

# Objektive Tunerqualifizierung

Entwicklung eines Klassifikators zur Störungserkennung  
in Audiosignalen

## Diplomarbeit

von

**Christian Siller**

ausgeführt am

Institut für Signalverarbeitung und Sprachkommunikation  
der Technischen Universität Graz

Vorstand: O. Univ.-Prof. Dipl.-Ing. Dr. techn. Gernot Kubin

Begutachter und Betreuer:

Dipl.-Ing. Dr. techn. Christian Feldbauer

Graz, 9. März 2010

## **Danksagung**

Während des Studiums und der Erarbeitung der vorliegenden Diplomarbeit haben mich viele Personen begleitet und bestärkt. Dafür gebührt ihnen mein Dank.

Besonders herzlich danke ich meinen Eltern, die mich in all diesen Jahren unterstützt haben.

Gute Einfälle sind Geschenke des Glücks.

Gotthold Ephraim Lessing

## **Kurzfassung**

Im Rahmen der Diplomarbeit „OBJEKTIVE TUNERQUALIFIZIERUNG – ENTWICKLUNG EINES KLASSIFIKATORS ZUR STÖRUNGSDETEKTION IN AUDIOSIGNALEN“ wird ein Konzept für einen Klassifikator entwickelt, der zur objektiven Bewertung und Gegenüberstellung von Rundfunkempfangsgeräten eingesetzt werden kann. Es wird gezeigt, welche Methoden es zur Detektion und Klassifikation gibt. Zudem wird ein Überblick über typische Features zur Beschreibung von Audiosignalen gegeben; eine Selektion daraus wird zur Feature-Extraktion herangezogen. Die gewonnenen Attribute der aufgezeichneten Signale sollen ein geeignetes Klassifikationsverfahren zur Detektion der auftretenden Störungen ermöglichen. Die Implementierung erfolgt in MATLAB<sup>®</sup>. In diesem Zusammenhang wird erörtert, welche Aufbereitung der Messdaten möglich bzw. notwendig ist, und wie die gesammelten Messdaten in eine Datenbank zusammengefasst werden können. Außerdem erfolgt die Auswertung der aus Testfahrten gewonnen Messdaten. Die Schlussfolgerungen aus den gewonnen Erkenntnissen stehen vor einem abschließenden Ausblick auf die Möglichkeiten von Klassifikationsverfahren.

## **Abstract**

The thesis „Objective Tuner Qualification – Development of a classifier for interference in audio signals“ is concerned with the detection and classification of interferences in audio signals. The aim is to develop an objective method for evaluating radio broadcast receivers. Known models should be investigated on their usability for the appropriate tasks. In addition pros and cons of these models should be found. An overview is given of frequently used audio features for pattern classification. Afterwards for comparison, a detection algorithm by means of some practical example will be developed. For this practical work a test setup for data acquisition should be found. The algorithms are implemented in MATLAB<sup>®</sup>. In this interrelation I will show how pre-processing of performance data it's possible and should be done and how all the measured data can be collected in a database for classification. In addition the performance of classification in a practical example will be evaluated. The summary from the collected technical expertise are listed before an overview of the possibilities for detection and classification algorithms.

## EIDESSTÄTTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Graz, am 09. März 2010

.....  
Christian Siller

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
	Marktvergleich . . . . .	2
	MSF-Testverfahren . . . . .	3
	Motivation . . . . .	4
<b>2</b>	<b>Mustererkennung</b>	<b>8</b>
2.1	Klassifikationsverfahren . . . . .	8
2.2	Der Entwicklungsprozess . . . . .	13
2.2.1	Datensammlung . . . . .	14
2.2.2	Datenvorverarbeitung . . . . .	14
2.2.3	Feature-Auswahl . . . . .	15
2.2.4	Feature-Extraktion . . . . .	15
2.2.5	Klassifikationsprozess . . . . .	15
2.2.6	Auswertung . . . . .	15
2.3	Audio-Features . . . . .	17
<b>3</b>	<b>Konzept</b>	<b>23</b>
3.1	Aufzeichnungsmethoden zur Datenerfassung . . . . .	24
3.2	Testsetup . . . . .	25
3.2.1	Voruntersuchung . . . . .	27
3.3	Datenerfassung . . . . .	29
3.3.1	Aufnahmesetup . . . . .	29
3.3.1.1	Mobile Prüfvorrichtung (DUT) . . . . .	29
3.3.1.2	Stationäre Referenz . . . . .	30
3.3.2	Messfahrt . . . . .	31
3.3.3	Datenaufbereitung . . . . .	32
3.3.4	Labelung . . . . .	33
<b>4</b>	<b>Implementierte Algorithmen</b>	<b>36</b>
4.1	Feature-Extraktion . . . . .	40

4.1.1	Melbank . . . . .	40
4.1.2	Kurzzeit-Fourier-Transformation . . . . .	41
4.1.3	Differenz von Test- und Referenzsignal . . . . .	42
4.1.4	Verwendete Feature-Vektoren . . . . .	46
4.2	Sortierung der Feature-Vektoren . . . . .	47
4.3	Wahl des Klassifikators . . . . .	48
4.3.1	Nachbearbeitung des Trainingsdatensatzes . . . . .	49
<b>5</b>	<b>Test des Klassifikators</b>	<b>51</b>
<b>6</b>	<b>Zusammenfassung und Ausblick</b>	<b>60</b>
6.1	Zusammenfassung . . . . .	60
6.2	Ausblick . . . . .	60
<b>A</b>	<b>Auswertung</b>	<b>62</b>
<b>B</b>	<b>UKW-Radio</b>	<b>76</b>
B.1	Einführung UKW-Rundfunk . . . . .	76
B.2	ISO-Stecker . . . . .	78
B.3	Teststrecke . . . . .	82
B.4	Universal Receiver Tester (URT) . . . . .	83
B.5	ServiceForce . . . . .	85
	<b>Abbildungsverzeichnis</b>	<b>86</b>
	<b>Tabellenverzeichnis</b>	<b>87</b>
	<b>Literaturverzeichnis</b>	<b>89</b>

# 1 Einleitung

UKW-Rundfunk wird weltweit seit Jahren im VHF-Band II zwischen 87,5 MHz und 108,0 MHz betrieben, wobei in manchen Ländern nur Teilbereiche davon genutzt werden. Lediglich Japan und die Länder des ehemaligen Ostblocks verwenden andere Frequenzbereiche. Als Modulationsart wird die Frequenzmodulation (FM) benutzt. Diese ermöglicht eine weniger störungsanfällige Übertragung von Rundfunksignalen, da sie im Vergleich zu der im Lang-, Mittel- und Kurzwellenbereich verwendeten Amplitudenmodulation (AM) recht unempfindlich gegenüber atmosphärischen Störungen ist. (Quelle: WIKIPEDIA)

Rundfunkempfangsgeräte werden sehr vielfältig eingesetzt. Eines der uns bekanntesten Einsatzgebiete ist der KFZ-Bereich. Nicht nur die technologische Weiterentwicklung der Rundfunktechnik (DAB-T, HD-Radio), sondern auch die steigenden Ansprüche der HörerInnen erfordern eine immer höhere Qualität der Rundfunkempfangsgeräte. Einzelkomponenten von Infotainmentanlagen wie DA-Wandler, Verstärker oder Lautsprecher lassen sich weitgehend problemlos auf ihre Übertragungseigenschaften hin untersuchen. Eine objektive Bewertung der Empfangsqualität eines Tuners ist bislang jedoch nur eingeschränkt möglich.

Es besteht zwar die Möglichkeit von Feldstärkemessungen im Zusammenhang mit einer HF-Messung am Eingangsteil des Tuners, jedoch kann diese Methode für sich nicht als Basis für eine objektive Aussage über die Qualität des tatsächlich akustisch wahrgenommenen Signals herangezogen werden. Auch bereits existierende Messeinrichtungen zur Tunerbewertung weisen hinsichtlich der Objektivität noch Schwachstellen auf, wie folgender Marktvergleich zeigt.

## Marktvergleich

Eine auf dem Markt erhältliche Messeinrichtung, die speziell zum Testen von KFZ-Tunern entwickelt wurde, ist der *Universal Radio Tester (URT)* von *Mindready*. Mit diesem modularen Testgerät ist es möglich, die verschiedenen Protokolle sowohl zu senden, als auch zu empfangen. Eine der mit diesem Gerät möglichen Messungen wird zur Abschirmung gegen äußere Einflüsse üblicherweise in einer EMV-Kabine durchgeführt. Dabei kann zum Beispiel die Sendeleistung am URT so lange reduziert werden, bis es zu Störungen im Radioempfang kommt. Gemessen wird mit geringer Sendeleistung, sodass das Gerät ohne zusätzliche Sendelizenz betrieben werden kann.

Eine Möglichkeit der Empfangsqualitätsmessung besteht darin, am Empfangsteil des *URT* die gewünschte Frequenz einzustellen und die gemessene Leistung im Analysator auszuwerten. Diese Messung findet üblicherweise im Freifeld statt.

Vorteile des *URT*:

- Messungen der Empfangsqualität mit beliebigem Sendematerial sind möglich.
- Die Messungen sind reproduzierbar.
- Die abgestrahlte Leistung ist bekannt.

Nachteile des *URT*:

- Es ist nicht möglich, reale Bedingungen nachzubilden.
- Eine objektive Aussage über das akustische Signal am Lautsprecher-Ausgang ist nicht möglich.
- Die Abstrahlung des Funkfeldes kann nicht beeinflusst werden. (Nahfeld/Fernfeld)

Weitere herstellerepezifische Informationen zu diesem Produkt finden sich im Anhang. [Anhang B.4: URT-AVERNA]

Die zweite Messeinrichtung, die sich der objektiven Tunerqualifizierung widmet, wurde vom Prüf- und Dienstleistungsunternehmen *ServiceForce-GmbH* beim Anwender- und Technologiekongress von National Instruments „Virtuelle Instrumente in der Praxis“ (VIP08) vorgestellt. Bei diesem Verfahren kommen Prozeduren nach ITU-R BS.1387-1 [ITU-R] zur Anwendung. Darin enthalten sind standardisierte Verfahren zur Bewertung von Audiocodecs (PESQ, PEAQ). Für die Messung wird ein Referenztuner im KFZ mitgeführt. Das Signal des Referenztuners wird von dem des Testtuners subtrahiert. Die Differenz der beiden Signale wird zur Bewertung herangezogen.

Vorteile des Messverfahrens:

- Eine Bewertung von Audiogeräten und -komponenten ist möglich.
- Die Bewertung des Audiosignals ist unmittelbar einsehbar (On-line-monitoring).
- Funktionstests von Audiosystemen sind möglich.

Nachteile des Messverfahrens:

- Die Messungen sind nicht reproduzierbar.
- Der Bewertungsmetrik liegt ein anderer Anwendungsbereich zugrunde, was einen systematischen Fehler im gesamten Messverfahren darstellt.

Detaillierte Informationen zu diesem Messverfahren finden sich im Anhang. [Anhang B.5: ServiceForce]

## **MSF-Testverfahren**

Magna Steyr Fahrzeugtechnik (MSF) hat im Laufe der vergangenen Jahre verschiedene Messfahrten durchgeführt, um eine Aussage über die Qualität von KFZ-Tunern treffen zu können. In diesem Zusammenhang ist zu erwähnen, dass auf einer der drei MAGNA-Teststrecken außerhalb von Graz (Grüner See) eine Feldstärkemessung vorhanden ist. Damit ist zwar eine bessere Beurteilung der Testbedingung möglich, letztendlich kann aber auch dadurch lediglich eine subjektive Aussage über die Qualität des Tuners getroffen werden. Die von MSF durchgeführten Testfahrten

zur Überprüfung der Empfangseigenschaften des KFZ-Tuners liefen nach folgendem Szenario ab:

Zwei Frequenzen – 89,20 MHz (HITRADIO Ö3) und 91,20 MHz (ÖSTERREICH 1) – werden *ohne* Alternativfrequenz<sup>1</sup> getestet. Dabei wird beim Durchfahren der Teststrecke von zwei bis drei Testpersonen vermerkt, ob auf einem Intervall von 100 m eine Störung wahrgenommen werden konnte, die

- kürzer als 1 Sekunde,
- genau 1 Sekunde oder
- länger als 1 Sekunde

dauert bzw. ob keine Störung aufgetreten ist. Durchgehendes Rauschen wird zusätzlich notiert.

Bei der Wahl der Teststrecke wurde darauf geachtet, dass sich der Empfang im Grenzbereich befindet, dass also weder optimale Empfangsqualität noch durchgehend schlechter Empfang auf der gesamten Teststrecke vorherrscht. Dazu wählte man eine öffentliche Strecke nahe Graz, die in einem Lageplan im Anhang ersichtlich ist. [Anhang B.3: Teststrecke]

## Motivation

### Warum testen

Mit fortschreitender technischer Entwicklung steigt der Qualitätsanspruch an Produkte. Gleichzeitig ist eine zunehmende Skepsis der Konsumenten gegenüber neuen, unbekanntem Produkten zu beobachten. In enger Wechselwirkung mit diesen Faktoren steht auf der Produzentenseite die Notwendigkeit, über die Qualität eines Produktes bestmöglich Bescheid zu wissen, um es gegebenenfalls verbessern und auf dem Markt besser platzieren zu können. Subjektive Aussagen über einzelne Produkte können dieser Problematik aber nicht gerecht werden. Vielmehr bedarf es dazu eines objektiven Vergleichs auf der Grundlage entsprechender Bewertungskriterien.

---

<sup>1</sup>Alternativfrequenz: Verwendung von jeweils nur einer Frequenz - durch Ausschalten des RDS

In unserem Fall bedeutet dies, dass eine objektive Bewertungsmetrik gefunden werden muss, mit der verschiedene UKW-Tuner getestet und einander gegenübergestellt werden können. Wie der oben angestellte Marktvergleich zeigt, gibt es zurzeit KEIN Gerät zur objektiven Tunerqualifizierung. Die bestehenden, subjektiven Testverfahren lassen objektive Vergleiche von unterschiedlichen UKW-Tunern nicht zu, sie ermöglichen lediglich eine Aussage über die Eigenschaften eines einzelnen Produktes. Da auch bei den bislang durchgeführten MSF-Testfahrten die Anzahl der Testpersonen sehr gering ist (in der Regel zwei bis drei), fällt diese Aussage außerdem sehr subjektiv aus. Ein Testverfahren zu entwickeln, das eine objektive Tunerqualifizierung ermöglicht, ist Ziel dieser Arbeit.

### **Warum klassifizieren**

Grundlage für eine objektive Bewertung ist es, die auftretenden Störfälle zu kennen. Dazu ist es notwendig, aus einer großen Datenmenge die möglichen Störfälle zu extrahieren und entsprechenden Klassen zuzuordnen. Dadurch wird die Möglichkeit zur Klassifikation und damit die Voraussetzung für eine automatische Detektion der gefundenen Störklassen am UKW-Tuner geschaffen.

### **Warum bewerten**

Eine Produktbewertung ist für den Endkunden ebenso bedeutsam wie für den Techniker oder für den Manager eines Unternehmens. Die dabei erforderliche Bewertungsskala muss den unterschiedlichen Bedürfnissen entsprechend detailliert sein und die jeweils relevanten Informationen bereitstellen. Die Basis für die Bewertung eines KFZ-Tuners ist die Klassifikation. Durch psychoakustische Tests kann eine Bewertung zusätzlich verfeinert werden. Es kann zum Beispiel eine einfache Bewertungsmetrik erstellt werden, so wie sie in Abbildung 1.1 dargestellt ist.

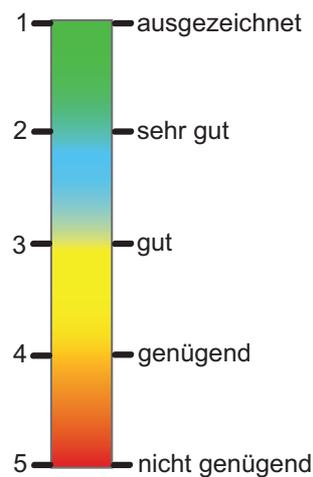


Abbildung 1.1: Einfache Bewertungsmetrik, nach Schulnotensystem zur Tunerqualifikation

## Zusammenfassung

Im Rahmen dieser Diplomarbeit wird ein Konzept für einen Klassifikator entwickelt, der zur objektiven Bewertung und Gegenüberstellung von Rundfunk empfangsgeräten eingesetzt werden kann.



Abbildung 1.2: Struktogramm Tunerqualifizierung

In Abbildung 1.2 ist ein schematischer Ablauf der Tunerqualifizierung dargestellt. Die Testfahrten und die Klassifikation sind Teil dieser Diplomarbeit. Der Schwerpunkt der Arbeit liegt auf der Auswahl von Attributen von Audiosignalen. Durch die Extraktion dieser Features soll ein geeignetes Klassifikationsverfahren zur Bewertung der auftretenden Störungen entwickelt werden. Damit wird hier auch eine Ausgangsbasis für weiterführende Arbeiten geschaffen, in denen diese Störklassen psychoakustischen Versuchen unterzogen werden, um eine objektive Bewertungsmetrik zu erstellen.

---

Das Gesamtprojekt *Objektive Tunerqualifizierung* wird in Zusammenarbeit mit Magna Steyr Fahrzeugtechnik (MSF), IVM-Technical Consulting (IVM-TC) und dem Institut für Signalverarbeitung und Sprachkommunikation (SPSC) an der Technischen Universität Graz durchgeführt.

## Gliederung

Der erste Abschnitt dieser Arbeit befasst sich mit der MUSTERERKENNUNG im Allgemeinen, um die theoretischen Grundlagen für die nachfolgenden Ausführungen zu schaffen. Dabei wird anhand eines Entwicklungsprozesses zur Mustererkennung dargelegt, welche Schritte im Einzelnen notwendig sind. Einer dieser Schritte ist die Feature-Auswahl. Für die speziell in der Audiosignalverarbeitung gängigen Features treffe ich bereits in diesem ausführlichen theoretischen Teil eine Vorauswahl.

Im Kapitel KONZEPT wird auf der Grundlage der bisherigen MSF-Testfahrten ein Konzept erstellt. Dabei wird das Testsetup zur Mustererkennung für eine Tunerqualifizierung beschrieben. In diesem Zusammenhang wird auch erläutert, warum bestimmte Entscheidungen getroffen werden. Außerdem wird gezeigt, wie Test- und Trainingsdaten aufbereitet werden, in welchem Zusammenhang sie zueinander stehen und wie sich daraus die Struktur der Datenbank ergibt.

Die Bedeutung der vorangestellten theoretischen Ausführungen wird im Kapitel IMPLEMENTIERTE ALGORITHMEN deutlich, da entsprechende Prinzipien die Grundlage für das hier entwickelte Klassifikationsverfahren zur Tunerqualifizierung darstellen.

Die Ergebnisse und deren Interpretation sind im Kapitel TEST DES KLASSIFIKATORS zu lesen. Hier wird gezeigt, wo die Stärken und Schwächen des Klassifikators liegen, und ausgeführt, welche Erkenntnisse aus den Ergebnissen zu gewinnen sind.

Im abschließenden 6. Kapitel steht neben einer ZUSAMMENFASSUNG der gewonnenen Erkenntnisse auch ein AUSBLICK auf mögliche Erweiterungen der Mustererkennung. Dabei wird aufgezeigt, wo die Anknüpfungspunkte für vertiefenden und weiterführenden Arbeiten zur objektiven Tunerqualifizierung liegen.

**Anmerkungen** Zur Auswertung und Implementierung des Klassifikationsverfahrens dient MATLAB<sup>®</sup> 2007a. Zum besseren Verständnis der vorliegenden Arbeit findet sich im Anhang eine zusammenfassende Erklärung der UKW-Rundfunktechnik. [Anhang B.1: Einführung UKW-Rundfunktechnik]

## 2 Mustererkennung

Klassifikationsverfahren sind Methoden und Kriterien zur Einteilung von Objekten in Klassen. Da Klassifikationsverfahren immer anwendungsbezogen sind, gibt es viele verschiedene Methoden. Bei den zu klassifizierenden Objekten kann es sich grundsätzlich um jede Art von Messgröße handeln. In der Wissenschaft spielen Klassifikationsverfahren u.a. bei der „Mustererkennung“ (Pattern Recognition) eine Rolle: Die „Datenobjekte“, die so genannten Patterns (Muster), werden entsprechenden Klassen zugeordnet, also klassifiziert. Von maschineller oder automatischer Klassifikation spricht man dann, wenn sich ein solches Verfahren als Algorithmus implementieren lässt. Die immense Bedeutung dieser automatischen Mustererkennung wird deutlich, wenn man sich deren breites Anwendungsspektrum vor Augen hält: Sprach- und Schrifterkennung (OCR), Gesichtserkennung, Identifikation von Fingerabdrücken, optische Zeichenerkennung, DNA-Reihen-Identifikation oder EEG und EKG sind nur einige der zahllosen Anwendungsgebiete.

Gerade weil es so viele Verfahren der automatischen Mustererkennung gibt, soll durch folgende Ausführungen klar werden, warum die Auswahl zur Algorithmusentwicklung auf ein bestimmtes Verfahren fiel.

### 2.1 Klassifikationsverfahren

Eine streng hierarchische Einteilung von Klassifikationsverfahren ist kaum möglich. Am besten lassen sie sich anhand verschiedener Eigenschaften einteilen:

- Manuelle und automatische Verfahren
- Statistische und verteilungsfreie Verfahren
- Parametrische und nicht-parametrische Verfahren
- Unüberwachte und Überwachte Verfahren

### **Manuelle und automatische Verfahren**

Die manuelle Klassifikation - von Experten durchgeführt - ist sehr präzise. Gleichzeitig ist sie aber auch mit großem zeitlichen und entsprechend hohem Kostenaufwand verbunden.

Bei automatischen Verfahren erfolgt der Klassifizierungsprozess durch entsprechende Software. Dieser Prozess kann als formale Methode des Entscheidens in neuen Situationen aufgrund erlernter Strukturen bezeichnet werden. Insofern ist die maschinelle Klassifikation ein Teilgebiet des maschinellen Lernens. [Bishop]

Im Speziellen handelt es sich bei der automatischen Klassifikation um die Erzeugung eines Algorithmus (lernender Algorithmus), der auf der Basis intellektuell klassifizierter Trainingsdaten neue Strukturen berechnet. Die so neu „erlernten“ Strukturen ermöglichen es einem weiteren Algorithmus (auswertender Algorithmus), neue, bisher unbekannte Fälle aufgrund der beobachteten Merkmale (Features) und deren Ausprägungen einer der bekannten Ziel-Klassen zuzuordnen.

### **Statistische und verteilungsfreie Verfahren**

Statistische Verfahren basieren auf Verteilungs- und Wahrscheinlichkeitsdichteberechnungen, während sich bei verteilungsfreien Verfahren klare Trennflächen zwischen den Klassen angeben lassen.

### **Parametrische und nicht-parametrische Verfahren**

Parametrische Verfahren beruhen auf parametrischen Wahrscheinlichkeitsdichten, während bei nicht-parametrischen Verfahren (z.B. Next-Neighbour- bzw. Nearest-Neighbour-Klassifikation) die Dichte meist aus einer Stichprobe geschätzt wird. [Bishop]

### **Unüberwachte und Überwachte Verfahren**

Bei unüberwachtem Lernen (unsupervised learning) gibt es keine vorgegebenen Klassen. Der Algorithmus versucht, die Trainingsdaten auf eine „natürliche“ Art zu gruppieren. Dabei führen unterschiedliche Klassifikationsalgorithmen oft zu unterschiedlichen Ergebnissen. Die Anzahl der Klassen kann vorgegeben sein oder vom Algorithmus selbst in eine von ihm für notwendig erachtete Anzahl gebracht werden. Hier wird klar, dass es eine hochkomplexe Aufgabe ist, einen effizienten Algorithmus zu entwerfen. Dabei ist es u.a. notwendig, die gesamte Aufgabe in kleine Teilgebiete zu isolieren, um dann die gewonnenen Ergebnisse auszuwerten und schließlich besser analysieren zu können.

Bei der überwachten Klassifikation (supervised learning) sind die Klassen vorgegeben. Der Klassifikator wird durch „lernen“ auf diese Klassen eingestellt. Die Trainingsdaten sind dabei Datenvektoren mit bekannter Klasseninformation. Aus einer Vielzahl an Messungen erstellt man eine Datenbank, wobei eine Unterteilung in Trainingsdaten und in Testdaten erfolgt. Für solch ein Klassifikationsverfahren ist folgende Vorgehensweise üblich [FrPer]:

- Trainingsdatensatz erstellen
- Trainingsdatensatz analysieren
- Klassifikator auswählen
- Klassifikator trainieren
- Genauigkeitsabschätzung (Anzahl der richtig klassifizierten Objekte ermitteln)
- Klassifikator ist einsatzbereit

Das Trainieren des Klassifikators, wird bei der überwachten Klassifikation meist von einer oder mehreren Testpersonen durchgeführt. Diese können sowohl Experten als auch eine Gruppe von Personen eines empirischen Tests sein. Bei dieser Art von Training werden die Daten den vorher definierten Klassen zugeordnet. Die Zuordnung geschieht beispielsweise in Zeitframes, die später den entsprechenden Feature-Vektoren zugeordnet werden. Dieser Vorgang wird allgemein als LABELUNG der Daten bezeichnet.

1. **Nearest-Neighbor-Klassifikator** Bei dem NN-Klassifikator wird der Testpunkt jener Klasse zugeordnet, die den Trainingspunkt beinhaltet, zu dem der geringste Abstand gefunden wird. Das Beispiel in Abbildung 2.2 zeigt die NN-Klassifizierung anhand eines zweidimensionalen Raumes zu einem Punkt des roten Clusters. Die Zuordnung erfolgt beispielsweise über die euklidischen Distanz.

Im allgemeinen Fall des  $n$ -dimensionalen Raumes  $R^n$  ist der euklidische Abstand definiert durch:

$$d(x, y) = \|\mathbf{x} - \mathbf{y}\|_2 = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} = \sqrt{\sum_i^n (x_i - y_i)^2} \quad (2.1)$$

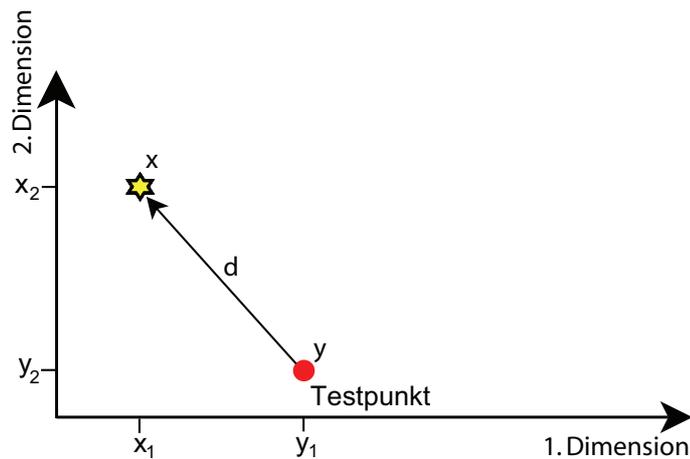


Abbildung 2.1: Euklidische Distanz in einem 2-D-Raum

Die beiden Punkte  $\mathbf{x}$  und  $\mathbf{y}$  sind durch folgende Koordinaten festgelegt:

$$\mathbf{x} = (x_1, x_2, \dots, x_n) \quad \text{und} \quad \mathbf{y} = (y_1, y_2, \dots, y_n)$$

Zur besseren Veranschaulichung ist der euklidische Abstand in Abbildung 2.1 am Beispiel eines zweidimensionalen Raumes dargestellt.

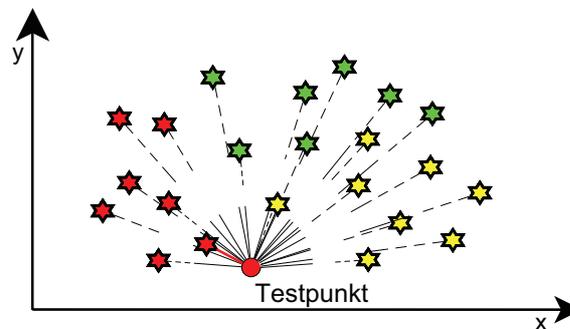


Abbildung 2.2: NN-Klassifikation in einem 2-D-Feature-Raum

### Vorgehensweise

Gehen wir davon aus, dass  $D_n = \mathbf{x}_1, \dots, \mathbf{x}_n$  einer Menge von Feature-Vektoren entspricht, die den gelabelten Daten zugeordnet wurden. Dann ist die Klassifikationsaufgabe für den NN-Klassifikator, den Testpunkt  $\mathbf{y}$  jener Klasse zuzuordnen, die jenen Feature-Vektor  $\mathbf{x}' \in D_n$  beinhaltet, zu dem der Testpunkt  $\mathbf{y}$  am nächsten liegt. Dies ist eine suboptimale Operation. Allerdings kann mit einer großen Trainingsdatenbank ein sehr gutes Ergebnis erzielt werden.

Der Vorteil eines NN-Klassifikators liegt in einer kurzen Erstellungsphase der Trainingsbasis. Dabei ist es nur mehr notwendig, die Feature-Vektoren den gelabelten Daten entsprechend zu sortieren. Die einmal erstellte Trainingsbasis kann dann für das Klassifikationsverfahren gespeichert werden. Da bei einem NN-Klassifikator die Testpunkte mit allen Trainingsdaten im Vektorraum verglichen werden müssen, ist die Klassifikation sehr rechenaufwendig. [Duda07]

2. **k-Nearest-Neighbor-Klassifikator** Den k-NN-Algorithmus können wir uns so folgendermaßen vorstellen: Aus den Trainingsvektoren wird der k-NN ermittelt; k wird ungerade gewählt. Aus den k-Samples wird die Anzahl der Vektoren  $k_i$  bestimmt, die den Klassen  $\omega_i$ , (mit  $i = 1, 2, \dots, M$ ) zugehören, die in diesem Feld liegen. Dabei gilt:  $\sum k_i = k$ .

Anschließend wird  $x$  der Klasse  $\omega_i$  mit dem Maximum an  $k_i$ -Auftreten zugeordnet. Dabei können verschiedene Distanzberechnungen angewandt werden, beispielsweise die euklidische Distanz. [TheoKon99]

3. **k-means** Der k-means-Algorithmus ist ein Verfahren, um ähnliche Objekte innerhalb einer Menge von Objekten zu k-Gruppen zusammenzufassen. Dieser Algorithmus, bei dem es sich um eine Implementierung eines EM-Algorithmus handelt, zeichnet sich durch seine Einfachheit aus. k-means-Algorithmen bestehen aber mehr durch Rechengeschwindigkeit als durch ausgesprochen präzise Ergebnisse.

Vor der Ausführung eines k-means-Algorithmus muss die Anzahl k der zu ermittelnden Gruppen festgelegt werden. Außerdem wird ein Hilfsalgorithmus benötigt, der den Schwerpunkt (das Zentrum) mehrerer Objekte ermittelt. Der k-means-Algorithmus selbst besteht aus fünf Schritten [FrPer]:

- (a) Bestimmung von k Cluster-Schwerpunkten.
- (b) Jedes Objekt wird jenem Cluster zugeordnet, dessen Schwerpunkt ihm am nächsten liegt.
- (c) Für jede Klasse wird der Schwerpunkt neu berechnet.
- (d) Basierend auf den neu berechneten Zentren werden die Objekte wieder wie in Schritt b auf die Cluster verteilt.
- (e) Wurde bei der Neuverteilung mindestens ein Objekt einem anderen Cluster zugeordnet, wiederhole man den Algorithmus beginnend mit Schritt c.

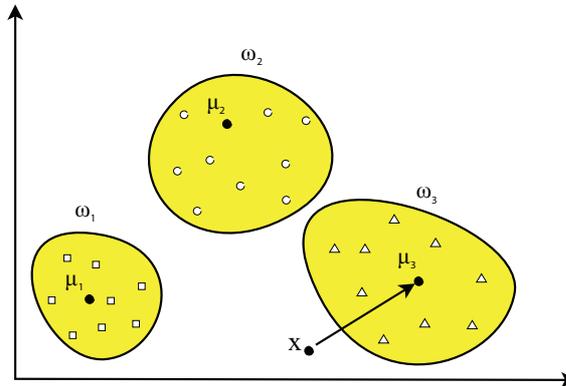


Abbildung 2.3: k-means Klassifizierung im 2-D Featureerraum

## 2.2 Der Entwicklungsprozess

Der Entwicklungsprozess eines Mustererkennungssystems schließt eine gewisse Anzahl von wiederholten Prozessen mit ein. Gehen wir von einem nicht-parametrischen Verfahren aus, so kann die gesamte Aufgabe in drei Phasen unterteilt werden, nämlich Sammlung, Vorverarbeitung und Klassifikation der Daten (siehe Abbildung 2.4).

In der ersten Phase werden die Daten akquiriert und über einen AD-Umsetzer in ein geeignetes Format für die spätere Weiterverarbeitung im Computer konvertiert. Dieser Entwicklungsschritt ist klar als Prozess anzusehen, der, je nach Anwendung, einige Stunden bis mehrere Jahre lang dauern kann.

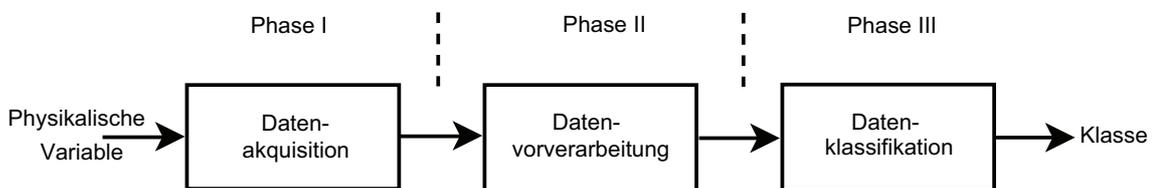


Abbildung 2.4: Prozess einer Mustererkennung

In der zweiten Phase (Vorverarbeitung) werden die gemessenen Daten vorverarbeitet, worauf die Feature-Extraktion folgt. Dann werden die Daten in Gruppen von charakteristischen Merkmalen für das Klassifikationsverfahren zusammengefasst.

Die dritte Phase stellt den eigentlichen Klassifikationsprozess dar, der in Form von Entscheidungsfunktionen ausgeführt wird. Mit einem Set von Features kann das „Muster“ erkannt und klassifiziert werden. Abbildung 2.5 ist eine schematische Darstellung einer Datenanalyse mit den drei Phasen. An den Punkten B, C, und D befinden wir uns im Pattern-Raum, im Feature-Raum bzw. im Klassifikationsraum. [ThuMid85]

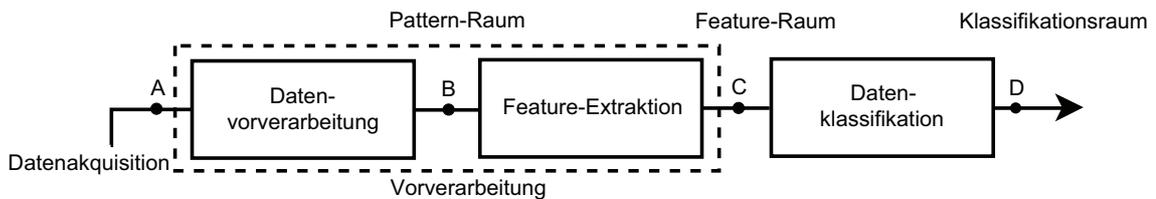


Abbildung 2.5: Vorverarbeitung zur Datenanalyse mit Feature-Extraktion

Im Folgenden werden die einzelnen Prozesse detaillierter beschrieben. [Duda]

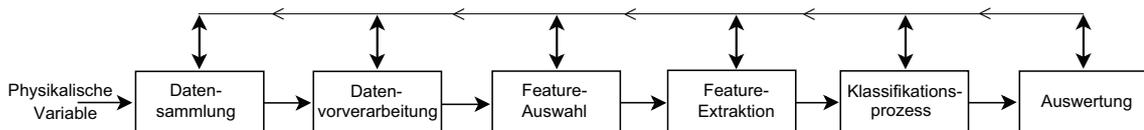


Abbildung 2.6: Entwicklungsprozess der Mustererkennung [TheoKon99]

### 2.2.1 Datensammlung

Die Datensammlung kann einen großen Teil der Kosten bei der Entwicklung eines Mustererkennungssystems ausmachen. Deshalb ist es ratsam, im Vorfeld eine Machbarkeitsstudie mit einer kleinen Datenmenge an „typischen“ Beispielen zu erstellen. Anschließend werden Effizienz und Zuverlässigkeit des Systems überprüft. Aus der Voruntersuchung ergibt sich ebenfalls die Größe der Datenmenge, die für eine aussagekräftige Analysen notwendig ist.

### 2.2.2 Datenvorverarbeitung

In diesem Schritt werden die Daten vereinfacht dargestellt, ohne dass dabei relevante Information verloren gehen. Diese bearbeiteten Daten werden dann an den „Feature-Extraktor“ geleitet. Die Vorverarbeitung kann vor allem dazu dienen, Störeinflüsse zu entfernen.

### 2.2.3 Feature-Auswahl

Die Auswahl von charakteristischen Features ist ein sehr kritischer Entwicklungsschritt und hängt sehr stark vom Einsatzgebiet der Mustererkennung ab. Von entsprechend großer Bedeutung ist das Vorwissen darüber. Grundsätzlich versucht man Features zu finden, die einfach zu extrahieren, invariant zu irrelevanten Transformationen und unempfindlich gegenüber Störeinflüssen sind. Da eine optimale Lösung nicht berechenbar ist, erfolgt die Auswahl der Features in der Regel basierend auf Erfahrungswerten bzw. nach heuristischen Methoden (trial and error).

### 2.2.4 Feature-Extraktion

Bei der Feature-Extraktion werden nur mehr Eigenschaften der Signale entnommen, wodurch auch eine Reduzierung der Datenmenge erfolgt. Ausschlaggebend für die Güte eines Klassifikators ist die Güte der Features, die zum Erkennen der Muster notwendig sind. Zuallererst ist es notwendig, messbare Größen zu finden, die das System beschreiben. Diese Größen sollen natürlich unempfindlich gegenüber äußeren Störeinflüssen sein und es ermöglichen, die unterschiedlichen Muster gut zu trennen. Eine detaillierte Aufstellung mit einer Auswahl an gängigen Features der Audiotechnik folgt in Abschnitt 2.3.

### 2.2.5 Klassifikationsprozess

Die Features, genauer gesagt deren Werte, werden im eigentlichen Klassifikationsprozess den vorher festgelegten Klassen zugeordnet. Hier wird die endgültige Entscheidung darüber getroffen, zu welcher Klasse das jeweilige Feature gehört. Dieser Prozess wird hier auch als TRAINING des Klassifikators bezeichnet. Wie bereits in Abschnitt 2.1 beschrieben, gibt es unterschiedlichste Klassifikationsverfahren. Es hat sich aber im Laufe der vergangenen Jahre herausgestellt, dass ein Training mit Trainingssamples eine der effizientesten Methoden ist.

### 2.2.6 Auswertung

Es ist essentiell, die Ergebnisse des Klassifikators auszuwerten, um eine Genauigkeitsabschätzung zu erhalten. So können Features auf Ihre Güte hin getestet, anschließend verbessert oder gegebenenfalls neue gesucht werden. Es ist also zwingend notwendig, geeignete Analysemethoden zu finden, um einerseits die Performance des

Klassifikators zu verbessern, andererseits aber auch, um vorhandene Schwachpunkte ausfindig zu machen.

### Testen des Klassifikationsverfahrens

Zum Testen des Klassifikators wird meist eine so genannte CONFUSION-MATRIX herangezogen. Mit dieser kann man die richtig erkannten Klassen, die falsch erkannten sowie die nicht erkannten Klassen auf „einen Blick“ analysieren und einfach auswerten. Die richtig erkannten Klassen sind in der Hauptdiagonale angeordnet, falsch erkannte und vertauschte Klassen sind außerhalb der Hauptdiagonale wiederzufinden.

	IST - KLASSEN					
S						
O						
L						
K						
L						
A						
S						
S						
E						
N						

Abbildung 2.7: Confusion-Marix mit Soll-Klassen aus Labelung und Ist-Klassen aus Klassifikation

## 2.3 Audio-Features

Zahlreiche Features zur Erkennung und Beschreibung von Audiosignalen kommen aus dem Bereich der Spracherkennung sowie aus der Klassifikation und der automatischen Identifikation von Musikinstrumenten. [BreMcKi01]

Features können nach folgenden Kriterien eingeteilt werden:

1. **Dynamik und Stationarität** Dynamische Features werden über einen längeren Zeitraum gesammelt. Im Gegensatz dazu stehen stationäre Features, die einen Wert über eine längere Zeitperiode zusammenfassen (beispielsweise Mittelwert, Standardabweichung, Ableitungen u.Ä.).
2. **Die Erstreckung über die Zeit**, in der das Feature gemessen wird. So kann beispielsweise nur der Einschwingvorgang eines Signal als Feature dienen, oder aber das gesamte Signal.

Features lassen sich aber auch dadurch unterscheiden, in welcher Domäne sie gebildet werden. [GePe]

- **Differenz zweier Signale im Zeitbereich** Voraussetzung dafür ist eine ausreichend genaue Phasenkohärenz zwischen den beiden Signalen. Erst dann kann Sample für Sample subtrahiert werden. Wenn die erforderliche Phasenkohärenz im Zeitbereich fehlt, kann die Subtraktion im Frequenzbereich erfolgen. Dabei wird nur das Amplitudenspektrum der beiden Signale subtrahiert (siehe auch Fast-Fourier-Transformation).

$$d(k) = x(k) - y(k) \tag{2.2}$$

$x(k)$  ... Testsignal

$y(k)$  ... Referenzsignal

- **Zeitliches Feature** Dafür können Autokorrelationskoeffizienten und die Rate der Nulldurchgänge (zero-crossing rate) bestimmt werden.
- **Energie Feature** Die Energie des Signals wird aus unterschiedlichen Zeitabschnitten berechnet (globale-, harmonische-, Rauschenergie).

Neben der Beschreibung eines Signals im Verlauf der Zeit, besteht auch die Möglichkeit einer Beschreibung des Signals durch sein Spektrum. Man spricht im Zusammenhang mit der Signalbeschreibung auch vom Frequenzinhalt eines Signals. Das bedeutet, ein Signal ist aus zahlreichen harmonischen Signalen verschiedener Frequenzen additiv zusammengesetzt (Fourier-Synthese). Umgekehrt bedeutet Frequenzanalyse (Spektralanalyse) das Zerlegen allgemeiner Signale in harmonische Grundkomponenten. [OS01]

- **Diskrete Fourier-Transformation**

Die gebräuchlichsten Spektralanalysemethoden benutzen die diskrete Fourier-Transformation. Gegenüber der klassischen Fourier-Transformation, die kontinuierliche Funktionen im Zeit- und Frequenzbereich miteinander verknüpft, wird entsprechend den Gleichungen 2.3 und 2.4, eine Verknüpfung von diskreten Folgen im Zeit- und Frequenzbereich als diskrete Fourier-Transformation (DFT) bezeichnet.

DFT-Transformationspaar:

$$x(k) = \frac{1}{N} \sum_{n=0}^{N-1} \underline{X}(n) \cdot e^{j(2\pi/N)n} \quad \text{für } k = 0, 1, 2, \dots, N-1 \quad (2.3)$$

$$\underline{X}(n) = \sum_{k=0}^{N-1} x(k) \cdot e^{-j(2\pi/N)k} \quad \text{für } n = 0, 1, 2, \dots, N-1 \quad (2.4)$$

$x(k)$  ..... Zeitdiskretes Signal  
 $\underline{X}(n)$  ..... Diskretes Spektrum  
 $e^{j(2\pi/N)n}$  ... komplexer Drehoperator

Das Spektrum  $X(n)$  lässt sich auch nach Betrag und Phase darstellen:

$$X(n) = |X(n)| \cdot e^{j \cdot \arg(X(n))} \quad (2.5)$$

$|X(n)|$  ..... Betrag des Spektrums  $X(n)$   
 $e^{j \cdot \arg(X(n))}$  ... Phase des Spektrums  $X(n)$

- **Kurzzeit-Fourier-Transformation**

Die Kurzzeit-Fourier-Transformation (STFT<sup>1</sup>) behebt ein Problem der Fouriertransformation, die keine Aussage über das zeitliche Auftreten einzelner Frequenzen oder Frequenzbereiche macht. Bei der Kurzzeit-Fourier-Transformation wird nur ein zeitlich begrenzter Abschnitt des Signals betrachtet. Dazu multipliziert man das Signal mit einer Fensterfunktion  $w$  und berechnet für das ausgeschnittene Signal die (klassischen) Fourier-Koeffizienten.

Wird das Signal „scharf“ gefenstert, z.B. durch eine Rechteckfunktion, ergeben sich Probleme an den Kanten jener Ausschnitte, an denen eine abrupte Änderung der Amplitude im Zeitbereich auftritt, so dass ein breites Spektrum im Frequenzbereich entsteht. Dieses kann durch die Wahl eines Zeitfensters (z.B. Hann-, Hamm-, Blackmannfenster) vermieden werden.

- **Harmonisches Feature** Harmonische Signale können am besten anhand ihrer Grundfrequenz beschrieben werden. Zur Berechnung der Grundfrequenz kann der Maximum-Likelihood-Algorithmus herangezogen werden. Zudem kann bei solchen Signalen immer die Energie der Harmonischen bzw. deren Verhältnis zueinander berechnet werden, beispielsweise das Verhältnis der geraden Harmonischen zu den ungeraden.
- **Spectral Flatness/Crest measure** Spectral Flatness ist ein Maß zur Charakterisierung des Spektrums. Ein hoher Wert gibt eine gleichverteilte Energie in allen Bändern an. Ein niedriger Wert hingegen charakterisiert ein Spektrum in dem die Energie nur in einigen Bändern liegt. Berechnet wird dieses Maß aus dem Verhältnis von geometrischem Mittelwert zu arithmetischem Mittelwert. Zusätzlich zu diesem Maß kann ein Crest measure berechnet werden. Dieses berechnet sich aus dem Verhältnis des maximalen Energiewertes eines Bandes zum arithmetischem Mittelwert aller Bänder.
- **Spektrale Gestalt des Signals** Die nach einer Kurzzeit-Fourieranalyse (STFT) ermittelte Energie bzw. die daraus gewonnenen MFC-Koeffizienten beschreiben die spektrale Gestalt des Signals.

---

<sup>1</sup>STFT: Short-Time-Fourier-Transformation

- **Mel-Skala** Um ein Audiosignal in der spektralen Domäne weiterverarbeiten zu können, ist es meist sinnvoll, die einzelnen Frequenzbins eines Signals in Frequenzbänder zusammenzufassen. Eine derartige Verarbeitung ist eine Anlehnung an die menschliche Wahrnehmung durch eine wahrnehmungsbezogene Gewichtung (perceptual weighting). Hierfür werden Bandpässe verwendet. Die Bandmittenfrequenzen  $b(m)$  der Mel-Filterbänke sind dabei nicht linear über das Frequenzspektrum verteilt. Auch die Breite der Filter ist abhängig von der Frequenz. Der Frequenzbereich unter 1 kHz wird von einem größeren Teil der Filterbänke abgedeckt. Die Bandmittenfrequenzen lassen sich mit

$$b(m) = 2595 \log \left( \frac{f}{700} + 1 \right) \quad (2.6)$$

berechnen.

Die Bandpässe sind Dreiecksfenster, die durch folgenden Zusammenhang beschrieben werden können. [Mel02]

$$H(k, m) = \begin{cases} 0 & \text{für } f(k) < f_c(m-1) \\ \frac{f(k)-f_c(m-1)}{f_c(m)-f_c(m-1)} & \text{für } f_c(m-1) \leq f(k) < f_c(m) \\ \frac{f_c(m+1)-f(k)}{f_c(m+1)-f_c(m)} & \text{für } f_c(m) \leq f(k) < f_c(m+1) \\ 0 & \text{für } f(k) \geq f_c(m+1) \end{cases} \quad (2.7)$$

$f_c$  ... Mittenfrequenz des zu berechnenden Bandes

Die wahrnehmungsgewichteten Werte an den Ausgängen dieser Mel-skalierten Filter lassen sich durch folgende Gleichung angeben [Bec99,ST95]

$$X'(m) = \sum_{n=0}^{N-1} |X(n)| \cdot H(n, m) \quad (2.8)$$

$X'(m)$  ..... Mel-gewichtetes Signal  $X(n)$

$|X(n)|$  ..... Betrag des diskreten Signals

$H(n, m)$  ... Mel-Filterbank

Im Frequenzbereich steht für die Feature-Extraktion unter anderem noch folgende Möglichkeit zur Verfügung:

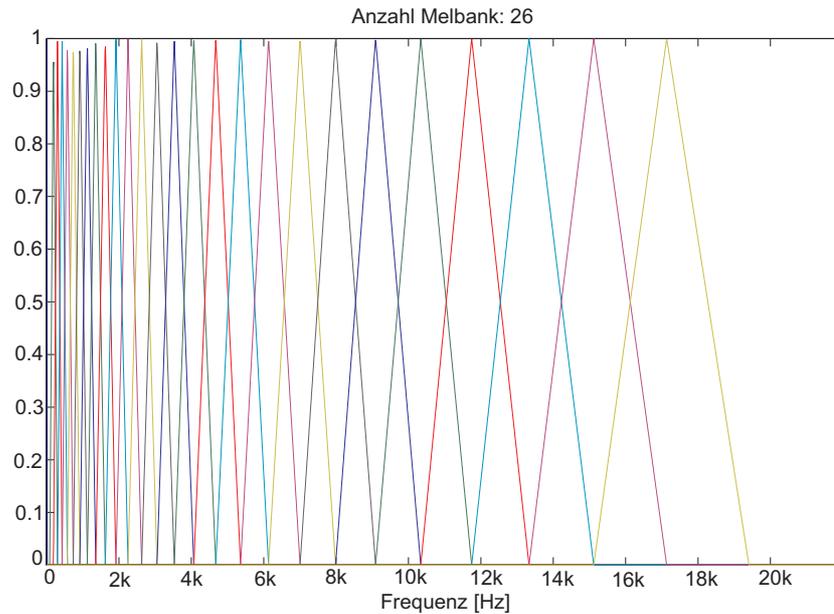


Abbildung 2.8: Melskala mit 26 Frequenzbändern

- **Differenzbildung zweier Signale im Frequenzbereich** Zunächst werden die Signale  $x(k)$  und  $y(k)$  mittels Short-Time-Fourier-Transformation in den Frequenzbereich gebracht. Anschließend werden die transformierten Signale  $X(n)$  und  $Y(n)$  mittels eines Filters (z.B. Terzbandfilter, Mel-Filter o.Ä.) in Frequenzbänder zusammengefasst.

$$X'(m) = \sum_{n=0}^{N-1} |X(n)| \cdot H(n, m) \quad (2.9)$$

$X'(m)$  ... Mel-gewichtetes Signal  $X(n)$

$|X(n)|$  ... Betrag des Diskreten Testsignals  $X(n)$

$$Y'(m) = \sum_{n=0}^{N-1} |Y(n)| \cdot H(n, m) \quad (2.10)$$

$Y'(m)$  ... Mel-gewichtetes Signal  $Y(n)$

$|Y(n)|$  ... Betrag des Diskreten Referenzsignals  $Y(n)$

Nun werden die erhaltenen Energien der jeweiligen Frequenzbänder voneinander subtrahiert.

$$D(m) = Y'(m) - X'(m) \quad (2.11)$$

$D(m)$  ... Differenz der entsprechenden Melbänder

- **Delta-Feature** Zu jedem der oben beschriebenen Features kann ein so genanntes Delta-Feature (Differenzkoeffizienten) berechnet werden, wodurch Unterschiede in benachbarten Zeitframes mit berücksichtigt werden können. Wie sich die Delta-Features zusammensetzen, ist in Abbildung 2.9 zu sehen. Die Differenz zum vorhergehenden Frame wird berechnet. Das Ergebnis daraus bildet den ersten Teil des neue Delta-Feature. Zudem wird das Frame aus dem ursprünglichen Feature-Vektor zum Delta-Feature hinzugefügt und bildet somit den zweiten Teil. Die Dimension erhöht sich somit um den Faktor zwei.

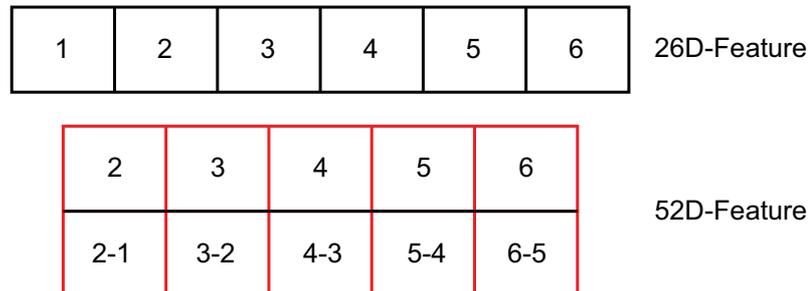


Abbildung 2.9: Zusammensetzung eines 52D-Delta-Feature aus 26D-Feature

Ein nicht zu unterschätzender Vorteil besteht darin, dass die bereits berechneten Features zur Bildung der Delta-Features herangezogen werden können und somit eine Klassifizierung unmittelbar möglich ist.

## 3 Konzept

Wie im einleitenden Kapitel bereits beschrieben wurde, hat Magna Steyr Fahrzeugtechnik (MSF) im Laufe der vergangenen Jahre Prüffahrten durchgeführt, um eine Aussage über die Qualität von KFZ-Tunern treffen zu können. Ausgehend von diesen manuellen Messungen, die rein subjektiven Charakter haben, soll nun ein Messsystem zur objektiven Bewertung von KFZ-Tunern entwickelt werden. Bei einem ersten Meeting aller Projektbeteiligten (MSF, IVM-TC, SPSC) ging es darum, die Kundenwünsche kennenzulernen und die entsprechenden Anforderungen an die zu realisierende Messeinrichtung zu definieren. Dazu wurden die von MSF seinerzeit angewandten Testverfahren begutachtet und die daraus gewonnenen Erkenntnisse diskutiert. Ein Kundenwunsch von grundlegender Bedeutung war es, die Tuner-Qualifizierung entgegen der ursprünglichen Intention nicht an Fahrzeug-Prototypen durchzuführen, sondern ein System zur Verfügung zu haben, das bei Neufahrzeugen universell einsetzbar ist. Das bedeutete, dass ein *nicht* invasives<sup>1</sup> Messsystem entwickelt werden musste, bei dem keine großen mechanische Eingriffe am Fahrzeug vorgenommen werden. Auf der Grundlage der bei diesem Meeting gewonnenen Erkenntnisse und der Analyse der vorangegangenen subjektiven MAGNA-Versuche sowie neuer Testfahrten wurden schließlich die Bedingungen geschaffen, die für die Entwicklung eines objektiven Verfahrens zur Bewertung eines KFZ-Tuners notwendig sind.

Ein erster Schritt auf dem Weg zur Entwicklung eines objektiven Messsystems bestand also in der eingehenden Analyse der bisherigen subjektiven Bewertung bei den MSF-Prüffahrten.

Bei aller Subjektivität dieser Testfahrten haben die MAGNA-Versuche und die folgenden Besprechungen rasch klar gemacht, dass es bei den Messungen auf die Häufigkeit sowie auf die Dauer der auftretenden Störungen ankommt. Das zu

---

<sup>1</sup>invasiv: keine mechanischen Eingriffe am KFZ erforderlich

entwickelnde objektive Messsystem muss also in der Lage sein, Störungen zu detektieren und anschließend zu klassifizieren.

Ein weiterer Anspruch an das zu realisierende System ist seine Adaptierbarkeit für diverse Sendeformate, da nicht nur im europäischen Raum getestet wird. Im Hinblick auf absehbare technische Entwicklungen ist es natürlich auch unerlässlich, dass das Produkt für digitales Radio (DAB-T, HD-Radio) erweiterbar ist.

### 3.1 Aufzeichnungsmethoden zur Datenerfassung

Nun galt es, ein geeignetes Verfahren zur Aufzeichnung der Testfahrten zu finden, wobei folgende Optionen zur Auswahl standen:

#### 1. Line-out am Tuner

Vorteile:

- optimales Messsignal des Tuners
- kein Eingriff am Fahrzeug notwendig

Nachteil:

- nicht an jedem KFZ-Tuner vorhanden

Bei neuen Fahrzeugen könnte man theoretisch zwar ein Line-out Signal über den CAN-Bus<sup>2</sup> bereitstellen; eine solche Lösung stellte sich in der Diskussion mit MSF-Mitarbeitern aber als kaum praktikabel heraus.

#### 2. Lautsprecher-Ausgang des KFZ-Tuners

Vorteile:

- kein Eingriff am Tuner notwendig
- der selbe Messpunkt für alle Fahrzeuge

Nachteile:

- Signalverfälschung durch Messwandler und Verstärker
- mechanischer Eingriff am Fahrzeug notwendig

---

<sup>2</sup>CAN-Bus: Controller Area Network - asynchrones, serielles Bussystem

### 3. Mikrofon im Fahrzeuginnenraum

Vorteil:

- kein Eingriff in das Fahrzeug notwendig

Nachteile:

- kleiner SNR durch Fahrzeug- und Umgebungsgeräusche
- adaptive Filter zum Extrahieren des Radio-Signals notwendig
- Messung der gesamten Signalkette inkl. Lautsprecher

Da ein universell einsetzbares Testsystem gesucht ist, wird der Lautsprecher-Ausgang des Tuners als Schnittstelle für die Aufzeichnung des Messsignals gewählt. So ist sichergestellt, dass für *alle* Testgeräte das selbe Setup verwendet werden kann und immer exakt derselbe Messpunkt gegeben ist – eine Voraussetzung für die objektive Bewertung.

Diese Variante stellt zudem eine praktikable Lösung dar, da es dabei nur eines kleinen Eingriffs in das KFZ bedarf. Der Aufnahmepegel kann mit Hilfe eines Messwandlers einfach an das Setup angepasst werden.

Damit ist einerseits eine vom Audio-Ausgang des Tuners unabhängige Messung möglich. Noch viel wichtiger aber ist es, dass auf diese Weise die Unabhängigkeit vom Sendeformat gewährleistet ist, was hinsichtlich der Erweiterbarkeit des Messsystems für künftige Sendeformate eine entscheidende Rolle spielt.

## 3.2 Testsetup

Nun galt es, ein geeignetes Testsetup zu entwerfen. Für die objektive Bewertung eines KFZ-Tuners wurde es für notwendig erachtet, mit einem Referenzsignal zu arbeiten. Aus der Differenz von Test- und Referenzsignal wird nämlich ein Fehler-signal gebildet, das dann als Ausgangsmaterial für die Feature-Extraktion dient, wie es in Abbildung 3.1 zu sehen ist. Bei der Wahl des „objektiven Maßes“ entschied man sich für ein stationäres Referenzsignal. Diese Referenzstation wird im Idealfall nahe der Teststrecke aufgebaut und steht im direkten Blickwinkel zum HF-Sender. Auf die Anforderungen der Referenzstation wird im Abschnitt 3.2.1.2 noch näher eingegangen.

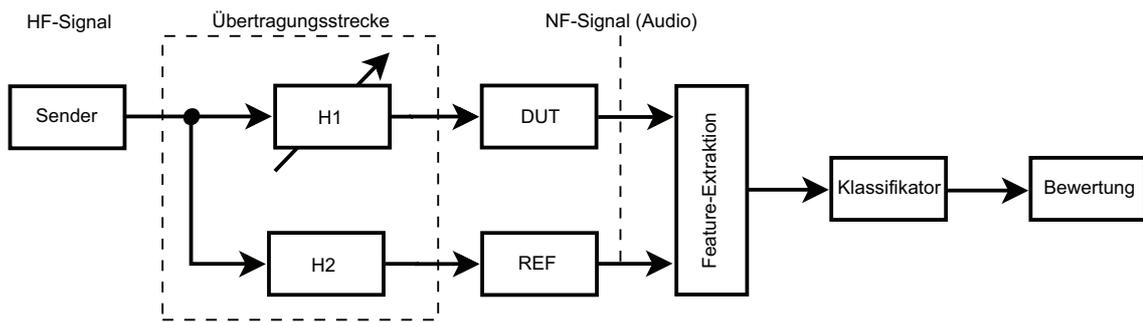


Abbildung 3.1: Testsetup mit HF-Sender und Übertragungsstrecke

Um einen geeigneten Klassifikator zu entwickeln, war eine Datenerhebung notwendig. Testfahrten mit einer Signalaufzeichnung waren erforderlich.

Wie es typischerweise bei technischen Entwicklungen der Fall ist, wurde in einer Machbarkeitsstudie eine Auswahl an repräsentativen Signalen aufgezeichnet. In dieser Voruntersuchung ging es aber nicht nur um die Aufzeichnung einer typischen Auswahl an Signalen; die Studie sollte vor allem auch der Entscheidungsfindung darüber dienen, mit welchem Aufnahmesetup in Zukunft gearbeitet wird. Dem ursprünglichen Konzept entsprechend, erfolgte die Aufzeichnung der Test- wie der Referenzsignale über einen direkten Anschluss am Tuner, so, wie es in den Abbildungen 3.2 und 3.3 dargestellt ist.

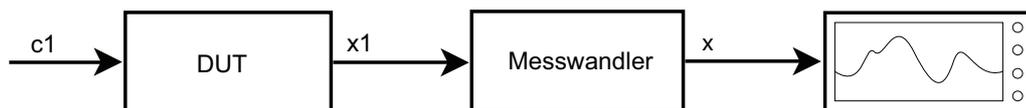


Abbildung 3.2: Testsignal - Device under Test (DUT)

- c1 ... HF-Signal am mobilen Empfänger
- x1 ... NF-Ausgang am Lautsprecherausgang des Testtuner
- x ... NF-Signal zur Aufnahme

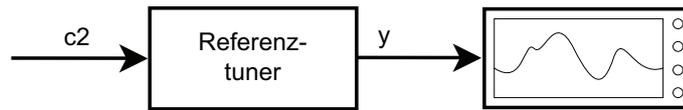


Abbildung 3.3: Referenzsignal aus Referenztuner

c2 ... HF-Signal am stationären Empfänger

y ... NF-Ausgang (Line-out) des Referenz-tuners

### 3.2.1 Voruntersuchung

Für die Aufzeichnungen stellte MSF ein Industrie-Laptop zur Verfügung. Die AD-Wandlung erfolgte über ein MOTU-Ultralite-Audiointerface; die Stromversorgung für Audiointerface und Tuner wurde über eine externe („erdfreie“) Autobatterie sichergestellt, um von äußeren Einflüssen entkoppelt und gegen entsprechende Störungen abgeschirmt zu sein. Auf diese Weise war eine dreistündige Aufzeichnung auf diversen Teststrecken ohne Unterbrechungen möglich. Getestet wurde mit einem eigens für diesen Zweck umgebauten BMW-Mini-Tuner.

Als Referenzstation diente ein Industrie-PC. Auch hier wurde die AD-Wandlung mit einem Audiointerface MOTU-Ultralite durchgeführt. Die Stromversorgung erfolgte im Gegensatz zum mobilen Aufnahmesetup über das Stromnetz. Aufgebaut wurde die Referenzstation auf dem Dach eines der Gebäude innerhalb des MAGNA-Geländes, mit direktem Blick zur Sendestation (Schöckel). Als Referenztuner stand ein BMW-Radio zur Verfügung.

Nach ersten Analysen (*Spektralanalyse*) in Adobe Audition 3 stellte sich heraus, dass *alle* Radiosender mit einem Pilotton von 19 kHz arbeiten, um das Stereosignal zu codieren: 19 kHz-Ton = Ja  $\hat{=}$  STEREO; 19 kHz-Ton = Nein  $\hat{=}$  MONO).

Es konnte gezeigt werden, dass beim Auftreten von Störungen sehr viel Energie im Bereich des Pilottons, also zwischen 16 und 21 kHz liegt.

#### Pilotton und Pegelanpassung

Dieser Vorversuch hat ergeben, dass der Pilotton zwar eine Hilfe zur Auswertung sein kann. Bis zur endgültigen Konzeption der Testverfahren aber stellte sich heraus, dass dieser Frequenzbereich für künftige Messsysteme nicht zur Evaluierung verwendet werden kann. Wie nämlich aus Hersteller-Datenblättern sowie aus Quellen über nachrichtentechnische Grundlagen (siehe Anhang B1: UKW-Rundfunk) hervorgeht, wird der Ausgang eines KFZ-Tuners bei einer Frequenz von 15 kHz Tiefpass gefiltert.



Abbildung 3.4: Referenzstation der Voruntersuchung

Nach diesen Erkenntnissen konnte nicht sichergestellt werden, dass der Pilotton zur Pegelanpassung und zur Störungsdetektion zur Verfügung steht. Eine Signalaufnahme mit Pilotton und dem Frequenzbereich bis 21 kHz würde einen Eingriff in die Hardware erfordern und erwies sich deshalb unter Berücksichtigung der Vorgabe als nicht zielführend.

So fiel die Wahl schließlich auf einen, der überwachten Mustererkennung entsprechenden Klassifikator. Da zur Entwicklung eines Klassifikators, der eine Störungsdetektion ermöglicht, eine große Datenmenge erforderlich ist, waren erneute Testfahrten notwendig.

Mit den gesammelten Daten sollte schließlich mindestens ein geeigneter Feature-Vektor gefunden werden, damit

- Störungen automatisch erkannt werden,
- die Störungen den vordefinierten Störklassen zugeordnet werden können,
- eine Aussage über Häufigkeit und Dauer der Störungen getroffen werden kann.

Gleichzeitig sollten die Ergebnisse dieser Arbeit als wissenschaftliche Basis für künftige Arbeiten bis hin zur Ausarbeitung einer qualitativen Bewertungsmetrik für KFZ-Tuner geeignet sein.

## 3.3 Datenerfassung

Um zu einem Aufnahmesetup zu gelangen, das auch für künftige Test- und Evaluierungsfahrten verwendet werden kann, wurden gewisse grundlegende Punkte berücksichtigt, die im folgenden Abschnitt beschrieben werden.

### 3.3.1 Aufnahmesetup

Um Daten für die Klassifizierung und Auswertung bzw. zum Entwurf geeigneter Algorithmen zu gewinnen, war es notwendig, ein praxisnahes Testsetup zusammenzustellen. Dabei wurde besonderes Augenmerk auf Mobilität und Zuverlässigkeit des Messequipments gelegt: einfacher Aufbau, intuitive Bedienung, Robustheit, hohe Speicherkapazität, Stromversorgungsmöglichkeit durch Akkubetrieb. Zudem ist es notwendig, eine qualitativ hochwertige und zuverlässige Referenzstation aufzubauen. Außerdem muss diese im Hinblick auf künftige KFZ-Tunerevaluierungen mobil gehalten werden, um Testfahrten gegebenenfalls auch in Kundennähe durchführen zu können.

Die bei der Voruntersuchung verwendete Ausführung der Referenzstation war zwar zuverlässig, erwies sich nach einigen Einsätzen aber als umständlich: Sie erfüllte *nicht* das Kriterium der Mobilität und Flexibilität. Deshalb wurde in eine flexible, hochqualitative Lösung für das Aufnahmesetup investiert.

#### 3.3.1.1 Mobile Prüfvorrichtung (DUT)

**Device under Test (DUT)** Wie bereits beschrieben, sollte das Testsignal am Lautsprecherausgang abgegriffen werden. Um die mechanischen Eingriffe klein zu halten, wurde für die Signalaufnahme ein eigenes Kabel angefertigt. Mit diesem Y-Kabel ist es möglich, das Lautsprechersignal am Tunerausgang abzugreifen und gleichzeitig den von der Boardelektronik abhängigen Tunerbetrieb aufrechtzuerhalten. Auch das Abschrauben der Lautsprecherabdeckung entfällt, sodass durch diese Lösung insgesamt eine rasche Anwendung in diversen Testfahrzeugen gewährleistet ist.

Für das mobile Setup fiel die Entscheidung auf einen Vier-Kanal-Digitalrecorder (EDIROL R-44) um künftigen Mehrkanalmessungen (verschiedene Tuner in einem Testsetup) zu entsprechen.

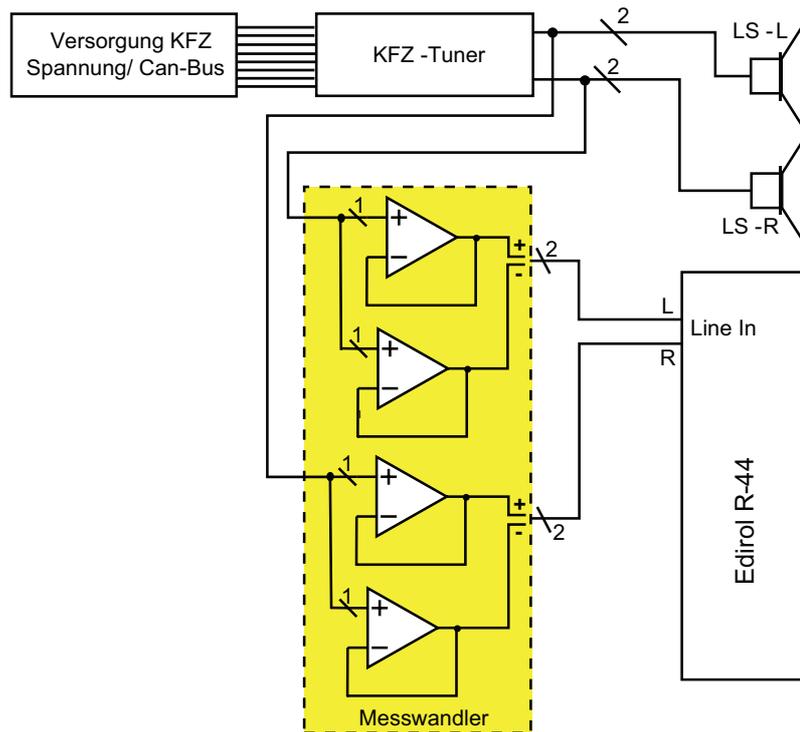


Abbildung 3.5: Elektrischer Verlauf des Messsignals

### 3.3.1.2 Stationäre Referenz

**Referenz tuner** Da bei der Referenzstation eine Zwei-Kanal-Version des Digitalrecorders ausreichend war, bot sich die Zwei-Kanal-Version des selben Herstellers wie beim mobilen Setup an (Edirol R-09HR). Als Referenz tuner standen sowohl KFZ-Tuner als auch Hifi-Tuner zur Auswahl, die Entscheidung fiel jedoch auf einen Hifi-Tuner (ONKYO T-4355). Der Vorteil dabei besteht darin, dass direkt über den Line-out des Tuners aufgenommen werden kann; ein zusätzlicher Messwandler, der das Messergebnis verfälschen könnte, ist damit nicht notwendig. Bereits bei den Voruntersuchungen hatte sich herausgestellt, dass es nicht ganz unproblematisch ist, eine von äußeren Bedingungen (Witterung, örtliche Gegebenheiten u.Ä.) *unabhängige* Referenz zu schaffen. Nach einem Testdurchlauf stand das Setup allerdings perfekt, und mit der von MSF zur Verfügung gestellten Kompetenz und der hochwertigen HF-Technik sowie der bereitgestellten Antenne konnte ein hochqualitatives Referenzsignal aufgezeichnet werden. So konnte beispielsweise sogar eine Klavier-sonate auf Ö1 rauschfrei aufgezeichnet werden, was u.a. auf die hochqualitativen Komponenten der Referenzstation zurückzuführen ist; in der Voruntersuchung mit einem KFZ-Tuner waren derartige störfreie Aufzeichnungen nicht immer möglich.

Eine hohe Empfangsqualität wurde durch die Verwendung einer  $\frac{\lambda}{4}$ -Antenne mit hochwertig abgeschirmtem Kabel sichergestellt (Abbildung 3.6). Um zusätzliche Störeinflüsse von vornherein auszuschließen, wurde der Referenztuner von MSF-Mitarbeitern in der EMV-Kammer auf elektromagnetische Verträglichkeit getestet.

Nun war man im Besitz einer flexiblen Referenzstation, die ohne Problem in der Nähe der Teststrecke aufgebaut werden konnte und auch alle anderen geforderten Ansprüche, vor allem eine hohe Empfangsqualität, erfüllte. Gleichzeitig wird dasselbe Aufzeichnungssystem verwendet wie im Testsetup, was in Bezug auf DA-Wandler und Speichermedium ein Kriterium der Datengüte ist.



Abbildung 3.6:  $\frac{\lambda}{4}$ -Antenne der Referenzstation

### 3.3.2 Messfahrt

Für die Aufnahme der Testdaten wurde die bereits aus der Voruntersuchung bekannte MSF-Teststrecke nahe Graz (siehe Anhang: Teststrecke) gewählt. Auf der Autobahn wurden zusätzliche Aufnahmen gemacht. Die Witterungsbedingungen waren bei sämtlichen Aufnahmen sehr gut (Sonnenschein, teilweise leicht bewölkt und mäßiger Wind). Am Ende der Messfahrt standen insgesamt 250 Minuten an Datenmaterial zur Verfügung.

### 3.3.3 Datenaufbereitung

Die gewonnenen Daten wurden in Trainings- und Testdaten aufgeteilt. Die Trainingsdaten werden zum Trainieren des in Kapitel 4 beschriebenen Algorithmus verwendet. Die Testdaten dienen zur Überprüfung der Güte und Genauigkeit des Klassifikators. Der Datensatz wurde einerseits in inhaltliche Kategorien, also Musik, Sprache, Doku (Dokumentation) und Jingles (Erkennungsmelodie/Text) unterteilt. Zudem wurde der Datensatz nach aufgezeichneten Sendern, also Hitradio Ö3 und Österreich 1 unterteilt.

Zum Trainieren des Algorithmus wird ein Trainingsdatensatz von etwa 75 Prozent der vorhandenen Daten verwendet. Bei der Auswahl ist darauf zu achten, dass ein repräsentativer Querschnitt aus dem vorhandenen Material im Trainingsmaterial enthalten ist. Die restlichen 25 Prozent dienen der Auswertung des Klassifikationsalgorithmus.

Praktisch bedeutet dies, dass aus beiden Sendeprogrammen (Österreich 1 und Hitradio Ö3) der gleiche Anteil an Trainingsdaten vorhanden ist. Entsprechendes gilt hinsichtlich der inhaltlichen Unterteilung: Für die Trainingsdaten steht aus den Bereichen Musik, Sprache, Doku und Jingle jeweils in etwa derselbe Anteil zur Verfügung.

Aus der Tabelle 3.1 ist die genaue inhaltliche Datenaufteilung ersichtlich.

	Trainingsdaten [min]	Testdaten [min]
Musik	115	43
Sprache	52	22
Doku	3	2
Jingle	7	3
<b>Gesamt</b>	<b>177</b>	<b>70</b>

Tabelle 3.1: Aufteilung des gesamten Datenmaterials

Für mögliche künftige Bewertungen der Störklassen (psychakustische Untersuchungen) liegt das Material somit in einer gut strukturierten Datenbank vor.

### 3.3.4 Labelung

Bei der Labelung (Etikettierung) geht es darum, die gesammelten Daten manuell den vorgegebenen Störklassen zuzuweisen. Dazu wurden die aufgezeichneten Signale in Zeitfenster (Frames) je 500 ms unterteilt, wodurch eine genügend feine zeitliche Auflösung sichergestellt ist. Jedes Frame wurde nun dahingehend „markiert“, ob eine Störung vorliegt. Lag eine Störung vor, wurde zusätzlich angegeben, um welche Störung es sich handelt. Dabei standen folgende Auswahlmöglichkeiten zur Verfügung:

- Störfrei
- Knacks
- Granulares Rauschen (Bröseln)
- Verzerrung
- Rauschen
- Aussetzer im Empfang

Mit enthalten ist auch die störfreie Klasse. Wäre ein störfreier Fall nämlich nicht in den Trainingsdaten mit enthalten, würde ein solcher nie erkannt werden. Beim Referenzsignal kann entsprechend den Ausführungen in Abschnitt 3.2.1.2 vorausgesetzt werden, dass es störfrei ist.

Aus der vorgegebenen Unterteilung der Daten ergibt sich eine zusammengehörende Struktur der Datenbank. Der Dateiname setzt sich zusammen aus Sender, Aufzeichnungsdatum, Inhalt und einer fortlaufenden Nummer. Ob es sich um das Test- oder das Referenzsignal handelt, ist an der vorangestellten Erweiterung zu erkennen; an der Dateiendung ist ersichtlich, ob es sich um ein aufgezeichnetes Signal oder um eine gelabelte Datei handelt. An folgendem Beispiel aus der Datenbank ist zu sehen, wie sich ein solches Datentripel im konkreten Fall zusammensetzt.

- Ref\_Sender\_Aufzeichnungsdatum\_InhaltNr.wav ... für das Referenzsignal  
**Beispiel** Ref\_Oe3\_250309\_Musik01.wav
- Test\_Sender\_Aufzeichnungsdatum\_InhaltNr.wav ... für das Testsignal  
**Beispiel** Test\_Oe3\_250309\_Musik01.wav
- Sender\_Aufzeichnungsdatum\_InhaltNr.xls ... für die Labelung  
**Beispiel** Oe3\_250309\_Musik01.xls

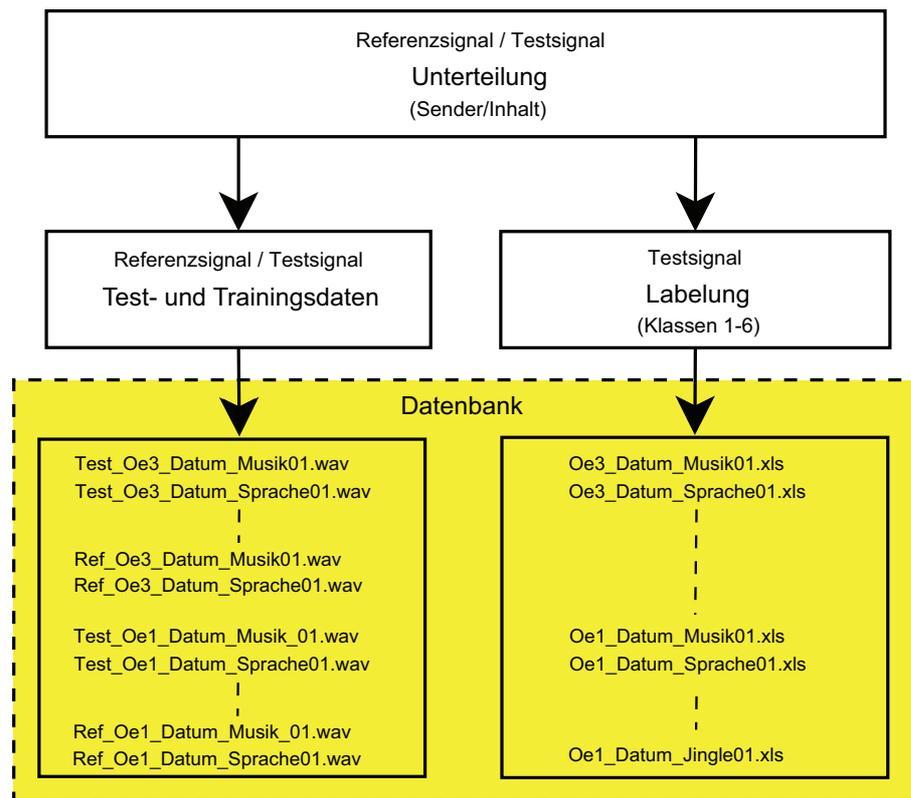


Abbildung 3.7: Struktur und Dateinamen nach Labelung

Nach der Labelung der Daten ist es möglich, einen Klassifikator zu erstellen, da erst mit dieser Labelung die Features den gefundenen Störklassen zugeordnet werden können. Nach erfolgreicher Labelung besteht der Datensatz für jedes Zeitframe aus einem „Datentripel“, wie es in Abbildung 3.8 zu sehen ist.

Zeit	Referenzsignal	Testsignal	Labelung
500 ms	Referenz	Test	Störklasse (1-6)
1000 ms	Referenz	Test	Störklasse (1-6)
1500 ms	Referenz	Test	Störklasse (1-6)
2000 ms	Referenz	Test	Störklasse (1-6)
⋮	⋮	⋮	⋮
N - 500 ms	Referenz	Test	Störklasse (1-6)
N ms	Referenz	Test	Störklasse (1-6)

Abbildung 3.8: Datentripel eines Files der Länge N ms

## 4 Implementierte Algorithmen

Ziel des Algorithmus zur Entwicklung des Klassifikators ist es, aus Testdaten Features zu extrahieren. Diese Features werden anschließend zur Klassifikation herangezogen (Abbildung 4.1).

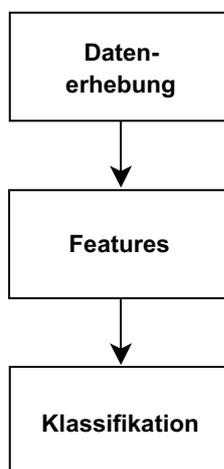


Abbildung 4.1: Algorithmus zur Klassifikation

Bereits in der Voruntersuchung mit einer Auswahl an Testdaten wurde ein Zeitversatz zwischen Test- und Referenzsignal festgestellt. Um die Signale zu synchronisieren, wurde eine Kreuzkorrelation durchgeführt. Allerdings war das Ergebnis nicht zufriedenstellend. Test- und Referenzsignal weisen neben dem zeitlichen Offset auch eine zeitliche Drift auf, d.h., die beiden Signale laufen mit fortschreitender Zeit auseinander. Der Offset der beiden Signale zueinander liegt bei nach einer Aufnahmezeitdauer von drei Minuten im Bereich von einer Sekunde.

Die zeitliche Drift ergibt sich aus Ungenauigkeiten der Samplefrequenz unterschiedlicher Aufnahmegeräte. Jedes Aufnahmegerät wird nämlich mit einem internen Quarz getaktet, aus dem die Samplefrequenz abgeleitet wird. Die Quarze in den verschiedenen Aufnahmegeräten unterscheiden sich aber durch Temperatur, Herstellungsverfahren und Alterungsprozess.

Da die Aufnahme von Test- und Referenzsignal an unterschiedlichen Orten stattfindet, ist eine Synchronisation der Geräte (z.B. Master Word Clock) nicht ohne weiteres möglich. Bei einem örtlich getrennten Aufnahmesetup ist deshalb eine Synchronisation im Anschluss an die Aufnahme erforderlich.

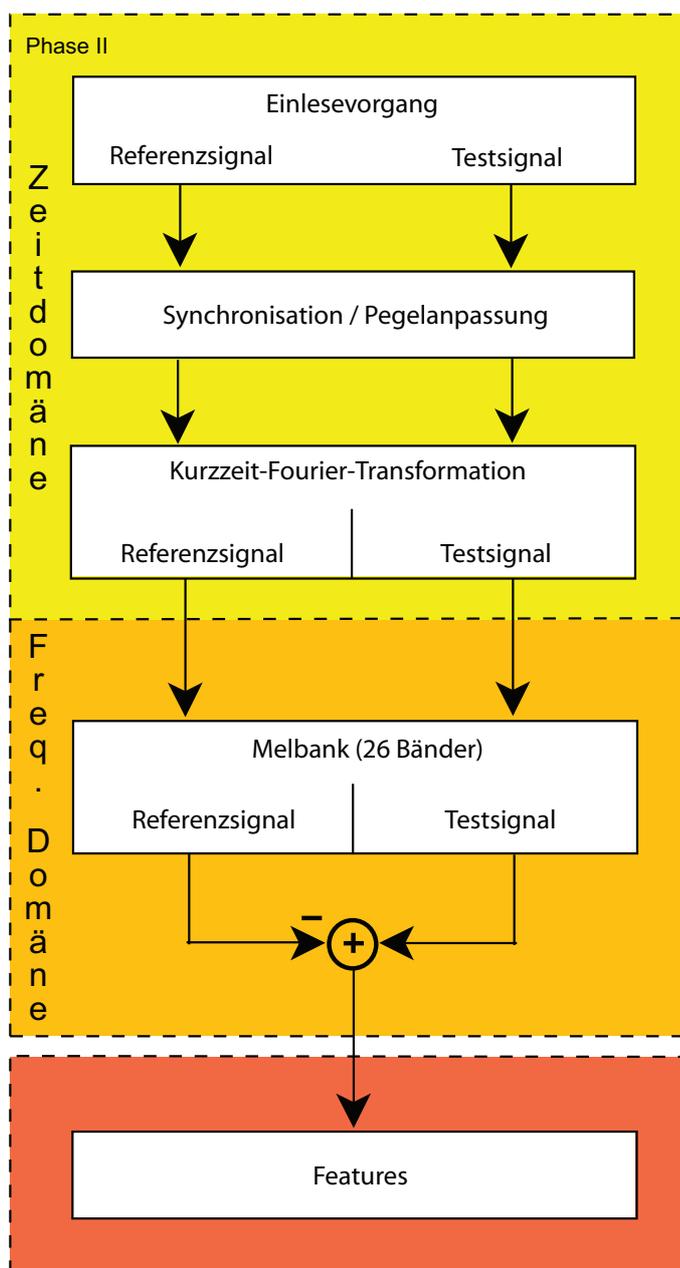


Abbildung 4.2: Algorithmus zur Feature-Extraktion

Aus zweierlei Gründen ist es nun sinnvoll, vom Zeitbereich in den Spektralbereich zu wechseln. Zum Einen ist es dort möglich, eine spektrale Verteilung der Signalenergie – Frame für Frame – zu betrachten, zum Anderen kann eine spektrale Einhüllende der Signalenergie mit reduzierter Genauigkeit berechnet werden. Die Phaseninkohärenz<sup>1</sup> der beiden Signale fällt dabei nicht mehr so signifikant ins Gewicht. In der spektralen Domäne kann nämlich der Betrag des Signals, der in direktem Zusammenhang zur Energie des Signals steht, getrennt von der Phasen betrachtet werden. Ein Struktogramm des Algorithmus ist in Abbildung 4.2 zu sehen.

## Synchronisation

Nachdem die Dateien eingelesen wurden, müssen die Signale zueinander synchronisiert werden. Da es, wie oben bereits ausgeführt, zu einer zusätzlichen zeitlichen Drift kommt, wird nicht ein einmaliger Offset der Zeitverschiebung berechnet, sondern in Zeitfenstern je 50 ms gearbeitet. Dabei wird für jedes Zeitfenster eine Kreuzkorrelation berechnet und der darin vorhandene Offset zum Referenzsignal addiert, um schließlich zeitsynchrone Signale zu erhalten. Erst danach ist eine Subtraktion des Referenzsignals vom Testsignal möglich. Ein schematischer Ablauf der Synchronisation ist in Abbildung 4.3 zu sehen.

In diesem Zusammenhang ist zu erwähnen, dass eine so genannte HOP-Size definiert werden muss, um das Signal zu synchronisieren. Dabei wird beim Einlesen der Daten jedes Zeitfenster um einen zusätzlichen Bereich (HOP-Size) zur zeitliche Verschiebung des Signals erweitert. Zur Synchronisation steht damit ein Datenbereich zur Verfügung, dessen Einlesegrenzen entsprechend dem Offset aus Berechnungen der Kreuzkorrelation angepasst werden können. Dadurch kann Frame für Frame synchron in einen Zwischenspeicher abgelegt werden, um die darin enthaltenen Daten dann als neue Referenzdatei zu speichern.

---

<sup>1</sup>Phaseninkohärenz: Diese kommt aufgrund der unterschiedlichen Übertragungsstrecken zustande

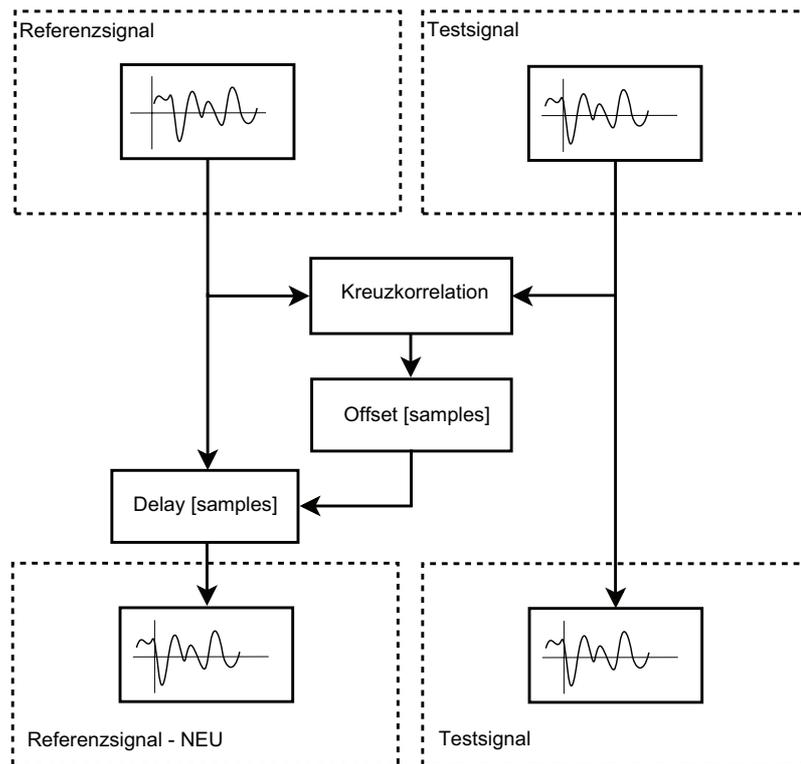


Abbildung 4.3: Synchronisation von Test- und Referenzsignal mit Kreuzkorrelation

### Pegelanpassung

Die Voruntersuchung hat auch ergeben, dass es für die praktische Anwendung nicht möglich sein wird, an den Aufnahmegegeräten für Test- und Referenzsignal dieselbe Pegeleinstellung vorzunehmen. Deshalb wurde eine Pegelanpassung implementiert, bei der aus beiden aufgezeichneten Signalen der RMS-Wert (Root Mean Square) berechnet wird. Das Verhältnis der beiden RMS-Werte ergibt den Korrekturfaktor, der dann zum Referenzsignal multipliziert wird. Im ersten Schritt wird aus Test- und Referenzsignal die Varianz  $\sigma_x^2 = Var(x)$  bzw.  $\sigma_y^2 = Var(y)$  berechnet, die dem RMS-Wert entspricht. Dabei gilt für das Testsignal der Zusammenhang

$$X_{RMS} = \sqrt{\sigma_x^2} \quad (4.1)$$

und entsprechend für das Referenzsignal

$$Y_{RMS} = \sqrt{\sigma_y^2} \quad (4.2)$$

Das pegelangepasste Signal ergibt sich aus

$$y = y * \frac{X_{RMS}}{Y_{RMS}} \quad (4.3)$$

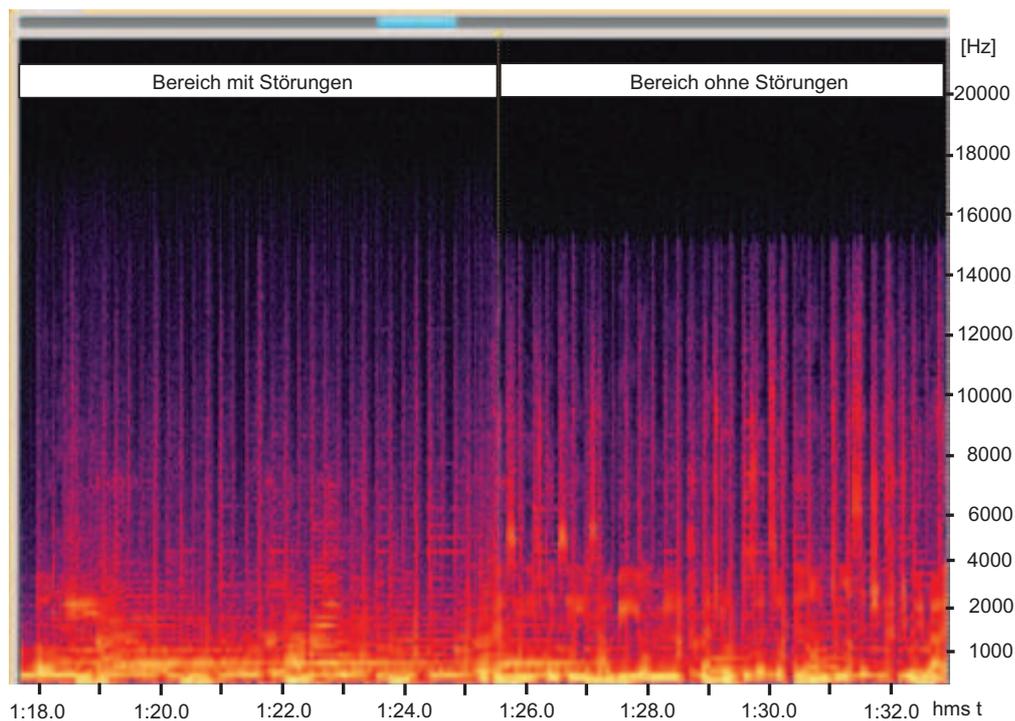


Abbildung 4.4: Im Spektrum zwischen 15 und 21 kHz sind Störungen ersichtlic

## 4.1 Feature-Extraktion

Aus den nun vorliegenden Daten können Features extrahiert werden.

### 4.1.1 Melbank

Im ersten Schritt zur Feature-Extraktion wird eine Melskala generiert. Für die Anwendung der Tunerqualifizierung wurde ein Frequenzbereich von 100 Hz bis 20 kHz gewählt und die Anzahl der Melfilterbänke auf 26 gesetzt. Das Signal des UKW-Tuners wird zwar Tiefpass gefiltert (vgl. Abschnitt 3.2.1) gefiltert, trotzdem ist beim Auftreten von Störungen ein Teil der Energie im Spektralbereich oberhalb von 16 kHz zu finden. Dies kann mit einfachen Spektralanalysemethoden belegt werden. Ein typisches Spektrum mit Störungen im besagtem Frequenzbereich ist in Abbildung 4.4 zu sehen. Man erhält eine Melbank wie sie in Kapitel 2.3 bereits beschrieben wurde.

### 4.1.2 Kurzzeit-Fourier-Transformation

Vor der Transformation in die spektrale Domäne wird das Signal in kurze Zeitblöcke zerlegt. Nach die Bufferung in Blöcke, die einem Vielfachen einer Zweierpotenz entsprechen, wird das Signal Fourier transformiert und mit einer Fensterfunktion (z.B. Hann-Fenster) gewichtet.

$$NFFT = 2^N \quad (4.4)$$

Zudem wird bei der Bufferung des Signal mit überlappenden Blöcken gearbeitet, um die gleiche Framelänge wie bei der Daten-Labelung zu erreichen (Abbildung 4.5). Für diese Anwendung wurde eine NFFT-Länge von 4096 gewählt. Somit ergeben sich sechs überlappende Blöcke innerhalb eines Zeitframes von 500 ms. Die Blocklänge entspricht dabei einer Dauer von 92 ms.

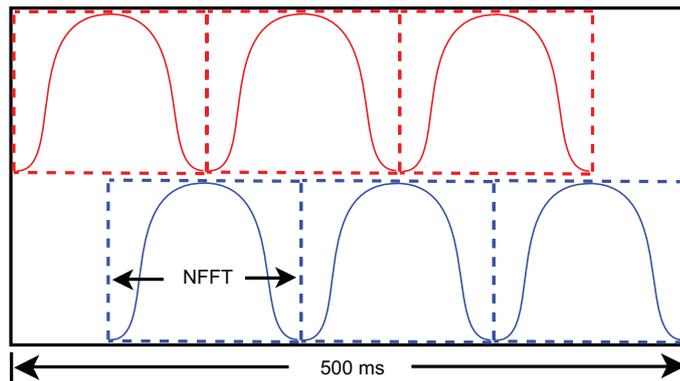


Abbildung 4.5: Überlappende Blöcke mit Fensterfunktion zur Glättung des Spektrums

Der nun vorliegenden Betrag des Fourier transformierten Signals wird mit den Bändern der Melkala multipliziert. Dies entspricht einer Gewichtung des Signals entsprechend der Mel-Filterbank wie es aus Gleichung 4.5 hervor geht.

$$E_m(n) = |X(n)| * H(m) \quad (4.5)$$

$E_m(n)$  ... Energie des Signals  $x(k)$  mit Mel-Gewichtung

$|X(n)|$  ... Betrag des Fourier transformierten Signals  $x(k)$

$H(m)$  ... Melbank

### 4.1.3 Differenz von Test- und Referenzsignal

Nun liegt das Mel-Spektrum, Frame für Frame, für Test- wie für Referenzsignal vor, wobei jedes Frame 26 Bänder der Melskala beinhaltet.

$$E(m) = \sum_{n=1}^{N-1} E_m(n) \quad m = 1, \dots, 26 \quad (4.6)$$

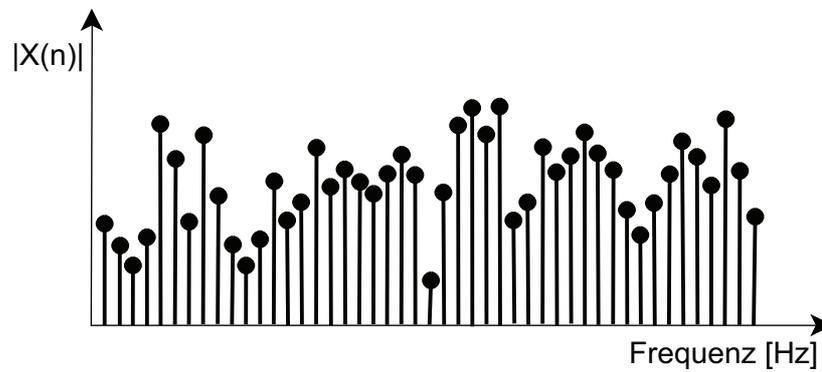
$E(m)$  ... Energie in einem Melband  $m$

$E_m(n)$  ... Mel-gewichtete Energie

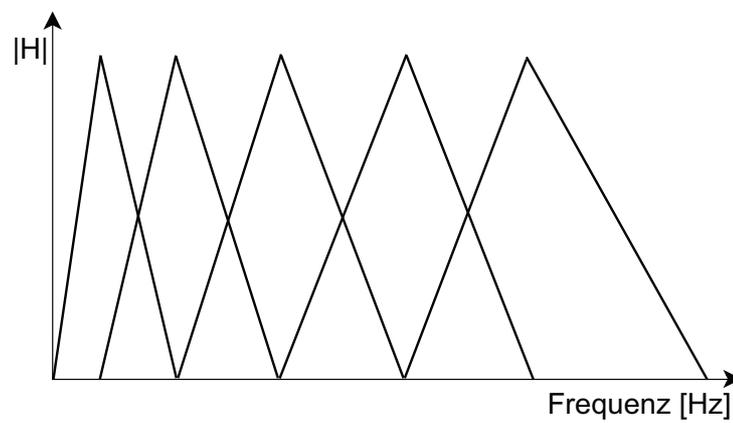
$N$  ..... Anzahl der FFT-Bins pro Melband – entspricht NFFT aus DFT

Die einzelnen Frequenzbins in den Melbändern von Test- und Referenzsignal können so aufsummiert werden. Am Ende stehen jeweils 26 Werte je Frame zur Verfügung.

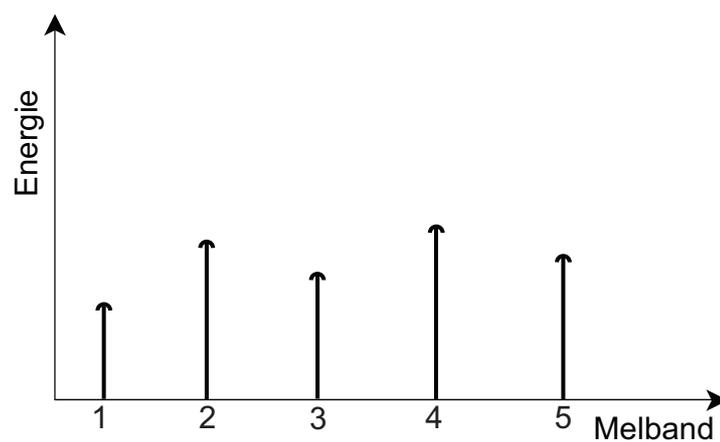
Zur besseren Veranschaulichung wird dieser Vorgang am Beispiel einer Melskala mit fünf Bändern beschrieben. (Abbildung 4.6)



a) Kurzzeit-Fourier-Transformation eines Signals



b) Melbank mit 5 Bänder



c) Aufsummierte Energie der jeweiligen Melbänder

Abbildung 4.6: Zusammenfassen der Energie in Melbänder

Die Energien der Bänder, wie sie in Abbildung 4.7 a) und 4.7 b) dargestellt ist, können dann voneinander subtrahiert werden. Allgemein gilt für die Energie der Differenz der Mel-gewichteten Signale folgender Zusammenhang:

$$Ed(m) = E_x(m) - E_y(m) \quad m = 1, \dots, 26 \quad (4.7)$$

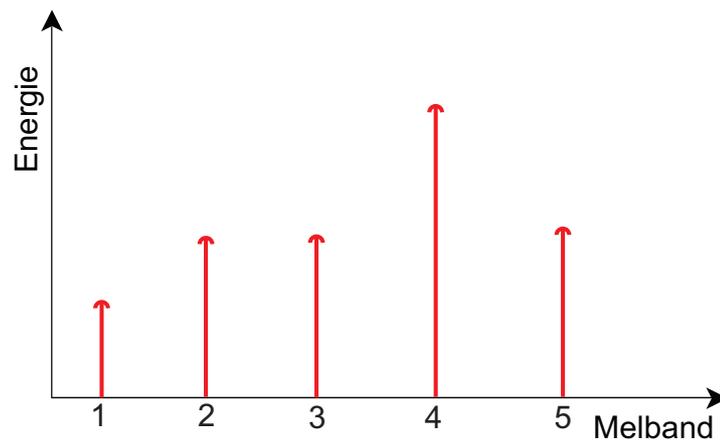
$Ed(m)$  ...Energie des Testsignals

$E_x(m)$  ...Energie des Referenzsignals

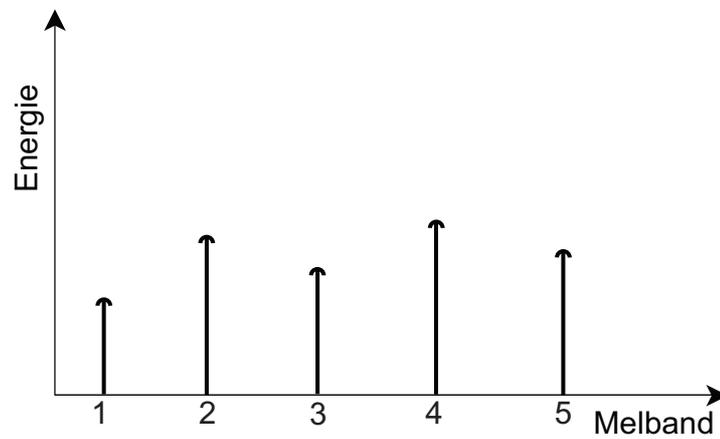
$E_y(m)$  ...Energie der Differenz aus entsprechenden Melbänder

Die Energie der Differenz wird quadriert. Das Ergebnis ist in Abbildung 4.7 c) dargestellt.

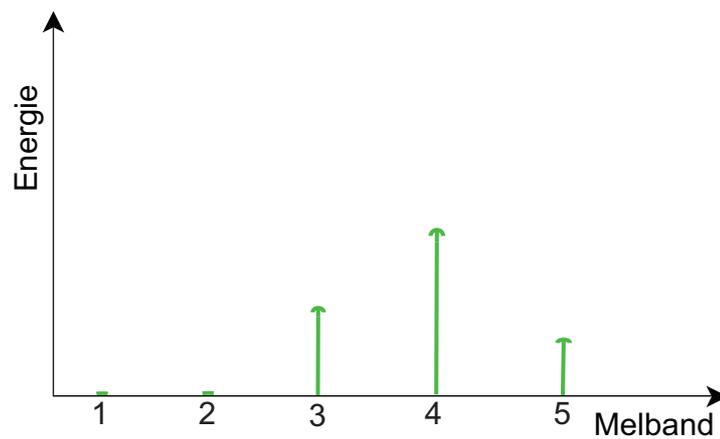
$$ENRG(m) = Ed(m)^2 \quad (4.8)$$



a) Energie des Testsignals der 5 Melbänder



b) Energie des Referenzsignal der 5 Melbänder



c) Energie des Differenzsignals der 5 Melbänder

Abbildung 4.7: Differenzbildung von Test- und Referenzsignal der einzelnen Melbänder

### Feature-Vektoren

Aus der Energie der gewichteten Differenz (Gleichung 4.8) können nun die Features extrahiert werden. Die verwendeten Features werden im folgenden Abschnitt genauer beschrieben.

#### 4.1.4 Verwendete Feature-Vektoren

- Summe über alle Bänder (1D-Feature-Vektor)  
Als erstes Feature steht die Summe der Energie über alle Bänder innerhalb eines Zeitframes zur Verfügung. Im folgenden wird diese Summe als `FEATURESUM` bezeichnet.
- Melband, in dem das Maximum liegt (2D-Feature-Vektor)  
Bei diesem Feature-Vektor wird das Band bestimmt, in dem der maximale Energiewert liegt. Beide Größen, also Maximalwert sowie das Melband, in dem dieser vorkommt, bilden den Feature-Vektor. Diese zweidimensionale Größe wird im folgenden als `FEATUREMAX` bezeichnet.
- Energien aller 26 Melbänder (26D-Feature-Vektor)  
Für diesen Feature-Vektor werden alle 26 Melbänder beibehalten. Die Energien aus der Differenzberechnung von Test- und Referenzsignal bilden somit den Feature-Vektor, der im folgenden als `FEATURE26` bezeichnet wird.

Weitere verwendete Features sind Delta-Features. Dabei wird zu jedem der drei oben beschriebenen Feature-Vektoren ein Differenz-Feature – Delta Feature – berechnet (vgl. Abschnitt 2.3). Daraus erhält man zusätzlich die drei folgenden Features:

- `DeltaFeatureSum` (2D-Feature-Vektor)  
Delta-Feature des `FeatureSum`
- `DeltaFatureMax` (4D-Feature-Vektor)  
Delta-Feature des `FeatureMax`
- `DeltaFeature26` (52D-Feature-Vektor)  
Delta-Feature des `Feature26`

Störklasse	Bezeichnung
1	Störfrei
2	Knacks
3	Granulares Rauschen (Bröseln)
4	Verzerrung
5	Rauschen
6	Aussetzer im Empfang

Tabelle 4.1: Störklassen zur Sortierung der Feature-Vektoren

## 4.2 Sortierung der Feature-Vektoren

Mit der Zuordnung der gefundenen Feature-Vektoren zu den Klassen (Clustering) ist das „Training“ innerhalb des Klassifikationsverfahrens abgeschlossen. Diese Sortierung der Feature-Vektoren ist durch das Vorhandensein von gelabelten Dateien möglich.

Die gewonnenen Daten wurden, wie in Abschnitt 3.3.4 bereits erläutert, anhand vorgegebener Zeitframes gelabelt: Jedem Frame wird ein Wert von 1 bis 6 zugewiesen, wobei jede Zahl einer Klasse der definierten Störklassen entspricht (Tabelle 4.1). Zum Erstellen der Klassen werden 75 Prozent der Testdaten aus der Datenbank verwendet. Zusammen mit den händisch gelabelten Daten können die extrahierten Features, Frame für Frame, den bekannten Klassen zugeordnet werden. Daraus ergibt sich eine Zusammengehörigkeit von beobachtbaren Features aus der Labelung, Feature-Vektoren die mit Hilfe des Algorithmus extrahiert werden, und der Klassenzahl. Daraus resultieren sechs, nachfolgend als Trainingsbasis bezeichnete Klassen. Wünschenswert ist eine klare Abgrenzung der unterschiedlichen Cluster. Eine derartige Abgrenzung ist jedoch meist nur theoretisch möglich. Aus diesem Grund gilt es, einen geeigneten Klassifikator zu finden, der auch bei teilweise überlappenden Clustern noch gute Ergebnisse liefert. Dieser Prozess ist anhand eines zweidimensionalen Feature-Raumes in Abbildung 4.8 dargestellt.

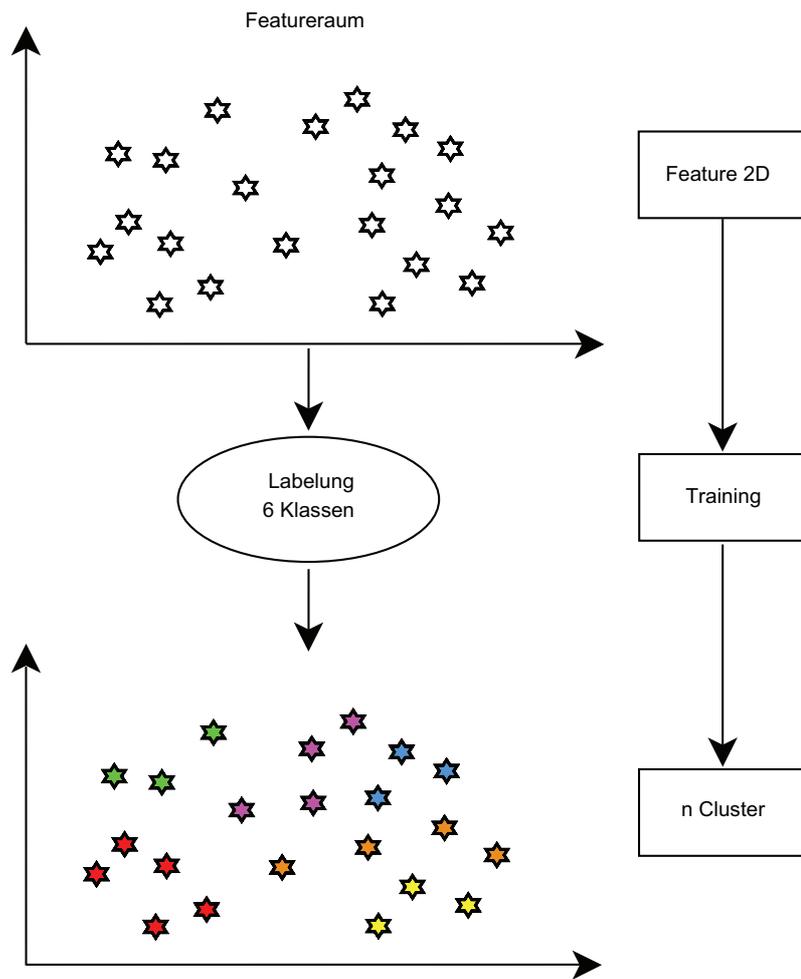


Abbildung 4.8: Clustering mit gelabelten Daten

### 4.3 Wahl des Klassifikators

Wie sich bei der Clustering der Trainingsdaten herausstellte, weichen die Clustergrößen der unterschiedlichen Störklassen voneinander ab. Dies erklärt sich aus der unterschiedlichen Häufigkeit, mit der verschiedene Störfälle innerhalb einer bestimmten Zeit auftreten. So tritt zum Beispiel Rauschen sehr viel öfter auf als beispielsweise Verzerrungen. Noch deutlicher ist der jeweilige Unterschied zwischen der störfreien Klasse und den fünf verschiedenen Störklassen; hier findet man sogar Verhältnisse von bis zu 1:1000 vor.

Nun könnte man zwar die Trainingsbasis auf gleich verteilte Cluster reduzieren, um Verfahren wie z.B. die k-NN-Klassifikation anwenden zu können. Allerdings weisen die Klassen „Aussetzer“ und „Verzerrung“ gar nur ein Dutzend

Trainingspunkte auf, weshalb die Wahl auf den NN-KLASSIFIKATOR fiel. Der mit einem NN-Klassifikator verbundene hohe Rechenaufwand spielt nur eine untergeordnete Rolle, weil bei der Auswertung der Empfangsqualität eines KFZ-Tuners die Ergebnisse nicht unmittelbar zur Verfügung stehen müssen.

### Vorgehensweise

Zur Klassifikation werden jene 25 Prozent der Daten herangezogen, die bereits als Testdaten definiert wurden (im Folgenden als Testbasis bezeichnet). Mit dieser Testbasis kann anschließend auch eine detaillierte Aussage über die Güte des Klassifikators getroffen werden, weil diese Testdaten ebenfalls Teil der gelabelten Datenbank sind.

Bei der Klassifikation werden die zu klassifizierenden Testdaten den ermittelten Störklassen zugeordnet; diese Berechnung geschieht über die euklidische Distanz.

### 4.3.1 Nachbearbeitung des Trainingsdatensatzes

Für den endgültigen Klassifikator wurden die Daten schließlich bezüglich der Randframes bzw. der überlappenden Blöcke sowie hinsichtlich der maximalen Anzahl von Vektoren je Klasse bearbeitet (Größe der Trainingsbasis).

#### Bufferblöcke

Feature-Vektoren der Randframes aus den gelabelten Blöcken (500 ms) wurden bei der Erstellung der Trainingsbasis aus zwei Gründen nicht berücksichtigt. Einerseits konnten dadurch die überlappenden Blöcke eliminiert werden die ohnedies zu einer möglicherweise anderen Störklasse gehören. Zum Anderen werden damit auch „Unschärfen“ der gelabelten Daten aus dem Trainingsdatensatz eliminiert. Dieser Vorgang ist in Abbildung 4.9 dargestellt.

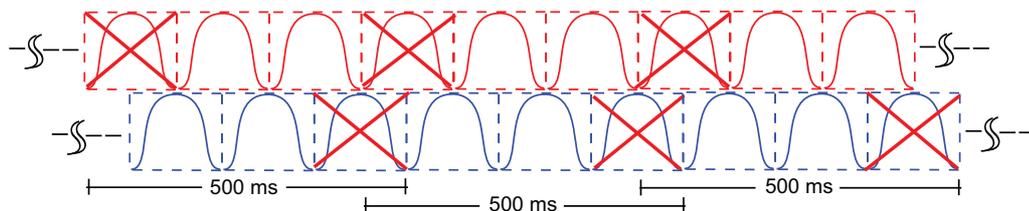


Abbildung 4.9: Überlappung der Frames bei der Bufferung

Sender	Trainingsbasis 2.000			Trainingsbasis 7.000		
	Feature26	FeatureMax	FeatureSum	Feature26	FeatureMax	FeatureSum
Ö1	17 %	35 %	34 %	29 %	54 %	60 %
Ö3	33 %	26 %	24 %	47 %	44 %	42 %

Tabelle 4.2: Anzahl richtig erkannter Frames in Abhängigkeit von der Größe der Trainingsbasis.

### Größe der Trainingsbasis

Wie aus Gleichung 4.9 hervorgeht, bedarf es für sechs Klassen, die aus einer entsprechend großen Datenmenge bestehen, einer sehr langen Rechenzeit, um sämtliche Vektoren der Testbasis einer Störklassen zuzuweisen. Um die Berechnungszeit bei der Klassifikation zu verkürzen, wurde zunächst mit maximal 2.000 Vektoren je Klasse gearbeitet. Diese relativ geringe Größe der Trainingsbasis wurde ursprünglich auch im Hinblick auf die Tatsache gewählt, dass die Anzahl der auftretenden störfreien Frames sehr stark überwiegt.

$$d(x, y) = \sqrt{\sum_i^n (x_i - y_i)^2} \quad (4.9)$$

Um die Reproduzierbarkeit des Klassifikators mit kleiner Trainingsbasis zu überprüfen, wurde die Anzahl der Trainingsbasis auf maximal 7.000 Vektoren hinaufgesetzt. Tatsächlich ist die Anzahl der richtig erkannten Frames bei der Trainingsbasis mit 2.000 Vektoren gegenüber jener mit 7.000 Vektoren erheblich geringer, wie aus der vergleichenden Tabelle in Abbildung 4.2 ersichtlich ist. Bei einer Wahl von zu wenig Trainingsvektoren sinkt demnach die Güte des Klassifikators.

Auf der Grundlage dieser Ergebnisse erfolgt die Klassifikation des Testdatensatzes im folgenden Abschnitt mit einer maximalen Anzahl von 7.000 Feature-Vektoren der Trainingsbasis.

## 5 Test des Klassifikators

In folgender Analyse werden die vom Algorithmus klassifizierten Testdaten mit der gelabelten Datenbank abgeglichen. Durch diesen Vergleich, der Frame für Frame durchgeführt wird, lässt sich die Qualität des Klassifikationsalgorithmus feststellen. Aufgrund der Struktur der gelabelten Datenbank ist bei der Auswertung der klassifizierten Testdaten folgendes zu beachten:

Wie in Abschnitt 4.1.2 beschrieben, arbeitet der Berechnungsalgorithmus zur Feature-Extraktion mit überlappenden Frames. Die gelabelten Daten liegen in Frames zu je 500 ms vor, die zur Datenclusterung zusätzlich in sechs Bufferblöcke unterteilt wurden. Durch diese kontinuierliche Bufferung des Signals überschneidet sich der erste und letzte Block eines jeden Frames. Deshalb durften die außen liegenden Blöcke jedes Frames NICHT zur Clusterung herangezogen werden. Ebenso wenig dürfen die „Rand-Blöcke“ bei der Klassifizierung mit evaluiert werden. Bei der Auswertung ist also unbedingt darauf zu achten, dass die sich überlappenden Blöcke benachbarter Zeitfenster (vgl. Abschnitt 4.3.1) NICHT mit ausgewertet werden. Dies würde nämlich zu einer Verfälschung der Ergebnisse führen.

### **Auswertung der Ergebnisse**

Bei der Analyse der ersten drei Feature-Räumen stellte sich heraus, dass es wenig zielführend ist, alle vorhandenen Dateien der Testbasis einzeln auszuwerten. In Tabelle 5.1 ist der Prozentanteil der richtig klassifizierten Frames zu sehen. Da innerhalb eines Features sehr unterschiedliche Trefferquoten zu erkennen sind, lässt sich aus dieser Darstellung nur schwer eine Aussage über die Güte des Klassifikators treffen: Eine sehr hohe Streuung lässt keine exakte Aussage über die Trefferquote des Klassifikators zu.

Um eine brauchbare Aussage über die Trefferquote machen zu können, ist es also notwendig, die Daten der Testbasis zusammenzufassen. Die bereits im Vorfeld durchgeführte Unterteilung der Daten erwies sich in diesem Zusammenhang als über-

Testdateien	Feature26	FeatureMax	FeatureSum
Oe1_100609_Musik02	16	35	25
Oe1_100609_Musik05	4	64	18
Oe1_100609_Musik08	12	46	30
Oe1_100609_Sprache02	40	35	42
Oe1_100609_Sprache06	40	36	43
Oe1_100609_Sprache09	29	39	43
Oe1_200509_Doku02	3	10	34
Oe1_200509_Musik05	4	13	22
Oe1_200509_Musik06	21	35	25
Oe1_200509_Musik10	0	8	15
Oe1_200509_Sprache04	4	8	20
Oe1_200509_Sprache07	12	8	31
Oe1_200509_Sprache08	0	12	50
Oe1_200509_Sprache12	13	11	40
Oe1_290509_Musik03	4	37	22
Oe1_290509_Sprache05	63	30	62
Oe3_100609_Jingle01	57	61	44
Oe3_100609_Jingle02	44	48	46
Oe3_100609_Musik02	18	24	23
Oe3_100609_Musik07	31	24	28
Oe3_100609_Musik10	36	8	25
Oe3_100609_Musik11	25	25	22
Oe3_100609_Musik14	20	37	14
Oe3_100609_Sprache02	32	47	39
Oe3_200509_Jingle01	9	11	56
Oe3_200509_Musik05	2	3	18
Oe3_200509_Musik07	11	10	24
Oe3_200509_Musik09	23	12	22
Oe3_200509_Sprache03	1	6	43
Oe3_290509_Musik04	91	95	55
Oe3_290509_Musik07	93	96	65
Oe3_290509_Sprache02	95	82	66

Tabelle 5.1: Auswertung der Testdateien — richtig erkannte Frames (Angaben in Prozent)

aus nützlich, weil dadurch eine Zusammenfassung nach Sendern oder nach Inhalten beliebig möglich ist.

Tabelle 5.2 zeigt den Prozentanteil der richtig klassifizierten Frames aus den sechs Störklassen für den eindimensionalen Feature-Raum (Summe über alle Bänder); dabei handelt es sich um eine Zusammenfassung nach Sendern.

Sender	FeatureSum
Ö1	50 %
Ö3	47 %

Tabelle 5.2: Richtig erkannte Klassen nach Sendern

Wie bereits erwähnt, können die Testdaten auch nach Inhalten zusammengefasst werden. Die Auswertungen mit den unterschiedlichen Zusammenfassungen der Ergebnisse befinden sich im Anhang (Anhang A: Auswertung).

Ein Vergleich der Klassifikationsergebnisse ist in Tabelle 5.3 zu sehen; verglichen werden hier die drei Feature-Vektoren: FeatureSum, FeatureMax und Feature26. (siehe Abschnitt 4.1.4)

Sender	Feature-Raum		
	Feature26	FeatureMax	FeatureSum
Ö1	29 %	54 %	50 %
Ö3	47 %	44 %	47 %

Tabelle 5.3: Auswertung der drei Feature-Räume im Vergleich

Die Ergebnisse der Klassifikation der beiden Feature-Räume FeatureSum und FeatureMax sind nahezu identisch, wobei die Trefferquote für den Sender Österreich 1 für beide Feature-Vektoren etwas höher liegt als für Hitradio Ö3. Hier fällt auf, dass ein Feature-Raum höherer Ordnung (Feature26) wider Erwarten zu schlechteren Ergebnissen führt. Eine Interpretation dieses Ergebnisses führt uns zu einem Schwachpunkt dieser Darstellungsmethode. Begründet liegt dieser in den nicht gleich verteilten Klassen:

Wie bereits erwähnt, beinhaltet sowohl die Trainings- als auch die Testbasis verhältnismäßig viele Frames aus Klasse 1, also aus der störfreien Klasse. Ein Klassifikator, der alle Frames der störfreien Klasse zuordnet, hat deshalb eine sehr hohe Trefferquote, weil der Großteil der Testframes dieser dominierenden störfreien Klasse tatsächlich zugehört. Allein darin liegen die besseren Ergebnisse des eindimensionalen Feature-Raumes begründet.

### Erweiterte Darstellungsmethode

Es ist also notwendig, eine zusätzliche Darstellungsmethode zu wählen, bei der nur die richtig klassifizierten Störklassen 2 bis 6 angegeben werden, um so aus den Ergebnissen eine Genauigkeitsabschätzung für den Klassifikator durchführen zu können. Ein Vergleich der drei Feature-Räume mit den Klassen 2 bis 6 ist in Tabelle 5.4 zu sehen.

Sender	Feature-Raum		
	Feature26	FeatureMax	FeatureSum
Ö1	27 %	31 %	39 %
Ö3	34 %	38 %	34 %

Tabelle 5.4: Auswertung der richtig erkannten Frames der Klassen 2-6 (Störungen)

In Tabelle 5.4 sind kaum mehr Unterschiede in den Ergebnissen der drei Feature-Räume zu erkennen. Um Aufschluss darüber zu erhalten, welche Klassen mit anderen vertauscht werden, bedarf es einer Darstellungsmethode bei der

- richtig klassifizierte Klassen erkannt werden,
- Verwechslungen der unterschiedlichen Klassen ersichtlich sind,
- falsch erkannte Klassen angezeigt werden und
- die störfreie Klasse sowie die Klassen (2 bis 6) mit Störungen zu erkennen sind.

Eine derartige Darstellungsmethode zur Genauigkeitsabschätzung ist die Confusion-Matrix, wie sie bereits in Abschnitt 2.4 erklärt wurde.

### Confusion-Matrix

Auch bei dieser Darstellungsmethode werden die Ergebnisse nach den gewünschten Kriterien<sup>1</sup> zusammengefasst, die für den eindimensionalen Feature-Vektor folgendermaßen aussehen:

---

<sup>1</sup>Kriterien: Vollständige Ergebnisse mit möglichen Zusammenfassungen im Anhang A: Auswertung

Confusion-Matrix Ö1

Ö1		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	4297	2690	3098	36	75	2
	Kl. 2	32	57	66	5	73	0
	Kl. 3	52	75	79	10	57	0
	Kl. 4	35	4	10	0	4	5
	Kl. 5	93	126	189	24	307	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix Ö3

Ö3		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	7339	3438	3055	11	1303	18
	Kl. 2	44	135	114	18	41	0
	Kl. 3	90	222	287	21	231	4
	Kl. 4	20	35	7	4	56	6
	Kl. 5	20	53	12	45	60	0
	Kl. 6	4	3	4	7	2	14

Tabelle 5.5: Erkannte Klassen mit eindimensionalem Feature

Die Confusion-Matrix beinhaltet sehr viele Werte, die außerhalb der Hauptdiagonale liegen. Das bedeutet, dass es viele falsch klassifizierte Frames gibt.

Auch für das zweidimensionale Feature sehen die Ergebnisse ähnlich aus (Tabelle 5.6). Allerdings ist bereits der Trend ersichtlich, dass sehr viele Werte in der ersten Zeile der Matrix liegen (Tabelle 5.6).

Confusion-Matrix Ö1

Ö1		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	6863	2017	2585	11	89	1
	Kl. 2	156	63	37	3	14	0
	Kl. 3	155	22	82	2	9	0
	Kl. 4	40	3	0	0	0	0
	Kl. 5	431	93	86	14	378	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix Ö3

Ö3		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	7865	2776	4193	10	1386	34
	Kl. 2	44	185	22	7	33	0
	Kl. 3	114	196	296	16	214	2
	Kl. 4	37	27	3	4	51	0
	Kl. 5	31	41	86	2	68	0
	Kl. 6	2	6	8	3	4	11

Tabelle 5.6: Erkannte Klassen mit zweidimensionalem Feature

Für einen weiteren Vergleich, aber auch, um diese neue Erkenntnis gegebenenfalls bestätigen zu können, wird die Klassifikation des 26-dimensionalen Feature-Raumes analysiert. Dabei erhält man folgende Ergebnisse (Tabelle 5.7):

Ö1		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	4061	3989	3956	8	1752	0
	Kl. 2	71	92	80	0	44	0
	Kl. 3	68	81	90	0	41	0
	Kl. 4	31	12	0	0	0	0
	Kl. 5	143	238	254	0	81	0
	Kl. 6	0	0	0	0	0	0

Ö3		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	8148	3264	3452	7	1372	21
	Kl. 2	36	171	48	5	31	0
	Kl. 3	98	397	277	2	94	0
	Kl. 4	41	30	9	0	42	0
	Kl. 5	23	131	50	3	21	0
	Kl. 6	4	8	16	0	5	1

Tabelle 5.7: Erkannte Klassen mit 26-dimensionalem Feature

Durch die Confusion-Matrix des 26-dimensionalen Features wird die Erkenntnis bestätigt werden, dass tatsächlich immer sehr viele Werte in der ersten Zeile der Matrix liegen. Dies bedeutet, dass Frames den Störclustern zugeordnet werden, in denen keine Störungen vorhanden sind, dass also Frames „falsch-positiv“ erkannt werden.

Diese Ergebnisse lassen darauf schließen, dass die Störungsdetektion an und für sich funktioniert, die als „falsch-positiv“ klassifizierten Frames die Ergebnisse aber verschlechtern.

Die folgenden Untersuchung der Delta-Features gibt Aufschluss darüber, in wie weit ein Vergleich benachbarten Frames das Klassifikationsverfahren verändert.

Für den 52-dimensionalen Feature-Raum beispielsweise liefern die Delta-Features folgende Ergebnisse:

Ö1		IST					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	3164	3606	3204	6	3777	0
	Kl. 2	54	92	76	0	64	0
	Kl. 3	69	66	67	0	78	0
	Kl. 4	23	6	0	0	0	0
	Kl. 5	111	224	150	0	231	0
	Kl. 6	0	0	0	0	0	0

Ö3		IST					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	7333	3304	3247	3	1242	35
	Kl. 2	38	127	120	0	49	0
	Kl. 3	79	403	242	1	11	0
	Kl. 4	52	6	0	0	0	0
	Kl. 5	36	75	62	0	35	0
	Kl. 6	4	0	1	0	0	15

Tabelle 5.8: Erkannte Klassen mit 52-dimensionalen Delta-Feature

Betrachtet man die Sub-Matrix (Abbildung 5.1) durch einen roten Rahmen gekennzeichnet), kann man den Gedanken entwickeln, aus dieser 6x6-Matrix eine 2x2-Matrix zu bilden. Dabei wird nur mehr danach unterschieden, ob eine Störung vorhanden ist (Klasse 0) oder nicht (Klasse 1). Die unterschiedlichen Störklassen (Klasse 2 bis 6) sind in Klasse 0 zusammengefasst.

Ö3		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o r t	Kl. 1	7333	3304	3247	3	1242	35
	Kl. 2	38	127	120	0	49	0
	Kl. 3	79	403	242	1	11	0
	Kl. 4	52	6	0	0	0	0
	Kl. 5	36	75	62	0	35	0
	Kl. 6	4	0	1	0	0	15

Ö3		I S T	
		Klasse 1	Klasse 0
S o r t	Kl. 1	7333	7831
	Kl. 0	209	1147

Abbildung 5.1: Confusion-Matrix mit Sub-Matrix

Bei einer solchen Zusammenfassung der Störklassen ergibt sich für den Sender Hitradio Ö3, dass aus den insgesamt 16.517 Frames die vorhandenen Störframes zu 85 Prozent richtig erkannt werden. Im Vergleich dazu liegt der Anteil der als „falsch-positiv“ erkannten Frames bei 52 Prozent. Ähnliche Ergebnisse erhält man für den Sender Österreich 1; hier liegt die Trefferquote für die richtig erkannten Störungen bei 82 Prozent, jene für die als „falsch-positiv“ erkannten Frames bei 46 Prozent.

# 6 Zusammenfassung und Ausblick

## 6.1 Zusammenfassung

Detektion und Klassifikation – was ist möglich?

Nach den Erkenntnissen aus der ausführlichen Analyse der Ergebnisse kann man festhalten, dass mit den ausgewählten Features und dem verwendeten Klassifikator Störungen zwar erkannt werden; aufgrund der hohen Anzahl der „falsch-positiv“ erkannten Frames, ist das Verfahren für eine praktische Anwendung jedoch noch nicht ausgereift. Um mit den vorhandenen Features dennoch eine hohe Trefferquote zu erzielen, ist die Verwendung von Delta-Features – vorzugsweise im 52-dimensionalen Raum – zu empfehlen.

In jedem Fall wird mit dem angewandten theoretischen Modell gezeigt, wie eine objektive Tunerqualifizierung möglich ist und wie reproduzierbare<sup>1</sup> Prüffahrten durchgeführt werden können. Gleichzeitig wurde ein Auswertalgorithmus hinsichtlich der konkreten Problemstellung dieser Arbeit, der Störungsdetektion bei KFZ-Tunern, entwickelt.

Es wurde also ein geeignetes Konzept mit dazugehörigem Messsetup für eine objektive Tunerqualifizierung ausgearbeitet. Das dem Algorithmus zugrunde liegende theoretische Modell eignet sich zur Weiterentwicklung.

## 6.2 Ausblick

Eine Weiterentwicklung des Klassifikationsverfahrens kann darin bestehen, zusätzliche Features zu finden, zu extrahieren und gegebenenfalls mit den bereits vorhandenen zu kombinieren.

---

<sup>1</sup>reproduzierbar: unter Berücksichtigung der Witterungsbedingung

Aufgrund des modularen Aufbaus des hier entwickelten Algorithmus können neue Features problemlos in die bereits vorhandene Struktur integriert werden.

Eine andere Entwicklungsmöglichkeit besteht darin, weitere Daten zu akquirieren, um statistische Klassifikationsverfahren zu ermöglichen.

Wird das Verfahren durch diese Erweiterungen dahingehend optimiert, dass die als „falsch-positiv“ klassifizierten Daten der „richtigen“, also der störfreien Klasse zugeordnet werden, können damit automatische Analysen und Auswertungen durchgeführt werden.

Wird für das hier verwendete Testsetup ein Klassifikationsverfahren gefunden, das eine gute Störungsdetektion ermöglicht, kann das Verfahren auch dahingehend untersucht werden, ob es mit einem neuen Ansatz – ohne Referenzsignal – für eine objektive Tunerqualifizierung herangezogen werden kann.

Schlussendlich sollte der Algorithmus von Matlab in eine niedrigere Programmiersprache portiert werden, um ihn dadurch hinsichtlich Speicherplatzmanagements und Rechengeschwindigkeit zu optimieren.

Dringend empfohlen sei schließlich die Durchführung psychoakustischer Tests. Erst nach der dadurch möglichen Erweiterung des Algorithmus um die Dimension des Störempfindens nämlich, kann eine endgültige Bewertungsmetrik erstellt werden.

# A Auswertung

Auswertung der Klassen 1-6:

Sender	Feature-Raum		
	Feature26	FeatureMax	FeatureSum
Ö1	29 %	54 %	50 %
Ö3	47 %	44 %	47 %

Tabelle A.1: Richtig erkannte Klassen (nach Sender)

Inhalt	Feature-Raum		
	Feature26	FeatureMax	FeatureSum
Musik	38 %	47 %	46 %
Sprache	45 %	56 %	60 %
Doku	9 %	32 %	61 %
Jingle	39 %	52 %	43 %

Tabelle A.2: Richtig erkannte Klassen (nach Inhalt)

Auswertung der Störklassen 2-6:

Sender	Feature-Raum		
	Feature26	FeatureMax	FeatureSum
Ö1	27 %	31 %	39 %
Ö3	34 %	38 %	34 %

Tabelle A.3: Richtig erkannte Störklassen (nach Sender)

Inhalt	Feature-Raum		
	Feature26	FeatureMax	FeatureSum
Musik	30 %	36 %	37 %
Sprache	37 %	33 %	33 %
Doku	29 %	31 %	42 %
Jingle	26 %	33 %	19 %

Tabelle A.4: Richtig erkannte Störklassen (nach Inhalt)

## Confusion-Matrix eindimensionales Feature

Confusion-Matrix Ö1

Ö1		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	4297	2690	3098	36	75	2
	Kl. 2	32	57	66	5	73	0
	Kl. 3	52	75	79	10	57	0
	Kl. 4	35	4	10	0	4	5
	Kl. 5	93	126	189	24	307	0
	Kl. 6	1	0	2	7	1	4

Confusion-Matrix Ö3

Ö3		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	7339	3438	3055	11	1303	18
	Kl. 2	44	135	114	18	41	0
	Kl. 3	90	222	287	21	231	4
	Kl. 4	20	35	7	4	56	6
	Kl. 5	20	53	12	45	60	0
	Kl. 6	4	3	4	7	2	14

Tabelle A.5: Confusion-Matrix des eindimensionalen Features (nach Sender)

Confusion-Matrix

Musik		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	8903	2949	4801	17	1457	33
	Kl. 2	83	110	138	3	75	3
	Kl. 3	171	213	330	7	225	2
	Kl. 4	4	8	5	0	0	0
	Kl. 5	384	108	173	2	89	0
	Kl. 6	2	6	9	0	3	0

Confusion-Matrix

Sprache		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	4797	1260	1431	2	805	0
	Kl. 2	97	26	30	5	14	0
	Kl. 3	87	18	21	4	14	0
	Kl. 4	2	5	9	1	56	0
	Kl. 5	61	23	21	79	11	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Doku		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	328	285	235	1	211	0
	Kl. 2	8	10	7	0	3	0
	Kl. 3	2	5	2	0	3	0
	Kl. 4	0	2	0	0	0	0
	Kl. 5	8	8	6	0	2	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Jingle		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	700	199	311	1	202	2
	Kl. 2	12	3	4	0	1	0
	Kl. 3	9	1	4	0	2	0
	Kl. 4	0	0	0	0	3	0
	Kl. 5	9	3	12	0	4	0
	Kl. 6	0	0	0	0	0	0

Tabelle A.6: Confusion-Matrix des eindimensionalen Features nach Inhalt

## Confusion-Matrix zweidimensionales Feature

Confusion-Matrix Ö1

Ö1		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	6863	2017	2585	11	89	1
	Kl. 2	156	63	37	3	14	0
	Kl. 3	155	22	82	2	9	0
	Kl. 4	40	3	0	0	0	0
	Kl. 5	431	93	86	14	378	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix Ö3

Ö3		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	7865	2776	4193	10	1386	34
	Kl. 2	44	185	22	7	33	0
	Kl. 3	114	196	296	16	214	2
	Kl. 4	37	27	3	4	51	0
	Kl. 5	31	41	86	2	68	0
	Kl. 6	2	6	9	0	2	1

Tabelle A.7: Confusion-Matrix des zweidimensionalen Features nach Sender

Confusion-Matrix

Musik		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	8903	2949	4801	17	1457	33
	Kl. 2	83	110	138	0	75	0
	Kl. 3	171	213	330	7	225	2
	Kl. 4	0	0	0	0	0	0
	Kl. 5	384	108	173	2	89	0
	Kl. 6	2	6	9	0	3	0

Confusion-Matrix

Sprache		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	4797	1260	1431	2	805	0
	Kl. 2	97	26	30	5	14	0
	Kl. 3	87	18	21	4	14	0
	Kl. 4	2	5	9	1	56	0
	Kl. 5	61	23	21	79	11	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Doku		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	328	285	235	1	211	0
	Kl. 2	8	10	7	0	3	0
	Kl. 3	2	5	2	0	3	0
	Kl. 4	0	2	0	0	0	0
	Kl. 5	8	8	6	0	2	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Jingle		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	700	199	311	1	202	2
	Kl. 2	12	3	4	0	1	0
	Kl. 3	9	1	4	0	2	0
	Kl. 4	0	0	0	0	3	0
	Kl. 5	9	3	12	0	4	0
	Kl. 6	0	0	0	0	0	0

Tabelle A.8: Confusion-Matrix des zweidimensionalen Features nach Inhalt

### Confusion-Matrix 26-dimensionales Feature

Confusion-Matrix Ö1

Ö1		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	4061	3989	3956	8	1752	0
	Kl. 2	71	92	80	0	44	0
	Kl. 3	68	81	90	0	41	0
	Kl. 4	31	12	0	0	0	0
	Kl. 5	143	238	254	0	81	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix Ö3

Ö3		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	7333	3304	3247	3	1242	35
	Kl. 2	38	127	120	0	49	0
	Kl. 3	79	403	242	1	111	0
	Kl. 4	52	3	0	0	0	0
	Kl. 5	36	75	62	0	35	0
	Kl. 6	4	0	1	0	0	15

Tabelle A.9: Confusion-Matrix des 26-dimensionalen Features nach Sender

Confusion-Matrix

Musik		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	7071	4411	5096	9	1538	35
	Kl. 2	40	155	153	0	58	0
	Kl. 3	77	448	299	1	123	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	137	249	283	0	87	0
	Kl. 6	4	0	1	0	0	15

Confusion-Matrix

Sprache		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	3716	2116	1479	1	983	0
	Kl. 2	59	55	28	0	25	0
	Kl. 3	61	34	26	0	19	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	28	47	22	0	19	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Doku		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	74	391	285	1	309	0
	Kl. 2	5	6	8	0	9	0
	Kl. 3	0	1	3	0	8	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	3	9	5	0	7	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Jingle		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	533	375	343	0	164	0
	Kl. 2	5	3	11	0	1	0
	Kl. 3	9	1	4	0	2	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	11	8	6	0	3	0
	Kl. 6	0	0	0	0	0	0

Tabelle A.10: Confusion-Matrix des 26-dimensionalen Features nach Inhalt

## Delta-Features

Auswertung der Klassen 1-6:

Sender	Feature-Raum		
	Delta26	DeltaMax	DeltaSum
Ö1	24 %	50 %	92 %
Ö3	45 %	40 %	92 %

Tabelle A.11: Richtig erkannte Klassen nach Sender

Inhalt	Feature-Raum		
	Delta26	DeltaMax	DeltaSum
Musik	35 %	42 %	90 %
Sprache	37 %	51 %	95 %
Doku	10 %	32 %	95 %
Jingle	40 %	49 %	96 %

Tabelle A.12: Richtig erkannte Klassen nach Inhalt

Auswertung der Störklassen 2-6:

Sender	Feature-Raum		
	Delta26	DeltaMax	DeltaSum
Ö1	38 %	39 %	0 %
Ö3	37 %	36 %	0 %

Tabelle A.13: Richtig erkannte Störklassen nach Sender

Inhalt	Feature-Raum		
	Delta26	DeltaMax	DeltaSum
Musik	37 %	38 %	0 %
Sprache	39 %	33 %	0 %
Doku	38 %	30 %	0 %
Jingle	41 %	36 %	0 %

Tabelle A.14: Richtig erkannte Störklassen nach Inhalt

### Confusion-Matrix

Zweidimensionaler Feature-Raum:

Confusion-Matrix Ö1

Ö1		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	3165	3341	5172	24	2063	1
	Kl. 2	69	74	97	1	46	0
	Kl. 3	75	66	96	32	41	0
	Kl. 4	0	55	20	4	42	0
	Kl. 5	132	156	315	2	111	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix Ö3

Ö3		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	9101	2259	2859	9	919	17
	Kl. 2	136	48	72	61	46	0
	Kl. 3	246	141	248	7	151	0
	Kl. 4	0	36	20	6	28	0
	Kl. 5	70	31	76	12	31	0
	Kl. 6	2	2	2	0	0	14

Tabelle A.15: Erkannte Klassen mit zweidimensionalen Delta-Feature nach Sender

Confusion-Matrix

Musik		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	8877	4132	2804	10	2325	12
	Kl. 2	225	84	56	0	41	0
	Kl. 3	267	220	236	0	223	2
	Kl. 4	0	0	0	0	0	0
	Kl. 5	156	157	220	0	222	1
	Kl. 6	0	3	8	0	4	5

Confusion-Matrix

Sprache		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	5223	1531	892	3	644	2
	Kl. 2	98	29	16	0	24	0
	Kl. 3	37	35	32	0	36	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	55	24	22	0	15	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Doku		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	665	189	109	1	96	0
	Kl. 2	19	3	4	0	2	0
	Kl. 3	8	2	0	0	2	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	11	4	5	0	4	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Jingle		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	738	312	192	3	173	0
	Kl. 2	14	4	1	0	1	0
	Kl. 3	4	4	6	0	2	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	4	15	7	0	2	0
	Kl. 6	0	0	0	0	0	0

Tabelle A.16: Erkannte Klassen mit zweidimensionalen Delta-Feature nach Inhalt

Vierdimensionaler Feature-Raum:

Confusion-Matrix Ö1

Ö1		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	7863	2017	2585	11	1289	1
	Kl. 2	156	44	57	0	30	0
	Kl. 3	155	41	61	2	21	0
	Kl. 4	40	5	0	0	0	0
	Kl. 5	431	101	126	0	58	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix Ö3

Ö3		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	6865	2676	4193	10	1386	34
	Kl. 2	44	105	122	0	63	0
	Kl. 3	114	196	296	5	223	2
	Kl. 4	37	7	0	0	0	0
	Kl. 5	31	41	86	2	48	0
	Kl. 6	2	6	9	0	3	0

Tabelle A.17: Erkannte Klassen mit vierdimensionalen Delta-Feature nach Sender

Confusion-Matrix

Musik		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	4754	4734	4745	72	3834	21
	Kl. 2	28	121	135	0	122	0
	Kl. 3	53	247	289	5	354	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	38	139	211	7	360	1
	Kl. 6	2	2	3	0	1	12

Confusion-Matrix

Sprache		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	2854	2082	1929	9	1421	0
	Kl. 2	44	39	45	0	39	0
	Kl. 3	22	37	43	0	38	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	19	35	36	0	26	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Doku		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	33	443	354	0	230	0
	Kl. 2	3	10	9	0	6	0
	Kl. 3	0	2	5	0	5	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	0	6	11	0	7	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Jingle		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	264	433	431	1	283	3
	Kl. 2	3	8	5	0	4	0
	Kl. 3	7	0	4	0	5	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	2	7	11	0	8	0
	Kl. 6	0	0	0	0	0	0

Tabelle A.18: Erkannte Klassen mit vierdimensionalen Delta-Feature nach Inhalt

52-dimensionaler Feature-Raum:

Confusion-Matrix Ö1

Ö1		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	3164	3606	3204	6	3777	0
	Kl. 2	54	92	76	0	64	0
	Kl. 3	69	66	67	0	78	0
	Kl. 4	23	6	0	0	0	0
	Kl. 5	111	224	150	0	231	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix Ö3

Ö3		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l l	Kl. 1	7333	3304	3247	3	1242	35
	Kl. 2	38	127	120	0	49	0
	Kl. 3	79	403	242	1	11	0
	Kl. 4	52	6	0	0	0	0
	Kl. 5	36	75	62	0	35	0
	Kl. 6	4	0	1	0	0	15

Tabelle A.19: Erkannte Klassen mit 52-dimensionalen Delta-Feature (nach Sender)

Confusion-Matrix

Musik		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	5410	5216	3569	40	3907	18
	Kl. 2	31	154	117	1	103	0
	Kl. 3	55	305	324	2	261	1
	Kl. 4	0	0	0	0	0	0
	Kl. 5	29	214	210	5	298	0
	Kl. 6	0	4	0	0	2	14

Confusion-Matrix

Sprache		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	2736	2614	992	10	1943	0
	Kl. 2	48	44	32	0	43	0
	Kl. 3	25	42	34	1	38	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	6	32	22	0	56	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Doku		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	4	270	251	0	535	0
	Kl. 2	0	10	6	0	12	0
	Kl. 3	0	5	3	0	4	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	0	7	4	0	13	0
	Kl. 6	0	0	0	0	0	0

Confusion-Matrix

Jingle		I S T					
		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
S o l i	Kl. 1	137	572	398	1	305	2
	Kl. 2	1	11	5	0	3	0
	Kl. 3	1	7	2	0	6	0
	Kl. 4	0	0	0	0	0	0
	Kl. 5	1	10	5	0	12	0
	Kl. 6	0	0	0	0	0	0

Tabelle A.20: Erkannte Klassen mit 52-dimensionalen Delta-Feature nach Inhalt

# B UKW-Radio

## B.1 Einführung UKW-Rundfunk

Für den UKW-Rundfunk wurde vor 1945 der Frequenzbereich zwischen 42 MHz und 50 MHz benutzt. Erst danach, als man die höheren Frequenzbereiche besser beherrschte, wurde der Frequenzbereich auf Vorschlag der Federal Communications Commission (FCC) auf den höheren, heute noch gültigen Frequenzbereich zwischen 88 MHz und 108 MHz angehoben.

Übertragungstechnisch arbeitet UKW-Rundfunk mit Frequenzmodulation und belegt eine Bandbreite von 75 kHz. Dabei wird das Audiosignal mit einer Bandbreite zwischen 30 Hz und 15 kHz übertragen. Der Modulationscode für UKW-Rundfunk ist F8E.

In den Anfangsjahren wurde das Audiosignal nur in Mono übertragen, später kam auch Stereo hinzu. Aus Gründen der Kompatibilität musste das Stereosignal innerhalb der vorhandenen Übertragungsbandbreite untergebracht werden. Gleichzeitig musste sichergestellt werden, dass ein Mono-Radio den rechten und linken Tonkanal gemeinsam wiedergibt, ein Stereo-Radio hingegen die beiden Stereokanäle getrennt wiedergeben kann.

Aus den genannten Gründen wird das Stereosignal in zwei Signale unterteilt: das Summensignal aus linkem und rechtem Stereokanal (L+R) und das Differenzsignal (L-R). Das (L+R)-Signal wird in Frequenzmodulation übertragen und ist damit mit Mono- und Stereogeräten empfangbar, das (L-R)-Signal wird in Amplitudenmodulation übertragen. Um Interferenzen und Störungen durch ein das Trägersignal zu unterbinden, sendet man einen Pilotton von 19 kHz, der damit oberhalb der Audio-bandbreite liegt. Dieser 19-kHz-Pilotton erfährt eine Frequenzverdoppelung und wird als Trägerfrequenz dem amplitudenmodulierten aber mit Trägerunterdrückung übertragenem (L-R)-Signal zugesetzt. Dadurch bilden sich die Frequenzbereiche für (L+R) zwischen 30 Hz und 15 kHz aus, während jene für die beiden Seitenbänder der

(L-R)-Signale zwischen 23 kHz und 38 kHz bzw. zwischen 38 kHz und 53 kHz liegen. In den 1970er Jahren hat man bei der dreifachen Frequenz des Pilottons, bei 57 kHz, die ARI-Information bzw. den Radiotext hinzugefügt. Empfangstechnisch gewinnt man die beiden Stereokanäle für links und rechts aus dem (L+R)-Summensignal und dem (L-R)-Differenzsignal durch Addition und Subtraktion.

(QUELLE: Itwissen.info)

## B.2 ISO-Stecker

### Pinbelegung des ISO-Stecker für KFZ-Tuner

Damit nicht jeder Autohersteller bei der Audio-Steckerbelegung sein eigenes Süppchen kocht, wurde eine Norm ins Leben gerufen. Mittlerweile wurde diese auch international von der ISO (International Standardizing Organization) aufgenommen. Mechanischer Aufbau, Größe und Form sowie Pinbelegung sind eigentlich eindeutig festgelegt. Eigentlich deshalb, weil einige Hersteller es sich nicht nehmen lassen, die ISO-Buchsen nach eigenem Ermessen zu beschalten. Audi, VW und Opel beispielsweise vertauschen schon fast in traditioneller Manier Dauer- und Schalt-Plus. Soll ein Nachrüst-Radio funktionieren, müssen Pin 4 und Pin 7 am ISO-Stromstecker vertauscht werden. Ansonsten reagiert das neu erworbene Steuergerät nicht auf das Einschalten durch die Zündung, und auch die Speicherdaten wie Radiosender etc. gehen bei abgezogenem Schlüssel verloren. Deshalb sehen die meisten Radiohersteller mittlerweile ISO-Kabel vor, bei denen sich Dauer- und Schaltplus mittels Rundstecker schnell vertauschen lassen. Die folgende Übersicht zeigt, wie die drei Blöcke der Radio-ISO-Buchsen belegt sind bzw. sein sollten.

#### Block A - Stromversorgung

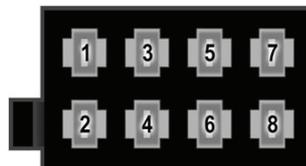


Abbildung B.1: ISO-Stecker; Block A - Pinbelegung

- Pin 1 bis 3 können durchaus vertauscht bzw. anders beschaltet sein; der freie Pin wird evtl. durch ein KFZ-spezifisches Bus-Signal belegt.
  - Pin 4 und 7 sind häufig vertauscht (z.B. Audi, VW und Opel).
  - Pin 5 ist bei neueren VW-Modellen mit einer weiteren Dauerstromversorgung beschaltet. Wird diese Verbindung bei Nachrüst-Radios nicht gekappt, kann das Gerät irreparabel beschädigt werden.
  - Pin 4 (bzw. 7): Bei einigen Automarken (z.B. Audi, VW) genügt bereits das Einstecken des Zündschlüssels, um die Schaltspannung zu aktivieren.
- Ob die eigentliche Stromversorgung über Dauer- oder Schaltplus erfolgt, hängt von Radio- und Autohersteller ab.

PIN	Bezeichnung	Beschreibung
1	Geschwindigkeitssignal	Ein pulsierendes Tachosignal sorgt für konstante Lautstärke oder bedient Navigationssysteme. Auch als Speed-Pulse, GALA (geschwindigkeitsabhängige Lautstärkeanpassung) oder SCV (speed controlled volume) bezeichnet.
2	Telefon-Mute	Schaltet das Radiogerät stumm. Setzt eine Freisprechanlage voraus, die bei einem Anruf Pin 2 auf Masse legt.
3	-	evtl. herstellerspezifisch belegt
4	Dauerplus (Klemme 30)	Hier liegt permanent die Batteriespannung an. Daten wie Radiosender, Klangeinstellungen, Uhrzeit etc. bleiben auch bei ausgeschaltetem Radio gespeichert.
5	(Antennen-) Remote	Gibt das Steuergerät hier 12 Volt aus, fährt die motorbetriebene Antenne aus bzw. bleibt ausgefahren. Auch Zusatzgeräte wie externe Verstärker oder Klangprozessoren etc. werden von dieser Leitung mit Strom versorgt. Einige Radios stellen diese Spannung nur zur Verfügung, solange Radio gehört wird oder auf Verkehrsfunk-Bereitschaft geschaltet ist. Das macht für die Antenne Sinn. Andere Geräte müssen dann aber durch das Schaltplus oder Radio-spezifische Netze (z.B. Alpine Ai-Net, Clarion Ce-NET, Pioneer P-Bus) aktiviert werden.
6	Beleuchtung	Über Pin 6 wird die Beleuchtung von Radiotasten und Display gesteuert.
7	Schaltplus (Klemme 15)	Liegen 12 Volt an, schaltet sich das Radiogerät ein. Wird auch oft allgemein als Zündungsplus bezeichnet. Bei einigen (v.a. japanischen) Autoherstellern gibt es jedoch noch vor der Zündungs-Schlüsselposition die Stellung ACC (engl. accessory = Zubehör).
8	Masse (Klemme 31)	Stellt Kontakt zur Fahrzeugkarosserie her. Ist damit gleichzeitig die Verbindung zum Minuspol der Batterie.

Tabelle B.1: ISO-Stecker; Block A - Stromversorgung

### Block B - Lautsprecher

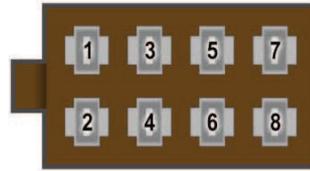


Abbildung B.2: ISO-Stecker; Block B - Pinbelegung

<b>PIN</b>	<b>Beschreibung</b>
1	rechts hinten +
2	rechts hinten -
3	rechts vorne +
4	rechts vorne -
5	links vorne +
6	links vorne -
7	links hinten +
8	links hinten -

Abbildung B.3: ISO-Stecker; Block B - Lautsprecher

Bei Block B – falls dieser vorhanden ist – herrscht bei den Autoherstellern hinsichtlich der Belegung ausnahmsweise Einigkeit.

### Block C - Erweiterungsanschluss

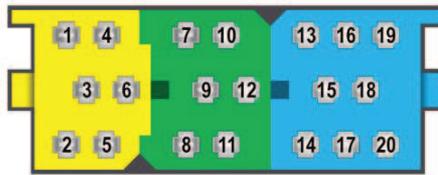


Abbildung B.4: ISO-Stecker; Block C - Pinbelegung

PIN	Beschreibung
1	links hinten Line-Out (VV)
2	rechts hinten Line-Out (VV)
3	Signal-Masse (VV)
4	links vorne Line-Out (VV)
5	rechts vorne Line-Out (VV)
6	Remote / Einschalt-Plus
7 - 20	evtl. herstellerspezifisch belegt (z.B. CD-Wechsler-Steuerung und Eingang, Telefon-Eingang, CAN-Bus etc.)

Abbildung B.5: ISO-Stecker; Block C - Erweiterungsanschluss

(Quelle: selfmadehifi.de)

## B.3 Teststrecke

MSF-Teststrecke nahe Graz (Raaba/Autal)



Abbildung B.6: Teststrecke zur KFZ-Tuner Bewertung

## B.4 Universal Receiver Tester (URT)

Verschiedenste Funktionen und Protokolle können hinsichtlich der Generierung und Analyse von HF Signalen, abhängig von Ausbaustufe und verwendeten Softwarepaketen, realisiert werden. Für flexible Test- und Validierungszwecke steht dem Anwender die komplette Bandbreite an Hard- und Softwarepaketen zur Verfügung. Durch den modularen und skalierbaren Aufbau ist der NOFFZ<sup>1</sup>-Autoradio-Tester für den schnellen und automatisierten Produktionstest in noch kompakterer Bauform verfügbar. In ein 8-Slot PXI-Chassis werden Generator, Up-Converter und wahlweise eine Analyseeinheit integriert.



Abbildung B.7: Messung mit dem URT

### Generierung von HF Signalen

Der neue Generator AWG PXIe-5441 zeichnet sich durch einen Onboard-Signalprozessor und integrierten Speicher aus. Abhängig vom verwendeten Softwarepaket können mit ihm die verschiedensten Sender (z.B. 6 UKW-Sender) und Protokolle (UKW/RDS, DAB, IBOC, SDARS) generiert werden. Ganze Senderlandschaften können unter Zuhilfenahme eines Massenspeichers sowohl für den Laborbetrieb als auch für den Produktionstest generiert werden. Hierbei werden kundenspezifische Senderlandschaften in Dateiform gemäß den Anforderungen kreiert und auf dem Massenspeicher gespeichert. Diese Sender lassen sich problemlos über die Generatoreinheit wiedergeben.

<sup>1</sup>NOFFZ: NOFFZ-ComputerTechnik GmbH; siehe auch Aversa (URT)

### **Analyse von HF-Signalen**

Die Analyseeinheit besteht aus einem Vektor-Signal-Analyser, der sich aus Down-converter und Digitizer mit Onboard-Signal-Processing zusammensetzt. Mit dieser Kombination können nicht nur gespeicherte Daten offline analysiert, sondern auch online-Tests durchgeführt werden — beispielsweise direkt auf einer Testfahrt.

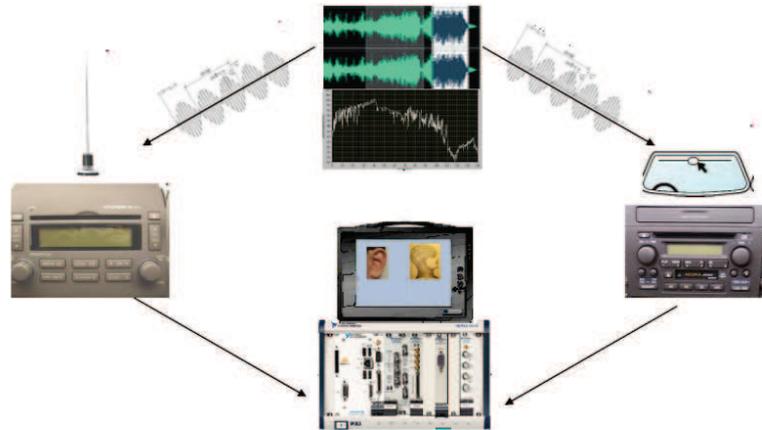
### **Aufnahme und Wiedergabe**

In Verbindung mit einem Massenspeicher über PXIe können aufwändige Validierungsfahrten erheblich reduziert werden. Während einer Testfahrt können nun alle Signale die beispielsweise im UKW-Band (87,5 - 108 MHz, jedoch max. 20 MHz Echtzeitbandbreite) liegen, inklusive RDS-Information mittels High Speed Streaming Technologie aufgenommen werden. Die Aufnahme kann über viele Stunden mit einer Datenrate von bis zu 100 MByte/s durchgeführt werden. Die Begrenzung der Aufnahmen liegt zwischen 99 kHz und 2.7 GHz mit einer Echtzeitbandbreite von 20 MHz. Es können also auch andere Rundfunknetze wie DAB, SDARS oder IBOC aufgezeichnet werden. Außerdem können mit der verwendeten Hardware und andere Funkprotokolle wie Bluetooth oder WLAN getestet werden.

Das gesamte System kann entsprechend den jeweiligen kundenspezifischen Anforderungen konfiguriert werden, von der Signalkonditionierung bis hin zur Datenspeicherung.

## B.5 ServiceForce

### Testarchitektur mit PEAQ



Quelle:VIP08 - ServiceForce

Abbildung B.8: Messung mit Messsetup von ServiceForce

### Testarchitektur mit PEAQ

- NI 9234 A/D-Wandler

- 4 Kanäle,  $\pm 5$  V, 51,2 kS/s pro Kanal, 24-bit-Auflösung



- Maximale Abtastrate von 51,2 kS/s, AC-gekoppelt (0,5 Hz)
  - Über Software wählbare IEPE-Signalkonditionierung (0 bzw. 2 mA)
  - 24-bit-Auflösung, Dynamikbereich von 102 dB, Antialiasing-Filter
  - 4 simultan abgetastete Analogeingänge, Eingangsbereich von  $\pm 5$  V

Quelle:VIP08 - ServiceForce

Abbildung B.9: Zugrundeliegendes Verfahren des Messsetup von ServiceForce

# Abbildungsverzeichnis

1.1	Einfache Bewertungsmetrik, nach Schulnotensystem zur Tunerqualifikation . . . . .	6
1.2	Struktogramm Tunerqualifizierung . . . . .	6
2.1	Euklidische Distanz in einem 2-D-Raum . . . . .	11
2.2	NN-Klassifikation in einem 2-D-Feature-Raum . . . . .	11
2.3	k-means Klassifizierung im 2-D-Feature-Raum . . . . .	13
2.4	Prozess einer Mustererkennung . . . . .	13
2.5	Vorverarbeitung zur Datenanalyse mit Feature-Extraktion . . . . .	14
2.6	Entwicklungsprozess der Mustererkennung [TheoKon99] . . . . .	14
2.7	Confusion-Matrix mit Soll-Klassen aus Labelung und Ist-Klassen aus Klassifikation . . . . .	16
2.8	Melskala mit 26 Frequenzbändern . . . . .	21
2.9	Zusammensetzung eines 52D-Delta-Feature aus 26D-Feature . . . . .	22
3.1	Testsetup mit HF-Sender und Übertragungsstrecke . . . . .	26
3.2	Testsignal - Device under Test (DUT) . . . . .	26
3.3	Referenzsignal aus Referenz tuner . . . . .	27
3.4	Referenzstation der Voruntersuchung . . . . .	28
3.5	Elektrischer Verlauf des Messsignals . . . . .	30
3.6	$\frac{\lambda}{4}$ -Antenne der Referenzstation . . . . .	31
3.7	Struktur und Dateinamen nach Labelung . . . . .	34
3.8	Datentripel eines Files der Länge N ms . . . . .	35
4.1	Algorithmus zur Klassifikation . . . . .	36
4.2	Algorithmus zur Feature-Extraktion . . . . .	37
4.3	Synchronisation von Test- und Referenzsignal mit Kreuzkorrelation . . . . .	39
4.4	Im Spektrum zwischen 15 und 21 kHz sind Störungen ersichtliche . . . . .	40
4.5	Überlappende Blöcke mit Fensterfunktion zur Glättung des Spektrums . . . . .	41

---

4.6	Zusammenfassen der Energie in Melbänder . . . . .	43
4.7	Differenzbildung von Test- und Referenzsignal der einzelnen Melbänder	45
4.8	Clustering mit gelabelten Daten . . . . .	48
4.9	Überlappung der Frames bei der Bufferung . . . . .	49
5.1	Confusion-Matrix mit Sub-Matrix . . . . .	59
B.1	ISO-Stecker; Block A - Pinbelegung . . . . .	78
B.2	ISO-Stecker; Block B - Pinbelegung . . . . .	80
B.3	ISO-Stecker; Block B - Lautsprecher . . . . .	80
B.4	ISO-Stecker; Block C - Pinbelegung . . . . .	81
B.5	ISO-Stecker; Block C - Erweiterungsanschluss . . . . .	81
B.6	Teststrecke zur KFZ-Tuner Bewertung . . . . .	82
B.7	Messung mit dem URT . . . . .	83
B.8	Messung mit Messsetup von ServiceForce . . . . .	85
B.9	Zugrundeliegendes Verfahren des Messsetups von ServiceForce . . . . .	85

# Tabellenverzeichnis

3.1	Aufteilung des gesamten Datenmaterials . . . . .	32
4.1	Störklassen zur Sortierung der Feature-Vektoren . . . . .	47
4.2	Anzahl richtig erkannter Frames in Abhängigkeit von der Größe der Trainingsbasis. . . . .	50
5.1	Auswertung der Testdateien — richtig erkannte Frames (Angaben in Prozent) . . . . .	52
5.2	Richtig erkannte Klassen nach Sendern . . . . .	53
5.3	Auswertung der drei Feature-Räume im Vergleich . . . . .	53
5.4	Auswertung der richtig erkannten Frames der Klassen 2-6 (Störungen) . . . . .	54
5.5	Erkannte Klassen mit eindimensionalem Feature . . . . .	55
5.6	Erkannte Klassen mit zweidimensionalem Feature . . . . .	56
5.7	Erkannte Klassen mit 26-dimensionalem Feature . . . . .	57
5.8	Erkannte Klassen mit 52-dimensionalen Delta-Feature . . . . .	58
A.1	Richtig erkannte Klassen (nach Sender) . . . . .	62
A.2	Richtig erkannte Klassen (nach Inhalt) . . . . .	62
A.3	Richtig erkannte Störklassen (nach Sender) . . . . .	62
A.4	Richtig erkannte Störklassen (nach Inhalt) . . . . .	63
A.5	Confusion-Matrix des eindimensionalen Features (nach Sender) . . . . .	63
A.6	Confusion-Matrix des eindimensionalen Features nach Inhalt . . . . .	64
A.7	Confusion-Matrix des zweidimensionalen Features nach Sender . . . . .	65
A.8	Confusion-Matrix des zweidimensionalen Features nach Inhalt . . . . .	66
A.9	Confusion-Matrix des 26-dimensionalen Features nach Sender . . . . .	67
A.10	Confusion-Matrix des 26-dimensionalen Features nach Inhalt . . . . .	68
A.11	Richtig erkannte Klassen nach Sender . . . . .	69
A.12	Richtig erkannte Klassen nach Inhalt . . . . .	69
A.13	Richtig erkannte Störklassen nach Sender . . . . .	69

---

A.14 Richtig erkannte Störklassen nach Inhalt . . . . .	70
A.15 Erkannte Klassen mit zweidimensionalen Delta-Feature nach Sender .	70
A.16 Erkannte Klassen mit zweidimensionalen Delta-Feature nach Inhalt .	71
A.17 Erkannte Klassen mit vierdimensionalen Delta-Feature nach Sender .	72
A.18 Erkannte Klassen mit vierdimensionalen Delta-Feature nach Inhalt .	73
A.19 Erkannte Klassen mit 52-dimensionalen Delta-Feature (nach Sender)	74
A.20 Erkannte Klassen mit 52-dimensionalen Delta-Feature nach Inhalt . .	75
B.1 ISO-Stecker; Block A - Stromversorgung . . . . .	79

# Literaturverzeichnis

- [OS01] Alan V. Oppenheim, Ronald W. Schaffer. *Zeitdiskrete Signalverarbeitung*, 3. Auflage, Oldenburg, 2001
- [KamKro02] Karl-Dirk Kammeyer, Kristian Kroschl. *Digitale Signalverarbeitung; Filterung und Spektralanalyse mit Matlab-Übungen*, 5. Auflage, 2002
- [Duda07] Richard O. Duda, Peter E. Hart, David G. Stork. *Pattern Classification*, Second Edition, 2002
- [TheoKon99] Sergios Theodoridis, Konstantinos Koutroumbas. *Pattern Classification*, 1999
- [ThuMid85] Marlin O. Thurston, William Middendorf. *Pattern Recognition*, 1985
- [Bishop] Christopher M. Bishop *Pattern Recognition and Machine Learning*, Springer Berlin, 2007
- [FrPer] Franz Pernkopf, *Klassifikation*, Technische Universität Graz, SPSC, 2002
- [ITU-R] International Telecommunication Union (1998-2002) *Methoden zur objektiven Messung der wahrgenommenen Audioqualität gemäß ITU-R BS.1387-1*, 1998
- [GePe] Geoffroy Peeters *A large set of audio features for sound description (similarity and classification) in the CUIDADO project*, IRCAM, Paris, April 2004
- [BreMcKi01] Jeroen Breebaart, Martin McKinney *Features for Audio Classification*, Philips Research Laboratories, Eindhoven, Netherlands, 2001

- [Mel02] Sigurdur Sigurdsson, Kaare Brandt Petersen, Tue Lehn-Schiøler *Mel Frequency Cepstral Coefficients: An Evaluation of Robustness of MP3 Encoded Music*, Technical University of Denmark, 2002
- [AudRec] Antti J. Eronen, Vesa T. Peltonen, Juha T. Tuomi, Anssi P. Klapuri, Seppo Fagerlund, Timo Sorsa, Gaetan Lorho, Jyri Huopaniemi *Audio-Based Context Recognition*, IEEE Vol. 14, NO. 1, January 2006
- [WesCox] Kris West, Stephen Cox *Finding an optimal segmentation for audio genre classification*, University of East Anglia, 2003
- [SpeCon] Dan-Ning Jiang, Lie Lu, Hong-Jiang Zhang, Jian-Hua Tao, Lian-Hong Cai *Music Type Classification by spectral contrast feature*, Tsinghua University, China; Department of Computer Science and Technology
- [spePeak] Miruslav Zivanovic, Axel Roebel, Xavier Rodet *A new approach to spectral peak classification*, University of Navarra; IRCAM, France
- [EsqKar] Paulo Esquef, Matti Karjalainen, Vesa Valimäki *Detection of clicks in audio signals using warped linear prediction*, Helsinki University of Technology; Tampere University of Technology
- [SpecEst] Kamil K Wojcicki, Benjamin J. Shannon, Kuldip K. Paliwal *Spectral Subtraction with Variance Reduced Noise Spectrum Estimates*, Griffith University, Nathan, Australia