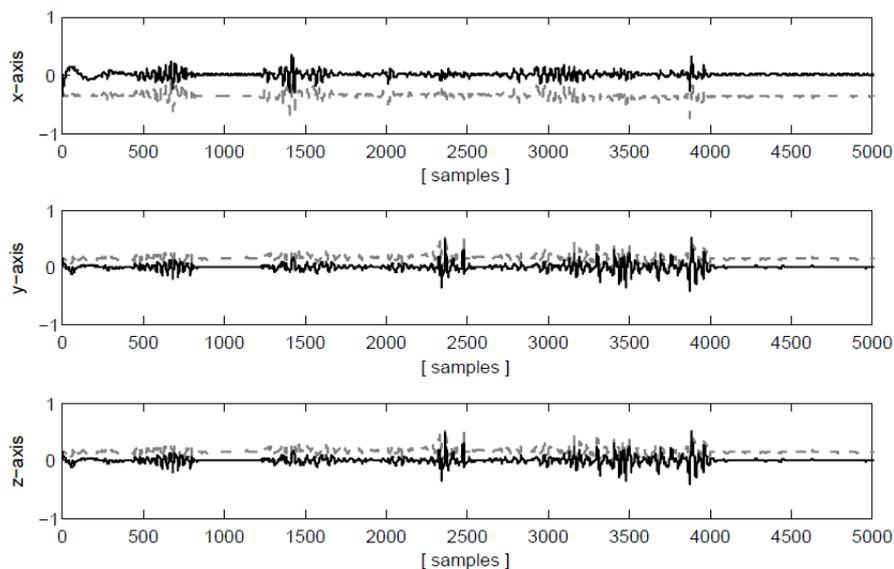


Abbildung 7.1: Sensorsignal vor (grau) und nach (schwarz) der DC-Filterung [17]

Der Umweg der Aufzeichnung der Daten über *Spike* wurde in der Testphase des Projekts verfolgt. In Zukunft wird diese Methode entfallen und ein Audiointerface wird die Daten an der BNC-Schnittstelle abgreifen (siehe Abschnitt 6.2.3) und in der Audiodatenrate entsprechend aufbereiten.

Ausgangsmaterial: Für alle Sonifikationen liegen sechs (2 x 3) Bewegungssignale in der Audiodatenrate von 44.1 kHz vor.

⁶„MATLAB[®] ist eine höhere Programmiersprache und interaktive Umgebung für numerische Berechnungen, Visualisierung und Programmierung. MATLAB[®] dient zur Datenanalyse, Algorithmenentwicklung und zur Erstellung von Modellen und Anwendungen.“ [16]

7.1 Automatic Gain Control

Die Dynamik der aufgezeichneten Tremorsignale variiert stark zwischen verschiedenen Messungen. Aus Erfahrung mit den Testsignalen, kann man Pegelunterschiede von $40dB$ bis $60dB$ erwarten.

Um alle Messsignale in einen ähnlichen Dynamikbereich zu transferieren wird die *Automatic Gain Control* durchgeführt.

Für jedes Eingangssignal eines Datensets wird mittels sogenanntem „Peak-Follower“ laufend der Spitzenwert bestimmt.

$$peak(x[n], t_c) = \begin{cases} x[n] & \text{für } x[n] > x[n-1] \\ peak(x[n-1], t_c)e^{(-1.0/t_c)} & \text{sonst} \end{cases} \quad (7.1)$$

Das Signal mit dem größten Pegel wird bestimmt und dient als Bezugssignal. Alle anderen Eingangssignale des Datensets werden nun über den Momentanspitzenwert des Bezugssignals normalisiert. Die Relationen der Tremorsignalamplituden zueinander bleiben somit erhalten, was ein wichtiges Kriterium für die folgende *principal component analysis* darstellt.

$$x_{agc}[n] = AGC \{x[n]\} = \frac{x[n]}{peak(x[n], t_c)^p} \quad (7.2)$$

Die Funktion $AGC \{x[n]\}$ beschreibt einen Kompressor. Der Wert des Exponenten p ist variabel einstellbar im Bereich 0,1 bis 1. Die Zeitkonstante t_c gibt die Zeit in Sekunden an, in der der Ausgangswert sich auf bis $-60dB$ an den Eingangswert genähert hat. Mit $p = 1$ werden alle Dynamikunterschiede zwischen den unterschiedlichen Tremorsignalen entfernt. Dies ist nicht wünschenswert, da es in der Diagnose durchaus von Interesse sein kann, wie sich die Pegelwerte unterschiedlicher Tremores verhalten. Der Wert ist daher mit $p = 0.8$ fix eingestellt um eine angemessene Kompression zu erhalten ohne dabei die signifikanten Dynamikunterschiede der einzelnen Tremores zu verlieren.

7.2 Principal Component Analysis (PCA)

Die *principal component analysis* ist eine Methode aus der multivariaten Statistik und erlaubt die Auswertung hochdimensionaler Datensätze. Sie besitzt besondere mathematische Eigenschaften, die es ermöglichen herkömmliche lineare Modelle zu berechnen, selbst wenn Variablen untereinander korrelieren.

Eines der größten Probleme von multivariaten Daten ist, dass sie nur schwer darstellbar bzw. fassbar sind. Dies führt oft dazu, dass eventuell bestehende Zusammenhänge nicht erkannt werden können.

Der zentrale Gedanke der *principal component analysis* ist nun, die Daten in ihren Dimensionen zu verringern und sie auf zB. eine zweidimensionale Ebene (hier „plane of movement“) oder einen eindimensionalen Vektor zu projizieren.

Zur Verbildlichung des Prozesses seien hier die grundlegenden Schritte der PCA kurz beschrieben:

1. Richtung der maximalen Streuung finden. Diese Richtung definiert die Richtung der ersten Achse.
2. Weitere Achse erzeugen. Diese steht normal auf die erste und wird so lange um die erste Achse rotiert, bis die Streuung entlang der neuen Achse maximal ist.
3. Dritte Achse hinzufügen. Diese steht wiederum normal auf die beiden anderen und zeigt in Richtung der verbleibenden Maximalstreuung.

Dieser Prozess resultiert in einer Rotation des Koordinatensystems in einer Art und Weise, dass die Achsen ein Maximum an Streuung entlang ihrer Richtung aufweisen. Diese Darstellung kann mathematisch als sogenanntes Eigenwertproblem aufgefasst werden. Die Eigenvektoren der Kovarianzmatrix bilden die Hauptkomponenten (*principal components*). Die korrespondierenden Eigenwerte weisen darauf hin, wie viel Information in den einzelnen Komponenten enthalten ist. (vgl. [14])

Der zugrundeliegende mathematische Prozess ist unter Abschnitt A beschrieben.

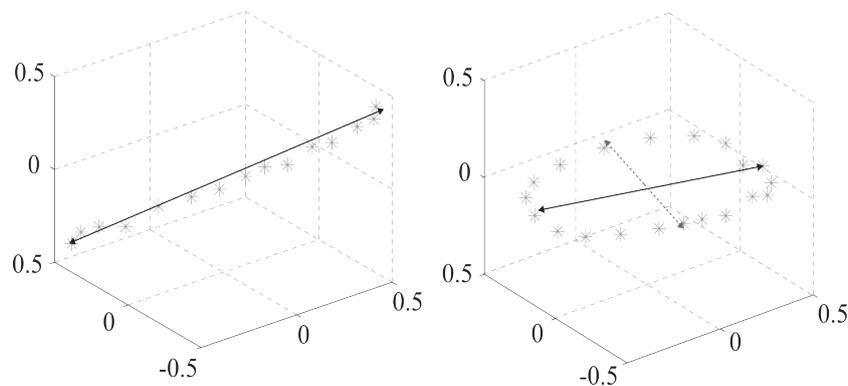


Abbildung 7.2: a) Richtung der maximalen Streuung einer rein translatorischen Bewegung, b) Richtungen der maximalen Streuung einer rotatorischen Bewegung [8]

Anwendung in der Sonifikation

Die PCA wird in den Sonifikationen in zweierlei Hinsicht angewendet:

1. **zur Reduktion der Dimensionen:** die dreidimensionalen Daten werden auf einen eindimensionalen Vektor projiziert. Das Ergebnis dieser Projektion wird in weiterer Folge als Eingangssignal für die Sonifikationen verwendet.

2. **zur Trennung in Rotation und Translation:** nach Auffinden der Hauptkomponenten (*principal components*) ist eine Mustererkennung möglich, die eine Unterscheidung in translatorische und rotatorische Bewegungsanteile erlaubt. Details dazu findet man unter A.2. Diese Art der Analyse der Bewegungsmuster ist neuartig und kann wichtige Informationen zu den Krankheitsbildern beisteuern. (vgl. [7])

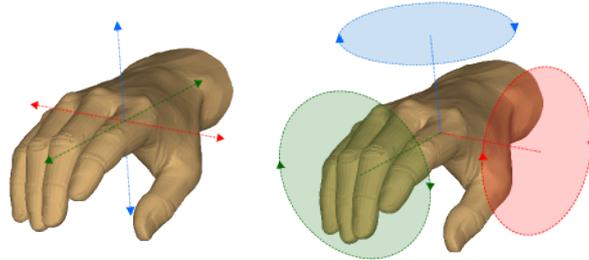


Abbildung 7.3: Translation und Rotation [8]

7.3 Amplitudenmodulation

Bei allen Sonifikationsumsetzungen wird am Ende der Signalstrecke eine Amplitudenmodulation mit der halbwellengleichgerichteten Version des Eingangssignals durchgeführt. Je nach Sonifikation ist das Eingangssignal aber ein anderes: bei „Vocoder“ (siehe Abschnitt 8.2.1) ist es beispielsweise das eindimensionale, aus der PCA stammende Signal $PCA_1[n]$, bei „Birds“ (siehe Abschnitt 8.2.2) ist es die jeweilige IMF. Bei der Amplitudenmodulation verändert sich die Amplitude des Sonifikationssignals (Trägersignal) proportional zu den Momentanamplitudenwerten des niederfrequenten Tremorsignals (Modulator).

Die Idee hinter dieser Form der Amplitudenmodulation ist, dass man auch direkt eine Information über das ursprüngliche Tremorsignal erhält. Nachdem die typischen Tremorfrequenzen aus Bewegungen entstehen, sind sie sehr tieffrequent und für das Ohr nicht hörbar. Mit der Amplitudenmodulation gelingt es, die Information der Bewegung der umgesetzten Sonifikation aufzuprägen und als Rhythmus wahrnehmbar zu machen.

Halbwellengleichrichtung

Die Halbwellengleichrichtung wird folgendermaßen beschrieben:

$$HW \{x(n)\} = \begin{cases} x[n] & \text{für } x[n] > 0 \\ 0 & \text{sonst} \end{cases} \quad (7.3)$$

Bei der Halbwellengleichrichtung werden negative Signalanteile verworfen. Der Rhythmus der Bewegung kommt somit deutlicher zur Geltung. Bei rein translatorischen Bewegungen entspräche das zB. nur der Abbewegung der Hände.

Glättung

Nachdem das halbwellengleichgerichtete Signal sehr steile Flanken aufweist, kommt es zu abrupten Signalunterbrechungen, die im Klang als „hart“ wahrgenommen werden. Zur Glättung wird eine Integration über das halbwellengleichgerichtete Signal durchgeführt, die für steigende und fallende Flanken je eine eigene Anstiegs- bzw. Abklingkonstante vorsieht. Es kommt zu einem „Verschmieren“ der Wellenberge. In der Benutzeroberfläche (siehe Abschnitt 9) ist der exponentielle Faktor und somit der Grad der „Härte“ einstellbar.

7.4 Frequenzmodulation

Neben der Amplitudenmodulation, welche fixer Bestandteil aller Sonifikationen ist, besteht die Möglichkeit eine Frequenzmodulation wahlweise dazuzuschalten (siehe Kapitel 9).

Bei der Frequenzmodulation verändert sich die Frequenz des Sonifikationssignals (Trägersignal) mit der Momentanfrequenz des niederfrequenten Tremorsignals (Modulator). Das heißt nun, dass die Information über die eigentliche Bewegung in der Frequenzänderung und somit Tonhöhenänderung steckt.

$$Son_{FM} \{x[n]\} = \sin \{ (f + kHW(AGC(x[n], 1))) 2\pi n \} \quad (7.4)$$

Als Modulator dient das aus der *Automatic Gain Control* stammende, komprimierte Signal $AGC \{x[n], p\}$. In diesem Fall wählt man $p=1$ um stets ein Signal zu erhalten, das auf den Bereich -1 und 1 begrenzt und dessen Maximalamplitude auf 1 normiert ist. Durch die Verwendung des halbwellengleichgerichteten Eingangssignals gibt es längere stille Phasen, wodurch sich die Tonhöhenänderungen besser verfolgen lassen.

7.5 Chorus

„Die in der Vorverarbeitungsstufe extrahierten Informationen über den rotatorischen und translatorischen Anteil des Tremorsignals (siehe Abschnitt A.2) können optional als zusätzlicher Parameter in die hier vorgestellten Sonifikationen integriert werden. Dieser Parameter steuert einen sogenannten „Chorus Effekt“ - also eine leicht zueinander verzögerte, doppelte Wiedergabe des Sonifikationssignals. Bei Signalen, die keinerlei rotatorischen Anteil aufweisen, ist dieser Effekt nicht hörbar, mit steigendem rotatorischem Anteil wird die Amplitude des verdoppelten Signals entsprechend erhöht bis bei Signalen mit rein rotatorischem Anteil ein deutlicher „Chorus“ zu hören ist.“ [1]

8 Sonifikationen

8.1 Erste Umsetzungen

In diesem Abschnitt werden zwei Sonifikationsmethoden beschrieben, die in der Anfangsphase des Projekts entstanden sind. Hier werden hauptsächlich die gewonnenen Informationen auf Basis der Leistungsspektren der Accelerometrie (siehe Abschnitt 3.2) verarbeitet.

Um die Echtzeitanwendung und die zeitlichen Zusammenhänge zu gewährleisten wird hier das Spektrum mittels FFT im „*sliding-window*“-Verfahren mit einer Fensterlänge von 1 Sekunde gewonnen. Auf diese Weise kann der Verlauf des Spektrums zeitlich verfolgt werden, die Auflösung ist jedoch durch die Fensterlänge begrenzt. (vgl. [7])

8.1.1 Karplus-Strong-Synthese

Die *Karplus-Strong-Synthese*⁷ beruht auf der Nachbildung eines physikalischen Modells zur Erzeugung von Gitarrenklängen bzw. von Klängen gezupfter Saiten.

Die Grundfrequenz des Systems ist hier proportional zur Mittenfrequenz des „Hauptpeaks“ gewählt. Das Anregungssignal ist eine Mischung aus Sägezahnsignal mit der selben Frequenz und rosa Rauschen. Die Halbwertsbreite des „Hauptpeaks“ wird genutzt um die Mischung des Anregungssignals und den Rückkopplungsfaktor der „Delay-Line“ zu kontrollieren.

Bei schmalen „Peaks“ steht der Klang des Sägezahnsignals im Vordergrund und der Rückkopplungsfaktor ist nahezu 1. Je breiter der „Peak“ wird, desto mehr Rauschen wird in das Anregungssignal gemischt und der Klang des Sägezahnsignals tritt in den Hintergrund. Der Rückkopplungsfaktor wird zu 0 wenn keine „Peaks“ im Spektrum vorhanden sind. Danach wird das Ausgangssignal Tiefpass gefiltert. Werden neben dem „Hauptpeak“ Harmonische detektiert, wird die Grenzfrequenz dieses Filters auf das achtfache der Grundfrequenz gesetzt, gibt es keine Harmonischen im Spektrum bleibt die Grundfrequenz erhalten. (vgl. [7])

Klangeindruck

Weist das Spektrum einen oder mehrere „Peaks“ auf, ist deutlich ein Ton zu hören. Beim Parkinson-Tremor, mit seinem harmonischen Spektrum ist der Ton für längere Zeit stabil wahrnehmbar. Beim psychogenen Tremor hingegen mischt sich der Sägezahnklang stärker dazu, außerdem bewegt sich die Grundfrequenz. Der essentielle Tremor liefert ein sehr rauschhaftes Klangbild in dem nur für kurze Augenblicke Töne wahrnehmbar sind.

⁷Nähere Informationen zur Karplus-Strong-Synthese sind unter „*Digital synthesis of plucked string and drum timbres*“ (K. Karplus, A. Strong, *Computer Music Journal*, vol. 7, no.2, pp. 43-55, 1983) zu finden.

8.1.2 Spectral Shaping

Das vorliegende Spektrum des Tremorsignals (0 Hz bis 50 Hz) dient bei dieser Sonifikationsmethode als „Filter“ und wird dabei mit dem Spektrum von rosa Rauschen (gesamter Hörbereich) multipliziert. Dabei ist der resultierende Klang abhängig von der Struktur des Tremorspektrums. Die Betragswerte des laufenden Tremorspektrums im Bereich von 0 bis 50 Hz „formen“ somit das Spektrum des Rosa Rauschens. Nach erfolgter Rücktransformation (*IFFT*) wird das Signal in den Zeitbereich gebracht und somit hörbar. (vgl. [7])

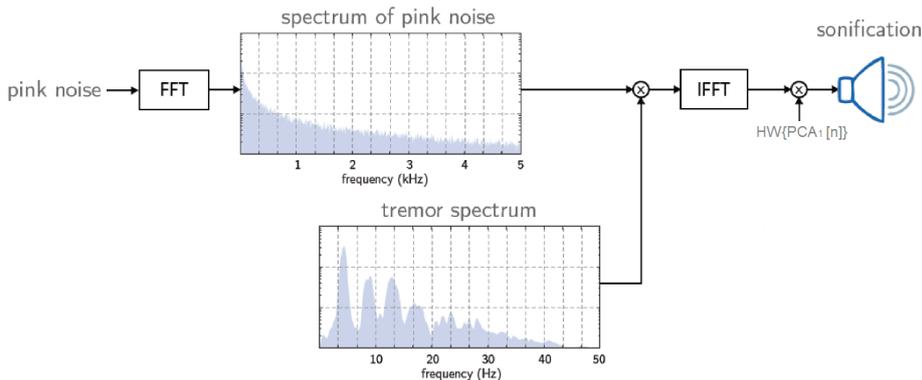


Abbildung 8.1: Flussdiagramm zur Sonifikation „*Spectral Shaping*“ [8]

Es gibt verschiedene Parameter, die eine Angleichung an die persönlichen Abhörgegewohnheiten erlaubt. Diese sind:

Offset: Das Tremorspektrum kann entlang der Frequenzachse so verschoben werden, dass es in einem angenehmen Abhörfrequenzbereich zu liegen kommt.

Stretch: Das Tremorspektrum kann entlang der Frequenzachse gestaucht oder gedehnt werden und wird somit auf unterschiedliche Art und Weise dem Rosa Spektrum aufgeprägt.

Scaling: Die Betragswerte des Spektrums können ebenfalls variiert werden.

Exponent: Der Kontrast zwischen den „Peaks“ kann variiert werden, so dass bei Bedarf der Verlauf des „Hauptpeaks“ einfacher verfolgt werden kann. Je höher der Exponent gewählt wird, desto stärker werden „Nebenpeaks“ unterdrückt und desto schmaler und spitzer bzw. „schärfer“ wird der „Hauptpeak“ abgebildet.

Klangeindruck

Der Klang der Sonifikation „*Spectral Shaping*“ ist durch die Wahl der verschiedenen Parameter sehr wandelbar. Grundsätzlich ergibt sich systembedingt ein sehr rauschhaftes Klangbild. Die Filterwirkung des Tremorspektrums auf den Klang ist für geübte Ohren gut nachvollziehbar.

Abhängig vom Tremorsignal können durch die Amplitudenmodulation am Ende der Signalstrecke Klangbeschreibungen gefunden werden wie zB. „Dampflokomotive“, „Regenprasseln“ und „Folien-Knistern“. Durch Variation des Parameters „Exponent“ lassen sich die rauschhaften Anteile unterdrücken und das Klangbild wird tonaler. Zur Klangvorstellung finden sich hier Beispiele wie „Grillenzirpen“ und „Fröschequaken“. Mit Verschiebung des Spektrums entlang der Frequenzachse („Offset“) kann man Klangbilder von „U-Bahn“ bis „Tafelquietschen“ realisieren.

8.2 Finale Umsetzungen

Im Laufe des Projekts wurden „Workshops“ mit den beteiligten ÄrztInnen durchgeführt, wobei die entwickelten Sonifikationen vorgestellt wurden. Nach gemeinsamen Auswertungen über Informationsgehalt und Klangeindruck wurde eine Vorauswahl getroffen. Im folgenden Abschnitt werden jene beiden Sonifikationen präsentiert, die im Zuge dieses Auswahlverfahrens als Diagnosewerkzeuge ausgewählt wurden.

8.2.1 „VOCODER“

Das Basismaterial für die Sonifikation sind die Bewegungsdaten, die aus der Accelerometrie vorliegen und nach den in Abschnitt 7 beschriebenen Methoden aufbereitet wurden.

Bei der Sonifikationsmethode „Vocoder“ werden aus dem Spektrum der Tremorsignale Informationen gewonnen, die als Parameter für die Sonifikation dienen. Die Idee dahinter ist, das Spektrum in mehrere Bänder zu unterteilen und die in den einzelnen Bändern enthaltene Energie auszuwerten.

Die folgende Aufzählung beschreibt schrittweise die zugrunde liegende Signalverarbeitung zur Umsetzung der Sonifikation „Vocoder“.

1) Automatic Gain Control

Bei der Umsetzung der Sonifikationsmethode „Vocoder“ wird zu Beginn die *Automatic Gain Control* (siehe Abschnitt 7.1) durchgeführt um die Dynamik der Signale aus unterschiedlichen Messungen in einen ähnlichen Dynamikbereich zu transferieren. Dies ist sinnvoll, da bei der Aufnahme von Daten die Pegel der Signale von Messung zu Messung bzw. von PatientIn zu PatientIn stark variieren. Damit die Pegelunterschiede aber erkennbar bleiben, da sie wichtige Informationen über die Stärke des Tremors liefern, wird ein Exponent $p < 1$ gewählt.

2) Downsampling

In Hinblick auf die später folgende *FFT* werden die Daten auf 1.024 Hz dezimiert. Es wird also eine Datenratenreduzierung vorgenommen.

3) PCA

Die *PCA* (siehe Abschnitt 7.2 und Appendix A) wird in diesem Fall durchgeführt um die drei Dimensionen der Sensordaten einer Hand auf eine zu reduzieren. Das Ausgangssignal der PCA dient als Eingangssignal für die Sonifikation und wird als $PCA_1[n]$ (siehe Gleichung A.4) bezeichnet.

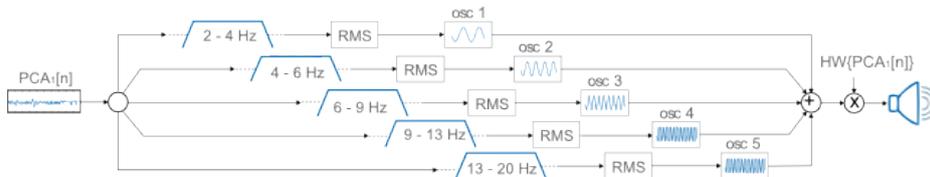


Abbildung 8.2: Flussdiagramm zur Sonifikation „Vocoder“ [8]

4) FFT

Um das Spektrum des Signals $PCA_1[n]$ zu erhalten, wird eine *Fast Fourier Transformation (FFT)* im „*sliding-window*“-Verfahren durchgeführt. Die Fensterfunktion hat eine Größe von 4.096-Punkten, wobei 1.024-Punkte ein „*Hanning-Window*“ darstellen und der Rest (3.072-Punkte) mittels „*zero-padding*“ ergänzt wird. In Abständen von 64 Samples (das entspricht $64/1024 = 62.5ms$) wird ein neues FFT-Spektrum ermittelt. Durch diese Wahl der Fensterform (langes „*zero-padding*“) erhält man ein „interpoliertes“, also geglättetes Spektrum. Würde man darauf verzichten, wären die „Peaks“ im Spektrum sehr spitz und schmal und eine Unterscheidung vor allem zwischen Parkinson und Psychogenem Tremor würde erschwert. Das Spektrum wird zusätzlich in der Benutzeroberfläche (siehe Abschnitt 9) visuell dargestellt.

5) Teilung in Frequenzbänder

Zur weiteren Analyse wird das Spektrum in fünf unterschiedlich breite Bänder unterteilt:

1. 2 - 4 Hz $\hat{=}$ bins 8 - 16
2. 4 - 6 Hz $\hat{=}$ bins 16 - 24
3. 6 - 9 Hz $\hat{=}$ bins 24 - 36
4. 9 - 13 Hz $\hat{=}$ bins 36 - 52
5. 13 - 20 Hz $\hat{=}$ bins 52 - 80

Mit der FFT-Fensterlänge von 4.096 Punkten erhält man für jeden FFT-Bin eine Breite von $1024/4096 = 0,25Hz$. Jeweils ein Bin eines Bandes überlappt mit dem Bin des darauffolgenden Bandes. Die Bänder in den tieferen Frequenzen sind schmal gehalten (2 Hz), da man in diesen Frequenzbereichen die meiste Energie erwartet. In den höheren Frequenzbereichen wurde die Bandbreite weiter gewählt, weil dort nur schwache Überhöhungen im Spektrum zu erwarten sind. Die Wahl der Mittenfrequenzen und der Bandbreite ergab sich aus den Erfahrungen mit den Spektren der unterschiedlichen Tremorarten, die im Laufe des Projekts genau analysiert wurden. Eine Variation der Parameter ist bei Bedarf möglich.

6) Energieverteilung in den Bändern

Die Parameter, welche letztendlich für die Sonifikation ausschlaggebend sind, werden im folgenden Schritt ermittelt.

$$s_{a,b}[n] = \left(\sum_{i=(a \cdot 4)-1}^{(b \cdot 4)+1} FFT_{PCA_1}[i] \right) + \frac{1}{2} \cdot FFT_{PCA_1}[a \cdot 4] + \frac{1}{2} \cdot FFT_{PCA_1}[b \cdot 4] \quad (8.1)$$

Es werden die Betragswerte des Spektrums in den einzelnen Bändern summiert, wobei a und b den Beginn und das Ende des entsprechenden Frequenzbereichs darstellen.

Mit $a \cdot 4$ und $b \cdot 4$ wird der dazugehörige Frequenzbin ermittelt. Da jeweils ein Bin des einen Bandes im darauffolgenden Band enthalten ist, gehen die halbierten Werte der überlappenden Bins der einzelnen Bänder in das Ergebnis ein.

7) Normierung der Energiewerte

Die maximale Energie eines Bandes wird zur Normierung der jeweils kleineren Energiewerte der anderen Bänder verwendet.

Bestimmen der Maximalenergie eines Bandes:

$$s_{a,b}^{max}[n] = \max(s_{2,4}[n], s_{4,6}[n], s_{6,9}[n], s_{9,13}[n], s_{13,20}[n]) \quad (8.2)$$

Normalisieren:

$$s_{a,b}^{norm}[n] = \frac{s_{a,b}[n]}{s_{a,b}^{max}[n]} \quad (8.3)$$

Man erhält somit Werte zwischen 0 und 1, die in weiterer Folge als Kontrollparameter verwendet werden.

8) Expansion

Nach der Normalisierung besteht die Möglichkeit diese Werte weiter zu modifizieren.

$$s_{a,b}^e[n] = \text{pow}(s_{a,b}^{norm}[n], p) \quad (8.4)$$

Durch Variation des Exponenten p lässt sich der Kontrast zwischen den Signalen der einzelnen Bänder verändern. Mit $p < 1$ erreicht man eine Kompression, mit $p > 1$ (*expansion*) besteht die Möglichkeit das energiereichste Band mit $s_{a,b}^{norm} = 1$ von allen anderen Bändern $s_{a,b}^{norm} < 1$ abzuheben und somit dessen Verlauf deutlicher verfolgen zu können.

9) Steuerung von Oszillatoren

Die Energiewerte aus dem vorhergehenden Schritt werden nun verwendet um die Ampli-

tude von fünf Sinusoszillatoren, die harmonisch zueinander gestimmt sind, zu steuern:

$$s_{voc}[n] = s_{2,4}^e[n]osc(f_1)[n] + s_{4,6}^e[n]osc(f_2)[n] + s_{6,9}^e[n]osc(f_3)[n] + s_{9,13}^e[n]osc(f_4)[n] + s_{13,20}^e[n]osc(f_5)[n] \quad (8.5)$$

Die Frequenzen der einzelnen Oszillatoren können beispielsweise folgendermaßen eingestellt werden:

- Oszillator 1: $f_1 = 200Hz$
- Oszillator 2: $f_2 = 400Hz$
- Oszillator 3: $f_3 = 600Hz$
- Oszillator 4: $f_4 = 800Hz$
- Oszillator 5: $f_5 = 1.000Hz$

Andere „Stimmungen“ können beliebig gewählt werden.

10) Amplitudenmodulation

Die Summe der fünf Oszillatorensignale wird nun mit dem halbwellengleichgerichteten Eingangssignal $HW \{PCA_1[n]\}$ amplitudenmoduliert:

$$SON_{voc}[n] = s_{voc}[n] \cdot HW \{PCA_1[n]\} \quad (8.6)$$

Klangeindruck

Der Klang der Sonifikation „Vocoder“ ist schwer mit bekannten Klängen zu vergleichen. Der Name „Vocoder“ stammt aus der Entwicklungsphase, in der mit Formantsynthese experimentiert wurde. Durch Einsatz der Sinusoszillatoren ist der Vergleich mit einem üblichen „Vocoder“ eher schwierig. Grundsätzlich lässt sich der Klang als sehr tonhaft beschreiben. Dieser Eindruck verstärkt sich je nach Anzahl der Harmonischen. Durch die Amplitudenmodulation wird der Eindruck des Rhythmus gewonnen.

Klangeigenschaften

Allgemein lassen sich die Klangeindrücke zu den einzelnen Tremorformen folgendermaßen beschreiben:

Parkinson:

- Ruhetremor: Typisches Kennzeichen ist, dass der Tremor verstärkt in der Ruheposition auftritt.
- stabiler Rhythmus: Deutliches Merkmal des Parkinson-Tremors ist der sehr regelmäßige und stabile Rhythmus.

- stabile Frequenz: Der Klang weist eine sehr stabile (oft tiefe) Frequenz auf, die kaum Schwankungen unterliegt.
- Harmonische: Zumeist sind die harmonischen Frequenzen deutlich herauszuhören. Es ergibt sich ein „vollerer“ Klang.
- Stabile Lautstärke: Das Tremorsignal wird überwiegend als gleichmäßig laut empfunden.
- Asynchronizität: Der Parkinson-Tremor differenziert sich von den anderen dadurch, dass nur ein Arm bzw. einer stärker als der andere, von den Zitterbewegungen betroffen ist.

Essentieller Tremor:

- Haltetremor: Der essentielle Tremor macht sich hauptsächlich in der Halteposition bemerkbar. Die beschriebenen Klangeigenschaften beziehen sich daher auf die Halteposition. In Ruhe ist er selten vorhanden und ist dort meist als sehr hochfrequentes, feines „Klirren“ wahrnehmbar.
- gestörter Rhythmus: Die Textur des Rhythmus' beim essentiellen Tremor ist grundsätzlich stabil. Sie weist oft eine Ähnlichkeit mit dem „Parkinson-Rhythmus“ auf. Man könnte sie, im Vergleich zum Parkinson, allerdings als „gestörter“ beschreiben.
- schwankende Frequenz: Die hörbare Frequenz des Hauptpeaks ist weniger stabil und schwankt leicht um die Mittenfrequenz.
- Harmonische: Die Anzahl an Harmonischen ist weniger stark ausgeprägt als beim Parkinson. Ein oder zwei Harmonische lassen sich oft ausmachen, die allerdings stärkeren Schwankungen unterliegen.
- Synchronizität: In der Regel sind beide Arme, mehr oder weniger, gleich betroffen.

Psychogener Tremor:

Die Klangeigenschaften des psychogenen Tremors sind schwierig zu beschreiben, da er eine Mischform der bekannten Tremorformen darstellt. Es können daher Eigenschaften auftreten die dem Parkinson- als auch dem essentiellen Tremor innewohnen. Als Hilfestellung dient daher oft der Vergleich zwischen den verschiedenen Aktivierungspositionen und der Synchronizität der Zitterbewegungen der Gliedmaßen.

- Mischform: Beim psychogenen Tremor lässt sich nicht pauschal feststellen in welcher Position die Zitterbewegung überwiegt.

- gestörter Rhythmus: Häufig ergibt sich ein sehr stabiler Rhythmus der an den Rhythmus des Parkinson-Tremors erinnert. Feine Schwankungen, die sich in Form von „Stolpern“ ergeben, können bei der Identifizierung hilfreich sein. Andererseits weist er auch oft eine stabile Textur, wie beim essentiellen Tremor auf.
- schwankende Frequenz: Die Frequenzschwankungen zeigen sich hier oft deutlich. Allerdings ergeben sich auch sehr stabile Frequenzen, die dem Parkinson-Tremor ähneln.
- Harmonische: Harmonische können sich zeigen, unterliegen aber häufig starken Schwankungen.
- Synchronizität: In der Regel sind beide Arme mehr oder weniger gleich betroffen.

8.2.2 „BIRDS“

Bei der Sonifikation „Birds“ wird im Gegensatz zu den bisher vorgestellten Sonifikationsmethoden ein anderer Ansatz gewählt. Hier werden, ohne Umweg über das Spektrum, direkt aus dem Tremorsignal Informationen zur weiteren Sonifikation gewonnen. Die Grundlage dazu liefert die sogenannte *Empirical Mode Decomposition*, die es ermöglicht, die dominanten, dem Signal innewohnenden Moden zu extrahieren.

Das Basismaterial sind auch hier die Bewegungsdaten, die aus der Accelerometrie vorliegen und nach den in Abschnitt 7 beschriebenen Methoden aufbereitet wurden.

Die folgende Aufzählung beschreibt schrittweise die zugrundeliegende Signalverarbeitung zur Umsetzung der Sonifikation „Birds“.

1) PCA

Aus der PCA (siehe Abschnitt A) wird direkt das eindimensionale Signal $PCA_1[n]$ gewonnen.

2) EMD

Mittels *Empirical Mode Decomposition* (EMD) werden die ersten fünf *intrinsic mode functions* (IMFs) aus $PCA_1[n]$ ermittelt (siehe Abschnitt B). Die *intrinsic mode functions* $IMF_2[n]$ bis $IMF_5[n]$ werden in weiterer Folge als Eingangsdaten für die Sonifikation verwendet. $IMF_1[n]$ ist jene mit der höchsten Frequenz und bildet eher Rauschen ab als eine relevante Information über den Tremor und wird daher verworfen. Je höher die Zahl der IMFs ist, desto tieffrequenter sind die darin enthaltenen Frequenzkomponenten.

3) Automatic Gain Control

Die *Automatic Gain Control* (siehe Abschnitt 7.1) wird nun auf die verbleibenden vier *IMFs* angewendet und die Dynamik verschiedener Eingangssignale somit eingengt.

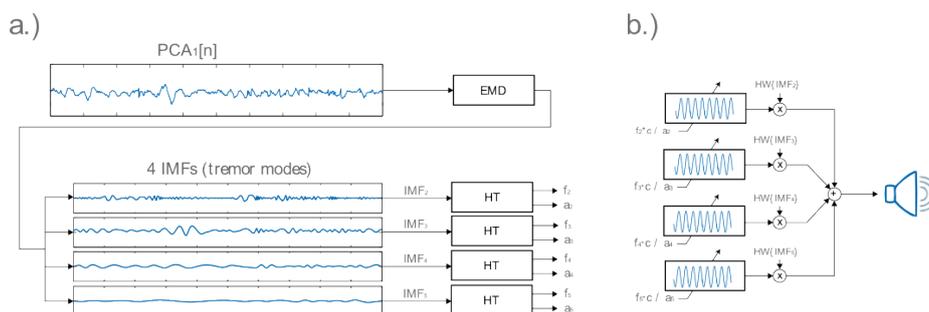


Abbildung 8.3: a) Gewinnung der Momentanamplitude und -frequenz aus den IMFs, b) Ansteuern von Sinusoszillatoren mit modifizierten IMFs [8]

4) Hilbert Transformation

Mittels Hilbert-Transformation (siehe Abschnitt B) wird für jede IMF zu jedem Zeitpunkt die Momentanamplitude $a_i[n]$ berechnet. Die Momentanfrequenz $f_i[n]$ wird mittels *gene-*

*ralized zero-crossing*⁸ ermittelt, da sich die Berechnung aus der Hilbert-Transformation als zu instabil erwiesen hat (siehe auch Abschnitt B). Die Frequenzen $f_i[n]$ werden mit einem konstanten Faktor c multipliziert, um die sehr kleinen Tremorfrequenzen (ca. 2Hz - 15Hz) in den Hörbereich zu transferieren. Mit dieser modifizierten Frequenz schwingt je ein Sinusoszillator, die Lautstärke jedes einzelnen Oszillators ist von der Momentanamplitude abhängig (siehe Abbildung 8.3b).

$$s_{bird_i}[n] = a_i \cdot osc(f_i[n]) \quad (8.7)$$

5) Amplitudenmodulation

Die Ausgangssignale der Oszillatoren werden wieder mit dem halbwellengleichgerichteten Eingangssignal, in diesem Fall mit der jeweiligen IMF, amplitudenmoduliert. Die Summe der vier Oszillatorsignale stellt das hörbare Ausgangssignal der Sonifikation dar.

$$Son_{bird}[n] = \sum_{i=2}^5 s_{bird_i}[n] \cdot HW \{IMF_i[n]\} \quad (8.8)$$

Klangeindruck

Durch die Ansteuerung von Sinusoszillatoren ergibt sich für jede IMF ein Klang der an Vogelgesang erinnert. Die Stimmlage jedes einzelnen Vogels ist von der Momentanfrequenz und die Lautstärke von der Momentanamplitude der dazugehörigen IMF abhängig. Der Eindruck der verschiedenen Geschwindigkeit, in der ein Vogel singt, wird von der Amplitudenmodulation bestimmt.

Klangeigenschaften

Allgemein lassen sich die Klangeindrücke zu den einzelnen Tremorformen folgendermaßen beschreiben:

Parkinson:

- Ruhetremor: Typisches Kennzeichen ist, dass der Tremor hauptsächlich in der Ruheposition auftritt.
- stabiler Rhythmus: Deutliches Merkmal des Parkinson-Tremors ist der sehr regelmäßige und stabile Rhythmus - das entspricht dem „Singen“ in gleichbleibendem Tempo.
- stabile Frequenz: Der Klang weist eine sehr stabile (oft tiefe) Frequenz auf, die kaum Schwankungen unterliegt - das entspricht einem „Singen“ auf gleicher Tonhöhe bzw. einem monotonem „Vogelgespräch“.

⁸Für nähere Informationen siehe „On instantaneous frequency“ (N. E. Huang, Z. Wu, S. R. Long, K. C. Arnold, X. Chen, K. Blank, *Advances in Adaptive Data Analysis*, Vol. 1, No. 2, 2009)

- eine dominante Mode: beim Parkinson tritt meist eine Mode (IMF) in den Vordergrund - es singt hauptsächlich ein „Vogel“.
- Asynchronizität: Der Parkinson-Tremor differenziert sich von den anderen dadurch, dass nur ein Arm bzw. einer stärker als der andere, von den Zitterbewegungen betroffen ist.

Essentieller Tremor:

- Haltetremor: Der essentielle Tremor macht sich hauptsächlich in der Halteposition bemerkbar. In Ruhe ist er selten vorhanden.
- gestörter Rhythmus: Die Textur des Rhythmus' beim essentiellen Tremor ist grundsätzlich stabil. Sie weist oft eine Ähnlichkeit mit dem Parkinson-„Rhythmus“ auf. Man könnte sie, im Vergleich zum Parkinson, allerdings als „gestörter“ bzw. als sich „eifrig“ unterhaltende Vögel beschreiben.
- Die hörbare Frequenz des „Hauptpeaks“ ist weniger stabil und schwankt leicht um die Mittenfrequenz - dieses Merkmal zeigt sich in der Sonifikation als „Singen“ in unterschiedlicher Tonhöhe bzw. einem „Vogelgespräch mit Nachdruck“.
- mehrere Moden: häufig steht eine dominante Mode (ein Vogel) im Vordergrund, die aber stärker als beim Parkinson von anderen „Vögeln“ begleitet wird.
- Synchronizität: In der Regel sind beide Arme ähnlich stark betroffen.

Psychogener Tremor:

Die Klangeigenschaften des psychogenen Tremors sind schwierig zu beschreiben, da er eine Mischform der bekannten Tremorformen darstellt. Es können daher Eigenschaften auftreten die dem Parkinson- als auch dem essentiellen Tremor innewohnen. Als Hilfestellung dient daher oft der Vergleich zwischen den verschiedenen Aktivierungspositionen und der Synchronizität der Zitterbewegungen der Gliedmaßen.

- Mischform: Beim psychogenen Tremor lässt sich nicht pauschal feststellen in welcher Position die Zitterbewegung überwiegt.
- gestörter Rhythmus: Häufig ergibt sich ein sehr stabiler Rhythmus der an den Rhythmus des Parkinson-Tremors erinnert. Andererseits weist er auch oft eine stabile Textur, wie beim essentiellen Tremor auf. Das „Wegbrechen“ der Textur und die darauffolgende Wiederholung des selben Musters können bei der Identifizierung hilfreich sein.
- schwankende Frequenz: Die Frequenzschwankungen zeigen sich hier oft deutlich. Allerdings ergeben sich auch sehr stabile Frequenzen, die dem Parkinson-Tremor ähneln.

- mehrere Moden: oft stellt sich eine Mode (ein Vogel) in den Vordergrund, die aber von zahlreichen anderen „Vögeln“ begleitet wird.
- Synchronizität: In der Regel sind beide Arme mehr oder weniger gleich betroffen.

Die Beschreibung der Klangeindrücke ist sehr von der persönlichen Wahrnehmung abhängig und wurde hier nach eigenem Empfinden bestmöglich dargestellt. Um sich einen eigenen Eindruck von den Klängen machen zu können, empfiehlt es sich die Klangbeispiele auf der dazugehörigen Projektwebpage⁹ anzuhören.

⁹<http://iem.at/?id=13661> (zuletzt aufgerufen am 31.10.2013)

9 Benutzeroberfläche

Die in PD¹⁰ umgesetzte Benutzeroberfläche der zur Wahl stehenden Sonifikationen erlaubt die Auswahl verschiedener Voreinstellungen, die Anpassung der Parameter an die persönlichen Abhörgehnheiten und die Orientierung anhand von Visualisierungen.

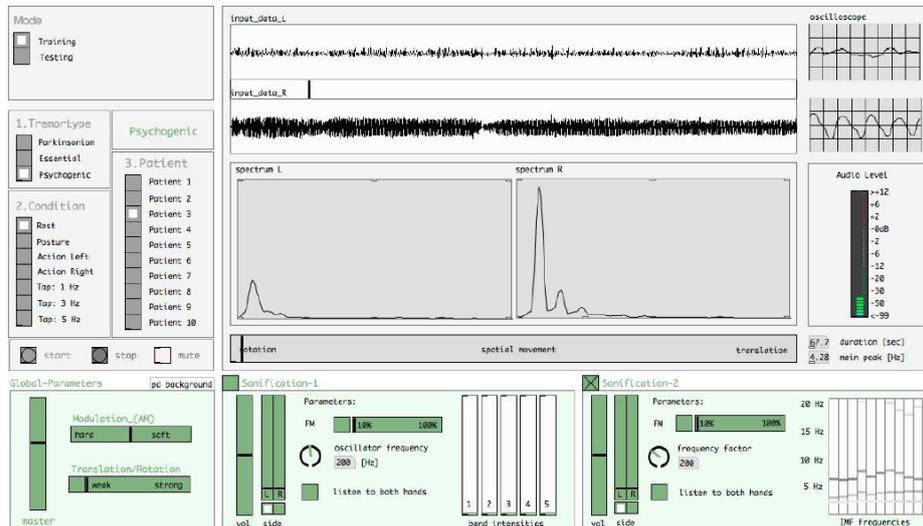


Abbildung 9.1: Benutzeroberfläche

Folgende Visualisierungen stehen zur Verfügung:

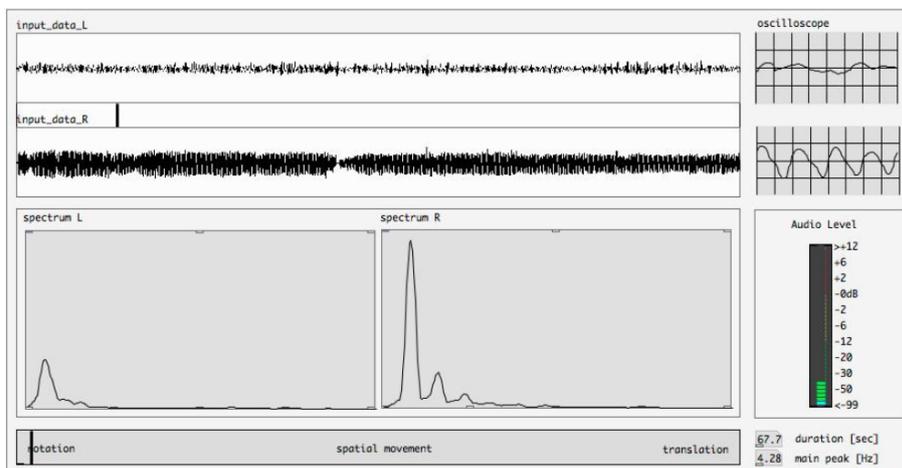


Abbildung 9.2: Grafische Darstellungsfelder

Wellenform: Darstellung der Wellenform der Eingangsdaten der linken (*input_data_l*) und der rechten (*input_data_r*) Hand. (links oben)

Oszilloskop: Vergrößerte Ansicht der Wellenform, die gerade durchlaufen wird. (rechts oben)

Audio Level: Überwachung der Lautstärkepegel. (rechts Mitte)

¹⁰Pure Data ist eine grafische Programmierumgebung für Audio-, Video- und Bildverarbeitung, die Echtzeitanwendungen erlaubt. (vgl. [15])

Spektrum: Darstellung des laufenden Spektrums der linken und rechten Hand. (Mitte)

Rotation - Translation: Grafische Darstellung des rotatorischen und translatorischen Anteils im Signal. (unten)

Duration: Gesamtlänge des Signals in Sekunden.(rechts unten)

Main Peak: Wert der momentanen Mittenfrequenz des Hauptpeaks. (rechts unten)

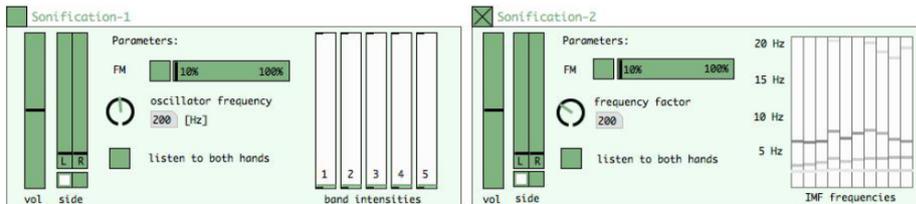


Abbildung 9.3: Bedienoberflächen der beiden Sonifikationen

Die Sonifikationsfelder erlauben das Umschalten zwischen Sonifikation 1 („Vocoder“) und Sonifikation 2 („Birds“). In beiden Feldern kann man zusätzlich zwischen linker und rechter Hand (**L/R**) umschalten oder mit **listen to both hands** die Daten beider Hände gleichzeitig hören.

Der Volumeregler (**Vol**) erlaubt die Lautstärkekontrolle der abgespielten Daten. Mit **FM** (siehe Abschnitt 7.4) hat man die Möglichkeit die Frequenzmodulation in beliebigen Anteilen dazuzuschalten. Mit **oscillator frequency** kann man die Grundfrequenz des untersten Bandes beliebig verschieben um eine Anpassung der Tonhöhe an die eigenen Hörgewohnheiten durchführen zu können. Diese Auswahl steht nur bei der Sonifikation 1 zur Verfügung. Zusätzlich gibt es eine Visualisierung zum Energieinhalt der einzelnen Frequenzbänder. Je größer die Energie in einem Frequenzband ist, desto höher steigt ein Balken in der entsprechenden Säule.

In der Sonifikation 2 kann man mit **frequency factor** auch hier eine Verschiebung der Tonhöhe vornehmen. In der Grafik sieht man in den „Zeitspalten“ wandernde Balken in unterschiedlichem Farbkontrast. Jede IMF ist durch einen einfärbigen Balken dargestellt. Je mehr unterschiedliche Balken es gibt (max. 4), desto mehr wandernde Querbalken befinden sich in der Grafik. Das Wandern der Balken in der Horizontalen entspricht der Änderung der Momentanfrequenz (vereinfachtes Hilbertspektrum) der einzelnen IMF.

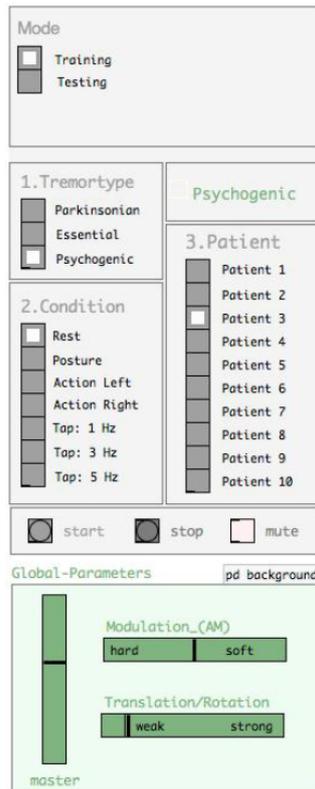


Abbildung 9.4: Globale Bedienelemente

Mode: Auswahl zwischen der Übungsumgebung (*Training*) oder der Testumgebung (*Testing*)

Tremortype: Auswahl zwischen den Tremorformen

Condition: Auswahl der Aktivierungsbedingungen

Patient: Im Trainingsmodus können hier die Daten der unterschiedlichen Patienten ausgewählt werden.

Start - Stop - Mute: Abspielen - Stoppen - Stummschalten der Audiodaten.

Unter **Global Parameters** kann man für beide Sonifikationen folgende Parameter beeinflussen:

Master: Globaler Lautstärkereglern

Modulation (AM): Einstellmöglichkeit der Abklingzeit zum Glätten des Halbwellengleichgerichteten Eingangssignals. Je kürzer die Abklingzeit, desto „härter“ (**hard**) klingt der Rhythmus, je länger desto „weicher“ (**soft**).

Translation - Rotation: Hinzuschalten des Chorus-Effekts und Regelung des Anteils.

10 Ausblick

Die Evaluierung¹¹ der verschiedenen Sonifikationen hat gezeigt, dass eine Unterscheidung der Tremorsignale durch hörendes Analysieren gut möglich ist und eine Zuordnung der Krankheitsformen zu den entsprechenden Tremores getroffen werden kann.

In Zukunft sollen die Daten des Beschleunigungssensors in ein Audio-Interface geführt werden, wobei die Daten direkt in der Audiodatenrate von 44.1 kHz abgetastet werden und ohne Umwege über Formatwandlung und Upsampling (siehe Abschnitt 7) von der Software weiterverarbeitet werden.

Eine Integration dieses Systems in den klinischen Bereich zur Erweiterung der bestehenden Diagnostikmittel kann ausschlaggebend für eine effiziente und richtige Diagnose sein. Die Auswertung der Information kann im Beisein des Patienten erfolgen und ersetzt womöglich die zeitaufwendige Auswertung der Video- und Spektraldaten im Nachhinein.

¹¹Zur Auswertung des Hörversuchs ist ein internes Dokument vorhanden, siehe „*Auswertung Hörversuch „Tremor hören“- Oktober 2012*“ (A. Wankhammer)

A Principal Component Analysis (PCA)

A.1 PCA zur Dimensionsreduzierung

Daten

Die Messdaten aus der Accelerometrie (siehe Abschnitt 3) liegen dreidimensional (x-, y- und z-Achse) vor und können als Spaltenvektor folgendermaßen angeschrieben werden:

$$a_n = \begin{pmatrix} a_{x_n} \\ a_{y_n} \\ a_{z_n} \end{pmatrix}$$

Kovarianzmatrix

Die Kovarianzmatrix für dreidimensionale Daten hat im allgemeinen folgendes Aussehen:

$$\Sigma = \begin{pmatrix} \text{cov}(x,x) & \text{cov}(x,y) & \text{cov}(x,z) \\ \text{cov}(y,x) & \text{cov}(y,y) & \text{cov}(y,z) \\ \text{cov}(z,x) & \text{cov}(z,y) & \text{cov}(z,z) \end{pmatrix} \quad (\text{A.1})$$

Für die Echtzeitanwendung werden zunächst die ersten 10 Samples ($N_c = 10$) des Eingangssignals verwendet um die Kovarianzmatrix Σ_1 zu bilden.

$$\Sigma_1 = (a_1, a_2, \dots, a_{N_c}) \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_{N_c}^T \end{pmatrix} \quad (\text{A.2})$$

Jedes weitere Sample wird mit einem α , das einen exponentiellen Abklingfaktor darstellt, gewichtet und in die laufende Kovarianzmatrix übernommen. Verallgemeinert ergibt sich:

$$\Sigma_n = (1 - \alpha)\Sigma_{n-1} + \alpha(a_n a_n^T) \quad (\text{A.3})$$

Für $\alpha \ll 1$ ergeben sich sehr stabile und sich langsam ändernde Kovarianzmatrizen. Es hat sich gezeigt, dass sich der Wert $\alpha = 0.05$ (das entspricht 0.5 Sekunden) gut eignet um die Änderungen im Signal angemessen zu übernehmen.

Eigenwerte und Eigenvektoren

Nach einem Diagonalisierungsprozess werden daraufhin laufend die Eigenvektoren $[\gamma_{1n}, \gamma_{2n}, \gamma_{3n}]$ und Eigenwerte $[\lambda_{1n}, \lambda_{2n}, \lambda_{3n}]$ von Σ_n ermittelt.

Principal Components

Die hier gefundenen Eigenvektoren sind die sogenannten „*principal components*“, ge-
reicht nach dem Eigenvektor mit dem größten Eigenwert.

Reduzierung der Dimension

Mit $\tilde{a}_{1_{PCA}}[n] = \gamma_1^T a_n$ erhält man die Projektion der dreidimensionalen Eingangssignale auf den Vektor mit der größten Varianz. Dieses $\tilde{a}_{1_{PCA}}[n]$ dient in den Sonifikationen als Eingangssignal und wird dort $PCA_1[n]$ bezeichnet.

$$\tilde{a}_{1_{PCA}}[n] = PCA_1[n] \quad (A.4)$$

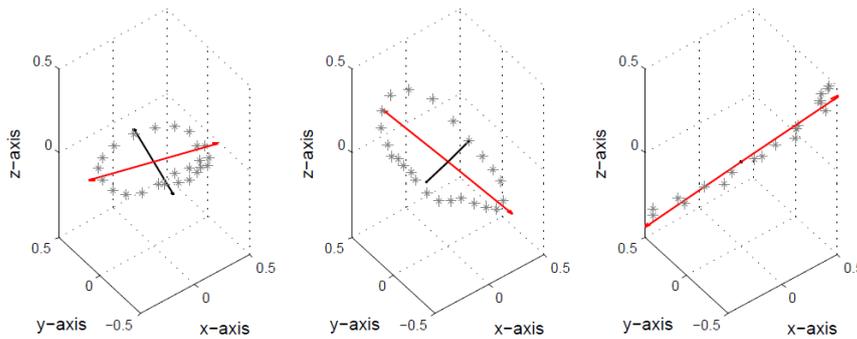


Abbildung A.1: Erste (rot) und zweite (schwarz) „principal component“ für unterschiedliche Eingangssignale: a) rotatorisch, b) translatorisch-rotatorisch, c) translatorisch [17]

(vgl. [6] und [17])

A.2 PCA zur Bestimmung von rotatorischen und translatorischen Signalanteilen

Die beiden „principal components“ γ_1 und γ_2 werden verwendet um die dreidimensionalen Eingangsdaten $a[n] = [x[n], y[n], z[n]]^T$ auf die sogenannte „plane of movement“ zu projizieren.

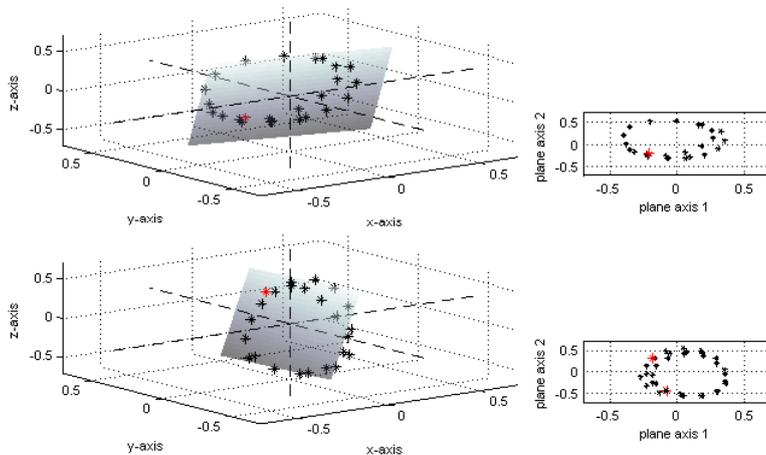


Abbildung A.2: „Planes of Movement“ [17]

Die resultierenden Eingangsdaten ergeben sich aus:

$$\tilde{a}_n = \begin{bmatrix} \tilde{a}_{1_{PCA}}[n] \\ \tilde{a}_{2_{PCA}}[n] \end{bmatrix} = [\gamma_{1n}, \gamma_{2n}]^T a_n$$

Die beiden Vektoren $\tilde{a}_{1_{PCA}}[n]$ und $\tilde{a}_{2_{PCA}}[n]$ stellen eine zweidimensionale Trajektorie dar, die wichtige Informationen über die rotatorischen und translatorischen Signalanteile der Tremorsignale gibt. Ein rein translatorisches Signal wird eine linienförmige Datenverteilung haben (siehe Abbildung A.1c), wo sich nur Vorzeichen und Betragswert des Vektors ändern. Ein rotatorisches Signal hingegen wird eine kreisförmige Datenverteilung haben (siehe Abbildung A.1a), wo sich hauptsächlich die Richtung der Vektoren ändert.

Aus diesen Beobachtungen heraus kann nun eine direkte Trennung von Translations- und Rotationssignal vorgenommen werden. Die erste „principal component“ kann also aus Rotation und Translation entstehen, wohingegen die zweite „principal component“ nur durch Anteile einer Rotationsbewegung gebildet werden kann.

$\tilde{a}_{2_{PCA}}[n] = r[n]$ stellt nun direkt den Anteil der Rotationsbewegung dar, nach Entfernen des rotatorischen Einflusses aus $\tilde{a}_{1_{PCA}}[n]$ wird der Anteil der Translationsbewegung $t[n]$ gebildet. (vgl. [7] und [17])

B Empirical Mode Decomposition (EMD)

B.1 Einführung

Die *Empirical Mode Decomposition* wurde von Huang et al. entwickelt. Wie in „*The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis*“ (siehe [9]) beschrieben, stellt sie eine Analysemethode für nicht-stationäre und nicht-lineare Signale dar. Komplizierte Datensätze können mit Hilfe der EMD in eine endliche (oft geringe) Anzahl sogenannter „intrinsic mode functions“, kurz IMFs genannt, zerlegt werden. Jede einzelne IMF repräsentiert eine, dem Signal innewohnende, Schwingungsmode.

Im Gegensatz zur Fourier-Analyse, wo das Signal in ein Set vorgefertigter Basisfunktionen zerlegt wird, werden bei der EMD die Basisfunktionen adaptiv aus dem Signal gewonnen. Durch Summierung der erhaltenen IMFs und dem resultierenden Restsignal, ist eine Rekonstruktion des Originalsignals verlustlos möglich.

Durch Einsatz der Hilbert-Transformation, im Zusammenhang mit der EMD als *Hilbert-Huang-Transformation* bezeichnet, kann für jede einzelne IMF die Momentanfrequenz und -amplitude bestimmt werden.

Vorteile der EMD:

- Rekonstruktion des Originalsignals ohne Verluste möglich
- Kein Vorwissen über die Qualität des Signals nötig (stationär, nicht stationär ...)
- Abbildung der „physikalischen Welt“
- Momentanattribute durch Hilbert-Transformation bestimmbar

Die Vorteile der Hilbert-Huang-Transformation gegenüber anderer Transformationsmethoden, haben die EMD zu einem mächtigen Analysewerkzeug gemacht.

Zum Überblick einige (wenige) Anwendungsbeispiele:

- Klimadatenentwicklung
- Analyse von Kurszeitreihen in der Finanzmathematik
- Erdbebenanalyse
- Vorzeitige Schadenserkennung bei Gebäuden und Maschinen
- Analysewerkzeug in der Medizin
- uvm.

B.2 Algorithmus

Folgende Grafik gibt einen allgemeinen Überblick über den, der EMD zugrundeliegenden Algorithmus. Die Beschreibung ist an „*Empirical Mode Decomposition - An Introduction*“ (siehe [10]) angelehnt:

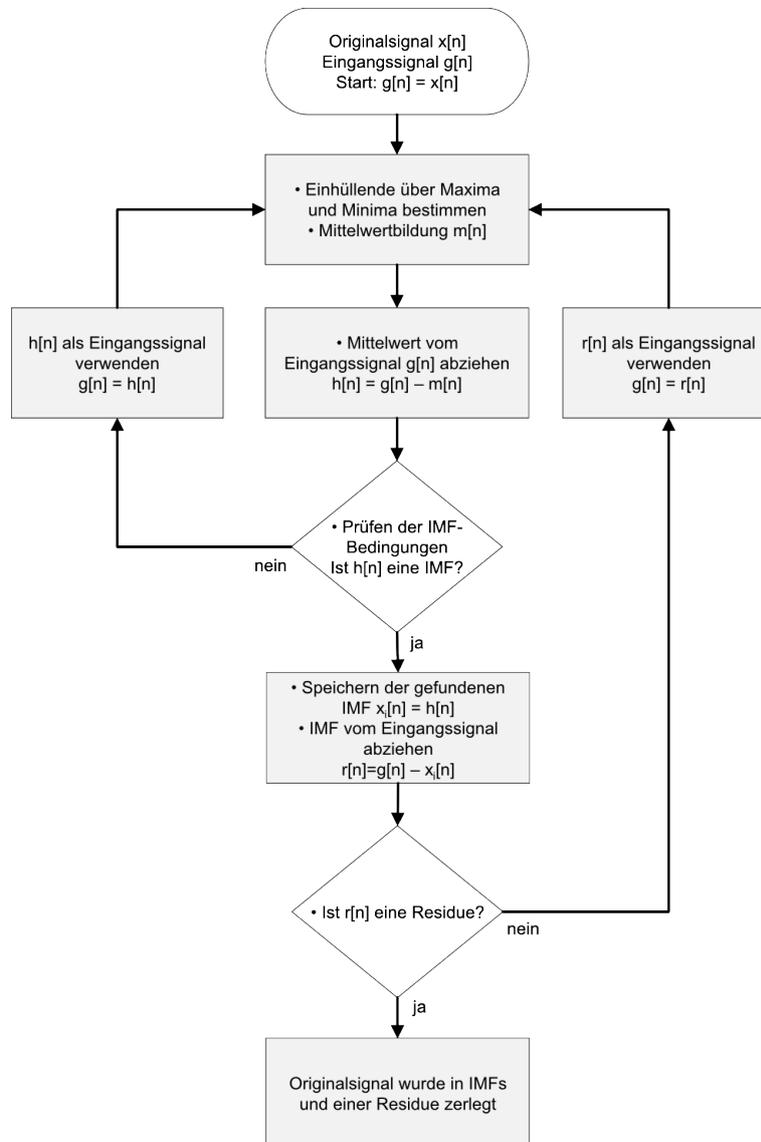


Abbildung B.1: Flussdiagramm des EMD-Algorithmus' [10]

Damit eine extrahierte Funktion als „intrinsic mode function“ (IMF) definiert werden kann, müssen zwei Bedingungen zutreffen:

1. Die Anzahl der Maxima und die der Nullpunktdurchgänge müssen gleich sein bzw. dürfen sich maximal um eins unterscheiden.
2. Der laufende Mittelwert, der über die Einhüllende der Maxima und Minima gebildet wird, muss Null sein.

Beschreibung des Vorgangs:

Das Auffinden der IMFs aus dem Originalsignal wird als „sifting“ (Sieben) bezeichnet. Hier werden schnell oszillierende Signalanteile, sogenannte „riding waves“, von den zugrundeliegenden, langsamer oszillierenden Anteilen getrennt.

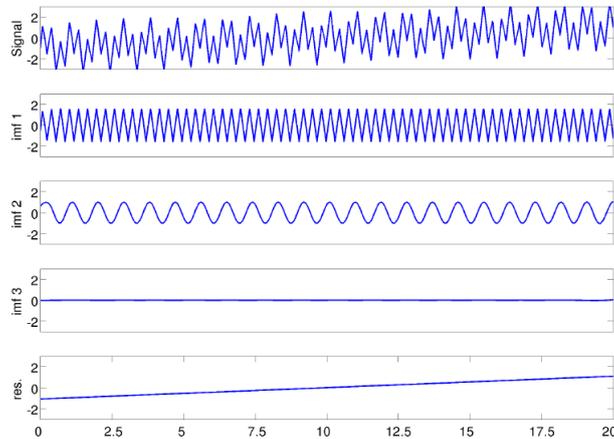


Abbildung B.2: Beispiel eines Signals bestehend aus: Sägezahn, Sinus und linearem Trend [10]

Der „Sifting-Prozess“

1. Auffinden der lokalen Maxima des Signals $x[n]$.
2. Bilden der Einhüllenden $u_1[n]$ über die Maxima mittels kubischer Spline-Interpolation.
3. Auffinden der lokalen Minima des Signals $x[n]$.
4. Bilden der Einhüllenden $l_1[n]$ über die Minima.
5. Bilden des laufenden Mittelwerts nach:

$$m_1[n] = \frac{l_1[n] + u_1[n]}{2} \quad (\text{B.1})$$

6. Der erhaltene Mittelwert $m_1[n]$ wird nun vom ursprünglichen Signal $x[n]$ abgezogen:

$$h_1[n] = x[n] - m_1[n] \quad (\text{B.2})$$

7. Das neu entstandene Signal $h_1[n]$ wird nun auf die Bedingungen der IMF geprüft. Sollte $h_1[n]$ diese erfüllen, stellt es eine „intrinsic mode function“ dar.

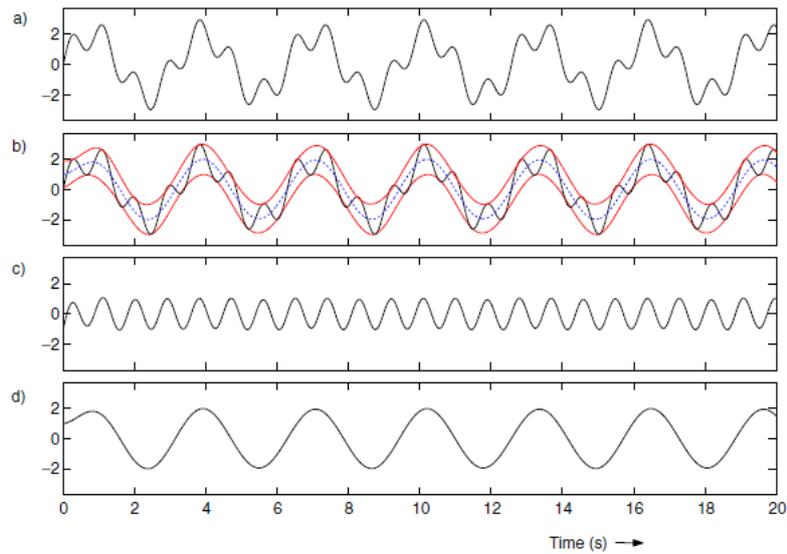


Abbildung B.3: a) Originalsignal, b) obere und untere Einhüllende (rot) und deren Mittelwert (blau), c) die erste IMF, d) das erste Restsignal [11]

8. Können die Bedingungen nicht eingehalten werden, erfolgt eine sogenannte „Siebung“, dh. der Vorgang der Mittelwertbildung und der Subtraktion wird nun mit $h_1[n]$ als „neues Eingangssignal“ durchgeführt.
9. Ein so gefundenes $h_{11}[n]$ wird ebenfalls auf die Bedingungen der IMF geprüft.
10. Der Vorgang der „Siebung“ wird so lange wiederholt, bis eine mögliche Komponente $h_{1k}[n]$ eine erste IMF $x_1[n]$ darstellt.
11. Die erste IMF wird nun vom Originalsignal $x[n]$ subtrahiert und ergibt das erste „Restsignal“:

$$r_1[n] = x[n] - x_1[n] \quad (\text{B.3})$$
12. Das Restsignal $r_1[n]$ stellt nun das neue „Ausgangsmaterial“ für den weiteren Zerlegungsprozess dar.
13. Das ursprüngliche Signal $x[n]$ wird nun in eine endliche Anzahl von IMFs $x_i[n]$ zerlegt.
14. Die „Siebung“ ist beendet, wenn das Restsignal $r_N[n]$ entweder eine konstante oder monotone Funktion darstellt.
15. Das Ursprungssignal $x[n]$ kann nun angeschrieben werden als:

$$x[n] = \sum_{i=1}^N x_i[n] + r_N[n] \quad (\text{B.4})$$

Die Hilbert-Transformation

Mittels Hilbert-Transformation erhält man für jede IMF: $z_i[n] = x_i[n] + jy_i[n] = a_i[n]e^{j\Phi_i[n]}$. Analytische Signale ermöglichen die Berechnung von Momentanattributen, die in weiterer Folge zur Sonifikation verwendet werden. (vgl. [9])

Folgende Momentanattribute lassen sich durch die HT berechnen:

$$\text{Momentanamplitude: } a_i[n] = \sqrt{x_i[n]^2 + y_i[n]^2}$$

$$\text{Momentanphase: } \phi_i[n] = \tan^{-1}\left(\frac{y_i[n]}{x_i[n]}\right)$$

$$\text{Momentanfrequenz: } f_i[n] = \frac{1}{2\pi} \frac{d\Phi_i}{dn}$$

B.3 Probleme in Hinblick auf die Online-EMD

Bei der Umsetzung der Sonifikation „Birds“, vor allem in Hinblick auf die Echtzeitanwendung der EMD, sind folgende Merkmale zu beachten:

Fixe Anzahl an „Siftings“

Aufgrund der nicht voraussagbaren Struktur des Eingangssignals, ist es nicht möglich abzuschätzen, wieviele *intrinsic mode functions (IMF)* dem Signal innewohnen. Es ergeben sich daher große zeitliche Unterschiede beim Auffinden der IMFs. Der „Sifting“-Prozess wird daher nicht vollständig durchlaufen und nach der Extraktion von 5 IMFs abgebrochen. Die daraufhin gefundenen IMFs besitzen daher nicht die „idealen“ Eigenschaften nach den oben beschriebenen Bedingungen. Die Momentanfrequenz wird aufgrund der sich ergebenden Instabilitäten nicht aus der Hilbert-Transformation, sondern mittels „*generalized zero-crossing*“ bestimmt.

Fixe Anzahl an IMFs

Aufgrund des zeitlichen Aufwands zur Bestimmung der IMFs, ist es nicht möglich alle, dem Signal innewohnende IMFs zu extrahieren. Die maximale Anzahl der extrahierten IMFs ist daher, wie in Abschnitt 8.2.2 erwähnt, auf fünf begrenzt.

Literatur

- [1] A. Wankhammer, *Akustisches Interface zur Tremoranalyse - Endbericht - Oktober 2012* - internes Dokument
- [2] Gregory Kramer, *Auditory display: Sonification, audification, and auditory interfaces*, 1994, ISBN 0201626039
- [3] Thomas Hermann, *Daten hören - Sonifikation zur explorativen Datenanalyse*, 2008 aus *Sound Studies: Traditionen - Methoden - Desiderate*, Holger Schulze (Hg.), transcript Verlag, ISBN 978-3-89942-894-0
- [4] Georg Thieme Verlag Stuttgart, *Leitlinien zur Diagnostik und Therapie in der Neurologie*, ISBN: 978-3-13-132414-6
- [5] <http://www.biometricsltd.com/k800.htm> (zuletzt aufgerufen am 31.10.2013)
- [6] Jonathan Shlens, *A tutorial on principal component analysis*, Systems Neurobiology Laboratory, University of California at San Diego (2005)
- [7] D. Pirrò, A. Wankhammer, P. Schwingenschuh, R.Höldrich, A. Sontacchi, *Acoustic interface for tremor analysis*, Presented at the 18th International Conference on Auditory Display (ICAD2012) on June 18-21, 2012 in Atlanta, Georgia - Paper
Siehe dazu auch: <http://hdl.handle.net/1853/44437> (zuletzt aufgerufen am 31.10.2013)
- [8] D. Pirrò, A. Wankhammer, P. Schwingenschuh, R.Höldrich, A. Sontacchi, *Acoustic interface for tremor analysis*, Presented at the 18th International Conference on Auditory Display (ICAD2012) on June 18-21, 2012 in Atlanta, Georgia - Poster
- [9] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung and H. H. Liu, *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis*, Proc. R. Soc. Lond. A (1998) 454, 903-995
- [10] A. Zeiler, R. Faltermeier, I. R. Keck and A. M. Tomé, *Empirical Mode Decomposition - An Introduction*, Member, IEEE, C.G. Puntonet and E. W. Lang, ISBN: 978-1-4244-6916-1
- [11] P. J. Oonincx, J.-P. Hermand, *Empirical Mode Decomposition of ocean acoustic data with constraint on the frequency range*, Proceedings of the Seventh European Conference on Underwater Acoustics, ECUA 2004 Delft, The Netherlands, 5-8 July, 2004

- [12] <http://www.science-products.com/Products/CatalogC/Interfaces/Power1401/Power1401.html>
(zuletzt aufgerufen am 31.10.2013)
- [13] <http://www.science-products.com/Products/CatalogC/Acq&AnaSoftware/Spike2/Spike2.html>
(zuletzt aufgerufen am 31.10.2013)
- [14] http://www.statistics4u.info/fundstat_germ/cc_pca.html (zuletzt aufgerufen am
31.10.2013)
- [15] <http://puredata.info/> (aufgerufen am 31.10.2013)
- [16] <http://www.mathworks.de/products/matlab/> (aufgerufen am 31.10.2013)
- [17] D. Pirrò, A. Wankhammer, *Acoustic interface for tremor analysis - Progress Report
01 - Oktober 2011* - internes Dokument

Abbildungsverzeichnis

3.1	Befestigung der Sensoren an der Hand	12
3.2	Aktion	13
3.3	Leistungsspektren von Parkinson (oben), Essentiellem Tremor (Mitte) und Psychogenem Tremor (unten); ptf...peak tremor frequency; hw...half width [7]	13
6.1	Messequipment an der Universitätsklinik für Neurologie	18
6.2	Beschreibung der Messgeräte	18
6.3	3-Achsen Accelerometer	19
6.4	Subject-Unit	20
6.5	Base-Unit	20
6.6	Split der analogen Kanäle an der BNC-Leiste	21
6.7	Power 1401 in der Signalkette	21
6.8	Datenerfassung mittels Spike [13]	22
7.1	Sensorsignal vor (grau) und nach (schwarz) der DC-Filterung [17]	23
7.2	a) Richtung der maximalen Streuung einer rein translatorischen Bewe- gung, b) Richtungen der maximalen Streuung einer rotatorischen Bewe- gung [8]	25
7.3	Translation und Rotation [8]	26
8.1	Flussdiagramm zur Sonifikation „Spectral Shaping“ [8]	29
8.2	Flussdiagramm zur Sonifikation „Vocoder“ [8]	32
8.3	a) Gewinnung der Momentanamplitude und -frequenz aus den IMFs, b) Ansteuern von Sinusoszillatoren mit modifizierten IMFs [8]	37
9.1	Benutzeroberfläche	41
9.2	Grafische Darstellungsfelder	41
9.3	Bedienoberflächen der beiden Sonifikationen	42
9.4	Globale Bedienelemente	43
A.1	Erste (rot) und zweite (schwarz) „principal component“ für unterschiedli- che Eingangssignale: a) rotatorisch, b) translatorisch-rotatorisch, c) trans- latorisch [17]	46
A.2	„Planes of Movement“ [17]	46
B.1	Flussdiagramm des EMD-Algorithmus' [10]	49
B.2	Beispiel eines Signals bestehend aus: Sägezahn, Sinus und linearem Trend [10]	50
B.3	a) Originalsignal, b) obere und untere Einhüllende (rot) und deren Mittel- wert (blau), c) die erste IMF, d) das erste Restsignal [11]	51