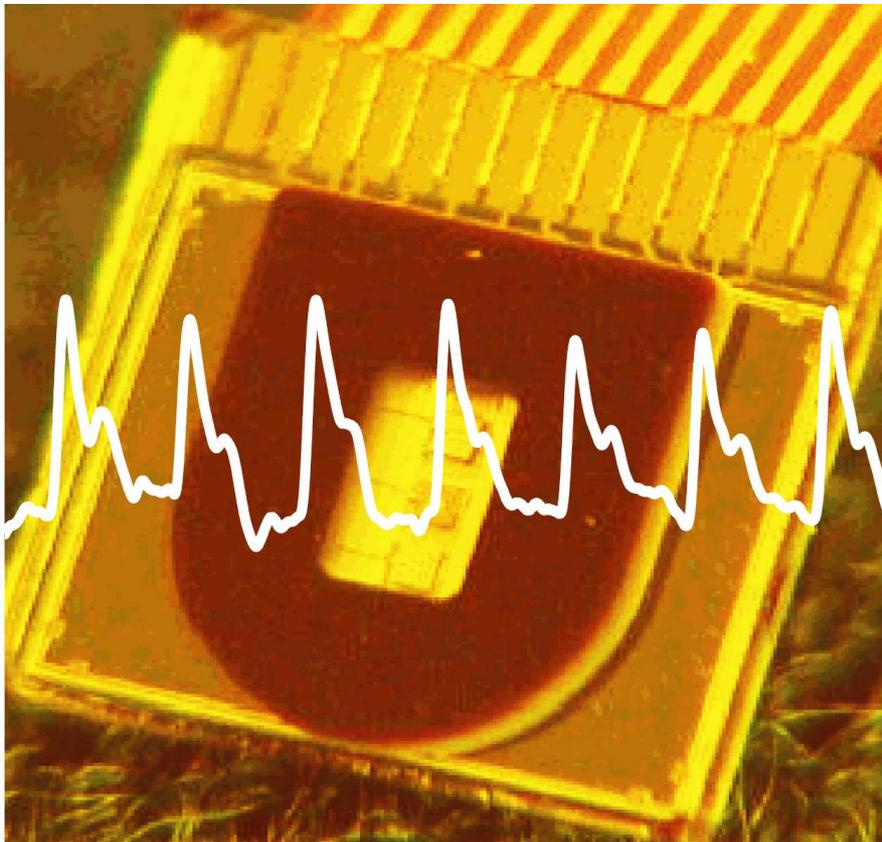


Fachhochschule Amberg-Weiden
Fachbereich Elektro- und Informationstechnik

Studiengang Elektrotechnik
mit Schwerpunkt Multimedialechnik

Diplomarbeit von
Andreas, T o b o l a



Verarbeitung von Signalen eines Remissionssensors
zum Einsatz in der Pulsoximetrie



Diplomand: Andreas Tobola
Student der Elektrotechnik mit Schwerpunkt
Multimediatechnik

Ausgeführt am Fraunhofer Institut für Integrierte Schaltungen
Abteilung Bildverarbeitung und Medizintechnik
Am Wolfsmantel 33
91058 Erlangen

Betreuer: Dipl.-Ing. Hans-Joachim Mörsdorf

in Zusammenarbeit mit Fachhochschule Amberg – Weiden
Hochschule für Technik und Wirtschaft
Fachbereich Elektro – und Informationstechnik
Kaiser-Wilhelm-Ring 23
92224 Amberg

Prüfer/Betreuer: Prof. Dr. Ulrich Vogl
Zweitprüfer/Betreuer: Prof. Dr. Alfred Höß

Thema: Verarbeitung von Signalen eines Remissionssensors zum
Einsatz in der Pulsoximetrie

Bearbeitungszeitraum: 01.11.2004 – 31.03.2005

**Fachhochschule Amberg – Weiden
Hochschule für Technik und Wirtschaft**



Fachbereich Elektro – und Informationstechnik
Studiengang: [Elektrotechnik](#)

Bestätigung gemäß § 31 Abs. 7 RaPO

Name und Vorname des Studenten: [Andreas Tobola](#)

Ich bestätige, dass ich die Diplomarbeit mit dem Titel:

[Verarbeitung von Signalen eines Remissionssensors zum Einsatz in der Pulsoximetrie](#)

selbständig verfaßt, noch nicht anderweitig für Prüfungszwecke vorgelegt, keine anderen als die angegebenen Quellen oder Hilfsmittel benützt sowie wörtliche und sinngemäße Zitate als solche gekennzeichnet habe.

Datum: [11.04.2005](#)

Unterschrift:

Fachhochschule Amberg – Weiden
Hochschule für Technik und Wirtschaft



Fachbereich Elektro – und Informationstechnik

Studiengang: [Elektrotechnik](#)

Diplomarbeit Zusammenfassung

Studentin / Student (Name, Vorname): [Andreas Tobola](#)

Studiengang / Studienschwerpunkt: AT EA AE MM

Semester der Anmeldung: 8

Aufgabensteller: [Prof. Dr. Ulrich Vogl](#)

Prof.: [Prof. Dr. Alfred Höß](#)

Durchgeführt in (Firma / Behörde / FH):

[Fraunhofer Institut für Integrierte Schaltungen](#)
[Abteilung Bildverarbeitung und Medizintechnik](#)
[Am Wolfsmantel 33](#)
[91058 Erlangen](#)

Betreuer in Firma / Behörde: [Dipl.-Ing. Hans-Joachim Mörsdorf](#)

Ausgabedatum: [01.11.2004](#)

Abgabedatum: [31.03.2005](#)

Semester der Abgabe: 8

Thema: [Verarbeitung von Signalen eines Remissionssensors zum Einsatz in der Pulsoximetrie](#)

Zusammenfassung: Ziel der Arbeit ist die Entwicklung eines Systems zur Bestimmung von Blutsauerstoffsättigung (SpO2) und Herzfrequenz aus den Signalen eines Remission-Sensors. Besonderes Augenmerk ist dabei auf die Robustheit des Verfahrens gegenüber Störungen wie z.B. Umgebungslicht oder Bewegungsartefakten zu richten. Dazu sind zum einen entsprechende Algorithmen zu entwickeln und zu analysieren (MATLAB), zum anderen ist die Hardware (Sensor, Analogteil, AD-Wandler, Mikrocontroller) zur Aufbereitung und Digitalisierung der Sensorsignale entsprechend anzupassen.

Schlüsselworte:

[Pulsoximetrie](#), [SpO2](#), [Sauerstoffsättigung](#), [Plethysmogramm](#), [Spread Spectrum](#), [TriCore](#)

Inhaltsverzeichnis

1 Einleitung.....	1
1.1 Motivation.....	1
1.2 Aufgabenstellung.....	3
1.3 Projektstadium.....	4
1.4 Ziel dieser Arbeit.....	5
2 Grundlagen.....	6
2.1 Physiologie	6
2.1.1 Pulsfrequenz.....	6
2.1.2 Blut.....	7
2.1.3 Hämoglobinvarianten.....	9
2.1.4 Funktionelle und fraktionelle Sauerstoffsättigung.....	10
2.2 Photoelektrische Plethysmographie (PPG).....	11
2.2.1 Definition.....	11
2.2.2 Transmissions-Sensoren.....	12
2.2.3 Remissions-Sensoren.....	15
2.3 Pulsoximetrie (SpO ₂).....	21
2.3.1 Einführung.....	21
2.3.2 Absorptionsspektren.....	22
2.3.3 Analytisches Modell für Transmission.....	24
2.3.4 Kalibrierung.....	27
2.3.5 Transmissionsmodell mit Remissions-Sensoren.....	27
2.3.5.1 Mit der Gleichung für Transmission 2.34.....	28
2.3.5.2 Mit der Gleichung für Remission 2.30.....	28
2.3.6 Einflüsse auf die Messgenauigkeit.....	29
2.3.6.1 Umgebungslicht.....	29
2.3.6.2 Elektromagnetische Felder.....	30
2.3.6.3 HF-Chirurgie.....	30
2.3.6.4 Nagellack.....	30
2.3.6.5 Bewegungsartefakte.....	30
2.3.6.6 Minderperfusion.....	30
2.3.6.7 Krankheitsbilder.....	30
2.3.6.8 Hautpigmentierung.....	31
2.3.6.9 Dysfunktionelles Hämoglobin	31
2.4 Multiplex-Verfahren.....	31
3 Systemüberblick.....	35

4 Algorithmus.....	37
4.1 Signalverarbeitung mit MATLAB.....	37
4.2 Chip-Folgen und Filter-Generierung.....	39
4.2.1 <i>Chip-Folgen</i>	39
4.2.1.1 Forderungen für die Chip-Folgen.....	39
4.2.1.2 Legendre-Folgen.....	40
4.2.2 <i>Extraktionsfilter</i>	42
4.2.3 <i>Filter der Frequenzweiche</i>	46
4.2.3.1 Interpolations-TP-Filter.....	46
4.2.3.2 Basis des Interpolations-TP-Filters.....	49
4.2.3.3 Interpolations-HP-Filter.....	50
4.3 Verarbeitungsschritte unter MATLAB.....	50
4.3.1 <i>Signal-Import in MATLAB</i>	51
4.3.2 <i>Korrelator und Sub sampler</i>	52
4.3.3 <i>Umgebungslichtsubtraktion</i>	55
4.3.4 <i>Blockbildung zur Blockverarbeitung</i>	57
4.3.5 <i>Frequenzweiche</i>	59
4.3.6 <i>Fensterung</i>	60
4.3.7 <i>Interference Canceler</i>	62
4.3.8 <i>Normierung</i>	64
4.3.9 <i>Fourier Transformation</i>	65
4.3.10 <i>Total Least-Squares Fit 1</i>	67
4.3.11 <i>Spektrale Maske</i>	71
4.3.11.1 Methode der Schwellwertfunktion.....	72
4.3.11.2 Methode der Harmonischen Beziehung.....	73
4.3.11.3 Verknüpfung der binären Masken.....	74
4.3.12 <i>Berechnung der Herzfrequenz</i>	75
4.3.13 <i>Total Least-Squares Fit 2</i>	76
4.3.14 <i>Kalibrierungsfunktion</i>	77
5 Hardware – Entwicklung.....	78
5.1 MORES ein Remissions-Sensor.....	78
5.2 TriCore – Entwicklungsumgebung.....	81
5.2.1 <i>TriCore 1765 – Mikrocontroller</i>	81
5.2.2 <i>Evaluation-Board (TriBoard)</i>	82
5.2.3 <i>Analog-Digital-Wandler</i>	83
5.3 Analoge Verarbeitung.....	84
5.3.1 <i>Ansteuerung der LEDs</i>	86
5.3.2 <i>Aufbereitung des Sensorsignals</i>	90
5.3.2.1 Transimpedanzwandler mit TP-Filter.....	90
5.3.2.2 Spannungsverstärker.....	93
5.3.2.3 Schutzschaltung für den AD-Wandler.....	94
5.3.2.4 Vermessung der Linearität.....	94
5.3.3 <i>Status-LEDs</i>	95
5.3.4 <i>Bestückung und Pin-Belegung</i>	96

6 Software – Entwicklung.....	99
6.1 TriCore-Programm.....	99
6.1.1 Entwicklungswerkzeuge.....	99
6.1.2 Funktionen der TriCore-Software.....	99
6.1.3 C-Programmmodule.....	101
6.1.4 Programmablauf.....	102
6.1.5 Laufzeitfehler-Behandlung.....	105
6.1.6 Kommunikationsprotokoll.....	105
6.2 Windows-Anwendung.....	107
6.2.1 Funktionalität.....	107
6.2.2 Übersicht der Klassen.....	109
6.2.3 TPD-Format.....	109
7 Untersuchungen und Ergebnisse.....	110
7.1 Black-Box zur Interferenz-Simulation.....	110
7.2 Einfluss von Störungen auf die Messgrößen.....	111
7.2.1 Test der Umgebungslichtsubtraktion.....	111
7.2.2 Störquellen im Übertragungsband.....	115
7.2.3 Breitbandige Störungen.....	120
7.2.4 100-Hz-Störer Unterdrückung.....	124
7.3 Referenzmessung der Herzfrequenz.....	125
8 Ausblick.....	126

Abkürzungsverzeichnis

ADC	Analog to Digital Converter
ASC	Asynchronous Serial Controller
AM	Amplitudenmodulation
bpm	Beats per minute (Schläge pro Minute)
CDMA	Code Division Multiple Access
CRC	Cyclic Redundancy Check
DFT	Discrete Fourier Transformation
DSSS	Direct Sequence Spread Spectrum
EKG	Elektrokardiogramm
EMV	Elektromagnetische Verträglichkeit
FDMA	Frequency Division Multiple Access
FM	Frequenzmodulation
GND	Massenpotential (engl. Ground)
GTPA	General Purpose Timer Array
GTPU	General Purpose Timer Unit
HP	Hochpass
FIR	Finite Response Filter
Hb	Hämoglobin
IIR	Infinite Response Filter
IR	Infrarot
I/O	Input / Output
PC	Personal Computer
PLL	Phase Locked Loop
RISC	Reduced Instruction Set Computing
SaO ₂	Arterielle Blutsauerstoffsättigung; alle übrigen Verfahren
SpO ₂	Arterielle Blutsauerstoffsättigung; ermittelt mittels Pulsoximeter
SIMD	Single Instruction, Multiple Data

SSC	Synchronous Serial Controller
Spektrogramm	Ein Plot in drei Dimensionen, in dem mehrere Spektren über die Zeit angezeigt werden.
SDK	Software Development Kit
SRAM	Static Data Memory
SNR	Signal-to-Noise Ratio (Signal- zu Rauschverhältnis)
SVD	Singular Value Decomposition
TDMA	Time Division Multiple Access; Zeitmultiplexverfahren
TP	Tiefpass
RS232	Norm einer asynchronen seriellen Schnittstelle
LED	Light Emitting Diode; lichtemittierende Diode; Leuchtdiode
LSB	Least Significant Bit/Byte
MAT-Datei	Binäre MATLAB®-Datei zum Speichern von Daten
M-Skript	MATLAB® – Stapelverarbeitungsdatei
MORES®	Micro-optical Remissions/Reflection Sensor system
NC	Not connected, Leitung nicht verbunden
MSB	Most Significant Bit/Byte
µC	Mikrocontroller

Notationen

A	Matrix A
A^{-1}	Inverse der Matrix A
$A^{\#}$	Pseudoinverse der Matrix A
A^{T}	Matrix A transponiert
A^*	Matrix A komplex konjugiert
$A^{\dagger} = A^{\text{T}*}$	Matrix A adjungiert
E	Eine Einheitsmatrix E hat Einsen in der Hauptdiagonalen und sonst Nullen
$[A]_{ik}$	Element der Matrix A in der i-ten Zeile und k-ten Spalte
\vec{a}	Vektor a
$A = \text{diag}(\vec{x})$	Matrix A bestehend aus Werten des Vektors x in der Hauptdiagonalen
$\text{rang}(A)$	Rang einer Matrix A
$A(k)$	Matrix A abhängig von der Zeitkonstante k
$\vec{a}(k)$	Vektor a abhängig von der Zeitkonstante k
Δ	Symmetrische Differenz
\forall	Für alle
$[a, b]$	Geschlossenes Intervall a bis b
$\{\dots\}$	Zahlenmenge
$\langle \vec{x}_1, \vec{x}_2 \rangle$	Skalarprodukt aus x_1 und x_2
$Ax \stackrel{!}{=} b$	Unter- oder überbestimmtes Gleichungssystem
$j = \sqrt{-1}$	Imaginäre Einheit
$M \downarrow$	Unterabtastung um den Faktor M
\log	Logarithmus zur Basis 10
\ln	Logarithmus zur Basis e

Glossar

Bewegungsartefakte	Mit Bewegungsartefakten sind hier die Störungen gemeint, die durch Bewegung des Patienten verursacht werden
Debug	Suchen und Entfernen von Fehlern aus Programmen
Diastole	Minimaler Arteriendruck
Figure	Graphischer Container unter MATLAB; enthält einen oder mehrere Graphen
Block	Signalausschnitt bei der Blockverarbeitung
Graph	Grafische Abbildung einer Funktion
invasiv	Eindringen in den Körper
Systole	Maximaler Arteriendruck
Perfusion	Durchblutung des Gewebes
Plethysmogramm	Pulswellenkurve, die Blutvolumenänderungen in Blutgefäßen widerspiegelt
pulsatil	Auf Kontraktionen des Herzens zurückzuführen

1 Einleitung

1.1 Motivation

Der technische Fortschritt hat die Überwachung kritisch kranker Patienten stark verändert. Nichts erinnert mehr an den Stand vor 25 Jahren. Heute wird der Patient über zahlreiche Kabel an Überwachungsgeräte angeschlossen. Diese erfassen über Sensoren zyklisch Messwerte und berechnen daraus aussagekräftige *Vitalparameter*, die auf den Zustand des Patienten schließen lassen. Beispiele für geläufige Vitalparameter sind: Körpertemperatur, Herzfrequenz, Blutdruck und die arterielle Blutsauerstoffsättigung. Vitalparameter spielen nicht nur bei der Diagnose eine entscheidende Rolle. Da die Überwachungsgeräte permanent Werte liefern und vorwiegend mit Alarmfunktionen versehen sind, spiegeln sie den aktuellen Zustand des Patienten wider und helfen lebensbedrohliche Veränderungen frühzeitig zu erkennen.

Beim Krankentransport, in der Notfallaufnahme, Chirurgie, Intensiv- und Pflegestation sind zuverlässige Aussagen über den Zustand des Patienten unverzichtbar. Auch außerhalb der klinischen Einrichtungen kann man einen Teil dieser Vitalparameter gut interpretieren. Der rasante Fortschritt der Informations- und Elektrotechnik lässt die Überwachungsgeräte kompakter und kostengünstiger werden. So findet man mittlerweile Uhren mit eingebautem EKG. Im Folgenden wird anhand von zwei Beispielen gezeigt, wie Überwachungsgeräte in der Sportmedizin und bei der Heimüberwachung eingesetzt werden können.

Beim Training im Fitnessstudio misst und überwacht ein Mikrocontroller die Pulsfrequenz des Sportlers und versucht diese in einem für die Fettverbrennung am besten geeigneten Bereich zu halten. Dies wird durch einen typischen Regelkreis realisiert. Um diesen zu schließen, gibt das Trainingsgerät Anweisungen an den Sportler:

- „Langsamer bewegen“, um die Pulsrate zu senken.
- „Schneller bewegen“, um die Pulsrate zu erhöhen.

Bei einigen Patienten ist ein langfristiger Krankenhausaufenthalt nicht zwingend notwendig, aber eine Überwachung dennoch erwünscht, falls ein wiederkehrender Ernstfall (Herzattacke, Atemnot etc.) zu erwarten ist. Betroffene Patienten können ihrem üblichen Tagesablauf nachgehen, während ein tragbares, kompaktes Gerät fortlaufend ihre Lebensfunktionen analysiert. Anomalien der *Vitalparameter* können so schon frühzeitig erkannt werden. Zusätzlich sind diese Geräte mit einem akustischen Alarm ausgestattet. Sie warnen vor einer bestehenden Gefahr, damit der Betroffene vor einem drohendem Zusammenbruch einen Arzt

benachrichtigen oder möglicherweise dem angekündigten Ernstfall mit entsprechendem Verhalten vorbeugen kann.

Die Ingenieure des *Fraunhofer Institutes für Integrierte Schaltungen* gehen einen Schritt weiter: Im Rahmen des Forschungsprojektes *Body Area Network (BAN)* können mittels Sensoren mehrere *Vitalparameter* eines Patienten über Funk an eine zentrale Einheit übermittelt werden. Im Notfall alarmiert diese Zentrale automatisch über eine Telefonleitung oder das Internet einen Arzt und übermittelt diesem bei Bedarf den Verlauf der *Vitalparameter*. Die Abb. 1.1 zeigt ein Anwendungsbeispiel des *BAN* zur Überwachung von Patienten in einer klinischen Einrichtung. Vitalparameter werden über Funk an eine nahegelegene Zentrale übermittelt. Von dort aus gelangen sie über ein lokales Netz (LAN) an den zuständigen Krankenpfleger oder Arzt. Abb. 1.2 zeigt ein Anwendungsbeispiel im Heimbereich. Auch hier werden die Vitalparameter zunächst an eine nahegelegene Station geleitet. Von dort aus gelangen sie über das Internet an eine medizinisch ausgebildete Zielgruppe.

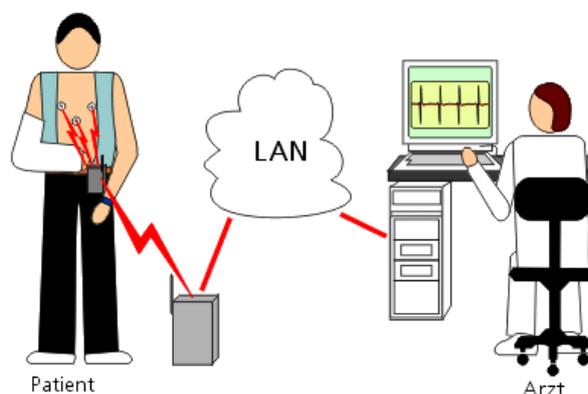


Abbildung 1.1: Einsatzgebiet des Body Area Network (BAN) im Krankenhaus
Quelle: www.iis.fraunhofer.de

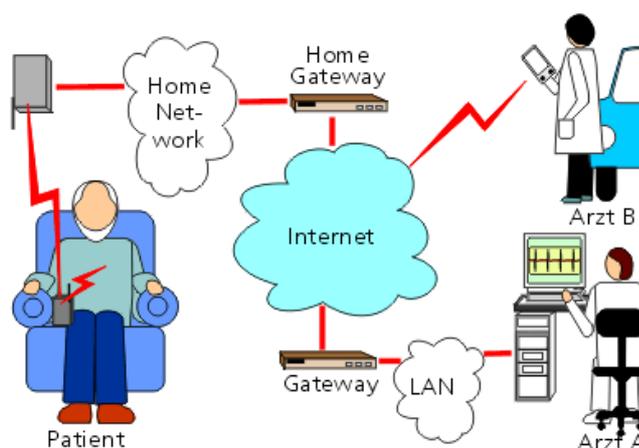


Abbildung 1.2: Einsatzgebiet des Body Area Network (BAN) während einer Heimüberwachung
Quelle: www.iis.fraunhofer.de

1.2 Aufgabenstellung

Zunächst steht die drahtlose Übertragung von Patientendaten im Rahmen des Projekts BAN im Vordergrund. Das Projekt lässt sich wegen seines Umfangs in viele Teilprojekte untergliedern. Unter anderem gehört dazu auch das Erforschen neuer und effizienterer Verfahren zur Erfassung diverser Vitalparameter. Ein solches Teilprojekt wurde zum Thema dieser Diplomarbeit, in der ein Verfahren zur möglichst störungsfreien Messung der arteriellen, funktionellen Sauerstoffsättigung mittels Pulsoximetrie entwickelt werden soll.

Ein Pulsoximeter ist ein Gerät zur Messung zweier Vitalparameter mittels eines optischen Sensors, der meistens am Finger oder Ohr läppchen angebracht wird. Die zwei Vitalparameter sind die arterielle Sauerstoffsättigung (SpO_2) und die Pulsfrequenz (HR), wobei die Messung der Sauerstoffsättigung in der Pulsoximetrie und damit auch im Rahmen dieser Arbeit im Vordergrund steht. Es wird sich später herausstellen, dass die Pulsfrequenz mit vergleichbar wenig Mehraufwand zusätzlich erfasst und zusammen mit der Sauerstoffsättigung über die Zeit aufgezeichnet werden kann.

Um diese zwei Werte erfassen zu können, wird ein geeigneter optischer Sensor benötigt. Am Fraunhofer Institut experimentierte man bereits mit dem Remissions-Sensors MORES des CiS Instituts für Mikrosensorik gGmbH und entwickelte dazu eine Anlogschaltung, mit der es möglich ist, die Signale des Sensors am Oszilloskop über die Zeit zu betrachten. Würden die Signale digitalisiert und in einen Rechner gespeichert werden, so könnte man daraus die gesuchten Größen berechnen. Da die Vorverarbeitung der Sensorsignale ausschließlich mittels analoger Schaltungen erfolgt, sind viele Bauelemente, und damit viel Platz erforderlich. Würde die Filterung digital auf einem Mikrocontroller erfolgen, so könnte Platz eingespart werden. Die erste Teilaufgabe besteht darin, die Eignung des Prototypen für die weitere Verarbeitung zu prüfen. Dazu sollen die Signale des Prototypen digitalisiert und auf ihre Eigenschaften untersucht werden.

Anzumerken ist, dass Pulsoximeter als sehr sensibel und stör anfällig bekannt sind. Diverse Störquellen führen vorwiegend zum Fehler in der SpO_2 -Messgröße. Solche Störeinflüsse können Bewegungsartefakte oder einfach nur die Raumbeleuchtung sein. Darüber hinaus berichten Ärzte sogar von Anzeigenausfall oder völlig falschen Werten. In einem weiteren Schritt wird ein Verfahren gesucht, das den Fehler in der Messgröße im Bezug auf Störquellen minimiert. Die Entwicklung eines solchen Verfahrens soll als Algorithmus zur Berechnung des SpO_2 -Wertes und der Herzfrequenz unter MATLAB[®] erfolgen.

1.3 Projektstadium

Bislang wurde ein Prototyp einer analogen Signalverarbeitungseinheit zur Verarbeitung von Signalen des MORES entwickelt. Die Hardware beinhaltet Treiberstufen zur Ansteuerung der drei Sensor-LEDs. Über eine Photodiode liefert der Sensor einen zur reflektierten Restintensität proportionalen Strom. Die LEDs werden von einem externen Mikrocontroller H8S/2329-F der Firma Renesas im Zeitmultiplex-Verfahren getaktet. Über einen Transimpedanzwandler gelangt das Signal als Spannung zu einem Demultiplexer, der als Zeitmultiplex-Demodulator dient. Die so in drei Kanäle gesplitteten Signale gelangen zu Sample&Hold-Gliedern. Von dort aus durchläuft jedes der drei Signale eine Reihe von analogen Filtern, einen Tiefpass-, einen 50Hz-Notch- und einen Hochpassfilter. Die A/D-Wandler des Mikrocontrollers sind mit 6 analogen Ausgängen der Schaltung verbunden. Es handelt sich hierbei um zwei Arten von Signalen für jeweils drei Wellenlängen, das Sensor-Rohsignal incl. DC-Komponente und die um einen einstellbaren Faktor verstärkte AC-Komponente. Die Signale werden von dem Mikrocontroller mit 10bit-Quantisierung abgetastet und an den PC weitergeleitet.

Die Schaltung verfügt über eine Umgebungslichteliminierung. Dies wird durch Hinzufügen von Dunkelphasen in die TDMA-Sequenz eingeleitet. In diesen Zeitfenstern sind alle LEDs dunkel. In diesen Dunkelphasen wird die Intensität des Umgebungslichtes erfasst und mittels einer Analogschaltung von jedem Kanal subtrahiert.

1.4 Ziel dieser Arbeit

Ziel dieser Arbeit ist die Entwicklung einer Verarbeitungskette vom Sensor bis zur Anzeige der Sauerstoffsättigung im Hinblick auf die Unterdrückung von Störeinflüssen.

Im ersten Schritt werden Signale mittels der bereits vorhandenen Schaltung digitalisiert und unter MATLAB auf ihre spektralen Eigenschaften untersucht (Anhang B).

Zur Unterdrückung von Störquellen im Vorfeld ist die Spreizspektrum-Modulation im Gegensatz zu der in Pulsoximetern üblichen TDMA weitaus besser geeignet. Es gilt, eine geeignete Variante der Spreizspektrum-Modulation zu finden (Kapitel 4).

Mit dem bereits am Fraunhofer IIS entwickelten Prototypen kann das Spreizspektrum-Modulationsverfahren nicht umgesetzt werden, da der Prototyp im Zeitmultiplex arbeitet. Aus diesem Grund muss neue Hardware entworfen und entwickelt werden. Die Entscheidung fiel wegen einem guten Preis/Leistungsverhältnis auf die Entwicklungsumgebung des DSP TriCore 1765 von Infineon. Zur Vervollständigung der Hardware ist eine Schaltung in Form eines aufsteckbaren Moduls für die TriCore-Entwicklungsplatine erforderlich. Das Modul, das dem Ansteuern der Leuchtdioden und der Vorverarbeitung der Sensor-Signale dient, wird entworfen und aufgebaut (Kapitel 5).

Die Entwicklung und Implementierung eines geeigneten Programmes für den TriCore-Mikrocontroller erfolgt in GNU-C. Dies beinhaltet im Wesentlichen die Ansteuerung des Sensors, die Digitalisierung der Sensor-Signale und die Kommunikation mit einem IBM-kompatiblen PC. Parallel dazu wird eine Windows-Anwendung zur Kommunikation mit dem TriCore-Evaluationboard entwickelt. Die Visualisierung und die Archivierung der Pulsoximetrie-Daten in Dateien zur weiteren Verarbeitung unter MATLAB ist Teil der Funktionalität dieser Software (Kapitel 6).

Neben dem in der Pulsoximetrie neuen Modulationsverfahren (Spreizspektrum), werden auch weitere Methoden zur Unterdrückung von Störeinflüssen untersucht und in MATLAB implementiert (Kapitel 4).

Im letzten Teil wird das Gerät unter verschiedenen Bedingungen auf Unterdrückung von Störeinflüssen untersucht (Kapitel 7).

2 Grundlagen

2.1 Physiologie

Im Folgenden soll der menschliche Körper als Messort untersucht und die Messgrößen definiert werden. Zusätzlich sind für die spätere Entwicklung Informationen zu den Toleranzen und Grenzwerten der Messgrößen, sowie Auflösung der Anzeigen von Bedeutung. Dies lässt sich anhand der physiologischen Grundlagen des Menschen erläutern.

2.1.1 Pulsfrequenz

Die *Pulsfrequenz* ist ein aussagekräftiger Indikator der Tätigkeit des Herzens. Das Herz ist eine Pumpe, die nicht kontinuierlich, sondern mit jeder Herzmuskelkontraktion stoßweise das Blut durch den Körper befördert. Die Pulswellen des Herzens sind an nahezu allen Stellen des Körpers detektierbar. Zur Erfassung dieser Pulswellen kann ein breites Spektrum an Verfahren eingesetzt werden (Beispiele in Tabelle 2.1). Die Pulsrate wird in Schlägen pro Minute (bpm) angegeben und liegt bei einem gesunden Menschen in Ruhe zwischen 50 und 80 bpm. Je nach körperlicher Verfassung kann der Ruhepuls variieren. So schlägt das trainierte Herz eines Marathonläufers im Ruhezustand nur 35mal in der Minute. Bei Neugeborenen dagegen sind Pulsraten von etwa 140 bpm normal. Unter Belastung erreicht ein 30-jähriger Mensch Pulsraten von bis zu 200 bpm. Mit dem Alter sinkt diese gesundheitsbedrohliche Obergrenze zunehmend ab. Zu berücksichtigen sind noch Ausnahmefälle, wie beim Vorhofflattern des Herzens. Hierbei sind Pulsraten bis zu 350 bpm zu verzeichnen. Diese Werte aus [MIZ] , [HB] und [BP] sind der erste Hinweis auf die Auslegung des Messbereiches zur Messung der Pulsfrequenz. Handelsübliche Pulsoximeter sind auf Pulsraten zwischen 30 und 300 bpm mit einem maximalen Fehler von +/-2% ausgelegt. Für die Auflösung der Anzeige hat sich 1 bpm etabliert.

Phys. Verfahren	Messgerät
elektrisch	Elektrokardiogramm (EKG)
optisch	Photoplethysmogramm (PPG)
mechanisch	Blutdruckmessgerät (RR)

Tabelle 2.1: Messverfahren zur Erfassung der Aktivität des Herzens

2.1.2 Blut

Während sich die *Pulsfrequenz* bereits mittels einer Uhr und dank der Sensibilität unserer Finger relativ präzise ermitteln lässt und deren Definition auch allgemein bekannt ist, stellt sich die Messung der *arteriellen Sauerstoffsättigung* als weit aufwändiger heraus. Desweiteren ist auch für Nicht-Mediziner Sachkenntnis über Blutkreislauf und Anatomie des menschlichen Körpers gefordert, um die Bedeutung der *arteriellen Sauerstoffsättigung* zu klären.

Der Blutkreislauf ist das wichtigste Transportsystem des menschlichen Körpers. Angetrieben vom Herz, werden Sauerstoff, Nährstoffe (wie Kohlenhydrate, Eiweiß und Fett), Hormone und Vitamine zu den Zellen im ganzen Körper transportiert. Abfallstoffe, wie das Kohlendioxid, werden wiederum zu den verschiedenen Ausscheidungsorganen befördert. Um leben zu können, benötigt jede Zelle unseres Körpers Energie. Diese gewinnt sie durch biologische Verbrennung von einfachem und komplexem Zucker. Formel (2.1) zeigt die Verbrennung von Monosaccharid. Bei dem hier vereinfacht beschriebenen chemischen Vorgang wird Sauerstoff verbraucht und Kohlenstoffdioxid produziert. Den Transport dieser Stoffe zu und von den einzelnen Zellen übernimmt das Blut über hin- und rückführende Blutbahnen. In den Lungen nimmt das Blut Sauerstoff auf und transportiert diesen über *Arterien* zu den Zellen des Körpers. In den *Venen* wird das Abbauprodukt Kohlenstoffdioxid von den Zellen zu den Lungen befördert [BP]. Dieser Sachverhalt lässt auf eine höhere Sauerstoffkonzentration im arteriellen als im venösen Blut schließen. Für die Messung der arteriellen Sauerstoffsättigung stellt dies zunächst ein Problem dar, geht man davon aus, dass sowohl venöse als auch arterielle Gefäße von dem in der Pulsoximetrie eingesetzten Licht erfasst werden. Eine Lösung wird im nächsten Kapitel gezeigt.



Das Blut besteht zu 44 % aus festen und zu 56 % aus flüssigen Bestandteilen [BP]. Die flüssigen Bestandteile werden Blutplasma genannt. Die festen Bestandteile lassen sich in Erythrozyten, Leukozyten und Thrombozyten unterteilen.

Blutplasma

Das Blutplasma besteht zu 93 % aus Wasser und zu 7 % aus gelösten Stoffen, aus Mineral- und Nährstoffen, Eiweißen und einem kleinen Anteil an Hormonen.

Erythrozyten (rote Blutkörperchen):

Die roten Blutkörperchen machen ca. 90 % der festen Blutbestandteile aus. Sie enthalten das Hämoglobin, dessen Aufgabe darin besteht, je nach Transportrichtung, Sauerstoff oder das Abfallprodukt Kohlendioxid an sich zu binden.

Leukozyten (weiße Blutkörperchen):

Die weißen Blutkörperchen dienen der Abwehr und Bekämpfung von Krankheitserregern und Fremdstoffen.

Thrombozyten (Blutplättchen):

Blutplättchen sind für die Blutgerinnung notwendig. Bei einem Schnitt in die Haut verhindern sie das Austreten größerer Blutmengen aus dem Körper, indem sie das Blut gerinnen lassen.

Blutzusammensetzung

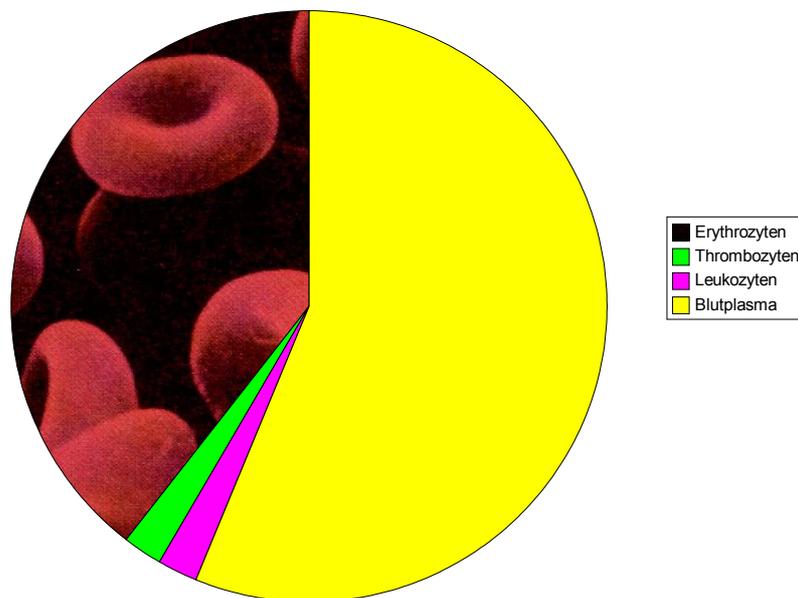


Abbildung 2.1: Zusammensetzung des Blutes

Die Erythrozyten oder genauer, die darin gebündelten Hämoglobinmoleküle sind für die Pulsoximetrie interessant und machen einen Großteil der festen Blutbestandteile aus. Die Überzahl an Erythrozyten ist ein Vorteil für die optische Messung der Sauerstoffsättigung. Jeder Erythrozyt (rotes Blutkörperchen) enthält etwa 280 Mio. Hämoglobinmoleküle. Das Hämoglobin wiederum setzt sich zusammen aus dem Farbstoff Häm und der Eiweißkomponente Globin. Der Farbstoff Häm verleiht dem sauerstoffreichen Blut seine markante rote Farbe. Sauerstoffarmes Blut ist dagegen dunkelrot gefärbt. Jedes Hämoglobinmolekül kann vier Sauerstoffmoleküle an sich binden (2.2). Der Sauerstoff verbindet sich in den Lungenkapillaren mit Hämoglobin. Die mit Sauerstoff angereicherten Blutkörperchen gelangen mit dem Blutstrom von der Lunge in das Herz und von dort über Arterien, die etwas kleineren Arteriolen und schließlich den Kapillaren¹ zu den einzelnen Zellen des Körpers. Hier gibt das Hämoglobin den

¹ In der Medizin werden die winzigen Blutgefäße, die Arterien und Venen miteinander verbinden, als Kapillaren bezeichnet. Sie haben eine Durchmesser von 0,01 bis 0,2 Millimetern und kommen zahlreich fast im gesamten Körper vor.

Sauerstoff ab und nimmt das Kohlendioxid auf. Die Erythrozyten führen es über die Venen zurück zum Herzen und weiter in den Lungenkreislauf. Dort wird wieder Kohlendioxid abgegeben und Sauerstoff aufgenommen.



2.1.3 Hämoglobinvarianten

Verbunden mit dem Eisenatom des Hämoglobinmoleküls, bestimmt ein Ligand² die Art des Hämoglobins und damit auch die spektralen Eigenschaften des Moleküls. In den Erythrozyten des arteriellen Blutes sind vier Varianten des Hämoglobins in nennenswerten Mengen vertreten. Tabelle 2.2 zeigt die am häufigsten im arteriellen Blutstrom vorkommenden Hämoglobine.

Bezeichnung	Kürzel	Ligand
Deoxyhämoglobin oder reduziertes Hämoglobin	HHb bzw. Hb	kein
Oxyhämoglobin	O ₂ Hb	Sauerstoff O ₂
Methämoglobin	MetHb	Hydroxid-Ion OH ⁻
Carboxyhämoglobin	COHb	Kohlenmonoxid CO

Tabelle 2.2: Funktionelle und dysfunktionelle Hämoglobine

Nicht jedes Hämoglobinmolekül wird in der Lunge oxygeniert³. Einige Hämoglobinmoleküle binden entweder keine oder andere Gase aus dem Atemluftgemisch. Der sauerstoffgesättigte Hämoglobinanteil wird als Oxyhämoglobin (O₂Hb) bezeichnet. Das ungebundene Hämoglobin wird reduziertes Hämoglobin oder Deoxyhämoglobin (HHb oder Hb) genannt. Die Hämoglobinderivate Carboxyhämoglobin (COHb) und Methämoglobin (MetHb) sind wegen ihrer stärkeren Bindung mit ihrem Ligand nicht mehr in der Lage Sauerstoff zu binden und werden daher als dysfunktionelles Hämoglobin oder Dyshämoglobine bezeichnet. Die Hämoglobine O₂Hb und Hb bezeichnet man als funktionelles Hämoglobin, weil sie ihre Funktion als Sauerstoffträger gerade erfüllen bzw. dazu in der Lage sind.

In der Regel ist der Anteil am dysfunktionellen Hämoglobin sehr gering. MetHb entsteht normalerweise nur in geringem Umfang (<1% des gesamten Hämoglobins [IMZ]) durch Oxidation des Eisens und gibt den so gebundenen Sauerstoff nicht wieder ab. Erst mit der Zerstörung des ausgedienten Erythrozyten verlässt das MetHb den Blutkreislauf.

2 Die Bezeichnung Ligand stammt aus der Komplexchemie und bezeichnet ein meist negativ geladenes oder ungeladenes Molekül, das über eine sogenannte kooperative Bindung an ein Zentralteilchen gebunden ist.

3 Oxygenierung bedeutet eine lockere, relativ leicht lösbare Bindung mit Sauerstoff. Nicht zu Verwechseln mit Oxidation.

Der Anteil an COHb liegt bei einem Nichtraucher gewöhnlich bei etwa 1% bis 3% des Gesamthämoglobins. Bei Rauchern liegt dieser Anteil grundsätzlich höher bei 5% bis 10% [IMZ] [AS]. Höhere Werte kommen durch Kohlenmonoxidvergiftungen zustande. Prinzipiell wird das COHb in den Lungenkapillaren zusammen mit dem Ausscheiden von Kohlendioxid wieder frei, geht aber zum Teil eine neue Bindung mit Hb ein. Die Ursache ist eine stärkere Affinität des Hämoglobins zu CO als zu O₂.

2.1.4 Funktionelle und fraktionelle Sauerstoffsättigung

Sauerstoffsättigung ist das Verhältnis zwischen oxygeniertem Hämoglobin zum Gesamthämoglobin. Für die exakte Messung der arteriellen Sauerstoffsättigung wird die Konzentration des gesamten Hämoglobins und die des Oxyhämoglobin O₂Hb des arteriellen Blutes benötigt. Der Quotient der Konzentrationen aus O₂Hb und der Summe aller Konzentrationen des Hämoglobins wird als fraktionelle Sauerstoffsättigung bezeichnet und wird üblicherweise in Prozent angegeben. Die Konzentrationen der Hämoglobine lassen sich exakt mittels einer Blutgasanalyse aus einer entnommenen Blutprobe ermitteln und damit die sog. fraktionelle Sauerstoffsättigung des arteriellen Blutes nach (2.3) berechnen.

$$SaO_2^{\text{frakt}} = \frac{c_{O_2Hb}}{c_{O_2Hb} + c_{HHb} + c_{COHb} + c_{MetHb}} \cdot 100\% \quad (2.3)$$

In der Praxis wird vorwiegend die funktionelle Sauerstoffsättigung ermittelt, bei der dysfunktionelles Hämoglobin nicht berücksichtigt wird (2.4).

$$SaO_2^{\text{func}} = \frac{c_{O_2Hb}}{c_{O_2Hb} + c_{HHb}} \cdot 100\% \quad (2.4)$$

Die meisten handelsüblichen Pulsoximeter arbeiten mit zwei Wellenlängen und sind damit nur in der Lage zwischen zwei Stoffen zu unterscheiden. Demnach wird nur die funktionelle Sauerstoffsättigung des arteriellen Blutes berechnet, d.h. nur das funktionelle Hämoglobin, also oxygeniertes und reduziertes Hämoglobin werden berücksichtigt. Einige Pulsoximeter arbeiten mit mehr als zwei Wellenlängen und können auf diese Weise auch dysfunktionelle Hämoglobine unterscheiden. Diese Funktion ist jedoch nicht Teil dieser Arbeit. Dennoch sei für die spätere Funktionskontrolle und ggf. Kalibrierung angemerkt, dass sich die Messwerte der funktionellen und fraktionellen Sauerstoffsättigung nur gering unterscheiden, solange der Anteil des dysfunktionellen Hämoglobins sehr niedrig im Verhältnis zum Gesamthämoglobin bleibt.

In der Medizin wird die speziell mit Pulsoximetern ermittelte Sauerstoffsättigung als SpO₂-Wert bezeichnet, anstelle der üblichen Bezeichnung (SaO₂) aller übrigen Messverfahren. Im Folgenden ist ausschließlich die Rede von SpO₂-Werten.

Der SpO₂-Wert kann physikalisch zwischen 0% und 100% liegen. Allerdings sind tiefere Werte als 70% sehr selten. Bei einem gesunden Erwachsenen liegt die Sauerstoffsättigung oberhalb von 90% [DOP]. Im oberen Bereich (70 bis 100%) werden die meisten Pulsoximeter mit einer Genauigkeit von +/- 2 Digits⁴ angegeben. Einige Hersteller von Pulsoximetern geben zusätzlich eine Genauigkeit von +/- 3 Digits im Bereich 50-70% an. Dies ergab eine Umfrage auf der Düsseldorfer Messe MEDICA 2004, bei der Hersteller ihre Pulsoximeter präsentierten. Grundsätzlich messen alle Geräte auch Sättigungen unterhalb von 70%, allerdings mit einem nicht spezifizierten Fehler. Zitat von Sigfried Kästle [Käs] bezüglich Auflösung und Genauigkeit von Pulsoximetern: „Eine Auflösung von 1% ist passend, denn sie zeigt noch kleine Trendänderungen an. Leider suggeriert sie eine Genauigkeit, die nicht erreicht wird.“

2.2 Photoelektrische Plethysmographie (PPG)

2.2.1 Definition

Unter photoelektrischer Plethysmographie versteht man die grafische Abbildung von Volumenänderungen in Blutgefäßen an einer bestimmten Stelle des Körpers. Jede Herzmuskelkontraktion führt zum Druckanstieg in den Arterien, über die kleineren Arteriolen, bis hin zu den feinsten arteriellen Kapillargefäßen. Die Wände arterieller Gefäße sind elastisch, was zur Erweiterung des Durchmessers dieser Gefäße mit jedem Pulsschlag und damit einer lokalen Blutvolumenänderung führt.

4 Ein Digit steht für die kleinste Schrittweite einer digitalen Anzeige.

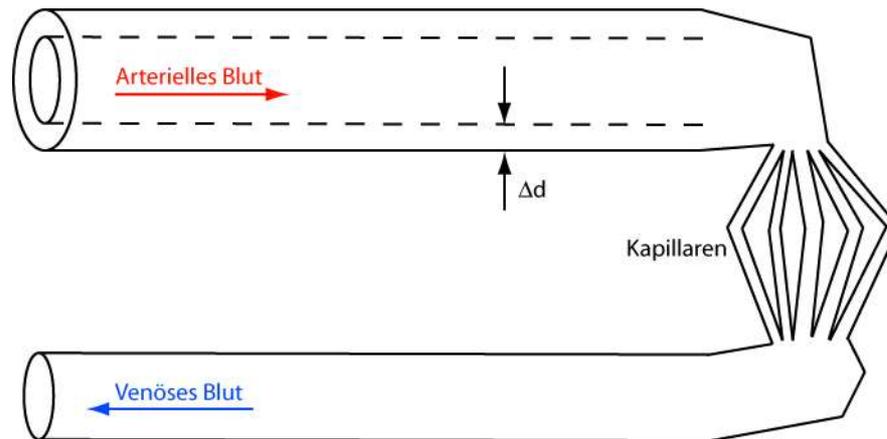


Abbildung 2.2: Schematisches Modell des Gefäßsystems

Die Volumenänderungen können mittels optischer Sensoren erfasst werden. Hierbei wird zwischen zwei Typen von Sensoren unterschieden: Den Transmissions- und den Remissions- bzw. Reflexions-Sensoren. Die Transmissions-Sensoren durchleuchten Gewebe und messen auf der gegenüberliegenden Seite die verbleibende Intensität, während Remissions-Sensoren in ein Körperteil hineinleuchten und das reflektierte Licht messen.

Die Plethysmographie ist die Vorstufe der Pulsoximetrie. Die hier erfassten Zusammenhänge zwischen den arteriellen Volumenänderungen und der Restintenzität dienen in Kapitel 2.3 als Berechnungsgrundlage für das Analytische Modell eines Pulsoximeters.

2.2.2 Transmissions-Sensoren

Für die weiteren Überlegungen betrachten wir das Modell aus Abb. 2.3, das einen am Finger angebrachten Transmission-Sensor repräsentiert. Hierbei handelt es sich um eine LED als konstante Lichtquelle und eine Photodiode, die auf der gegenüberliegenden Seite des Fingers die verbleibende Intensität erfasst. Durchstrahlt man Gewebe mittels monochromatischen Lichtes, so erfährt dieses Licht eine Dämpfung, die abhängig von der Wellenlänge ist. Das Licht, das auf der gegenüberliegenden Seite des Körperteils austritt, setzt sich aus einem konstanten und einem mit den Volumenänderungen der arteriellen Blutgefäße modulierten Anteil zusammen. Zeichnet man die resultierende Lichtintensität am Austrittsort über die Zeit auf, so erhält man damit ein sog. Plethysmogramm, das arteriellen Volumenänderungen an jener Körperstelle entspricht.

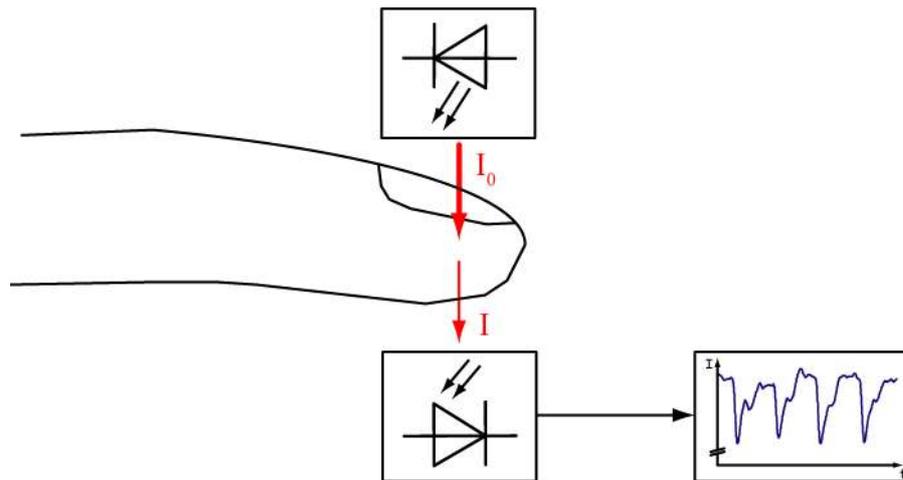


Abbildung 2.3: Modell der Transmission-Plethysmographie

Die Beziehung zwischen Anfangsintensität I_0 und der Restintensität hängt vom Absorptionsgrad der einzelnen Schichten im Gewebe ab. Diese Zusammenhänge lassen sich durch das Lambert-Beer'sche Gesetz beschreiben, allerdings nur unter der Annahme, dass keine Streuung im Gewebe vorliegt; tatsächlich sind Streuungseffekte vorhanden.

$$I = I_0 e^{-A} \quad (2.5)$$

- A: Absorptionsgrad
- I_0 : Auf das Gewebe einfallende Intensität
- I: Lichtintensität nach Durchqueren des Gewebes

Die einfallende Strahlung durchquert größtenteils Haut, Knochen, Bindegewebe, Fettgewebe, Muskelgewebe, venöses und arterielles Blut. Der Absorptionsgrad im Gewebe lässt sich als eine Überlagerung aus mehreren Teilabsorptionen der durchquerten Schichten beschreiben.

$$A = A_{Haut} + A_{Knochen} + A_{Bindegewebe} + A_{Fett} + A_{Muskel} + A_{Venen} + A_{Arterien} \quad (2.6)$$

Jede Teilabsorption einer Schicht i hängt von den molaren Extinktionskoeffizienten ϵ_i , deren Konzentration c_i und deren Schichtdicke x_i ab. Die molaren Extinktionskoeffizienten ϵ_i hängen von den spektralen Eigenschaften des Stoffes und der Wellenlänge der Strahlung, mit der der Stoff durchstrahlt wird, ab.

$$A_i = \varepsilon_{i,\lambda} c_i x_i \quad (2.7)$$

- A_i : Absorptionsgrad des Stoffes i
 ε_i : Extinktionskoeffizient des Stoffes i
 c_i : Konzentration des Stoffes i
 x_i : Schichtdicke des Stoffes i
 λ : Wellenlänge der Strahlung

Befindet sich der durchstrahlte Körperteil (z.B. Finger) in Ruhe, so sind alle Schichtdicken bis auf die des arteriellen Blutes relativ konstant. Die pulsatile Druckschwankungen in den Venen sind gegenüber denen in den Arterien vergleichsweise gering, so dass fast ausschließlich nur die Volumenänderungen des arteriellen Blutes zum pulsatilem Anteil des resultierenden Signals beitragen. Fassen wir die Gleichungen 2.5 bis 2.7 und die bisherigen Erkenntnisse zusammen, so können wir die minimale und maximale Intensität am Austrittsort jeweils für Diastole 2.8 und Systole 2.9 beschreiben.

$$I_{max} = I_0 e^{-\left(\sum_{i=1}^n \varepsilon_{i,\lambda} c_i x_i\right)} e^{-\left[(\varepsilon_{Hb,\lambda} c_{Hb} + \varepsilon_{O_2Hb,\lambda} c_{O_2Hb}) x_{min}\right]} \quad (2.8)$$

$$I_{min} = I_0 e^{-\left(\sum_{i=1}^n \varepsilon_{i,\lambda} c_i x_i\right)} e^{-\left[(\varepsilon_{Hb,\lambda} c_{Hb} + \varepsilon_{O_2Hb,\lambda} c_{O_2Hb}) x_{max}\right]} \quad (2.9)$$

- i : Über die Zeit konstante Schicht i
 I_0 : Auf die Probe einfallende Lichtintensität
 I_{max} : Lichtintensität nach Durchqueren aller Schichten während einer Diastole
 I_{min} : Lichtintensität nach Durchqueren aller Schichten während einer Systole
 x_i : Schichtdicke des Stoffes i
 x_{min} : Schichtdicke des arteriellen Blutes während einer Diastole
 x_{max} : Schichtdicke des arteriellen Blutes während einer Systole
 c_i : Konzentration des Stoffes i
 c_{Hb} : Konzentration des reduzierten Hämoglobins
 c_{O_2Hb} : Konzentration des oxygenierten Hämoglobins
 ε_i : Extinktionskoeffizient des Stoffes i
 ε_{Hb} : Extinktionskoeffizient des reduzierten Hämoglobins
 ε_{O_2Hb} : Extinktionskoeffizient des oxygenierten Hämoglobins

Der Verlauf der Intensität kann auch als Funktion der Zeit geschrieben werden, so dass zu jedem Zeitpunkt t die gemessene Intensität ermittelt werden kann.

$$I(t) = I_0 e^{-\left(\sum_{i=1}^n \varepsilon_{i,\lambda} c_i x_i\right)} e^{-\left[\left(\varepsilon_{Hb,\lambda} c_{Hb} + \varepsilon_{O2Hb,\lambda} c_{O2Hb}\right) x_A(t)\right]} \quad (2.10)$$

Abb. 2.4 zeigt den Verlauf der Intensität I am Austrittsort über die Zeit in Quantisierungsstufen eines AD-Wandlers. Deutlich erkennbar ist der geringe pulsatile Anteil (AC-Komponente) im Verhältnis zum konstanten Anteil (DC-Komponente). Je nach Messort liegt das Verhältnis AC/DC in der Größenordnung 0,01% bis 5%. Für einen maximalen Wechselanteil werden Messorte mit einem dichten Geflecht an Kapillaren an der Oberfläche der Haut bevorzugt. Dafür eignen sich Fingerkuppen und Ohrläppchen am besten.

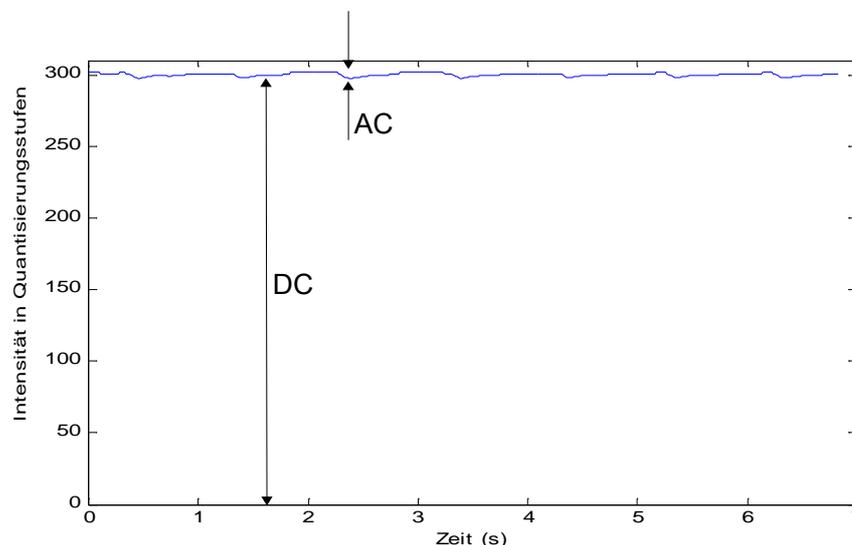


Abbildung 2.4: Verlauf der Intensität

Wegen dem kleinem AC/DC-Verhältnis wird in einem Plethysmogramm nur der Wechselanteil angezeigt. Für die Pulsoximetrie dagegen sind beide Teile (AC und DC) interessant. Die Gleichung 2.13 ist die Grundlage für ein analytisches Modell eines Pulsoximeters mit einem Transmissions-Sensor.

2.2.3 Remissions-Sensoren

An Fingerkuppen als auch an Ohrläppchen angebrachte Transmissions-Sensoren wirken störend. An Stellen wie Handrücken oder Handgelenk würden sie die Bewegungsfreiheit des Patienten weniger einschränken. Transmissions-Sensoren lassen sich auch an Handgelenken anbringen, sie benötigen allerdings ein Vielfaches an Strahlungsleistung, um die betreffenden Körperregionen zu durchleuchten. Ein Transmissions-Sensor in Form einer Uhr, die am

Handgelenk angebracht, zwischen Elle und Speiche hindurchleuchtet, wäre denkbar, ist jedoch wegen der hohen Leistungsaufnahme kaum vertretbar. An den meisten Stellen des menschlichen Körpers lassen sich keine Transmissions-Sensoren mit Wellenlängen zwischen 450nm und 960nm einsetzen, da die Absorption schon aufgrund der Schichtdicke zu hoch ist. Abhilfe schafft hier die Reflexions- bzw. Remissions-Plethysmographie (Abb. 2.5). Dabei dringen Photonen der LED zunächst in das Gewebe ein und werden beim Durchqueren der Schichten unterschiedlich absorbiert. Die Strahlung wird zum großen Teil an der Hautoberfläche, an Knochen und verbleibendem Gewebe reflektiert, legt wieder den selben Weg zurück, und kann dann auf derselben Körperseite gemessen werden.

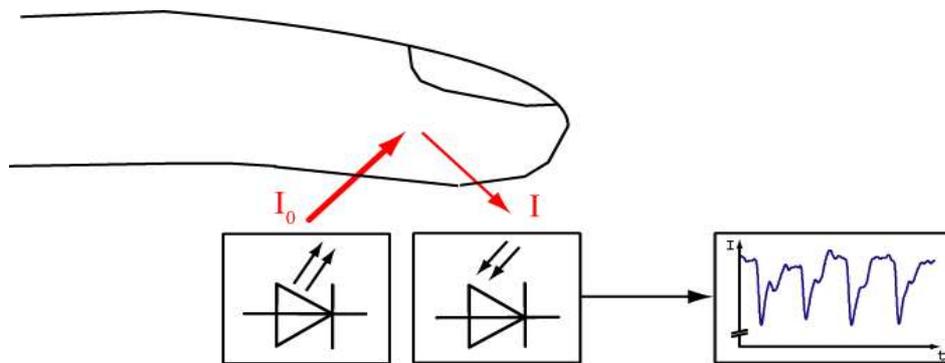


Abbildung 2.5: Modell der Reflexions- bzw. Remissions-Plethysmographie

In der Literatur lässt sich kein Nachweis finden, inwiefern das Verfahren der Transmission auf das der Remission tatsächlich übertragbar ist. Generell wird das Modell der Transmission für Reflexions-Sensoren direkt übernommen, mit der These, dass sich beide Verfahren nur in der Kalibrierungsfunktion und den Messorten unterscheiden. Dennoch sei der Fehler im SpO_2 -Wert größer als bei Transmission. In der Praxis seien Transmissions-Sensoren als zuverlässiger bekannt. [DOP] Für Transmissions-Sensoren wurde bereits der Verlauf der Intensität über die Zeit in Abhängigkeit von allen relevanten Faktoren durch Gleichung 2.10 definiert. Das gleiche soll für Reflexions-Sensoren geschehen, um zu klären, ob und wenn ja, inwiefern sich beide Verfahren unterscheiden.

Das Lambert-Beer'sche Gesetz gilt nur für Transmission, deshalb muss ein Gesetz für Reflexion hergeleitet werden. Dazu soll zunächst ein alternatives Modell für das Lambert-Beer'sche Gesetz erstellt werden, um daraus ein vergleichbares Modell für Reflexion abzuleiten.

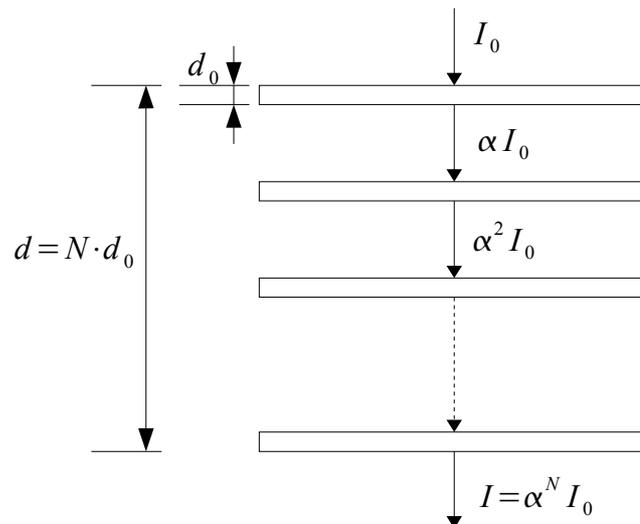


Abbildung 2.6: Das Schichtenmodell des arteriellen Blutes unter Verwendung von Transmissions-Sensoren

Das Modell der Transmission besteht aus N gleichen Schichten der Dicke d_0 . Nicht zu verwechseln mit den Gewebeschichten aus Kapitel 2.2.2. Alle N Schichten zusammen ergeben die Dicke d . Die Anfangsintensität I_0 wird nach Durchqueren der ersten Schicht um einen Faktor $\alpha < 1$ gedämpft. Nach Durchqueren der zweiten Schicht wird die Intensität αI_0 um den Faktor α gedämpft, was $\alpha^2 I_0$ ergibt. Nach N Schichten beträgt die Restintensität $I = \alpha^N I_0$. Der Zusammenhang zum Lambert-Beer'schen Gesetz kann durch die folgende Gleichung hergestellt werden:

$$I = I_0 \alpha^N = I_0 e^{N \ln(\alpha)} = I_0 e^{\frac{d \ln(\alpha)}{d_0}} = I_0 e^{-d c \epsilon_\lambda} \quad (2.11)$$

Zwischen der Konzentration und dem wellenlängenabhängigen Extinktionskoeffizienten aus dem Lambert-Beer'schen Gesetz lässt sich ein direkter Zusammenhang zu dem Dämpfungsfaktor α und der Schichtdicke d_0 aus dem Schichtenmodell der Transmission aufstellen:

$$-c \epsilon_\lambda = \frac{\ln(\alpha)}{d_0} \quad (2.12)$$

Passend dazu kann ein Schichtenmodell für Reflexion erstellt werden.

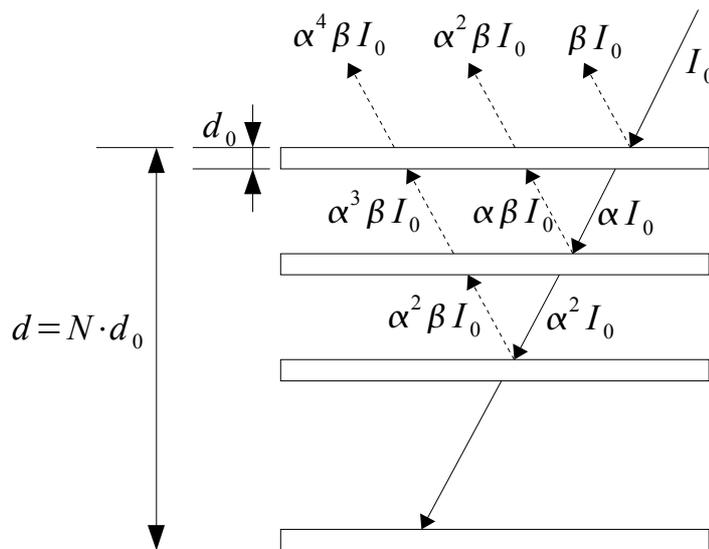


Abbildung 2.7: Das Schichtenmodell des arteriellen Blutes unter Verwendung von Remissions-Sensoren

Das Modell besteht wiederum aus N gleichen Schichten der Dicke d_0 . Alle N Schichten zusammen ergeben die Dicke d . Das Modell ist nach demselben Prinzip, wie das der Transmission aufgebaut, nur dass hier ein Teil der Intensität an den Schichten reflektiert wird. Die Anfangsintensität I_0 wird nach Durchqueren der ersten Schicht um einen Schwächungsfaktor $\alpha < 1$ gedämpft. Ein Teil wird an der Schicht mit dem Faktor β reflektiert. Die Restintensität αI_0 dringt in tiefere Schichten ein, wird mit Faktor β reflektiert und nimmt anschließend den gleichen Weg zurück usw. Die Restintensität lässt sich, unter der Bedingung, dass es sich nicht um kohärente Strahlung handelt, aufaddieren und als geometrische Reihe schreiben:

$$I = I_0 \beta + \alpha \beta I_0 + \alpha^2 \beta I_0 + \alpha^3 \beta I_0 + \dots + \alpha^{2(N-1)} \beta I_0 \quad (2.13)$$

Wird $I_0 \beta$ ausgeklammert ergibt sich:

$$I = I_0 \beta \sum_{k=0}^{N-1} \alpha^{2k} \quad (2.14)$$

Aus der Formelsammlung [For] geht hervor:

$$\sum_{k=0}^{N-1} x^k = \frac{1-x^N}{1-x} \quad (2.15)$$

Lässt sich umformen zu:

$$\sum_{k=0}^{N-1} x^{2k} = \frac{1-x^{2N}}{1-x^2} \quad (2.16)$$

Ein Vergleich von Gl. 2.14 und 2.16 ergibt:

$$I = I_0 \beta \frac{1 - \alpha^{2N}}{1 - \alpha^2} \quad (2.17)$$

Äquivalent zu Gleichung 2.11 kann man α^N entwickeln zu:

$$\alpha^{2N} = e^{2N \ln(\alpha)} = e^{\frac{2d}{d_0} \ln(\alpha)} \quad (2.18)$$

Mit Gleichung 2.12 lässt sich schreiben:

$$e^{\frac{2d}{d_0} \ln(\alpha)} = e^{-2dc\epsilon_\lambda} \quad (2.19)$$

Aus Gleichung 2.18 und 2.19 folgt:

$$\alpha^{2N} = e^{-2dc\epsilon_\lambda} \quad (2.20)$$

In Gleichung 2.17 eingesetzt:

$$I = I_0 \frac{\beta}{1 - \alpha^2} (1 - e^{-2dc\epsilon_\lambda}) \quad (2.21)$$

Schließlich kann man den unbekanntem, konstanten Teil bestehend aus dem Absorptionsfaktoren zu einer unbekanntem, wellenlängenabhängigen Konstante γ zusammenfassen:

$$I = I_0 \gamma_\lambda (1 - e^{-2dc\epsilon_\lambda}) \quad (2.22)$$

Gleichung 2.22 beschreibt die Intensität des durch Streuung an dem Hämoglobin reflektierten Lichtes, in Abhängigkeit von der Schichtdicke d des arteriellen Blutes. Weiterhin wird ein neues Modell des Gewebes zur Untersuchung der Remission betrachtet. Das Modell aus Abb. 2.8 besteht aus einer konstanten Gewebeschicht (z.B. Haut) und der veränderlichen Schicht (arterielles Blut). Das Licht, das an der ersten Schicht reflektiert wird, ist über die Zeit konstant, da sich die obere Schicht nicht verändert. Die verbleibende Intensität wird allein durch den wellenlängenabhängigen Schwächungsfaktor β' beschrieben. Das Licht, das die arteriellen Blutgefäße erreicht, und davon reflektiert wird, kann durch das Modell aus Abb. 2.7 beschrieben werden, und damit durch Gleichung 2.22. Da das Licht nicht kohärent ist, lassen sich die Intensitäten der konstanten und der über die Zeit veränderlichen Schicht aufaddieren:

$$I_R(t) = \beta'_\lambda I_{0,\lambda} + \gamma_\lambda I_{0,\lambda} (1 - e^{-2 \cdot (\epsilon_{Hb,\lambda} c_{Hb} + \epsilon_{O2Hb,\lambda} c_{O2Hb}) x_A(t)}) \quad (2.23)$$

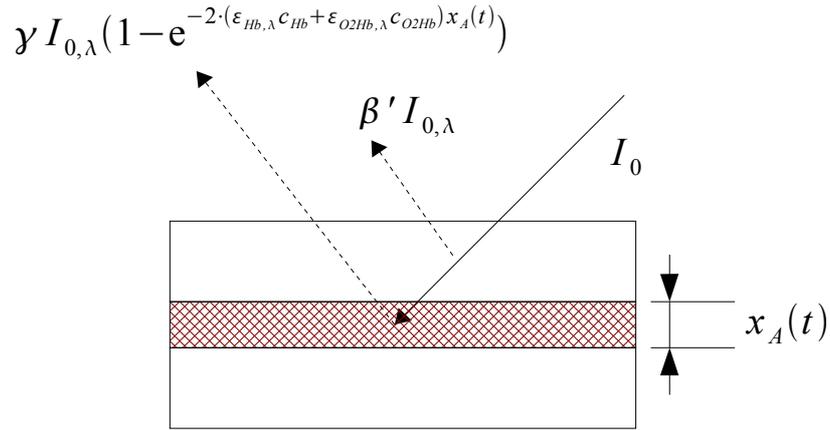


Abbildung 2.8: Das Schichtenmodell des Gewebes unter Verwendung von Remissions-Sensoren bestehend aus

Die funktionellen Absorptionsfaktoren des Hämoglobins werden zur besseren Übersicht zu einem Faktor α_{func} zusammengefasst:

$$\alpha_{\text{func}} = \varepsilon_{\text{Hb},\lambda} c_{\text{Hb}} + \varepsilon_{\text{O2Hb},\lambda} c_{\text{O2Hb}} \quad (2.24)$$

$$I_{\text{R}}(t) = \beta' I_{0,\lambda} + \gamma_{\lambda} I_{0,\lambda} (1 - e^{-2\alpha_{\text{func}} x_A(t)}) \quad (2.25)$$

Unter der Annahme, dass $-2\alpha_{\text{func}} x_A(t)$ für alle t viel kleiner als 1 bleibt, was Messungen belegten, lässt sich die e-Funktion linearisieren:

$$I_{\text{R}}(t) = \beta' I_{0,\lambda} + \gamma_{\lambda} I_{0,\lambda} (1 - (1 - 2\alpha_{\text{func}} x_A(t))) \quad (2.26)$$

$$I_{\text{R}}(t) = I_{0,\lambda} (\beta' + \gamma_{\lambda} 2\alpha_{\text{func}} x_A(t)) \quad (2.27)$$

Mit

$$A_{\lambda} = I_{0,\lambda} \beta' \quad (2.28)$$

$$B_{\lambda} = 2I_{0,\lambda} \gamma_{\lambda} \quad (2.29)$$

als Faktoren kann man die Intensität bei der Remission schreiben als:

$$I_{\text{R}}(t) = A_{\lambda} + B_{\lambda} \alpha_{\text{func}} x_A(t) \quad (2.30)$$

Äquivalent zur Remission kann die Gleichung der Transmission 2.10 in die gleiche Form gebracht werden. Gleichung 2.10 lautet mit der Zusammenfassung aus Gl. 2.24:

$$I_{\text{T}}(t) = I_{0,\lambda} e^{-\left(\sum_{i=1}^n \varepsilon_{i,\lambda} c_i x_i\right)} e^{-\alpha_{\text{func}} x_A(t)} \quad (2.31)$$

Weiterhin wird der zeitlich konstante Schwächungsanteil zusammen mit der konstanten Anfangsintensität zu einem wellenlängenabhängigen Faktor C zusammengefasst:

$$I_T(t) = C_\lambda e^{-\alpha_{\text{func}} x_A(t)} \quad (2.32)$$

Unter der Annahme, dass auch hier der Exponent viel kleiner als 1 für alle t bleibt, lässt sich die e-Funktion linearisieren.

$$I_T(t) = C_\lambda (1 - \alpha_{\text{func}} x_A(t)) \quad (2.33)$$

$$I_T(t) = C_\lambda - C_\lambda \alpha_{\text{func}} x_A(t) \quad (2.34)$$

Die vereinfachten Gleichungen der Remission 2.30 und der Transmission 2.34 unterscheiden sich in ihren Faktoren. In der Remission existiert neben dem Absorptionsfaktor B auch ein Faktor A, der von der Beschaffenheit der äußersten Schicht abhängt. Welche Folgen dieser Unterschied für die Pulsoximetrie bei Verwendung des Transformationsverfahrens für Remissions-Sensoren hat, wird in Kapitel 2.3.5.2 untersucht.

2.3 Pulsoximetrie (SpO₂)

2.3.1 Einführung

Mediziner bezeichnen die Entwicklung der sogenannten Pulsoximetrie als wichtigsten Fortschritt beim Monitoring⁵ von Patienten. Über die Haut werden mittels optischer Sensoren kontinuierlich sowohl Sauerstoffsättigung als auch Pulsfrequenz eines Patienten ermittelt und als Größen auf einer Anzeige dargestellt. Hierbei findet, wie bei der Plethysmographie, keinerlei mechanischer Eingriff in den Körper statt, wie es bei herkömmlichen invasiven Verfahren zur Ermittlung von Sauerstoffsättigung der Fall ist. Invasive Verfahren basieren auf stichprobenartigen Blutentnahmen, die im Labor ausgewertet werden müssen. Dies ist eine Belastung für Patienten und klinisches Personal. Zusätzlich dauert die Auswertung ab Blutentnahme bis zur Analyse im Labor sehr lange und ist dadurch nur bei Verdacht auf niedrige Sättigungen üblich. Wegen ihrer leichten Handhabung können Pulsoximeter bei jeder Gelegenheit angebracht werden. Die preiswerten Geräte vermindern durch verbreiteten Einsatz das Risiko einer unentdeckten Unterbrechung des Sauerstofftransports zu den Zellen, die schon über wenige Sekunden zum irreparablen Massensterben von Hirnzellen und damit zum Hirntod führen kann.

5 Die kontinuierliche Überwachung eines Patienten wird in der Medizin als Monitoring bezeichnet.

Ende der 80er Jahre hat das System zunächst in der Anästhesie Einzug gefunden, während heute die Pulsoximetrie allgemein etabliert und auch bei der Erstversorgung, Intensivmedizin und der postoperativen Patientenüberwachung unentbehrlich geworden ist.

2.3.2 Absorptionsspektren

Das grundlegende Prinzip der Pulsoximetrie liegt in der Absorptionsspektroskopie⁶. Diese wiederum beruht im Wesentlichen auf dem Lambert-Beer'schen Gesetz. Die unterschiedlichen Absorptionseigenschaften des reduzierten und oxygenierten Hämoglobins ermöglichen eine Bestimmung der relativen Konzentrationen von HHb und O₂Hb und damit die optische Messung der Sauerstoffsättigung. Hierbei baut das Messprinzip auf der photoelektrischen Plethysmographie auf, allerdings unter Verwendung mehrerer Wellenlängen. Zur Messung der funktionellen Sauerstoffsättigung (definiert in Kapitel 2.1.4) arbeiten Pulsoximeter mit zwei monochromatischen Lichtquellen verschiedener Wellenlängen. In der Regel verwenden Hersteller handelsüblicher Pulsoximeter die Wellenlängen 660nm (rot) und 940nm (infrarot) oder die alternative Kombination 660nm und 880nm. Aus dem Absorptionsspektrum in Abb. 2.9 lassen sich die Extinktionskoeffizienten zu jeder Wellenlänge für das reduzierte und oxygenierte Hämoglobin ablesen.

Das Wellenlängenpaar ist so gewählt, dass die Absorption beim Übergang von vollkommen reduziertem Hämoglobin (SpO₂ = 0%) zum vollkommen oxygeniertem Hämoglobin (SpO₂ = 100%) maximal variiert und zugleich die Extinktionskoeffizienten jeder Wellenlänge linear unabhängig sind. Um zugleich Toleranzen im Wellenlängenmaximum der LEDs zu kompensieren, werden Wellenlängen in einem flach verlaufenden Bereich im Absorptionsspektrum bevorzugt. Diese, und damit auch die Absorptionsgrade dürfen nicht zu hoch sein, so dass die Restintensität nach der Absorption noch messbar ist. Wegen der roten Hautpigmentierung ist die Absorption von Strahlung mit Wellenlängen unter 600nm zu stark. Das ist an den Extinktionskoeffizienten in diesem Bereich ersichtlich. Demnach sind Wellenlängen unterhalb dieser Grenze nicht geeignet.

Unter Verwendung einer der Wellenlängenpaare {660, 880} oder {660, 940} gilt für den funktionellen SpO₂-Wert, dass erhöhte COHb-Anteile zu einem fälschlich hohen SpO₂-Wert führen [AS]. Die Ursache ist, dass die Absorption des COHb bei 660nm (COHb nicht in Abb. 2.9) gleich der Absorption von O₂Hb ist, während im infraroten Bereich die niedrige Absorption des COHb einfach vernachlässigbar ist. Daraus lässt sich bei höheren COHb-Konzentrationen für den funktionellen SpO₂-Wert ein zunehmender Fehler mit sinkender Sauerstoffsättigung ableiten. A. Schmitz [AS] bezeichnet den Fehler bei hohen Sauerstoffsättigungen als minimal, während diesem bei niedrigen Sauerstoffsättigungen mehr Beachtung geschenkt werden sollte.

⁶ Absorptionsspektroskopie ist die Messung der Lichtschwächung, die von der gewählten Wellenlänge des Lichts und den Eigenschaften der untersuchten Substanz abhängt [PH].

2.3 Pulsoximetrie (SpO₂)

Die Kurven des HHb und O₂Hb kreuzen sich an den sog. isobestischen Punkten (z.B. 805nm) im Spektrum. Für die isobestischen Wellenlängen sind die Extinktionskoeffizienten des funktionellen Hämoglobins gleich und damit die Absorption bei dieser Wellenlänge unabhängig von der funktionellen Sauerstoffsättigung. Diese Wellenlängen werden oftmals genutzt um zusätzlich die relative Konzentration an Dyshämoglobin und somit die fraktionelle Sättigung zu ermitteln. Die Hinzunahme speziell eines der isobestischen Punkte zur Berechnung der fraktionellen Sauerstoffsättigung ist nicht zwingend notwendig, erleichtert aber die mathematische Herleitung.

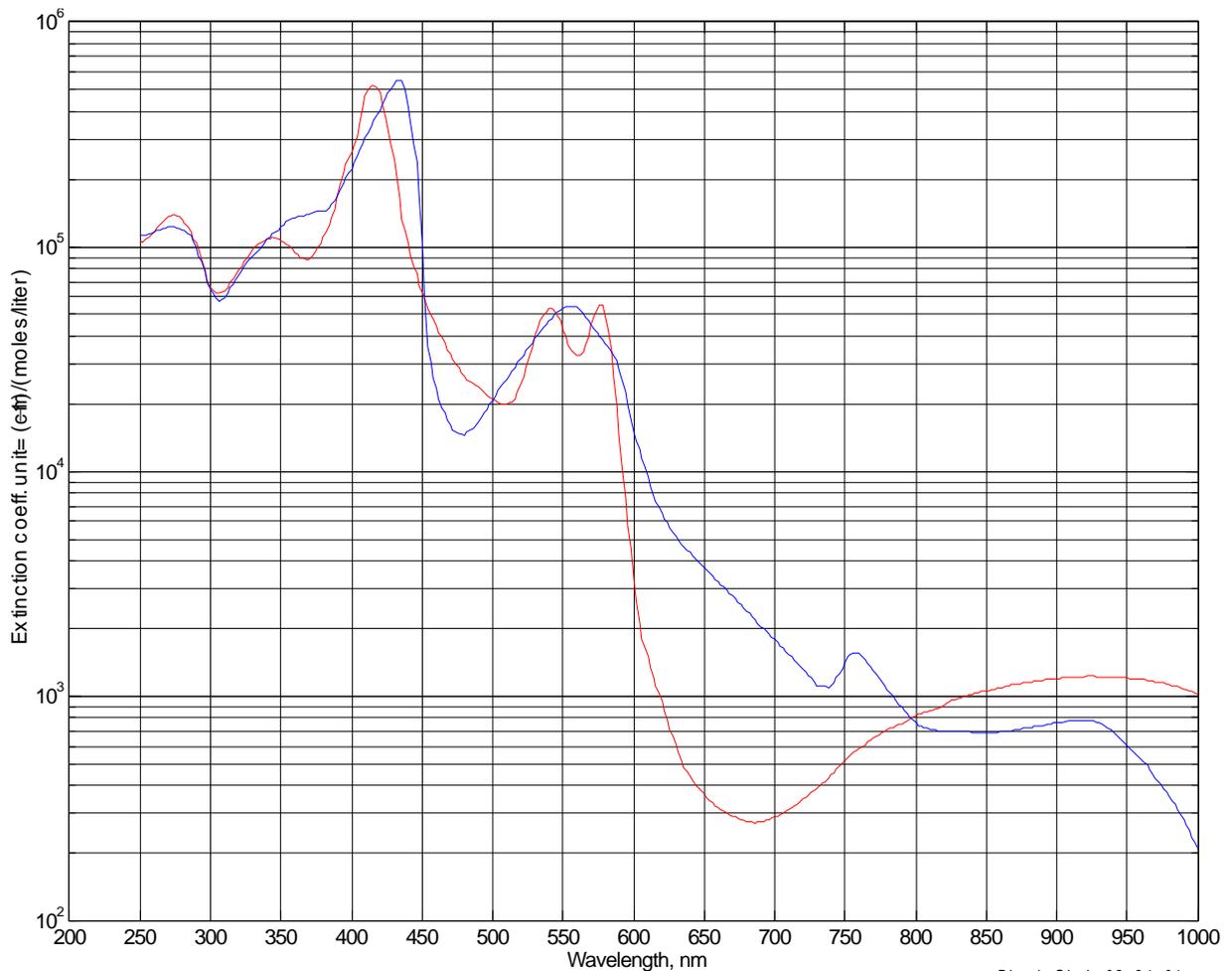


Abbildung 2.9: Absorptionsspektren des reduzierten (blau) und oxygenierten (rot) Hämoglobins
Quelle: Paper - Emory School of Medicine – Department of Anesthesiology
Pulse Oxygenation Falls with Arm Position Change
James E. Cooke, M.D.- Jay W. Johansen, M.D.

Bhavin Shah 08-24-01

2.3.3 Analytisches Modell für Transmission

Abb. 2.10 zeigt ein Modell zur Vermessung der Lichtabsorption über die Zeit in zwei Wellenlängen (rot und infrarot). Die beiden Plethysmogramme werden zur Berechnung des SpO₂-Wertes benötigt. In der Praxis ist es wichtig, dass die Leuchtdioden so nahe wie möglich aneinander liegen, damit das Licht beider LEDs denselben Weg nimmt. Die Photonen beider LEDs treffen nach wellenlängenabhängiger Absorption auf der Photodiode auf. Da die beiden Plethysmogramme zu jeder Wellenlänge separat benötigt werden, müssen die Intensitäten (Rot und Infrarot) mittels Modulation voneinander getrennt werden. Dabei steuert ein Mikrocontroller beide LEDs über Treiberstufen an und tastet gleichzeitig das analog aufbereitete Signal der Photodiode ab. Nach Demodulation sind die beiden Plethysmogramme (Rot und Infrarot) verfügbar.

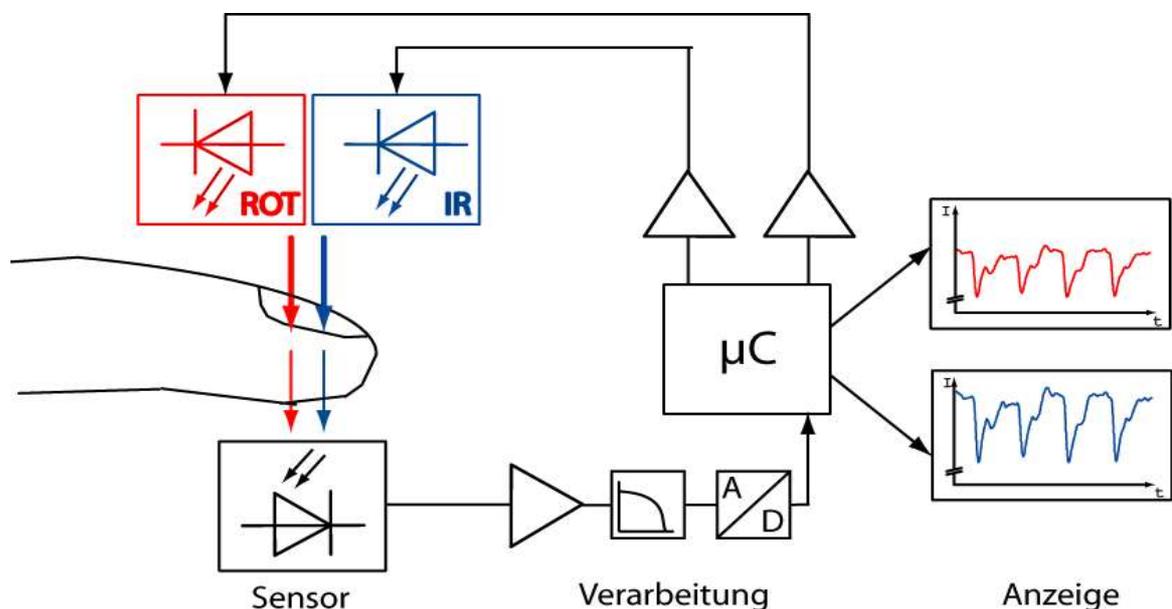


Abbildung 2.10: Modell zum Erfassen der Lichtabsorption zweier Wellenlängen. Ein Mikrocontroller übernimmt die Modulation und Demodulation und liefert jeweils ein Signal pro Wellenlänge. Die Graphen zeigen den Wechselanteil der Signale.

Eine wichtige Erkenntnis bei der Messung der arteriellen Sauerstoffsättigung, wie wir sie heute kennen, ist, dass nur die arteriellen Volumenänderungen zum pulsatilem Anteil im Plethysmogramm beitragen. Damit lassen sich Absorptionen, die durch arterielle Volumenänderung verursacht wurden, von allen anderen Gewebeschichten unterscheiden. Unter Berücksichtigung nur einer Wellenlänge, war es möglich, anhand der Gleichungen 2.8 und 2.9 zu zeigen, wie die gemessene Intensität von den Minima und Maxima der arteriellen Schicht abhängt. Diese Gleichungen beinhalten noch viele unbekannte Faktoren: z.B. die einstrahlende Intensität I_0 der LED. Die Empfindlichkeit der Photodiode ist abhängig von der Wellenlänge und damit zunächst nicht bekannt. Zudem sind die Absorptionseigenschaften der Gewebeschichten von Patient zu Patient unterschiedlich. Das Signal muss normiert werden, um diese Faktoren zu eliminieren. Betrachten wir dazu allein die arterielle Schichtdickenänderung

$\Delta d := d_{\max} - d_{\min}$ unter der Annahme, dass sich die Schichtdicke aller anderen Gewebeanteile nicht verändert hat, können wir den Quotienten der Gleichungen 2.8 und 2.9 bilden:

$$\frac{I_{\min}}{I_{\max}} = \frac{e^{-[(\epsilon_{Hb} c_{Hb} + \epsilon_{O_2Hb} c_{O_2Hb}) d_{\max}]}}{e^{-[(\epsilon_{Hb} c_{Hb} + \epsilon_{O_2Hb} c_{O_2Hb}) d_{\min}]}} \quad (2.35)$$

Damit können wir durch Normierung die unbekannt Parameter der konstanten Gewebeschichten, der unbekannt einstrahlenden Intensität I_0 der LED und der spektralen Nichtlinearität der Photodiode eliminieren. Weiterhin kann Gleichung 2.35 mit $\Delta d = d_{\max} - d_{\min}$ umgeformt werden zu:

$$N := \ln\left(\frac{I_{\min}}{I_{\max}}\right) = -(\epsilon_{Hb} c_{Hb} + \epsilon_{O_2Hb} c_{O_2Hb}) \Delta d \quad (2.36)$$

Mit 2.36 wird es möglich, die Konzentrationen als Funktion der minimalen und maximalen Intensitäten und der unbekannt arteriellen Schichtdickenänderung Δd der arteriellen Gefäße zu schreiben. Die ersten Pulsoximeter arbeiteten nicht nach diesem Prinzip und mussten deshalb für jeden Patienten neu kalibriert werden.

Unter Verwendung einer zweiten Wellenlänge kann der letzte unbekannt Parameter Δd aus der Gleichung 2.9 eliminiert werden. Dazu dient das zweite Plethysmogramm, aus dem sich äquivalent die Intensitätsextrema I_{\min} und I_{\max} ablesen lassen. Die Gleichung 2.9 wird zwei mal benötigt und deshalb mit dem Index der Wellenlänge versehen, um die so entstandenen Gleichungen voneinander unterscheiden zu können.

Für die Normierte der erste Wellenlänge (z.B. rot) gilt:

$$N_{\lambda 1} := \ln\left(\frac{I_{\min, \lambda 1}}{I_{\max, \lambda 1}}\right) = -(\epsilon_{Hb, \lambda 1} c_{Hb} + \epsilon_{O_2Hb, \lambda 1} c_{O_2Hb}) \Delta d \quad (2.37)$$

Für die Normierte der zweiten Wellenlänge (z.B. infrarot) gilt:

$$N_{\lambda 2} := \ln \left(\frac{I_{min, \lambda 2}}{I_{max, \lambda 2}} \right) = -(\varepsilon_{Hb, \lambda 2} c_{Hb} + \varepsilon_{O2Hb, \lambda 2} c_{O2Hb}) \Delta d \quad (2.38)$$

Mit dem Verhältnis ρ aus den Gleichungen 2.37 und 2.38 ergibt sich ein Zusammenhang zwischen den vier bekannten Intensitätsextrema aus den beiden Plethysmogrammen und den zwei Konzentrationen des funktionellen Hämoglobins.

$$\rho := \frac{\ln \left(\frac{I_{min, \lambda 1}}{I_{max, \lambda 1}} \right)}{\ln \left(\frac{I_{min, \lambda 2}}{I_{max, \lambda 2}} \right)} = \frac{(\varepsilon_{Hb, \lambda 1} c_{Hb} + \varepsilon_{O2Hb, \lambda 1} c_{O2Hb})}{(\varepsilon_{Hb, \lambda 2} c_{Hb} + \varepsilon_{O2Hb, \lambda 2} c_{O2Hb})} \quad (2.39)$$

Unter der Annahme, dass das Verhältnis der Extrema aufgrund des hohen DC-Anteils (Kapitel 2.2.2) nahe an 1 liegt, kann der Logarithmus linearisiert werden. Weiter lässt sich das Signal durch einen Wechselanteil $AC = I_{max} - I_{min}$ und einen Gleichanteil $DC = (I_{min} + I_{max})/2$ beschreiben [Käs].

$$\rho = \frac{\frac{AC_{\lambda 1}}{DC_{\lambda 1}}}{\frac{AC_{\lambda 2}}{DC_{\lambda 2}}} \quad (2.40)$$

Mit dem Verhältnis ρ aus Gleichung 2.39 und der Gleichung der funktionellen Sauerstoffsättigung 2.41

$$SpO_2 = \frac{c_{O2Hb}}{c_{O2Hb} + c_{HHb}} \cdot 100\% \quad (2.41)$$

lässt sich die funktionelle Sauerstoffsättigung als Funktion von ρ schreiben:

$$SpO_2(\rho) := \frac{\rho \cdot \varepsilon_{Hb, \lambda 2} - \varepsilon_{Hb, \lambda 1}}{\rho \cdot (\varepsilon_{Hb, \lambda 2} - \varepsilon_{O2Hb, \lambda 2}) + \varepsilon_{O2Hb, \lambda 1} - \varepsilon_{O2Hb, \lambda 1}} \quad (2.42)$$

2.3.4 Kalibrierung

Die Extinktionskoeffizienten aus Gleichung 2.39 könnte man theoretisch aus Abb. 2.9 entnehmen. In der Praxis führt die als nicht vorhanden angenommene Streuung zu Abweichungen in der Messung der Sauerstoffsättigung mittels Transmission nach einem reinen analytischen Modell. Abgesehen davon strahlen LEDs kein monochromatisches Licht ab. Dieser Fehler bleibt allerdings nur systemabhängig und variiert nicht von Patient zu Patient. Demnach ist lediglich eine einmalige Kalibrierung des Gerätes mit Sensor erforderlich. Ersetzt man die Extinktionskoeffizienten durch Konstanten, so erhält man eine Gleichung mit vier unbekanntem Konstanten für die Sauerstoffsättigung als Funktion des Verhältnisses R.

$$SpO_2(\rho) = \frac{\rho \cdot k_1 - k_2}{\rho \cdot k_3 + k_4} \quad \text{mit } k_i \in \mathbb{R} \quad (2.43)$$

Die Konstanten k_i können mit Hinzunahme invasiver Verfahren als Referenz empirisch ermittelt werden. Da jede empirisch ermittelte Kalibrierungsfunktion im Intervall

$$SpO_2 \in [0,100] \quad (2.44)$$

stets stetig ist, somit kein Pol existiert, können die unbekanntem Konstanten auf drei reduziert werden. Man erhält damit für die Sauerstoffsättigung eine Parabel als Funktion des Verhältnisses R:

$$SpO_2(\rho) = \rho^2 \cdot c_1 + \rho \cdot c_2 + c_3 \quad \text{mit } c_i \in \mathbb{R} \quad (2.45)$$

2.3.5 Transmissionsmodell mit Remissions-Sensoren

Um die Frage zu beantworten, inwiefern das Verfahren der Transmission im Einsatz mit Remissions-Sensoren übertragbar ist, müssen die vereinfachten Gleichungen der Remission 2.30 und der Transmission 2.34 jeweils in das analytische Modell der Transmission eingesetzt werden. Mit den beiden Gleichungen 2.30 und 2.34 wird zum Vergleich jeweils die Normierung nach dem Modell der Transmission durchgeführt. Dazu müssen die Gleichungen in ihren AC-Anteil 2.46 und DC-Anteil 2.47 aufgeteilt werden.

2.3.5.1 Mit der Gleichung für Transmission 2.34

$$AC = - C_{\lambda} \alpha_{\text{func}} x_A(t) \quad (2.46)$$

$$DC = C_{\lambda} \quad (2.47)$$

Normierung:

$$N = \frac{AC_{\lambda}}{DC_{\lambda}} = \frac{-C_{\lambda} \alpha_{\text{func}} x_A(t)}{C_{\lambda}} = -\alpha_{\text{func}} x_A(t) \quad (2.48)$$

Verhältnis ρ nach Gleichung 2.40:

$$\rho = \frac{-\alpha_{\text{func},\lambda 1} x_A(t)}{-\alpha_{\text{func},\lambda 2} x_A(t)} = \frac{\alpha_{\text{func},\lambda 1}}{\alpha_{\text{func},\lambda 2}} \quad (2.49)$$

2.3.5.2 Mit der Gleichung für Remission 2.30

$$AC = B_{\lambda} \alpha_{\text{func}} x_A(t) \quad (2.50)$$

$$DC = A_{\lambda} \quad (2.51)$$

Normierung:

$$N = \frac{AC_{\lambda}}{DC_{\lambda}} = \frac{B_{\lambda}}{A_{\lambda}} \alpha_{\text{func}} x_A(t) \quad (2.52)$$

In der Gleichung der Remission existieren zwei wellenlängenabhängige Faktoren A und B, die nach Einsetzen in die Gleichung der Normierung nicht wegfallen. Der Quotient aus B und A wird zu einem wellenlängenabhängigen Faktor D zusammengefasst.

$$D_{\lambda} = \frac{B_{\lambda}}{A_{\lambda}} \quad (2.53)$$

$$N = D_{\lambda} \alpha_{\text{func}} x_A(t) \quad (2.54)$$

Verhältnis ρ nach Gleichung 2.40:

$$\rho = \frac{D_{\lambda 1} \alpha_{\text{func}, \lambda 1}}{D_{\lambda 2} \alpha_{\text{func}, \lambda 2}} \quad (2.55)$$

Demnach kommt beim Einsatz von Remissions-Sensoren ein Faktor hinzu, der von den Absorptionseigenschaften der ersten Schicht, an der das Licht reflektiert wird, abhängig ist. Die Absorptionseigenschaften könnten je nach Messort und Hauttyp unterschiedlich sein. Die Abhängigkeit des Verhältnisses ρ von den Absorptionseigenschaften der ersten Hautschichten könnte eine Kalibrierung je nach Messort und Hauttyp erfordern. Dieser Faktor kann nur unter der Bedingung

$$D_{\lambda 1} \simeq D_{\lambda 2} \quad (2.56)$$

vernachlässigt werden. Das bedeutet, dass die Absorptionseigenschaften der ersten Schicht unabhängig von den eingesetzten Wellenlängen sein sollten; bzw. die Absorptionseigenschaften sollten möglichst gleich für beide Wellenlängen sein. Dann lässt sich das Modell der Transmission als Näherung für Remissions-Sensoren verwenden.

$$\rho = \frac{D_{\lambda 1} \alpha_{\text{func}, \lambda 1}}{D_{\lambda 2} \alpha_{\text{func}, \lambda 2}} \approx \frac{\alpha_{\text{func}, \lambda 1}}{\alpha_{\text{func}, \lambda 2}} \quad (2.57)$$

2.3.6 Einflüsse auf die Messgenauigkeit

2.3.6.1 Umgebungslicht

Umgebungslicht kann in alternierende Lichtquellen (wie die Operationstischbeleuchtung [DOP], Monitore und Leuchtstofflampen) oder konstante Lichtquellen (wie das Sonnenlicht) unterteilt werden. Die meisten Sensoren sind gegen Einstrahlung geschützt. Dennoch dringt Licht in die Schutzhülle der Sensoren ein. Das Sonnenlicht kann mit seiner hohen Intensität zum Übersteuern von analogen Verstärkern führen. Ein primäres Problem sind alternierende Lichtquellen.

2.3.6.2 Elektromagnetische Felder

Grundsätzlich stellen alle elektromagnetischen Felder ein Problem für Pulsoximeter dar. Vor allem beim Einsatz der Magnetresonanztomographie zeigen Pulsoximeter falsche Werte [DOP] an.

2.3.6.3 HF-Chirurgie

In der HF-Chirurgie wird Gewebe mittels hochfrequenter Spannung geschnitten [ME]. Diese hohen Frequenzen verursachen Induktionen in Leitungen der Pulsoximeter und können so deren Funktion stören.

2.3.6.4 Nagellack

Blauer, grüner, brauner und schwarzer Nagellack muss entfernt werden [DOP]. Diese Farben absorbieren das rote und infrarote Licht zu stark. Bei rosa oder rotem Lack gibt es weniger Probleme [AS].

2.3.6.5 Bewegungsartefakte

Schon kleinste Bewegungen führen im Verhältnis zum pulsatilen Anteil zu extrem hohen Ausschlägen im Sensorsignal. Bewegungsartefakte müssen mit allen Mitteln der DSV bekämpft werden.

2.3.6.6 Minderperfusion

Niedrige Durchblutung hat viele Ursachen und führt zu einem schwachen pulsatilen Signalanteil, so dass die Bestimmung der Sauerstoffsättigung schwierig oder nicht mehr möglich wird. Vor allem in Schockzuständen oder Unterkühlung, gerade dann, wenn es nötig wird ein Pulsoximeter anzulegen, sinkt die Durchblutung der Extremitäten, wie z.B. der Arme bis zu den Fingern.

2.3.6.7 Krankheitsbilder

Anämie ist ein Krankheitsbild, das auf einer Missbildung der Erythrozyten zurückzuführen ist. Das Absorptionsspektrum missgebildeter Erythrozyten unterscheidet sich von dem gesunder Erythrozyten. Die Farbänderung führt zu falschen SpO₂-Werten.

Herzrhythmus-Störungen können bewirken, dass der Algorithmus, keine harmonische Beziehung der Frequenzen des pulsatilen Signals im Spektrum findet, was zu Fehlern in der Messgröße oder im Extremfall zum Ausfall der Anzeige führen kann.

2.3.6.8 Hautpigmentierung

Mit dem hier vorgestellten analytischen Modell (mit Normierung nach Formel 2.36) beeinflusst die Hautpigmentierung die SpO₂-Werte nicht. Seit dem Aufkommen von Reflexions-Sensoren ist die Hautpigmentierung wieder im Gespräch. Bisher existiert kein von der Hautpigmentierung unabhängiges Modell für die Reflexion. Das Modell der Transmission kann nur mit Bedacht in der Reflexions-Pulsoximetrie eingesetzt werden.

2.3.6.9 Dysfunktionelles Hämoglobin

Vor allem bei Rauchern führt eine hohe Konzentration an COHb zu einer fälschlich hohen funktionellen Sauerstoffsättigung. Hohe Konzentrationen an MetHb sind dagegen sehr selten und tragen zu einem fälschlich niedrigen SpO₂-Wert bei [AS].

2.4 Multiplex-Verfahren

In der Pulsoximetrie verfügen Transmissions- und Remissions-Sensoren über mehrere LEDs (Sender) und nur eine Photodiode (Empfänger). Vergleichen lässt sich das Modell eines Pulsoximeters (Abb. 2.10) mit mehreren Radiostationen, die sich einen Kanal teilen, und mit einem einzigen Rundfunkempfänger, der in der Lage sein sollte jede Station separat zu empfangen. Der einzige Unterschied ist, dass die Radiostationen eine Information senden, und deshalb ihren Träger schon in der Sendestation modulieren, während das Pulsoximeter die Vermessung des Kanals mit unterschiedlichen Wellenlängen vornimmt. Der Träger wird also erst im Kanal mit dem pulsatilen Anteil nach dem Lambert-Beer'schen Gesetz (Gleichung 2.5) moduliert. Prinzipiell wäre es möglich die Kanäle anhand der Wellenlängen zu unterscheiden, z.B. durch Farbfilter an mehreren Photodioden. Da dies auf der Seite der Photodiode technisch aufwändig ist, müssen die Intensitäten der LEDs moduliert werden. Nur dann ist eine Unterscheidung der Wellenlängen mittels einer einzigen breitbandigen Photodiode möglich.

Man kann sagen, dass zu diesem Zeitpunkt alle bekannten Pulsoximeter TDMA einsetzen. Dabei wird jeder Sensor-LED ein Zeitfenster zugewiesen, in dem diese eingeschaltet wird. Nacheinander leuchten die LEDs für eine kurze Zeit auf. Oftmals versucht man durch Hinzufügen zusätzlicher Dunkelphasen, in denen keine der LEDs leuchtet, den durch Umgebungslicht verursachten Signalanteil zu messen und anschließend vom Nutzsignal zu subtrahieren. Einer der Vorteile dieses Verfahrens ist die einfache Implementierung des Verfahrens, wenn Sender und Empfänger einmal synchronisiert sind. Dies ist in der Pulsoximetrie der Fall, da nur ein zentraler Mikrocontroller für LEDs und Photodiode eingesetzt wird. Ein weiterer Vorteil von TDMA ist, dass zum Ansteuern der LEDs nur zwei Intensitäten genügen (ein/aus). Dies erspart die D/A-Wandler und vereinfacht den Bau der LED-Treiberstufen, da nunmehr der nichtlinearen Strom-zu-Intensität-Kennlinie der LEDs keine Beachtung geschenkt werden muss.

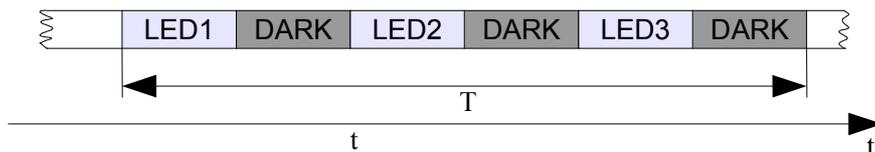


Abbildung 2.11: Zeitfenster in TDMA, wie es verbreitet in Pulsoximetern eingesetzt wird.

In der Literatur ist oftmals die Rede von Abweichungen der SpO_2 -Werte von der tatsächlichen Sauerstoffsättigung, die durch Umgebungslicht [AS] oder Hochfrequenz-Chirurgie [Käs] verursacht worden sind. In der HF-Chirurgie wird Gewebe mittels hochfrequenter Spannung geschnitten [ME]. Diese hohen Frequenzen verursachen Induktionen in Leitungen der Pulsoximeter und können so deren Funktion stören. Die optischen Störeinflüsse können weitgehend unterdrückt werden, da die Sensoren gegen Einstrahlung von außen geschützt sind. Dennoch tritt Umgebungslicht in die Hülle des Sensors ein. Die Subtraktion des Anteils an Umgebungslicht, durch hinzufügen von Dunkelphasen, verbessert deutlich die Signalqualität. Allerdings verbleiben Störungsartefakte, die zu einem falschen SpO_2 -Wert führen. Bislang wurden zahlreiche Versuche unternommen, die durch Leuchtstofflampen, Infrarotwärmelampen, Operationstischbeleuchtung und Monitore verursachten Störungen aus dem Nutzsignal zu entfernen.

Eine Möglichkeit die Störungen im Voraus zu unterbinden, ist die Wahl eines effizienteren Modulationsverfahrens. Ein solches Verfahren ist beispielsweise CDMA, das bereits in den modernen Mobilfunksystemen wie UMTS eingesetzt wird. CDMA ist eine speziell für den Mobilfunk entwickelte Form der Spreizspektrum-Modulation. Die fundamentalen Ansätze dieser Technologie wurden schon Mitte der 40er Jahre vom Militär erforscht [KS]. Es bestand ein Interesse an einem Modulationsverfahren, das die Energie, die in einem Signal mit einer Information steckt, gleichmäßig auf einen breiten Frequenzbereich zu verteilen, anstatt das Signal einfach aus dem Basisband zu verschieben, wie es bei AM der Fall ist. Wenn die Gesamtenergie auf einen größeren Bereich verteilt wird, so bleibt das Spreizsignal im Spektrum eher unentdeckt als ein amplituden- oder frequenzmodulierter Sinus-Träger. Dadurch, dass die Information auf ein breites Spektrum verteilt wird, tragen Störsignale weniger zu einem Fehler im Nutzsignal bei. Am besten lässt sich das Prinzip des Spreizens als Vergleich zur Amplitudenmodulation erklären.

Betrachten wir das Spektrum des Signals im Basisband, wie es bei der reinen Plethysmographie (vgl. Kapitel 2.2.2) entstehen könnte. Zu erwarten ist ein hoher Gleichanteil, und im Verhältnis dazu ein kleiner pulsatile Anteil, der aus einer Grundwelle und einigen schnell abklingenden Oberwellen besteht. Abb. 2.12 zeigt, welchen Bereich bis zur Frequenz f_b das Signal im Basisband einnehmen könnte.

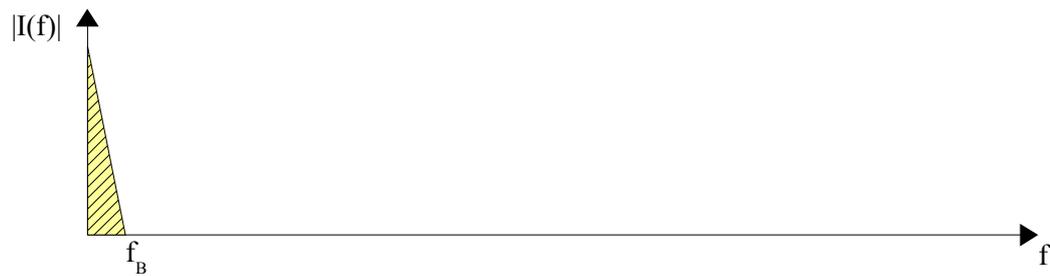


Abbildung 2.12: Spektrum des Signals im Basisband für positive Frequenzen

Versuchen wir einen der Kanäle auf einen Sinus-Träger zu modulieren. Dazu muss die Intensität der LED sinusförmig bei einer Frequenz f_{AM} oszillieren. Die dafür notwendige Beimischung des Nutzsignals läuft nach dem Lambert-Beer'schen Gesetz im Gewebe ab.

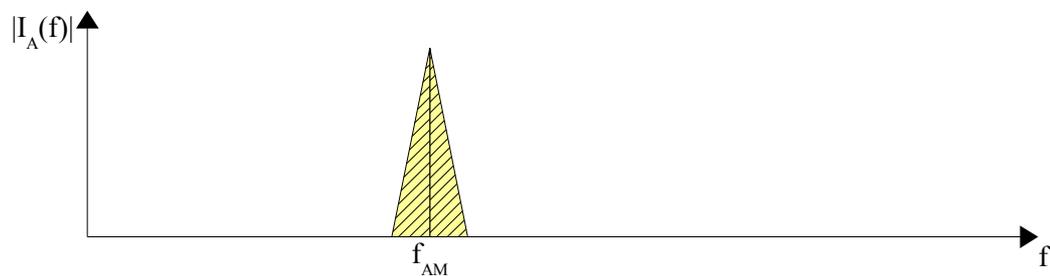


Abbildung 2.13: Spektrum des amplitudenmodulierten Signal für positive Frequenzen

Wir erhalten eine Verschiebung des Signals aus dem Basisband. Unter Hinzunahme einer zweiten Trägerfrequenz lassen sich die Signale der zwei LEDs im Spektrum unterscheiden. Allerdings nur, wenn nicht gerade eine Störung mit einer Frequenz von $f_{AM}-f_B$ bis $f_{AM}+f_B$ zum Signal addiert wird.

Der Zweck der Spreizspektrum-Modulation ist, die Information auf das gesamte Spektrum möglichst gleichmäßig zu verteilen. Dazu wird das Nutzsignal mit einem scheinbar stochastischen Träger-Signal multipliziert, das einem Rauschen ähnelt. Da die Modulation diskret abläuft, handelt es sich bei der Trägerfunktion um eine endliche Folge von Zahlen, die endlos wiederholt wird. Folgen dieser Art besitzen besondere spektrale Eigenschaften. So besteht das Spektrum eines solchen Signals aus vielen äquidistanten Peaks⁷ gleicher Länge und niedriger Leistung. Wird das Nutzsignal mit einem solchen Träger multipliziert, so wird dieses Signal auf viele kleine Träger amplitudenmoduliert.

⁷ Ein Peak bezeichnet ein spitzenartiges Maximum

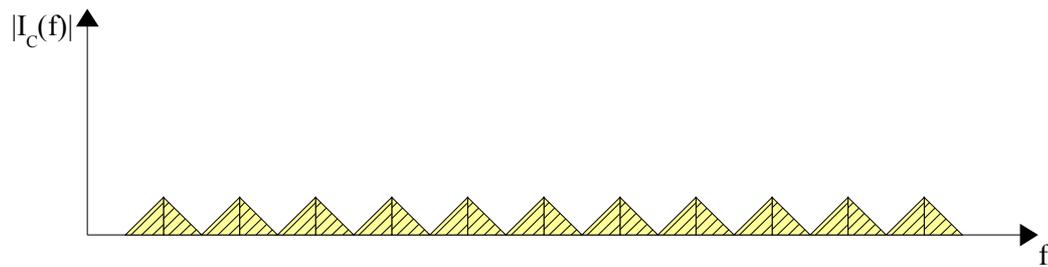


Abbildung 2.14: Spektrum des gespreizten Signals für positive Frequenzen

Um so ein Signal demodulieren oder ins Basisband entspreizen zu können, muss der Empfänger über dieselbe Träger-Funktion verfügen und zeitsynchron mit dem Sender arbeiten. Die Vorteile des Verfahrens sind anhand des Spektrums ersichtlich, denn unabhängig von dem Frequenzbereich einer Störung, bleiben einige Träger zur Rekonstruktion des Nutzsignals übrig. Auch im Zeitmultiplex wird ein Basisbandsignal auf ein Spektrum verteilt. Allerdings nicht gleichmäßig, so dass vor allem die stärker gewichteten Träger an einer Stelle im Spektrum liegen, und dadurch das Übertragungssystem stör anfälliger gegenüber schmalbandigen Störungen wird. Aus diesem Grund soll der Versuch gestartet werden, die Störungen im Vorfeld durch Einsatz eines der Spreizspektrum-Modulationsverfahren zu unterdrücken.

Wenn der gesamte Frequenz- und Zeitbereich mit einem Spreizsignal belegt ist, muss man sich die Frage stellen, wie es möglich ist, zwei solcher Sender unter diesen Voraussetzungen in einem Kanal unterscheiden zu können, so dass eine Rekonstruktion beider Signale möglich wird. Dazu müssen die Folgen, die zur Modulation mit den Nutzsignalen benötigt werden, orthogonal zueinander sein. Es gibt verschiedene Ansätze entsprechende Folgen zu generieren. In der Nachrichtentechnik verwendet man vorwiegend endliche Folgen, diese werden Chip-Folgen genannt. Jede Chip-Folge kann für einen Kanal verwendet werden. Die Fachliteratur [Sch] [KS] beschreibt bereits eine Reihe von Gesetzen, um Folgen mit speziellen Eigenschaften zu bilden. Mehrere orthogonale Folgen werden zu einer Familie zusammengefasst.

Grundsätzlich muss eine Folge c_i mit allen anderen Folgen aus ihrer Familie orthogonal sein, damit ein Übersprechen zwischen den Kanälen verhindert wird. Es muss also für die Skalarprodukte einer Familie mit N Folgen gelten:

$$\langle c_i, c_j \rangle = 0, \quad i \neq j; \quad \forall i, j \quad (2.58)$$

Weiterhin sollte die Autokorrelationsfunktion Φ_{ii} aller Folgen c_i ein deutliches Maximum für den ersten Koeffizienten ergeben und damit folgende Bedingung erfüllen:

$$\Phi_{ii}(0) \gg \Phi_{ii}(k), \quad k \neq 0; \quad \forall k; \quad \forall i \quad (2.59)$$

Welche Eigenschaften für die speziell hier vorgestellten Anwendungen von Vorteil sind und welche Folgen-Familie schließlich gewählt wird, ist das Thema von Kapitel 4.2.1.1.

3 Systemüberblick

Das gesamte System vom Sensor bis zur Anzeige der Vitalparameter lässt sich in separate Einheiten unterteilen. Ein Teil der Verarbeitung läuft in Echtzeit auf dem TriCore – Mikrocontroller, ein anderer Teil wiederum auf dem PC ab. Abb. 3.1 veranschaulicht die Beziehung aller Hardware- und Software-Komponenten, die in den folgenden Kapiteln 4, 5 und 6 näher erläutert werden.

Der Sensor (MORES) steht am Anfang der Verarbeitungsschritte und ist direkt mit dem analogen Modul verbunden. Dieses Modul ist ein steckbarer Aufsatz für die TriCore-Entwicklungsumgebung (TriBoard) und enthält drei LED-Treiber, sowie die analoge Aufbereitung des Sensorsignals zur Digitalisierung. Details zur Hardware sind Inhalt von Kapitel 5.

Der TriCore-Mikrocontroller, der sich auf dem TriBoard-Evaluation-Kit befindet, übernimmt die Verarbeitung in Echtzeit. Über die LED-Treiber werden die Sensor-LEDs mit beliebigen binären Folgen getaktet. Gleichzeitig wird das analog aufbereitete Signal mittels des integrierten A/D-Wandlers digitalisiert. Damit steht das Sensor-Signal in diskreter Form zunächst im Speicher des Mikrocontrollers zur Verfügung. Die Rohdaten können zur weiteren Verarbeitung über RS232 an den PC gesandt werden. Optional kann das Sensor-Signal demoduliert und zum PC gesandt werden. In Kapitel 6 wird dieses Thema behandelt. Auf der PC-Seite nimmt ein Windows-Programm die über die serielle Schnittstelle ankommenden Daten entgegen und zeigt diese in Echtzeit in einem Graphen an. Diese Daten können das „rohe“ Sensor-Signal oder schon die demodulierten Kanäle sein. Das Programm verfügt über eine Funktion zum Speichern der Daten als Datei mit dem Suffix TPD (Tricore Pulsoxymetry Data). Die Funktionalität der Software wird in Kapitel 6 beschrieben.

Dateien diesen Formats können in MATLAB importiert und weiter verarbeitet werden. Dies geschieht durch ein MATLAB-Skript, das letztendlich die SpO2-Werte und Pulsfrequenz über die Zeit liefert. Das Thema wird ausführlich in Kapitel 4 behandelt.

Auf der rechten Seite der Abb. 3.1 befinden sich die Werkzeuge, die zum Entwickeln und Debuggen der Programme für TriCore und PC verwendet wurden. Zu diesem Thema befindet sich eine Funktionsbeschreibung in Kapitel 6.

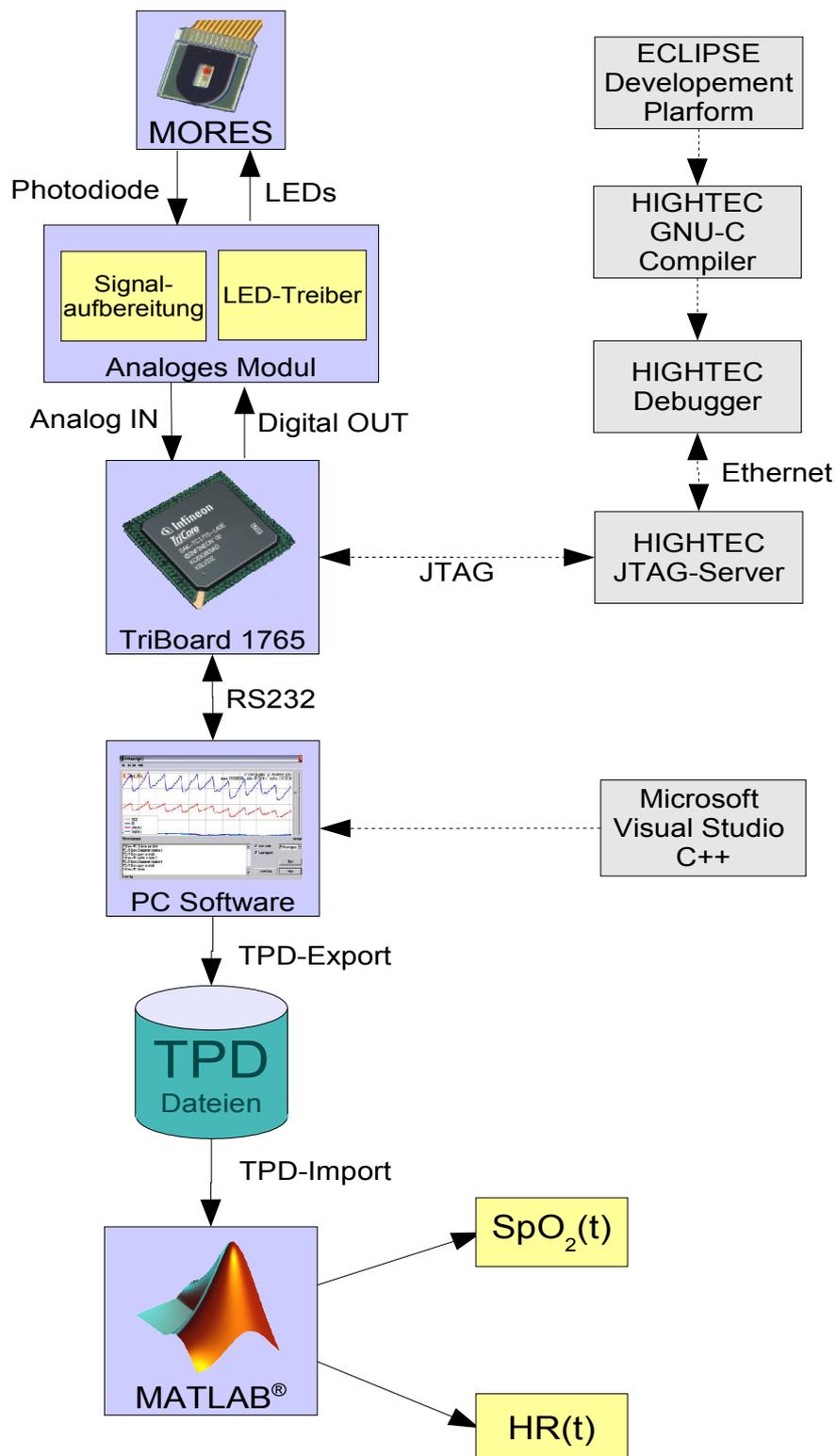


Abbildung 3.1: Übersicht aller Systemkomponenten vom Sensor bis zur Anzeige der Vitalparameter

4 Algorithmus

4.1 Signalverarbeitung mit MATLAB

Der größte Teil des Algorithmus ist unter MATLAB implementiert worden. Für die Verarbeitung unter MATLAB wird das Sensorsignal in diskreter Form benötigt. Dies geschieht durch Import des Sensorsignals als TPD-Datei. Als Beschreibung für den Weg vom Sensor bis zur TPD-Datei soll zunächst Abb. 3.1 ausreichen. Die dafür notwendige Hardware und Software wird in den Kapiteln 5 und 6 beschrieben.

Das diskrete Sensorsignal wird als TPD-Datei in MATLAB importiert und kann von hier aus weiterverarbeitet werden. Die Verarbeitung des Sensorsignals ist Teil dieses Kapitels und ist zur Übersicht als Blockschema 4.1 dargestellt. Dieses Blockschema kann als Referenz in diesem Kapitel genommen werden. Dort wird auf all die Blöcke der gesamten Verarbeitungskette unter MATLAB im Detail eingegangen. Die Verarbeitung des Sensorsignals erfolgt durch das M-Skript *spo2_test3.m*. An einigen Stellen erzeugt das M-Skript Graphen, die zu einer Figure zusammengefasst werden. Die Figures 1 bis 6 und 8 sind in der Abbildung 4.1 an den entsprechenden Stellen eingezeichnet. Figure 7 zeigt eine Kontrollfunktion, die nicht zur Verarbeitung gehört, und somit im Blockschema des Algorithmus keinen Platz findet.

Alle Filter und Chip-Folgen liegen als MAT-Datei vor und müssen daher nicht bei jedem Start des Hauptprogrammes *spo2_test3.m* neu generiert werden. Diese Filter wurden mit dem M-Skript *gen_spo2.m* erzeugt und können jederzeit neu generiert werden.

<i>M-Skript</i>	<i>Funktion</i>
<i>spo2_test3.m</i>	Berechnen der Vitalparameter
<i>gen_spo2.m</i>	Chip-Folgen und Filter-Generierung

Tabelle 4.1: Hauptprogramme unter MATLAB

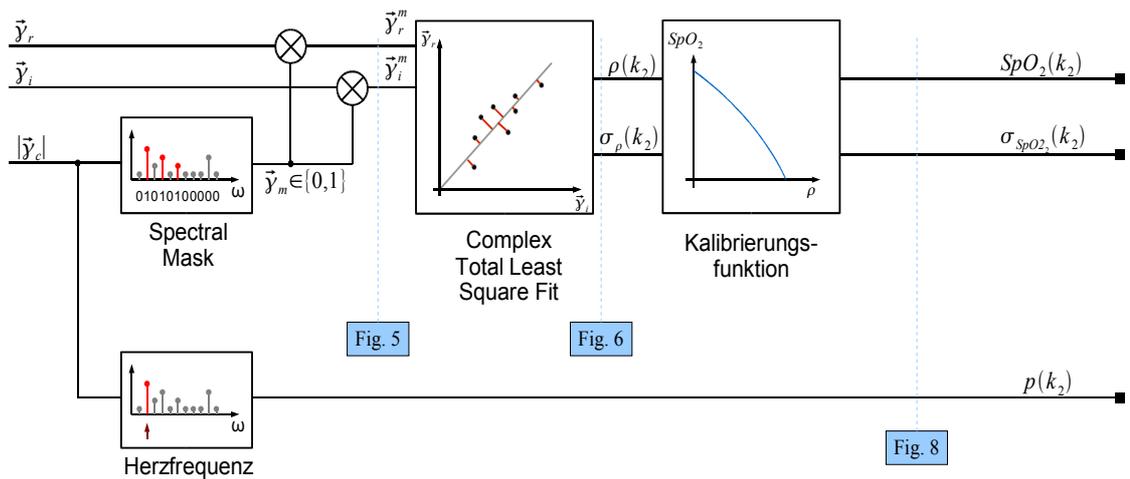
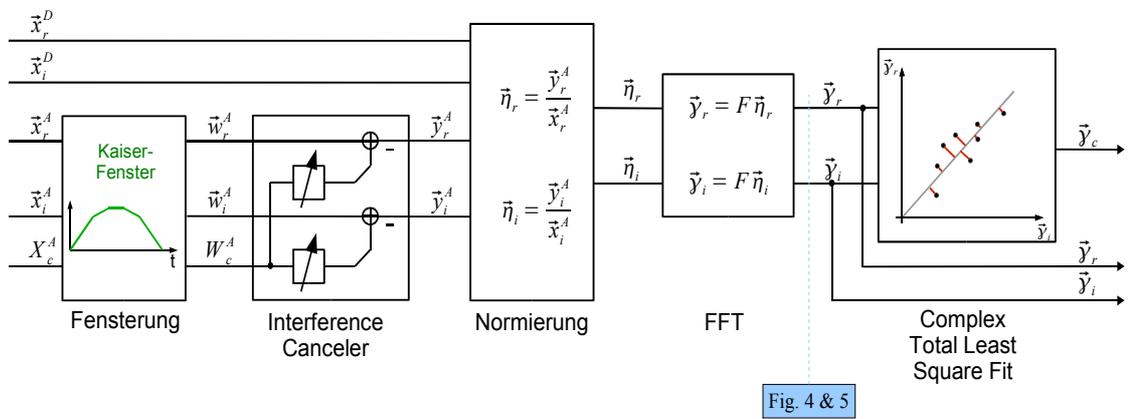
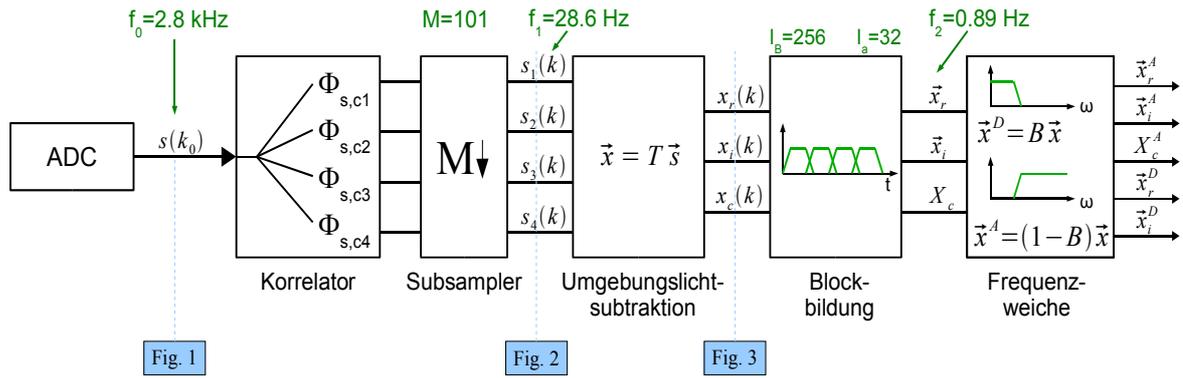


Abbildung 4.1: Blockschema des Algorithmus unter MATLAB

4.2 Chip-Folgen und Filter-Generierung

4.2.1 Chip-Folgen

4.2.1.1 Forderungen für die Chip-Folgen

Vor der Suche nach geeigneten Chip-Folgen müssen zunächst Überlegungen angestellt werden, welche Eigenschaften die gesuchte Folgen-Familie erfüllen muss, um die Effizienz des Modulationsverfahrens speziell in unserer Anwendung zu maximieren. Aus dem Angebot in der Literatur soll eine Familie, bestehend aus mindesten drei Chip-Folgen c_i , mit den folgenden Kriterien ausgesucht werden.

Grundsätzlich müssen die Folgen orthogonal sein, damit nach der Modulation eine Kanaltrennung und damit eine Demodulation ohne Übersprechen möglich wird. Es muss also für eine Familie mit N Folgen gelten:

$$\langle c_i, c_k \rangle = 0, \quad \text{mit } i \neq k; \quad \forall i, k \quad (4.1)$$

Es werden nur zwei Zustände der LEDs (EIN und AUS) zugelassen. Ein Vorteil ist, dass sich der LED-Treiber wegen der zwei Zustände sehr einfach gestalten lässt, weil die Nichtlinearität der Leuchtdioden nicht mehr beachtet werden muss. Ein weiterer Grund ist, dass zum Ansteuern nunmehr ein digitaler Ausgang pro LED genügt. Demnach sollte die gesuchte Familie aus binären Folgen bestehen und damit die Bedingung

$$c_i \in \{-1, 1\}, \quad \forall i \quad (4.2)$$

erfüllen. Da eine negative Intensität physikalisch nicht möglich ist, können wir die Folgen anpassen, so dass gilt:

$$c_i \in \{0, 1\}, \quad \forall i \quad (4.3)$$

Weiterhin sollen alle Folgen c_i den gleichen Energiegehalt besitzen und mit der periodischen Autokorrelationsfunktion Φ_{ii} die folgende Bedingung erfüllen [KS]:

$$\Phi_{ii}(0) = E_c, \quad \forall i \quad (4.4)$$

Damit schmalbandige Störer unabhängig von ihrer Frequenz ins Gewicht fallen, fordern wir für jede Chipfolge eine gleichverteilte Fouriertransformierte. Aus 4.3 wird klar, dass die Folge nicht DC-frei ist. Dem DC-Anteil soll zunächst keine Beachtung geschenkt werden. Für die DFT von c_i können wir schreiben:

$$C_i(k) = \sum_{n=0}^{N-1} c_i(n) e^{-2\pi \frac{j}{N} n k} \quad (4.5)$$

Mit 4.5 lautet die Bedingung für eine Gleichverteilung im Spektrum:

$$|C_i(k)| = K, \quad \forall k; \quad k \neq 0; \quad \forall i \quad (4.6)$$

4.2.1.2 Legendre-Folgen

M. R. Schröder beschreibt in [Sch] das Bildungsgesetz zum Generieren der sog. Legendre – Folgen, die die oben aufgestellten Eigenschaften erfüllen und gute Auto- und Kreuzkorrelationseigenschaften besitzen. Schröder [Sch] beschreibt das Bildungsgesetz sowohl für ternäre Folgen mit den Elementen $\{0, 1, -1\}$, als auch für binäre Folgen mit den Elementen $\{-1, 1\}$. Das MATLAB-Skript *leg_series.m* erstellt eine Legendre - Familie bestehend aus vier binären Folgen. Da Legendre – Folgen für alle Längen l_c , mit l_c als Primzahl existieren, muss für die Länge der Folgen eine Primzahl als Parameter an das Skript übergeben werden.

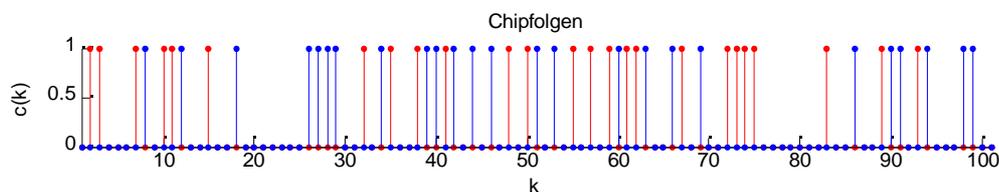


Abbildung 4.2: Zwei der insgesamt vier binären Legendre – Folgen der Länge $l_c = 101$, wie sie für die Kanäle Infrarot (blau) und Rot (rot) in dem hier entwickelten Pulsoximeter eingesetzt werden.

Mit den vier Folgen können insgesamt vier Signale unterschieden werden. Zwei dieser Folgen sind sogleich den beiden LEDs zugeordnet worden, so dass zwei Kanäle für andere Zwecke verbleiben. In Abb. 4.2 können die Schaltzustände (ein=1, aus=0) der LEDs direkt abgelesen werden. Dabei fällt auf, dass jeweils nur eine der LEDs zum gleichen Zeitpunkt k eingeschaltet ist, was ein eindeutiger Beweis für die Orthogonalität der Folgen ist. Diese Eigenschaft gilt für alle vier Folgen.

Für die spektralen Eigenschaften aller Folgen forderten wir eine Gleichverteilung im Spektrum, um die Störerunterdrückung nach dem Prinzip aus Kapitel 2.4 zu gewährleisten. Den Beweis liefert die FFT, deren Beträge für alle vier Folgen identisch sind. Damit ist das Spektrum gleich verteilt und die Gleichung 4.6 erfüllt. Zugleich lässt sich aus dem für alle Folgen c_i identischen Spektrum auf einen identischen Energiegehalt jeder Folge schließen und damit die Gleichung 4.4 erfüllen.

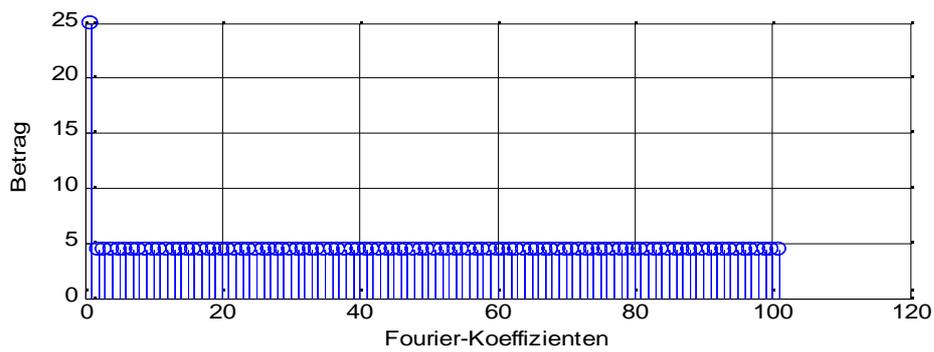


Abbildung 4.3: Betrag der FFT einer beliebigen Folge c_i

Da jede Folge c_i periodisch fortgesetzt wird, erhält man äquidistante, gleichhohe Peaks im Spektrum (Abb 4.4). Jeder dieser Peaks kann als Träger gesehen werden, um den das Nutzsignal im Gewebe amplitudenmoduliert wird. Der hohe DC-Anteil entsteht durch Verwendung binärer Folgen, deren Wahl wiederum auf die Technik zurückzuführen ist. Negative Intensitäten sind mit LEDs grundsätzlich nicht realisierbar.

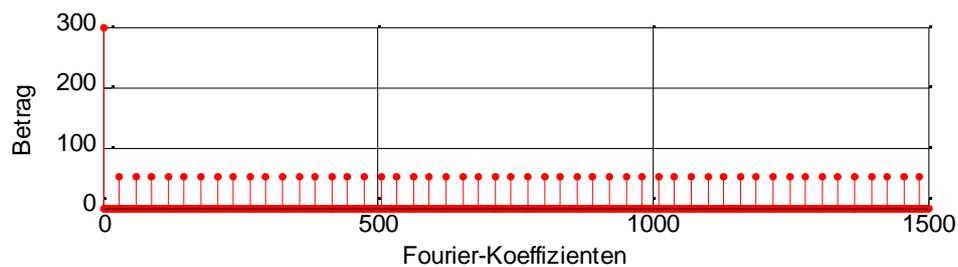


Abbildung 4.4: Betrag der FFT einer periodisch fortgesetzten Folge c_i

Den verbleibenden zwei Folgen werden Dunkelphasen zugeordnet. In diesen Dunkelphasen wird lediglich die Umgebungslichtintensität gemessen, die als Störsignal interpretiert wird, und von dem Nutzsignal subtrahiert werden muss. Daraus folgt, dass zwei Signale mit Umgebungslichtintensitäten hinzukommen, die sich durchaus voneinander unter bestimmten Bedingungen unterscheiden, wie wir sehen und davon Gebrauch machen werden. Die folgende Tabelle zeigt die Zuordnung aller verfügbaren Chip-Folgen zu den Signalen.

Chipfolge	Signalzuordnung
c_1	$s_1(k)$ Dunkelphase 1
c_2	$s_2(k)$ LED 660nm
c_3	$s_3(k)$ LED 880nm
c_4	$s_4(k)$ Dunkelphase 2

Tabelle 4.2: Kanalzuordnung

4.2.2 Extraktionsfilter

Für die weiteren Betrachtungen nehmen wir an, dass ein Sensorsignal $s(k_0)$ vorliegt, das bereits alle vier Signale aus Tabelle 4.2 gemischt enthält. Dieses Signal ist mit der Frequenz f_0 getaktet, wodurch wir jede $1/f_0$ Sekunden einen Wert erhalten. Wir benötigen für jeden der vier Kanäle ein entsprechendes Matched-Filter zur Extraktion der Signale aus dem Sensorsignal.

Ein solches Matched-Filter Φ_i lässt sich zunächst näherungsweise als eine Chip-Folge c_i schreiben, mit der das Sensorsignal zyklisch jede 101 Samples (für Chip-Folgen der Länge $l_c = 101$) multipliziert wird. Die so entstandenen Werte werden aufaddiert, also es entsteht eine Skalarmultiplikation der Form:

$$\tilde{s}_i(k) = \sum_{n=0}^{l_c-1} \tilde{\Phi}_i(n) \cdot s(k \cdot l_c + n), \quad \tilde{\Phi}_i = c_i; \quad \forall i \quad (4.7)$$

- c_i : Chip-Folge
- $\tilde{\Phi}_i$: Matched-Filter
- i : Kanal i
- s : Sensorsignal
- l_c : Länge einer Chip-Folge
- \tilde{s}_i : entspreiztes Nutzsinal
- k : Diskrete Zeitvariable

Dabei ist darauf zu achten, dass das Sensorsignal zum Zeitpunkt l_c mit dem ersten Chip beginnt. Sender und Empfänger müssen also perfekt synchronisiert sein. Diese Skalarmultiplikation führt blockweise zu einer Entspreizung des Nutzsignals ins Basisband. Zugleich erhält man eine Unterabtastung mit dem Faktor $l_c = 101$ für das Nutzsinal. Um Aliasing zu vermeiden, muss die Bandbreite eines Signals vor jeder Unterabtastung reduziert werden. Demnach benötigen wir noch ein Anti-Aliasing-Filter, und können das Matched-Filter nach der Gleichung 4.7 ohne dieses TP-Filter nicht in dieser Form verwenden. Bei einer Unterabtastung mit Faktor l_c kann man für die Abtastfrequenz des Nutzsignals f_1 schreiben:

$$f_1 := \frac{f_0}{l_c} \quad (4.8)$$

- f_1 : Abtastfrequenz des Sensorsignals
- f_0 : Abtastfrequenz des Nutzsignals
- l_c : Länge einer Chip-Folge

Bei einem perfekten TP-Filter würde man dessen Grenzfrequenz der halben Abtastfrequenz f_1 gleichsetzen. Da so ein Filter unendlicher Ordnung wäre, und damit nicht realisierbar ist, muss die Grenzfrequenz geschickt gewählt werden, so dass keine unerwünschten Aliasing-

Effekte beim Unterabtasten entstehen. In der Regel geht man einen Kompromiss zwischen maximalem Durchlassbereich und kleinstem Aliasing-Bereich nach dem folgenden Schema ein:

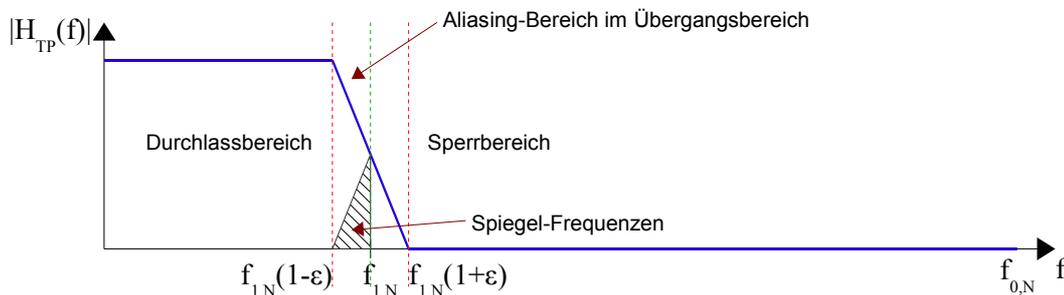


Abbildung 4.5: Design eines TP-Filters für Unterabtastung mit Aliasing-Bereich. Nach der ersten Nyquist-Bedingung gilt für die Grenzfrequenzen: $f_{1,N} := f_1 / 2$ und $f_{0,N} := f_0 / 2$

Der Übergangsbereich hat die Breite ϵ und hängt von der Filterordnung ab. Die Ordnung eines solchen Filters sollte möglichst groß sein, um den Übergangsbereich und damit den Aliasing-Bereich, möglichst klein zu bekommen. Ein solches TP-Filter mit Filterordnung 4039 könnte so aussehen:

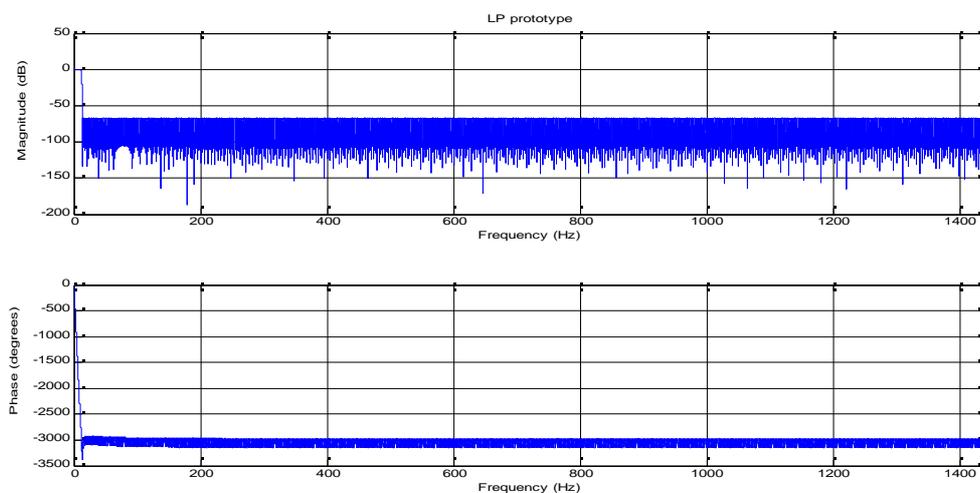


Abbildung 4.6: TP-Filter-Prototyp für die Unterabtastung

Wir können das Matched-Filter und das TP-Filter für Unterabtastung durch Faltung zu einem Extraktionfilter zusammenfassen. Da das Filter aus 4040 Koeffizienten besteht, muss die Chipfolge durch periodisches Fortsetzen auf dieselbe Länge gebracht werden. Anschließend lassen sich TP-Filter-Koeffizienten mit den Chip-Folgen im Zeitbereich multiplizieren. Betrachtet man die Übertragungsfunktion eines solchen Extraktionfilters, so fällt auf, dass durch die Faltung nunmehr 15dB Dämpfung im Sperrbereich verbleiben. Das bedeutet, dass alle Signale im Frequenzbereich zwischen $f_{1,N}$ und $f_{0,N}$, die nicht mit den Chip-Folgen korreliert sind, mit

15dB gedämpft werden und als Aliasing ins Nutzband, also zwischen 0 Hz und $f_{1,N}$ fallen. Eine höhere Dämpfung kann durch längere Chip-Folgen erreicht werden.

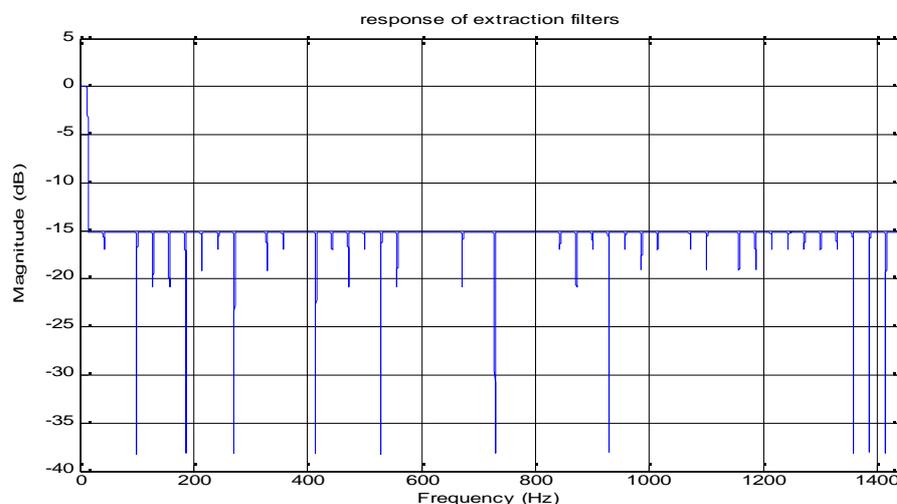


Abbildung 4.7: Extraktion – Filter mit 15dB Dämpfung

Untersuchungen haben gezeigt, dass Störungen starker Amplitude überwiegend auf künstliche Beleuchtung zurückzuführen sind. In Europa beträgt die Netzfrequenz 50 Hz, demnach liegt die Grundwelle der Leistung (bzw. Intensität) dieser Lampen bei 100 Hz, und deren Oberwellen liegen bei dem Vielfachen von 100 Hz. Es besteht die Gefahr, dass die Dämpfung des Extraktionfilters im Sperrbereich für starke Störer nicht ausreichen könnte. Aufgrund dieser Erkenntnisse ist es ratsam, Frequenzen beim Vielfachen von 100 Hz durch Anpassung der Eigenschaften des Extraktionfilters zu unterdrücken. Eine Möglichkeit wären Notches an den besagten Frequenzen durch erneute Faltung mit einem Notch-Filter.

Viel eleganter ist, dies durch geschickte Wahl der Abtastfrequenz f_0 und durch Verschieben der Sperrzone des TP-Filter-Prototypen zu erreichen. Aus dem Sperrbereich fallen alle Frequenzen größer als $f_{1,N}$ mit 15dB Dämpfung als Aliasing ins Nutzband unterhalb $f_{1,N}$. Wie schon angekündigt, kann man die Abtastfrequenz f_0 so wählen, dass die 100-Hz-Vielfachen mit 15dB Dämpfung an die Ränder des Nutzbandes fallen, also bei 0 bzw. $f_{1,N}$. Wählt man $f_0 = 2.886$ kHz, so fallen die Frequenzen $(2n-1)*100$ Hz als Aliasing bei $f_{1,N}$ und alle Frequenzen $2n*100$ Hz als Aliasing bei 0 Hz (also als DC-Offset) in das Nutzband, für alle

$$n \in \mathbb{N} / \{0\} \quad (4.9)$$

In der Regel ist gegen die Wahl der Grenzfrequenz nach dem Schema in Abb. 4.5 nichts auszusetzen. Da sich die 100-Hz-Vielfachen bei $f_{1,N}$ aufaddieren, ist es sinnvoll hier eine Sperrzone zu setzen. Dazu kann die Sperrzone nach links verschoben werden, wie in Abb. 4.8 gezeigt. Die Grundwelle bei 100 Hz und die ungeraden Vielfachen fallen in diesen Bereich. Die geraden Vielfachen von 100 Hz addieren sich zu jedem Kanal als DC-Offset. Speziell in dieser

Anwendung fällt der damit entstandene Fehler im ohnehin relativ hohen DC-Anteil des Nutzsignals kaum ins Gewicht.

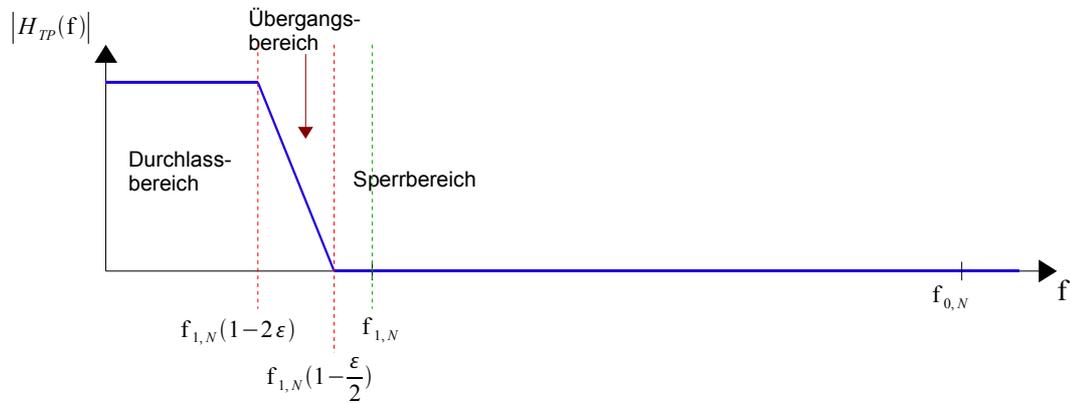


Abbildung 4.8: Design eines TP – Filters für Unterabtastung mit extra Sperrbereich

Die Faltung eines jeden Matched-Filters mit dem TP-Filter für Unterabtastung ergibt vier neue Extraktionsfilter, die neben der Extraktion des Nutzsignals aus dem Sensorsignal zusätzlich Störer beim Vielfachen von 100 Hz eliminieren. Die Dämpfung sonstiger, mit den Chip-Folgen unkorrelierter Signale, beträgt weiterhin 15dB. Abb. 4.9 zeigt, dass neben 100 Hz auch weitere Frequenzen unterdrückt werden. Dies muss aufgrund des Filterdesigns und der Wahl der Abtastfrequenz f_0 in Kauf genommen werden, und hat zur Folge, dass die gespreizte Information an diesen Stellen verloren geht. Dies wiederum bedeutet eine identische Dämpfung im Nutzsingnal für alle vier Kanäle.

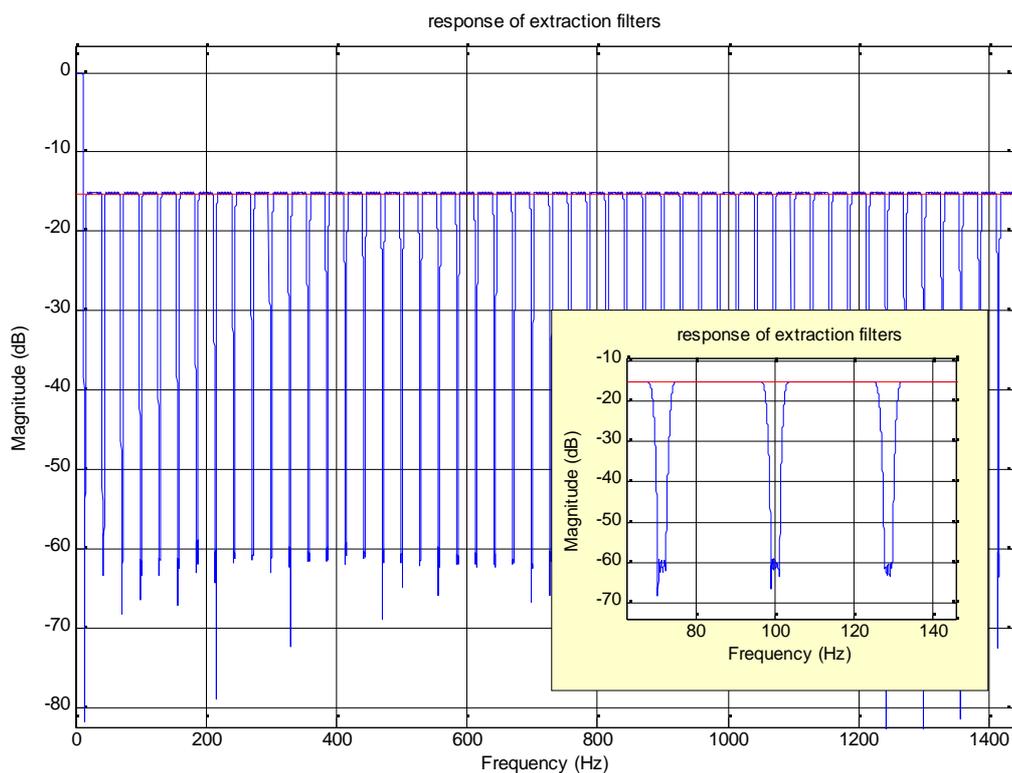


Abbildung 4.9: Übertragungsfunktion der Extraktion – Filter mit 15dB Dämpfung und extra 100-Hz-Unterdrückung; Vergrößerung im Bereich von 100 Hz.

Die so generierten Extraktionsfilter Φ_i mit $i = \{1,2,3,4\}$ können im Korrelator (siehe Blockschema Abb. 4.1) zum Demodulieren der vier Nutzsingale eingesetzt werden.

4.2.3 Filter der Frequenzweiche

Zur Trennung des pulsartigen Anteils von der DC-Komponente, entsprechend für die Normierung nach Gleichung 2.36, wird eine Frequenzweiche (vgl. Abb. 4.1) benötigt. Die Frequenzweiche besteht aus einem TP-Filter B und einem dazu inversen HP-Filter 1-B. Die minimale Pulsfrequenz ist nach den Untersuchungen in Kapitel 2.1.1 bei 0.5 Hz festgelegt worden. Die Trennfrequenz dieser Weiche liegt folglich unterhalb 0.5 Hz, so dass kein pulsartiger Anteil mehr in der DC-Komponente erkennbar ist. Zugleich darf die Trennfrequenz nicht zu niedrig sein, um noch rechtzeitig den Trend des Signals verfolgen zu können.

4.2.3.1 Interpolations-TP-Filter

Das TP-Filter B arbeitet blockweise, wobei ein Block ein Signalausschnitt ist. Es handelt sich demnach um kein gewöhnliches FIR-Filter, das für jedes Sample am Eingang ein Sample am Ausgang liefert, sondern ein Filter, das pro Block am Eingang der Frequenzweiche einen TP-

gefilterten Block errechnet. Es wird nun ein Block der Länge $l_b=256$ betrachtet. So ein Block kann als Vektor x geschrieben werden. Dieser Vektor ist ein Element des Raumes V :

$$\vec{x} \in V = \mathbb{R}^{256} \quad (4.10)$$

So gesehen repräsentiert V den Raum, in dem sich alle Signale x darstellen lassen. Zu diesem Vektorraum V existiert ein Untervektorraum W , zu dem eine Basis L bekannt ist. Durch die Basis L lässt sich der Vektorraum W vollständig beschreiben, da jede Linearkombination der Basis L ein Element des Raumes W ist.

$$\vec{a}L \in W \quad (4.11)$$

Die Basis L besteht aus einem n -Tupel linear unabhängiger Vektoren, also spannt die Basis L einen n -dimensionalen Vektorraum W auf.

$$L \in W(l_b \times n, \mathbb{R}) \quad (4.12)$$

n ist die Dimension des Raumes, und mit $n < l_b$ ist die Dimension des Raumes W kleiner als die des Raumes V .

$$W \subset V \quad (4.13)$$

Der Vektor x sei nun ein Ausschnitt aus einem Eingangssignal. Welche Signalform einen Vektor x repräsentiert, ist durch die Menge der möglichen Vektoren des Raumes V gegeben. Die durch den Vektor x repräsentierten Signale werden so klassifiziert, dass man drei Aussagen über ein Signal x treffen kann: Zum Einen kann ein Signal x hohe Frequenzen enthalten, zum Zweiten tiefe Frequenzen und zum Dritten beide Frequenzkomponenten. Wenn ein Signal, das durch einen Vektor x repräsentiert wird, Element des Raumes V ist, dann können alle drei Aussagen zutreffend sein. Die Basis L sei nun so gewählt, dass sich mit dieser Basis nur tiefe Frequenzen darstellen lassen. Schlussfolgernd lässt sich sagen, dass ein Vektor, der nur tiefe Frequenzen enthält, ein Element des Raumes W ist.

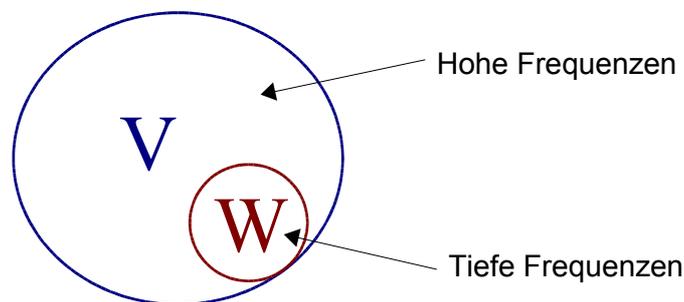


Abbildung 4.10: Schema des Raumes V und seines Unterraumes W

Gesucht ist eine Abbildung des Raumes V auf W , so dass sich alle Vektoren des Raumes V auf den Raum W projizieren lassen:

$$f_B : V \rightarrow W \quad (4.14)$$

Mit dieser Abbildung ließe sich ein Filter realisieren, das mögliche Frequenzen des Raumes V auf den Raum W projiziert, also ein TP-Filter.

Als Ansatz wird eine Linearkombination der Basis L gesucht, die einen Vektor x am besten beschreibt, während der Vektor x weiterhin Element des Raumes V sei. Da diese Linearkombination nicht existiert, ist die Gleichung nicht lösbar; Es handelt sich hierbei um ein überbestimmtes Gleichungssystem der Form:

$$\vec{x} \stackrel{!}{=} L\vec{a} \quad (4.15)$$

Es existieren mehr Gleichungen als Unbekannte, deshalb gibt es in diesem Fall keine Lösung für a . Also sucht man nach einem a , das am besten in das überbestimmte Gleichungssystem passt, also ein lineares least-square Problem, bei dem eine Linearkombination a zu L gesucht wird, für die der minimale quadratische Fehler (least square fit) [NUM] gilt:

$$\|L\vec{a}_m - \vec{x}\| \rightarrow \textit{Minimum} \quad (4.16)$$

Die Gleichung lässt sich erweitern mit L^T zu

$$L^T \vec{x} = L^T L \vec{a}_m \quad (4.17)$$

und umformen zu

$$(L^T L)^{-1} L^T \vec{x} = \vec{a}_m \quad (4.18)$$

Der Term zur linken Seite entspricht der Pseudo-Inversen von L . Pseudo-Inverse, weil L in einem überbestimmten Gleichungssystem nicht invertierbar ist, jedoch lässt sich der Term $(L^T L)^{-1} L^T$ als eine Art Inverse von L sehen, worauf man den Term $(L^T L)^{-1} L^T$ als die Pseudo-Inverse $L^\#$ bezeichnet [For]:

$$L^\# \vec{x} = \vec{a}_m \quad (4.19)$$

Damit erhält man für den Vektor a_m eine Linearkombination von L , die die Bedingung 4.16 am besten erfüllt. Somit wurde x aus V auf W projiziert. Da sich mit L nur tiefe Frequenzkomponenten darstellen lassen, ist das Resultat

$$\vec{y} := L\vec{a}_m \quad (4.20)$$

ein Signal y , das nurmehr tiefe Frequenzen enthält. Fassen wir nun die Gleichungen 4.19 und 4.20 zusammen, dann lässt sich eine Projektion des Raumes V auf W als Abbildung schreiben, wobei x die Eingangsgröße und y die Ausgangsgröße ist.

$$\vec{y} = LL^\# \vec{x} \quad (4.21)$$

L und L# lassen sich noch zu B zusammenfassen, was in der MATLAB-Simulation auch gemacht worden ist.

$$\vec{y} = B \vec{x} \quad (4.22)$$

Bei der Implementierung auf ein Echtzeitsystem ist von Gl. 4.22 abzuraten, da diese im Verhältnis zu Gl. 4.21 mehr Operationen liefert. Die Ersparnis an Operationen lässt sich als Funktion von Blocklänge und Anzahl der Basisvektoren von L schreiben.

$$n_o = \frac{l_B}{2} \cdot n \quad (4.23)$$

- n: Anzahl linear unabhängiger Vektoren der Basis L (Dimension von W)
- n_o: Ersparnis an Operationen unter Implementierung von Gl. 4.21 anstelle von Gl. 4.22
- l_B: Länge eines Blocks x (Dimension von V)

4.2.3.2 Basis des Interpolations-TP-Filters

Bisher wurden die Basisvektoren der Basis L so gewählt, dass sich mit allen Linearkombinationen von L nur Signale tiefer Frequenzen beschreiben lassen. Wie sieht eine Basis mit dieser Eigenschaft aus? Zunächst wird ein entsprechendes FIR-Filter mit der gewünschten Grenzfrequenz entworfen, dessen Koeffizienten typischerweise einer sinc-Funktion entsprechen. Jeder der Basisvektoren von L repräsentiert eine solche sinc-Funktion, die mit jedem weiteren Basisvektor um 180 Grad versetzt wird. Der Rang einer solchen Basis bestimmt die Filterordnung dieses Filters. Zusätzlich wird noch ein Vektor mit identischen Elementen > 0 hinzugefügt, mit dem sich ein DC-Offset beschreiben lässt.

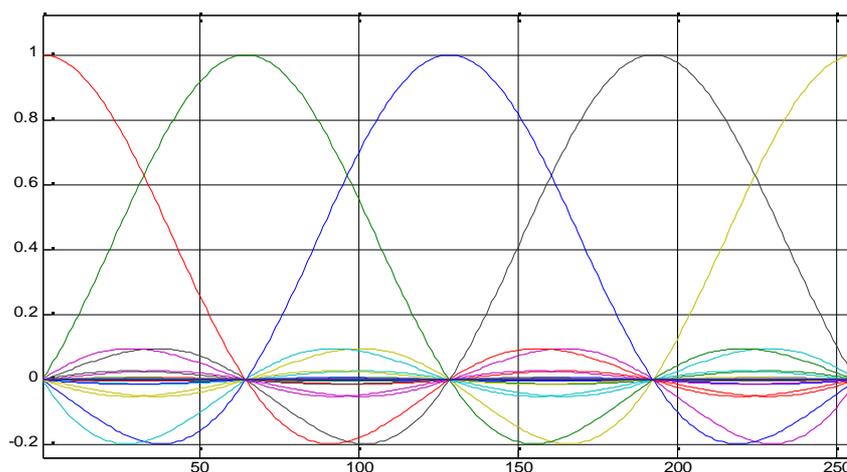


Abbildung 4.11: Basisvektoren der Basis L

4.2.3.3 Interpolations-HP-Filter

Ein zu B inverses HP-Filter wird ebenfalls benötigt, um den pulsatilen Anteil aus den an Eingängen der Frequenzweiche anliegenden Signalen zu extrahieren. Ein solches Filter muss nicht erneut hergeleitet und berechnet werden. Es genügt lediglich B von der Einheitsmatrix zu subtrahieren und man erhält das Inverse Filter zu B, ein HP-Filter.

$$A = E - B \quad (4.24)$$

Die Filterung erfolgt mit A äquivalent zur Gl. 4.22, wobei bei der Portierung auf einem Echtzeitsystem auch hier mit der Implementierung nach Gl. 4.21 Rechenzeit eingespart werden kann. Eine bessere Alternative, die lediglich 256 Operationen erfordert, ist die Subtraktion des Vektors y vom Vektor x aus der Gl. 4.22. Als Beweis kann Gl. 4.22 auf beiden Seiten mit dem Vektor x erweitert und umgeformt werden zu:

$$\vec{x} - \vec{y} = (E - B)\vec{x} \quad (4.25)$$

4.3 Verarbeitungsschritte unter MATLAB

Sind alle Filter generiert und liegen anschließend Test-Episoden mit Sensorsignalen im TPD-Format vor, so sind alle Voraussetzungen für die Berechnung der Vitalparameter unter MATLAB gegeben. Zum Algorithmus gehört die Demodulation der Signale, und zugleich die Unterdrückung von Störungen. Beides ist hier eng miteinander verknüpft worden, da der zentrale Gedanke dieser Arbeit die Unterdrückung von Störungen im voraus ist. Hierzu wird ein Teil der Störungen bereits durch das Modulationsverfahren aus dem Nutzsignal entfernt. Anschließend wird auch versucht, durch Einsatz zusätzlicher Algorithmen, den verbleibenden Rest der Störungen im Nutzsignal zu minimieren. Die Berechnung wird durch Start des M-Skriptes *spo2_test3.m* angestoßen. Der Funktionsaufruf unter MATLAB lautet:

```
spo2_test3(filename, lf_flag, ifc_flag, maskflag, ifc_order)
```

Einige Parameter wurden zur Übergabe an die Funktion herausgeführt. Dies ermöglicht das Ab- oder Zuschalten verschiedener Einheiten zur Störerunterdrückung und das Wechseln zwischen Test-Episoden, ohne den Code zu editieren.

Parameter	Datentyp	Funktion
filename	String	Absoluter oder relativer Pfad zu einer TPD-Datei
lf_flag	Integer	Baseline-Subtraktion; 1=ein, 0=aus
ifc_flag	Integer	Interference canceler; 1=ein, 0=aus
maskflag	Integer	Frequency mask; 1=ein, 0=aus
ifc_order	Integer	Interference canceler filter order

Tabelle 4.3: Parameter des Hauptprogrammes zur Berechnung der Sauerstoffsättigung

Vor der Verarbeitung lädt das M-Skript die zuvor generierten Filter aus einer MAT-Datei. Die Datei `syst101_100Hz.mat` enthält beispielsweise einen Datensatz an Extraction-Filtern mit zusätzlicher 100-Hz-Unterdrückung. Es ist möglich, zwischen Filter-Datensätzen zu wechseln, es können aber auch neue Filter generiert werden. Es ist darauf zu achten, dass die Test-Episoden bezüglich der Chip-Folgen mit den Extraction-Filtern übereinstimmen, so dass eine Rekonstruktion der Signale möglich wird. Weiterhin ist die Abtastfrequenz unter Einsatz eines Filters mit 100-Hz-Unterdrückung zu beachten, da ein 100-Hz-Filter im voraus für eine feste Abtastfrequenz generiert wird.

In den folgenden Kapiteln wird der gesamte Algorithmus blockweise anhand des Blockschemas in Abb. 4.1 erläutert. Die Symbole und Bezeichnungen werden strikt nach dem Blockschema eingehalten.

4.3.1 Signal-Import in MATLAB

Die Abtastpunkte des Senorsignals $s(k_0)$ liegen als Datei im TPD-Format vor. TPD-Dateien werden mittels des M-Skripts `read_tpd2.m` geladen. Da inzwischen zwei Versionen von TPD-Dateien existieren, ist das Skript abwärtskompatibel und erkennt automatisch die Version der TPD-Datei. Das TPD-Format beinhaltet neben dem Datensatz zusätzliche Informationen, wie Abtastfrequenz f_0 oder Anzahl der Abtastpunkte, die als Variablen für die weitere Verarbeitung benötigt werden.

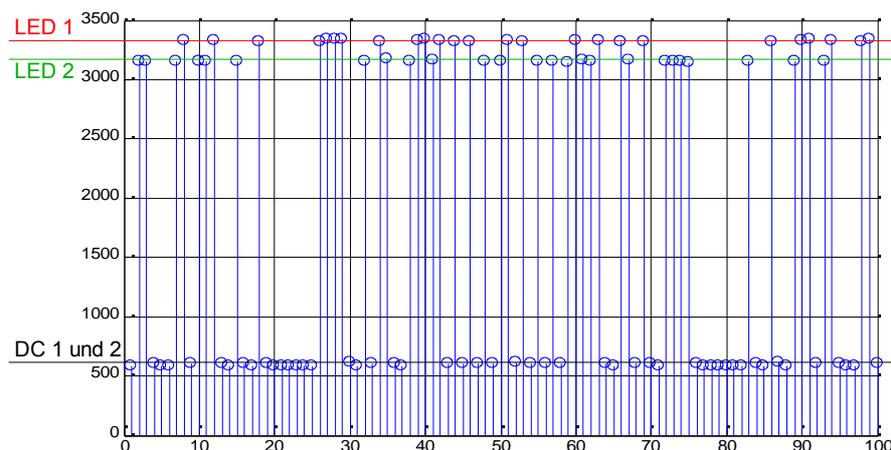


Abbildung 4.12: Figure 1 aus der Simulation unter MATLAB:
Ausschnitt über 100 Samples aus Figure 1, die das diskrete
Sensorsignal zeigt. Eingezeichnet sind die Signalpegel der beiden
LEDs und der beiden Dunkelphasen.

In der MATLAB-Simulation zeigt Figure 1 das Sensorsignal unverarbeitet, so wie es vom AD-Wandler abgetastet wurde. In der Abb. 4.12 wird deutlich, dass die Intensitäten der LEDs angepasst wurden, so dass die beiden Signale in etwa die selbe Amplitude im Photostrom verursachen. In der Regel würde hier das Signal der infraroten LED deutlich dominieren, was zum Übersprechen in den Rot-Kanal führen könnte.

4.3.2 Korrelator und Subsampler

Dem Korrelator liegt das Sensorsignal $s(k_0)$ vor, das nun zu vier Kanälen demoduliert und unterabgetastet werden soll. Dabei ist k_0 die diskrete Zeitkonstante des Sensorsignals. Korrelator und Subsampler können zu einer Einheit, bestehend aus vier Subsampling-Filtern, zusammengefasst werden. In Kapitel 4.2.2 ist die Generierung der Koeffizienten der vier Extraktionfilter ausführlich behandelt worden. Ein Extraktionfilter Φ_i beinhaltet bereits ein für die Unterabtastung notwendiges TP-Filter, und zugleich ein Matched-Filter zum Entspreizen eines Spreizsignals i ins Basisband. Ein solches Subsampling-Filter hat die Form:

$$s_i(k) = \sum_{n=0}^{l_\phi-1} \Phi_i(n) \cdot s(k \cdot l_c + n), \quad \forall i \quad (4.26)$$

- Φ_i Koeffizienten eines Extraktionsfilter
- l_ϕ : Anzahl der Filterkoeffizienten eines Filters Φ_i
- l_c : Länge einer Chip-Folge
- i : Kanal i
- s : Sensorsignal
- s_i : entspreiztes Nutzsignal
- k : Diskrete Zeitvariable

Das Filter setzt mindestens $l_\phi=4040$ Samples im Speicher voraus, erst dann kann mit der Demodulation begonnen werden. Nach Eintreffen weiterer $l_c=101$ Samples kann das nächste Sample-Quadrupel $s_i(k)$ für alle i dekodiert werden. Unter MATLAB stehen alle Samples von Anfang an zur Verfügung. Für die spätere Implementierung auf einen Mikrocontroller bedeutet das eine einmalige Wartezeit bis zur Funktionsbereitschaft von max. 2 Sekunden, abhängig von der Abtastrate. Die 2 Sekunden sind absolut vertretbar.

Nach der Extraktion aus dem Sensorsignal liegen vier Signale $s_i(k)$ mit einer Abtastfrequenz f_1 vor. Mit der Chip-Folgen-Länge l_c folgt für die Abtastfrequenz f_1 nach der Unterabtastung:

$$f_1 = \frac{f_0}{l_c} \quad (4.27)$$

Als erstes Resultat sind die Störungen oberhalb des Nutzbandes, also Störungen oberhalb der halben Abtastfrequenz f_1 , mit 15dB gedämpft ins Nutzband gespiegelt worden. Gleichzeitig sollten alle Vielfachen von 100 Hz eliminiert bzw. als DC-Offset aufaddiert worden sein.

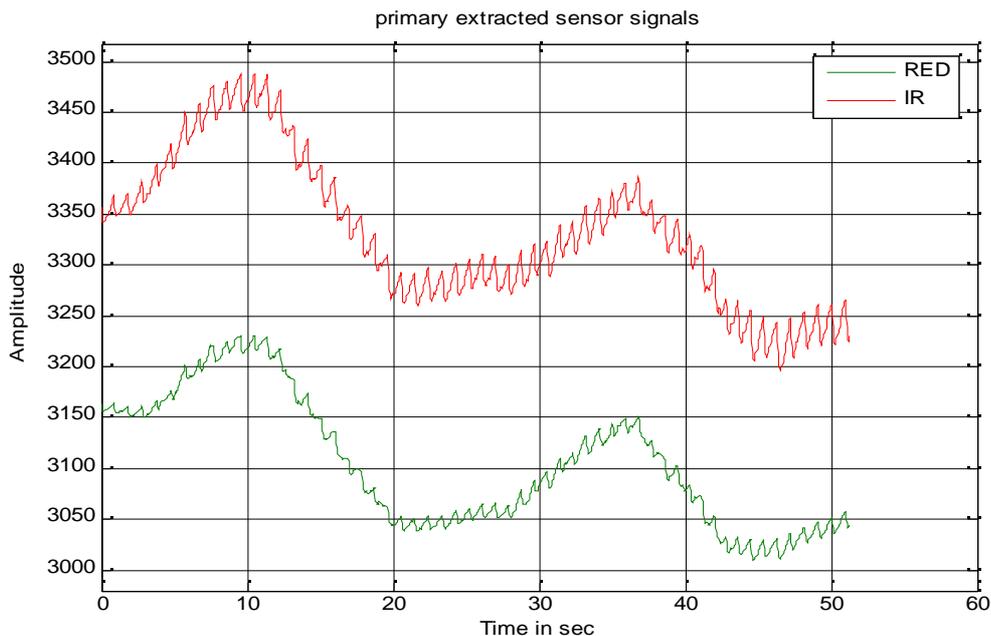


Abbildung 4.13: Figure 2 aus der Simulation unter MATLAB:
 Die Signale $s_2(k)$ und $s_3(k)$ direkt nach der Extraktion aus dem Sensorsignal. $s_2(k)$ entspricht dem roten Kanal und $s_3(k)$ entspricht dem infraroten Kanal. Den Signalen sind Störungen additiv überlagert. 100-Hz-Störer wurden bereits gefiltert bzw. als DC-Offset gespiegelt

Die Störungen unterhalb der halben Abtastfrequenz f_1 sind additiv allen vier Kanäle überlagert worden, genauso ist in der Baseline (Dunkelphase) auch schon der reine Störungsanteil sichtbar.

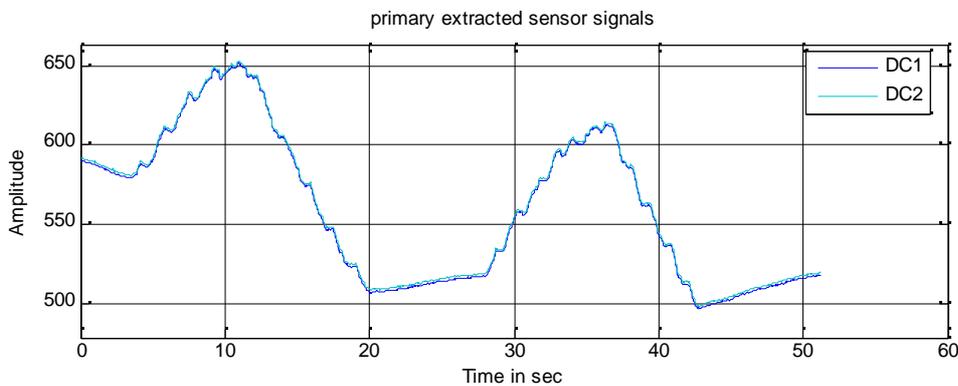


Abbildung 4.14: Figure 2 aus der Simulation unter MATLAB:
 Dunkelphasen $s_1(k)$ und $s_4(k)$ nach der Extraktion aus dem Sensorsignal.

4.3.3 Umgebungslichtsubtraktion

Nach dem Subsampler liegen die Signale $s_i(k)$ vor, die zur weiteren Betrachtung zu einem Spaltenvektor \vec{s} zusammengefasst werden.

$$\vec{s}(k) = \begin{bmatrix} s_1(k) \\ s_2(k) \\ s_3(k) \\ s_4(k) \end{bmatrix} \quad (4.28)$$

Die Aufgabe des Moduls zur Umgebungslichtsubtraktion ist, wie der Name verrät, die Subtraktion des Umgebungslichtanteils vom Nutzsignal, aber zugleich die Generierung eines Referenzsignals für den adaptiven Interference-Canceler. Alle Operationen dieses Moduls lassen sich durch eine Matrix T beschreiben.

$$\tilde{T} = \begin{bmatrix} -0.5 & 1 & 0 & -0.5 \\ -0.5 & 0 & 1 & -0.5 \\ 0.5 & 0 & 0 & -0.5 \end{bmatrix} \quad (4.29)$$

Nach Multiplikation der Eingangsgrößen mit T

$$\vec{x}(k) = T \vec{s}(k) \quad (4.30)$$

entsteht ein Vektor x , bestehend aus den zwei Kanälen (Rot und Infrarot) und einem Referenzsignal, das als Referenz für den adaptiven Interference-Canceler dienen wird.

$$\vec{x}(k) = \begin{bmatrix} x_r(k) \\ x_i(k) \\ x_c(k) \end{bmatrix} \quad (4.31)$$

Den Signalen Rot x_r und Infrarot x_i ist bereits der Anteil an Umgebungslicht, als Mittelung der beiden Dunkelphasen, subtrahiert worden. Es ist keine exakte Mittelung, da sich unter Betrachtung des Spektrums herausstellt, dass die Frequenzen unterschiedlich gewichtet subtrahiert werden. Das bedeutet, dass der Umgebungslichtanteil von dem Nutzsignal mit einem Fehler subtrahiert wird, wie in Abb. 4.15 dargestellt. Die Abbildung zeigt die DFT der Chip-Folgen mit der Gleichung:

$$X_1 = DFT \left(\vec{c}_2 - \frac{\vec{c}_1 + \vec{c}_4}{2} \right) \quad (4.32)$$

Um dieses Problem zu beseitigen, wurden die Dunkelphasen linear gewichtet und anschließend vom Nutzsignal subtrahiert. Die Gewichtung wurde iterativ ermittelt.

$$x_1(k) = s_2(k) - \frac{s_1(k) \cdot 0.475 + s_4(k) \cdot 0.525}{2} \quad (4.33)$$

Für den Kanal INFRAROT müssen die Gewichtungsfaktoren vertauscht werden:

$$x_2(k) = s_3(k) - \frac{s_1(k) \cdot 0.525 + s_4(k) \cdot 0.475}{2}$$

Die Gewichtung sorgt für eine gleichverteilte FFT und damit eine korrekte Subtraktion der Umgebungslichtintensität.

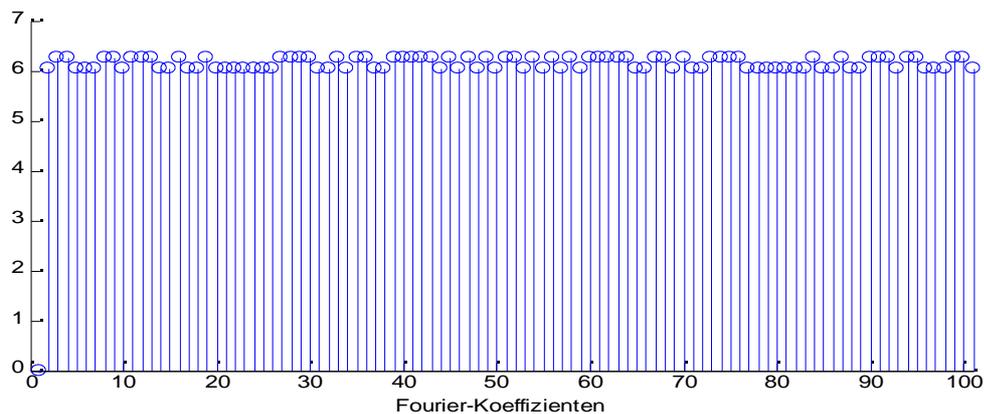


Abbildung 4.15: Problem bei der Subtraktion des Mittelwertes der Dunkelphasen am Beispiel für Kanal ROT Berechnet nach Gleichung 4.32.

Die Matrix T nach der Korrektur:

$$T = \begin{bmatrix} -0.475 & 1 & 0 & -0.525 \\ -0.525 & 0 & 1 & -0.475 \\ 0.5 & 0 & 0 & -0.5 \end{bmatrix} \quad (4.34)$$

Wir unterscheiden zwischen zwei Frequenzbändern in die ein Störer fallen kann. Zum Einen ist es das Band oberhalb der halben Abtastfrequenz f_1 , das Nutzband. Zum Anderen ist es das Band oberhalb dieser Frequenz, das Übertragungsband. Störungsbedingte Frequenzkomponenten, die in das Nutzband fallen, können mittels der Dunkelphasen-Subtraktion aus den zwei Nutzsignalen (Rot und Infrarot) entfernt werden. Diese Frequenzen sind sowohl in Phase als auch Amplitude gleich, und treten deshalb nicht im Referenzsignal auf. Das Referenzsignal ist lediglich die Differenz der beiden Dunkelphasen. Ein Störer im Nutzband ergibt demnach durchgehend Null für das Referenzsignal. Ein Störer im Nutzband könnte eine Lichtquelle sein, die durch das Gewebe vom Sensor erfasst, und deren Intensität mit den Volumenänderungen des arteriellen Blutes moduliert wird. Diese Anteile dürfen im Referenzsignal nicht auftauchen, sonst würde ein Interference-Canceler die gewünschte Information (den pulsatilen Anteil) herausfiltern.

Im Gegensatz dazu kann ein Störer in das Übertragungsband fallen. In diesem Fall setzt die Dämpfung des TP-Filters der Unterabtastung ein, was zunächst dazu führt, dass die Störung gedämpft in das Nutzband fällt. Zusätzlich ist dies mit einer Phasenverschiebung verbunden, die für jeden Kanal unterschiedlich ist. Die Differenz der beiden Dunkelphasen ergibt damit nicht mehr Null, sondern ein Signal, dessen Frequenzkomponenten die gespiegelten Frequenzen des

Störers aus dem Übertragungsband sind. Also ein perfektes Referenzsignal für einen adaptiven Interference-Canceler, um auch die verbleibenden Störungen aus dem Übertragungsband zu beseitigen.

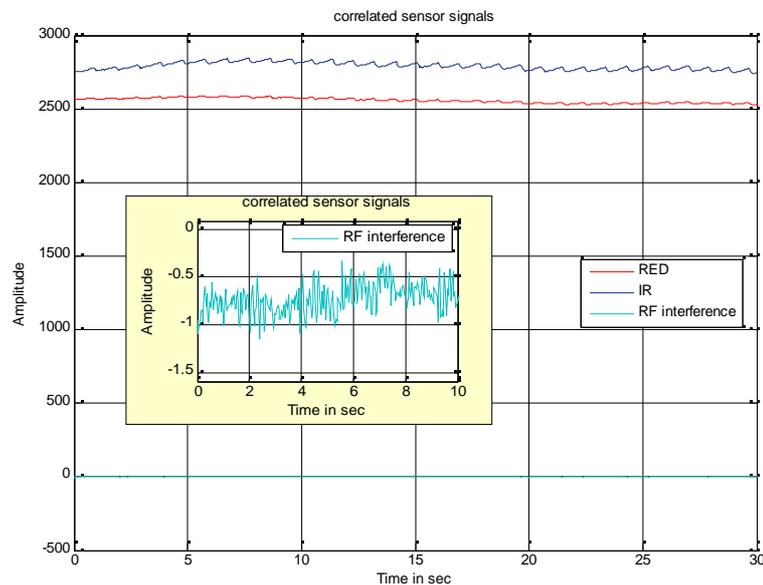


Abbildung 4.16: Figure 3 aus der Simulation unter MATLAB:
Die Signale x_r und x_i nach der Umgebungslichtsubtraktion
und das Referenzsignal x_c mit Vergrößerung.

Der Nachteil der Umgebungslichtsubtraktion sind 3dB Dämpfung, die durch die Subtraktion verloren gehen. Das bedeutet, dass Störer aus dem Übertragungsband nunmehr mit 12dB Dämpfung ins Nutzband gespiegelt werden und nicht, wie zuvor, mit 15dB Dämpfung.

Einer der Übergabeparameter an das M-Skript spo2_test3.m ist das lf_flag, es dient dem Ab- oder Zuschalten der Umgebungslichtsubtraktion.

4.3.4 Blockbildung zur Blockverarbeitung

Die weiteren Schritte erfolgen in Blockverarbeitung. Die Signale werden dazu in Blöcke (engl. auch Frames genannt) gleicher Länge l_b unterteilt, wobei nach jeden l_a Samples ein Block aus einem Eingangssignal herausgenommen wird, und damit für $l_a < l_b$ in jedem Block ein Teil der Information seines vorhergehenden Blocks steckt. Damit werden die Übergänge zwischen den Blöcken verwischt. Üblicherweise wird diese Blockbildung vor einer Fensterung mit anschließender FFT zur Verarbeitung im Frequenzbereich benötigt [MAT] [Grü]. Auch in diesem Algorithmus folgen Fensterung und FFT, allerdings nicht gleich unmittelbar nach der Blockbildung.

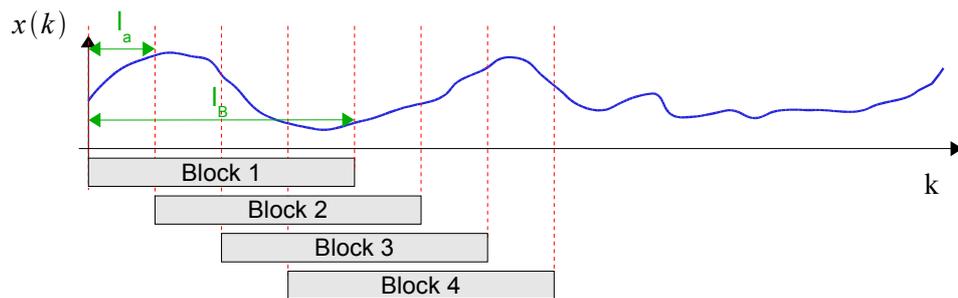


Abbildung 4.17: Blockbildung vor der Blockverarbeitung, vollständig beschrieben durch die Parameter l_B und l_a .

Die Abbildung zeigt, wie in Sprungweiten von l_a Samples, ein Block der Länge l_B aus x „herausgenommen“ wird. Die Blockbildung erfolgt für alle drei Eingangsgrößen $x_r(k)$, $x_i(k)$ und $x_c(k)$. Damit erhält man für die Signale $x_r(k)$ und $x_i(k)$ jeweils einen Block im Takt f_2 . Die Frequenz f_2 hängt unmittelbar von der Frequenz f_1 und dem Parameter l_a ab.

$$f_2 = \frac{f_1}{l_a} \quad (4.35)$$

Da es sich unter MATLAB um Offline-Verarbeitung handelt, liegen bereits alle Signale x für alle k vor. In einem Echtzeitsystem würde jede $1/f_2$ Sekunden ein Block erzeugt werden, was derzeit nach Gl. 4.35 in etwa einer Sekunde entspricht. In dem Blockschema des Offline-Systems wurde wegen einer besseren Übersicht die Abhängigkeit der Größen von der diskreten Zeitvariable k_2 weggelassen. Und deswegen soll für die Blockverarbeitung gelten: Falls keine Abhängigkeit von einer diskreten Zeitvariable angegeben ist, so ist ein Vektor oder eine Matrix abhängig von k_2 .

Im Blockschema aus Abb. 4.1 sind die Vektoren x_r und x_i für die Blöcke der beiden Kanäle Rot und Infrarot als Ausgangsgrößen abgebildet. Zudem wird eine Matrix X_c ausgegeben, die als Referenz für den adaptiven Interference-Canceler erzeugt wird. Wie zuvor festgelegt, sind all die Ausgangsgrößen von k_2 abhängig. Die Vektoren x_r und x_i wurden dabei exakt nach dem Prinzip aus Abb. 4.17 generiert.

Die Matrix X_c besteht aus mehreren solcher Blöcke, die aus $x_c(k)$ nach demselben Prinzip zur Blockbildung erzeugt wurden, allerdings sind diese Blöcke jeweils um ein Sample zueinander verschoben. Die Spaltenvektoren der Matrix X_c sind somit zeitversetzte Blöcke aus dem Referenzsignal $x_c(k)$. Die folgende Gleichung beschreibt die Matrix X_c vollständig:

$$\begin{aligned}
 [X_c]_{ij}(k_2) &= x_c(i+k_2 \cdot l_a + j) \\
 \forall j &\in \{0, 1, \dots, N_{\text{ifc}}\} \\
 \forall i &\in \{0, 1, \dots, l_B - 1\}
 \end{aligned}
 \tag{4.36}$$

- x_c : Interferenz-Referenzsignal
- k_2 : Diskrete Zeitvariable in der Blockverarbeitung
- l_B : Länge eines Blocks
- l_a : Sprungkonstante bei der Blockbildung
- N_{ifc} : Filterordnung des adaptiven Interference-Cancelers

Bereits hier wird die Filterordnung des adaptiven Interference-Cancelers bestimmt. Der Übergabeparameter `ifc_order` an das M-Skript `spo2_test3.m` ist die Filterordnung N_{ifc} aus der Gleichung 4.36.

4.3.5 Frequenzweiche

Die Aufgabe der Frequenzweiche ist die Filterung des DC-Anteils und des pulsatilen Anteils aus den anliegenden Eingangssignalen. Die Trennfrequenz dieser Weiche liegt unterhalb von 0.5 Hz. Hierzu wurde schon bereits in Kapitel 4.2.3 ein TP-Filter B zur Blockverarbeitung erstellt und dessen Arbeitsweise erläutert.

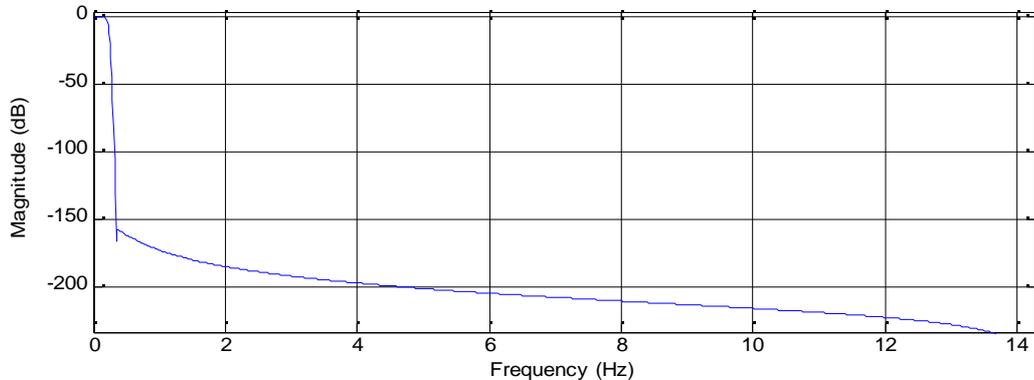


Abbildung 4.18: Frequenzgang des TP-Filters B mit -20dB bei 0.25 Hz und bereits -80dB bei 0.3 Hz

Für jeden Vektor \vec{x}_r und \vec{x}_i am Eingang der Frequenzweiche wird eine TP-Filterung und HP-Filterung durchgeführt:

$$\vec{x}_r^D = B \vec{x}_r \tag{4.37}$$

$$\vec{x}_i^D = B \vec{x}_i \tag{4.38}$$

$$\vec{x}_r^A = A \vec{x}_r \tag{4.39}$$

$$\vec{x}_i^A = A \vec{x}_i \tag{4.40}$$

Auch für alle Spaltenvektoren aus X_c erfolgt eine HP-Filterung.

$$X_c^A = A \cdot X_c \quad (4.41)$$

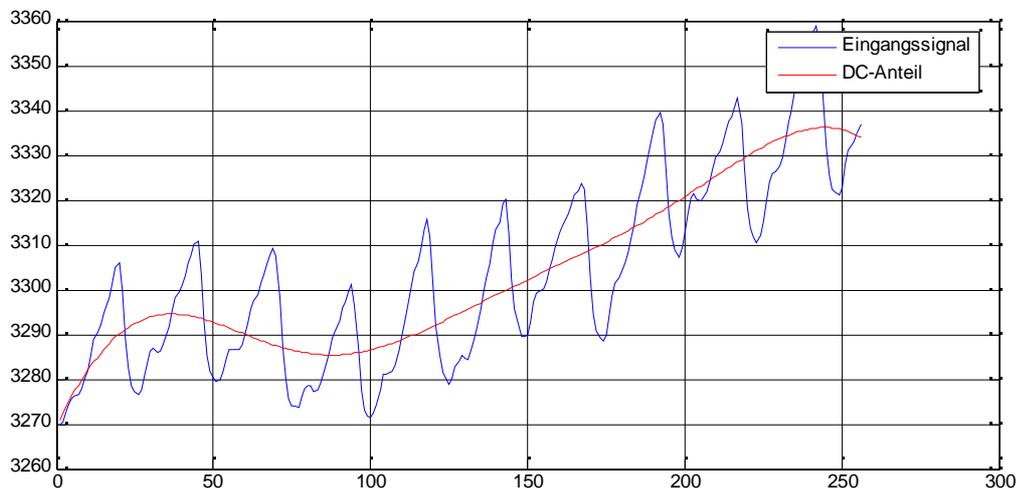


Abbildung 4.19: Beispiel: Ein Block aus dem Infrarot-Kanal (blau), bestehend aus 256 Samples, ergibt mit der Matrix B multipliziert, einen TP-gefilterten Block (rot).

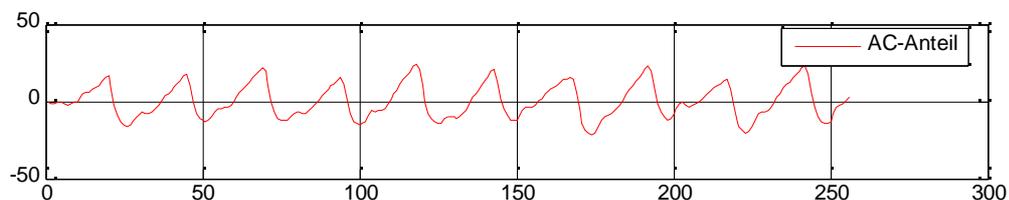


Abbildung 4.20: Das gleiche Eingangssignal, wie in Abb.4.19, mit der Matrix $A = E - B$ multipliziert, ergibt einen HP-gefilterten Block (rot).

4.3.6 Fensterung

Da die Fourier-Transformation eigentlich für das Zeitintervall $]-\infty, +\infty[$ definiert ist, und hier wegen der DFT mit kurzen Zeitausschnitten gearbeitet wird, entstehen nach der Transformation unerwünschte Nebeneffekte. Wird ein Signalauschnitt aus einem Signalverlauf entnommen, spricht man bereits von der sog. Fensterung mit einem Rechteck-Fenster. Es gibt nur einen Fall, in dem ein Rechteck-Fenster dazu geeignet ist, darauf eine DFT durchzuführen. Dies ist der Fall, wenn das Zeitsignal periodisch ist und die DFT eine oder ein Vielfaches der Periode dieses Signals erfasst [Grü]. In allen anderen Fällen führt die Fourier-Transformation mit einem Rechteck-Fenster zur Verschmierung der Frequenzkomponenten im Spektrum. Das kommt durch die Faltung der Fensterfunktion mit den Peaks des eigentlichen Signals im Frequenzbereich zustande. Zur Lösung dieses Problems werden Fensterfunktionen mit

schneller abklingendem Spektrum verwendet. Vor einer DFT muss also ein Signalausschnitt mit einer geeigneten Fensterfunktion multipliziert werden.

Ein sehr oft zu diesem Zweck verwendetes Fenster ist das Kaiser-Fenster, das oft auch Kaiser-Bessel-Fenster genannt wird. Um genauer zu sein, lassen sich die unerwünschten Effekte mit diesem Fenster am besten unterbinden. Mit dem β -Parameter kann das Kaiser-Fenster variabel, je nach Anforderung, angepasst werden. Je größer der β -Parameter, desto kleiner der sog. Leckeffekt [Grü], auf Kosten der spektralen Auflösung. Die Berechnungsvorschrift und die Eigenschaften sind der Literatur zu entnehmen [Grü] [MAT].

Die Fensterung erfolgt mit einem Kaiser-Fenster mit I_B Punkten und dem β -Parameter 8. Der β -Parameter wurde empirisch bestimmt, so dass Schwingungen, die in Anwesenheit einer im Spektrum benachbarten Schwingung deutlich höherer Amplitude, noch erkennbar werden, und gleichzeitig noch genügend spektrale Auflösung vorhanden ist.

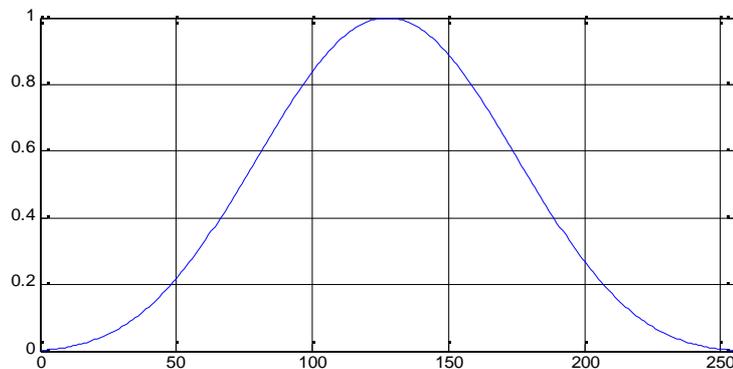


Abbildung 4.21 Kaiser-Bessel-Fenster ζ_{Kaiser} mit $I_B=256$ Punkten und einem $\beta = 8$. Generiert unter MATLAB mit dem Befehl `kaiser(256,8)`.

Nur die pulsartigen Signale aus den Gleichungen 4.39 bis 4.41 werden mit diesem Kaiser-Bessel-Fenster ζ_{Kaiser} im Zeitbereich multipliziert. Die Elemente des Vektors ζ_{Kaiser} können in die Diagonale einer Matrix W übertragen werden:

$$K = \text{diag}(\vec{\zeta}_{\text{Kaiser}}) \quad (4.42)$$

Die AC-Komponenten der beiden LED-Kanäle werden mit dem Fenster gewichtet.

$$\vec{w}_r^A = K \vec{x}_r^A \quad (4.43)$$

$$\vec{w}_i^A = K \vec{x}_i^A \quad (4.44)$$

Auch die Spaltenvektoren der Matrix, die als Interferenz-Referenz dient, werden gewichtet.

$$W_c^A = K X_c^A \quad (4.45)$$

4.3.7 Interference Canceler

Der Interference-Canceler ist ein adaptives Filter, dessen Aufgabe ist, Störungen zu entfernen, die im Übertragungsband lagen und nach der Korrelation mit Dämpfung in das Nutzband gespiegelt worden sind. Aus den Dunkelphasen wurde oben ein Referenzsignal extrahiert, das der gespiegelten Störung im Nutzband entspricht. Dieses Referenzsignal unterscheidet sich in Phase und Amplitude von den in beiden Nutzsignalen (Rot und Infrarot) überlagerten Störungen. Mit dem hier vorgestellten adaptiven Filter ist es möglich, die unerwünschten Spiegelfrequenzen anhand des Referenzsignals aus den Nutzsignalen herauszufiltern.

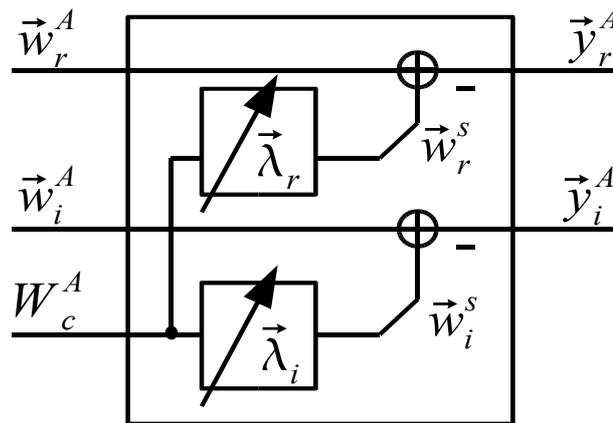


Abbildung 4.22: Modell des adaptiven Filters mit Eingangsgrößen (links) und Ausgangsgrößen (rechts)

Die Filterung findet auch hier in Blockverarbeitung statt. Oben wurde bereits eine Matrix X_c erstellt und mit der Fensterung zur Matrix W_c nach außen hin schwächer gewichtet. Diese Basis dient als Referenz der Störung für das Adaptive Filter. Die Basis W_c spannt einen Raum S auf. S ist der Raum des Störers. Damit ist jede Linearkombination aus der zum gegenwärtigen Zeitpunkt gegebenen Basis ein potentielles Störsignal.

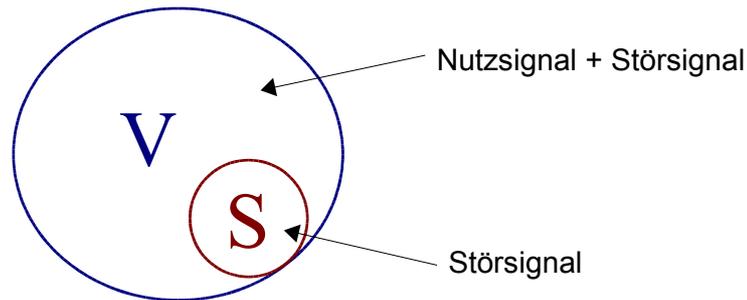
Das Eingangssignal soll Element eines weiteren Vektorraumes V sein, wobei der Vektorraum S ein Untervektorraum von V ist.

$$S \subset V \quad (4.46)$$

Somit kann ein Signal, das Element des Vektorraumes V ist, sowohl Nutzsignal, Störsignal oder beides gleichzeitig sein. Gesucht ist eine Abbildung in Form einer Projektion des Raumes V auf die Differenz der Räume $V \setminus S$, so dass das Eingangssignal am besten ohne Störung rekonstruiert werden kann.

$$V \rightarrow (V \setminus S) \quad (4.47)$$

Nach diesem Prinzip lässt sich die Störung aus dem Nutzsignal entfernen.

Abbildung 4.23: Schema des Raumes V und seines Unterraumes S

Eines der Eingangssignale sei Element des Vektorraumes V :

$$\vec{w}^A \in V = \mathbb{R}^{256} \quad (4.48)$$

Mit der Linearkombination

$$\vec{w}^S = \vec{\lambda} W_c^A \in S \quad (4.49)$$

ließe sich ausschließlich das Störsignal rekonstruieren.

Gesucht ist zunächst eine Linearkombination der Basis W_c , die das Eingangssignal am besten wiedergibt, also eine Rekonstruktion der Störung, so wie sie dem Eingangssignal überlagert wurde.

$$\vec{w}^A \stackrel{!}{=} W_c^A \vec{\lambda} \quad (4.50)$$

Es existieren mehr Gleichungen als Unbekannte, deshalb gibt es in diesem Fall keine exakte Lösung. Ein Vektor λ , der am besten in das überbestimmte Gleichungssystem passt sei gesucht.

$$\|W_c^A \vec{\lambda} - \vec{w}^A\| \rightarrow \mathbf{Minimum} \quad (4.51)$$

Dieses Problem lässt sich mit der Pseudoinversen angehen. Damit erhält man mit dem Vektor λ eine Linearkombination von W_c , mit der sich der Störer beschreiben lässt:

$$(W_c^A)^\# \vec{w}^A = \vec{\lambda} \quad (4.52)$$

Mit der Projektion

$$f_B : V \rightarrow S \quad (4.53)$$

lässt sich der Störer aus dem Eingangssignal rekonstruieren:

$$\vec{w}^S = W_c^A (W_c^A)^\# \vec{w}^A \quad (4.54)$$

Weiterhin kann aus Abb. 4.22 entnommen werden, dass die Differenz aus geschätztem Störer und Eingangssignal das gefilterte Signal ergibt:

$$\vec{y}^A = \vec{w}^A - \vec{w}^s \quad (4.55)$$

Gleichung 4.54 eingesetzt in Gleichung 4.55 ergibt das adaptiv gefilterte Signal:

$$\vec{y}^A = \vec{w}^A - W_c^A (W_c^A)^\# \vec{w}^A \quad (4.56)$$

Dies entspricht der Projektion:

$$f_B : V \rightarrow (V \setminus S) \quad (4.57)$$

Eine weitere Umformung zeigt den adaptiven Interference-Canceler, so wie er für einen Kanal in der Simulation unter MATLAB eingesetzt wird.

$$\vec{y}^A = \vec{w}^A \left(E - W_c^A (W_c^A)^\# \right) \quad (4.58)$$

Die Unterdrückung des Störers gilt nur unter der Annahme, dass die Basis W_c tatsächlich nur den Raum des Störers aufspannt, sonst werden auch gewünschte Signalanteile herausgefiltert.

4.3.8 Normierung

Die Normierung basiert auf dem analytischen Modell aus Kapitel 2.3.3 und ist hier durch die Gleichungen

$$\vec{\eta}_i = \frac{\vec{y}_i^A}{\vec{x}_i^D} \quad \text{und} \quad \vec{\eta}_r = \frac{\vec{y}_r^A}{\vec{x}_r^D} \quad (4.59)$$

beschrieben. Die Division erfolgt punktweise. Die Eingangsvektoren y sind zuvor durch die Fensterung gewichtet worden, demnach wird diese Gewichtung auch auf das normierte Signal übertragen.

4.3.9 Fourier Transformation

Die normierten und gewichteten Signale sind proportional zu den arteriellen Volumenänderungen, so dass genau diese normierten Signale im Spektrum bearbeitet werden sollen. Unter MATLAB wird jeder Block mit einer Länge von $l_B = 256$ Samples mittels der FFT in den Frequenzbereich transformiert.

Die DFT ist definiert mit:

$$y(k) = \sum_{n=0}^{N-1} \eta(n) (\omega_N)^{nk} \quad (4.60)$$

Wobei der Drehfaktor ω_N schon zuvor als Konstante berechnet werden kann, da $N = l_B$ in dem hier vorgestellten SpO₂-Algorithmus konstant ist:

$$\omega_N = e^{-j \frac{2\pi}{N}} \quad (4.61)$$

Nach der Transformation in den Frequenzbereich liegen jeweils zwei Spektren zu den Kanälen Rot und Infrarot vor. Die DFT liefert mit $l_B = 256$ und $k = [0, 255]$ insgesamt 256 Fourierkoeffizienten im Spektrum. Da das Zeitsignal mit

$$\vec{\eta} \in \mathbb{R}^{l_B} \quad (4.62)$$

reell ist, sind davon $l_B/2 = 128$ Koeffizienten für die weitere Verarbeitung relevant. Die restlichen 128 Koeffizienten enthalten dieselbe Information. Es reicht also, wenn die DFT für $k = [0, 127]$ berechnet wird.

Da N bekannt ist, lässt sich auch der Term $(\omega_N)^{nk}$ für alle Kombinationen n und k berechnen, mit $k = [0, 127]$ und $n = [0, 255]$. Dadurch ließe sich die DFT durch eine Matrix F mit komplexen Elementen ersetzen:

$$[F]_{kn} = (\omega_N)^{nk} \quad (4.63)$$

In einem Echtzeitsystem würden die komplexen Elemente der Matrix F abgespeichert werden, und durch Multiplikation der Matrix F mit dem Zeitsignal könnte die Fouriertransformierte berechnet werden:

$$\vec{y} = F \vec{\eta} \quad (4.64)$$

Mit einer 128×256 -Matrix ergibt das zusammen $N^2 = \text{MAC-Operationen}^8$ für den Imaginär- und den Realteil. Dies ist im Vergleich zu Gleichung 4.60 eine Ersparnis an Rechenzeit, da fortan komplexe und zeitaufwendige Operationen entfallen. Noch schneller lässt sich die DFT mit einem der FFT-Algorithmen berechnen. Hier reichen $N \cdot \log_2(N)$ Operationen aus [Grü]. Die

8 Der MAC-Befehl ist eine Operation, die heute in keinem digitalen Signalprozessor fehlt. Mit einer Anweisung wird eine Multiplikation mit einer anschließenden Addition durchgeführt.

Fourier-Transformierte wird jeweils für Rot und Infrarot benötigt, worauf zwei Vektoren mit komplexen Spektren zur weiteren Verarbeitung im Frequenzbereich entstehen.

$$\vec{y}_r \in \mathbb{C}^{128}, \quad \vec{y}_i \in \mathbb{C}^{128} \quad \text{für } l_B = 256 \quad (4.65)$$

Nach der Fouriertransformation können die Spektren in verschiedenen Ansichten dargestellt werden. Eine Möglichkeit ist, das Spektrum für einen Block zu zeigen, wie in Abb. 4.24. Möchte man den Verlauf der Spektren über die Zeit sehen, dann ist das Spektrogramm aus Abb. 4.25 am besten geeignet. Schließlich stellt S. Kästle in [Käs] seine Spektren in einem Nadeldiagramm (X-Y-Plot) dar, wie in Abb. 4.26 für den Fall von Störungen und Abb. 4.27 für den Fall mit weniger gestörten Signalen (Rot und Infrarot). Ohne Einfluss von Störungen würde ein solches Nadeldiagramm eine Gerade zeigen, deren Steigung das Ratio ρ wäre.

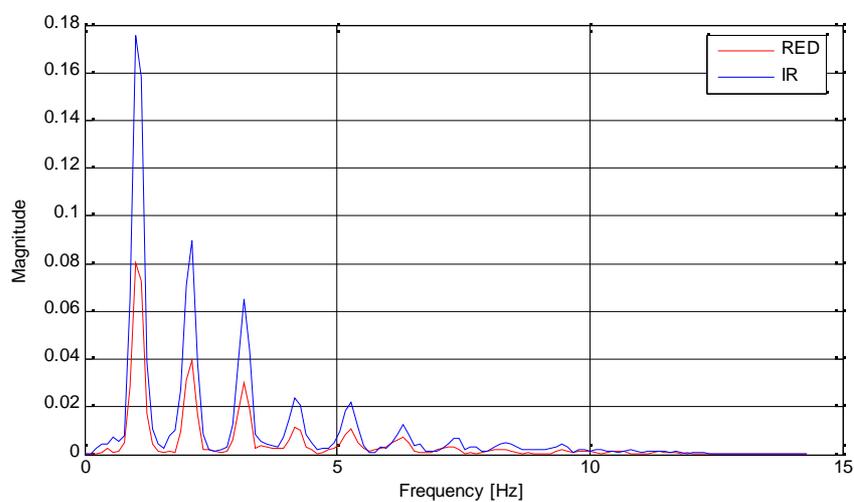


Abbildung 4.24: Zwei Spektren der normierten Signale Rot und Infrarot für einen Block. Das Spektrum zeigt ein Signal bei guten Bedingungen.

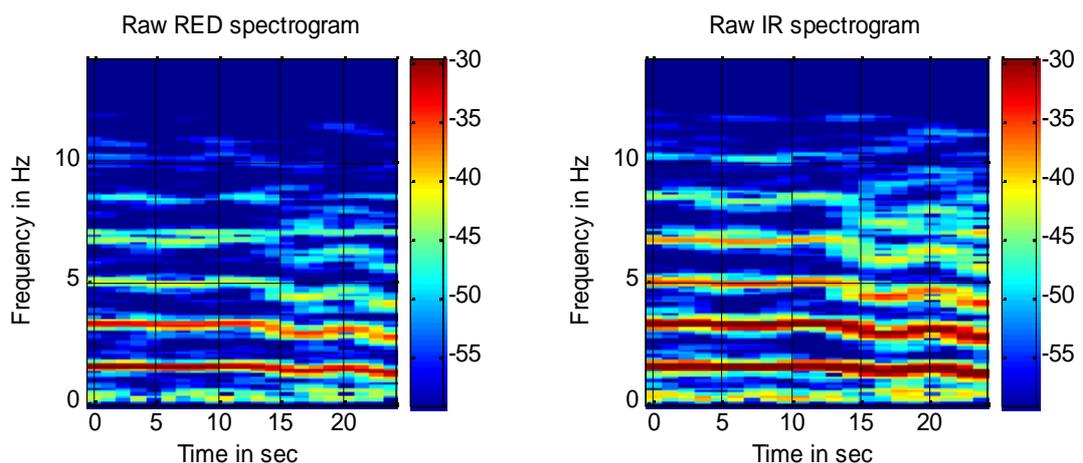


Abbildung 4.25: Figure 4 aus der Simulation unter MATLAB zeigt ein Spektrogramm je Kanal

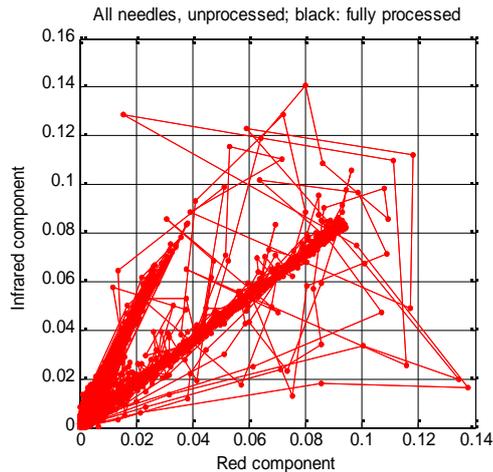


Abbildung 4.26: Nadeldiagramm mit Störung

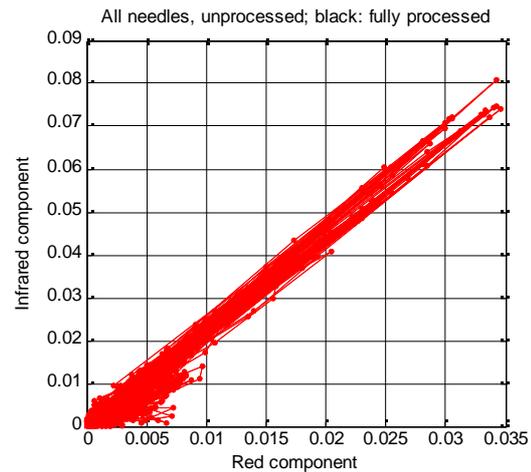


Abbildung 4.27: Nadeldiagramm ohne Störung

4.3.10 Total Least-Squares Fit 1

Die nächsten beiden Algorithmen, die zur Generierung der Spektralen Maske 4.3.11 und Bestimmung der Herzfrequenz 4.3.12 dienen, benötigen ein Spektrum als Berechnungsgrundlage. Da zwei verschiedene Spektren vorliegen, wird in diesem Modul versucht, ein Referenz-Spektrum aus den beiden Spektren zu extrahieren. Nach der DFT liegen die zwei Spektren vor, deren Zeitsignale ohne Einfluss von Störungen perfekt miteinander korreliert und deren Spektren linear abhängig wären. Die Spektren werden hier als Vektoren geschrieben, und man kann versuchen, das eine Spektrum als Linearkombination des anderen auszudrücken:

$$\vec{y}_r \stackrel{!}{=} \rho \vec{y}_i \quad (4.66)$$

Das Gleichungssystem ist überbestimmt und nicht lösbar. Nebenbei angemerkt: Wäre die Gleichung lösbar, also würden keine Störungen vorliegen, dann wäre der Skalar ρ schon das Verhältnis, das mithilfe der Kalibrierungsfunktion den gewünschten SpO_2 -Wert liefert. Man kann die Gleichung nicht lösen, es existiert aber ein ρ , das am besten in die Gleichung passt. Das Problem kann veranschaulicht werden, indem die beiden Spektren in einem x-y-Plot aufgetragen werden. So ein IR- vs RED-Spektrum-Plot wird auch in Figure 6 während der Simulation mit dem M-Skript `spo2_test3.m` für alle Blöcke k_2 generiert.

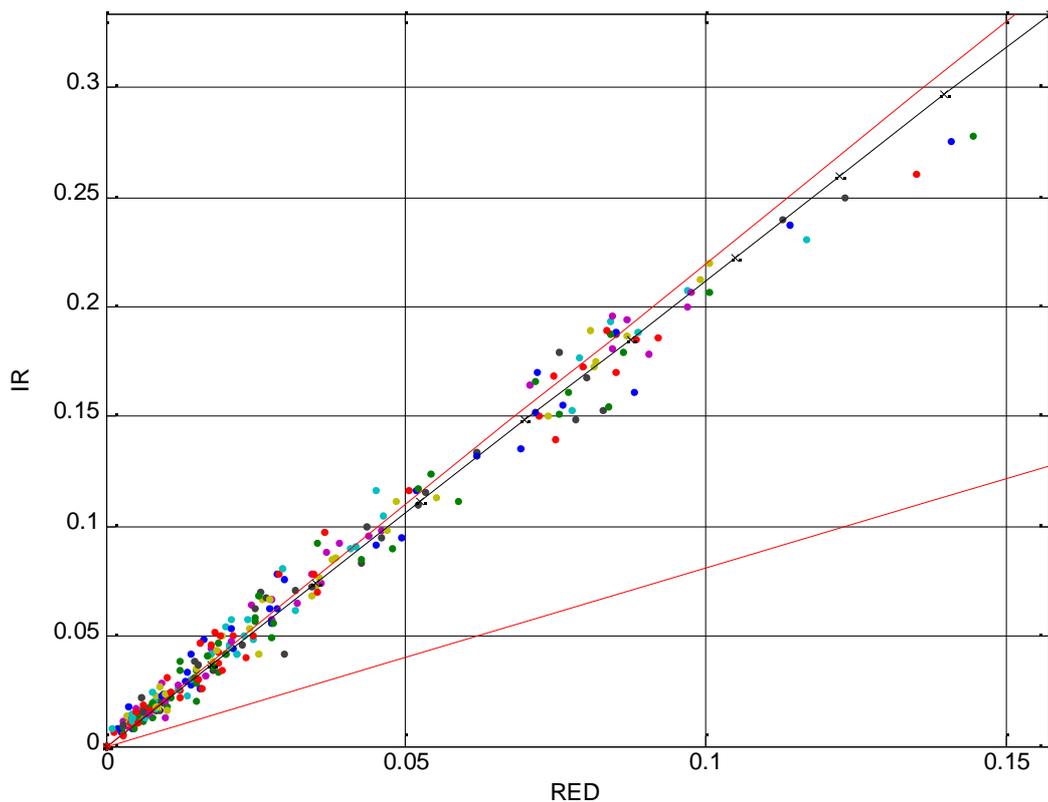


Abbildung 4.28: IR- vs RED-Spektrum zu mehreren Zeitpunkten k_2 . Die roten Linien sind die Spo2-Grenzen 100% (oben) und 70% (unten). Die schwarze Linie hat die Steigung ρ , die mit einem Total Least-Squares Fit ermittelt wurde.

Aufgrund von Störungen sind die Spektren unterschiedlich. Wären die Spektren identisch, würden die Punkte im Graphen oben eine perfekte Ursprungsgerade⁹ ergeben. Zur Lösung dieses Problems, soll eine Linearkombination ρ gesucht werden, die am besten in Gleichung 4.66 passt. Allerdings muss die Gleichung um zwei Störvektoren \vec{i}_r und \vec{i}_i erweitert werden, da aufgrund von Bewegungsartefakten oder weißem Rauschen zwei unterschiedliche Störsignale zu jedem Kanal additiv überlagert wurden. Die Spektren können in einen reinen, störungsfreien Puls-Signal-Anteil ($\vec{y}_{r,p}$ $\vec{y}_{i,p}$) und einen additiven Störungsanteil (\vec{i}_r \vec{i}_i) zerlegt werden.

$$\vec{y}_r = \vec{y}_{r,p} + \vec{i}_r \quad \text{und} \quad \vec{y}_i = \vec{y}_{i,p} + \vec{i}_i \quad (4.67)$$

Das überbestimmte Gleichungssystem lautet dann:

$$\vec{i}_r + \vec{y}_{r,p} = \rho(\vec{y}_{i,p} + \vec{i}_i) \quad (4.68)$$

⁹ Gerade, die einen Punkt im Ursprung (0,0) hat

Das Problem lässt sich mit der Methode der kleinsten Quadrate (Least-Squares Fit) nicht lösen, da dies nur dann möglich wäre, wenn nur eins der beiden Signale mit Störungen überlagert worden wäre.

$$\vec{i} + \vec{y}_{r,p} \stackrel{!}{=} \rho \vec{y}_{i,p} \quad (4.69)$$

Zum Vergleich wird in Abb. 4.25 der Unterschied zwischen dem *Least-Squares Fit* und dem Total Least-Squares Fit veranschaulicht. In beiden Fällen wird versucht, die Summe der Quadrate der rot eingezeichneten Distanzen zu den Punkten zu minimieren, wobei weiter vom Ursprung liegende Punkte stärker gewichtet werden, was durchaus gewünscht ist, wenn man bedenkt, dass Koeffizienten, wie die der Grundwelle, höher gewertet werden müssen, als die der Oberwellen. Die Steigung der Geraden ist in beiden Fällen das Verhältnis ρ , allerdings bringt nur ein Total Least-Squares Fit die beste Näherung, weil nur hier die tatsächliche Distanz zu den Punkten berücksichtigt wird. Anders als bei einem reinen Least-Squares Fit, wird hier nur die vertikale Distanz zu den Punkten minimiert.

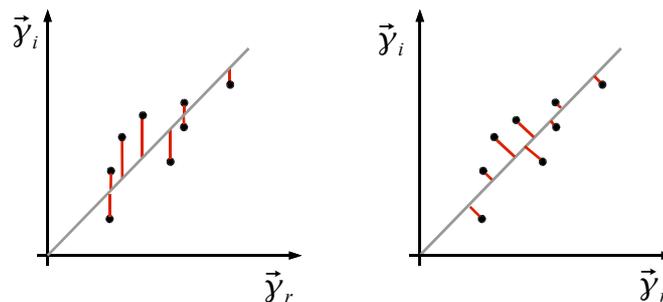


Abbildung 4.29: Unterschied zwischen einem Least-Squares Fit als Minimierung der vertikalen Distanz und einem Total Least-Squares Fit als Minimierung der tatsächlichen Distanz zur Linie.

Zur Lösung dieses Problems nach Gl. 4.68 eignet sich die Singularwert-Zerlegung (engl. SVD = Singular Value Decomposition), mit der sich so ein Total Least-Squares Fit vollziehen lässt [NUM]. SVD basiert auf der Zerlegung einer beliebigen Matrix A in drei Matrizen:

$$A = U \Sigma V^+ \quad (4.70)$$

Das Produkt der drei Matrizen der SVD ergibt wieder die Matrix A . Mit dieser Zerlegung lässt sich eine Vielzahl von Problemen aus der linearen Algebra lösen. Für die hier vorgestellte Problematik ist speziell die Matrix Σ von Bedeutung. Die diagonalen Elemente der Matrix Σ werden Singularwerte genannt. Zur weiteren Untersuchung wird eine Matrix A aus den zwei Spaltenvektoren zusammengesetzt, die jeweils das Spektrum der ungestörten Signale mit ihren additiven Störern repräsentieren:

$$A = \begin{bmatrix} \vec{y}_r & \vec{y}_i \end{bmatrix} \in \mathbb{C}^{128 \times 2} \quad (4.71)$$

Der Rang dieser Matrix ist mit absolut hoher Wahrscheinlichkeit 2, da es aufgrund der additiven Störer absolut unwahrscheinlich ist, dass die Spaltenvektoren linear abhängig sind. Mittels der SVD lässt sich nicht nur der Rang einer Matrix ermitteln, es lässt sich auch untersuchen, wie nahe (nach der L_2 -Norm) eine Matrix A am niedrigeren Rang ist. Und schließlich kann zur Matrix A eine möglichst ähnliche Matrix gefunden werden, deren Rang niedriger als der der Matrix A ist. Ließe sich der Rang der Matrix A auf 1 reduzieren, so würden die beiden Spaltenvektoren linear abhängig werden, und die Spektren wären bis auf den Faktor ρ gleich. Eines der Spektren könnte dann zur weiteren Berechnung genommen werden, da es fortan die Information beider Kanäle tragen würde. Führt man für A eine SVD durch, so erhält man für Σ eine Matrix der Form:

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$$

Aus [NUM] geht hervor, dass der Rang einer Matrix A der Anzahl aller von Null verschiedenen Singularwerte entspricht. Mit $\text{rang}(A) = 2$ sind beide Singularwerte größer 0:

$$\Sigma_1 \geq \Sigma_2 > 0 \quad (4.72)$$

Das Element Σ_2 verrät bereits, wie nahe die Matrix Σ am nächsten niedrigen Rang ist, also in diesem Fall, wie nahe Matrix Σ am Rang 1 ist. Daraus folgt: Je mehr Störungen den Spektren überlagert sind, desto größer ist Σ_2 . Der Rang kann auf 1 reduziert werden, indem $\Sigma_2 = 0$ gesetzt wird [NUM]. Demnach können alle Störungen, die in den beiden Spektren verschieden sind, mit der Reduktion auf Rang 1 eliminiert werden. Die Rekonstruktion der Matrix A ergibt eine neue Matrix mit dem Rang 1:

$$\tilde{A} = U \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} V^+ \quad (4.73)$$

Das Resultat ist ein komplexer *Total Least-Squares Fit* der beiden mit Störungen überlagerten Spaltenvektoren. Zur Erinnerung: Die beiden Spaltenvektoren sind komplexe Spektren. Die zwei so errechneten Vektoren sind linear abhängig, und stellen zugleich die beste Näherung an die mit Störungen überlagerten Spektren dar. Das bedeutet, dass alle Störungen entfernt wurden, die in beiden Spektren unterschiedlich waren, wie z.B. weißes Rauschen.

$$\tilde{A} = [\vec{y}_{r,f} \quad \vec{y}_{i,f}] \in \mathbb{C}^{128 \times 2} \quad (4.74)$$

Die Gleichung 4.75 ist lösbar und liefert das Verhältnis ρ zur Berechnung der Sauerstoffsättigung.

$$\vec{y}_{r,f} = \rho \vec{y}_{i,f} \quad (4.75)$$

Das Verhältnis wird an dieser Stelle noch nicht verwendet. Es besteht noch ein Problem, wenn sich eine Störung hoher Amplitude in beiden Spektren identisch überlagert. Diese Art von

Störungen zu filtern, ist Aufgabe der *Spektralen Maske*. Als Berechnungsgrundlage benötigt die Spektrale Maske ein Referenz-Spektrum, das repräsentativ für beide Kanäle ist. Genau dafür kann eines der Spektren aus 4.74 genommen werden, z.B. das IR-Spektrum:

$$\vec{\gamma}_c = \vec{\gamma}_{i,f} \in \mathbb{C}^{128} \quad (4.76)$$

4.3.11 Spektrale Maske

Die Funktion der Spektralen Maske ist, die Fourier-Koeffizienten des Puls-Signals im Spektrum zu suchen, um alle anderen Koeffizienten auf Null zu setzen, die nicht zum Puls-Signal gehören. Denn alle anderen Koeffizienten tragen zu einem falschen Ratio bei. Dies ist am besten ersichtlich, wenn man zunächst punktweise den Quotienten der beiden gestörten Spektren (Rot und Infrarot) bildet, so dass ein Ratio über der Frequenz entsteht, wie in Abb. 4.30 oben. Vergleicht man anschließend das Ratio mit dem Referenz-Spektrum in Abb. 4.30 unten über mehrere Zeitfenster, so wird deutlich, dass das Ratio nur über den Frequenzkomponenten des Puls-Signals korrekt ist, und für all diese Frequenzen gleich ist.

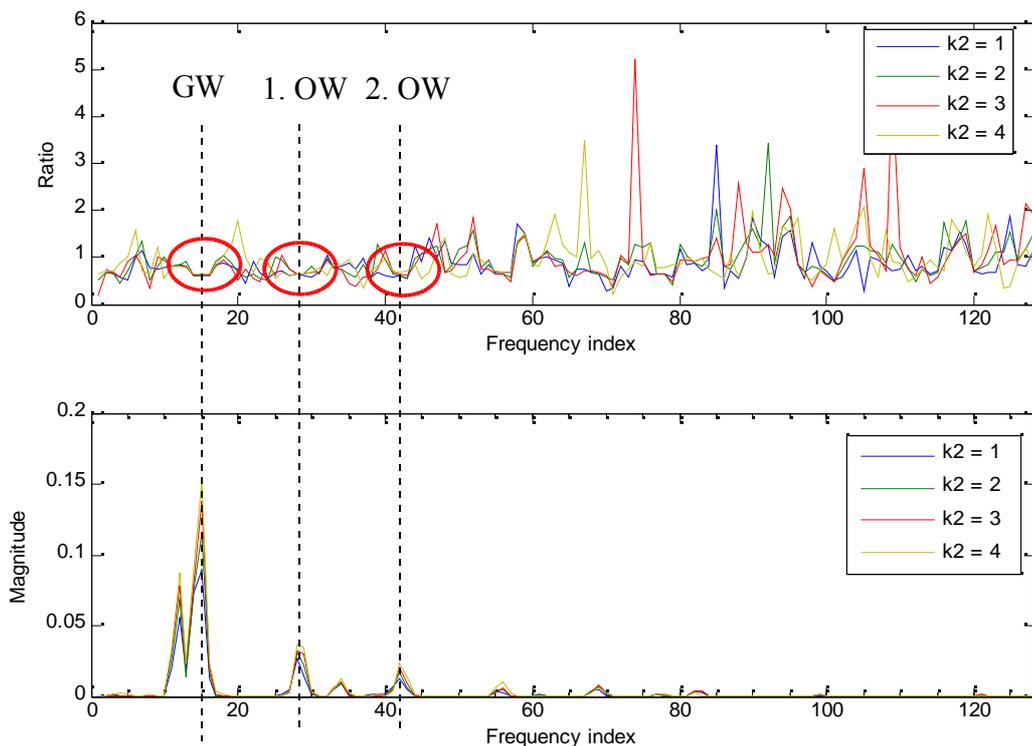


Abbildung 4.30: Zusammenhang zwischen dem Ratio und der Amplitude im Spektrum, betrachtet über die Frequenz und zu vier Zeitpunkten $k_2 = [1, 4]$. Oben das Ratio aus den Vektoren der beiden Spektren; Unten das Referenzspektrum γ_c

Dies führt zunächst zu keinem nennenswerten Fehler in der Berechnung des Ratio mit der SVD, denn bei einem *Least-Squares Fit* wird das Ratio bei höheren Amplituden stärker gewichtet als bei niedrigen Amplituden. Problematisch wird es nur dann, wenn die Amplituden der Störungen größer als die der Pulswelle werden, so wie in Abb. 4.31 dargestellt. Dann wird das Ratio bei den Frequenzen der Störer stärker gewichtet, als die der Pulswelle. Das Ratio bei den Frequenzen der Störer ist undefiniert und hat keinen Bezug zur Sauerstoffsättigung. Erfahrungsgemäß beträgt das Ratio bei solch dominanten Peaks im Spektrum 1, denn die Störung ist meistens auf Lichtquellen zurückzuführen, die in beiden Spektren mit der gleichen Amplitude vorzufinden sind. Ein Ratio bedeutet je nach Kalibrierungskurve etwa 80% Sauerstoffsättigung. Ohne der hier vorgestellten Maske würden im Spektrum dominant auftretende Störungen zu einem falschen SpO_2 -Wert von etwa 80% führen, ohne dabei die tatsächliche Sättigung relevant zu berücksichtigen.

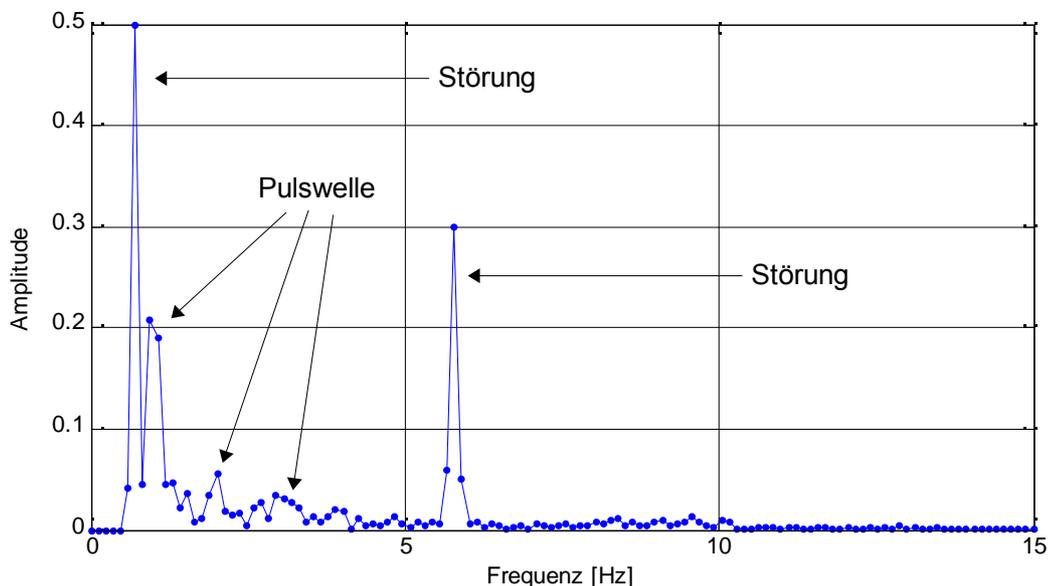


Abbildung 4.31: Beispiel für das Spektrum eines stark gestörten Signals

Es lohnt sich nach einem Algorithmus zu suchen, der in der Lage ist, die Frequenzkomponenten der Pulswelle von denen der Störer zu unterscheiden. Der Algorithmus, hier genannt *Spectral Mask*, setzt sich aus zwei Teilalgorithmen zusammen, die beide eine binäre Maske mit den Elementen $\{0,1\}$ liefern, mit der das Spektrum punktweise multipliziert werden kann, um so die nicht zum Puls-Signal gehörenden Fourier-Koeffizienten zu unterdrücken. Die Implementierung unter MATLAB ist das M-Skript *tresmask.m*

4.3.11.1 Methode der Schwellwertfunktion

Der zentrale Gedanke in diesem Algorithmus beruht auf der Annahme, dass die Koeffizienten, die auf das Puls-Signal zurückzuführen sind, in der Regel auffällig aus dem Spektrum herausragen, so wie es in Abb. 4.30 unten der Fall ist. Die Idee aus [Käs] ist, das Spektrum

kurzfristig als Zeitsignal zu behandeln, und es mit einem Tiefpass-FIR-Filter zu filtern. Dabei spielen die Parameter des Filters keine große Rolle. Die normierte Grenzfrequenz bei 0.3 sowie die 5. Ordnung des Filters wurden empirisch ermittelt, bis sich der gewünschte Effekt wie in Abb. 4.32 gezeigt hat.

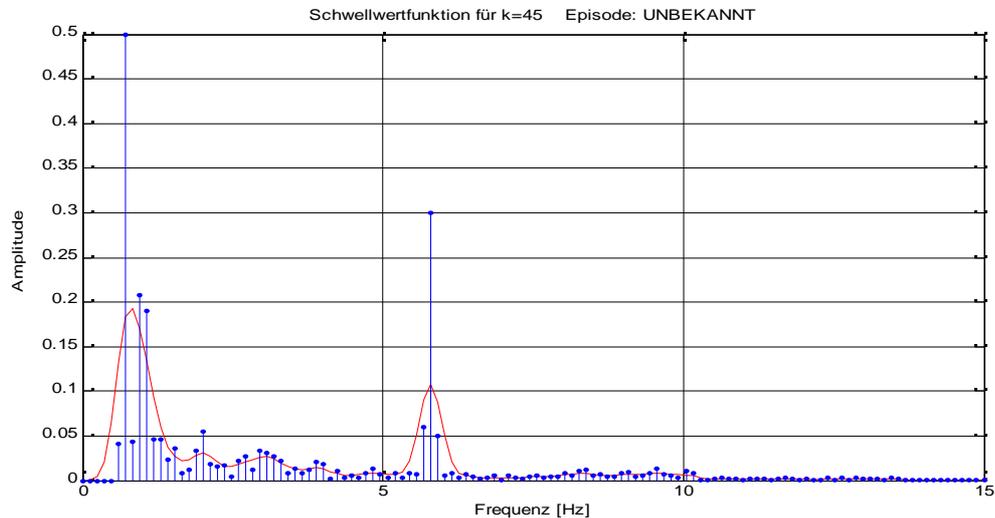


Abbildung 4.32: Schwellwertfunktion (rot) und Referenz-Spektrum (blau).

Alle Koeffizienten des Referenz-Spektrums, die oberhalb der Schwellwertfunktion liegen, werden ausgewählt. Das bedeutet, dass eine binäre Maske mit den Elementen $\{0,1\}$ erstellt wird, die dieselbe Länge, wie das Spektrum hat, und immer dort eine Eins hat, wo der Koeffizient des Referenz-Spektrums die Schwellwertfunktion übertagt.

Wie in Abb. 4.32 dargestellt, funktioniert diese Methode für Störer wesentlich kleinerer Amplitude als die des Puls-Signals. Die Methode versagt, sobald ein Störer im Spektrum als alleinstehender Peak auffällt.

4.3.11.2 Methode der Harmonischen Beziehung

Der Algorithmus nach der Methode der harmonischen Beziehung basiert auf Erkenntnissen aus Untersuchungen zahlreicher Puls-Signale auf ihre spektralen Eigenschaften. Die fundamentale Erkenntnis ist die harmonische Beziehung der drei relevanten Frequenzen (f_g Grundwelle, f_{o1} erste Oberwelle und f_{o2} zweite Oberwelle):

$$f_{o1} = 2 \cdot f_g, \quad f_{o2} = 3 \cdot f_g \quad (4.77)$$

Dieser Beziehung lässt sich für eine beliebige Frequenz innerhalb des gegebenen Referenz-Spektrums suchen.

$$|\vec{y}_c| \in \mathbb{R}^{128} \quad (4.78)$$

Dazu wurde ein M-Skript *spo2mask.m* entwickelt, das als Übergabeparameter eine Herzfrequenz erhält. Ohne zu „wissen“, ob bei dieser Frequenz tatsächlich die Grundwelle des Puls-Signals liegt, legt das M-Skript eine binäre Maske über das Spektrum nach der Beziehung aus 4.77, so dass alle Frequenzen ausserhalb der Maske auf 0 gesetzt werden. Aus den verbleibenden Koeffizienten wird eine Summe gebildet, und diese der vorgegebenen Frequenz zugeordnet. Dieser Vorgang wird mit dem M-Skript *spo2fcs.m* für alle möglichen Herzfrequenzen [30,300] Hz wiederholt, und anschließend die Frequenz selektiert, bei der die Summe maximal wird.

Bei der Maximum-Suche werden die Werte nicht beachtet, für die die betreffende Iteration die Bedingung des abklingenden Spektrums eines Puls-Signals nicht erfüllt.

$$|X(f_g)| > |X(f_{o1})| > |X(f_{o2})| \quad (4.79)$$

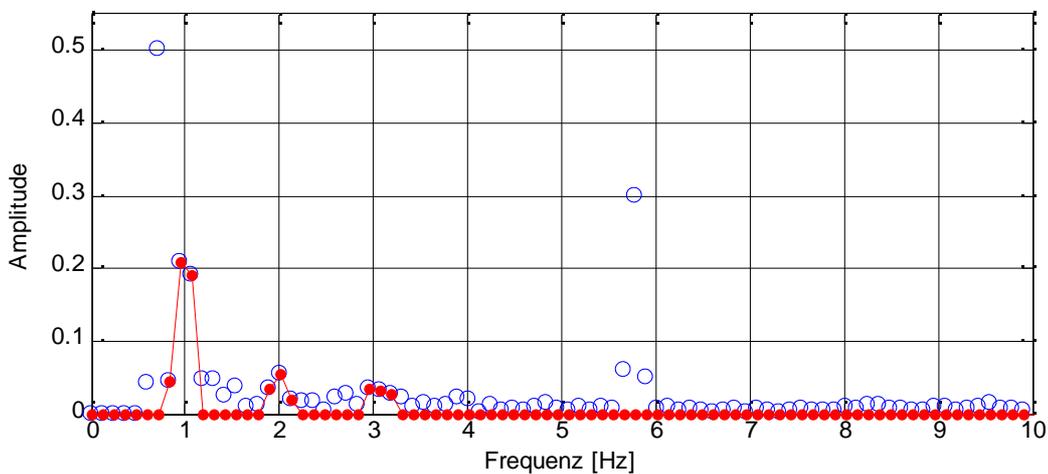


Abbildung 4.33: Referenz-Spektrum (blau) und die nach Anwendung der binären Maske ausgewählten Koeffizienten. Die beiden Störer sind im neuen Spektrum (rot) nicht mehr vorzufinden.

4.3.11.3 Verknüpfung der binären Masken

Vergleicht man Abb. 4.32 und 4.33, so selektiert die Maske nach der Methode der harmonischen Beziehung das reine Puls-Signal wesentlich effizienter. Dennoch hat die Schwellwertfunktion-Methode den Vorteil, dass nur die deutlich herausragenden Peaks selektiert werden. Um beide Vorteile zu vereinen, werden die beiden binären Masken mittels UND-Verknüpfung zu einer binären Maske γ_m verknüpft.

$$\vec{\gamma}_m = (\vec{\gamma}_{Mask1} \wedge \vec{\gamma}_{Mask2}) \in \{0,1\} \quad (4.80)$$

Die beiden komplexen Spektren werden anschließend mit der Maske γ_m punktweise multipliziert.

$$\vec{y}_r^m = \vec{y}_m \cdot \vec{y}_r, \quad \vec{y}_i^m = \vec{y}_m \cdot \vec{y}_i \quad (4.81)$$

Über den Übergabeparameter *maskflag* an das M-Skript *spo2_test3.m* kann die Maske ab- oder -zugeschaltet werden.

4.3.12 Berechnung der Herzfrequenz

Zum Ermitteln der Herzfrequenz aus dem Referenz-Spektrum γ_c dient das M-Skript *hr.m*. Die Frequenz der Grundwelle ist die Herzfrequenz. Üblicherweise gibt man die Herzfrequenz in Schlägen pro Minute an, so dass die Frequenz mit dem Faktor 60 multipliziert, die Herzrate (eng. Heart rate) in bpm ergibt.

Das M-Skript greift auf die Frequenz der Grundwelle zurück, die schon beim Erstellen der Masken näherungsweise ermittelt wurde. Näherungsweise, weil sich aus einem Spektrum mit 128 Punkten und mit einer reinen Maximum-Suche keine Auflösung von einem bpm realisieren lässt, wie in Kap. 2.1 gefordert. Mit einer Abtastfrequenz $f_1 = 28.5743$ Hz, ergibt sich für die maximale Frequenz im Spektrum 14.2872 Hz. Mit 128 Punkten im Spektrum, ergibt sich eine Auflösung von 6.7bpm:

$$\Delta HR = \frac{f_1 \cdot 60}{2 \cdot 128} \approx 6.7 \text{ bpm} \quad (4.82)$$

Damit erreicht man die selbst gestellte Forderung, mit einer Auflösung von einem bpm, nicht. Die gesuchte Information steckt dennoch im Spektrum. Die Grundwelle ist im Spektrum auf mehrere Koeffizienten verteilt. Ob die unmittelbaren Koeffizienten links oder rechts vom Maximum der Grundwelle höher liegen, hängt davon ab, ob die Grundwelle tatsächlich mehr links oder rechts vom Grundwellenmaximum im Spektrum liegt. Die Schwerpunktsfrequenz der Grundwelle kann mit der Gleichung

$$f_g = \frac{\sum_{i=f_u}^{f_o} i \cdot \gamma_c(i)}{\sum_{i=f_u}^{f_o} \gamma_c(i)} \quad (4.83)$$

aus dem Referenz-Spektrum γ_c gewonnen werden. Dadurch lassen sich weitaus höhere Auflösungen erreichen [Käs]. Die untere Fußfrequenz f_u und die obere Fußfrequenz f_o können anhand des Maximums der Grundwelle gefunden werden. Man nimmt dazu 1 oder 2 Koeffizienten jeweils links und rechts vom Maximum der Grundwelle zur Berechnung nach der Gleichung der Schwerpunktsfrequenz.

4.3.13 Total Least-Squares Fit 2

Wie schon oben erwähnt, könnte man das Verhältnis ρ schon aus dem ersten Total Least-Squares Fit berechnen, allerdings führen auch hier Störungen hoher Amplitude zu einem Fehler in der Messgröße. Nach der Multiplikation mit der Maske wurden nur die relevanten Frequenzkomponenten selektiert. Nun erfolgt erneut ein komplexer *Total Least-Squares Fit* mittels Singularwertzerlegung (SVD), nach dem gleichen Prinzip wie in Kap. 4.3.10. Das unbestimmte Gleichungssystem lautet:

$$\vec{y}_r^m \stackrel{!}{=} \rho \vec{y}_i^m \quad (4.84)$$

Mit der Matrix A , die wiederum aus zwei komplexen Vektoren besteht

$$A = \begin{bmatrix} \vec{y}_r^m & \vec{y}_i^m \end{bmatrix} \in \mathbb{C}^{128 \times 2} \quad (4.85)$$

wird die Singularwert-Zerlegung durchgeführt.

$$A = U \Sigma V^+ \quad (4.86)$$

Der Least-Squares Fit führt zur Eliminierung der noch verbleibenden Differenzen aus den beiden komplexen Spektren. Dies kann erreicht werden, indem alle Singularwerte aus Σ bis auf den Hauptwert $[\Sigma]_{11}$ auf Null gesetzt werden [NUM] :

$$\tilde{A} = U \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} V^+ \quad (4.87)$$

Es entsteht eine neue Matrix A mit dem Rang 1. Diese Matrix enthält die komplexen Spektren, deren Differenzen beseitigt wurden.

$$\tilde{A} = \begin{bmatrix} \vec{y}_{r,f} & \vec{y}_{i,f} \end{bmatrix} \in \mathbb{C}^{128 \times 2} \quad (4.88)$$

So dass gilt:

$$\vec{y}_{r,f} = \rho \vec{y}_{i,f} \quad (4.89)$$

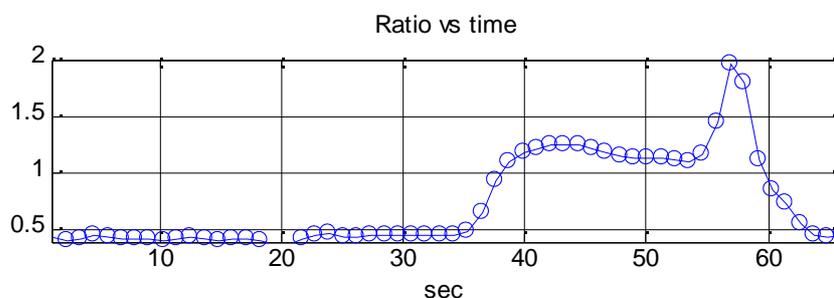


Abbildung 4.34: Das Verhältnis ρ über die Zeit

Auf diese Art lässt sich zu jedem Spektren-Paar das Verhältnis ρ errechnen. Mit der Kalibrierungsfunktion lässt sich anschließend die Sauerstoffsättigung errechnen.

4.3.14 Kalibrierungsfunktion

Die Kalibrierungsfunktion liefert SpO_2 -Werte zum Verhältnis ρ nach Gleichung 2.39. Da das Gerät nicht kalibriert wurde, und das Verhältnis ρ zum Testen des Algorithmus ungeeignet erscheint, wurde eine Kalibrierungsfunktion aus [Käs] entnommen. Das bedeutet, dass die Kalibrierungsfunktion nicht auf exakt das hier entwickelte Gerät, in Kombination mit dem verwendeten Sensor, abgestimmt ist. Damit sind die SpO_2 -Werte mit einem Fehler behaftet, bieten aber genügend Aussagekraft, um den Algorithmus auf die Unterdrückung von Störungen zu testen. Letztendlich geht es um die relativen Abweichungen des SpO_2 -Wertes beim Eintreten einer Störung.

Die Kalibrierungsfunktion nach [Käs] ist eine Parabel der Form:

$$SpO_2(\rho) = 115 - 25\rho - 8\rho^2 \quad (4.90)$$

Bei gesunden Probanden erhält man mit dieser Funktion, mit dem hier vorgestellten System, bestehend aus Hardware und Sensor, einen SpO_2 -Wert von etwa 102 %. Um diesen optischen Fehler auszugleichen, wird eine alternative Funktion verwendet:

$$SpO_2(\rho) = 113 - 25\rho - 8\rho^2 \quad (4.91)$$

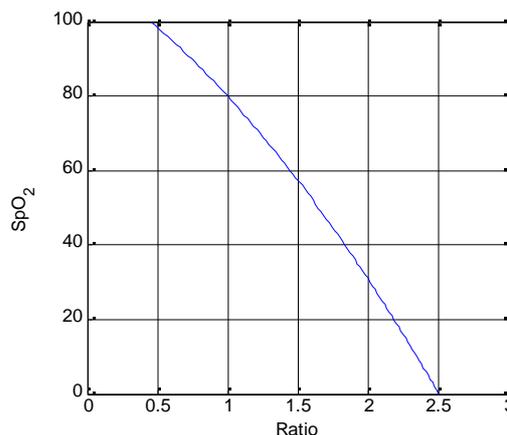


Abbildung 4.35: Kalibrierungsfunktion

Mit beiden Funktionen lassen sich relative Abweichungen der Sauerstoffsättigung in Abhängigkeit von Störeinflüssen beschreiben.

5 Hardware – Entwicklung

5.1 MORES ein Remissions-Sensor

Der Mikro-Optische Remissions-Sensor (MORES) hat eine zentrale Bedeutung in dieser Arbeit. MORES wurde vorwiegend für medizinische Diagnostik durch das CiS Institut für Mikrosensorik in Kooperation mit MIRA GmbH Berlin entwickelt. Mit seinen geringen Abmessungen $8 \times 8 \times 1,3 \text{ mm}^3$ und dem geringen Gewicht, eignet sich MORES hervorragend zum Einsatz in kompakten Pulsoximetern.

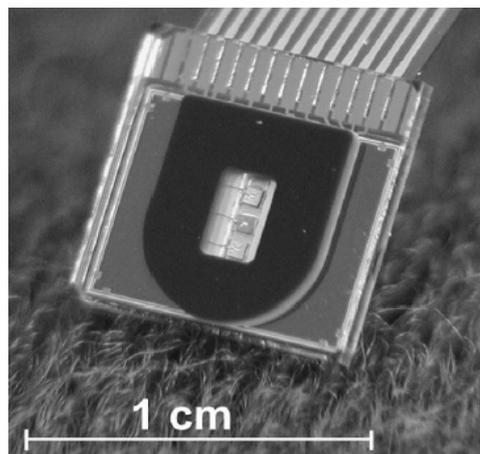


Abbildung 5.1: MORES
Quelle: MORES-Datenblatt

In einem Remissions-Sensor sind Sender und Empfänger zum Vermessen der Lichtabsorption eines Objekts untergebracht. Durch bis zu drei LEDs, verschiedener, frei wählbarer Wellenlängen im Bereich 450 bis 890 nm, ist die Vermessung der Lichtabsorption auch in Abhängigkeit der Wellenlänge möglich. Die reflektierte Intensität kann mittels der auf dem Sensor integrierten zwei pin-Dioden¹⁰ erfasst werden. In dieser Arbeit soll auf dem bereits vorliegenden Typ MORES01-3 aufgebaut werden. Zur Messung der funktionellen Sauerstoffsättigung ist MORES01-3 mit geeigneten Wellenlängen konfiguriert. Zwei davon werden für die funktionelle Sauerstoffsättigung benötigt. Die dritte Wellenlänge könnte später zur Messung des COHb-Anteils und der Bestimmung der fraktionellen Sauerstoffsättigung hinzugenommen werden.

¹⁰ Der wesentliche Vorteil von pin-Dioden ist eine Grenzfrequenz von bis zu 1GHz, während diese bei Photodioden bei 10 MHz liegt [TS].

Bezeichnung	Spektralbereiche	Wellenlängen-Maximum	Toleranz
RED	rot	660nm	+/-5nm
RIR	rot / nahinfrarot	760nm	+/-8nm
IR	infrarot	880nm	+/-15nm

Tabelle 5.1: Wellenlängen des MORES01-3

Als optische Empfänger sind zwei pin-Dioden auf dem Substrat des MORES01-3 integriert, so dass sie vor direktem Übersprechen der LED-Strahlung im Wellenlängenbereich 660 bis 880 nm mit einer Dämpfung von Faktor 40 bis 80 geschützt sind. Dämpfungen Faktor $>10^3$ sind laut Datenblatt auf Anforderung möglich. Bei Anwendung auf der Haut, führt die optische Barriere zu einer Vermessung der Absorption in tieferen Gewebeschichten. Dies ist ein positiver Effekt, denn die ersten Schichten der Haut sind kaum mit Kapillaren versehen, so dass die Vermessung der Absorption erst in tieferen Regionen sinnvoll ist, wobei die exakte Messtiefe abhängig vom Messort und Proband ist.

Die zwei Photodioden des Sensors können zu einem stärkeren Signal zusammengefasst oder getrennt zur Detektion von relativen Rotationseffekten erfasst werden. In dieser Arbeit werden die Photoströme durch Parallelschalten der pin-Dioden aufaddiert. In diesem Entwicklungsstadium wäre es noch zu früh, sich auf die Rotationseffekte zu konzentrieren.

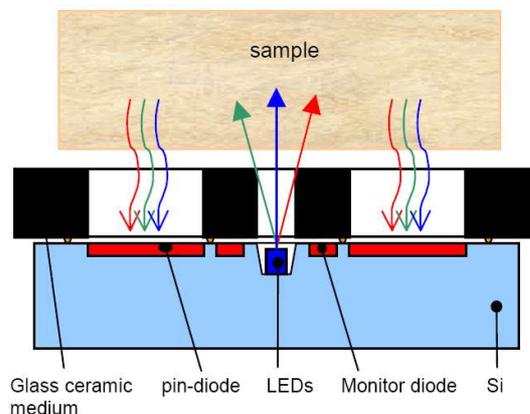


Abbildung 5.2: MORES-Querschnitt
Quelle: MORES-Datenblatt

Weiterhin ist auf dem Substrat eine Monitor-Diode integriert, die ohne optische Barrieren in der Nähe der LEDs angebracht ist. Die Monitor-Diode kann verwendet werden, um die absolute Strahlungsintensität der LEDs vor dem Eindringen in das Messobjekt zu vermessen. Dabei wurde auf minimale Toleranzen der gesamten, im System befindlichen Photoelemente geachtet. Nach dem hier verwendeten analytischen Modell, ist der Einsatz einer solchen Monitor-Diode nicht notwendig.

MORES verfügt zusätzlich über einen Temperatursensor, der zur Messung der lokalen Körpertemperatur verwendet werden kann. Es wäre vorstellbar, ein mit dem Sensor ausgestattetes Pulsoximeter um diesen Vitalparameter zu erweitern. Zudem besteht ein Zusammenhang zwischen Körpertemperatur und Extinktionskoeffizienten [DOP].

Die Anschlüsse des MORES sind über ein feines Flachbandkabel herausgeführt, das zum Einklemmen in einem Flex-Verbinder vorgesehen ist. Um an die feinen Anschlüsse heranzukommen, wird das Flachbandkabel in einen Flexverbinder-Adapter eingeklemmt.

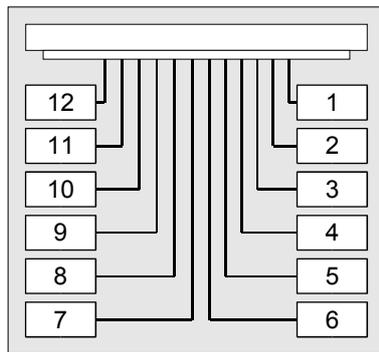


Abbildung 5.3: Flexverbinder - Adapter

Pin	Belegung
1	LED 660 nm, Anode
2	LED 760 nm, Kathode
3	LED 880 nm, Kathode
4	GND (LED)
5	GND (LED)
6	GND (Schutzring)
7	Si-Substrat, Kathode
8	Si-Substrat, Kathode
9	Pin-Diode 1, Anode
10	Monitordiode, Anode
11	Temperatur-Diode, Anode
12	Pin-Diode 2, Anode

Tabelle 5.2: Anschlussbelegung des Adapters in Verbindung mit MORES

5.2 TriCore – Entwicklungsumgebung

5.2.1 TriCore 1765 – Mikrocontroller

Der TriCore 1765 ist ein 32-bit-Mikrocontroller aus Infineons AUTMDO-Familie, der primär zum Einsatz in der Automobiltechnik, und im Allgemeinen für Applikationen entwickelt wurde, die hohe Performance in Echtzeitverarbeitung erfordern. Dazu wurden im TriCore, wie der Name deutet, drei Konzepte vereint: Eine Mikrocontroller-, eine RISC- und ein DSP-Architektur. Die DSP-Architektur beinhaltet vor allem den zum Entwickeln effizienter Algorithmen unentbehrlichen MAC-Befehl, die Parallelausführung via SMID und verschiedene Adressierungsarten (z.B. Ring-Buffer oder Bit-Reverse). Mit RISC-Architektur lassen sich Optimierungsstrategien von C-Compilern besser umsetzen. Der Mikrocontroller verfügt über eine vier Stufen-Pipeline, wie sie schon in dem Infineon C166 eingesetzt wurde. Mit einem zusätzlichen 16-Bit-Satz der wichtigsten Befehle, ist es möglich, Leistungsaufnahme und Befehlsspeicherbedarf zu senken, falls 16 Bit Operationen genügen. Alle Operationen sind für Integer und Fixed-Point ausgelegt. Zusätzlich ist eine Floating-Point-Einheit mit einem kleinen Befehlssatz integriert. Ein schnelles Interrupt-System, mit dem sich Interrupt-Prioritäten in Gruppen definieren lassen, erlaubt eine flexible Kontrolle der Peripherie. Intern verfügt der TC1765 über 48kByte an SRAM für zeitkritischen Befehlscode oder Daten.

Der TC1765 ist für einen Automobil Temperaturbereich von -40°C bis +125°C bei 40MHz ausgelegt, und ist in einem P-LBGA-260-Gehäuse verbaut. Die maximale Leistungsaufnahme beträgt nach Datenblatt max. 1.2W bei 2.5V Betriebsspannung und 40MHz. Per Software lässt sich der Mikrocontroller in stromsparende Zustände versetzen.

Es sind interne Taktfrequenzen bis zu 40 MHz mittels dem eingebauten PLL möglich. In der hier vorgestellten Anwendung wird der TC1765 extern mit 16MHz getaktet und läuft intern mit 40MHz. Mittels der PLL kann die interne Taktfrequenz via Software variiert werden.

Auf dem Chip sind zahlreiche Peripherie-Einheiten integriert. Aufgrund der Vielzahl werden hier nur die für das Projekt relevanten genannt. Dazu gehören:

- General Purpose Timer Unit (GPTU) mit drei 32-bit Timern/Zählern, die sich nochmals in kleinere Timer separieren lassen;
- zwei synchrone serielle Schnittstellen (SSC0, SSC1) zur Kommunikation mit externer Peripherie, wie z.B. AD-Wandlern oder einstellbaren Widerständen;
- zwei serielle, asynchrone Schnittstellen (ASC0, ASC1) mit variabler Baudrate, wovon eine während der Entwicklungsphase zur Kommunikation mit dem PC verwendet wird;
- zwei AD-Wandler-Einheiten (ADC0, ADC1) mit jeweils 16 Kanälen über analoge Multiplexer, wovon in diesem Projekt nur einer der Eingänge mit der maximalen Quantisierung von 12 Bit betrieben wird;
- die insgesamt 77 digitalen Ein- und Ausgänge.

5.2.2 Evaluation-Board (TriBoard)

Das 100mm x 160mm große Evaluation-Board (TriBoard 1765) besteht aus dem Mikrocontroller TriCore 1765 von Infineon und externer Peripherie, wie 1 MByte RAM, 2 MByte FLASH, RS232-Schnittstelle und einem JTAG-Port.

Die asynchrone Schnittstelle (ASC0) ist über den Pegelwandler-IC MAX232 auf eine 9-polige SUB-D-Buchse als RS232-Schnittstelle herausgeführt. Hier wird ein PC angeschlossen, um Daten zwischen dem Mikrocontroller und dem PC austauschen zu können. Die maximale Übertragungsgeschwindigkeit ist durch den UART im PC auf 115200 BAUD begrenzt. Das sind 11520 Byte pro Sekunde, wenn pro Byte ein Start-, Stop-Bit und 8 Bit an Daten anfallen.

Die Entwicklungsplatine wird über den JTAG-Port mit dem Parallelport des PCs verbunden. Damit lässt sich der Code in das RAM auf der Entwicklungsplatine hochladen, um das laufende Programm zu debuggen.

Alle Anschlüsse des Mikrocontrollers sind über vier 80-pin-Verbinder jeweils an der unteren und oberen Seite der Platine herausgeführt. Alle Anschlüsse sind funktionsgemäß auf die vier Verbinder (X801 bis X804) aufgeteilt.

- X801: Adressbus, Datenbus und Chip-Select-Leitungen
- X802: Synchrone Serielle Schnittstelle (SSC) oder alternativ GPIO (Ports 0 und 5)
- X803: AD-Wandler, Referenzspannungen für AD-Wandler oder alternativ GPIO (Port 0)
- X804: GTPU, GTPA oder alternativ GPIO (Ports 1 bis 4)

Jeder Anschluss der internen Peripherie kann alternativ als General Purpose I/O (GPIO) konfiguriert werden, also als einer der 77 digitalen Ein- oder Ausgänge. Die 77 GPIO sind wiederum zu 5 parallelen Ports gruppiert.

Port	Port-Breite	Alternative Funktion
P0	16 Bit	GPTU, ASC0, SSC0, CAN, ADC0, ADC1
P1	16 Bit	GPTA
P2	16 Bit	GPTA
P3	16 Bit	GPTA
P4	8 Bit	GPTA
P5	5 Bit	ASC1, SSC1

Tabelle 5.3: Die alternativen Funktionen der General Purpose I/O

Die Entwicklungsplatine wird mit einer stabilisierten Spannung zwischen 6V und 60V gespeist. Nach drei Spannungswandlern stehen intern 5V, 3.3V und 2.5V für Mikrocontroller und Peripherie zur Verfügung. Auf dem Board befinden sich drei grüne LEDs, die den gültigen

Zustand der drei Spannungen anzeigen. Beim Anlegen der Betriebsspannung wird ein systemweiter Power-On-Reset durchgeführt.

Eine gelbe LED steht zur freien Disposition, und ist über das Board fest mit P0.5 des Mikrocontrollers verbunden. Meint man einen bestimmten Anschluss eines Ports, z.B. den Anschluss 0 an Port 2, lautet die Notation: P2.0

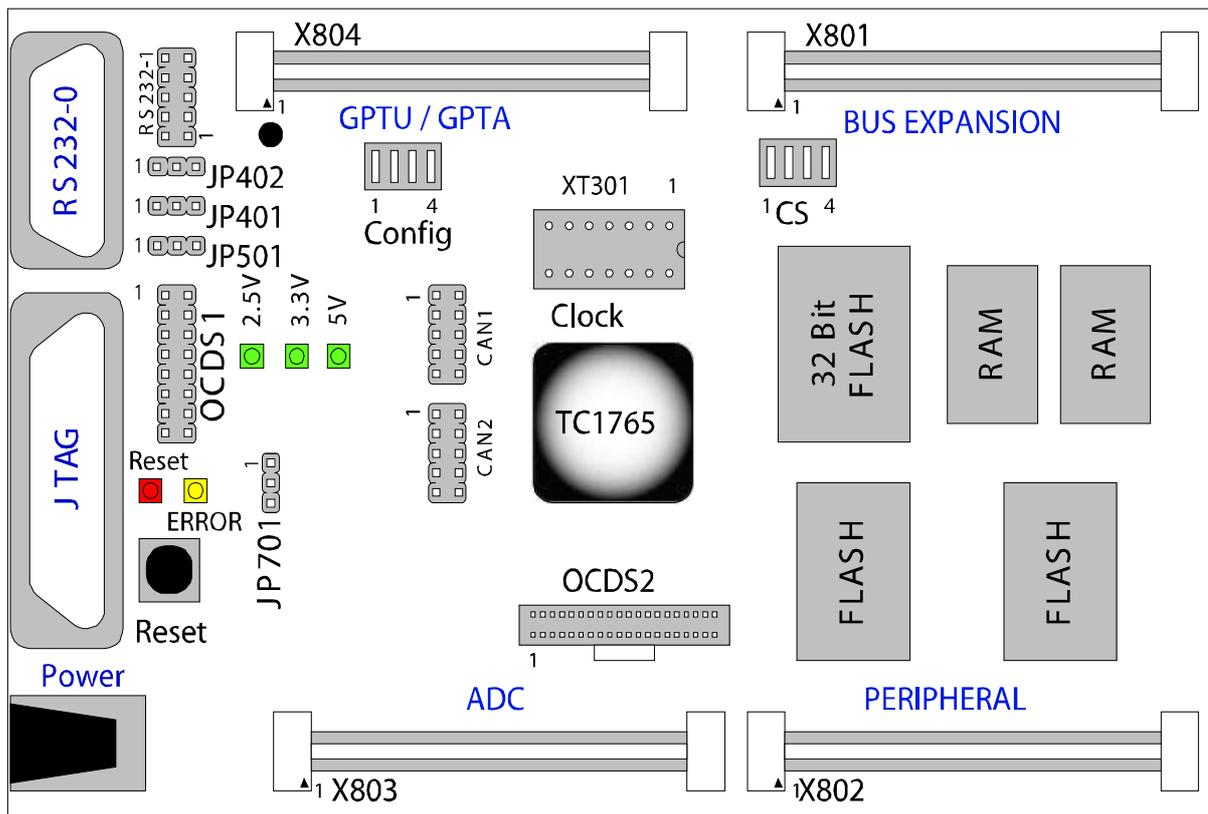


Abbildung 5.4: Das TriBoard mit dem TriCore 1765 und externer Peripherie, den 80-poligen Verbindern X801 bis X804, RS232-Anschluss, JTAG-Interface zum Debuggen, Stromversorgung und status LEDs

Quelle. TriBoard 1765 Hardware Manual

Alle Angaben zur Entwicklungsplatine stammen aus dem Handbuch TriBoard TC1765 - Hardware Manual TC1765-200.

5.2.3 Analog-Digital-Wandler

Ein im Mikrocontroller integrierter AD-Wandler verringert Entwicklungszeit und Produktionskosten. Der TriCore 1765 verfügt über zwei AD-Wandler mit variabler Quantisierung von 8, 10 und 12 Bit und einem Sample-and-Hold-Glied. Prinzipiell verfügt jeder AD-Wandler über einen zuschaltbaren Multiplexer, der in diesem Projekt per Software überbrückt wird, da nur ein Kanal des AD-Wandlers benötigt wird. Die Abtastung kann durch Software-Anweisung

oder für präzise Äquidistanz via Timer-Interrupt getriggert werden. Wird eine AD-Wandlung eingeleitet, so wird ein Kondensator des betreffenden AD-Wandlers mit dem gewünschten Analog-Eingang verbunden. Die Zeit, um die Kapazität zu laden, wird Sample-Zeit t_s bezeichnet. Im nächsten Schritt wird der Kondensator vom Eingang getrennt und es erfolgt eine AD-Wandlung der am Kondensator anliegenden Spannung. Die Dauer der gesamten AD-Wandlung ist die Zeit t_c (conversion time). Die Sample-Zeit errechnet sich aus dem Takt f_{BC} der gesamten Peripherie-Einheit ADC0, zu der neben dem AD-Wandler auch Kontroll-Logik gehört. Dieser Takt f_{BC} wird per Software mit Faktor 2 auf die Hälfte des Systemtaktes geteilt, also wird die AD-Wandler-Einheit ADC0 mit $f_{BC} = 20\text{MHz}$ getaktet. Die Sample-Zeit lässt sich unabhängig davon via Software mit dem Parameter STC variieren; dieser wird auf 0 gesetzt. Das ergibt mit der folgenden Gleichung eine Sample-Zeit von $1.2\ \mu\text{s}$.

$$t_s = (STC + 2) \cdot \frac{4}{f_{BC}} \quad (5.1)$$

Um die gesamte Wandlungszeit zu berechnen muss noch die Taktfrequenz f_{DIV} bekannt sein, mit der allein der AD-Wandler getaktet wird. Dieser Takt wird per Software auf 6.667MHz eingestellt. Die Gleichung 5.2 gilt für eine 12-Bit-Quantisierung und ergibt mit der hier vorgestellten Konfiguration eine AD-Wandlungszeit von $9.7\ \mu\text{s}$.

$$t_c = t_s + \frac{56}{f_{BC}} + \frac{2}{f_{DIV}} \quad (5.2)$$

Daraus folgt: Mit einer maximalen Quantisierung von 12 Bit, benötigt ein AD-Wandler $9.7\ \mu\text{s}$. Damit sind, unter Verwendung eines einzigen analogen Eingangs, Abtastfrequenzen bis zu 100kHz möglich. Alle Angaben zum AD-Wandler stammen aus dem Handbuch *TC1765 - Peripheral Units*.

5.3 Analoge Verarbeitung

Ohne analoge Aufbereitung ist DSV nicht möglich, dennoch wird versucht, die Zahl der analogen Komponenten, soweit es möglich ist, minimal zu halten, um primär Kosten in der Serienfertigung zu sparen. Für den Sensor MORES wurde ein Analog-Modul konzipiert, das im Wesentlichen aus einem LED-Treiber, einem Strom-zu-Spannung-Wandler und einem Verstärker besteht. Diese Elektronik wurde auf einer durchkontaktierten 1.27mm -Lochraster-Platine untergebracht, die über die Verbinder X801 und X802 auf das TriBoard gesteckt wird. Vorwiegend wurden SMD-Bauteile verwendet, die sich auf der Oberseite der Platine befinden. Auf der Unterseite wurden die Bauteile mit Lackdraht verdrahtet. Dabei wurde darauf geachtet, dass sich zu Gunsten der EMV digitale und analoge Leitungen nicht kreuzen. Das steckbare Modul erlaubt es schnell an beide Seiten der Elektronik zu gelangen um Änderungen vorzunehmen, dennoch ist das gesamte Gerät auf einem Chassis untergebracht.

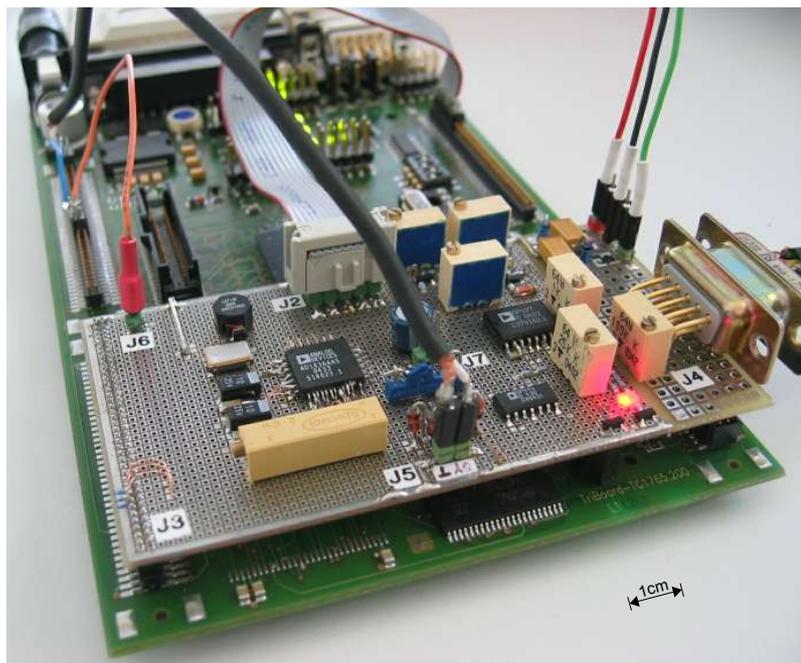


Abbildung 5.5: TriBoard mit dem analogen SpO₂-Modul

Für die analogen Komponenten wird eine separate Spannungsversorgung von -5V und 5V verwendet. Der wesentliche Grund für die Wahl der negativen Spannung ist das Design des MORES mit seinen LEDs, die abwechselnd mit Kathode oder Anode an GND verbunden sind. Die Schaltung der Spannungsversorgung wurde einfach gehalten, da davon ausgegangen wird, dass sich das Gerät während der Entwicklungszeit an einem stabilisierten Labornetzteil befindet. Mit den Zener-Dioden D7 und D8 wurde lediglich für einen Überspannungs- und Verpolungsschutz gesorgt. Damit dieser Schutz gewährleistet wird, muss das Labornetzteil bei Strömen >20mA abschalten oder zumindest den Strom begrenzen. Die Kondensatoren C1 und C2 dienen durch Kompensation von Stromspitzen einer lokalen Spannungsstabilisierung. Die Versorgungsspannungen werden am Anschluss J1 angeschlossen.

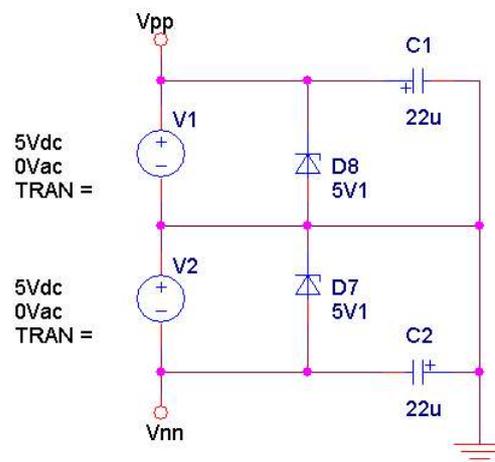


Abbildung 5.6: Spannungsversorgung der analogen Schaltkreise

Die maximale Leistungsaufnahme der analogen Schaltkreise beträgt 135mW. Zusammen mit dem Mikrocontroller (ohne externe Peripherie) beträgt die maximal zu erwartende Leistungsaufnahme bei Vollast 1.3W.

5.3.1 Ansteuerung der LEDs

Für alle LEDs wurden Treiber entwickelt und getestet, obwohl für die funktionelle Sauerstoffsättigung zwei Wellenlängen ausreichen. Die 760nm-LED ist bereits mit einem Ausgang des Mikrocontrollers verbunden, wird aber nicht im Algorithmus berücksichtigt, und bleibt somit permanent dunkel.

Das Design des MORES macht es erforderlich, negative und positive Spannungen zum Betreiben der LEDs zu verwenden, weil eine LED mit der Kathode an Masse gebunden ist, während die anderen zwei LEDs mit den Anoden an Masse gebunden sind. Diese Verbindung befindet sich auf dem Si-Substrat und konnte deshalb nicht aufgetrennt werden.

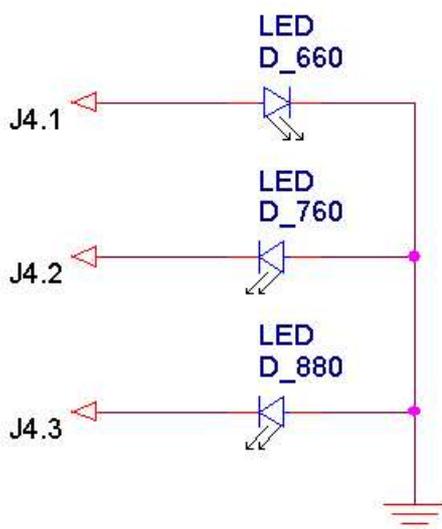


Abbildung 5.7: MORES01-3

Die LED-Treiber werden mittels zwei Bipolar-Transistoren realisiert. Die Transistoren werden als Schalter betrieben. Die Strombegrenzung erfolgt rein durch Vorwiderstände an den LEDs. Probleme mit dieser Schaltung könnten entstehen, wenn die Intensität der LEDs zu stark in Frequenz und Amplitude schwankt. Das ist ein Kriterium für die konstante Stromquelle, das es zu untersuchen galt.

Mit der Strombegrenzung durch einen Widerstand wird der Strom von Schwankungen der Versorgungsspannung und Temperatur abhängig. Die lineare Abhängigkeit beträgt im Mittel für alle LED-Treiber 6.5mA/V. Das bedeutet für Spannungsschwankungen von 50mV, dass sich der Strom um 300µA verändert, was zu einer gerade noch bemerkbaren Schwankung der Intensität

der LEDs führt, die sich nach Reflexion im Photostrom der pin-Diode bemerkbar macht, was je nach Frequenz zum Fehler der Messgröße führen kann.

Die Frequenzanfälligkeit kann definiert werden. Nach dem analytischen Modell, so wie es hier verwendet wird, ist eine konstante Intensität der LEDs gefordert. Die absolute Intensität einer LED und die Differenz der Intensitäten der LEDs ist dabei nicht von Bedeutung, da dies durch die Normierung im analytischen Modell bereits berücksichtigt wird. Wichtig ist nur, dass sich die Intensität der LEDs nicht zu schnell ändert, d.h. die Intensität darf nicht in dem AC-Anteil des Algorithmus bemerkbar auffallen. Dazu müsste die Intensität mit Frequenzen > 0.5 Hz alternieren, was hier mit Schwankungen der Umgebungstemperatur nicht erreicht wird. Das einzige Problem bei der Strombegrenzung mit Widerständen könnte also entstehen, wenn die Versorgungsspannung mit Frequenzen > 0.5 Hz und zugleich mit Amplituden in der Größenordnung von > 50 mV schwankt, wie bereits oben definiert. Es ist also dafür zu sorgen, dass die Spannungsversorgung diese Kriterien nicht verletzt. Dies wurde mit Labornetzteilen und stabilisierenden Kapazitäten erfüllt, und als fertiges Produkt würden später Batterien als konstante Spannungsquelle dienen.

Ein anderer Aspekt ist die Zerstörung von Bauteilen durch zu hohe Ströme, deren Ursache in der Temperaturabhängigkeit der Bauteile liegt. Es konnte gezeigt werden, dass Temperaturschwankungen in der Größenordnung $\pm 5^\circ\text{C}$ den Vorwärtsstrom der LEDs nicht messbar veränderten. Zum Betrieb in Räumen bei 15 bis 25°C ist die Funktion der Schaltung zunächst sichergestellt, was zum Testen des Algorithmus genügt.

Die Vorwiderstände der LEDs können durch ihre Vorwärtsspannung U_F und dem maximalen Vorwärtsstrom I_F der LEDs bestimmt werden. Dabei wird auch die Collector-Emitter-Sättigungsspannung berücksichtigt, die bei den hier verwendeten Transistoren BC848 und BC858 (vergleichbar mit den bedrahteten Typen BC547 und BC557), bei einer Temperatur von 25°C bei $U_{CESat} = 100$ mV liegt. Die Widerstände wurden mit der folgenden Gleichung näherungsweise bestimmt und in einen festen und einstellbaren Widerstand aufgeteilt, so dass eine Einstellung der exakten Ströme mittels 2-Kanal-Oszilloskop möglich wird.

$$R_v = \frac{U_B - U_{CESat} - U_F}{I_F} \quad (5.3)$$

LED	I_F [mA]	U_F [V]
660nm	20	1.9
760nm	20	1.7
880nm	20	1.35

Tabelle 5.4: Vorwärtsspannung U_F und maximaler Vorwärtsstrom I_F der MORES-LEDs nach Angaben des Herstellers

Über den 10pol-Verbinder J2 sind alle Treiberstufen an digitale Ausgänge des Mikrocontrollers angeschlossen. Nähere Angaben zu Verbindern können Kapitel 5.3.4 entnommen werden. Die 660nm-LED wird mit einer positiven Spannung angesteuert, dazu dient ein PNP-Transistor, der mit einem Basiswiderstand von $6.81k\Omega$ bei einem LO-Pegel (TTL) des Mikrocontrollers in Sättigung geht, und so der zuvor eingestellte Kollektorstrom fließen kann. Der genaue Kollektor (= Diodenstrom) kann mittels des Widerstandes R_2 für die 660nm-LED eingestellt werden. Die 660nm-LED wurde sicherheitshalber kurz vor dem Maximum (20mA) auf 19.8mA eingestellt.

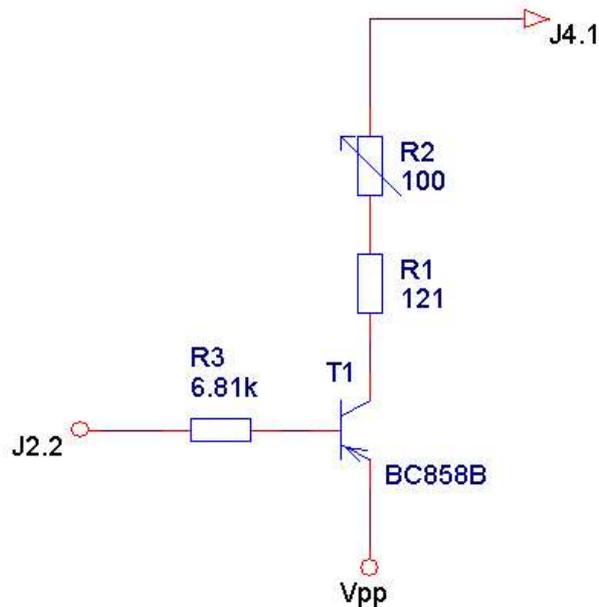


Abbildung 5.8: Die Treiberstufe für die rote LED mit einer Wellenlänge von 660nm

Die 760nm- und die 880nm-LED werden bei Anlegen einer negativen Spannung aktiv. Dazu wurde ein NPN-Transistor verwendet. Damit dieser mit TTL-Pegeln von 0V und 5V angesteuert werden kann, musste eine Z-Diode eingesetzt werden, um das Spannungspotential an der Basis um ca. 5V zu senken. Damit lassen sich Basis-Emitter-Spannungen von 5V und 0V einstellen. Bei 5V Basis-Emitter-Spannung geht der Transistor in Sättigung. Damit aktiviert ein HI-Pegel des Mikrocontrollers die LED. Der Strom der 760nm-LED kann mit R_4 eingestellt werden. Die LED wird nicht zum gegenwärtigen Entwicklungsstand im Algorithmus berücksichtigt, bleibt damit während der Messung dunkel.

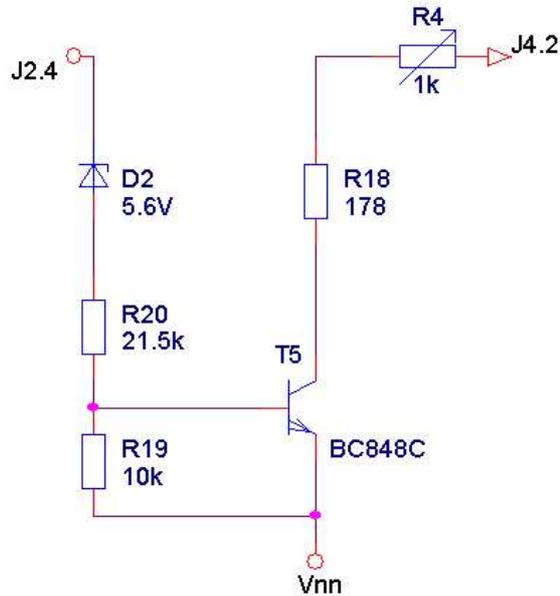


Abbildung 5.9: Die Treiberstufe für die Infrarot-LED mit einer Wellenlänge von 760nm

Nach dem gleichen Prinzip wurde die Schaltung der 880nm-LED aufgebaut. Die 880nm-LED wurde auf 2.2mA eingestellt. Der Strom kann mit R7 variiert werden.

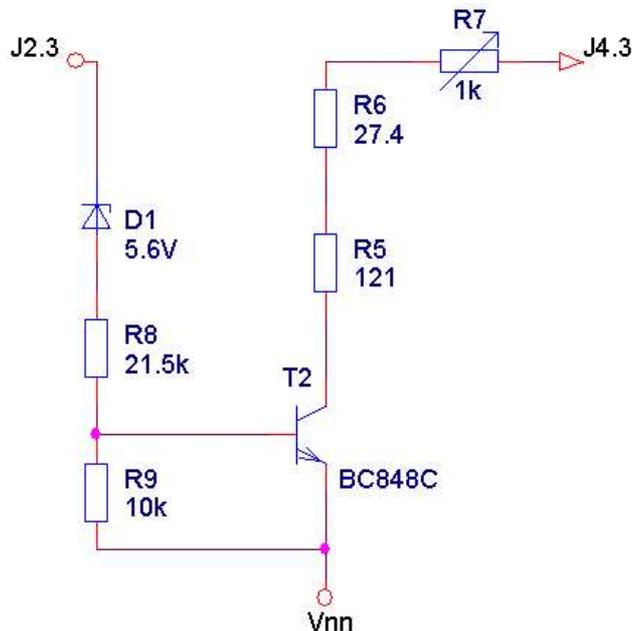


Abbildung 5.10: Die Treiberstufe für die Infrarot-LED mit einer Wellenlänge von 880nm

Der Grund für die Wahl eines niedrigeren Stromes der 880nm-LED ist zunächst auf die wellenlängenabhängige Empfindlichkeit der pin-Diode zurückzuführen. Höhere Wellenlängen verursachen einen höheren Photostrom, was von 660nm auf 880nm etwa den Faktor 10 des Photostroms ausmacht. Um dies zu kompensieren, wurde die Intensität der 880nm-LED mit Absenken des Diodenstroms gedrosselt, so dass nach Reflexion an der Haut, unabhängig

davon welche LED aktiv ist, etwa der gleiche Photostrom an der pin-Diode entsteht. Der Grund dafür ist das hier eingesetzte Multiplex-Verfahren, bei dem ein „Sender“ mit 10-facher Sendeleistung, stärker in andere Kanäle übersprechen würde, so dass die Qualität der entspreizten Signale aller anderen Kanäle beeinträchtigt wird.

5.3.2 Aufbereitung des Sensorsignals

5.3.2.1 Transimpedanzwandler mit TP-Filter

Die vom Körper zurückgestrahlte Restintensität der LEDs fällt auf die beiden pin-Dioden. Die pin-Dioden wurden parallel geschaltet, die Ströme beider LEDs addieren sich dabei zu einem Strom i_s auf. Die auf die aktive Dioden-Fläche einfallende Strahlungsintensität ist proportional zu dem von der pin-Diode erzeugten Strom. Um die Intensität als Signal digitalisieren zu können, muss der Strom proportional in eine Spannung konvertiert werden. Dafür eignet sich die typische Operationsverstärker-Schaltung des Transimpedanzwandlers [TS].

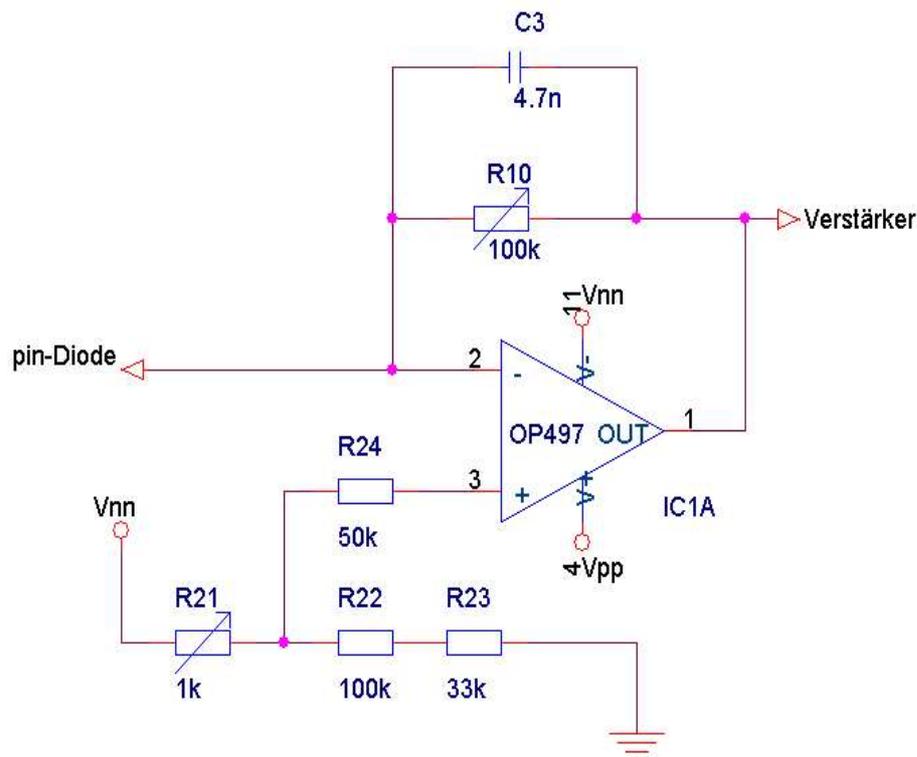


Abbildung 5.11: Transimpedanzwandler mit Kapazität

Das Strom-zu-Ausgangsspannung-Verhältnis ist hier durch den Widerstand R_{10} definiert.

$$U_S(t) = -R_{10} \cdot i_S(t) \quad (5.4)$$

Mit dem Takt der LEDs ist die Spannung U_s einem Rechteck-Signal ähnlich. Mit der hier vorgestellten Schaltung, in Verbindung mit dem Sensor MORES, gab es ein Stabilitätsproblem, das sich je nach Verstärkungsfaktor R_{10} als Überschwingen mit Amplituden von bis zu 70% des Nutzsignals bemerkbar machte. Das Frequenzmaximum der Schwingung lag bei 200kHz. Die Ursache dafür ist eine 180-Grad-Phasendrehung bei etwa 200kHz, während sich die Verstärkung noch oberhalb der 0dB Linie befindet. Um dieses Problem zu beseitigen, wurde ein Kondensator C_3 parallel zu R_{10} angeschlossen, um ab 200Hz eine Verstärkung unter 0dB zu garantieren. Um gleichzeitig dem hochfrequenten Rauschen und dem Aliasing beim Digitalisieren vorzubeugen, wurde die Grenzfrequenz des so entstandenen TP-Filters 1. Ordnung niedriger gewählt, als es zur Stabilisierung des rückgekoppelten Verstärkers erforderlich wäre. Die 3dB-Grenzfrequenz des RC-Tiefpasses berechnet sich nach der Gleichung:

$$f_g = \frac{1}{2\pi R_{10} C_3} \quad (5.5)$$

Mit $R_{10} = 7k\Omega$ und $C_3 = 4.7nF$ liegt die 3dB-Grenzfrequenz bei 4.8kHz. Die Abtastrate des AD-Wandlers $f_0 = 2886$ ist in Kapitel 4.2.2 gewählt worden. Das bedeutet für die Abtastung, dass Frequenzen ab 1.5kHz bis 4.8kHz ungedämpft als Aliasing in das Band 0Hz bis 1443Hz hineinfallen. Bei 4.8kHz greift dann das TP-Filter, startend bei 3dB mit 20dB pro Dekade. Der Grund für die Inkaufnahme von Aliasing sind die Probleme mit nichtlinearen analogen Filtern, die in der DSV zu nicht nachvollziehbaren Effekten führen können. Um diesen Effekten aus dem Weg zu gehen, wurde ein Aliasing in Kauf genommen. Aliasing ist wegen einer reinen Spiegelung eher nachvollziehbar als unterschiedliche Gruppenlaufzeiten.

Ein weiterer Grund für die im Verhältnis zu der halben Abtastrate f_0 hohen Grenzfrequenz f_g , ist die Abtastung, mit der hier gearbeitet wird. Abb. 5.12 zeigt das Sensorsignal, so wie es nach dem Transimpedanzwandler aussehen könnte. Die rot eingezeichneten Zeitpunkte markieren den Chip-Takt, bei dem die LEDs abhängig von den Chip-Folgen ihre Schaltzustände wechseln. Die hellblauen Linien markieren den Zeitpunkt der Abtastung, der nach dem zweiten Drittel der Chip-Periode folgt. Der Kondensator C_3 wurde so gewählt, dass das Sensorsignal vor dem Abtastzeitpunkt einen stabilen Zustand erreicht. Stabil bedeutet, dass sich das Signal von der Abtastung bis zum nächsten Chip nicht um mehr als eine Quantisierungsstufe ändern darf. Bei einem 12-bit AD-Wandler und einem Eingangsspannungsbereich von 0V bis 5V entspricht eine Quantisierungsstufe 1.2mV.

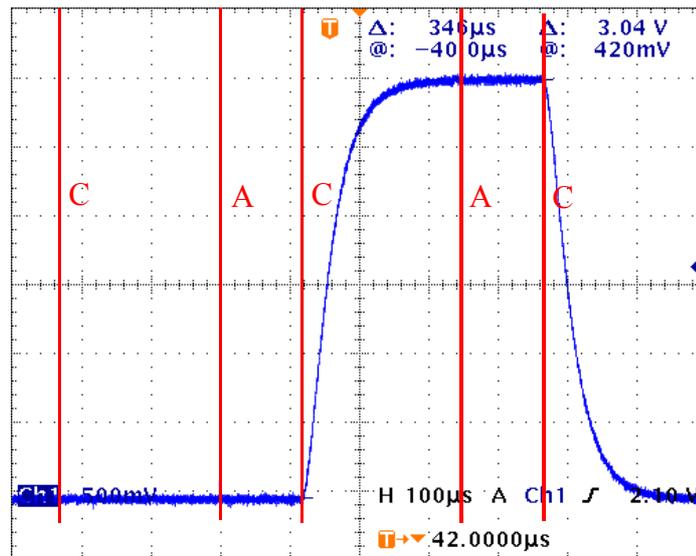


Abbildung 5.12: Das Sensorsignal mit Markierungen zu den Zeitpunkten zum Wechsel der Chips (C) und den Abtastpunkten (A)

Der Operationsverstärker OP497 ist zum Betrieb als Photostrom-Verstärker in der Transimpedanzwandler-Beschaltung geeignet. Die Offsetspannung beträgt maximal $50\mu\text{V}$. Die niedrige Stromaufnahme von $625\mu\text{A}$ spielt in der Pulsoximetrie eine wesentliche Rolle. Der Nachteil des OP497 ist seine maximale Aussteuerung am Ausgang. Bei einer bipolaren Versorgungsspannung von 5V sind das -4V bis 4V .

Mit R_{21} lässt sich die Photodiode mit einer kleinen Spannung vorspannen, und damit ein Arbeitspunkt der gesamten Schaltung einstellen. Der Grund für die AP-Einstellung sind die nichtlinearen Effekte der Operationsverstärker am Rande ihrer Ausgangsaussteuerung. Bei der Vermessung der Umgebungslicht-Intensität arbeiten die Operationsverstärker im nichtlinearen, unteren Grenzbereich. Ohne diese AP-Einstellung entstehen Fehler bei der Umgebungslicht-Subtraktion. Mit der AP-Einstellung wurde dieses Problem behoben.

5.3.2.2 Spannungsverstärker

Das Signal des Transimpedanzwandlers wird mit einem OP496 auf max. 5V verstärkt. Da es sich um einen unipolaren Rail-to-Rail-Verstärker handelt, sind mit der aktuellen Beschaltung Ausgangsspannungen von 36mV bis 4.92V möglich. Der Verstärkungsfaktor kann mit R_{12} im Intervall von 26 bis 37 eingestellt werden.

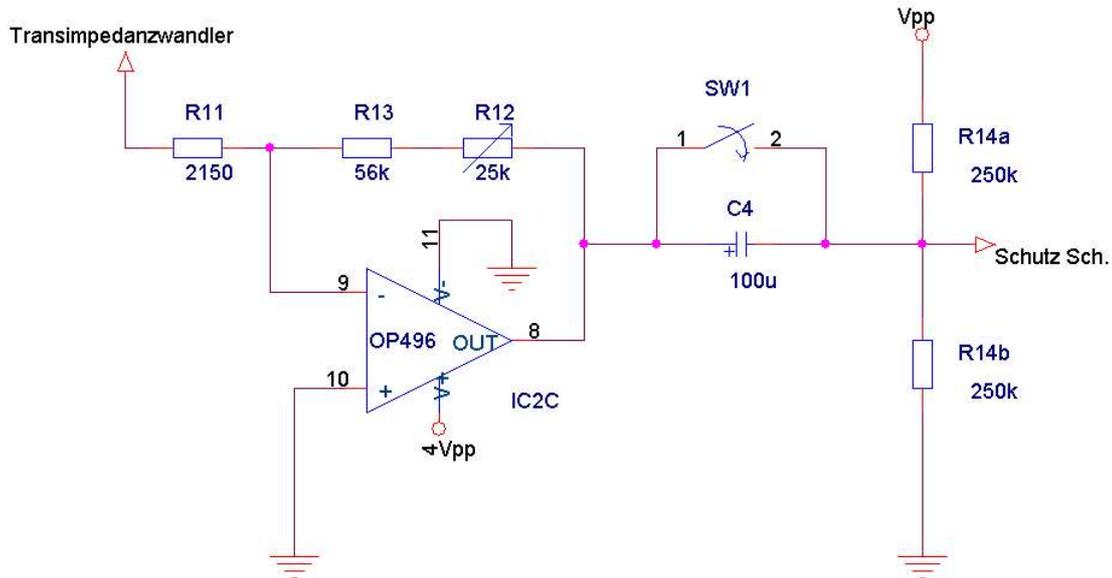


Abbildung 5.13: Verstärkerschaltung mit zuschaltbarer AC-Entkopplung

Nach dem Operationsverstärker folgt eine AC-Entkopplung mit dem Kondensator C_4 und dem Präzisionspotentiometer R_{14} , woraus eine Zeitkonstante von $2 \cdot R_{14} \cdot C_4 = 0.05 \text{ s}$ folgt. Diese Funktion war in einer früheren Entwicklungsphase zum Testen des Algorithmus auf automatischen Ausgleich des Offsets eingeführt worden, der durch starke Umgebungslichteinstrahlung (z.B. Sonnenlicht) entstehen könnte. Der Ausgleich musste langsam erfolgen – deshalb die kleine Zeitkonstante – so dass sich die Nachregelung nicht im AC-Anteil des Algorithmus bemerkbar macht. Diese Funktion wird zum gegenwärtigen Entwicklungsstand nicht benötigt, somit muss C_4 mit dem Schalter SW1 kurzgeschlossen werden.

5.3.2.3 Schutzschaltung für den AD-Wandler

Die Schutzschaltung wurde als Überspannungsschutz des AD-Wandlers integriert. Der AD-Wandler darf nach Spezifikation nicht über die als Referenz angegebene Spannung übersteuert werden. Die Referenzspannungen liegen bei 0V und bei 5V. Die Schaltung begrenzt ein Eingangssignal im Spannungsbereich von 280mV bis 4.82V.

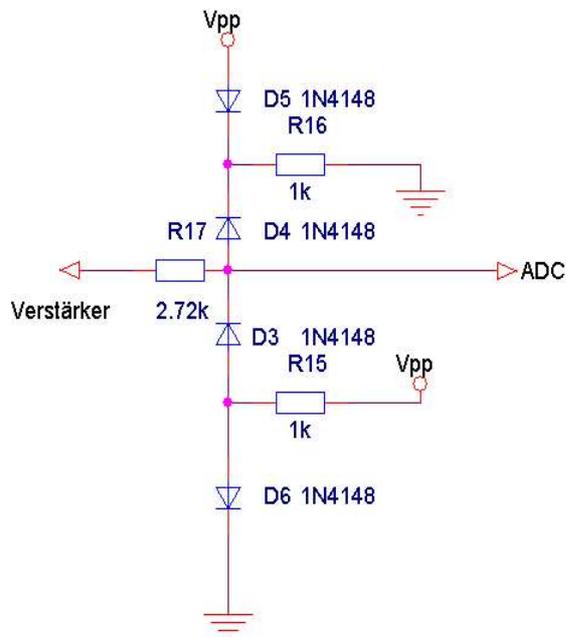


Abbildung 5.14: Schutzschaltung

5.3.2.4 Vermessung der Linearität

Die analoge Verarbeitungskette wurde auf Linearität vermessen. Dazu wurde die Intensität, bestehend aus einem Wechselanteil und einem DC-Anteil, variiert. Die Amplitude des Wechselanteils blieb während der Messung konstant, und der DC-Anteil wurde stufenweise erhöht. Das Ergebnis: Die vermessene Amplitude des Wechselanteils war für alle DC-Werte gleich, was ein Zeichen für die Linearität eines Systems ist.

5.3.3 Status-LEDs

Die Dual-LED dient zum Anzeigen von Zuständen des Pulsoximeters und besteht aus den Dioden D7a (rot) und D7b (grün).

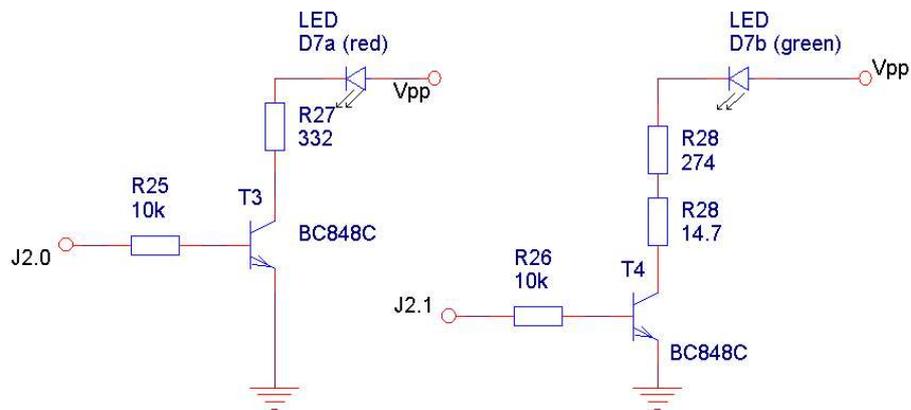


Abbildung 5.15: Treiber der Status-LEDs

5.3.4 Bestückung und Pin-Belegung

Der Bestückungsplan in der Abb. 5.16 zeigt die Position aller hier zur analogen Verarbeitung verwendeten Bauelemente. Die Bezeichnungen der Bauelemente sind in diesem Dokument einheitlich. Dem IC3 und seiner Beschaltung braucht keine Beachtung geschenkt werden, da er nicht mehr benötigt wird.

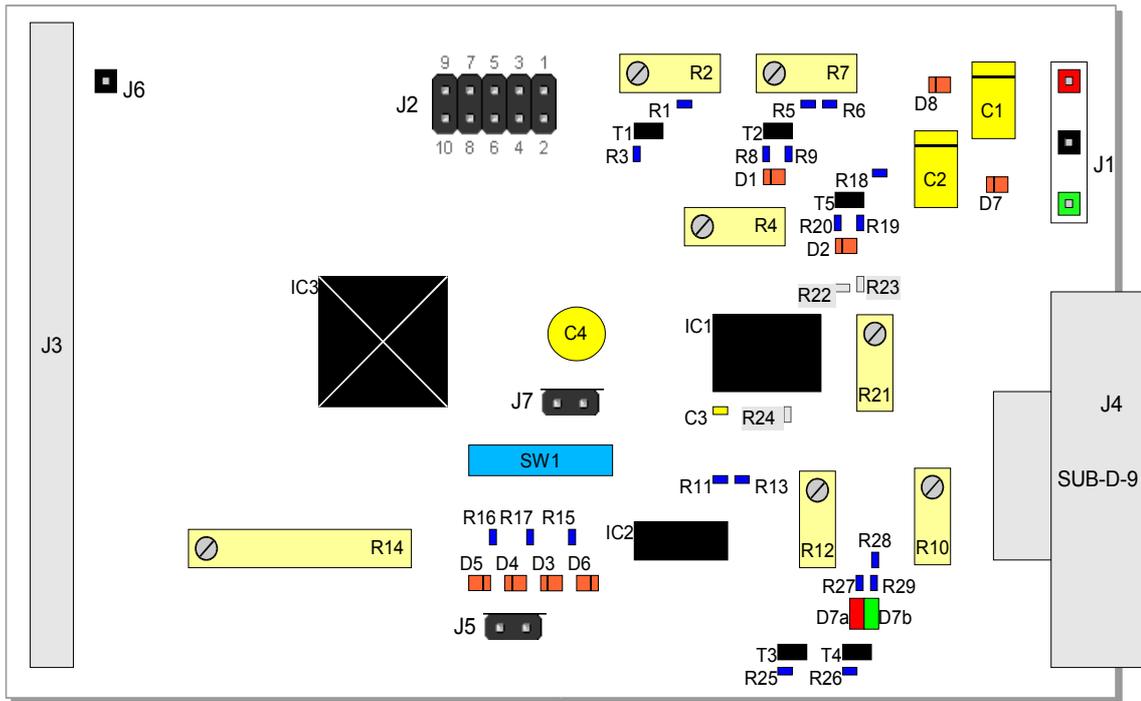


Abbildung 5.16: Bestückungsplan der Platine

Am Anschluss J1 wird die Stromversorgung angelegt. Dabei ist darauf zu achten, dass das Modul ordentlich auf dem TriBoard mit dem Verbinder J3 verbunden ist, bevor die Spannung angelegt wird. Das Anlegen der Spannung des TriBoards sollte vor dem Anlegen der Spannung des analogen Moduls erfolgen.

Pin	Verwendung
ROT	5V
SCHWARZ	GND
GRÜN	-5V

Tabelle 5.5: Belegung des J1-Verbinders auf der Analogen

Der J2-Verbinder wird über ein 9-poliges Flachbandkabel an die Pins 5,7,9,11,13,15,17,19 und 21 des 80-poligen-Verbinders X804 angeschlossen. Diese sind als TTL-Ausgänge konfiguriert.

Pin	Port des μC	Verwendung
1	NC	-
2	P1.0	rote Status-LED
3	P1.1	grüne Status-LED
4	P1.2	MORES-LED 660nm
5	P1.3	MORES-LED 880nm
6	P1.4	MORES-LED 760nm
7	P1.5	NC
8	P1.6	NC
9	P1.7	NC
10	P1.8	NC

Tabelle 5.6: Belegung des J2-Verbinders (NC = Not connected)

Der Verbinder J3 ist 80-polig und führt alle 80 Signalleitungen des Verbinders X802 auf das Modul. Der Verbinder wurde für frühere Zwecke verwendet, er bietet nunmehr Halt auf dem TriBoard und eine Verbindung mit GND.

J4 ist die Schnittstelle zum MORES, ein 9-poliger SUB-D-Verbinder, so wie er in der Pulsoximetrie üblich ist.

Pin	Verwendung
1	660nm-LED
2	760nm-LED
3	880nm-LED
4	GND
5	Pin-Diode
6	-5V
7	NC
8	NC
9	NC

Tabelle 5.7: Belegung des J4-Verbinders

J5 ist der Signalausgang zum AD-Wandler. Das TriBoard wird über ein Koaxialkabel an die Pins 1 und 5 des Verbinders X803 mit J5 verbunden. Zudem müssen die Pins 43, 44 und 47 kurzgeschlossen sein. Damit wird die untere Referenzspannung des AD-Wandler auf GND gesetzt. Die Pins 45 und 46 müssen mit dem Verbinder J6 verbunden sein. J6 ist die Referenzspannung des Analog-Moduls für den AD-Wandler.

5.3 Analoge Verarbeitung

Pin	Verwendung
1	Analog-GND
2	Signalausgang zum ADC

Tabelle 5.8: Belegung des J5-Verbinders

6 Software – Entwicklung

6.1 TriCore-Programm

6.1.1 Entwicklungswerkzeuge

Das Programm für den Mikrocontroller TriCore 1765 wurde unter Windows mit den kostenfreien *High-Tec GNU-Entwicklungswerkzeugen*¹¹ entwickelt. Der Code wurde mit dem High-Tec GNU C Cross Compiler übersetzt und vom Linker zu einer elf-Datei gebunden. Das TriBoard verfügt über ein JTAG-Interface, das mit dem Parallel-Port eines PCs verbunden werden kann. Ein JTAG-Treiber wurde auf dem PC installiert. Ein mitgelieferter JTAG-Server greift auf den Treiber zu und öffnet einen TCP/IP-Port. Auf diesem Port ist es schließlich möglich, den im Produktumfang enthaltenen Debugger zu verbinden, und somit eine Verbindung zum TriCore zu erstellen. Über den Debugger lässt sich Maschinencode in das SRAM auf dem TriBoard hochladen.

Die mitgelieferte Entwicklungsumgebung zum Editieren des Quellcodes wurde durch das *Open-Source Entwicklungswerkzeug Eclipse*¹² ersetzt. Dazu wurde Eclipse mit dem Plugin zum Editieren von C-Code erweitert.

Eine Hilfe bot das von Infineon mitgelieferte Programm DAvE, mit dem sich hardware-spezifische Einstellungen über Windows-Dialoge vornehmen lassen, und schließlich daraus C-Code generieren lässt, der allerdings für den GNU-Compiler angepasst werden muss.

6.1.2 Funktionen der TriCore-Software

Die Hauptaufgaben des Mikrocontrollers sind die Ansteuerung der MORES-LEDs und das Digitalisieren des Sensorsignals. Der Algorithmus wurde bereits bis zur Umgebunglichtsubtraktion (Kapitel 4.3.3) implementiert. Um Daten weiter bearbeiten zu

¹¹ Nähere Informationen können der Homepage <http://www.hitec-rt.com> der Firma High-Tec entnommen werden.

¹² Eclipse und die C-Erweiterung sind frei für kommerzielle Zwecke und können auf <http://www.eclipse.org/platform> heruntergeladen werden.

können, ist das TriBoard mit dem PC über die RS232-Schnittstelle verbunden. TriCore kommuniziert über diese Schnittstelle mit dem PC in beiden Richtungen. Auf der PC-Seite können wahlweise die Sensor-Abtastpunkte oder die bereits demodulierten, entstörten vier Kanäle (vgl. Kapitel 4.3.3) empfangen werden.

Der Mikrocontroller kann in vier Betriebszustände versetzt werden. Über die serielle Schnittstelle kann der PC Anweisungen an den Mikrocontroller versenden und dessen Betriebszustand wechseln. Im Betriebszustand *rmHalt* wartet der Mikrocontroller auf Anweisungen vom PC, dies ist der Zustand zu Programmbeginn. Im Betriebszustand *rmWarmUp* werden die LEDs mit dem Chip-Folgen getaktet, ohne das Sensorsignal aufzuzeichnen. Dieser Modus wird zum Einschwingen der analogen Elemente verwendet. Im Betriebszustand *rmRawData* werden die LEDs mit dem Chip-Folgen getaktet, das Sensorsignal digitalisiert und zum PC gesandt. Im Betriebszustand *rmDemod* werden anstatt des Sensor-Signals die bereits demodulierten und entstörte Signale zum PC gesandt.

Nummer	Betriebsart	Beschreibung
0	rmHalt	Warten auf Anweisungen
1	rmWarmUp	Einschwingen der analogen Elemente
2	rmRawData	Sensorsignal an PC senden
3	rmDemod	Demodulierte Signale an PC senden

Tabelle 6.1: Betriebszustände des TriCore-Mikrocontrollers

Die Status-LEDs (D7a und D7b aus in Abb. 5.16) geben Auskunft über den Betriebszustand des Mikrocontrollers. Der Mikrocontroller befindet sich bei Rot in *rmHalt*, bei Rot/Grün in *rmWarmUp* und bei Grün in *rmRawData* oder *rmDemod*. Leuchtet keiner der Status-LEDs, so ist das TriCore-Programm nicht gestartet.

6.1.3 C-Programmmodule

Quellcode-Modul	Beschreibung
SpO2.c	Das Hauptprogramm besteht aus der Konfiguration des TriCore 1765, der Hauptschleife und der Initialisierung aller Softwarekomponenten
timer.c timer.h	Der <i>Timer</i> übernimmt die Initialisierung, das Schalten der MORES-LEDs und das Timing der Abtastung
comm.c comm.h	Die Kommunikationseinheit verwaltet die ankommenden und abgehenden Datenströme von und zum PC und interpretiert die Befehle des PCs und antwortet auf seine Anfragen
receiver.c receiver.h	Der <i>Receiver</i> übernimmt die Initialisierung des AD-Wandlers, die Abtastung und die Zwischenspeicherung der Abtastpunkte
demodulator.c demodulator.h	Der <i>Demodulator</i> dient der Entspreizung des Signals, dem Subsampling und der Umgebungslicht-Subtraktion; Es ist die Implementierung bis zum Kapitel 4.3.3
truart.c truart.h	Initialisierung und Zugriff auf die asynchrone Schnittstelle zum PC
led.c led.h	I/O-Initialisierung für die MORES- und Status-LEDs
tc_error.c tc_error.h	Fehlerbehandlung

Tabelle 6.2: Programmmodule

Quellcode-Modul	Beschreibung
tc_error_def.c	Definition von Fehlermeldungen und Fehlernummern *
runmode.h	Definition der Betriebszustände *
demod.h	Filterkoeffizienten des Demodulators; wurde durch das M-Skript <i>demodgen.m</i> erzeugt
chips.h	Chip-Folgen wurde durch das M-Skript <i>chipgenp2.m</i> erzeugt

Tabelle 6.3: Module mit Daten und Definitionen

* wird auch von der PC-Anwendung kompiliert

6.1.4 Programmablauf

Das Programm beginnt im Modul SpO2.c, das als Flussdiagramm in Abb. 6.1 dargestellt ist. Hier werden digitale Ausgänge definiert, der globale Interrupt freigegeben, die Betriebsart *rmHalt* gewählt, der *Receiver* und damit der AD-Wandler konfiguriert, die serielle Schnittstelle zu PC eingerichtet, und schließlich ein Timer konfiguriert und gestartet. Anschließend folgt eine Endlosschleife, in der abwechselnd der *Demodulator* oder die *Kommunikationseinheit* ihre „Aufgaben“ erledigen können, falls Aufgaben anstehen. Die beiden Programmteile (Demodulation und das Senden von Daten zum PC) nehmen in regelmäßigen Abständen relativ lange Verarbeitungszeiten in Anspruch. Sie sind außerdem nicht zeitkritisch. Deshalb wurden sie auf unterster Prioritätsebene in eine Endlosschleife eingebaut. Sie können von jeder Interrupt-Routine unterbrochen werden und laufen asynchron zu allen anderen Vorgängen in diesem Programm.

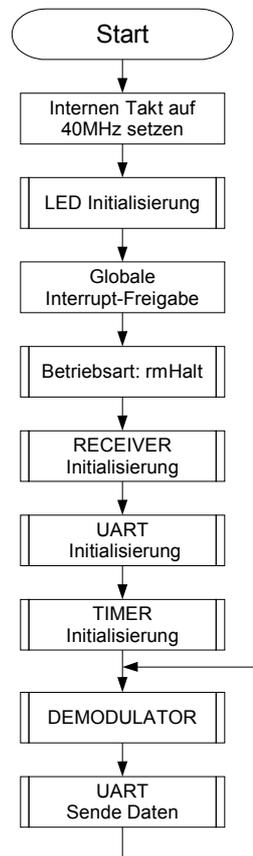


Abbildung 6.1: Flussdiagramm des Moduls SpO2.c

Nach der Initialisierung startet der Timer mit dreifacher Abtastfrequenz f_0 . Der Interrupt des Timers hat die höchste Priorität im System. Bei einem Timer-Interrupt wird die Timer-Interrupt-Routine aufgerufen. Ein Flussdiagramm zu der Timer-Interrupt-Routine ist in Abb. 6.2 dargestellt. Beim Einsprung in diesen Programmteil wird zunächst der Betriebszustand auf

rmHalt untersucht. Ist dieser *rmHalt*, so findet ein Rücksprung statt. Weiterhin wurde ein Einsprung in die Timer-Interrupt-Routine in drei Sequenzen aufgeteilt. Senden, Warten und Empfangen. Die Sequenz wird bei jedem Durchlauf der Routine gewechselt.

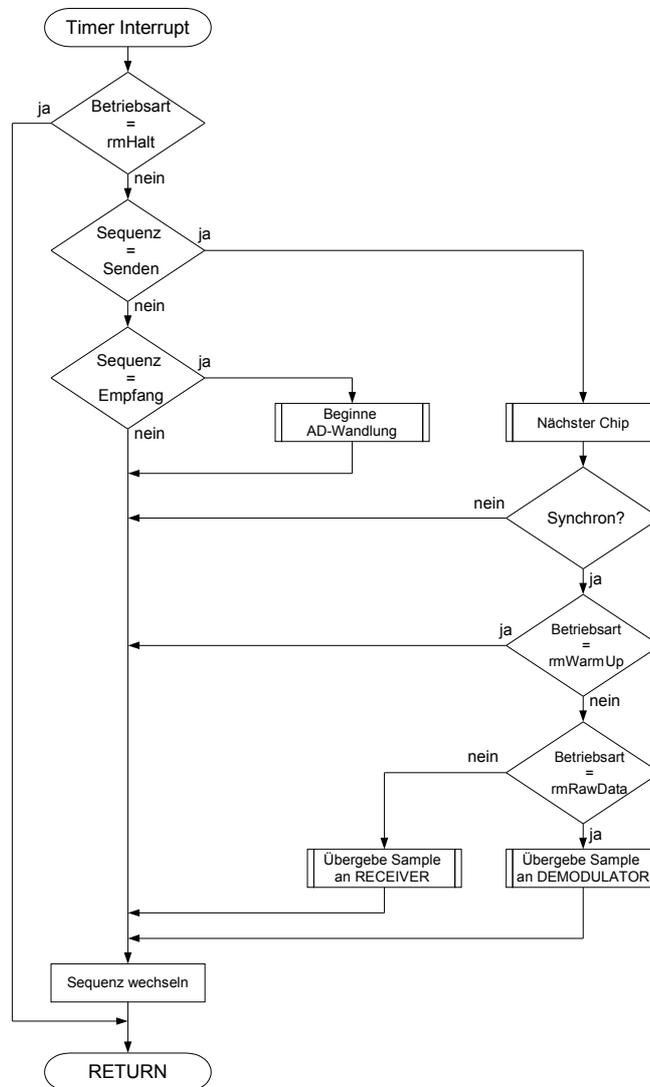


Abbildung 6.2: Flussdiagramm der Timer-Interrupt-Routine

Bei der Übergabe eines Samples erfolgt beim *Demodulator* zunächst keine Demodulation. Der *Demodulator* hat die Aufgabe, nach Eintreffen neuer $I_c=101$ Samples, ein Skalarprodukt nach der Definition in Kapitel 4.3.2 durchzuführen. Der *Demodulator* erzeugt jede 101 Samples insgesamt 4 Samples (Rot, Infrarot, Dunkel1, Dunkel2) des demodulierten Signals. Dazu benötigt der Modulator ein Ring-Buffer nach der Formel 4.26. Anschließend erfolgt die Umgebunglichtsubtraktion (vgl. Kapitel 4.3.3), bei der nochmals 3 Samples (Rot, Infrarot, Referenz-Signal) erzeugt werden. Diese 7 Samples werden in ein Doppel-Puffer gepackt, und sobald dieses voll ist, bemerkt das die *Kommunikationseinheit* und schickt den Inhalt des

Puffers über die serielle Leitung zum PC. Der Demodulator arbeitet nur, wenn ihm Samples zugeführt werden. Diese kommen vom *Receiver*, aber nur im Betriebszustand *rmDemod*, sonst werden dem *Demodulator* keine Samples zugeführt. Die Demodulation läuft asynchron zum Timer und wird vom Hauptprogramm aufgerufen.

Hat der Demodulator seine Aufgabe erledigt, so kann anschließend die *Kommunikationseinheit* überprüfen, ob etwas zum Senden vorliegt. Das könnte ein volles Demodulator-Sende-Puffer, ein volles Receiver-Sende-Puffer, eine anstehende Antwort an den PC oder eine Nachricht über einen Fehler an den PC sein.

Der Empfang von Daten erfolgt auch durch die *Kommunikationseinheit*. Beim Eintreffen eines neuen Zeichens wird ein *RX-Interrupt* ausgelöst. Erfolgt ein *RX-Interrupt*, so wird das gerade eingetroffene Zeichen zusammen mit dem vorhergehenden Zeichen analysiert. Ergeben die beiden Zeichen einen Befehl, so wird dieser ausgeführt, ergeben sie eine Anfrage, so wird diese zum Beantworten an die Sende-Einheit übergeben.

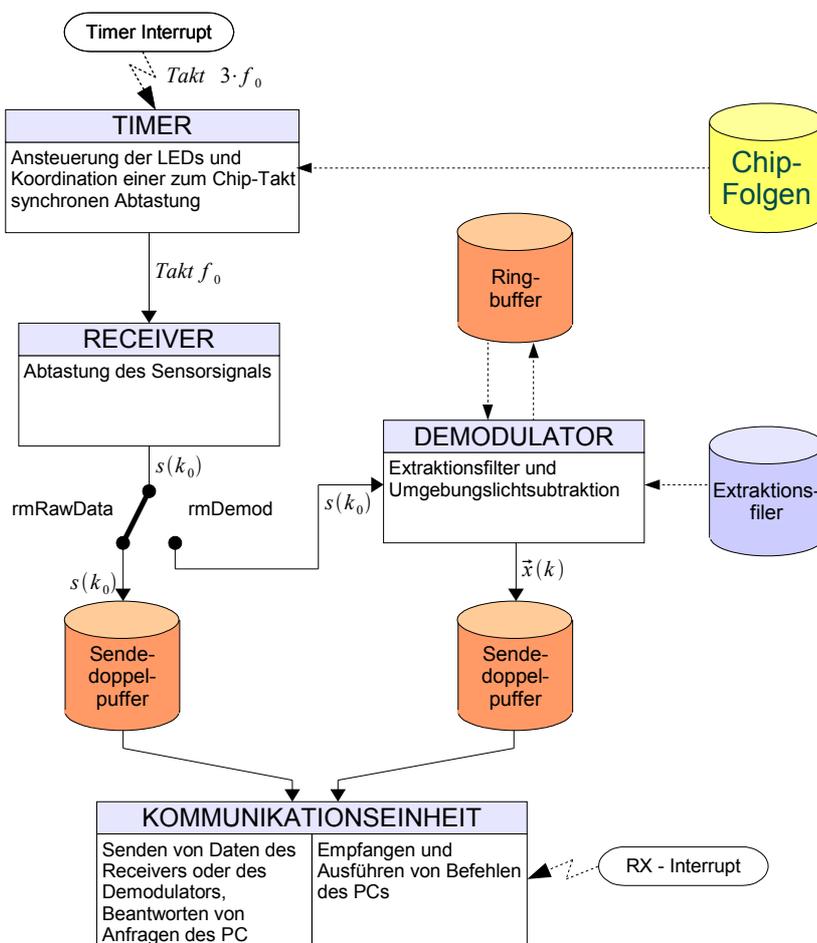


Abbildung 6.3: Schematische Darstellung aller Programmkomponenten

Im Betriebsmodus *rmRawData* und *rmDemod* arbeitet der *Receiver*, dessen Aufgabe die AD-Wandlung und Weiterleitung des diskreten Sensor-Signals $s(k_0)$ ist. Im Betriebszustand *rmRawData* werden die Samples $s(k_0)$ in einem Doppel-Buffer zwischengespeichert. Ist das Puffer voll, bemerkt das die *Kommunikationseinheit* und sendet den Inhalt an den PC. Im Betriebsmodus *rmDemod* werden die Samples $s(k_0)$ an den Demodulator übergeben, anstatt es im Sende-Doppel-Puffer zu speichern.

6.1.5 Laufzeitfehler-Behandlung

TriCore ist mit Abfragen ausgestattet, die zur Fehlerdetektion dienen und wenig Prozessorlast benötigen. Die möglichen Fehler wurden durchnummeriert und zu jeder Fehlernummer eine Fehlermeldung deklariert. Damit der Speicher des Mikrocontrollers nicht mit Fehlermeldungen überflutet wird, wurden nur Fehlernummern ins Programm kompiliert. Wird in einem der Programmteile ein Fehler entdeckt, so wird als erste Reaktion die gelbe LED am TriBoard (siehe Abb. 5.4) eingeschaltet. Diese LED stand zur freien Disposition. Im zweiten Schritt wird die Fehlernummer an den PC gesandt. Quittiert der PC den Empfang der Fehlernummer, so erlischt die gelbe LED auf dem TriBoard. Die gelbe LED ist damit ein Indikator dafür, dass ein Fehler eingetreten ist, und dieses Ereignis nicht dem PC gemeldet werden konnte.

6.1.6 Kommunikationsprotokoll

Der PC und der Mikrocontroller kommunizieren über die serielle Schnittstelle mit einer Geschwindigkeit von 115200 Baud. Das sind 11520 Byte pro Sekunde bei einem Start-, Stop-Bit und 8 Bit an Daten. Das Protokoll wurde aufgrund der kleinen Anzahl an Befehlen und der niedrigen Fehlerrate bei der Übertragung einfach gehalten. Da es sich um eine sequentielle Übertragung von jeweils einem Byte (aus 8 Bit) handelt, muss ein entsprechender Algorithmus auf der Senderseite die Daten serialisieren, und auf der Empfängerseite ein Algorithmus den sequentiellen Datenstrom zu Daten zusammenbauen. Dies ist Aufgabe der Kommunikationseinheit (comm.c).

Bei der Kommunikation zwischen Mikrocontroller und PC wird zwischen Anfragen, Kommandos und Sample-Daten unterschieden. Die Samples werden zu Blöcken zusammengefasst. Jeder Block beginnt mit einem Erkennungszeichen, gefolgt von Daten, und endet mit einer Checksumme. Die Länge eines Blocks ist nicht spezifiziert, sie kann über das Protokoll abgefragt werden. Ein Kommando oder eine Anfrage besteht aus zwei Byte, auf die eine vorzeichenbehaftete 32bit-Integer-Zahl (bestehend aus 4 Byte) oder ein einziges Byte folgen kann.

Datenstrom	Beschreibung
cR <byte>	Betriebszustand senden
cG <int>	Abtastfrequenz der Daten im Block senden
cB <int>	Länge eines Blocks in Byte senden
e <int>	Fehlernummer senden
ds <data>...<data><CRC>	Block mit Samples und einen Byte als Checksumme senden

Tabelle 6.4: Übertragungsprotokoll des Mikrocontrollers and den PC

Datenstrom	Beschreibung
cR <byte>	Betriebsart ändern
cE	Quittierung eines Fehlers
qR	Anfrage über Betriebszustand
qG	Die Abtastfrequenz des Sensorsignals abfragen
qB	Die Blocklänge beim Übertragen des Sensorsignals abfragen
qg	Die Abtastfrequenz der demodulierten Signale abfragen
qb	Die Blocklänge beim Übertragen der demodulierten Signale abfragen

Tabelle 6.5: Übertragungsprotokoll des PCs an den Mikrocontroller

Die Daten innerhalb eines Blocks werden immer mit dem minderwertigsten Byte voraus gesendet und enden mit dem höchstwertigsten Byte. Bei der Übertragung des Sensorsignals werden 16 Bit unsigned Integer gesendet. Bei der Übertragung von den demodulierten Daten werden 32 Bit Integer nacheinander gesendet. In der Reihenfolge:

1. Rot
2. Infrarot
3. Dunkelphase 1
4. Dunkelphase 2
5. Rot nach Umgebungslichtsubtraktion
6. Infrarot nach Umgebungslichtsubtraktion
7. Referenzsignal für den adaptiven Interference-Canceler

Von jedem Byte aus einem Block wird bitweise eine XOR-Verknüpfung mit dem vorhergehendem Byte durchgeführt. So entsteht die Checksumme. Die bitweise XOR-Verknüpfung des ersten Bytes erfolgt mit 0.

6.2 Windows-Anwendung

6.2.1 Funktionalität

Die Hauptfunktionen des Windows-Programmes sind, die Kommunikation über die serielle Schnittstelle mit dem Mikrocontroller und der Export der Daten als TPD-Datei. TPD steht für *Tricore Pulsoxymetry Data* und wurde speziell für dieses Projekt definiert.

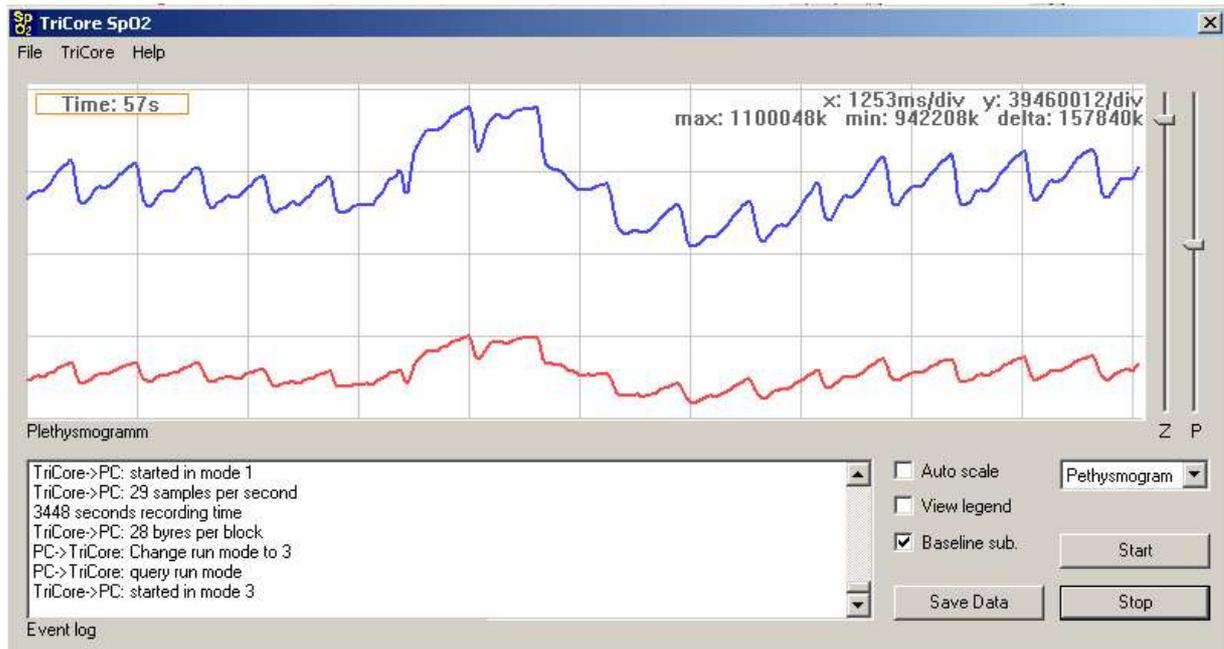


Abbildung 6.4: Screenshot der Windows-Anwendung zum Speichern der Signale des Mikrocontrollers als TPD-Datei

Das Programm verfügt über ein grafisches Steuerelement, in dem die Signale betrachtet werden können. Links oben in diesem Display werden der Füllstand des internen Puffers und die vergangene Aufnahmezeit angezeigt. Rechts oben im Display sind Angaben zum Display, ähnlich wie bei einem Oszilloskop, die Zeit-Auflösung und die Amplituden-Auflösung. Eine Zeile darunter werden die maximale und minimale im Bild sichtbare Amplitude angezeigt, sowie die Differenz daraus. Dies dient als Orientierungshilfe, da das Display vertikal skalierbar ist. Mit dem Z-Schieber kann die Größe des vertikalen Ausschnitts verändert werden, und mit dem P-Schieber wird die Position des Ausschnitts variiert. Ist die Check-Box¹³ „Auto scale“ markiert, so skaliert sich das Display automatisch auf das im aktuellen Zeitausschnitt sichtbare Signal. Ist die Check-Box „View legend“ markiert, so wird eine Legende zu den sichtbaren Signalen angezeigt.

13 Der Ausdruck Check-Box bezeichnet ein Steuerelement unter Windows, bei dem ein Haken durch Anklicken sichtbar wird und beim erneuten Anklicken wieder verschwindet. Neben dem Haken befindet sich meist ein Text, der die Funktion benennt, die so aktiviert bzw. deaktiviert werden kann.

Im Auswahlménü kann zwischen „Plethysmogramm,“ oder „Sensor signal“ gewáhlt werden. Mit „Plethysmogramm“ wird der Mikrocontroller in den Betriebszustand *rmDemod* versetzt. Mit „Sensor signal“ wird der Mikrocontroller in den Betriebszustand *rmRawData* versetzt. Die Betriebszustände des Mikrocontrollers wurden in Kapitel 6.1.2 erläutert.

Mit den Schaltflächen „Start“ und „Stop“ kann ein Messvorgang gestartet bzw. gestoppt werden. Der Mikrocontroller wird beim Start für eine Sekunde in den Betriebszustand *rmWarmUp* versetzt, anschließend in den im Auswahlménü definierten Zustand versetzt. Von jetzt an sendet der Mikrocontroller Daten, die im Arbeitsspeicher des PCs abgelegt werden, solange der reservierte Speicherplatz reicht oder bis der Benutzer die Messung beendet.

Alle Ereignisse werden in dem Listenfeld „Event log“ und der Datei event.log protokolliert. Die Protokoll-Datei befindet sich im Anwendungspfad. Das Ereignis-Protokoll enthält auch die ausformulierten Fehler des Mikrocontrollers. Die Fehlermeldungen enthalten bereits Hinweise zur Fehlerursache oder -beseitigung. Fehler seitens der PC-Anwendung tauchen auch im Protokoll auf, wie z.B. CRC-Fehler nach der Übertragung über die serielle Schnittstelle.

Beim Betätigen der Schaltfläche „Save data“ öffnet sich ein Windows Datei-Dialog zum Exportieren der aufgezeichneten Daten als TPD-Datei.

Die Check-Box „Baseline sub.“ ist nur im Betriebszustand *rmDemod* (entspricht im Auswahlménü der Option „Plethysmogram“). Ist die Check-Box nicht markiert, so sind im Display die Signale

- Rot
- Infrarot
- Dunkelphase 1
- Dunkelphase 2

sichtbar. Ist die Check-Box markiert, so sind die Signale

- Rot nach Umgebungslicht-Subtraktion
- Infrarot nach Umgebungslicht-Subtraktion
- Referenzsignal für den adaptiven Interference-Canceler

Das Programm wurde unter Microsoft Windows 2000 Professional 5.0 Service pack 4 auf einem PC mit 1.6 GHz Intel Pentium 4 Prozessor und 512 MB Arbeitsspeicher entwickelt.

6.2.2 Übersicht der Klassen

Klasse	Funktion
CSequencer	Ablaufsteuerung, bestehend aus zwei Threads; ein Thread für Grafik (15 Aufrufe pro Sekunde); ein zweiter Thread zur Rekonstruktion des sequentiellen Datenstromes zu Samples und zur Kommunikation mit dem Mikrocontroller (500 Aufrufe pro Sekunde)
CGraph	Visualisierung der Signale eines CDataSets
CDataSet	Universeller Signal-Container für mehrere Kanäle
CSerial	Zugriff auf die Serielle Schnittstelle
CMemDC	<i>Anti-Clipping memory context</i> von Keith Rule, verhindert das Flimmern des Graphen
CCommandLine	Kommandozeilen-Funktionen von Carlos Antollini, wird lediglich benötigt um den Pfad der Anwendung zu bestimmen

Tabelle 6.6: Klassenübersicht der PC-Anwendung

6.2.3 TPD-Format

TPD-Dateien bestehen aus einem Kopf und aus Abtastpunkten. Der Kopf beginnt mit der Zeichenkette „SpO2 data set“, gefolgt von der Version der TPD-Datei (aktuell 2), einer zusätzlichen Information über den Datensatz, der Anzahl der Kanäle, der Abtastrate und der Anzahl der Abtastpunkte je Kanal. Nach dem Kopf folgen zeilenweise die Abtastpunkte. Bei mehreren Kanälen werden die Signale durch Kommas in einer Zeile getrennt.

```
SpO2 data set
@VERSION:    2
@INFO:       RAW DATA
@NUM_CHAN:   1
@SAM_RATE:   2886
@NUM_OF_SAM: 232401
388
2427
2426
2421
2419
...
```

7 Untersuchungen und Ergebnisse

7.1 Black-Box zur Interferenz-Simulation

Um die Beständigkeit des Algorithmus in verschiedenen Umgebungslicht-Situationen zu testen, wurde eine Black-Box für den Sensor entwickelt. Das Gehäuse ist gegen Eindringen von Umgebungslicht optisch abgeschlossen, um undefinierbare Störquellen auszuschließen. Die Black-Box enthält zwei LEDs, mit denen Störquellen simuliert werden sollen. Eine Infrarot-LED strahlt von oben und eine weiße LED strahlt von der Seite auf den Finger. Die Strahlung der oberen LED dringt durch den Finger bis zum Sensor vor und ist somit zum Teil mit dem Puls-Signal moduliert. Die Strahlung der LED an der Seite dringt, je nach Lage des Fingers, zum Teil durch den Finger und zum Teil direkt zum Sensor vor. Strahlung, die direkt auf den Sensor trifft, ist nicht mit arteriellen Volumenänderungen moduliert. Die Black-Box ist schwarz, um die Reflexionen im Gehäuse gering zu halten.



Abbildung 7.1: Black-Box (geöffnet) zur Simulation von Umgebungslichtinterferenzen

Mit der Black-Box werden gezielt Störquellen simuliert. Dazu werden die LEDs mit einem Funktionsgenerator betrieben. Mit der Black-Box können zum Vergleich auch Test-Episoden ohne Störeinflüsse erstellt werden.

7.2 Einfluss von Störungen auf die Messgrößen

7.2.1 Test der Umgebungslichtsubtraktion

Bei der Umgebungslichtsubtraktion nach dem Prinzip aus Kapitel 4.3.3, wird die Intensität des Umgebungslichtes separat gemessen. Ist die Umgebungslichtsubtraktion zugeschaltet, so wird der Anteil an Umgebungslicht von den Signalen (Rot und Infrarot) subtrahiert. Die Umgebungslichtsubtraktion funktioniert nur für Signale im Nutzband (von 0 bis 14,29Hz). Die Umgebungslichtsubtraktion soll mit der Episode 9 getestet werden. Bei der Episode 9 wird davon ausgegangen, dass es sich um einen gesunden Probanden handelt, von dem in der Regel eine konstante Sauerstoffsättigung bei 100% zu erwarten ist.

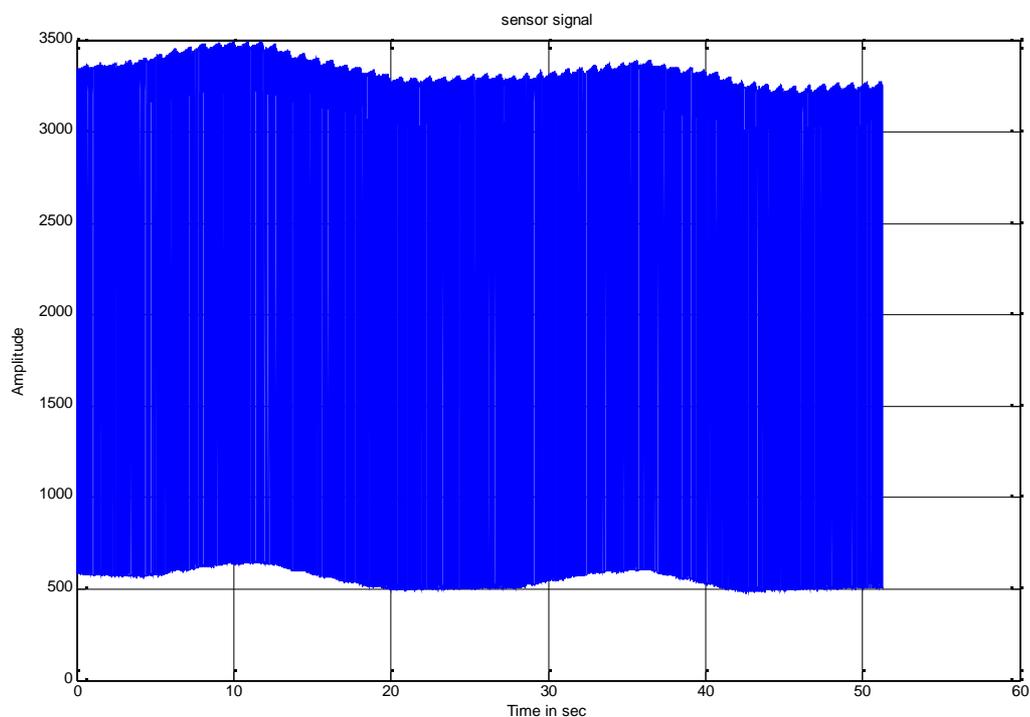


Abbildung 7.2: Das Sensorsignal (Figure 1) aus der Test-Episode 9 mit einer langsam alternierenden Umgebungslichtintensität.

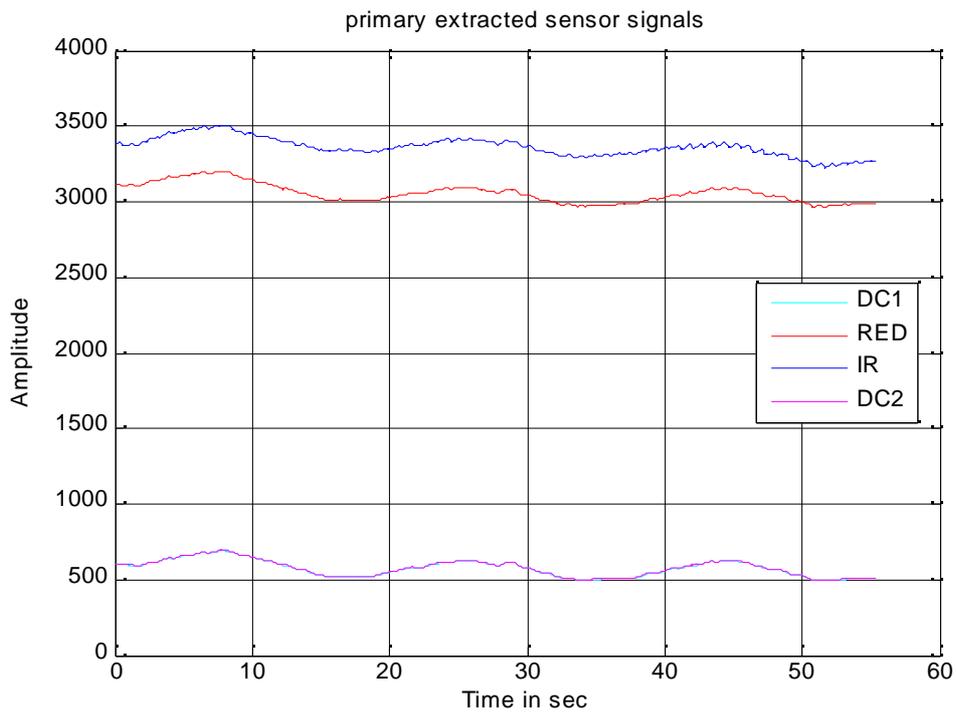


Abbildung 7.3: Die Signale nach der Demodulation (Figure 2). Die Messung in den Dunkelphasen (DC1 und DC2) zeigt den reinen Umgebungslichtanteil, der auch den Nutzsignalen (Rot und Infrarot) überlagert ist.

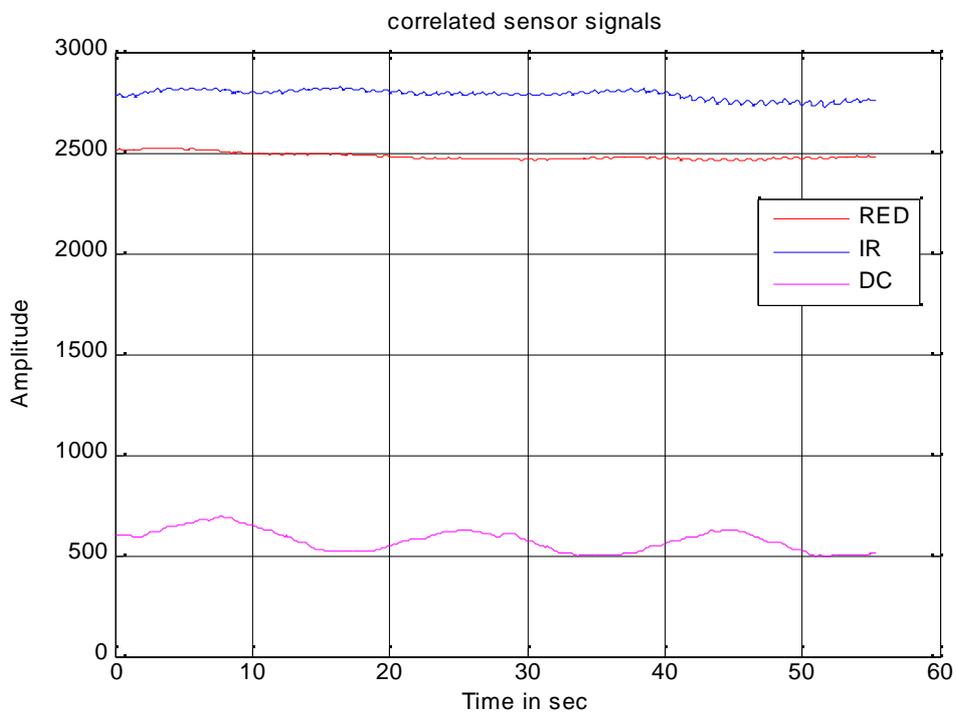


Abbildung 7.4: Die Linearkombination der Dunkelphasen DC1 und DC2 nach Kapitel 4.3.3 wurde als Umgebungslichtanteil DC deklariert, und ist hier bereits von den Nutzsignalen subtrahiert worden. Das Resultat der Umgebungslichtsubtraktion sind die Signale (Rot und Infrarot) ohne Umgebungslichtanteil.

7.2 Einfluss von Störungen auf die Messgrößen

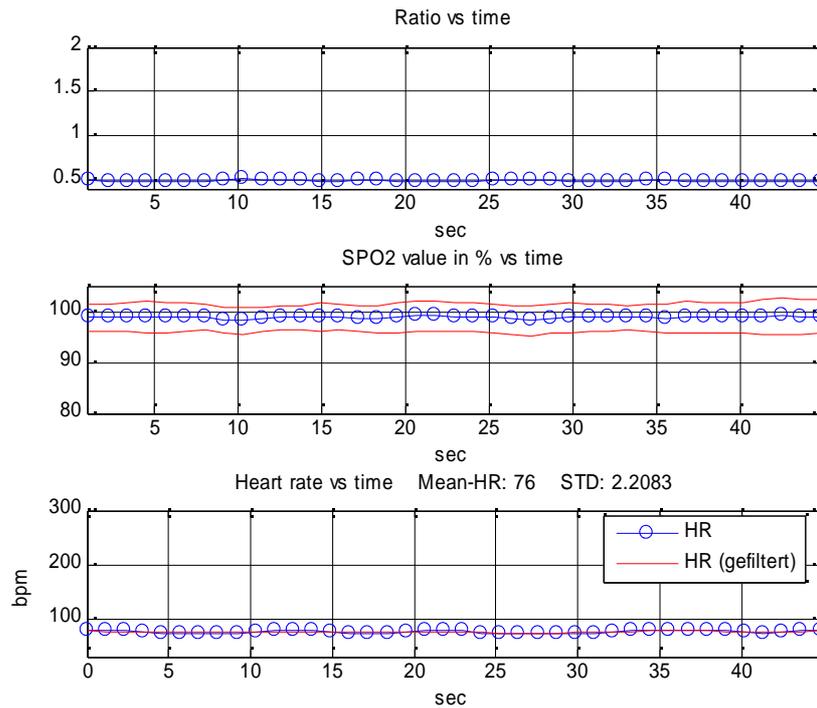


Abbildung 7.5: Die Abbildung zeigt den Verlauf des Verhältnisses ρ , des SpO_2 -Wertes und der Herzfrequenz über 40 Sekunden der Test-Episode 9 mit zugeschalteter Umgebungslichtsubtraktion.

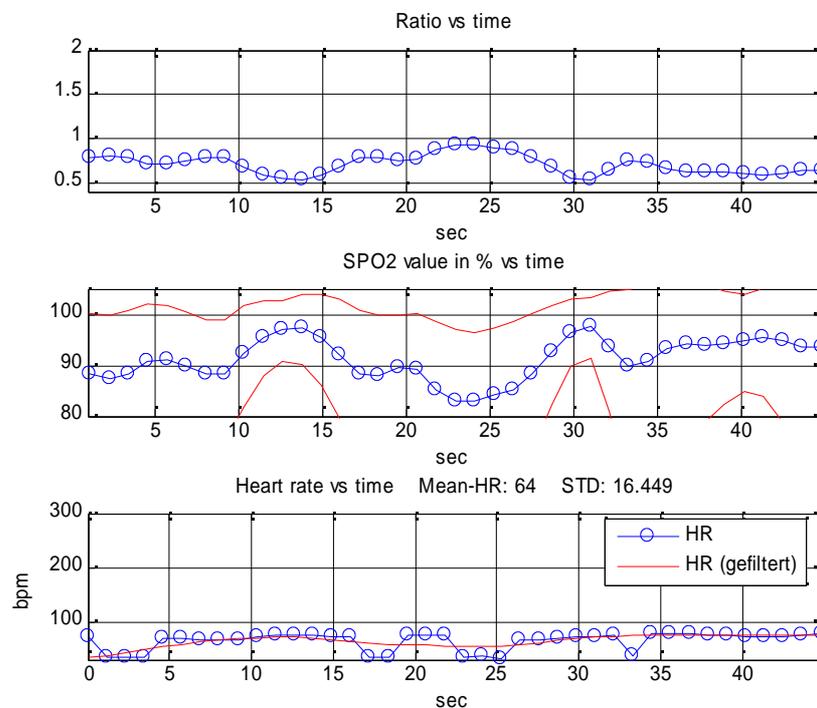


Abbildung 7.6: Zum Vergleich die Episode 9 ohne Umgebungslichtsubtraktion nach dem Prinzip aus Kapitel 4.3.3. Der SpO_2 -Wert ist erkennbar abhängig von der Umgebungslichtintensität.

Der SpO₂-Wert sinkt mit zunehmender Intensität, wenn die Umgebungslichtsubtraktion nicht aktiviert ist. Um genau zu sein, der Wert geht gegen etwa 80%. Das liegt zunächst daran, dass der Störer in diesem Fall unter 0.5 Hz liegt und somit in die DC-Komponente fällt. Der Effekt lässt sich anhand der Gleichung 2.39 aus dem analytischen Modell erklären. Die folgende Grenzwertrechnung zeigt mit x als der Intensität des niederfrequenten Störers, dass das Ratio für niederfrequente Störer gegen 1 geht.

$$\lim_{x \rightarrow +\infty} \frac{\frac{AC_{\lambda 1}}{DC_{\lambda 1} + x}}{\frac{AC_{\lambda 2}}{DC_{\lambda 2} + x}} = 1 \quad (7.1)$$

Mit der Kalibrierungsfunktion aus Kapitel 4.3.14 erhält man etwa 80% für ein Verhältnis von 1. Daraus folgt: Ohne die Umgebungslichtsubtraktion geht die Sauerstoffsättigung gegen 80%. In dem hier vorgestellten System bleiben die Vitalparameter bei Intensitätsschwankungen dieser Art stabil.

Auf die Herzfrequenz haben Lichtquellen niedriger Frequenz (<0.5 Hz) weniger Einfluss, da die Herzfrequenz aus dem AC-Band ermittelt wird.

7.2.2 Störquellen im Übertragungsband

Befindet sich ein Störer im Übertragungsband, d.h. liegen seine Frequenzen oberhalb von der halben Abtastfrequenz f_s , so zeigt die Umgebungslichtsubtraktion nicht den gewünschten Effekt. Dazu wurde der Interference Canceler entwickelt. Diese Frequenzen werden mit 12dB Dämpfung in das Nutzband als Aliasing gespiegelt. Da sich das TP-Filter der Unterabtastung im aktiven Bereich (Sperrbereich) befindet, kommt es zu unterschiedlicher Phasenverschiebung in allen vier Signalen. Derselbe Effekt wurde zum Erzeugen eines Referenzsignals für ein adaptives Filter ausgenutzt. Die Differenz der Signale DC1 und DC2 ergibt bei einem Störer aus dem Übertragungsband ein Signal, das Frequenzen des ins Nutzband gespiegelten Störers enthält. Das adaptive Filter aus Kapitel 4.3.7 passt die Phase und die Amplitude des Referenzsignals an und subtrahiert die im Referenzsignal vorkommenden Frequenzen aus dem Nutzsignal.

Eine Extremsituation soll simuliert werden: In der Test-Episode 18 ist dem Puls-Signal durch eine von oben einstrahlende Infrarot-LED mit einem Sweep (0 bis 30 Hz) hoher Amplitude, überlagert worden. Anhand dieser Episode kann die Funktion der Umgebungslichtsubtraktion zusammen mit der Funktion des adaptiven Filters gezeigt werden. Auch hier wird ein SpO_2 -Wert von etwa 100% erwartet.

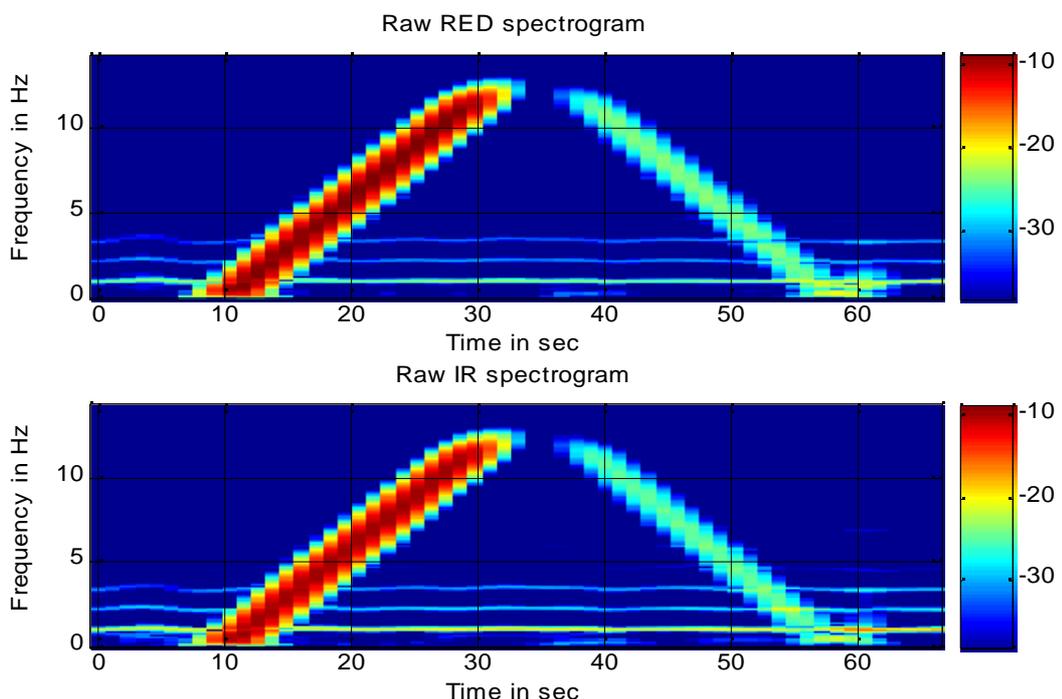


Abbildung 7.7: Die Spektren der demodulierten Signale (Rot und Infrarot) der Episode 18 ohne Umgebungslichtsubtraktion, Interference Canceler und der Spektralen Maske. Bis zum Zeitpunkt $t=31$ s befindet sich der Störer im Nutzband. Schließlich verschwindet das Signal in der Aliasing-Zone, die für die 100 Hz-Unterdrückung geschaffen wurde. Ab dem Zeitpunkt $t=39$ s wird der Störer mit 12dB Dämpfung als Aliasing in das Nutzband gespiegelt.

7.2 Einfluss von Störungen auf die Messgrößen

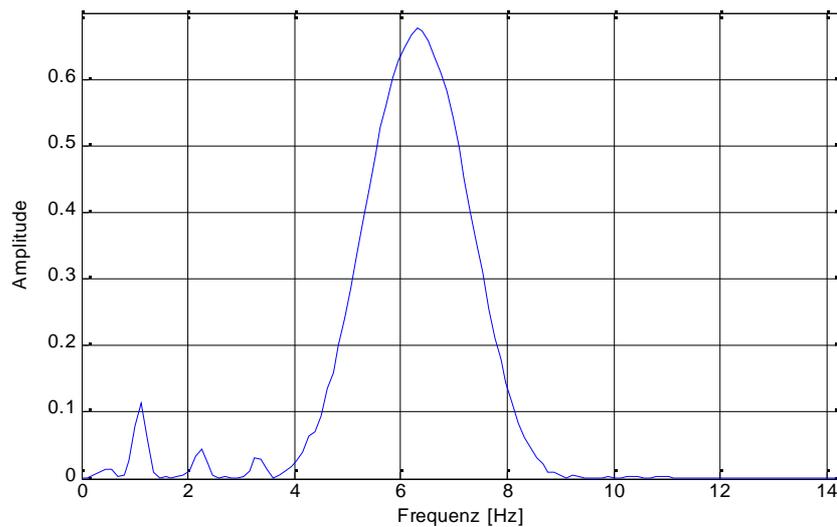


Abbildung 7.8: Ein Spektrum aus dem Spektrogramm aus Abbildung 7.7 zum Zeitpunkt $t=20$ s zeigt den Störer im Verhältnis zum Puls-Signal auf einer linearen Skala. Das Störer-zu-Nutzsignal-Verhältnis beträgt hier 15.5 dB.

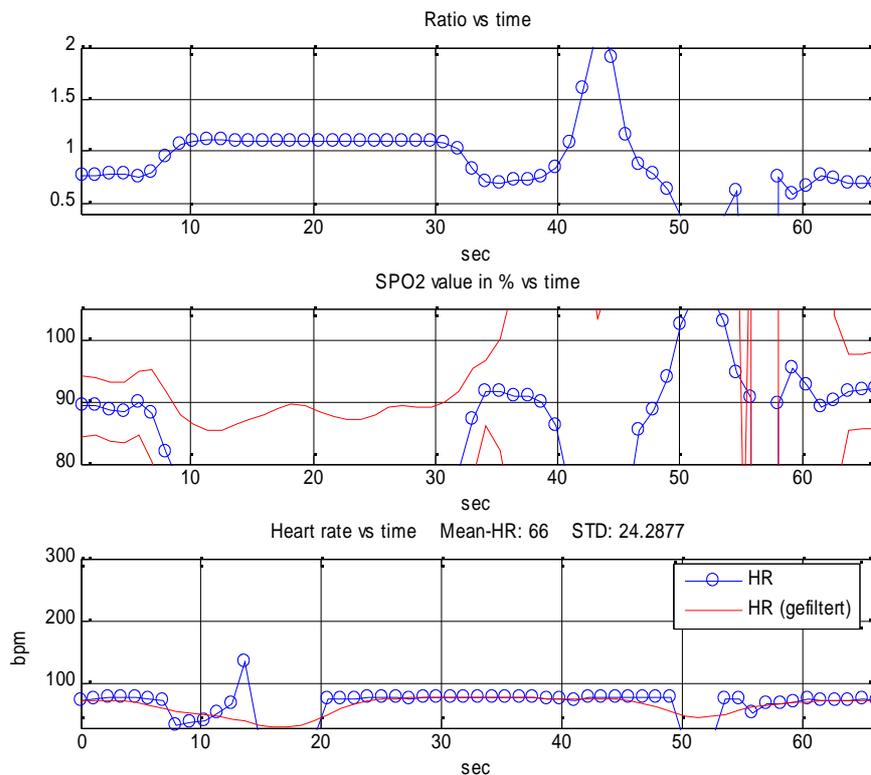


Abbildung 7.9: Die verhältnismäßig hohe Amplitude des Störers führt ohne weitere Vorkehrungsmaßnahmen zu einem unbrauchbaren SpO_2 -Wert über die gesamte Zeitspanne der Messung.

7.2 Einfluss von Störungen auf die Messgrößen

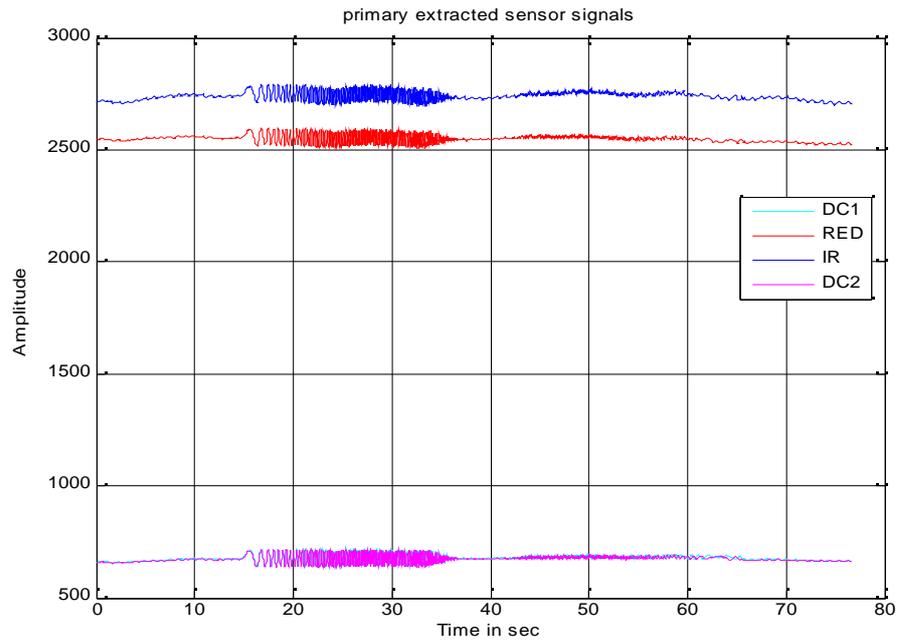


Abbildung 7.10: Im Zeitbereich ist der Umgebungslichtanteil der Störung aus dem Nutzband in den beiden Dunkelphasen DC1 und DC2 erkennbar.

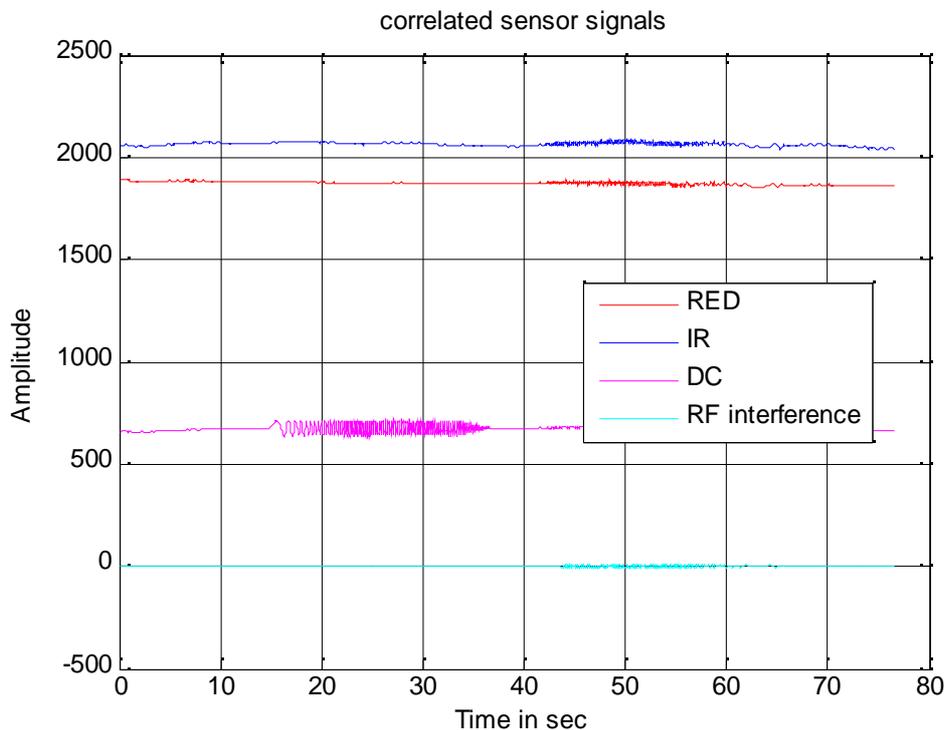


Abbildung 7.11: Mit dem Zuschalten der Umgebungslichtsubtraktion kann der Anteil an Umgebungslicht bis zu der halben Abtastfrequenz t_s subtrahiert werden. Die Abbildung zeigt die beiden Signale (RED und IR) nach der Umgebungslichtsubtraktion, den Umgebungslichtanteil DC und das Referenzsignal für den Interference Canceler.

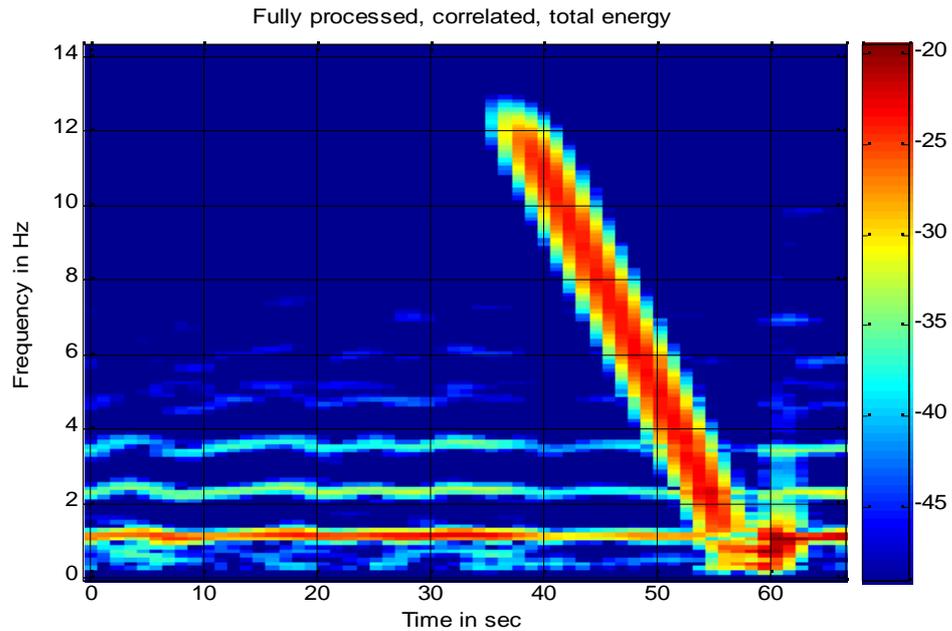


Abbildung 7.12: Der Plot entspricht Figure 6 in der Simulation und zeigt das Referenzspektrum (Kap.4.3.12) der Episode 18. Nach der Umgebungslichtsubtraktion ist der Sweep bis zur 35. Sekunde nicht mehr im Spektrum nachweisbar. Selbst die Frequenzen, die dem Puls-Signal überlagert waren, sind rekonstruiert worden.

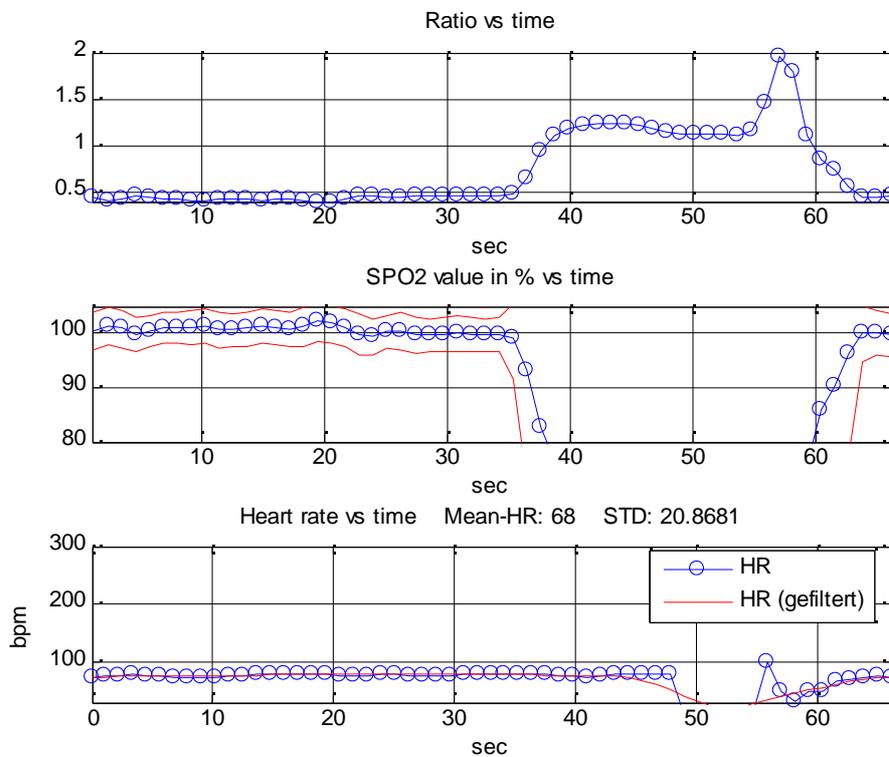


Abbildung 7.13: Das Resultat der Umgebungslichtsubtraktion ist 100% Sauerstoffsättigung bis zur 35. Sekunde.

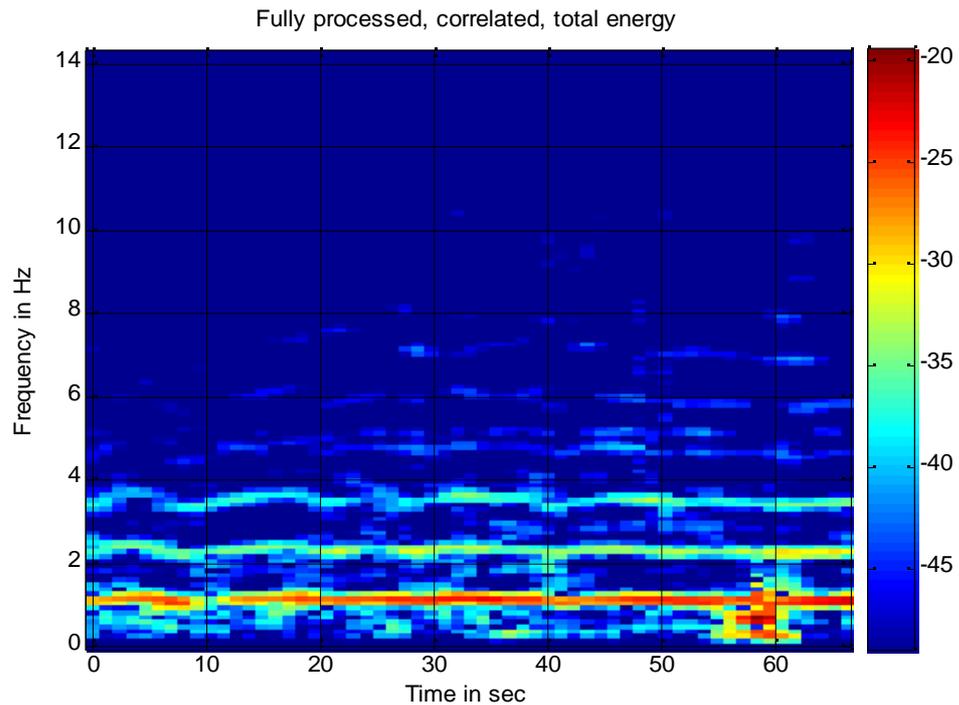


Abbildung 7.14: Nach Zuschalten des Interference Cancelers mit der Filterordnung 10, können anhand des Referenzsignals auch durch Aliasing gespiegelte Frequenzen herausgefiltert werden. In der 57. Sekunde ist ein Artefakt zu sehen, das nicht vom Interference Canceler entfernt werden konnte.

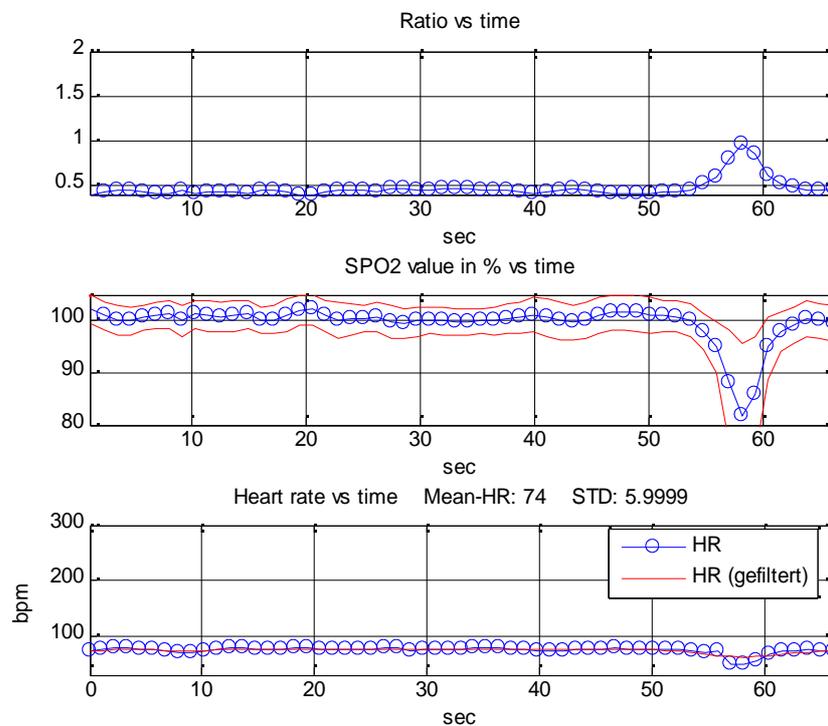


Abbildung 7.15: Das Resultat ist ein SpO_2 -Wert von 100% mit einem kleinen Einbruch über ca. 10 Sekunden. Ein gutes Ergebnis in Anbetracht des Störer-zu-Nutzsignal-Verhältnisses aus Abb. 7.8.

7.2.3 Breitbandige Störungen

Episode 16 ist bis zur 17. Sekunde störungsfrei. Dann wird ein 8-Hz-Sinus-Signal mit einem DC-Offset zugeschaltet. Das Zuschalten ergibt einen Sprung, der sich breitbandig im Spektrum zeigt. Zur 40. Sekunde wird das Signal samt Offset abgeschaltet. Ein erneuter Sprung macht sich im Spektrum bemerkbar. Auch hier wird eine Sauerstoffsättigung von 100% erwartet.

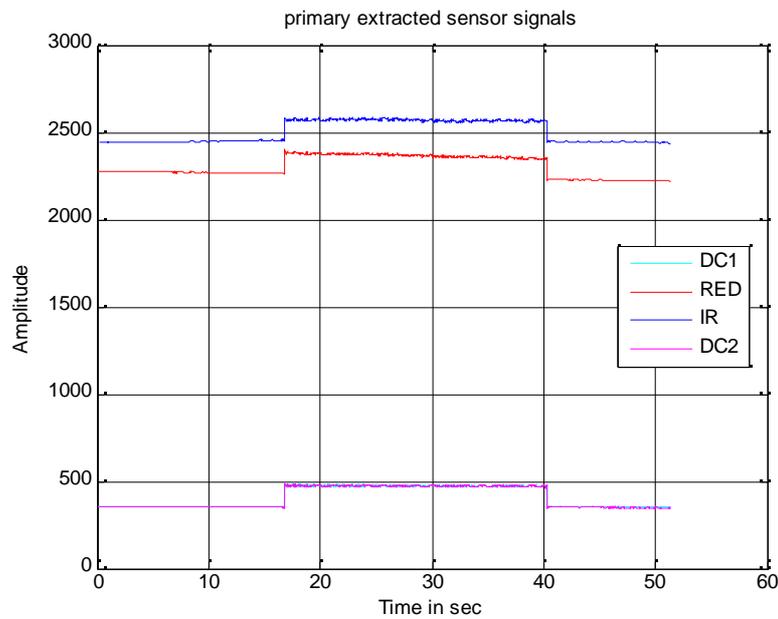


Abbildung 7.16: Der Sprung erkennbar an den Nutzsignalen (RED u. IR) und den Dunkelphasen DC1 und DC2

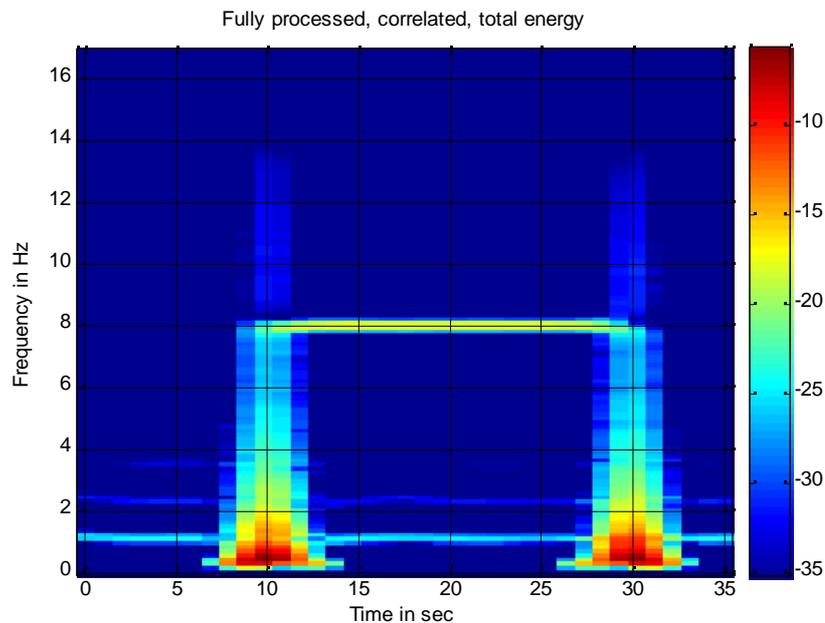


Abbildung 7.17. Das Spektrum der Episode 16

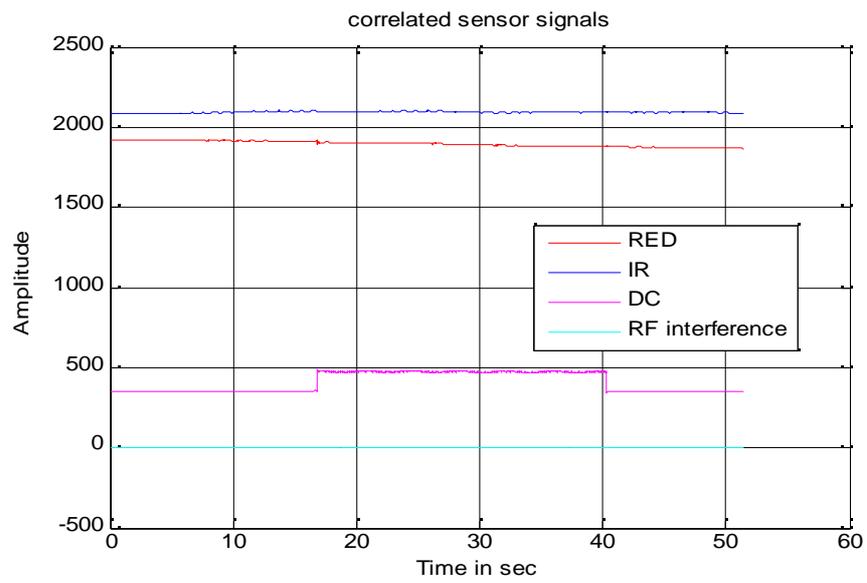


Abbildung 7.18: Die Signale RED u. IR nach Anwendung der Umgebungslichtsubtraktion

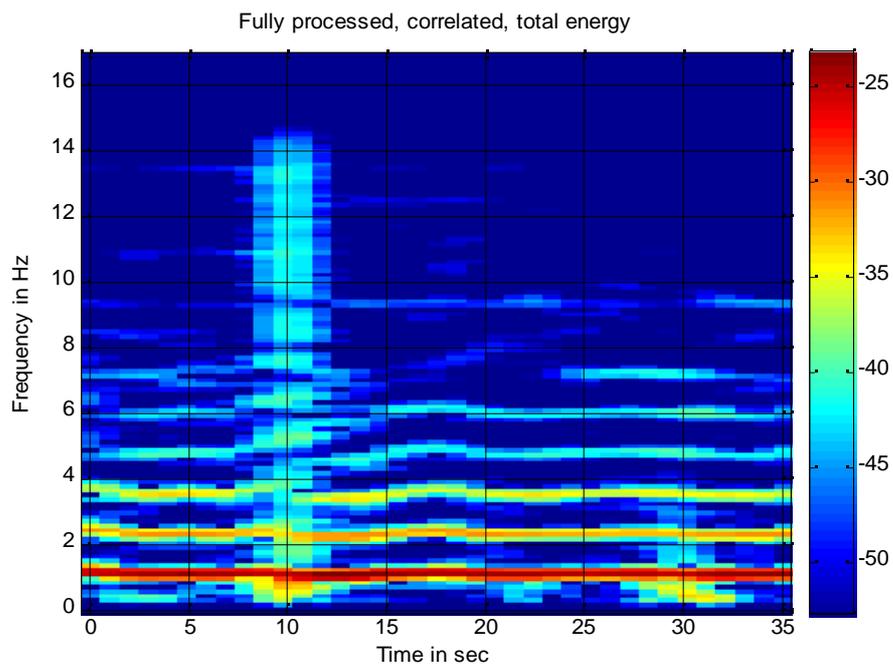


Abbildung 7.19: Spektrum nach Anwendung der Umgebungslichtsubtraktion

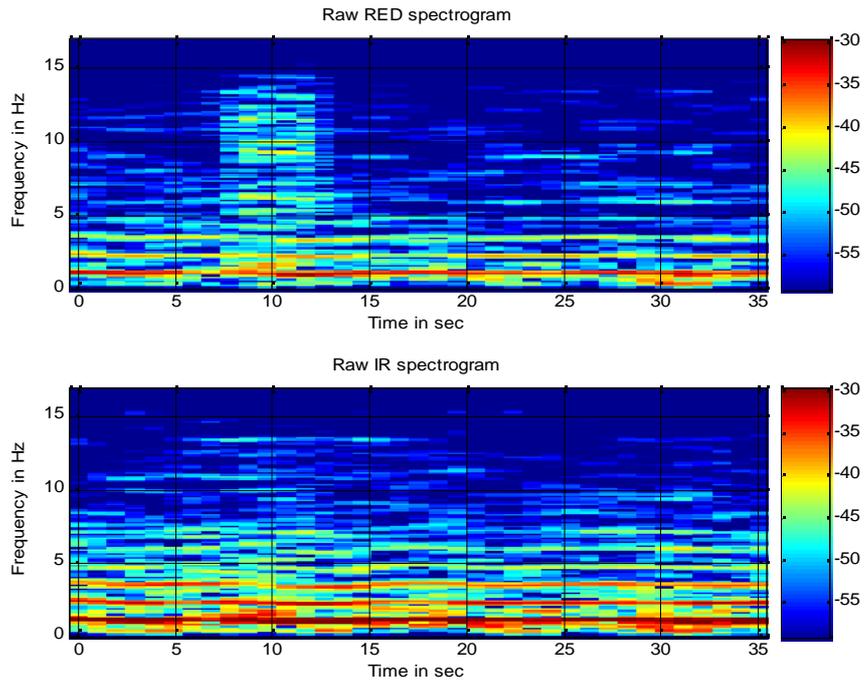


Abbildung 7.20: Spektrum nach Anwendung der Umgebungslichtsubtraktion und dem Interference Canceler (20. Ordnung)

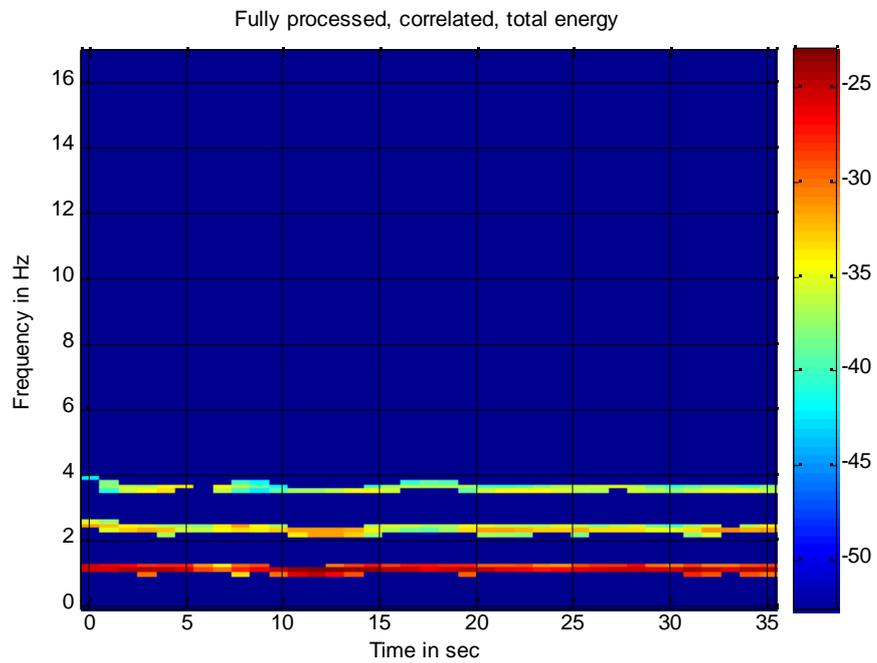


Abbildung 7.21: Spektrum nach Anwendung der Umgebungslichtsubtraktion, des Interference Cancelers (20. Ordnung) und der Spektralen Maske nach dem Prinzip aus Kap. 4.3.11

Abb. 7.22 zeigt, wie das Zuschalten der Filter zu einem stabilen SpO₂-Wert führt, der erwartungsgemäß bei 100% liegt. Zunächst ist mittels der Umgebungslichtsubtraktion der niederfrequente Anteil an Störungen entfernt worden. Durch den Einschalt- und Ausschaltspung ist eine breitbandige Störung bis ins Übertragungsband entstanden, die allein mit der Umgebungslichtsubtraktion nicht aus dem Nutzsignal entfernt werden kann. Mit dem Interference Canceler konnte ein Teil der hochfrequenten Störungen, die ins Nutzband gespiegelt worden sind, eliminiert werden. Schließlich wurden die zur Messung der Sauerstoffsättigung relevanten Frequenzen mittels der Spektralen Maske selektiert. Das Resultat ist ein über die gesamte Messdauer konstanter SpO₂-Wert.

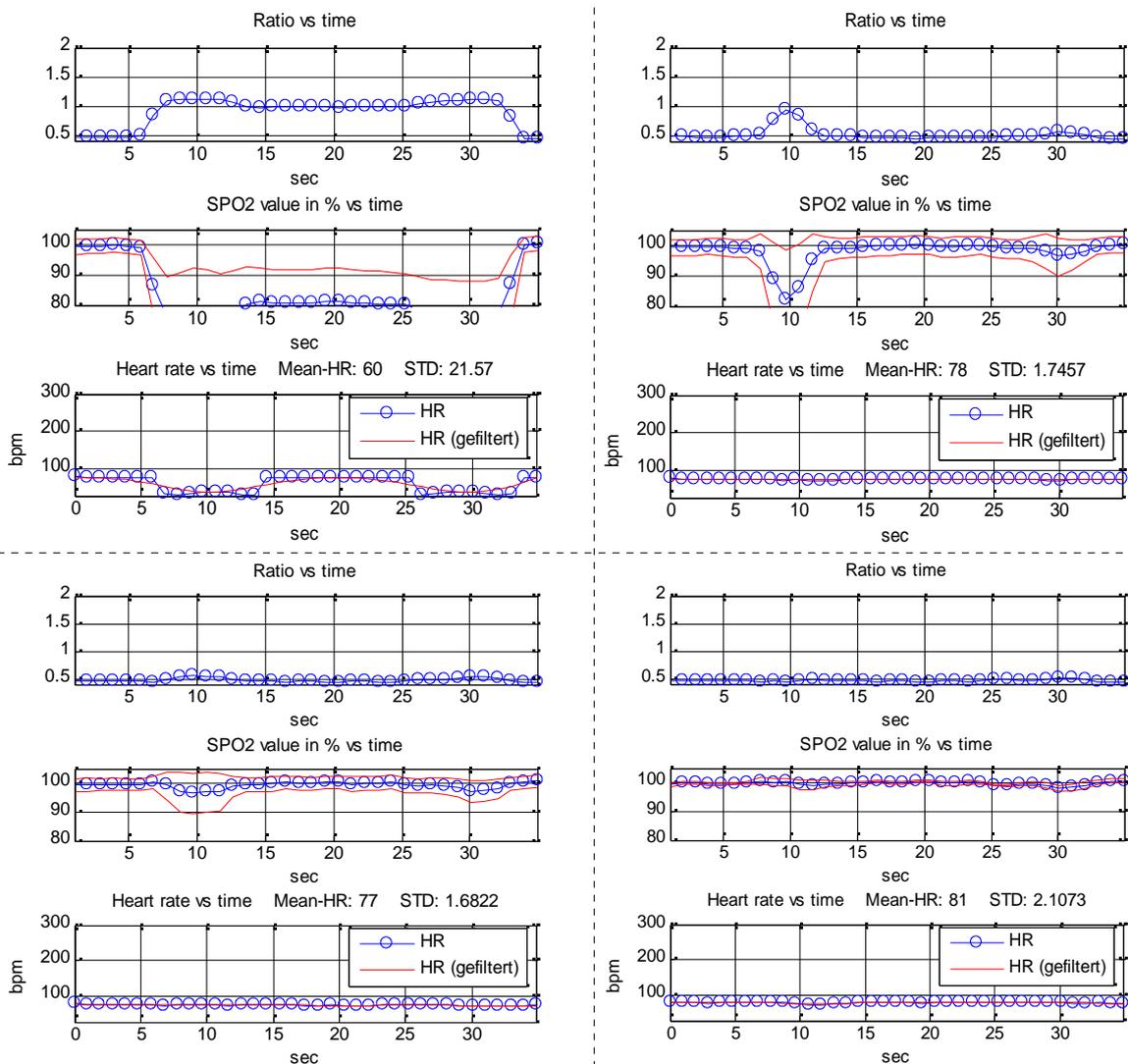


Abbildung 7.22: 1. Links oben die Vitalparameter zum Signal, ohne die zuschaltbare Filterung
 2. Rechts oben die Vitalparameter zum Signal nach Anwendung der Umgebungslichtsubtraktion
 3. Links unten die Vitalparameter zum Signal nach Anwendung der Umgebungslichtsubtraktion und des Interference Cancelers
 4. Rechts unten die Vitalparameter zum Signal nach Anwendung der Umgebungslichtsubtraktion, des Interference Cancelers und der Spektralen Maske

7.2.4 100-Hz-Störer Unterdrückung

Zur Untersuchung der 100-Hz-Störer-Unterdrückung werden alle zuschaltbaren Filter deaktiviert. Abb. 7.23 zeigt das Spektrum des Sensorsignals aus der Test-Episode 29. Im Spektrum wirken die äquidistanten Träger des Spreizsignals sehr klein im Verhältnis zu der 100Hz-Störung. Die Störung besteht aus einer Grundwelle bei 100Hz, einer Oberwelle bei 200 Hz und einer Oberwelle bei 400Hz. (Die Oberwelle bei 300 Hz ist kaum erkennbar). Das Verhältnis der Grundwelle des Störers zu einem der Träger aus dem Spreizspektrum beträgt rund 36dB. Mit dem 100-Hz-Filter aus Kapitel 4.2.2 wird der Störer herausgefiltert und der SpO₂-Wert bleibt über die Messzeit konstant.

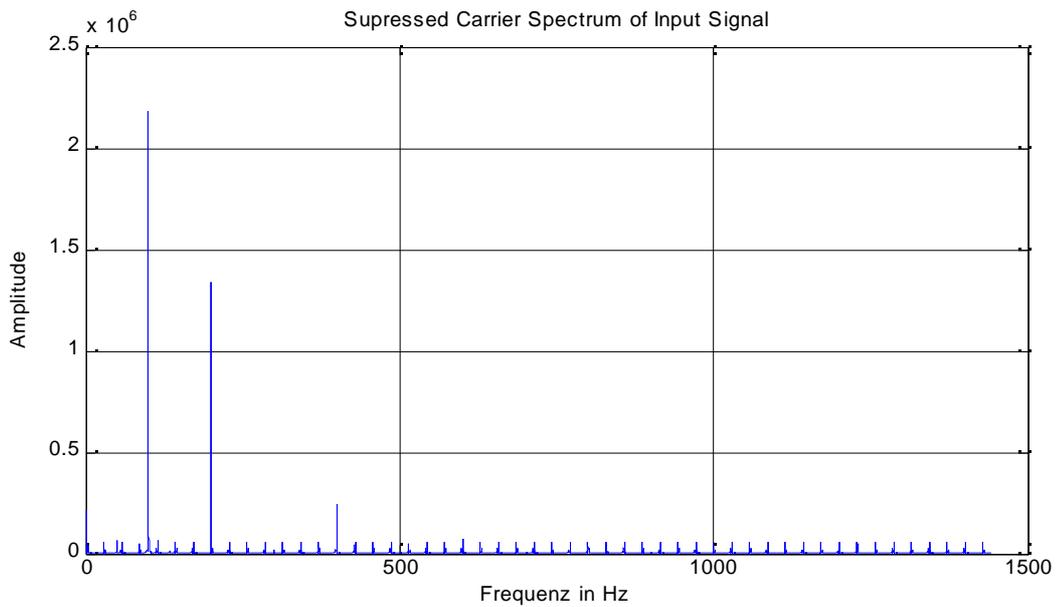


Abbildung 7.23: Spektrum des Sensorsignals aus der Episode 29

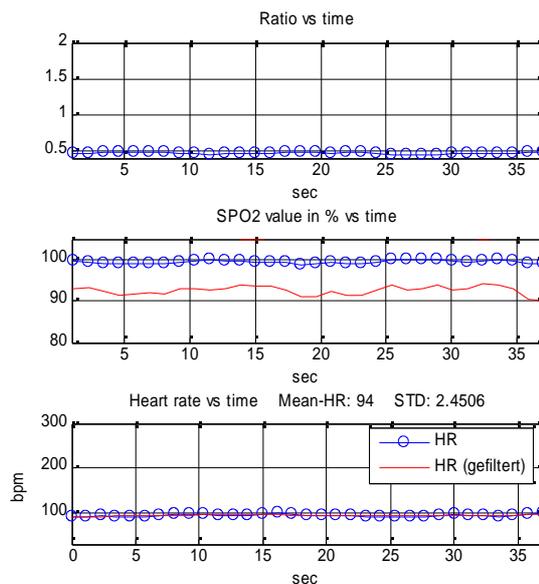


Abbildung 7.24: Verlauf der Vitalparameter

7.3 Referenzmessung der Herzfrequenz

Die Referenz-Herzfrequenz wurde mit einer Zählung der Pulse an der Halsschlagader und unter Auszählung der Schwingungen im Zeitsignal bestimmt. Der Vergleich zeigte Abweichungen bis zu 6.5%. Der Fehler kann durch weitere Optimierungsmaßnahmen an dem Skript *hr.m* minimiert werden.

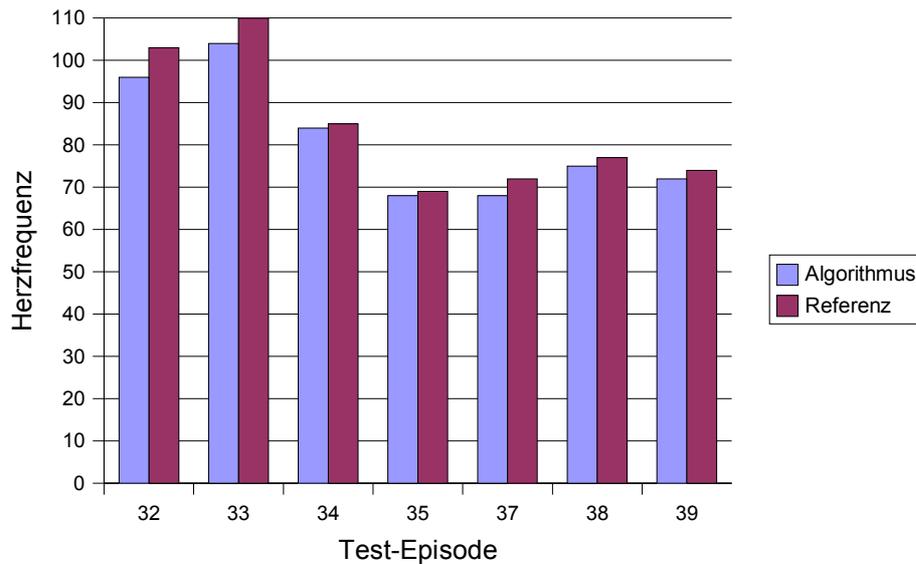


Abbildung 7.25: Mit dem Algorithmus aus Kap. 4.3.12 gemessene Herzfrequenz im Vergleich zur gezählten Herzfrequenz

8 Ausblick

Das hier vorgestellte System bietet eine Grundlage zur Entwicklung eines produktreifen Pulsoximeters. Im nächsten Schritt müssen weitere Algorithmen zur Unterdrückung von Bewegungsartefakten integriert werden. Bisher trägt die Singularwertzerlegung (Kap. 4.3.13) und die Spektrale Maske (Kap. 4.3.11) einen geringen Schutz in Anwesenheit starker Bewegungsartefakte. Hier könnte ein „Faden-Algorithmus“ zu guten Ergebnissen führen, wie er in [Käs] beschrieben wird. Die Entwicklung der Algorithmen sollte weiterhin unter MATLAB erfolgen.

Eine Kalibrierung des Systems ist unumgänglich, um akkurate Aussagen über die Qualität der eingesetzten Verfahren zur Störungsunterdrückung jeglicher Art treffen zu können. Die Kalibrierung kann in der Klinik erfolgen. Besser geeignet ist ein Modell des Blutkreislaufs mit einer variablen Sauerstoffanreicherung zur Kalibrierung und zum Testen von Pulsoximetern.

Ein wichtiger Punkt ist der Funktionsumfang und die damit verbundene Leistungsaufnahme des Mikrocontrollers. Es steht bislang nicht fest, ob der Mikrocontroller TriCore zur Vereinigung mehrerer Messsysteme eingesetzt wird. Für diesen Fall ist der TC1765 ausreichend mit Ressourcen und Peripherie-Einheiten ausgestattet. Aber allein zur Messung der Sauerstoffsättigung ist der Mikrocontroller überdimensioniert.

Die Hardware wurde bislang minimal gehalten, um nicht nachvollziehbaren Effekten, die durch Einsatz analoger Bauelemente entstehen, während der Entwicklung der digitalen Algorithmen, aus dem Weg zu gehen. An der Elektronik lassen sich noch Verbesserungen zu Gunsten der Signalqualität vornehmen.

Der Spreizspektrum-Demodulator und die Umgebungslichtsubtraktion sind bereits auf das Echtzeitsystem portiert worden. Da noch kein Algorithmus zur Unterdrückung von Bewegungsartefakten existiert, wäre die Portierung des bereits vorhandenen SpO_2 -Algorithmus verfrüht. Es kann zu Änderungen des vorhandenen Algorithmus kommen. Der nächste Schritt sollte die Entwicklung eines Algorithmus zur Unterdrückung von Bewegungsartefakten sein. Erst dann ist das Spektrum an Methoden zur Störungsbeseitigung abgedeckt und der Algorithmus zur Portierung auf ein Echtzeitsystem komplett.

Literaturverzeichnis

- [MIZ] Konrad Kunsch, Der Mensch in Zahlen, 1996 (ISBN: 3-437-25200-3)
- [HB] Eberhardt L. Betz, Herz und Blutkreislauf, 1991 (ISBN: 3-89330-754-0)
- [BP] Birbaumer Schmidt, Biologische Psychologie, 1996 (ISBN: 3-540-59427-2)
- [IMZ] Isabella Maria Zimmermann, Entwicklung eines Kalibrationsverfahrens für ein fetales Pulsoximeter zur Anwendung sub partu, 2001 (Dissertation)
- [AS] Anke Schmitz, Pulsoximetrie unter der Geburt, 1995, Universität Erlangen (Dissertation)
- [DOP] J G Webster, Design of pulse oximeters, 1997 (ISBN: 0-7503-0467-7)
- [Käs] Sigfried Kästle, Ein Algorithmus zur zuverlässigen Verarbeitung von Pulsoximetrie-Signalen bei schwierigen Störverhältnissen, 1999 (ISBN: 3-89722-274-4)
- [For] L. Rade, B. Westerhagen, Mathematische Formeln, 2000 (ISBN: 3-540-67505-1)
- [PH] Hering Martin Stohrer, Physik für Ingenieure, 2002 (ISBN: 3-540-42964-6)
- [ME] Josef Eichmeier, Medizinische Elektrotechnik, 3. Auflage, 1996 (ISBN: 3-540-61499-0)
- [KS] Hans Dieter Lüke, Korrelationssignale, 1992 (ISBN: 3-540-54579-4)
- [Sch] M. R. Schröder, Number Theory in Science and Communication, Third Edition, (ISBN: 3-540-15800-6)
- [NUM] William H. Press, Numerical recipes in C, 1999 (ISBN: 0-521-43108-5)
- [MAT] Gerhard Doblinger, MATLAB - Programmierung in der digitalen Signalverarbeitung, 2001 (ISBN: 3-935340-02-8)
- [Grü] Daniel Ch. von Grünigen, Digitale Signalverarbeitung, 1993 (ISBN: 3-8007-1971-1)
- [TS] Tietze Schenk, Halbleiter Schaltungstechnik 12. Auflage, 2001 (ISBN: 3-540-42849-6)

Anhang A: Verzeichnis der Test-Episoden

Nr.	Datei	Blackbox	Proband	Beschreibung
1	IR_10Hz_rect(oben).tpd	Ja	5	Von oben mit Rechtecksignal über IR-LED bei 10 Hz
2	IR_110Hz_sin(oben).tpd	Ja	5	Von oben mit Rechtecksignal über IR-LED bei 110 Hz
3	UVO_IR_LED_ein_aus(oben).tpd	Ja	5	Von der Seite durch Sprung über IR-LED
4	UVO_blue_10Hz(seitlich).tpd	Ja	5	Von der Seite durch Blaue LED bei 10 Hz
5	UVO_in_black_box_ohne_Stoerung.tpd	Ja	5	Keine
6	UVO_rote_LED_variiert(oben).tpd	Ja	5	Rote LED von oben; Langsame Variation der Intensität
7	UVO_weisse_LED_100Hz(seitlich).tpd	Ja	5	Weisse LED seitlich; 100 Hz Sinus
8	UVO_weisse_LED_10Hz(seitlich).tpd	Ja	5	Weisse LED seitlich; 10 Hz Sinus
9	UVO_weisse_LED_variiert(seitlich).tpd	Ja	5	Weisse LED seitlich; Langsame Variation der Intensität
10	White_1000Hz_sin(seite).tpd	Ja	4	Weisse LED seitlich; 1000 Hz Sinus
11	White_100Hz_rect(seite).tpd	Ja	4	Weisse LED seitlich; 100 Hz Rechteck
12	Ohne_Signal.tpd	Ja	4	Keine
13	WHITE_SWEEP_2(seite).tpd	Ja	4	Weisse LED seitlich; Sweep
14	white_SWEEP(seite).tpd	Ja	4	Weisse LED seitlich; Sweep
15	IR_8Hz_sin(oben).tpd	Ja	5	IR LED von oben; 8 Hz Sinus
16	IR_8Hz_sin(oben_vorspann).tpd	Ja	5	IR LED von oben; 8 Hz Sinus
17	IR_sweep2_0-30Hz.tpd	Ja	5	IR LED von oben; 0-30 Hz Sweep
18	IR_sweep3_0-30Hz.tpd	Ja	5	IR LED von oben; 0-30 Hz Sweep
19	IR_sweep_0-30Hz.tpd	Ja	5	IR LED von oben; 0-30 Hz Sweep
20	IR_sweep_20-40Hz.tpd	Ja	5	IR LED von oben; 20-40 Hz Sweep
21	IR_sweep_90-110Hz.tpd	Ja	5	IR LED von oben; 90-110 Hz Sweep
22	White_8Hz_sin(seitlich).tpd	Ja	5	Weisse LED seitlich; 8 Hz Sinus
23	Daylight.tpd	Nein	4	Tageslicht evtl. Monitor (EIZO)
Ab hier mit Offset -12mV am Transimpedanzwandler				
24	O_White_10Hz.tpd	Ja	4	Weisse LED seitlich; 10 Hz Sinus
25	O_SWEEP.tpd	Ja	4	Weisse LED seitlich; Sweep 0.1 - 100 Hz Sinus
26	O_Pulse_Per5s_DCyc5%.tpd	Ja	4	Weisse LED seitlich; Pulse mit Periode 5Hz, duty cycle 5%
27	NS_100Hz_1.tpd	Ja	4	100 Hz Störer mit Oberwellen, weisse LED, Finger geneigt
28	NS_100Hz_2.tpd	Ja	4	100 Hz Störer mit Oberwellen, weisse LED, Finger geneigt
29	NS_100Hz_3.tpd	Ja	4	100 Hz Störer mit Oberwellen, weisse LED, Finger geneigt
30	NS_200Hz_1.tpd	Ja	4	200 Hz Störer mit Oberwellen, weisse LED, Finger geneigt
31	NS_200Hz_2.tpd	Ja	4	200 Hz Störer mit Oberwellen, weisse LED, Finger geneigt
32	A1.tpd	Ja	4	94bpm mit Zählung
33	A2.tpd	Ja	4	104bpm mit Zählung
34	HF_1.tpd	Ja	4	84bpm mit Zählung
35	HF_2.tpd	Ja	8	68bpm mit Zählung
36	HF_3.tpd	Ja	10	
37	HF_4.tpd	Ja	9	68bpm mit Zählung
38	HF_5.tpd	Ja	6	75bpm mit Zählung
39	HF_6.tpd	Ja	7	72bpm mit Zählung

Anhang B: Signalanalyse

Mit dem bereits vorhandenen Prototypen wurden die Signale auf ihre Eigenschaften hin untersucht. Dazu wurden nacheinander alle DC- und AC-Signale mit 12-Bit Quantisierung 600 Samples pro Sekunde abgetastet. Bei der Abtastung ist das erste Nyquist-Kriterium mit einem Anti-Aliasing-Filter bei einer Grenzfrequenz $f_g = 200$ Hz erfüllt worden. Ein solches Filter zweiter Ordnung mit Butterworth-Charakteristik und Einfachmitkopplung, ist durch die Parameter $a_1 = 1.4142$ und $b_1 = 1$ beschrieben. Die Bauelemente $C = 4.7000e-008$ und $R_4 = 10k$ wurden festgelegt. Somit folgt für $R_1 = R_2 = 1.6931\Omega$ und $R_3 = 5858\Omega$. Die Berechnung kann [TS] entnommen werden.

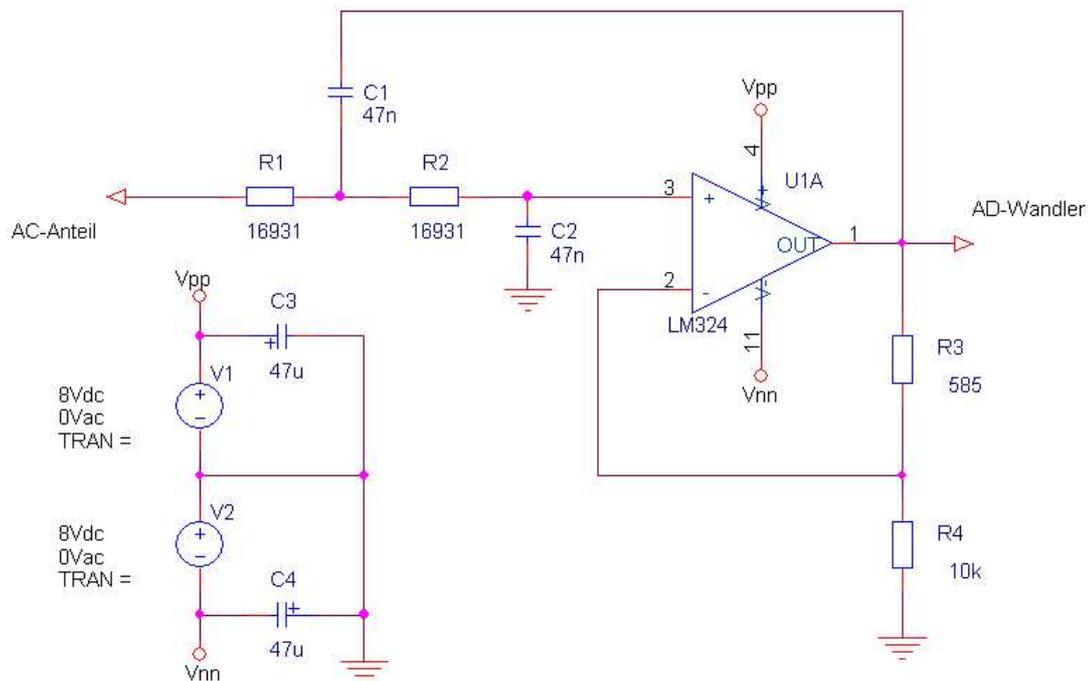


Abbildung 8.1: Butterworth-TP-Filter mit einer Grenzfrequenz $f_g = 200$ Hz

Die Datenerfassung erfolgt mit LabJack, einer externen Datenerfassungseinheit via USB. Zu diesem Zweck wurde ein M-Skript `spo2_rd.m` geschrieben, mit dem sich die Abtastpunkte direkt in MATLAB importieren lassen. Die Untersuchung im Zeitbereich zeigte, dass im Zeitmultiplex mindestens 16-bit-Quantisierung notwendig wären, um den DC-Anteil und den pulsatilem Anteil gemeinsam abzutasten. Damit ließe sich der pulsatile Anteil bei voller Signalaussteuerung am AD-Wandler mit 6 bis 8 bit quantisieren.

Im Frequenzbereich wurden die Frequenzen untersucht, die auf arterielle Volumenänderungen zurückzuführen sind. Das Signal besteht aus einer Grundwelle bei f_g und seinen Oberwellen bei $f_{o1}, f_{o2}, \dots, f_{oN}$. Bei störungsfreien Bedingungen beträgt das SNR für die Grundwelle bis zu 25dB. Die erste und zweite Oberwelle kann noch gut im Spektrum erkannt werden, während ab der dritten Oberwelle das SNR nunmehr 0 bis maximal 6 dB beträgt. Deshalb fällt der Fokus auf die Grundwelle und die zwei nächsten Oberwellen.

Für die Frequenzen gilt immer die harmonische Beziehung:

$$f_{o1} = 2 \cdot f_g, \quad f_{o2} = 3 \cdot f_g \quad (8.1)$$

Das Spektrum eines ungestörten Signals ist stets abklingend, so dass folgendes gilt:

$$|X(f_g)| > |X(f_{o1})| > |X(f_{o2})| \quad (8.2)$$

Die Wahl der zulässigen Herzrate im Intervall [30, 300] bpm bedeutet für die Frequenz der Grundwelle einen zulässigen Bereich von [0.5, 5] Hz. Mit Gl. 8.1 folgt für die 2. Oberwelle ein Intervall [1.5, 15] Hz. Mit Minimum und Maximum aus allen Frequenzintervallen folgt für den gesamten Frequenzbereich des Nutzsignals im Basisband ein Intervall [0.5, 15] Hz. Dieses Intervall muss für die Auslegung der Systemparameter berücksichtigt werden, so dass die Pulsrate bis zu 300 bpm berechnet werden kann. Demnach lässt sich ein Plethysmogramm bereits ausreichend mit einer Abtastrate von 30 Samples pro Sekunde darstellen.