

Diplomarbeit

Akustische Güteprüfung
durch Schallemissions-
untersuchung mit
Fuzzy Clusteranalyse

Christian Leubner

Diplomarbeit
am Fachbereich Informatik
der Universität Dortmund

20. Juli 1999

Betreuer:

Prof. Dr. Bernd Reusch
Dipl.-Inform. Lars Hildebrand

Inhaltsverzeichnis

1	Einleitung	4
1.1	Akustische Güteprüfung	4
1.2	Vorteile eines objektiven Prüfsystems	5
1.3	Akustische Güteprüfung bei Mannesmann Dematic	6
2	Schall und seine Darstellung	8
2.1	Harmonische Schwingungen	8
2.2	Frequenzbetrachtung	9
2.3	Schallstärke	12
2.4	Das menschliche Gehör	13
2.5	Digitale Schallaufnahme	15
3	Einführung in Mustererkennungssysteme	16
3.1	Aufbau	16
3.2	Aktueller Stand und Fokussierung	18
3.3	Verwendung von Fuzzy Logik	19
3.4	Technische Entwicklung	22
4	Herleitung des Mustererkennungssystems	24
4.1	Grundlegende Betrachtung von Merkmalen	24
4.2	Signalverarbeitung	28
4.2.1	Leistungsdichtespektrum	29

4.2.2	Normierung	31
4.3	Unschärfes Klassifikationskonzept	33
4.3.1	Distanzma und Zugehörigkeitsfunktion	33
4.3.2	Hauptachsentransformation in das Cluster	36
4.3.3	Entscheidung über das Klassifikationsergebnis	41
4.4	Alternative Klassifikationskonzepte	42
4.4.1	Linearer Klassifikator	42
4.4.2	Allgemeine Entscheidungsfunktionen	43
4.5	Merkmalsextraktion	44
4.5.1	Gütekriterien	45
4.5.2	Auswahlverfahren	49
4.5.3	Laufzeitbetrachtung	53
4.5.4	Überadaption	57
5	Software	60
5.1	WAV Analyse	60
5.1.1	Funktionsumfang	60
5.1.2	Bedienung	61
5.2	Klassifikations Tool	65
5.2.1	Funktionsumfang	65
5.2.2	Bedienung	65
5.3	Implementierung	69
5.3.1	Die Klassen <i>CMerkmalssatz</i> und <i>CAuswahl</i>	70
5.3.2	Klasse <i>CKlasse</i>	70
5.3.3	Klasse <i>CKlassifikator</i>	70
5.3.4	Die Klassen <i>CWAVFile</i> und <i>CBMPFile</i>	71

6	Umsetzung des Verfahrens von Mannesmann Dematic	72
6.1	Aufgabenstellung	72
6.1.1	Ausgangssituation	72
6.1.2	Versuchsaufbau	73
6.1.3	Signalverarbeitung	75
6.1.4	Merkmalsextraktion	76
6.2	Ergebnisse	77
6.2.1	Merkmalsanzahl N	82
6.2.2	Parameter a	82
6.2.3	Parameter b	83
6.2.4	Parameter t	83
6.3	Ergebnis	84
7	Ausblick	85

Kapitel 1

Einleitung

1.1 Akustische Güteprüfung

Bei der akustischen Güteprüfung handelt es sich um ein Verfahren, das durch die Untersuchung der Schallemission eines Testobjektes eine Qualitätsbeurteilung ermöglicht. Der Einsatz von akustischen Güteprüfmethoden ist überall dort möglich, wo sich die Qualität eines Testobjektes auf seine Schallemission auswirkt. Bei einigen Produkten oder Verschleißteilen, die im Betrieb oder durch Einwirkung von außen ein Geräusch erzeugen, kann ein Prüfer durch genaues Hinhören bereits erkennen, ob das Testobjekt einwandfrei ist oder ob ein Schaden vorliegt. Durch die Zuhilfenahme von technischen Geräten kann auch eine Beurteilung von Schwingungen im für Menschen nicht-hörbaren Bereich (zum Beispiel Ultraschall) erfolgen. Ziel ist in allen Fällen die Qualitätsbeurteilung von Testobjekten und gegebenenfalls die Einteilung der Testobjekte in Fehlerklassen.

Die akustische Güteprüfung wird vor allem bei der Endprüfung von maschinell hergestellten Produkten eingesetzt. Anhand direkter Schallemission wie zum Beispiel bei Motoren oder durch eine extern herbeigeführte Schallemission wie in dem hier untersuchten Fall von Mannesmann Dematic, wird durch einen in der Regel menschlichen Prüfer eine Aussage über die Fertigungsqualität des Produktes gefällt. Die Schallemission ist dabei lediglich als Übermittler eines physikalischen Zustandes zu betrachten. Ziel ist es, eine Klassifizierung der Objekte durchzuführen, um defekte oder fehlerhafte Prüfobjekte zu erkennen.

In der Vergangenheit wurde häufig erfolgreich versucht, bei akustischen Güteprüfungen den menschlichen, subjektiv urteilenden Prüfer durch ein automatisiertes Verfahren zu erset-

zen. Bereits 1983 wurde am Fraunhofer Institut für Informations- und Datenverarbeitung, Karlsruhe, ein Prüfsystem für Elektromotoren entwickelt, das seinerzeit noch mit speziellen Hardwareanfertigungen realisiert wurde [9]. Später wurden für weitere Anwendungen, wie zum Beispiel die Diagnose von Schäden an Hydraulikpumpen [14] oder die Gütesicherung bei Laserpunktschweißungen [16], Konzepte entworfen und realisiert. Einen guten Überblick für den Bereich der Diagnose- und Überwachungssysteme liefert Strackeljan in seiner Dissertation [17].

1.2 Vorteile eines objektiven Prüfsystems

Akustische Prüfverfahren werden größtenteils durch menschliche Prüfer durchgeführt, die aufgrund ihrer Erfahrung die Schallemission beurteilen. Dadurch unterliegen die Ergebnisse aber auch der persönlichen Subjektivität des durchführenden Prüfers. Da sich menschliche Hörempfindungen im Laufe der Zeit ändern, kann es vorkommen, daß das selbe Prüfobjekt durch den selben Prüfer nach einigen Monaten unterschiedlich beurteilt wird. Gerade im Bereich der nicht melodischen Geräusche und Klänge ist das menschliche Erinnerungsvermögen wenig ausgeprägt [9]. Unterschiedliche Beurteilungen treten insbesondere auch dann auf, wenn mehrere Prüfer die gleichen Objekte untersuchen und zu unterschiedlichen Prüfergebnissen gelangen. Durch die Verwendung eines einheitlichen Prüfsystems erhöht sich mithin die Prüfsicherheit, da subjektive Einflüsse ausgeblendet werden können.

Für ein objektives Prüfsystem sprechen weitere Gründe. Zum einen ist der Schutz, vor allem der Gesundheit, des Prüfers zu nennen. Akustische Güteprüfungen müssen häufig in lauten Umgebungen durchgeführt werden. Dabei wird vom Prüfer verlangt, leise Geräusche seines Testobjektes wahrzunehmen, wodurch es ihm nicht möglich ist, beispielsweise einen Gehörschutz zu tragen. Des weiteren kann es vorkommen, daß der Prüfer sich mit seinem Kopf heißen oder schnell rotierenden Objekten nähern muß, so daß Unfälle begünstigt werden. Auch der Zeitdruck bei industriellen Fertigungen kann zu Stresssituationen des Prüfers führen, da er sich dem einzelnen Objekt unter Umständen nicht lange genug widmen kann, wodurch Fehlentscheidungen begünstigt werden. Für ein automatisiertes Prüfsystem sprechen außerdem Gründe der Produktivitätssteigerung: Der Experte kann sich anderen Aufgaben widmen. Seine Abwesenheit durch Urlaub oder Krankheit führt nicht mehr zu Engpässen.

1.3 Akustische Güteprüfung bei Mannesmann Dematic

Bei der Firma Mannesmann Dematic, Wetter/Ruhr, wird neben verschiedenen anderen Güteprüfverfahren auch ein akustisches eingesetzt. Dieses kommt stichprobenartig nach dem Schweißen von Kranbahnschienen mittels Widerstandspunktschweißen zum Einsatz. Die Schnittansicht einer Kranbahnschiene kann Abbildung 1.1 entnommen werden, Abbildung 1.2 zeigt eine Frontalansicht. Der schwarze Kreis markiert in beiden Abbildungen den Schweißpunkt.

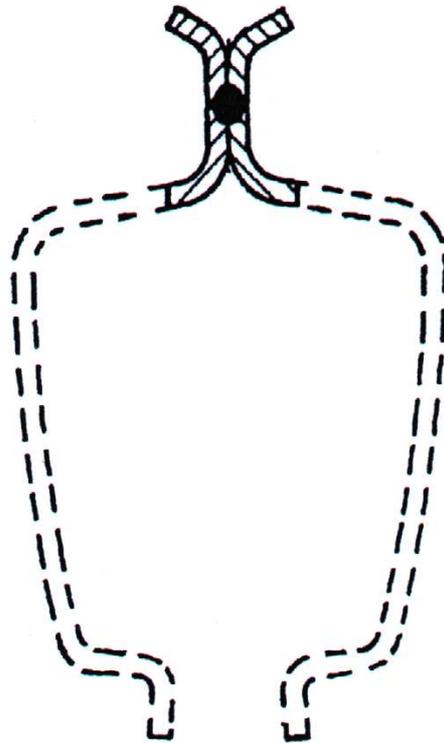


Abbildung 1.1: Schnittansicht der Kranbahnschiene

Bei der akustischen Güteprüfung wird die Schiene mittels eines Brecheisens nicht zerstörend auseinandergebogen. Wenn die Schweißverbindung fehlerfrei ist, tritt keine Schallemission auf. Im Fehlerfall kann jedoch ein charakteristisches "Knistern" wahrgenommen werden, das auf eine schadhafte Verbindung hindeutet.

Bei der Konzeption eines automatisierten Prüfsystems ist den speziellen Rahmenbedingungen dieses Prüfverfahrens Rechnung zu tragen. Im Gegensatz zu anderen bisher in der Literatur untersuchten Verfahren, ist hier nur im Fehlerfall eine Schallemission wahrzunehmen, nämlich das charakteristische Knistern. Auf den ersten Blick scheint die Aufgabe dadurch

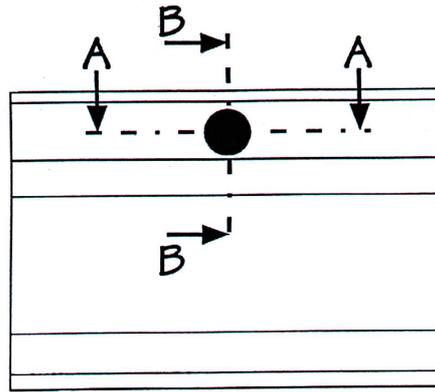


Abbildung 1.2: Frontalansicht der Kranbahnschiene, die Buchstaben A und B markieren hier die Schnittebenen für zerstörende Prüfverfahren

trivialer zu werden, bei genauerer Betrachtung stellt sich allerdings heraus, daß weitere Rahmenbedingungen berücksichtigt werden müssen, die einer simplen Lösung im Wege stehen. In ganz erheblichem Maße ist hier vor allem die vielfältige und teils sehr laute Geräuschkulisse in einer großen Produktionshalle zu nennen. Somit reicht es für ein automatisches Prüfsystem nicht aus, lediglich zwischen Knistern und Stille zu unterscheiden, da ständig sehr laute Nebengeräusche auftreten können. Mithin muß das Ziel sein, ein System zu entwickeln, daß zwischen Knistern und möglichem Umgebungslärm unterscheiden kann. Durch diese Notwendigkeit, das spezifische Knistern erkennen zu müssen, ist die Problemstellung bei der Realisierung ähnlich wie bei anderen aus der Literatur bekannten Verfahren.

Kapitel 2

Schall und seine Darstellung

Schall entsteht immer dort, wo sich etwas bewegt. Seine Ursache ist immer eine mechanische Schwingung, egal ob es sich um Töne, Klänge oder Geräusche allgemeiner Art handelt. Daher kann die Akustik, die Lehre vom Schall, als ein Spezialgebiet der allgemeinen Schwingungslehre betrachtet werden [2].

2.1 Harmonische Schwingungen

In der Schwingungslehre werden häufig die *harmonischen Schwingungen* studiert, die sich aus der gleichförmigen Kreisbewegung eines Körpers ergeben und dementsprechend mit der Sinus Funktion beschrieben werden können:

$$y = A \sin \varphi \tag{2.1}$$

A bezeichnet hier die *Amplitude*, definiert als größter anzunehmender Wert der harmonischen Schwingung. Der Phasenwinkel φ ist von der Zeit abhängig und gibt die aktuelle Position der Kreisbewegung an. Um φ in Abhängigkeit von der Zeit auszudrücken, ist zu berücksichtigen, daß ein voller Kreisumlauf dem Winkel 2π entspricht. Dessen Dauer wird mit T Sekunden bezeichnet. Außerdem beachtet man einen möglichen Nullphasenwinkel φ_0 , der den Phasenwinkel zum Zeitpunkt $t = 0$ angibt. Dann läßt sich φ zur Zeit t Sekunden angeben als:

$$\varphi(t) = \frac{2\pi}{T}t + \varphi_0 \tag{2.2}$$

Die *Frequenz* f einer harmonischen Schwingung bezeichnet die Anzahl der Kreisumläufe pro Sekunde und wird in der Einheit Hertz, abgekürzt Hz, angegeben. Die Dauer eines einzelnen

Kreisumlaufs T und die Frequenz f stehen offensichtlich in der Beziehung

$$f = \frac{1}{T} \quad (2.3)$$

Alles in allem kann somit die Schwingungsgleichung für eine harmonische Schwingung durch die Funktion

$$y(t) = A \sin(2\pi ft + \varphi_0) \quad (2.4)$$

angegeben werden. Darin sind, noch einmal zusammengefaßt:

y : Elongation

A : Amplitude der Schwingung

f : Frequenz

t : Zeit

φ_0 : Nullphasenwinkel

Da die Elongation $y(t)$ bei der harmonischen Schwingung dem Sinus der Zeit proportional ist, spricht man auch von einer *Sinusschwingung*.

2.2 Frequenzbetrachtung

Komplexe Geräusche, wie die hier zu untersuchenden Schallemissionen, bestehen aus einer Vielzahl von Schwingungen mit unterschiedlichen Frequenzen, die sich überlagern. Abbildung 2.1 zeigt drei harmonische Schwingungen mit verschiedenen Frequenzen und Amplituden, sowie das Ergebnis, wenn diese sich überlagern, was mathematisch der Addition der Einzelfunktionen entspricht [14].

Eine komplizierte Gesamtschwingung wieder in ihre Einzelschwingungen zu zerlegen, ist wesentlich aufwendiger. Allerdings hat der französische Physiker und Mathematiker *J.B.J. Fourier* (1768 - 1830) gezeigt, daß sich jede noch so komplizierte Schwingung als Summe von harmonischen Schwingungen darstellen läßt. Die Umrechnung erfolgt mit dem nach ihm benannten und weithin bekannten *Fast Fourier Transform (FFT)* Algorithmus, der es ermöglicht, eine Schwingung nicht nur als Funktion der Zeit darzustellen, sondern auch als Funktion der Frequenz. Das so zu gewinnende *Frequenzspektrum* erlaubt die Betrachtung einer Schallemission aus einer völlig anderen Perspektive. [12]

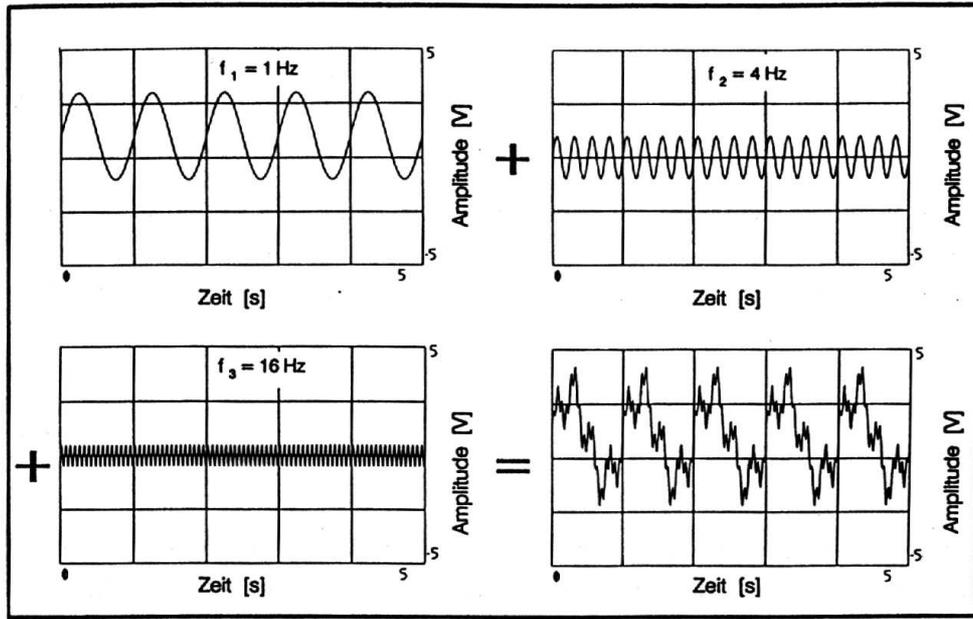


Abbildung 2.1: Überlagerung verschiedener harmonischer Schwingungen

Amplituden-Darstellung von WAV-File: "D:\Diplom\Dematic\knister01.wav"

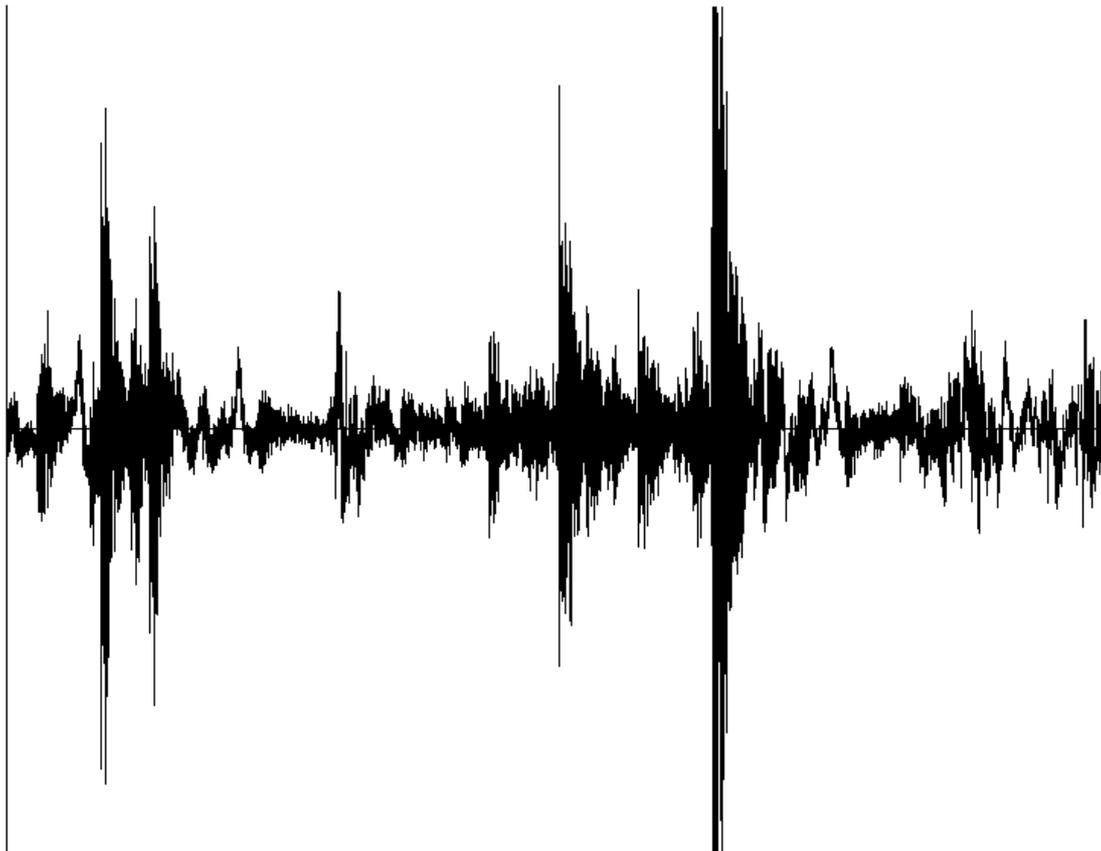


Abbildung 2.2: Zeitsignal eines charakteristischen Knisterns

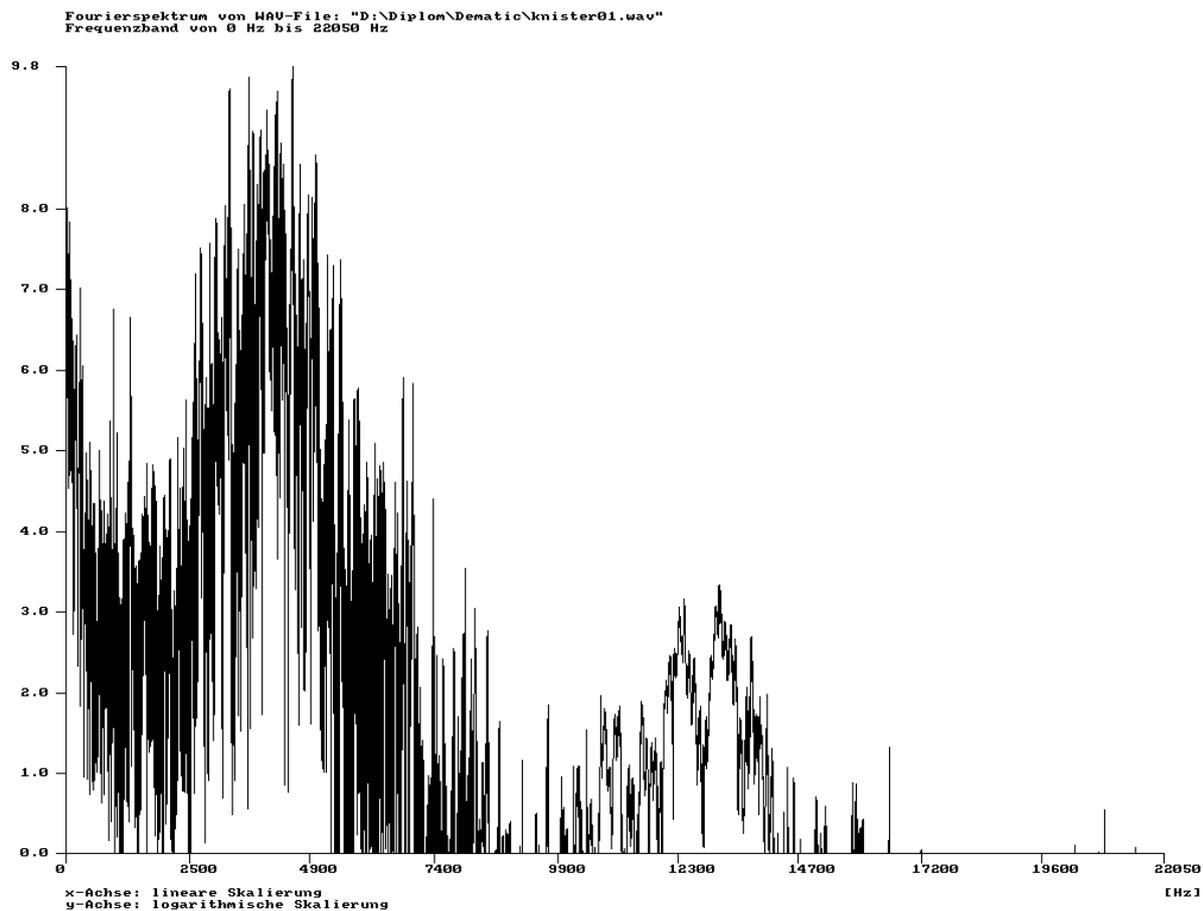


Abbildung 2.3: Frequenzspektrum eines charakteristischen Knisterns

Abbildung 2.2 zeigt beispielhaft das Zeitsignal eines charakteristischen Knisterns, das aus einer digitalen Aufnahme bei Mannesmann Dematic stammt¹. Abbildung 2.3 zeigt dazu ein mittels Fourier Transformation erzeugtes Frequenzspektrum, das aus der Aufnahme berechnet wurde. Im konkreten Fall deutet das Knistern auf eine fehlerhafte Schweißung hin. Es sind deutliche Amplitudenunterschiede zu erkennen, allerdings ist die Schallemission nicht auf ein schmales Frequenzband beschränkt.

2.3 Schallstärke

Schallwellen sind Träger von Energie, der sogenannten *Schallenergie*. Diese wird von der Schallquelle weg transportiert. Die gesamte, in einer Sekunde abgegebene Energie einer Schallquelle, wird als *Schalleistung* bezeichnet und in der Einheit Watt gemessen.

Die Schallenergie, die auf einer senkrecht zur Quelle stehenden Fläche pro Zeiteinheit auftrifft, wird als *Schallstärke* bezeichnet. Sie erhält das Symbol J und die Einheit Watt/cm^2 . Das menschliche Ohr reagiert bereits bei einer Schallstärke von nur $10^{-16} \text{ Watt}/\text{cm}^2$ und löst eine Hörempfindung aus. Die Schmerzgrenze liegt mit $10^{-3} \text{ Watt}/\text{cm}^2$ um einige 10er Potenzen höher. Um dieses große Intervall rechnerisch einfacher erfassen zu können, wird das logarithmische Dezibelmaß verwendet. Die Schallstärkeangabe wird dabei auf die Hörschwelle $10^{-16} \text{ Watt}/\text{cm}^2$, die mit J_0 bezeichnet wird, bezogen. Man erhält die Dezibelangabe für eine bestimmte Schallstärke J , den sogenannten *Schallstärkepegel* L aus der Beziehung

$$L = 10 \log_{10} \frac{J}{J_0} \text{dB} \quad (2.5)$$

Demnach entspricht eine Verzehnfachung der Schallstärke einer Zunahme des Schallstärkepegels um 10 dB. An der Hörschwelle beträgt er folglich 0 dB und an der Schmerzgrenze entsprechend 130 dB. Wichtig ist, daß die jeweilige Bezugsgröße mit angegeben wird, hier also $10^{-16} \text{ Watt}/\text{cm}^2$.

Bei digitalen Aufzeichnungen wird häufig mit der Dezibelangabe dBFS² gearbeitet, die als Bezugsgröße die digitale Vollaussteuerung, bei 16 Bit also den positiven Maximalwert 32767, hat.

¹Auf der dieser Arbeit beigelegten CD-ROM ist das beispielhaft verwendete Knistern im Verzeichnis "Dematic" als "knister01.wav" zu finden.

²FS: Full Scale

2.4 Das menschliche Gehör

Anatomisch wird das menschliche Ohr in drei Abschnitte unterteilt:

- das äußere Ohr, zu dem auch Ohrmuschel und Gehörgang zählen,
- das Mittelohr mit Trommelfell, Gehörknöchelchenkette und Ohrtrompete,
- sowie das Innenohr.

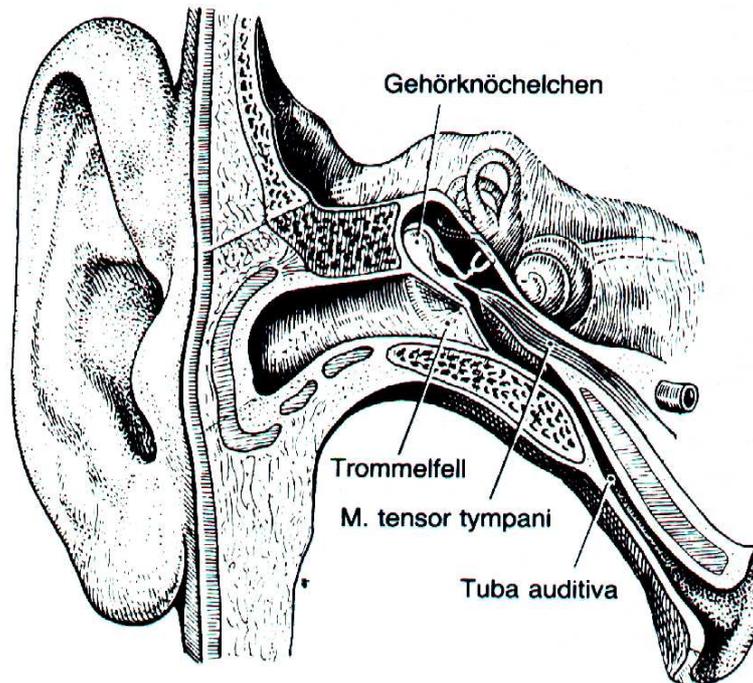


Abbildung 2.4: Äußeres Ohr, Mittel- und Innenohr in schematischer Darstellung

Abbildung 2.4 zeigt eine schematische Darstellung. Schallwellen, die am Ohr eintreffen, gelangen über den äußeren Gehörgang zum Trommelfell, das durch die Schalldruckschwankungen in Schwingungen versetzt wird. Über die Gehörknöchelchenkette wird die Schallenergie an das Innenohr weitergegeben, wo die eigentliche Reizaufnahme erfolgt. Das Innenohr liegt in einer harten Knochenkapsel und besteht aus einem System von Gängen und Hohlräumen, das als *Labyrinth* bezeichnet wird. Dort ist in der sogenannten Schnecke (*Cochlea*) neben dem Gleichgewichtsorgan auch das *Cortische Organ*, das Hörempfindungen verarbeitet, zu finden. Das Cortische Organ besteht aus Rezeptorzellen, die als innere und äußere Haarzellen bezeichnet werden. Die Haarzellen reagieren auf für sie spezifische Frequenzen, so daß es zu einer Entmischung der eingehenden Schallwellen kommt. Dieser Vorgang wird als *Frequenzdispersion* bezeichnet. Wird eine Haarzelle mit ihrer spezifischen Frequenz beschallt, löst sie

ein Aktionspotential aus, das über entsprechende Nervenfasern an das Gehirn weitergeleitet wird. Während es dem Gehirn möglich ist, die Frequenz aufgrund der auslösenden Haarzelle zuzuordnen, wird die Lautstärke durch die Anzahl der Nervenimpulse übermittelt.

Durch beidohriges Hören ist es dem Menschen möglich, auch bei fixierter Kopfstellung die Richtung eines Geräusches zu erkennen. Befindet sich eine Schallquelle seitlich des Kopfes, treffen die Schallwellen zu unterschiedlichen Zeitpunkten und mit unterschiedlicher Intensität an den beiden Ohren an. Richtungsunterschiede von 3° können sicher wahrgenommen werden. Auch die Entfernung einer Schallquelle kann der Mensch unter Ausnutzung seiner Erfahrungen schätzen. Die Lautstärke eines Schalls nimmt ungefähr mit dem Quadrat der Entfernung ab. Des weiteren ändert sich mit zunehmender Entfernung das Frequenzspektrum des Schalls, da hohe Frequenzen stärker als tiefe gedämpft werden. Durch im Laufe seines Lebens gesammelte Erfahrungen kann der Mensch durch diese Informationen die Entfernung einer Schallquelle schätzen. Eine Anwendung ist beispielsweise die Schätzung der Entfernung eines Gewitters, indem die Zeitspanne zwischen Blitz und Donner abgewartet wird. [18]

Das menschliche Gehör ist in der Lage, Frequenzen zwischen ungefähr 16 Hz und 20000 Hz wahrzunehmen. Die obere Hörgrenze nimmt mit zunehmendem Alter ab, bei 35jährigen liegt sie bei ungefähr 15000 Hz und bei 60jährigen noch bei 5000 Hz. Die größte Empfindlichkeit besteht im Bereich von 1000 bis 2000 Hz, hier genügt die minimale Schallstärke von 10^{-16} Watt/cm², die zuvor bereits als Hörschwelle bezeichnet wurde, um eine Hörempfindung auszulösen.

Um einen Bezug zwischen Frequenz und wahrgenommener Tonhöhe herzustellen, sei auf die Tonsysteme aus der Musik hingewiesen. Das Klein-c "c" ist durch eine Frequenz von 128 Hz definiert. Mit jeder weiteren Oktave verdoppelt sich die Frequenz, so daß es sich um eine exponentielle Steigerung handelt. Das eine Oktave höher liegende *eingestrichene-c* "c₁" hat die Frequenz 256 Hz und so weiter. Bezieht man diese Betrachtungsweise auf die Abnahme des menschlichen Hörvermögens im Alter, so erscheint der Rückgang von 20000 auf 5000 Hz als obere Hörgrenze zwar gewaltig, bedeutet aber nur den Verlust von zwei Oktaven. [2]

2.5 Digitale Schallaufnahme

Charakteristisch für eine digitale Schallaufnahme (Samplen) ist das Abtasten der Schall-schwingung zu diskreten Zeitpunkten. Bei der Compact Disc (CD) werden beispielsweise zwei Schwingungen (stereo) durch 44100 Amplitudenwerte (Samples) pro Sekunde festgelegt. Die Anzahl der Abtastungen pro Sekunde wird als *Samplerate* bezeichnet und erhält ebenfalls die Einheit Hertz, hier also 44.1 kHz. Zur Darstellung eines Amplitudenwertes wird dabei ein vorzeichenbehafteter 16 Bit Wert genutzt. Insgesamt können ganze Zahlen im Intervall von -32768 bis 32767 unterschieden werden.

Bei der Samplerate muß das *Nyquist Theorem* berücksichtigt werden. Die kritische Nyquist Frequenz ist bei diskret gesampelten Daten wie im hier vorliegenden Fall stets zu beachten: Sei Δ das Zeitintervall zwischen zwei Abtastzeitpunkten, was dem Reziproken der Samplerate entspricht. Dann ist die Nyquist Frequenz f_c gegeben durch:

$$f_c = \frac{1}{2\Delta} \quad (2.6)$$

Wenn eine Schwingung der Nyquist Frequenz f_c gesampelt wird, erfolgen genau zwei Ab-tastungen pro Sinuswelle. Wenn zu einem Abtastzeitpunkt der positive Spitzenwert der Welle erfaßt wird, wird beim nächsten Zeitpunkt der negative Spitzenwert erfaßt und so weiter. Genauso gut können die Abtastzeitpunkte die Welle aber auch bei ihrem Schnitt mit der x-Achse, also mit dem Wert 0, erfassen, der zweimal pro Welle angenommen wird. Konkret bedeutet dies, daß Frequenzen, die größer oder gleich der Nyquist Frequenz sind, nicht mehr sinnvoll erfaßt werden. Bei der hier verwendeten Samplerate von 44100 Hz, ergibt sich eine Nyquist Frequenz von 22050 Hz. Diese liegt deutlich höher als die höchste vom menschlichen Ohr wahrnehmbare Frequenz von 20000 Hz, die ohnehin mit zunehmendem Alter rasch weiter absinkt (siehe 2.4). [12]

Da die akustische Güteprüfung bei Mannesmann Dematic von Menschen durchgeführt wird, liegen alle benötigten Informationen zur Beurteilung der Schallemission offenbar im hörbaren Bereich. Es kann daher auf Aufnahmegeräte, die im Ultraschallbereich arbeiten, verzichtet werden. Die Abtastqualität, die durch Standard PC-Hardware erreicht wird (annähernd CD-Qualität), reicht hier völlig aus.

Kapitel 3

Einführung in Mustererkennungssysteme

3.1 Aufbau

Ziel und Aufgabe eines Mustererkennungssystems ist es, Objekte oder Vorgänge auf bestimmte Eigenschaften (Muster) hin zu überprüfen und aufgrund von Ähnlichkeitsvergleichen einer von mehreren zuvor definierten Objektmengen (Klassen) zuzuordnen. Die eigentliche Zuordnung eines Objektes zu einer Klasse wird als *Klassifikationsprozeß* bezeichnet.

Bei der akustischen Güteprüfung wird die Schallemission eines Objektes als Grundlage für die Mustererkennung genutzt. Nicht von Bedeutung ist dabei, ob die Schallemission durch das Objekt selbst erzeugt wird wie zum Beispiel das Laufgeräusch eines Motors [9] oder ob sie extern herbeigeführt wird wie bei dem hier untersuchten Verfahren von Mannesmann Dematic durch Dehnung mit einem Brecheisen. Die Schallemission wird analysiert und gegebenenfalls aus der zeitlichen Betrachtung in andere Darstellungen wie zum Beispiel das Frequenzspektrum überführt.

Formal können folgende Aufgaben eines Mustererkennungssystems genannt werden:

- **Signalerfassung und Aufbereitung**

Bei der akustischen Güteprüfung fällt hierunter die Aufzeichnung und die anschließende einheitliche Aufbereitung der Schallemission eines Objektes. Da im Bereich der Datenerfassung die Grundlage für spätere Ähnlichkeitsvergleiche geschaffen wird, gilt es, in diesem Bereich besondere Sorgfalt walten zu lassen. Fehler in der Signalerfassung können auch durch leistungsfähige Klassifikationsstufen nicht mehr ausgeglichen wer-

den¹.

- **Merkmalsextraktion**

Das Herausfinden von charakteristischen Merkmalen in den Schallemissionen der Objekte einer Klasse, die eine möglichst gute Abgrenzung gegenüber Objekten anderer Klassen ermöglichen, wird als *Merkmalsextraktion* bezeichnet. Grundlage für die Ermittlung von charakteristischen Merkmalen ist eine Menge von Schallemissionen, deren Klassenzugehörigkeit bekannt ist und die im folgenden als *Lernmenge* bezeichnet wird.

- **Klassifikation**

Auf der Grundlage der ermittelten charakteristischen Merkmale der verschiedenen Klassen werden neue, unbekannte Objekte durch ihre Schallemission beurteilt. Die Schallemissionen neuer Objekte werden unter identischen Rahmenbedingungen aufgezeichnet und verarbeitet wie die Schallemissionen der Lernmenge, die für die Merkmalsextraktion verwendet wurden. Das unbekannte Objekt wird derjenigen Klasse zugeordnet, der es aufgrund der Merkmale am ähnlichsten ist. Bereits an dieser Stelle sei darauf hingewiesen, daß sich die Methoden der Fuzzy Logik zur Modellierung solcher Ähnlichkeiten, die grundsätzlich unscharfer Natur sind, anbieten.

Anhand dieser Aufgaben kann der Einsatz eines Mustererkennungssystems in zwei Phasen unterteilt werden:

- **Lernphase**

In der Lernphase werden die Schallemissionen einer möglichst großen Anzahl von Objekten erfaßt. Deren Klassenzugehörigkeit bestimmt der menschliche Experte aufgrund seiner Erfahrung. Auf der Grundlage dieser sogenannten *Stichproben*, die zusammen die Lernmenge bilden, wird die Merkmalsextraktion vorgenommen. Die so gefundenen Merkmale sollen die verschiedenen Klassen möglichst gut gegeneinander abgrenzen. Im Fall Mannesmann Dematic sind zwei mögliche Klassen zu unterscheiden, zum einen das "Knistern", stellvertretend für eine fehlerhafte Schweißung, und zum anderen der "Lärm", stellvertretend für beliebige Nebengeräusche, die in einer großen Fabrikhalle auftreten. Wird "Lärm" erkannt, wird im Umkehrschluß angenommen, daß kein "Knistern" vom Testobjekt ausgeht und daher eine gute Schweißung vorliegt.

- **Einsatzphase**

In der Einsatzphase wird die Klassifikation neuer Objekte anhand ihrer Schallemissio-

¹siehe Seite 10 in [17]

nen durchgeführt.

In der Literatur ist eine große Menge von verfügbaren Klassifikationskonzepten bekannt. Ein ausführlicher Überblick über die klassischen Verfahren finden sich bei Niemann [11], eine Zusammenfassung der neueren Forschungsergebnisse liefert Strackeljan [17].

3.2 Aktueller Stand und Fokussierung

Akustische Güteprüfverfahren werden häufig mit Hilfe eines Mustererkennungssystems realisiert. Unterschiedliche Vorgehen sind in der Literatur vor allem im Bereich der Merkmalsextraktion zu finden. Viele Methoden wurden für eine spezifische Aufgabenstellung entwickelt und sind daher nur bedingt bei der Realisierung von anderen akustischen Prüfverfahren übertragbar. Grundlage bei der Lösung ist meistens eine Vorstellung über die Entstehung charakteristischer Geräusche. Als Beispiele seien hier Lagerschäden oder Anstreifvorgänge von Elektromotoren genannt, die in [9] untersucht wurden. Durch die genaue Vorstellung, wie das charakteristische Geräusch entsteht und wie sich der spezifische Schaden auf die Schallemission auswirkt, wurden dort Merkmale gesucht und gefunden. Auch in [16] werden charakteristische Geräusche, diesmal bei Laserpunktschweißungen, mit konkreten Ursachen für deren Entstehung in Zusammenhang gebracht. Ein anderer, etwas allgemeinerer Ansatz wird in [6] verfolgt. Dort werden verschiedene Attribute wie zum Beispiel Brummen, Rattern, Dröhnen definiert, deren jeweilige Ausprägung zur Bestimmung charakteristischer Merkmale genutzt werden.

All diese Verfahren setzen im Bereich der Merkmalsextraktion auf das menschliche Vermögen, für das System objektive und prüfbare Kriterien zu finden, anhand derer die Klassifikation durchgeführt werden kann. Diese, prinzipiell wie Expertensysteme arbeitenden Mustererkennungssysteme konnten zwar in den genannten Fällen erfolgreich realisiert werden, allerdings fehlen die dafür notwendigen Voraussetzungen im hier vorliegenden Fall der Güteprüfung von Mannesmann Dematic. Zum einen handelt es sich beim "Knistern" um ein sehr komplexes Geräusch, das nicht durch bestimmte sprachliche Attribute charakterisierbar ist. Zum anderen fehlt eine Vorstellung über die spezifische Entstehung des Geräusches und vor allem, wie daraus Regeln für die Klassifikation hergeleitet werden können.

Strackeljan weist in seiner Dissertation [17] darauf hin, daß es zur Zeit kaum möglich ist, eine Beziehung zwischen den Amplituden eines Schallsignals und spezifischen Fehlerursachen

herzustellen². Daher entwickelt Strackeljan in [17] gezielt einen allgemeinen, methodischen Ansatz für ein Mustererkennungssystem, dessen Merkmalsextraktion vollständig automatisiert abläuft. Merkmale werden ausschließlich in der Frequenzdarstellung gesucht, die sich aus der Zeitdarstellung mittels Fourier Transformation herleiten läßt. Daraus ergibt sich eine in der Regel sehr große Anzahl von möglichen Merkmalen, die mehr oder weniger charakteristisch für die Klassenzugehörigkeit eines Objektes sein können. Mit Hilfe einer Menge von Stichproben, das heißt einer Menge von Schallemissionen, deren Klassenzugehörigkeit bekannt ist, wird nach aussagekräftigen, für eine Klasse charakteristischen Merkmalen gesucht. Da dieser Ansatz bewußt ohne spezielle Kenntnisse über das spezifische Problem auskommen will, kann im Vorfeld keine Reduktion der Merkmalsanzahl vorgenommen werden. Daher handelt es sich um einen sehr rechenintensiven Vorgang, eine geeignete Teilmenge von Merkmalen zu ermitteln. Um n' Merkmale aus einer Menge von n zur Verfügung stehenden Merkmalen zu bestimmen, sind

$$\binom{n}{n'} = \frac{n!}{n'! \cdot (n - n')!} \quad (3.1)$$

mögliche Teilmengen zu überprüfen. Da dies nur in Ausnahmefällen möglich sein wird, wird ein Algorithmus vorgestellt, der eine suboptimale Merkmalsauswahl findet. Für die Realisierung werden Methoden der Fuzzy Logik eingesetzt, die in der Literatur bereits seit langem und häufig in Verbindung mit Mustererkennungssystemen zu finden sind³.

Da das vorgestellte Verfahren völlig unabhängig von der Art der zu klassifizierenden Geräusche ist, scheint es für die Umsetzung des Prüfverfahren von Mannesmann Dematic besonders geeignet zu sein, weswegen vor allem dieser Ansatz hier weiterverfolgt wird.

3.3 Verwendung von Fuzzy Logik

Fuzzy Logik eignet sich besonders gut für die Modellierung von Zuständen oder Beschreibungen, die nicht mathematisch präzise faßbar sind. Beispielhaft seien hier die bereits erwähnten Ähnlichkeitsvergleiche bei der Klassifikation genannt (siehe 3.1). Offensichtlich ist hier eine scharfe Trennung von "ist ähnlich" und "ist nicht ähnlich" wenig hilfreich. Vielmehr ist es

²siehe dort Seite 15

³Fuzzy Logik in Verbindung mit Klassifikationsverfahren wurde bereits 1980 in einem Forschungsbericht im Auftrag des Bundesministeriums für Verteidigung [7] untersucht. Weitere Literaturhinweise können auf Seite 15 in [17] gefunden werden.

erforderlich, den Grad der Ähnlichkeit fassen zu können. Dafür wird auf die Mittel der Fuzzy Logik zurückgegriffen. Für umfangreiche Einführungen in die Fuzzy Logik sei auf [8] oder [19] verwiesen.

Eine Fuzzy Menge F (engl. *Fuzzy Set*) über einem Universum U ist standardmäßig folgendermaßen definiert:

$$F : U \rightarrow [0, 1], \quad (3.2)$$

wobei sich "Standard" auf das gewählte Einheitsintervall bezieht.

Charakterisiert wird eine Fuzzy Menge F durch ihre sogenannte *Zugehörigkeitsfunktion* μ_F

$$\mu_F : U \rightarrow [0, 1]. \quad (3.3)$$

Thiele macht in [19] darauf aufmerksam, daß es mathematisch sinnlos ist, zwischen einer "Fuzzy Menge F über U " und einer "Zugehörigkeitsfunktion μ_F von F " zu unterscheiden⁴. Daher wird im folgenden der Fokus auf die eine Fuzzy Menge charakterisierende Zugehörigkeitsfunktion gesetzt.

Häufig werden Fuzzy Mengen genutzt, um umgangssprachliche Begriffe wie zum Beispiel "groß", "klein", "kalt" oder "warm" zu modellieren. Anschaulich gibt der Wert $\mu_F(x)$ für ein Element $x \in U$ den Grad der Erfüllung eines solchen Begriffs an, wobei hohe Zugehörigkeitswerte einen hohen Erfüllungsgrad repräsentieren. Beispielhaft sei hier eine mögliche Modellierung von "ähnlich" mit

$$x \in U; \quad \delta, a \in \mathbf{R}; \quad U = \mathbf{R} \quad (3.4)$$

dargestellt. An die Zugehörigkeitsfunktion $\mu_{\text{ähnlich}}(x)$ wird die Anforderung gestellt, für Werte x , die "ähnlich" a sind, also ungefähr dem Wert a entsprechen, hohe Zugehörigkeiten zu liefern. Für den Fall $x = a$ wird dabei eine 1 erwartet, je größer die Differenz von x zu a ist, desto geringer soll die Zugehörigkeit sein. Die Funktion wird abschnittsweise definiert

⁴siehe Seite 13 in [19]

und durch die Parameter a und δ festgelegt:

$$\mu_{\text{ähnlich}}(x) = \begin{cases} 0 & \text{für } x \leq a - \delta \\ 2 \left(\frac{x-a+\delta}{\delta} \right)^2 & \text{für } a - \delta < x \leq a - \frac{\delta}{2} \\ 1 - 2 \left(\frac{a-x}{\delta} \right)^2 & \text{für } a - \frac{\delta}{2} < x < a \\ 1 - 2 \left(\frac{x-a}{\delta} \right)^2 & \text{für } a \leq x \leq a + \frac{\delta}{2} \\ 2 \left(\frac{a-x+\delta}{\delta} \right)^2 & \text{für } a + \frac{\delta}{2} < x < a + \delta \\ 0 & \text{für } x \geq a + \delta \end{cases} \quad (3.5)$$

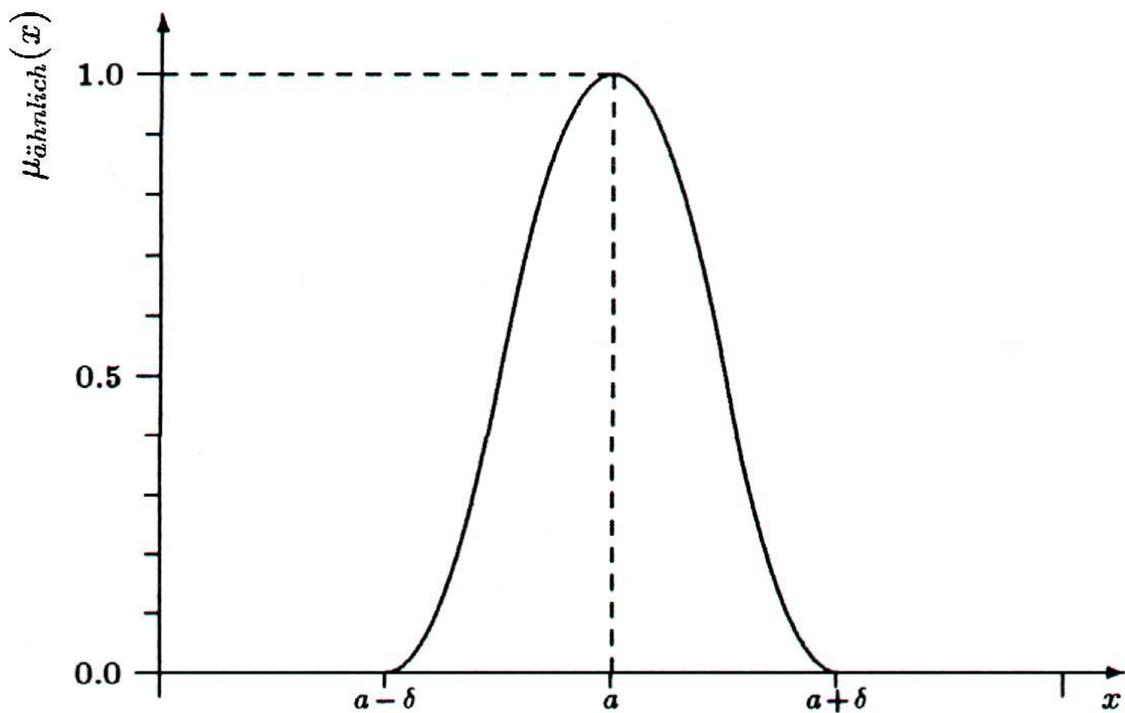


Abbildung 3.1: Modellierung von Ähnlichkeit [8]

Abbildung 3.1 zeigt $\mu_{\text{ähnlich}}(x)$, offensichtlich werden die gestellten Forderungen erfüllt: Für den Fall $x = a$ wird die Zugehörigkeit 1 geliefert. Innerhalb des Intervalls $[a - \delta, a + \delta]$ nimmt die Zugehörigkeit für Werte x mit zunehmender Differenz zu a zunächst langsam, dann rasant ab. Zu den Intervallrändern hin geht die Zugehörigkeit dann wieder langsamer gegen 0. Gilt für die Differenz $|a - x| \geq \delta$ wird die Zugehörigkeit 0 ermittelt, was im Sinne der Semantik von $\mu_{\text{ähnlich}}$ als "nicht ähnlich" angesehen werden kann. Die genaue Gestalt einer Zugehörigkeitsfunktion für einen spezifischen Anwendungsfall kann nicht allgemeingültig angegeben werden. Gegebenenfalls muß durch Ausprobieren die geeignetste Funktion herausgefunden werden. Im vorliegenden Fall kann beispielsweise durch den Parameter δ die

Funktion verändert werden. Je größer δ gewählt wird, desto größer ist das Intervall um a , für das Zugehörigkeitswerte größer 0 errechnet werden. Somit kann über den Parameter δ anschaulich die Toleranz der Funktion festgelegt werden. Da von Zugehörigkeitsfunktionen häufig eine ähnliche Gestalt verlangt wird, werden zum Beispiel in [8] weitere parametrierbare Zugehörigkeitsfunktionen vorgestellt ⁵.

Fuzzy Mengen sind wesentlich besser geeignet, linguistische Terme wie zum Beispiel "ähnlich" zu beschreiben, als dies mit den Mitteln der klassischen, zweiwertigen Logik möglich ist. Daher werden die in einem Mustererkennungssystem erforderlichen Bewertungen von zum Beispiel Ähnlichkeit (siehe 3.1) oder die Beurteilung der Güte eines Klassifikators etc. mit den Mitteln der Fuzzy Logik realisiert.

3.4 Technische Entwicklung

Interessant zu beobachten ist auch die technische Entwicklung im Bereich der akustischen Güteprüfung. In den frühen Veröffentlichungen wie zum Beispiel [9], 1983, wird noch ein enormer Aufwand betrieben, um eine Lösung hervorzubringen, die aus einer spezifischen, dem Problem angepaßten Hardware besteht. In [14], 1990, wird bereits auf die Rolle des Personal Computers verwiesen. Wenngleich die Frequenzanalyse dort ebenfalls noch mit externer Hardware vorgenommen wird, geschieht die umfangreiche Auswertung bereits mit einem PC, wobei vor allem die beliebigen Ausgabemöglichkeiten positiv hervorgehoben werden. Es wird bereits darauf hingewiesen, daß auch die Aufzeichnung und Frequenzanalyse mit einer Einschubkarte für den PC gelöst werden könnte, allerdings die Rechenzeiten und die Auflösung der Meßdaten schnell an Grenzen stoßen würden.

Die hier entwickelte Anwendung ist auf einem aktuellen PC Modell mit handelsüblicher Soundkarte betreibbar. Da die Güteprüfung von einem Menschen durchführbar ist, wird auf Untersuchungen im nicht hörbaren Bereich (außerhalb von 16 - 20000 Hz) verzichtet. Die digitalen Aufnahmen vor Ort bei Mannesmann wurden mit einem Laptop vorgenommen, der die Daten direkt auf Festplatte gespeichert hat. Auf zusätzliche Hardware wurde verzichtet. Abbildung 3.2 zeigt den prinzipiellen Aufbau eines Arbeitsplatzes.

⁵siehe Kapitel 2 in [8]

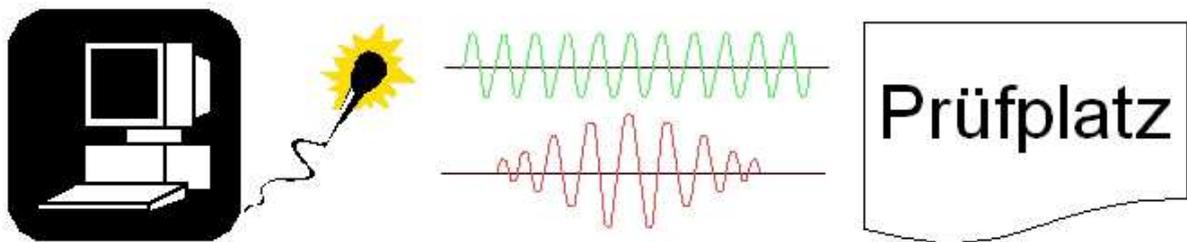


Abbildung 3.2: Schematischer Aufbau des Prüfsystems

Kapitel 4

Herleitung des Mustererkennungssystems

4.1 Grundlegende Betrachtung von Merkmalen

Nachdem bisher viele Begriffe recht intuitiv verwendet wurden, sollen nun einige formale Definitionen erfolgen. Für das entwickelte Merkmalsextraktionsverfahren sowie den Klassifikationsvorgang wird zunächst vom Ursprung und der Bedeutung der Merkmale abstrahiert.

Definition 1 : (*Merkmalsatz*)

Ein Merkmalsatz \vec{M} sei durch einen Vektor folgender Art gegeben:

$$\vec{M} = (m_1, m_2, \dots, m_{N_M})$$

mit

$$\forall m_i \in \mathbf{R} : \quad m_i \in [0, 1],$$

$$1 \leq i \leq N_M, \quad i, N_M \in \mathbf{N}$$

*Die Elemente m_i werden **Merkmalsausprägung** des Merkmals i genannt.*

Ein solcher Merkmalsatz soll genutzt werden, um alle möglicherweise relevanten Merkmale einer Schallemission zu repräsentieren. Während der Merkmalsextraktion soll eine in der Regel kleinere Anzahl von Merkmalen ermittelt werden, anhand derer eine Entscheidung über die Klassenzugehörigkeit eines Testobjektes gefällt werden kann. Dazu sei eine Auswahl von Merkmalen aus einem Merkmalsatz wie folgt definiert:

Definition 2 : (Merkmalsauswahl)

Gegeben sei ein Merkmalssatz

$$\vec{M} = (m_1, m_2, \dots, m_{N_M})$$

laut Definition 1. Dann wird ein Vektor

$$\vec{A} = (a_1, a_2, \dots, a_N)$$

mit

$$\forall i \in \mathbf{N} : 1 \leq i \leq N : 1 \leq a_i \leq N_M, \quad a_i \in \mathbf{N}$$

$$N \leq N_M, \quad N \in \mathbf{N}$$

sowie

$$a_j \neq a_k \quad \forall j, k \in \mathbf{N} : 1 \leq j, k \leq N$$

eine **Merkmalsauswahl** auf den Merkmalssatz \vec{M} genannt. Die Auswahl ist gegeben durch

$$\vec{M}(\vec{A}) = (m_{a_1}, m_{a_2}, \dots, m_{a_N})$$

Bei $\vec{M}(\vec{A})$ handelt es sich wieder um einen gültigen Merkmalssatz laut Definition, jedoch mit N und nicht mit N_M Merkmalen. Durch eine Merkmalsauswahl kann ein Merkmalssatz auf eine kleinere Anzahl von Merkmalen reduziert werden.

Außerdem sei als Grundlage für die Merkmalsextraktion eine Lernbasis folgendermaßen definiert:

Definition 3 : (Lernbasis)

Gegeben sei eine Menge L von Stichproben (auch als Lernmenge L bezeichnet), repräsentiert durch Merkmalssätze:

$$L = \{\vec{M}_1, \vec{M}_2, \dots, \vec{M}_{N_S}\}$$

N_S bezeichne die Anzahl der Stichproben, aus denen die Merkmalssätze gebildet wurden, und N_{M_i} bezeichne die Anzahl der Merkmale des Merkmalssatzes \vec{M}_i ($1 \leq i \leq N_S, i \in \mathbf{N}$). Insbesondere gelte:

$$N_{M_i} = N_{M_j} \quad \text{für} \quad \forall i, j \in \mathbf{N} : 1 \leq i, j \leq N_S$$

Alle Merkmalssätze \vec{M}_i besitzen also die gleiche Anzahl an Merkmalen. Des Weiteren sei eine Menge von Klassen

$$C = \{\Omega_1, \Omega_2, \dots, \Omega_{N_C}\}$$

gegeben, wobei $N_C \in \mathbf{N}$ die Anzahl der Klassen bezeichne, sowie eine Abbildung

$$\lambda : L \rightarrow C,$$

die jeden Merkmalssatz \vec{M}_i einer Klasse zuordnet.

Dann wird $B = (L, C, \lambda)$ als **Lernbasis** bezeichnet.

Im Rahmen der Signalerfassung und Aufbereitung muß für jede Stichprobe S_i ein Merkmalssatz \vec{M}_i mit N_{M_i} Merkmalen auf immer identische Art und Weise erzeugt werden. Dies gilt zum einen für die Lernphase, in der ein Merkmalssatz \vec{M}_i die Stichprobe S_i repräsentiert und in der Lernbasis zur Menge L gehört. Zum anderen gilt dies für die Einsatzphase, wo für neue und unbekannte Testobjekte eine Klassifikation vorgenommen werden soll. Die dabei erzeugten Merkmalssätze müssen auf identische Art wie die Merkmalssätze der Stichproben erzeugt werden. Wie bereits in Kapitel 3.1 erläutert, ist in diesem Zusammenhang unter anderem zu beachten, daß eine einheitliche Erfassung und Aufzeichnung der Schallsignale sowohl in der Lern- als auch in der Einsatzphase sichergestellt ist.

Die zuvor definierte Lernbasis stellt die Grundlage für die Merkmalsextraktion dar. Es ist eine Menge L von Stichproben gegeben, eine Menge C von Klassen sowie eine Abbildung λ , die jede Stichprobe einer Klasse Ω zuordnet ($\Omega \in C$). Da diese Abbildung λ elementweise vom menschlichen Experten festgelegt wird, gibt sie das menschliche Expertenwissen wieder und bildet eine wesentliche Grundlage des Mustererkennungssystems.

Auf der Grundlage einer Lernbasis soll durch das Merkmalsextraktionsverfahren ein Klassifikator gefunden werden, der folgendermaßen definiert ist:

Definition 4 : (Klassifikator)

Gegeben sei eine Lernbasis $B = (L, C, \lambda)$ mit

$$L = \{\vec{M}_1, \vec{M}_2, \dots, \vec{M}_{N_S}\}$$

und

$$C = \{\Omega_1, \Omega_2, \dots, \Omega_{N_C}\}$$

und

$$\lambda : L \rightarrow C$$

gemäß Definition 3.

Außerdem sei eine Merkmalsauswahl \vec{A} mit N Merkmalen gegeben, so daß es sich um eine gültige Merkmalsauswahl für die Merkmalssätze \vec{M}_i der Menge L gemäß Definition 2 handelt.

Θ_N bezeichne die Menge aller Merkmalssätze mit N Merkmalen, so daß eine Abbildung

$$\kappa : \Theta_N \rightarrow C$$

gegeben ist.

Dann heißt $K = (\vec{A}, \kappa)$ ein **Klassifikator zur Lernbasis B mit N Merkmalen**.

Ein Klassifikator $K = (\vec{A}, \kappa)$ führt die Klassifikation durch, indem er durch die Abbildung κ einen beliebigen Merkmalssatz $\vec{M} \in \Theta_N$, reduziert durch eine Auswahl \vec{A} , einer Klasse Ω zuordnet:

$$\kappa(\vec{M}(\vec{A})) = \Omega \quad (4.1)$$

Es sei auf den Unterschied zwischen den Abbildungen λ einer Lernbasis und κ eines Klassifikators aufmerksam gemacht: Die Abbildung λ wird elementweise durch einen menschlichen Experten festgelegt und bildet lediglich die Menge L der bekannten Merkmalssätze auf die Klassen ab. So wird das menschliche Expertenwissen bezüglich der Klassenzugehörigkeit der einzelnen Stichproben der Lernmenge L abgebildet. Im Gegensatz dazu leistet die Abbildung κ die Zuordnung eines beliebigen Merkmalssatzes zu einer Klasse, sprich die Klassifikation. Ursprungsmenge ist in diesem Fall Θ_N , die Menge aller Merkmalssätze mit N Merkmalen. Daher werden die Merkmalssätze mit N_M Merkmalen zuvor durch die Auswahl \vec{A} auf N Merkmale reduziert, so daß auch sie Elemente von Θ_N sind. Typischerweise reduziert die Auswahl auf eine Anzahl von fünf bis 20 Merkmale, wohingegen ein Merkmalssatz beispielsweise über mehrere hundert Merkmale verfügen kann ¹.

Da λ in der Lernbasis das Expertenwissen repräsentiert, sollte die Abbildung κ eines idealen Klassifikators für die Menge der bekannten Merkmalssätze L der Lernbasis dasselbe Klassifikationsergebnis, wie durch λ vorgegeben, liefern. Idealerweise gilt:

$$\kappa(\vec{M}_i(\vec{A})) = \lambda(\vec{M}_i) \quad \forall \vec{M}_i \in L \quad (4.2)$$

¹siehe S. 64 in [17]

Die Abbildung κ kann angewendet werden, da

$$\forall \vec{M}_i \in L : \quad \vec{M}_i(\vec{A}) \in \Theta_N \quad (4.3)$$

gilt. Die Elemente von L können durch die Auswahl \vec{A} in Elemente von Θ_N überführt werden. Dies folgt unmittelbar aus den Definitionen. Der Vorgang der Gegenüberprüfung von κ durch die Lernbasis gemäß Gleichung 4.2 wird *Reklassifikation* genannt. Dabei werden alle Stichproben, repräsentiert durch ihre Merkmalsätze in der Lernmenge L , hinsichtlich der Gültigkeit von Gleichung 4.2 überprüft. Über den Anteil der durch $K = (\vec{A}, \kappa)$ nicht richtig reklassifizierten Stichproben kann eine Aussage über die Güte des Klassifikators K getroffen werden.

Das formale Ziel des Merkmalsextraktionsverfahren besteht im Finden einer geeigneten Abbildung κ sowie einer geeigneten Merkmalsauswahl \vec{A} , die zusammen einen möglichst guten Klassifikator $K = (\vec{A}, \kappa)$ bilden.

4.2 Signalverarbeitung

Die bisher vorgenommenen Definitionen liefern eine allgemeine Grundlage für ein Mustererkennungssystem. Da es sich um ein akustisches Güteprüfverfahren handelt, werden als Datengrundlage ausschließlich die Schallemissionen der Objekte verwendet. Folgerichtig wird vor allem das Frequenzspektrum der Schallemissionen untersucht. Die bisher vorgenommenen Definitionen lassen jedoch auch die Nutzung beliebiger anderer Daten der Objekte zu, solange sie den in Definition 1 aufgeführten Restriktionen genügen.

In diesem Abschnitt soll zunächst eine allgemeine Vorgehensweise zur Verarbeitung der Schallemissionen aufgezeigt werden. Ausgangsbasis sind die digital aufgezeichneten Schallemissionen, die aus einer endlichen, aber unterschiedlichen Anzahl von diskreten Werten bestehen. Ziel der Signalverarbeitung ist die Erstellung von Merkmalsätzen gemäß Definition 1. Von besonderer Bedeutung ist, daß alle Schallemissionen sowohl die Stichproben innerhalb der Lernphase als auch zu klassifizierende Testobjekte in der Einsatzphase einer absolut einheitlichen Signalerfassung und -verarbeitung unterzogen werden. Zur Erstellung einer Lernbasis gemäß Definition 3 ist des weiteren auf eine einheitliche Anzahl N_M von Merkmalen in den einzelnen Merkmalsätzen zu achten, zumal die digital aufgezeichneten Schallemissionen unterschiedlich lang sein können.

4.2.1 Leistungsdichtespektrum

In Kapitel 2.2 wurde die *Fast Fourier Transformation (FFT)* bereits als geeignetes Mittel genannt, Schallsignale aus ihrer zeitlichen Darstellung in ihre Frequenzdarstellung zu überführen². Der FFT-Algorithmus nach Cooley und Tukey hat eine Laufzeit von $O(N \log_2 N)$. Neben der direkten Arbeit mit den Ergebnissen des FFT Algorithmus, sind auch andere Frequenzdarstellungen bekannt, beispielsweise das Leistungsdichtespektrum³. Bei dieser Methode wird der diskreten Struktur der Ein- und Ausgabedaten Rechnung getragen. Folgendes Szenario sei angenommen:

Eine digitale Schallaufnahme bestehe aus N Samples, die mit einer Samplingrate von f_S Hz aufgezeichnet wurden. Unter Berücksichtigung des Nyquist Theorems⁴ sind somit Frequenzen im Intervall von 0 bis $\frac{f_S}{2}$ Hz erfaßt. Der FFT Algorithmus überführt die N Samples aus der Zeit- in die Frequenzdarstellung, so daß nach der Durchführung N Amplitudenwerte zu den Frequenzen f_0, f_1, \dots, f_{N-1} ermittelt wurden. Der Amplitudenwert zu einer Frequenz f_k repräsentiert allerdings nicht nur genau die Frequenz f_k , sondern vielmehr ein Intervall von Frequenzen (Frequenzband) um f_k herum. Da insgesamt N Werte bestimmt wurden, ergeben sich N Frequenzbänder von jeweils $\frac{f_S}{2N}$ Hz, die durch einen Amplitudenwert repräsentiert werden. Daher sollte dieser Amplitudenwert zu f_k eine Art Mittelwert über das ihn umgebende Frequenzband darstellen und nicht den genauen Wert zu einer Frequenz f_k liefern, wie man dies im Gegensatz dazu von einer kontinuierlichen Funktion erwarten würde. Dies ist jedoch nicht der Fall. Vielmehr tritt eine sogenannte Leckage (engl. *leakage*) auf, so daß sich die Amplituden einer Frequenz nicht nur innerhalb ihres Frequenzbandes darstellen, sondern auch in ihren benachbarten Bändern und darüber hinaus. Erklärbar ist dieser unerwünschte Sachverhalt durch Randeffekte (engl. *edge effects*), die dadurch entstehen, daß es sich um eine diskrete, nur in einem bestimmten Zeitintervall definierte Ausgangsfunktion handelt und nicht um eine kontinuierliche Funktion, die in \mathbf{R} definiert ist. Durch den abrupten Beginn und das abrupte Ende an den Intervallgrenzen enthält die Fourier Transformation Anteile an hohen Frequenzen.

Diesem unerwünschten Sachverhalt wird mit einer Methode namens *Data Windowing* begegnet. Die Hauptziele einer Windowingfunktion sind ein sanftes Ansteigen und ein sanftes Ausklingen der diskreten Zeitdaten. Die Quelldaten werden mit einer Windowfunktion vor

²Eine typische FFT Implementierung kann in [12], S. 496-591, gefunden werden.

³engl. Power Spectral Density (PSD)

⁴siehe Kapitel 2.5

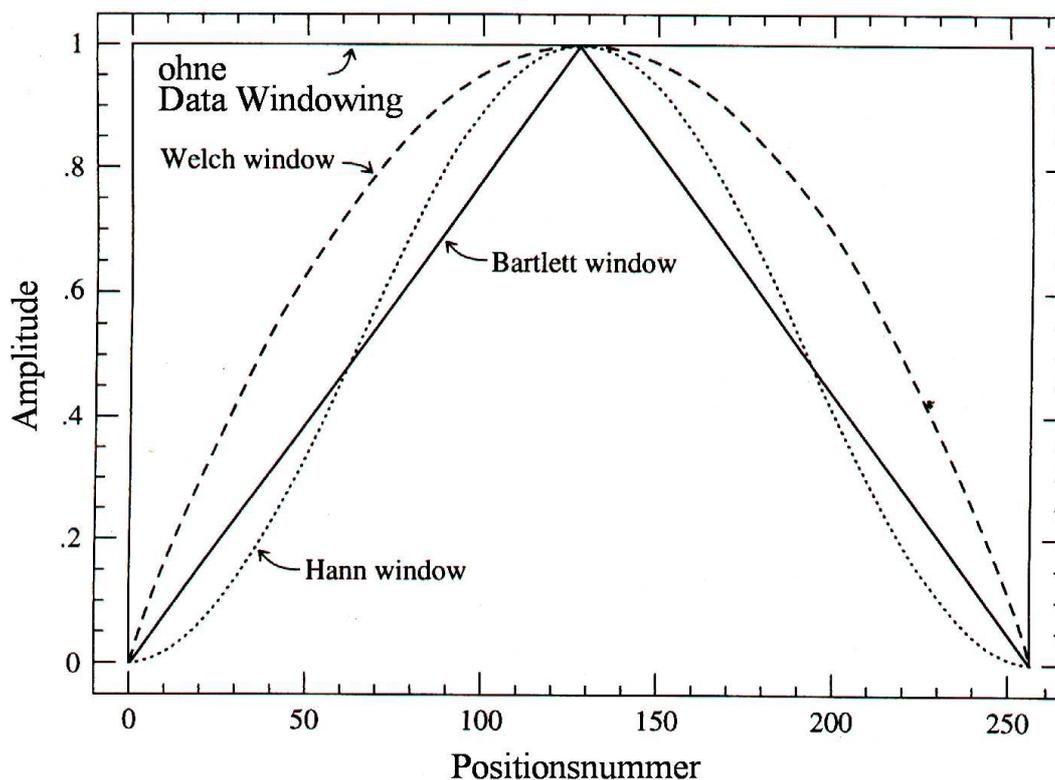


Abbildung 4.1: Geeignete Funktionen für Data Windowing

der FFT Berechnung multipliziert. Beispielfür andere sei hier die *Bartlett Window* Funktion angegeben:

$$w(j) = 1 - \left| \frac{j - \frac{1}{2}N}{\frac{1}{2}N} \right| \quad j \in \mathbf{N}, \quad 0 \leq j \leq N - 1 \quad (4.4)$$

Abbildung 4.1 zeigt neben der Bartlett Funktion auch weitere Windowing Funktionen für $N = 256$. Bei der Berechnung des Leistungsdichtespektrums werden die diskreten Zeitdaten in Segmente eingeteilt, die einzeln verarbeitet und deren einzelnen Ergebnisse aufsummiert werden. Optional können sich die Segmente auch überlappen.

Abbildung 4.2 zeigt die Auswirkung der Leakage eines Signals in andere Frequenzbänder hinein. Es ist zu erkennen, daß beispielsweise die Bartlett Funktion die Leakage deutlich reduziert.

Mehr Informationen über die Berechnung des Leistungsdichtespektrums und Data Windowing sowie einen entsprechenden Algorithmus können in [12] gefunden werden. Der dort vorgestellte Algorithmus leistet des weiteren die Berechnung des Leistungsdichtespektrums mit einer fest vorgegebenen Anzahl von Ausgabewerten, die unabhängig von der Anzahl der Eingabewerte ist. Durch die einheitliche Vorgabe des Parameters ist sichergestellt, daß die

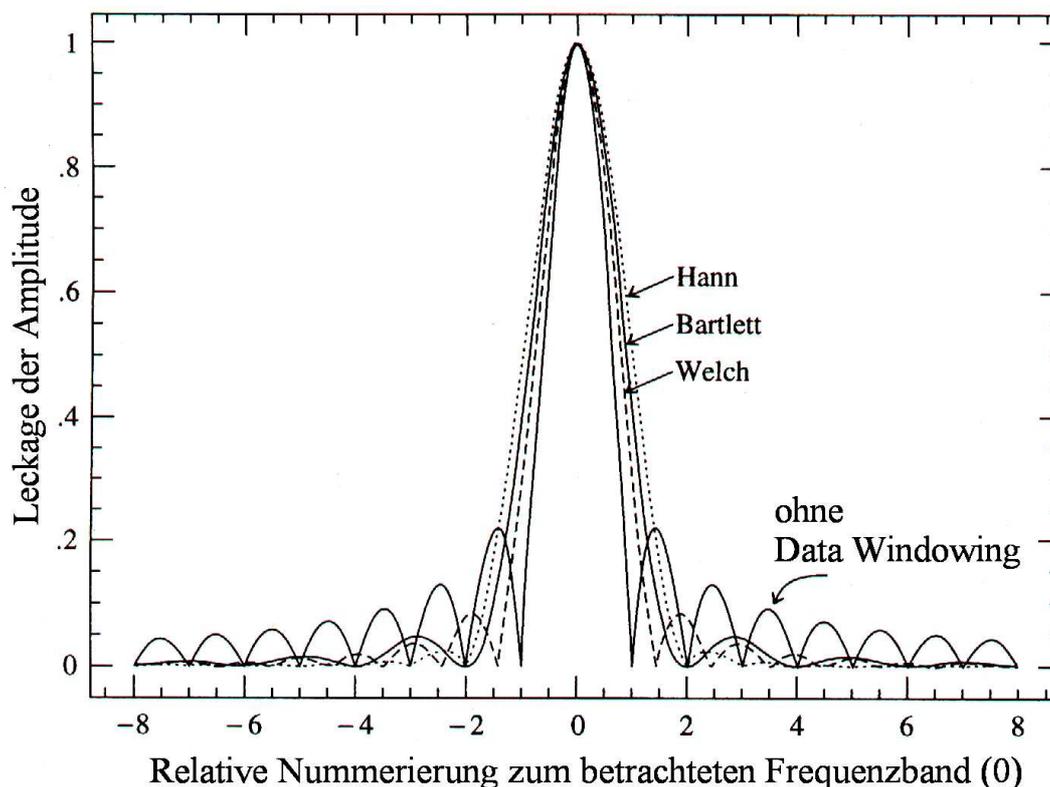


Abbildung 4.2: Leckage eines Signals in andere Frequenzintervalle

Ausgangsdaten aller Schallemissionen aus der gleichen Anzahl von Werten bestehen, was bei der Erstellung von Merkmalsätzen zur Verwendung in einer Lernbasis unter anderem eingangs gefordert wurde.

4.2.2 Normierung

In Definition 1 wurde gefordert, daß die Merkmalsausprägungen im Intervall $[0, 1]$ liegen. Da dies in der Regel nicht automatisch der Fall ist, muß eine Normierung der Daten auf dieses Intervall stattfinden. Die vorgestellte Rechenvorschrift ⁵ ist allgemein anwendbar, daher auch für die Ergebnisse aus der Ermittlung des Leistungsdichtespektrums möglich:

$$m_{js}^T = \frac{m_{js} - m_{jmin}}{m_{jmax} - m_{jmin}} \quad (4.5)$$

wobei

m_{js}^T : transformiertes j -tes Merkmal der s -ten Stichprobe S

⁵siehe S. 35 in [17]

- m_{jmin} : minimale Ausprägung des j-ten Merkmales über alle S
 m_{jmax} : maximale Ausprägung des j-ten Merkmales über alle S

Anschaulich geschieht durch diese Normierung folgendes: Für jedes einzelne Merkmal wird die minimale und die maximale Merkmalsausprägung über alle Stichproben bestimmt. Die Merkmalsausprägungen werden so in das Intervall $[0, 1]$ abgebildet, daß die minimale Ausprägung 0 und die maximale Ausprägung 1 beträgt. Dadurch ist sichergestellt, daß alle Merkmalsausprägungen im Intervall $[0, 1]$ liegen und sich die Werte über das ganze Intervall verteilen. [17]

Je nach Art der verwendeten Daten kann es vorkommen, daß diese sich besser durch ein logarithmisches Maß beschreiben lassen (siehe zum Beispiel Kapitel 2.3). Um in einem solchen Fall eine bessere Verteilung der Stichproben im Merkmalsraum zu erreichen, kann die Normierung beispielsweise so durchgeführt werden:

$$m_{js}^T = \frac{\log_b(m_{js}) - m_{jlogmin}}{m_{jlogmax} - m_{jlogmin}} \quad (4.6)$$

wobei

- m_{js}^T : transformiertes j-tes Merkmal der s-ten Stichprobe S
 $m_{jlogmin}$: minimale Ausprägung des logarithmierten j-ten Merkmales über alle S
 $m_{jlogmax}$: maximale Ausprägung des logarithmierten j-ten Merkmales über alle S
 b : geeignete Basis zum Logarithmieren

Die Normierung auf das Intervall $[0, 1]$ hat zum einen praktische Vorteile bei der Weiterverarbeitung in Fuzzy Zugehörigkeitsfunktionen, zum anderen wird bei den vorgeschlagenen Rechenvorschriften 4.5 und 4.6 die einheitliche Skalierung aufgehoben, so daß die Ausprägungen aller Merkmale den ganzen Intervallbereich nutzen.

Insgesamt ist mit der Berechnung des Leistungsdichtespektrums und den vorgestellten Normierungsvorschriften eine einheitliche Erstellung von Merkmalssätzen entsprechend den eingangs formulierten Restriktionen sowohl in der Lern- als auch in der Einsatzphase möglich. Unter Zuziehung des menschlichen Experten kann eine Lernbasis B gemäß Definition 3 erstellt werden. Voraussetzung ist eine Menge von Klassen

$$C = \{\Omega_1, \Omega_2, \dots, \Omega_{N_C}\}$$

und eine Menge von N_S Testobjekten, deren Klassenzugehörigkeit bekannt ist. Durch das vorgestellte Verfahren werden die Schallemissionen mittels Berechnung des Leistungsdichtespektrums und anschließender Normierung in gültige Merkmalsätze überführt, zusammengefaßt in der Menge:

$$L = \{\vec{M}_1, \vec{M}_2, \dots, \vec{M}_{N_S}\}$$

Mit Hilfe des menschlichen Experten wird die Abbildung λ elementweise festgelegt

$$\lambda : L \rightarrow C,$$

so daß anschließend eine den Definitionen entsprechende Lernbasis $B = (L, C, \lambda)$ vorliegt.

4.3 Unscharfes Klassifikationskonzept

Nachdem dargestellt wurde, auf welche Art und Weise eine gültige Lernbasis B erstellt werden kann, soll nun erörtert werden, wie anhand eines Klassifikator $K = (\vec{A}, \kappa)$ ein Merkmalsatz \vec{M} einer Klasse Ω zugeordnet wird. In diesem Abschnitt soll eine geeignete Abbildung κ gefunden werden, die unabhängig von den spezifischen Merkmalsätzen ist. Zu diesem Zweck wird zunächst davon ausgegangen, daß eine geeignete Auswahl \vec{A} vorhanden ist, die den zugrundeliegenden Modellvorstellungen entspricht. In der Lernphase des Mustererkennungssystem wird es Ziel der Merkmalsextraktion sein, zu einem gegebenen κ eine passende Auswahl \vec{A} zu finden. Diese sei hier zunächst als geeignet gegeben angenommen. Da es sich laut Definition 2 auch bei Merkmalsauswahlen $\vec{M}(\vec{A})$ um gültige Merkmalsätze handelt, wird im folgenden allgemein von Merkmalsätzen \vec{M} gesprochen, die N Merkmale enthalten sollen. Für die weiteren Erläuterungen sei außerdem eine Menge

$$C = \{\Omega_1, \Omega_2, \dots, \Omega_{N_C}\}$$

von Klassen angenommen.

4.3.1 Distanzmaß und Zugehörigkeitsfunktion

Ein häufiger Ansatz bei Klassifikationsaufgaben ist die Auffassung der Merkmalsätze als Punktvektoren im N -dimensionalen Raum (Merkmalsraum), wobei dies anschaulich darstellbar nur für $N \leq 3$ möglich ist. Die gesuchte Charakteristik in Merkmalen beruht auf der

in vielen schwingungstechnischen Arbeiten beobachteten Tatsache, daß die Merkmalsausprägungen eines Merkmals für eine Klasse Ω_k ähnliche Werte aufweisen ⁶. Dadurch kommt es im Merkmalsraum zu einer Häufung von Punkten, die als *Cluster* bezeichnet werden. Ein häufig zu findender Ansatz die Zugehörigkeit eines zu klassifizierenden Merkmalsatzes zu ermitteln, besteht darin, die räumliche Distanz zu den Clustern der verschiedenen Klassen zu bestimmen. Als Klassifikationsergebnis wird die Klasse ermittelt, deren Cluster die geringste Distanz zum Testobjekt aufweist. Diese Art von Klassifikatoren werden *Abstandsklassifikatoren* ⁷ genannt. Häufig eingesetzt wird die euklidische Distanz, die für zwei Merkmalsätze \vec{M}_l und \vec{M}_m folgendermaßen bestimmt wird:

$$d_{ml} = |\vec{M}_l - \vec{M}_m| = \sqrt{\sum_{n=1}^N (m_{nm} - m_{nl})^2} \quad (4.7)$$

Dabei entspricht

$$\begin{aligned} \vec{M}_l, \vec{M}_m & : \text{ Merkmalsätze mit } N \text{ Merkmalen} \\ m_{nl}, m_{nm} & : \text{ n-tes Merkmal des Merkmalsatzes } \vec{M}_l \text{ bzw. } \vec{M}_m \end{aligned}$$

Da ein Cluster der Klasse Ω_k ($k = 1, \dots, N_C$) aus einer größeren Anzahl von Stichproben besteht, bietet sich für die Abstandsbestimmung zu einem zu klassifizierenden Merkmalsatz die Berechnung des Schwerpunktvektors $\vec{r}_k \in \mathbf{R}^N$ des Clusters an:

$$\vec{r}_k = \frac{1}{N_{Sk}} \sum_{s=1}^{N_{Sk}} \vec{M}_{ks} \quad (4.8)$$

Dabei entspricht

$$\begin{aligned} N_{Sk} & : \text{ Anzahl der Stichproben zur Klasse } \Omega_k \\ \vec{M}_{ks} & : \text{ Merkmalsatz der Stichprobe } s, \text{ die zur Klasse } \Omega_k \text{ gehört} \end{aligned}$$

⁶ siehe S. 28 in [17]

⁷ siehe S.45, 46 in [13] oder S. 27 in [17]

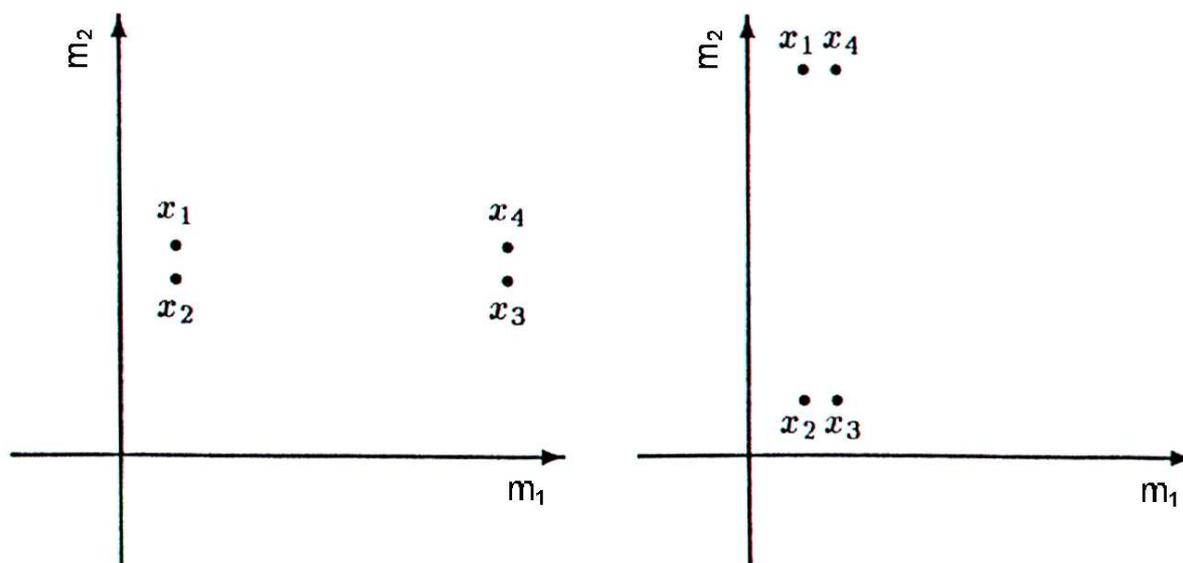


Abbildung 4.3: Schematische Darstellung von Clustern bei unterschiedlicher Skalierung

Bereits in Kapitel 3.3 wurde dargelegt, daß zur Modellierung von Ähnlichkeit zwischen einem zu klassifizierenden Objekt und den bekannten Lernstichproben die Mittel der Fuzzy Logik eingesetzt werden sollen. Aufgrund der räumlichen Distanz, berechnet durch beispielsweise Gleichung 4.7, eines zu klassifizierenden Merkmalsatzes \vec{M}_x und dem Schwerpunktvektor \vec{r}_k nach Gleichung 4.8 kann eine Zugehörigkeitsfunktion $\mu_{\text{ähnlich}}$ festgelegt werden. Grundsätzlich geeignet dafür ist eine Funktion der Art

$$\mu_{\text{ähnlich}}(x) = \frac{1}{1 + d_{xr}^2} \quad \text{mit } d_{xr} = |\vec{M}_x - \vec{r}_k| \text{ entsprechend Gleichung 4.7} \quad (4.9)$$

Diese Funktion ordnet einer kleinen Distanz eine hohe Zugehörigkeit bzw. Ähnlichkeit zu. Bei zunehmendem Abstand nimmt die Zugehörigkeit quadratisch ab. Somit erfüllt die Funktion die durch die Semantik von "ähnlich" entstehenden Anforderungen. Welches Distanzmaß letztendlich zur Anwendung kommt, ist an dieser Stelle unerheblich. [17]

Nachdem die Bedeutung der räumlichen Distanz von Merkmalen zueinander als wichtiges Indiz zur Analyse von Clustern dargestellt wurde, kann ein weiterer Grund für die Normierung der Ausgangsdaten aufgezeigt werden (siehe Kapitel 4.2.2): Je nach Art der vorliegenden Daten kann es vorkommen, daß diese unterschiedlich skaliert sind. Werden Daten mit unterschiedlicher Skalierung in ein Koordinatensystem eingetragen, können sich grundverschiedene Darstellungen ergeben, die zu gegensätzlichen Annahmen über die Zusammensetzung der Cluster führen.

Abbildung 4.3 zeigt beispielhaft die Auswirkung von unterschiedlichen Skalierungen. Im

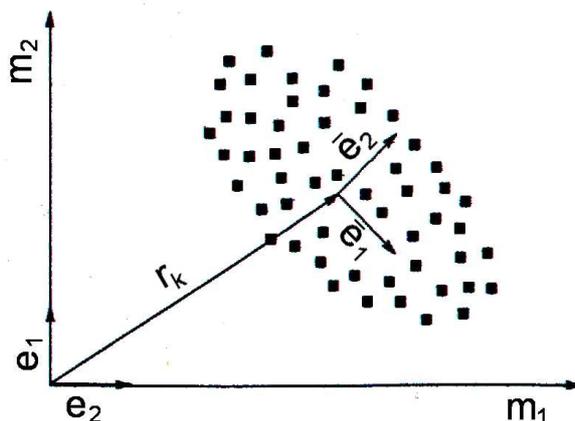


Abbildung 4.4: Beispiel für eine Hauptachsentransformation mit $N = 2$

linken Koordinatensystem bilden offenbar die Punkte (x_1, x_2) und (x_3, x_4) bezogen auf ihre räumliche Distanz ein Cluster. Wird die Skalierung geändert (rechtes Koordinatensystem), bilden (x_1, x_4) und (x_2, x_3) ein Cluster. Um die Vergleichbarkeit der Daten zu gewährleisten, sollte daher eine geeignete Skalierung vorgenommen werden. [5]

4.3.2 Hauptachsentransformation in das Cluster

In vielen schwingungstechnischen Arbeiten wurde beobachtet, daß das Cluster, das von den Merkmalsätzen einer Klasse ausgebildet wird, häufig eine ellipsoide bzw. hyperellipsoide Form im \mathbf{R}^N mit beliebiger Orientierung hat. Ein geeignetes Distanzmaß sollte diesen Umstand berücksichtigen, so daß verschiedene randständige Stichproben eines Clusterellipsoids ähnliche Distanzen zum Schwerpunkt aufweisen. Das in Gleichung 4.7 vorgestellte Distanzmaß leistet dies nicht, da alle Raumachsen gleich bewertet werden. Um der ellipsoiden Struktur und der Lage und Orientierung des Clusters im Raum gerecht zu werden, wird ein transformiertes Koordinatensystem genutzt, dessen Basisvektoren den Hauptachsen des Clusterellipsoids entsprechen. Die geforderte ähnliche Distanz von randständigen Stichproben zum Clusterschwerpunkt läßt sich bei der Berechnung der transformierten Basisvektoren erreichen, indem die Streuung über die jeweilige Raumachse berücksichtigt wird. Dieses Verfahren wird *Hauptachsentransformation* genannt.

Abbildung 4.4 zeigt beispielhaft eine Hauptachsentransformation für Merkmalsätze mit $N = 2$. Die Basisvektoren \vec{e}_1 und \vec{e}_2 bilden die kanonische Basis. Die Merkmalsausprägungen der Merkmale m_1 und m_2 der zur Klasse Ω_k gehörenden Stichproben bilden ein Clusterellip-

soid. Durch die Hauptachsentransformation entsteht das transformierte Koordinatensystem, dessen Ursprung dem Clusterschwerpunkt \vec{r}_k entspricht und dessen Basisvektoren \vec{e}_1 und \vec{e}_2 den Hauptachsen des Clusterellipsoids entsprechen [17].

Grundlage für die Berechnung des transformierten Koordinatensystems ist die Stichprobenkovarianzmatrix, die die statistischen Abhängigkeiten zwischen den einzelnen Stichproben wiedergibt. Die elementweise Definition der $N \times N$ -Kovarianzmatrix V_k für eine Klasse Ω_k lautet

$$v_{kji} = \frac{1}{N_{S_k} - 1} \sum_{s=1}^{N_{S_k}} (m_{ksj} - r_{kj})(m_{ksi} - r_{ki}) \quad (4.10)$$

Dabei bedeutet

- v_{kji} : Matricelement ji der Kovarianzmatrix V_k
- N_{S_k} : Anzahl der Stichproben von Klasse Ω_k
- r_{ki}, r_{kj} : i -te bzw. j -te Komponente des Schwerpunktvektors r_k
- m_{ksi}, m_{ksj} : i -tes bzw. j -tes Merkmal des Merkmalsatzes M_s der Klasse zur Ω_k gehört
- k : Bezug zur Klasse Ω_k

Zur rechnerischen Bestimmung der transformierten Basisvektoren wird die Kovarianzmatrix V_k in eine Diagonalmatrix Δ_k überführt:

$$V_k \Rightarrow \Delta_k \quad (4.11)$$

Für die Elemente der Diagonalmatrix Δ_k wird

$$\delta_{kij} = 0 \quad \text{für} \quad i \neq j \quad (4.12)$$

gefordert. Die Transformation erfolgt durch die Lösung des Eigenwertproblems

$$V_k \cdot \vec{e}_{kn} = \delta_{kn} \cdot \vec{e}_{kn} \quad (4.13)$$

bzw.

$$(V_k - \delta_{kn} \cdot E) \cdot \vec{e}_{kn} = \vec{0} \quad (4.14)$$

mit

- δ_{kn} : Eigenwerte der Matrix V_k
 E : kanonische Einheitsmatrix
 \vec{e}_{kn} : orthonormales Eigenvektorsystem, $n = 1, \dots, N$

Die Eigenvektoren \vec{e}_{kn} geben bezogen auf die kanonische Basis die Richtungen der transformierten Koordinatenachsen an (siehe hierzu auch Abbildung 4.4). Die Eigenwerte δ_{kn} bilden die Diagonalelemente der transformierten $N \times N$ -Kovarianzmatrix Δ_k :

$$\Delta_k = \begin{pmatrix} \delta_{k1} & 0 & \dots & 0 \\ 0 & \delta_{k2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \delta_{kN} \end{pmatrix} \quad (4.15)$$

Ursprung des transformierten Koordinatensystems ist der Schwerpunktvektor \vec{r}_k der jeweiligen Klasse Ω_k . Für das transformierte Koordinatensystem gilt unter anderem:

- Die Komponenten von Δ_k sind unkorreliert.
- Die Eigenwerte δ_{kn} geben die Varianz in den Koordinatenrichtungen der Vektoren \vec{e}_{kn} an.
- Bei Eigenwerten $\delta_{kn} = 0$ gilt, daß die Stichproben der Klasse Ω_k in dieser Raumrichtung keine Streuung besitzen und daher nicht weiter betrachtet werden müssen.

Eine Zugehörigkeitsfunktion μ_k , die die Ähnlichkeit eines Testobjekts, dargestellt durch den Merkmalsatz \vec{M} , zur Klasse Ω_k auf der Grundlage der Distanz im transformierten Koordinatensystem bestimmt, kann so definiert werden ⁸:

$$\mu_k(\vec{M}) = \frac{1}{1 + \frac{1}{t} \cdot \sum_{n=1}^N \frac{1}{\delta_{kn}} \cdot ((\vec{M} - \vec{r}_k) \cdot \vec{e}_{kn})^2} \quad (4.16)$$

mit $k = 1, \dots, N_C$

Über den Parameter t kann die Gestalt der Zugehörigkeitsfunktion verändert werden. Bei Werten $t \leq 2$ hat die Funktion einen eher schmalen Bereich mit hohen Zugehörigkeiten, mit zunehmender Distanz nimmt die Zugehörigkeit rasch ab. Bei größerem t liefert die

⁸siehe S. 32 in [17]

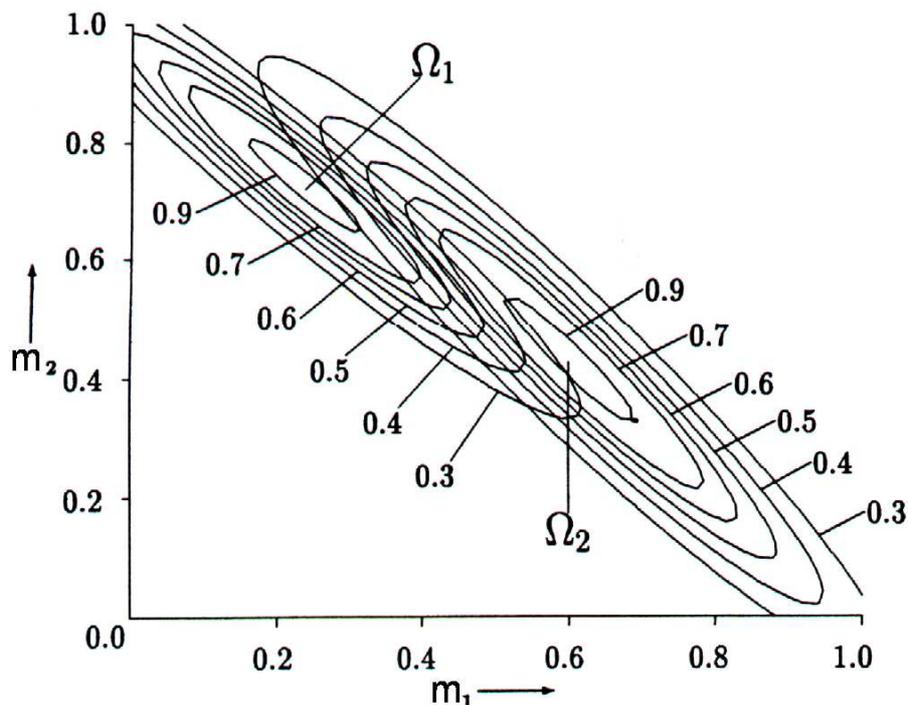


Abbildung 4.5: Linien konstanter Zugehörigkeit bei zwei Klassen Ω_1 und Ω_2 (Beispiel)

Funktion auch zu größeren Distanzen noch hohe Zugehörigkeitswerte. Wie der Parameter t gut gewählt wird, kann a priori nicht gesagt werden. Letztendlich kann die Bewertung einer spezifischen Wahl von t nur über die Leistungsfähigkeit des Klassifikators beurteilt werden. Strackeljan schlägt in [17] die Wahl von $t = 3$ vor.

Abbildung 4.5 zeigt beispielhaft zwei Klassen Ω_1 und Ω_2 mit $N = 2$. Um die Clusterschwerpunkte herum sind Linien gleicher Zugehörigkeitswerte zur jeweiligen Klasse eingezeichnet. Offensichtlich muß es zwischen den beiden Klassen im Merkmalsraum eine Überschneidung geben, auf der Stichproben sowohl zu Ω_1 als auch zu Ω_2 gleiche Zugehörigkeitswerte liefern. Diese Trennkurve zwischen Ω_1 und Ω_2 zeigt Abbildung 4.6. An einigen Stellen sind die identischen Zugehörigkeitswerte vermerkt. Durch die Gleichheit der Zugehörigkeitswerte zu beiden Klassen erfolgt eine räumliche Aufteilung durch die Trennkurve. Stichproben, deren Merkmalsatz im Merkmalsraum links der Trennkurve liegen, werden der Klasse Ω_1 zugeordnet, Merkmalsätze rechts der Trennkurve der Klasse Ω_2 .

Abbildung 4.7 zeigt die beiden Zugehörigkeitsfunktionen in einem dreidimensionalen Gittermodell. Auf der dritten Achse wurde jeweils das Maximum der beiden Zugehörigkeitswerte aufgetragen.

Meistens werden in der Fuzzy Logik N eindimensionale Zugehörigkeitsfunktionen aufge-

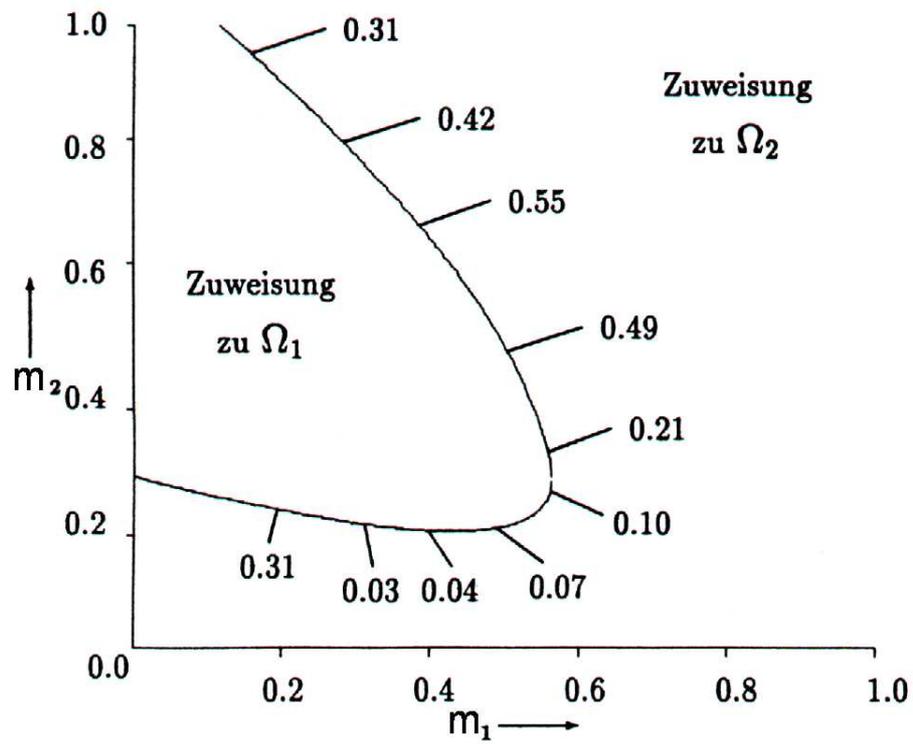


Abbildung 4.6: Trennkurve zwischen den Klasse Ω_1 und Ω_2 (Beispiel)

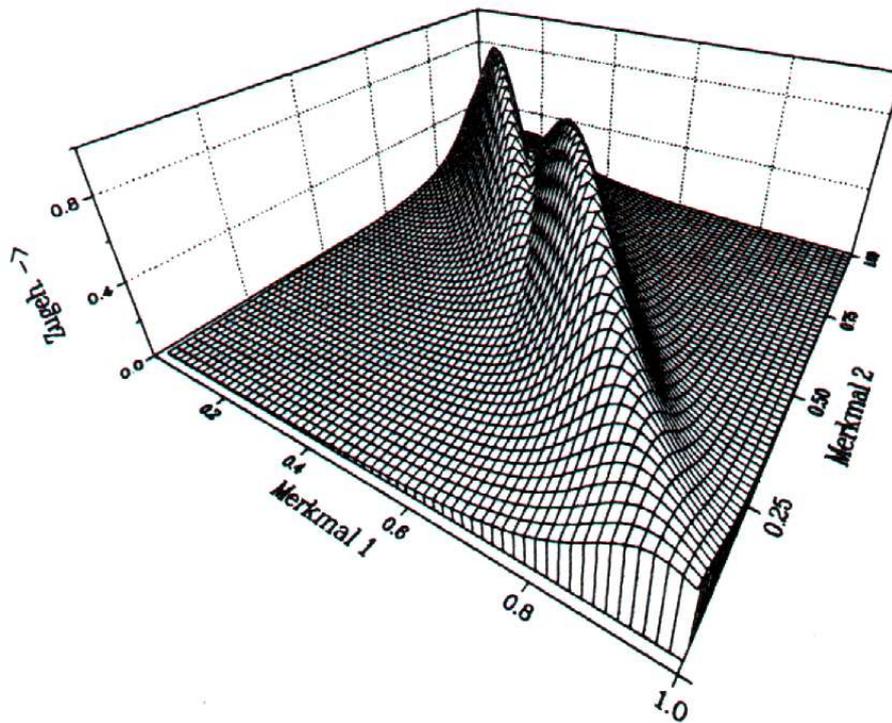


Abbildung 4.7: Zugehörigkeitswerte im 3D-Gittermodell

stellt, statt wie hier eine N -dimensionale. Diese N eindimensionalen Zugehörigkeitsfunktionen müssen nach ihrer einzelnen Auswertung mit einem Fuzzy Operator wie zum Beispiel der max-Funktion aggregiert und zu einer unscharfen Gesamtaussage ausgewertet werden [8, 19]. Strackeljan zeigt in [17], daß bei der Verwendung von Standard Fuzzy Operatoren im Vergleich zur Zugehörigkeitsfunktion nach 4.16 deutlich schlechtere Ergebnisse erzielt werden können. Als vorteilhaft wird die mathematisch geschlossene Formulierung bei N -dimensionalen Merkmalsätzen erachtet.

Mehr Informationen zum Thema Varianz und Kovarianz kann in [4], mehr zur Hauptachsentransformation in [13, 17] gefunden werden.

4.3.3 Entscheidung über das Klassifikationsergebnis

Die vorgestellte Zugehörigkeitsfunktion ermöglicht die Beurteilung eines Testobjekts, indem dessen Lage im Merkmalsraum mit der Lage der Cluster der jeweiligen Klasse verglichen wird. Dabei wird der ellipsoiden Struktur der Cluster Rechnung getragen. Je mehr der Merkmalsatz eines Testobjektes im Zentrum des Clusters einer Klasse liegt, desto größer ist die ermittelte Zugehörigkeit zu ihr.

Als Klassifikationsergebnis wird diejenige Klasse Ω_x ermittelt, für die gilt:

$$\mu_x = \max(\mu_1, \mu_2, \dots, \mu_{N_C}) \quad \text{mit } 1 \leq x \leq N_C \quad (4.17)$$

Ein Testobjekt wird der Klasse zugeordnet, für die die maximale Zugehörigkeit ermittelt wurde.

Wenn für jede Klasse Ω_x (mit $x = 1, \dots, N_C$) eine Zugehörigkeitsfunktion μ_x aufgestellt ist, kann die anfänglich geforderte Abbildung κ eines Klassifikators K auf der Grundlage des vorgestellten Klassifikationskonzeptes folgendermaßen formuliert werden:

$$\kappa(\vec{M}) = \Omega_x \quad (4.18)$$

mit

$$x \in \mathbf{N} : 1 \leq x \leq N_C : \quad \mu_x(\vec{M}) = \max(\mu_1(\vec{M}), \mu_2(\vec{M}), \dots, \mu_{N_C}(\vec{M}))$$

4.4 Alternative Klassifikationskonzepte

Das in Kapitel 4.3 vorgestellte Klassifikationskonzept basiert im wesentlichen auf einer Zugehörigkeitsfunktion, die für jede einzelne Klasse die Zugehörigkeit eines Testobjektes bestimmt und als Ergebnis diejenige Klasse ermittelt, zu der das Testobjekt die größte Zugehörigkeit hat. Ein anderer Ansatz ist die Aufstellung einer *Entscheidungsfunktion*, die den Merkmalsraum durch eine Gerade (bei $N = 2$) bzw. durch eine Hyperebene (bei mehrdimensionalen Merkmalsräumen) aufteilt. [13]

4.4.1 Linearer Klassifikator

Die einfachste Form einer Entscheidungsfunktion ist eine lineare Funktion. In einem einfachen Szenario mit 2 Klassen Ω_1 und Ω_2 und N Merkmalen kann die Entscheidungsfunktion angegeben werden als

$$t(\vec{M}) = w_0 + w_1 m_1 + w_2 m_2 + \dots + w_N m_N = w_0 + \sum_{n=1}^N w_n m_n \quad (4.19)$$

Als Klassifikationsregel für einen Merkmalsatz \vec{M} wird festgelegt:

$$\vec{M} \rightarrow \Omega_1 \quad \text{wenn} \quad t(\vec{M}) > 0 \quad (4.20)$$

$$\vec{M} \rightarrow \Omega_2 \quad \text{wenn} \quad t(\vec{M}) < 0 \quad (4.21)$$

Die Entscheidungsfunktion ist gegeben durch $t(\vec{M}) = 0$ und bildet eine Gerade ($N = 2$) bzw. eine Hyperebene im mehrdimensionalen Raum.

Bei der Erweiterung auf mehrere Klassen bieten sich verschiedene Möglichkeiten zur Realisierung an. Die Auswahl der Vorgehensweise sollte dem Anwendungsfall angemessen sein und kann allgemeingültig nicht beantwortet werden. Beispielhaft sei hier eine Möglichkeit dargestellt:

Jede Klasse wird durch eine spezifische Entscheidungsfunktion von allen anderen Klassen separiert. Bei N_C Klassen ergeben sich dadurch N_C Entscheidungsfunktionen t_i der Art von 4.19 mit ($i = 1, \dots, N_C$). Die Entscheidung über die Klassenzugehörigkeit eines Merkmalsatzes \vec{M} kann über folgende Regelung getroffen werden:

$$\vec{M} \rightarrow \Omega_i \quad \left\{ \begin{array}{l} \text{wenn} \quad t_i(\vec{M}) > 0 \\ \text{wenn} \quad t_j(\vec{M}) < 0 \quad \forall j = 1, \dots, N_C \end{array} \right. \quad (4.22)$$

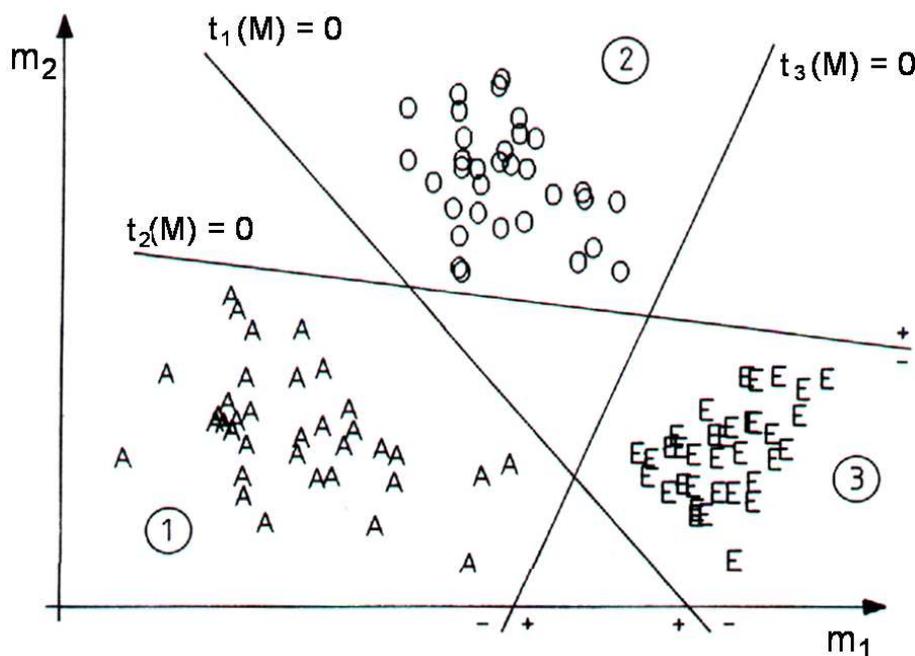


Abbildung 4.8: Linearer Klassifikator für $N = 2$ und $N_C = 3$

Abbildung 4.8 zeigt ein mögliches Szenario mit 3 Klassen im zweidimensionalen Merkmalsraum ($N = 2$). Die Entscheidungsfunktionen t_1, t_2 und t_3 erlauben eine klare Trennung der einzelnen Cluster. Das durch die Entscheidungsfunktionen gebildete Dreieck bildet einen Bereich aus, in dem keine Klassenzugehörigkeit für einen Merkmalsatz \vec{M} ermittelt werden kann. [13]

4.4.2 Allgemeine Entscheidungsfunktionen

Wenn die Clusterverteilung komplex ist, kann es sein, daß eine lineare Entscheidungsfunktion die Trennung der einzelnen Cluster nicht mehr ermöglicht. Die Entscheidungsfunktion kann dahingehend erweitert werden, daß beliebige Funktionen von \vec{M} verwendet werden, von denen lediglich gefordert wird, daß sie reell und einwertig sind:

$$t(\vec{M}) = w_0 + w_1 f_1(\vec{M}) + w_2 f_2(\vec{M}) + \dots + w_N f_N(\vec{M}) \tag{4.23}$$

Die einzelnen Funktionen f_1, \dots, f_N können als eine Basis angesehen werden, aus der durch Linearkombination die Entscheidungsfunktion $t(\vec{M})$ zusammengesetzt wird. Durch die Forderung, daß f_1, \dots, f_N einwertig sein sollen, können allgemeine Entscheidungsfunktionen auf lineare zurückgeführt werden. Dies wird deutlich, indem ein Merkmalsatz \vec{M} folgenderma-

ßen transformiert wird:

$$\vec{M}^* = \begin{pmatrix} f_1(\vec{M}) \\ f_2(\vec{M}) \\ \vdots \\ f_N(\vec{M}) \end{pmatrix} \quad (4.24)$$

Im transformierten Merkmalsraum liegt wieder ein linearer Klassifikator vor, so daß die Feststellungen über lineare Klassifikatoren hier gültig sind. Die Trennflächen einer allgemeinen Entscheidungsfunktion sind allgemeine Kegelschnitte wie zum Beispiel Ebenen, Paraboloiden, Ellipsoide oder Hyperboloide. [13]

4.5 Merkmalsextraktion

In Kapitel 3.1 wurden bereits die Ziele der Merkmalsextraktion formuliert und im Gesamtkontext eines Mustererkennungssystems erläutert. Auf der in Kapitel 4.1 geschaffenen formalen Basis dient die Merkmalsextraktion der Ermittlung einer geeigneten Merkmalsauswahl \vec{A} , die zusammen mit einer zuvor festgelegten Abbildung κ einen Klassifikator $K = (\vec{A}, \kappa)$ laut Definition 4 bildet. Die Reduzierung auf eine Auswahl von Merkmalen ist vorteilhaft, da zum einen weniger Merkmale verarbeitet werden müssen und zum anderen nur Informationen behandelt werden, die tatsächlich charakteristisch sind [13].

Die Merkmalsextraktion ist ein elementarer Bestandteil der Lernphase und findet auf der Grundlage der vorgegebenen Stichproben statt, deren Klassenzugehörigkeit bekannt ist. Formal wurden diese Grundlagen durch die Definition einer Lernbasis (Definition 3) zusammengefaßt. Wie für ein akustisches Güteprüfverfahren eine Lernbasis erstellt werden kann, wurde in Kapitel 4.2 erläutert. Bei Mustererkennungssystemen aus dem Bereich der akustischen Güteprüfung wurden in der Literatur häufig Merkmale nicht durch einen automatisierten Prozeß herausgefunden. In [9] waren beispielsweise die physikalischen Ursachen für bestimmte Fehler und deren spezifische Geräusche bekannt. Durch diese Modellvorstellungen konnten Merkmale für die Schallemissionsuntersuchung vorgegeben werden, die eine Klassifikation erlaubten. Auch in [16] waren physikalische Ursache und Auswirkung auf die Schallemission nachvollziehbar. Diese Vorgehensweise ist jedoch nicht übertragbar, wenn solche Kenntnisse fehlen. Handelt es sich bei den Schallemissionen zudem um sehr komplexe Geräusche, die sich über ein breites Frequenzband erstrecken, wird es für einen Menschen zunehmend schwieriger, Merkmale zu benennen, die für eine spezifische Klasse charakteristisch sind. In [6]

werden Schallemissionen durch eine Auswahl von Vokabeln charakterisiert. Voraussetzung hierfür ist jedoch, daß genügend Vokabeln mit einer der Schallemission angepaßten Semantik zur Verfügung stehen. Das charakteristische, bei Mannesmann Dematic ausschlaggebende, "Knistern" entzieht sich einer solchen Charakterisierbarkeit.

Da dem System "von außen" keine charakteristischen Merkmale bekannt gemacht werden können, ist es erforderlich, ein automatisiertes Suchverfahren mit einem Algorithmus einzusetzen, wie zum Beispiel in [17] oder [13] vorgestellt. Es werden zunächst alle vorliegenden Informationen erfaßt, hier das komplette verfügbare Frequenzspektrum, wobei im Sinne der allgemein gehaltenen Definitionen in Kapitel 4.1 prinzipiell auch andere Informationen zum Einsatz kommen können, solange deren Darstellung in Form von Merkmalsätzen gemäß Definition 1 möglich ist. Zwangsläufig werden auch irrelevante Informationen erfaßt. Diese gilt es bei der Merkmalsextraktion zu finden und durch deren Ausschluß eine Dimensionsreduktion vom N_M -dimensionalen auf den N -dimensionalen Merkmalsraum zu erreichen. Um eine Merkmalsauswahl \vec{A} durch einen Algorithmus bewerten zu können, wird ein geeignetes Gütekriterium benötigt, das Aufschluß über die Qualität von \vec{A} liefert. Ein Algorithmus, der anhand einer vorgegebenen Lernbasis B und einer ebenfalls vorgegebenen Abbildung κ eine im Sinne eines beliebigen Gütekriteriums geeignete Auswahl \vec{A} findet, wird \vec{A} in Richtung einer möglichst maximalen Gütebewertung optimieren.

4.5.1 Gütekriterien

Grundlage für die Bewertung der Güte eines Klassifikators $K = (\vec{A}, \kappa)$ ist häufig die Betrachtung der Klassifikationsergebnisse [13]. Die Abbildung κ wird bei der Suche einer geeigneten Auswahl \vec{A} als gegeben angenommen und entstammt den Modellvorstellungen über die Beschaffenheit von Clustern (siehe Kapitel 4.3).

Für die weiteren Ausführungen sei eine Lernbasis $B = (L, C, \lambda)$ gemäß Definition 3 angenommen.

Reklassifikationsfehler

Ein häufig eingesetztes Verfahren zur Bewertung eines Klassifikators $K = (\vec{A}, \kappa)$ besteht darin, ihn auf die Merkmalsätze der Menge L der Lernbasis B anzuwenden. Der Quotient

$$G_R^N = 1 - \frac{\text{richtig klassifizierte Stichproben der Menge } L}{\text{Gesamtanzahl Stichproben } N_S} = 1 - \frac{S_K}{N_S} \quad (4.25)$$

gibt die Qualität des Klassifikators, angewendet auf die Lernmenge L , an und wird *Reklassifikationsfehler* genannt. Die Nennung von N in G_R^N soll wie bisher auf die Anzahl der verwendeten Merkmale hindeuten. Da die Werte für G_R^N im Intervall von $[0, 1]$ liegen, kann die Funktion $G_R^N(S_K)$ auch als Fuzzy Zugehörigkeitsfunktion zu umgangssprachlichen Begriffen wie zum Beispiel "geeignet" oder "gut" betrachtet werden.

Die Nutzung des Reklassifikationsfehlers als wichtiges Gütekriterium entspricht den Vorstellungen von L als Lernmenge und führt zu einer Optimierung von K in Richtung der Abbildung λ , die das menschliche Expertenwissen repräsentiert. Bei einem Reklassifikationsfehler von 0.0 - alle Stichproben werden durch K richtig klassifiziert - entspricht K dem als ideal bezeichneten Klassifikator laut Gleichung 4.2 in Kapitel 4.1. In diesem Fall klassifiziert K alle Stichproben aus der Lernmenge L so, wie es der menschliche Experte vorgegeben hat. Implizit wird erwartet, daß neue und unbekannte Testobjekte in der Einsatzphase durch K ebenfalls so klassifiziert werden, wie es der menschliche Experte getan hätte. Daher kann die tatsächliche Leistungsfähigkeit eines Klassifikators letztendlich nur über seine Klassifikationsergebnisse in der Einsatzphase beurteilt werden, die zum Zeitpunkt der Lernphase jedoch noch nicht zur Verfügung stehen können. Vorteilhaft bei der Verwendung des Reklassifikationsfehlers ist, daß bereits eine geringe Anzahl an Stichproben $N_S < 50$ in der Lernmenge L ausreicht, um aussagekräftige Werte für G_R^N zu erhalten [17].

Bei dem vorgestellten Klassifikationskonzept aus Kapitel 4.3 wird die Entscheidung über das Klassifikationsergebnis durch die Betrachtung der maximalen Zugehörigkeit gefällt, die auch als *Hauptzugehörigkeit* $\hat{\mu}$ bezeichnet wird. Der zweitgrößte Zugehörigkeitswert, der neben $\hat{\mu}$ ermittelt wurde, wird als größte *Nebenzugehörigkeit* $\check{\mu}$ bezeichnet. Bei der Betrachtung des Reklassifikationsfehlers wird nur berücksichtigt, ob die Hauptzugehörigkeit $\hat{\mu}$ für die richtige, durch λ in der Lernbasis vorgegebene Klasse berechnet wurde. Die absolute Größe der Hauptzugehörigkeit $\hat{\mu}$ wird allerdings nicht mit einbezogen. Im semantischen Sinne der Zugehörigkeitsfunktion würde eine von der absoluten Größe betrachtete kleine Hauptzugehörigkeit $\hat{\mu}$ für eine unsichere Entscheidung sprechen, da sich die Stichprobe scheinbar keiner Klasse mit hoher Sicherheit zuordnen läßt. Möglich ist die Einführung einer Grenzzugehörigkeit μ_{Grenz} , so daß für Stichproben $M_i \in L$, für die

$$\mu(M_i) \leq \mu_{\text{Grenz}} \quad \text{mit beispielsweise } \mu_{\text{Grenz}} = 0.1 \quad (4.26)$$

gilt, kein Klassifikationsergebnis ermittelt wird bzw. die Klassifikation von M_i als "falsch" gewertet wird.

Mit dem Quotienten

$$V_\mu = \frac{\check{\mu}}{\hat{\mu}}, \quad (4.27)$$

der Werte im Intervall von $[0, 1]$ annimmt, kann dagegen das Verhältnis von Haupt- und größter Nebenzugehörigkeit in Betracht gezogen werden. Hohe Werte von V_μ sprechen für eine geringe Differenz zwischen der Hauptzugehörigkeit $\hat{\mu}$ und der größten Nebenzugehörigkeit $\check{\mu}$. Daraus läßt sich ebenfalls eine gewisse Unsicherheit über das getroffene Klassifikationsergebnis ableiten, da die Zugehörigkeit des Testobjektes zu mindestens zwei Klassen beinahe gleich hoch ist. [17]

Zugehörigkeitsdifferenz

Um ein Maß für die Deutlichkeit der Klassifikationsentscheidung zu erhalten, kann die Differenz der Hauptzugehörigkeit und der größten Nebenzugehörigkeit betrachtet werden:

$$G_D(\vec{M}) = \hat{\mu}(\vec{M}) - \check{\mu}(\vec{M}) \quad (4.28)$$

Je größer G_D wird, desto eindeutiger ist das Klassifikationsergebnis. Es ist zu beachten, daß bei kleinen Werten für die Hauptzugehörigkeit $\hat{\mu}$ auch G_D eher klein ist, da das Verhältnis von V_μ (nach Gleichung 4.27) bei der Berechnung von G_D nicht einbezogen wird.

Trennschärfe

Allgemeiner kann die Zugehörigkeitsdifferenz als Trennschärfe betrachtet werden, die nicht nur die größte, sondern alle Nebenzugehörigkeiten einbezieht. Dazu werden die Zugehörigkeitsdifferenzen über alle Klassen gemittelt:

$$G_T(\vec{M}) = \frac{\sum_{k=1}^{N_C} \hat{\mu}(\vec{M}) - \mu_k(\vec{M})}{\hat{\mu}(\vec{M})(N_C - 1)} \quad (4.29)$$

Dabei bezeichne $\mu_k(\vec{M})$ die Zugehörigkeit von \vec{M} zur Klasse Ω_k (mit $k = 1, \dots, N_C$). Auch die Funktion G_T liefert Werte im Intervall $[0, 1]$ und kann als Fuzzy Zugehörigkeitsfunktion zu dem semantischen Begriff "eindeutig" betrachtet werden.

Durch die Einbeziehung des Durchschnitts über alle Klassenzugehörigkeiten wird das nahe Beieinanderliegen von der Hauptzugehörigkeit $\hat{\mu}(\vec{M})$ zur größten Nebenzugehörigkeit $\check{\mu}(\vec{M})$ nach unten relativiert, sofern die restlichen Nebenzugehörigkeiten deutlich kleiner sind. Die Tabelle verdeutlicht exemplarisch für zwei angenommene Merkmalsätze \vec{M}_i mit $i = 1, 2$ diese Situation:

i	$\mu_1(\vec{M}_i)$	$\mu_2(\vec{M}_i)$	$\mu_3(\vec{M}_i)$	$\mu_4(\vec{M}_i)$	$G_T(\vec{M}_i)$	$G_V(\vec{M}_i)$
1	0.95	0.05	0.05	0.05	0.947	0.947
2	0.80	0.75	0.10	0.05	0.625	0.0625

Tabelle 4.5.1: Beispiele für die Trennschärfe G_T und G_V

Für das Szenario $i = 1$ liefert $G_T(\vec{M}_1)$ mit 0.947 ein vernünftiges Ergebnis, da die Differenz zu allen Nebenzugehörigkeiten groß ist. Die Trennschärfe $G_T(\vec{M}_2)$ ist mit 0.625 höher als es die Differenz von nur 0.05 zwischen Haupt- und größter Nebenzugehörigkeit rechtfertigen würde. Da die Nebenzugehörigkeiten von μ_3 und μ_4 jedoch sehr gering ausfallen und bei ihnen eine deutliche Trennung zur Hauptzugehörigkeit vorliegt, liefert $G_T(\vec{M}_2)$ einen recht hohen Wert. Dieser Nachteil läßt sich beseitigen, indem lediglich das Verhältnis von Haupt- und größter Nebenzugehörigkeit betrachtet wird, wie es bereits in Gleichung 4.27 geschehen ist. Der Quotient kann geeignet in eine Zugehörigkeitsfunktion eingebracht werden:

$$G_V(\vec{M}) = 1 - \frac{\check{\mu}(\vec{M})}{\hat{\mu}(\vec{M})} \quad (4.30)$$

Tabelle 4.5.1 zeigt beispielhaft die Werte für G_V im Vergleich zu G_T . Da beim Beispiel $i = 2$ die Eindeutigkeit der Klassentrennung nicht vorhanden ist, liefert G_V im Gegensatz zu G_T wie verlangt einen geringen Wert. [17]

Gemittelte Trennschärfe

Im Gegensatz zum Reklassifikationsfehler bezieht sich die Trennschärfe als Gütekriterium lediglich auf ein einziges Klassifikationsergebnis. Zur Verwendung der Trennschärfe als Gütekriterium innerhalb der Merkmalsextraktion ist es sinnvoll, die Trennschärfe über alle richtig reklassifizierten Stichproben zu mitteln. Die gemittelte Trennschärfe G_{Vm} wird wie folgt errechnet:

$$G_{Vm} = \sum_{\vec{M} \in L_R} \frac{G_V(\vec{M})}{|L_R|} \quad (4.31)$$

wobei

$$L_R \subseteq L \quad : \quad \text{Menge der richtig reklassifizierten Stichproben}$$

$$|L_R| \leq N_S \quad : \quad \text{Anzahl der richtig reklassifizierten Stichproben}$$

Die Optimierung der zu findenden Merkmalsauswahl \vec{A} anhand der gemittelten Trennschärfe G_{V_m} führt zu einer möglichst hohen Trennschärfe G_V gemäß Gleichung 4.30 für die Elemente der Stichprobenmenge L .

Denkbar ist eine Verwendung der gemittelten Trennschärfe G_{V_m} im Zusammenspiel mit dem Reklassifikationsfehler G_R . Wenn die Anzahl der Stichproben N_S der Lernmenge L eher klein ist, können sich für den Reklassifikationsfehler schnell identische Werte ergeben. In diesem Fall kann anhand der gemittelten Trennschärfe eine zusätzliche Beurteilung über eine Auswahl \vec{A} ermittelt werden.

Tabelle 4.5.1 zeigt noch einmal im Überblick die vorgestellten Gütekriterien sowie deren Anwendungsmöglichkeit:

Bezeichnung	Beurteilung von
Reklassifikationsfehler G_R	Lernmenge
Zugehörigkeitsdifferenz G_D	einzelnes Klassifikationsergebnis
Trennschärfe G_V	einzelnes Klassifikationsergebnis
Gemittelte Trennschärfe G_{V_m}	Lernmenge

Tabelle 4.5.1: Die Gütekriterien im Überblick

4.5.2 Auswahlverfahren

In Kapitel 3.2 wurde bereits ein Grundproblem, das sich bei der Ermittlung einer geeigneten Merkmalsauswahl stellt, umrissen. Es sei angenommen, daß in einer Lernbasis $B = (L, C, \lambda)$ die Merkmalsätze $M_i \in L$ aus N_M Merkmalen bestehen. Zu B und einer gegebenen Abbildung κ wird nun eine Merkmalsauswahl \vec{A} mit N Merkmalen gesucht, die zusammen mit κ

den Klassifikator $K = (\vec{A}, \kappa)$ bildet. Da keine a priori Kenntnisse über geeignete Merkmale vorliegen, würde eine vollständige Untersuchung aller möglichen Auswahlen \vec{A} genau

$$\binom{N_M}{N} = \frac{N_M!}{N! \cdot (N_M - N)!} \quad (4.32)$$

Analysen erforderlich machen. Bereits bei recht kleinen Werten für N und N_M ist es nicht mehr praktikabel alle möglichen Auswahlen einzeln zu bewerten. Bei $N_M = 400$ und $N = 10$ existieren $\binom{400}{10} = 2,58 \cdot 10^{19}$ verschiedene Auswahlen⁹. Diese alle bezüglich ihrer Tauglichkeit zu überprüfen, führt schnell zu unpraktikablen Rechenzeiten. Bei Verwendung des in Kapitel 4.3 dargestellten Klassifikationskonzeptes mit dem Reklassifikationsfehler als Gütekriterium muß für jede Auswahl folgendes bestimmt werden:

1. Bestimmung der Clusterschwerpunkte \vec{r}_k für jede Klasse Ω_k
2. Erstellen der Kovarianzmatrix V_k für jede Klasse Ω_k
3. Bestimmen der Eigenwerte (Hauptachsentransformation)
4. Reklassifikation aller zur Lernmenge L gehörenden Merkmalsätze
5. Bestimmung des Reklassifikationsfehlers G_R^N

Um nicht alle möglichen Merkmalsauswahlen bewerten zu müssen, wird ein Algorithmus verwendet, der als Ergebnis eine suboptimale Merkmalsauswahl \vec{A} liefert. [13, 17]

Der Algorithmus sieht eine sukzessive Ausweitung der Merkmalsauswahl um jeweils ein Merkmal vor, bis schrittweise eine Auswahl \vec{A} mit N Merkmalen gefunden wurde. Da die Anzahl der Merkmale einer Merkmalsauswahl bisher immer mit N bezeichnet wurde, die Anzahl N hier aber erst das Ergebnis des Algorithmus ist, wird im folgenden i als laufender Index für die Merkmalsanzahl verwendet. Der Ablauf des Algorithmus sieht wie folgt aus:

Initialisierungsphase

Für eine gegebene Anzahl a von Merkmalen werden alle $\binom{N_M}{a}$ Merkmalskombinationen mit $i = a$ Merkmalen überprüft (mit $a \in \mathbf{N}$). Jede dieser Auswahlen wird anhand eines Gütekriteriums beurteilt, das zu einer Bewertung der Auswahl führt. Die Merkmalsauswahlen werden in absteigender Ordnung gemäß des Gütekriteriums sortiert und in einer Bestenliste BL_i festgehalten, wobei die Auswahlen $\vec{A}_{j,i}$ als ersten Index ihre Position in der Liste und als zweiten Index ihren Berechnungsschritt führen sollen. Nach der Initialisierungsphase wird

⁹In [17] wird für $N = 5-20$ Merkmale aus einer Menge von $N_M = 400$ angenommen (siehe dort S. 64).

die beste Auswahl mit $\vec{A}_{1,a}$, die zweitbeste mit $\vec{A}_{2,a}$ und die b -te beste Auswahl mit $\vec{A}_{b,a}$ bezeichnet. Für die dazwischen liegenden Auswahlen gilt entsprechendes.

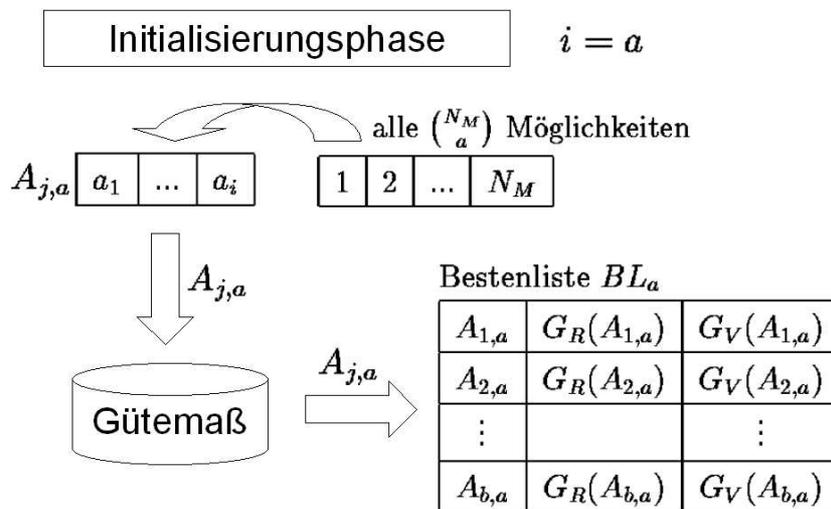


Abbildung 4.9: Schematischer Ablauf der Initialisierungsphase

Abbildung 4.9 zeigt schematisch den Ablauf der Initialisierungsphase.

Sukzessive Erweiterung um ein Merkmal

In einer Schleife wird i schrittweise um 1 erhöht, so daß $i = i + 1$ mit $a < i \leq N$ gilt.

Im Schritt i werden sukzessive die besten b Merkmalsauswahlen $\vec{A}_{j,i-1}$ aus BL_{i-1} , die im Schritt $i - 1$ ermittelt wurden und $i - 1$ Merkmale enthalten, als Grundlage genutzt (mit $1 \leq j \leq b$, $b \in \mathbf{N}$). Sind in BL_{i-1} weniger als b Auswahlen vorhanden, so wird nur die tatsächlich vorhandene Anzahl genutzt.

Jede dieser Auswahlen $\vec{A}_{j,i-1}$ wird um jedes mögliche i -te Merkmal erweitert, so daß es sich bei $\vec{A}_{j,i-1}$ nach dem Hinzufügen des i -ten Merkmals um eine gültige Merkmalsauswahl mit i Merkmalen laut Definition 2 handelt und insbesondere kein Merkmal doppelt in der Auswahl vorkommt. Für die Erweiterung einer Auswahl um das i -te Merkmal stehen bei einer Gesamtanzahl von N_M Merkmalen $N_M - i + 1$ Merkmale zur Verfügung. Insgesamt müssen im Schritt i maximal $b \cdot (N_M - i + 1)$ Merkmalsauswahlen entsprechend des festgelegten Gütekriteriums bewertet werden. Wie in der Initialisierungsphase werden die Merkmalsauswahlen in absteigender Ordnung gemäß ihrer Bewertung in der Bestenliste BL_i sortiert.

Die Schleife wird solange fortgesetzt bis $i = N$ gilt.

Abbildung 4.10 zeigt schematisch den Ablauf der sukzessiven Erweiterung.

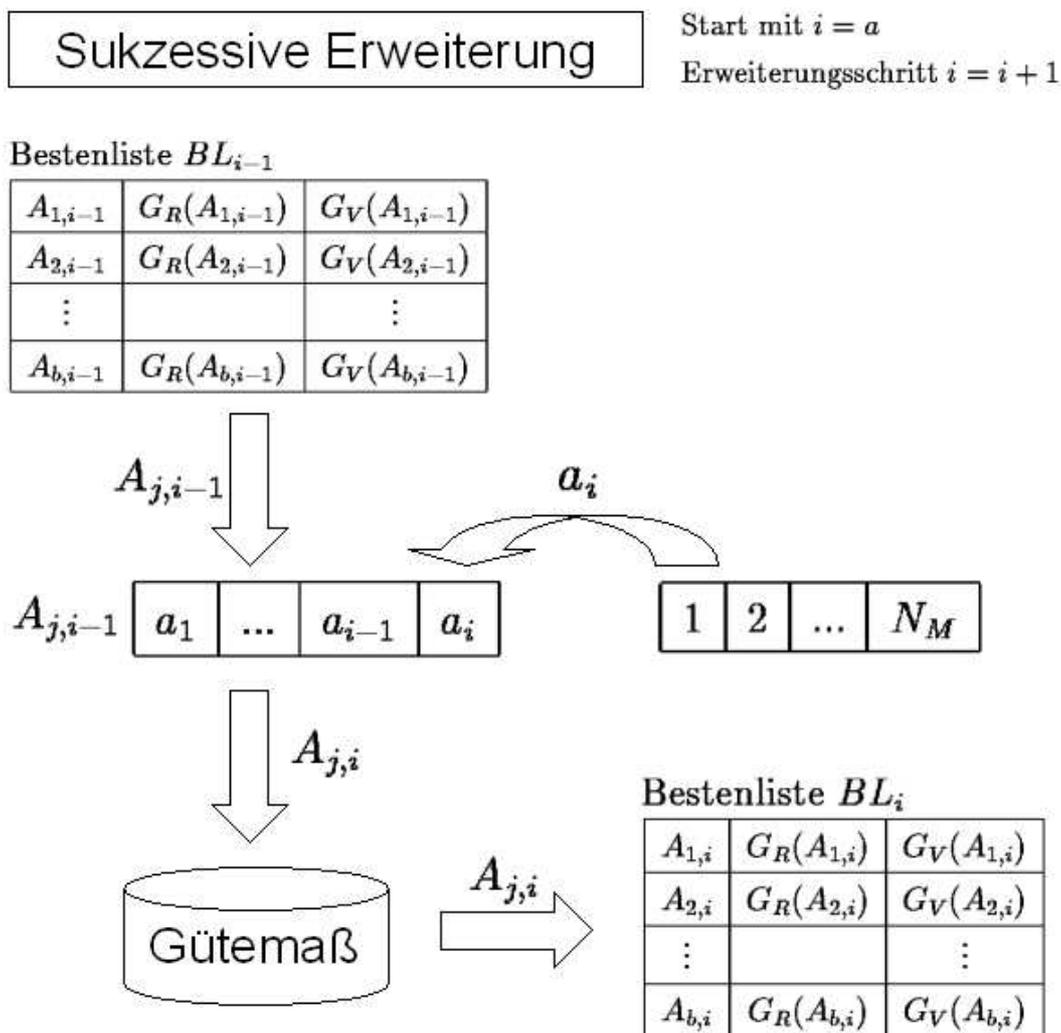


Abbildung 4.10: Schematischer Ablauf eines Schrittes der Erweiterungsphase

Ergebnis

Nach Beendigung der Schleife mit $i = N$ liefert der Algorithmus eine Bestenliste BL_N , die in absteigender Ordnung die besten Merkmalsauswahlen mit N Merkmalen enthält, die im Rahmen des suboptimalen Verfahrens überprüft wurden. Das oberste bzw. erste Element der Liste kann als Lösung der Aufgabe, eine suboptimale Merkmalsauswahl \vec{A} zu finden, angesehen werden.

Der Algorithmus kann auf verschiedene Arten parametrisiert werden:

- Der Parameter a gibt an, bis zu welcher Merkmalsanzahl sämtliche möglichen Merkmalsauswahlen überprüft werden sollen. Bei $a = 3$ werden beispielsweise alle möglichen $\binom{N_M}{3}$ Merkmalsauswahlen mit drei Merkmalen überprüft. Für a muß gelten:

$1 \leq a \leq N$. Im Fall $a = N$ wird bereits in der Initialisierungsphase die komplette Untersuchung aller möglichen $\binom{N_M}{N}$ Merkmalsauswahlen mit N Merkmalen vorgenommen. Die Schleife wird in diesem Fall nicht mehr durchgeführt, da unmittelbar die Abbruchbedingung $i > N$ vorliegt.

- Der Parameter b bestimmt ab der Merkmalsanzahl $i > a$, wieviele der besten Merkmalsauswahlen mit $i - 1$ Merkmalen aus der Bestenliste BL_{i-1} im Schritt i weiterverwendet werden. Nur sie dienen als Grundlage zur Erweiterung um ein i -tes Merkmal. Für b muß $b \geq 1$ gelten.

Des weiteren kann der Algorithmus durch das verwendete Klassifikationskonzept sowie das eingesetzte Gütekriterium modifiziert werden. Die Anzahl der Merkmale N_M in den Merkmalsätzen wird im Rahmen der Signalerfassung und -verarbeitung festgelegt und ist entsprechend Definition 3 innerhalb einer Lernbasis für alle Merkmalsätze gleich.

4.5.3 Laufzeitbetrachtung

Die Laufzeit des Algorithmus hängt von verschiedenen Einflußgrößen ab. Zum einen ist die Laufzeit für die Errechnung des Gütekriteriums einer einzigen zu bewertenden Auswahl zu betrachten, zum anderen stellt sich die Frage, wie häufig eine solche Bewertung vorgenommen wird.

Für die Bewertung einer einzigen Auswahl nach dem in Kapitel 4.3 vorgestellten Klassifikationskonzept und der Verwendung des Reklassifikationsfehlers als Gütekriterium sind die in Abschnitt 4.5.2 genannten Schritte mit folgenden Laufzeiten durchzuführen:

1. Bestimmung der Clusterschwerpunkte für jede Klasse: $O(N)$
2. Erstellen der Kovarianzmatrix für jede Klasse: $O(N^2)$
3. Bestimmen der Eigenwerte (Hauptachsentransformation): $O(N^3)$ (bei Verwendung des Jacobi Algorithmus [12])
4. Reklassifikation aller zur Lernmenge L gehörenden Merkmalsätze und Bestimmung des Reklassifikationsfehlers G_R^N : $O(N_S)$

Insgesamt ist für eine einzelne Bewertung einer Auswahl \vec{A} eine Laufzeit von $O(N^3)$ zu veranschlagen. Bleibt zu ermitteln, wie viele solcher Bewertungen vorzunehmen sind. Dies ist von den verschiedenen Parametern abhängig: Zum einen der Anzahl der zur Verfügung

stehenden Merkmale N_M aus der Lernmenge, zum anderen der Anzahl N der Merkmale in der zu findenden Merkmalsauswahl \vec{A} . Bei einer vollständigen Überprüfung aller möglichen Merkmalsauswahlen mit N Merkmalen sind $\binom{N_M}{N}$ mögliche Merkmalsauswahlen zu überprüfen. Durch den vorgestellten Algorithmus wird jedoch eine erhebliche Einschränkung der zu untersuchenden Merkmalsauswahlen vorgenommen. Zum einen ist die Laufzeit der Initialisierungsphase zu bestimmen, in der alle Merkmalsauswahlen mit a Merkmalen untersucht werden. Dort müssen insgesamt

$$\binom{N_M}{a}$$

Bewertungen vorgenommen werden. Zum anderen ist die Laufzeit der Schleife zu betrachten, in der sukzessive die Anzahl der Merkmale erhöht wird. Als Grundlage für die Erweiterung werden jedoch nur die besten b Merkmalsauswahlen aus dem vorangegangenen Schritt bzw. aus der Initialisierungsphase genutzt. Sind aus dem vorherigen Schritt weniger als b Merkmalsauswahlen in die Bestenliste aufgenommen worden, wird nur die tatsächlich vorhandene Anzahl an Merkmalsauswahlen als Grundlage zur Erweiterung genutzt. Daher sind die in den folgenden Überlegungen angegebenen Laufzeiten stets als obere Grenze zu betrachten. Nach der Initialisierungsphase werden in jedem Schritt i (für $a < i \leq N$)

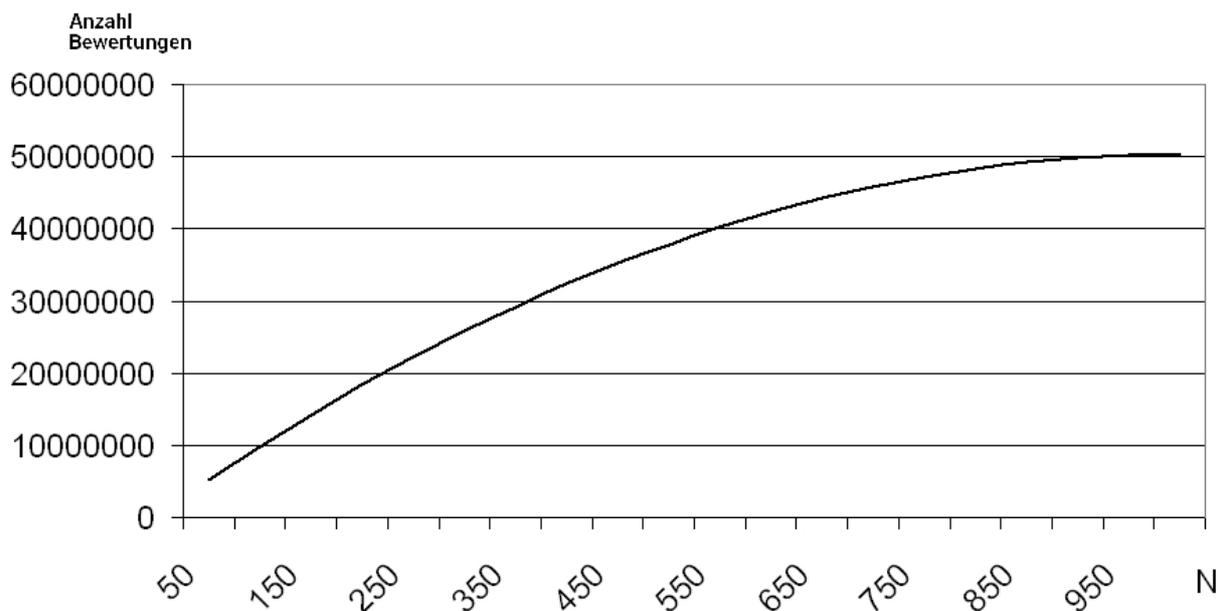
$$b \cdot (N_M - (i - 1)) = b \cdot (N_M - i + 1) \tag{4.33}$$

Bewertungen durchgeführt. Der Ausdruck $N_M - i + 1$ entspricht der Anzahl der Erweiterungsmöglichkeiten im Schritt i . Als Gesamtanzahl der Bewertungen ergibt sich:

$$\begin{aligned} \textbf{Initialisierung} & : \binom{N_M}{a} = \frac{N_M!}{a! \cdot (N_M - a)!} \\ \textbf{Schleife} & : \\ i = a + 1 & : b \cdot (N_M - (a + 1) - 1) \quad \text{mit } N = a + 1 \text{ Merkmalen} \\ i = a + 2 & : b \cdot (N_M - (a + 2) - 1) \quad \text{mit } N = a + 2 \text{ Merkmalen} \\ i = a + 3 & : b \cdot (N_M - (a + 3) - 1) \quad \text{mit } N = a + 3 \text{ Merkmalen} \\ & \vdots \\ i = N & : b \cdot (N_M - (N) - 1) \quad \text{mit } N \text{ Merkmalen} \end{aligned}$$

oder unter Berücksichtigung der Initialisierungsphase in mathematisch geschlossener Form [15] :

$$\binom{N_M}{a} + \sum_{i=a+1}^N b \cdot (N_M - (i - 1))$$


 Abbildung 4.11: Bewertungen in Abhängigkeit von N mit $a = 2, b = 100, N_M = 1000$

$$\begin{aligned}
 &= \binom{N_M}{a} + \sum_{i=a+1}^N b \cdot N_M - b \cdot i + b \\
 &= \binom{N_M}{a} + N_M b (N - a) + b(N - a) - \sum_{i=a+1}^N b \cdot i \\
 &= \binom{N_M}{a} + N N_M b - N_M a b + N b - a b - \frac{1}{2} b (N - a)(N + a + 1) \\
 &= \binom{N_M}{a} + N N_M b + N b - N_M a b - a b - \frac{1}{2} N^2 b - \frac{1}{2} N b + \frac{1}{2} a^2 b + \frac{1}{2} a b \\
 &= \binom{N_M}{a} + \frac{1}{2} b \cdot (2 N N_M - N^2 + N - 2 N_M a - a + a^2) \tag{4.34}
 \end{aligned}$$

Die hergeleitete Gleichung erlaubt es, die Anzahl der Bewertungen für vorgegebene N, N_M, a, b im voraus zu ermitteln. Während es bei der Berechnung der Laufzeit einer einzelnen Bewertung Sinn gemacht hat, die Laufzeit alleine in Abhängigkeit von N zu ermitteln, ist dies bei der Betrachtung der Schleifendurchläufe in Frage zu stellen, da eine starke Abhängigkeit von N und N_M durch $N \leq N_M$ besteht.

Abbildung 4.11 zeigt die Anzahl der Bewertungen nach Gleichung 4.34 in Abhängigkeit von N mit den beispielhaften Parametern $a = 2, b = 100, N_M = 1000$. Für $N \rightarrow N_M$ läuft die Funktion asymptotisch gegen eine Gerade. Dies folgt offensichtlich aus der Zusammensetzung der einzelnen Summanden, aus denen die geschlossene Darstellung abgeleitet wurde:

$$b \cdot (N_M - N + 1) \tag{4.35}$$

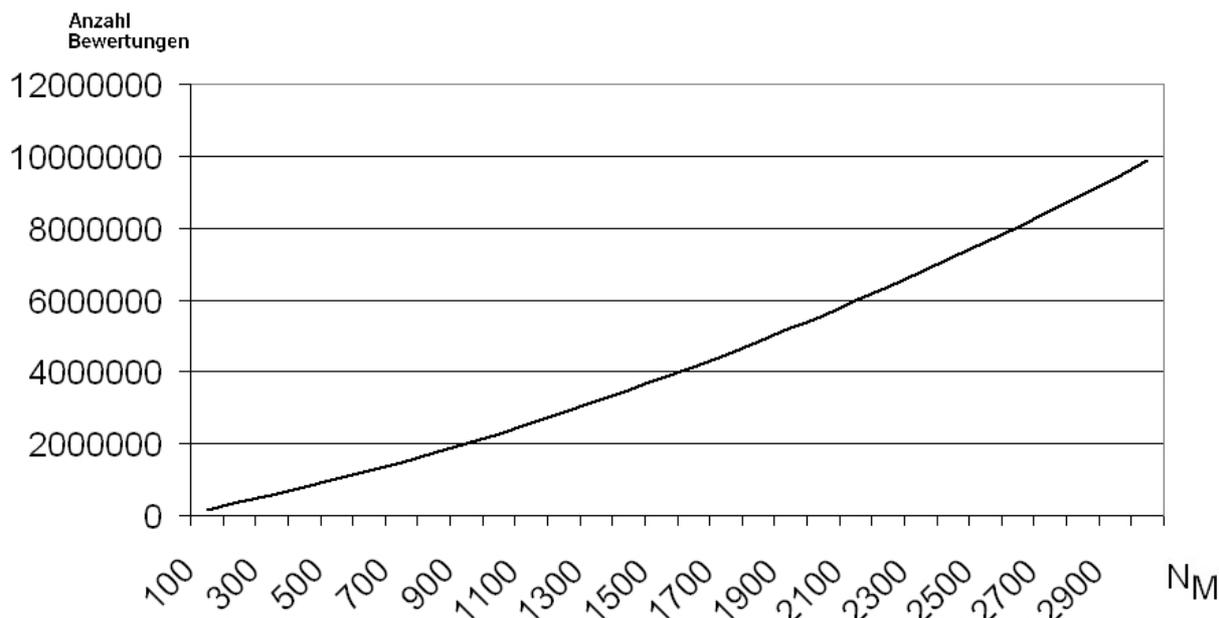


Abbildung 4.12: Bewertungen in Abhängigkeit von N_M mit $a = 2, b = 100, N = 20$

Läuft N gegen N_M werden die Summanden immer kleiner, bis sich für $N = N_M$ nach Gleichung 4.34:

$$\begin{aligned} & \binom{N_M}{a} + \frac{1}{2}b \cdot (2N_M^2 - N_M^2 + N_M - 2N_M a - a + a^2) \\ &= \binom{N_M}{a} + \frac{1}{2}b \cdot ((N_M - a)^2 + N_M - a) \end{aligned} \tag{4.36}$$

als maximale Anzahl der Schleifendurchläufe ergibt, für das Rechenbeispiel $5.035 \cdot 10^7$ Bewertungen.

Da N durch N_M beschränkt ist, besteht ein alternativer Ansatz in der Betrachtung der Laufzeit in Abhängigkeit von N_M bei fest gewähltem N . Abbildung 4.12 läßt ein exponentielles Wachstum der benötigten Bewertungen für steigende N_M erkennen ($N = 20, a = 2, b = 100$). Bei genauer Betrachtung von Gleichung 4.34 ist dieses exponentielle Wachstum durch die Initialisierungsphase mit $\binom{N_M}{a}$ begründet. Setzt man $a = 1$, so daß dieser Summand konstant N_M wird, bleibt ein lineares Wachstum wie es Abbildung 4.13 zeigt.

Die Differenz zwischen der Untersuchung aller möglichen $\binom{N_M}{N}$ Merkmalsauswahlen und der eingeschränkten Betrachtung durch den Algorithmus ist beträchtlich. Im bereits oben genutzten Beispielszenario mit den Parametern $N_M = 1000, a = 2, b = 100$ stehen sich folgende Anzahlen an vorzunehmenden Bewertungen gegenüber:

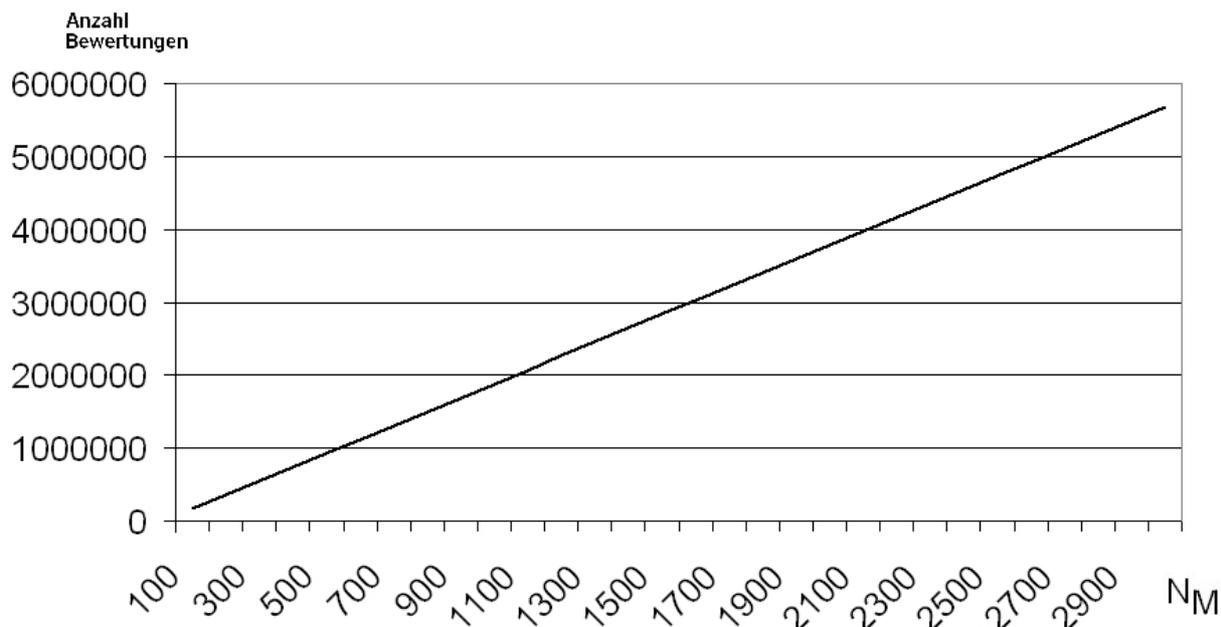


Abbildung 4.13: Bewertungen in Abhängigkeit von N_M mit $a = 1, b = 100, N = 20$

N	$\binom{N_M}{N}$	nach Algorithmus
2	499500	499500
3	166167000	599300
5	$8.250 \cdot 10^{12}$	798600
10	$2.658 \cdot 10^{23}$	1295100
15	$6.881 \cdot 10^{32}$	1789100
20	$3.395 \cdot 10^{41}$	2280600

4.5.4 Überadaption

Das Ziel der Merkmalsextraktion besteht im Finden einer geeigneten Auswahl \vec{A} , die in Verbindung mit einer aus einem Klassifikationskonzept stammenden Abbildung κ einen Klassifikator $K = (\vec{A}, \kappa)$ bildet. Dieser elementare Bestandteil der Lernphase wird im wesentlichen durch das verwendete Gütekriterium beeinflusst. Der Versuch, eine bessere Auswahl \vec{A} zu finden, führt zu einer Optimierung des Gütekriteriums, die nicht gleichzusetzen ist mit der Optimierung des gesamten Mustererkennungssystems.

Durch die Verwendung des Reklassifikationsfehlers (siehe Kapitel 4.5.1) wird die Auswahl \vec{A} so optimiert, daß die Stichprobenmenge L der Lernbasis durch den Klassifikator $K = (\vec{A}, \kappa)$ möglichst gut reklassifiziert wird. Problematisch ist, daß der Klassifikator K sich zu sehr an

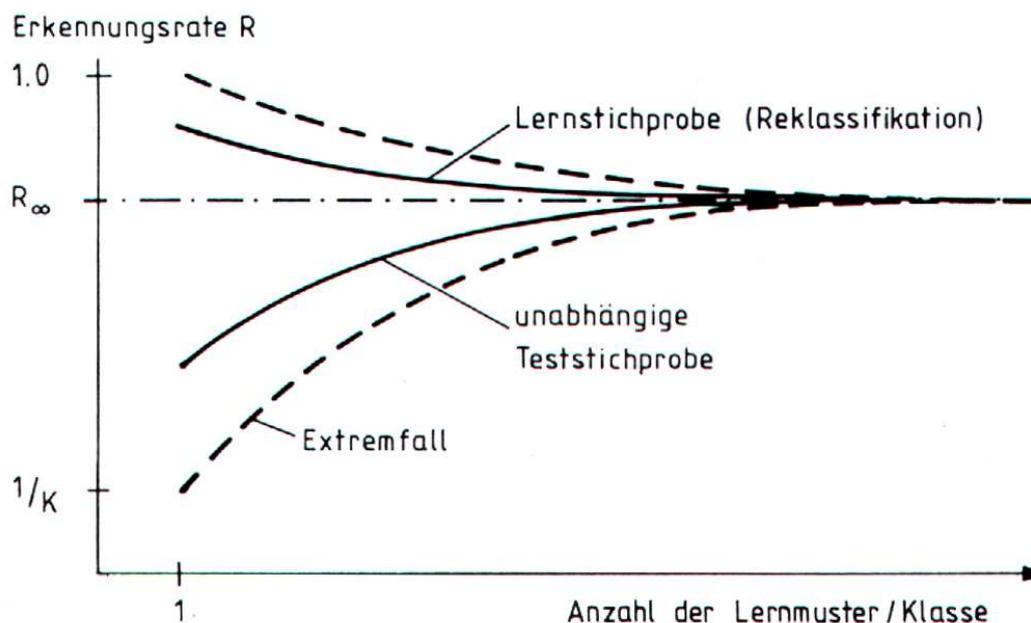


Abbildung 4.14: Abhängigkeit der Erkennungsrate zur Anzahl der Stichproben (schematisch)

die Eigenschaften der Lernmenge L anpassen kann. Dadurch ist es möglich, daß zwar die Lernmenge besonders gut reklassifiziert wird, in der späteren Einsatzphase aber schlechtere Ergebnisse erzielt werden als mit einem Klassifikator, der nicht so sehr an die Lernmenge angepaßt ist. Eine solche sogenannte *Überadaption* kann vor allem dann auftreten, wenn die Anzahl der Stichproben pro Klasse Ω_k in der Lernphase klein ist.

Abbildung 4.14 zeigt diesen Sachverhalt in schematischer Darstellung. Bei einer geringen Anzahl an Lernstichproben pro Klasse (mindestens eine) wird tendenziell ein besseres Reklassifikationsergebnis erzielt als mit einer größeren Anzahl an Stichproben. Benutzt man den ermittelten Klassifikator jedoch mit unabhängigen weiteren Teststichproben werden tendenziell eher schlechtere Klassifikationsergebnisse erzielt, als wenn eine größere Anzahl von Stichproben zur Merkmalsextraktion zur Verfügung gestanden hätte. Je größer die Anzahl der Merkmalsätze in der Lernmenge ist, desto schlechter werden die Reklassifikationsergebnisse und desto besser die Ergebnisse in der Einsatzphase. Im Extremfall werden alle Stichproben richtig reklassifiziert, die Erkennung der weiteren Teststichproben sinkt aber auf die Ratewahrscheinlichkeit

Durch den Einsatz von weiteren Teststichproben, die nicht Elemente der Menge L der Lernbasis sind, deren Klassenzugehörigkeit jedoch bekannt ist, kann eine unabhängige Bewertung eines Klassifikators K durchgeführt werden, indem diese Teststichproben ebenfalls reklassi-

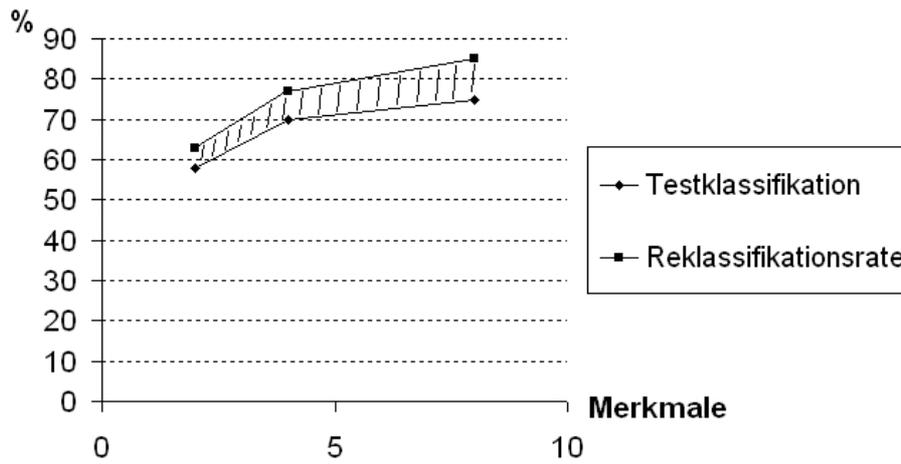


Abbildung 4.15: Diskrepanz zwischen Reklassifikation der Lernmenge und der unabhängigen Teststichproben (schematisch)

fiziert werden.

Abbildung 4.15 zeigt schematisch die Diskrepanz die sich zwischen der Reklassifikation der Lernmenge L einer Lernbasis und weiteren unabhängigen Teststichproben ergeben kann. Durch den Vergleich der Klassifikationsergebnisse kann ein Maß für die Überadaption abgeleitet werden. [13]

Kapitel 5

Software

Die vorgestellten Konzepte und der Algorithmus zur suboptimalen Merkmalsextraktion wurden im Rahmen dieser Arbeit in zwei Softwaretools umgesetzt:

- WAV Analyse
- KlassifikationsTool

Beide Programme wurden in der Programmiersprache C++ unter Microsoft Windows NT4.0 mit der Entwicklungsumgebung Microsoft Visual C++ 5.0 erstellt. Beiden Programmen liegen zum Teil die gleichen Klassen zu Grunde.

5.1 WAV Analyse

5.1.1 Funktionsumfang

Das Programm "*WAV Analyse*" bietet verschiedene Analyse- und Visualisierungsmöglichkeiten für digitale Audiodateien im WAV-Format ¹ an. Die Leistungspunkte im einzelnen sind:

- Einladen von WAV-Dateien mit 44.1 kHz, 16 Bit, mono
- Erzeugen und Mischen von Schallsignalen mit vorgegebenen Frequenzen
- Herausfiltern bzw. Ausblenden eines beliebigen Frequenzbandes

¹Digitales Audioformat der Firma Microsoft. Ohne Zusatzsoftware abspielbar unter Microsoft Windows NT4.0/95/98.

- Abspeichern von gegebenenfalls modifizierten Dateien

Die Visualisierungsfunktionen sehen alle als Ausgabe eine Bilddatei im BMP-Format ² vor. Folgende Ausgaben sind möglich:

- Amplitudenansicht; Darstellung des zeitlichen Schallsignals
- Spektralansicht; Darstellung des Leistungsdichtespektrums
- 3D-Spektralansicht; Darstellung des Leistungsdichtespektrums über kurze Zeitintervalle in einer dreidimensionalen Ansicht

Das Programm eignet sich vor allem zur Visualisierung von Schallemissionen und kann beispielsweise verdeutlichen, daß der Einsatz von automatisierten Merkmalsextraktionsverfahren erforderlich ist, da von einem Menschen keine charakteristischen Frequenzen erkannt werden bzw. diese in der Visualisierung nicht sichtbar werden.

5.1.2 Bedienung

Das Programm wird durch eine grafische Benutzeroberfläche bedient, die in Abbildung 5.1 dargestellt ist. Die genannten Funktionen sind über die selbsterklärenden Buttons zugänglich. Ein Teil der Funktionen kann über die direkt in der Nähe ihrer Buttons angeordneten Textfelder parametrisiert werden.

Es bestehen zwei verschiedene Wege eine WAV-Datei in den Arbeitsspeicher zu bekommen. Zum einen kann eine neue WAV-Datei erzeugt werden. Dies erfolgt über den Button "WAV-Datei erzeugen", wobei die gewünschte Anzahl der Samples im benachbarten Textfeld angegeben werden kann. Zum anderen kann eine bestehende WAV-Datei eingeladen werden. Dazu wird der Button "WAV-Datei laden" betätigt. Beide Funktionen unterstützen wie eingangs beschrieben ausschließlich das WAV-Format: 44100 Hz, 16 Bit, mono.

Zum Erzeugen eines Signals (beispielsweise in eine zuvor erzeugte leere WAV-Datei hinein) wird die gewünschte Frequenz, die Amplitude des Signals (aus dem 16-Bit Wertebereich von -32768 bis 32767), die gewünschte Startposition innerhalb der WAV-Datei in Samples sowie die gewünschte Länge des Signals ebenfalls in Samples angegeben. Das Mischen von verschiedenen Signalen ist durch mehrmalige Anwendung der Funktion "Signal erzeugen" möglich.

²Microsoft Bitmap Format, verlustfreies und unkomprimiertes Bildformat. Programme zur Anzeige von BMP-Bildern sind standardmäßig unter Windows NT4.0/95/98 vorhanden.

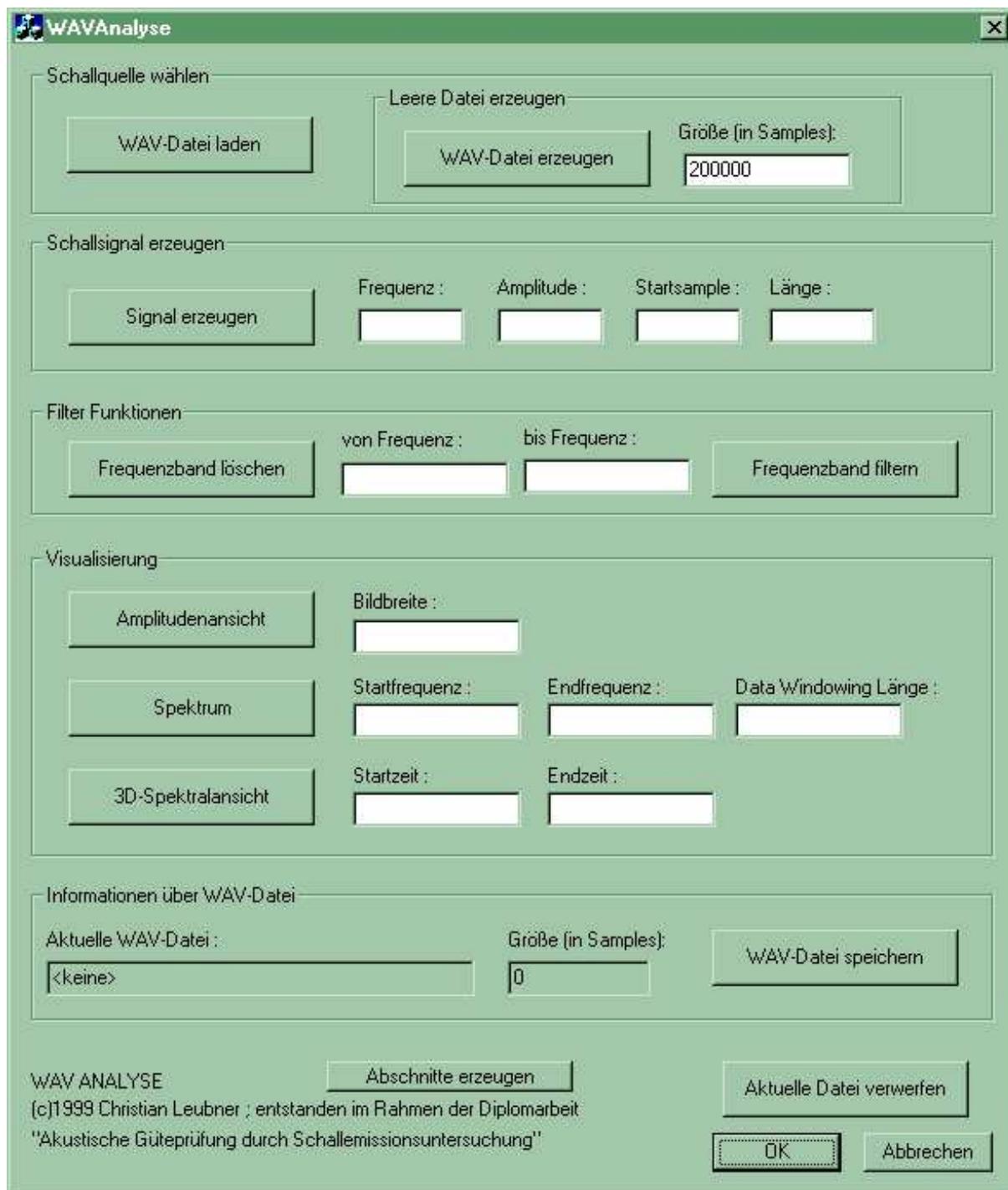


Abbildung 5.1: Oberfläche des Programms "WAV Analyse"

Die Funktionen "Frequenzband löschen" sowie "Frequenzband filtern" teilen sich beide die Textfelder "von Frequenz" und "bis Frequenz", über die das gewünschte Frequenzband eingestellt werden kann. Nach Anwendung der Funktion "Frequenzband löschen" sind alle Frequenzen des angegebenen Frequenzbandes aus dem Gesamtsignal entfernt. Im Gegensatz dazu sind nach Anwendung von "Frequenzband filtern" alle Frequenzen **außerhalb** des angegebenen Frequenzbandes aus dem Gesamtsignal entfernt. Neu erzeugte oder modifizierte WAV-Dateien können über den Button "WAV-Datei speichern" abgespeichert werden.

Die Visualisierungsfunktionen können ebenfalls über die zugeordneten Textfelder parametrisiert werden. Allerdings sind alle drei Funktionen auch ohne Angabe von Parametern durchführbar. In diesem Fall werden Standardwerte angenommen. Prinzipiell sind für die Visualisierungsfunktionen die Parameter von oben nach unten verfügbar. Das heißt, die Funktion "Amplitudenansicht", die ein Bild des Zeitsignals erzeugt, kann ausschließlich durch das Textfeld "Bildbreite" beeinflusst werden. Standardmäßig wird bei keiner Eingabe eine Bildbreite von 800 Punkten angenommen, die Bildhöhe wird im Verhältnis 4:3 zur Breite festgelegt. Die Funktion "Spektrum", die das Leistungsdichtespektrum als Bilddatei abspeichert, kann neben der Vorgabe der Bildbreite auch bezüglich eines Frequenzintervalls und der Data Windowing Länge (siehe Kapitel 4.2.1) parametrisiert werden. Wird eine Start- und eine Endfrequenz vorgegeben, wird lediglich dieses Intervall im Ausgabebild dargestellt. Werden keine Werte angegeben, wird das gesamte Frequenzspektrum - bei 44.1 kHz Samplingrate von 0 bis 22050 Hz - und eine Data Windowing Länge von 65536 verwendet.

Neben den bisher erläuterten Parametern kann die Funktion "3D-Spektralansicht" mit einer vorgegebenen Start- und Endzeit arbeiten. Bei der 3D-Spektralansicht wird über mehrere Zeitintervalle das Leistungsdichtespektrum bestimmt und in einem dreidimensionalen Koordinatensystem dargestellt. Auf der x-Achse ist die Frequenz und auf der y-Achse die Amplitude aufgetragen. Zusätzlich ist auf der z-Achse ein Zeitintervall angegeben, zu dem das Leistungsdichtespektrum bestimmt wurde. Über diese Art der Darstellung kann die Veränderung des Spektrums in Abhängigkeit von der Zeit veranschaulicht werden. Abbildung 5.2 zeigt beispielhaft diese Art der Darstellung.

3d Spektralansicht von "D:\Diplom\Dematic\eisen-raus02.wav"
x-Achse: Frequenzband von 0 Hz bis 22050 Hz
z-Achse: Zeitintervall von 0.000 s bis 1.088 s

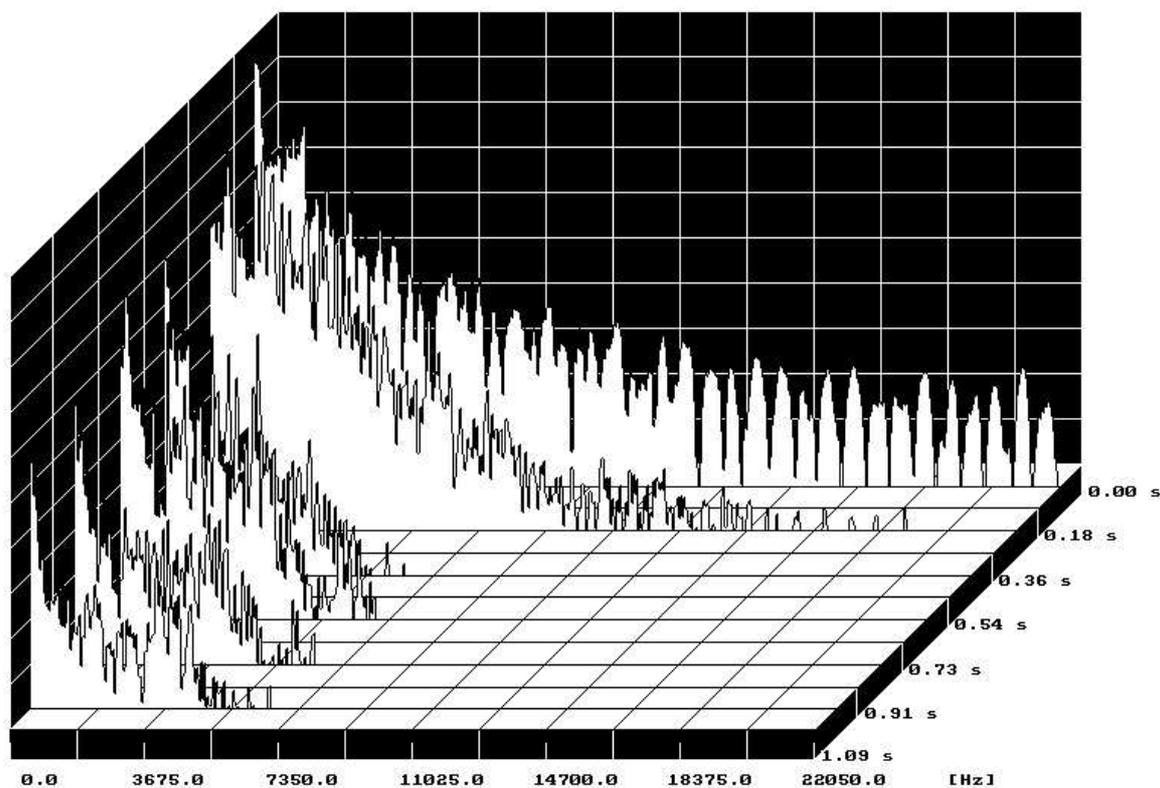


Abbildung 5.2: 3D-Spektralansicht

5.2 Klassifikations Tool

5.2.1 Funktionsumfang

Bei dem Programm "Klassifikations Tool" handelt es sich um die Implementierung eines Mustererkennungssystems wie es in dieser Arbeit besprochen wird. Das Programm unterstützt bzw. leistet alle dafür nötigen Schritte:

- **Signalverarbeitung**

Überführen einer digitalen Schallaufnahme in Merkmalsätze entsprechend den in Kapitel 4.2.1 und 4.2.2 vorgestellten Verfahren.

- **Merkmalsextraktion**

Durchführen der Merkmalsextraktion mittels des in Kapitel 4.5 vorgestellten suboptimalen Auswahlverfahrens auf der Grundlage des unscharfen Klassifikationskonzeptes nach Kapitel 4.3.

- **Klassifikation**

Durchführen der Klassifikation für beliebige digitale Schallaufnahmen, die im WAV-Format (44100 Hz, 16 Bit, mono) vorliegen.

5.2.2 Bedienung

Das Programm "Klassifikations Tool" wird über eine grafische Benutzeroberfläche bedient, die in Abbildung 5.3 gezeigt ist. Die Erstellung eines Klassifikators durch den Anwender erfolgt prinzipiell wie in Kapitel 3.1 erläutert.

Erstellen einer Lernbasis

Zunächst ist es erforderlich, die zu verwendenden Klassen dem Programm bekannt zu machen. Dies erfolgt durch Eingabe eines Namens in das Textfeld "Klassenname" und anschließendes Betätigen des Buttons "Klasse erstellen". In der Listbox "Bekannte Klassen" wird daraufhin der Name der neuen Klasse angezeigt, geführt von einer laufenden Indexnummer.

Im nächsten Schritt können jeder Klasse ihre Schallemissionen als Stichproben zugeordnet werden. Dazu wird aus der Listbox "Bekannte Klassen" mittels einfachem Mausclick

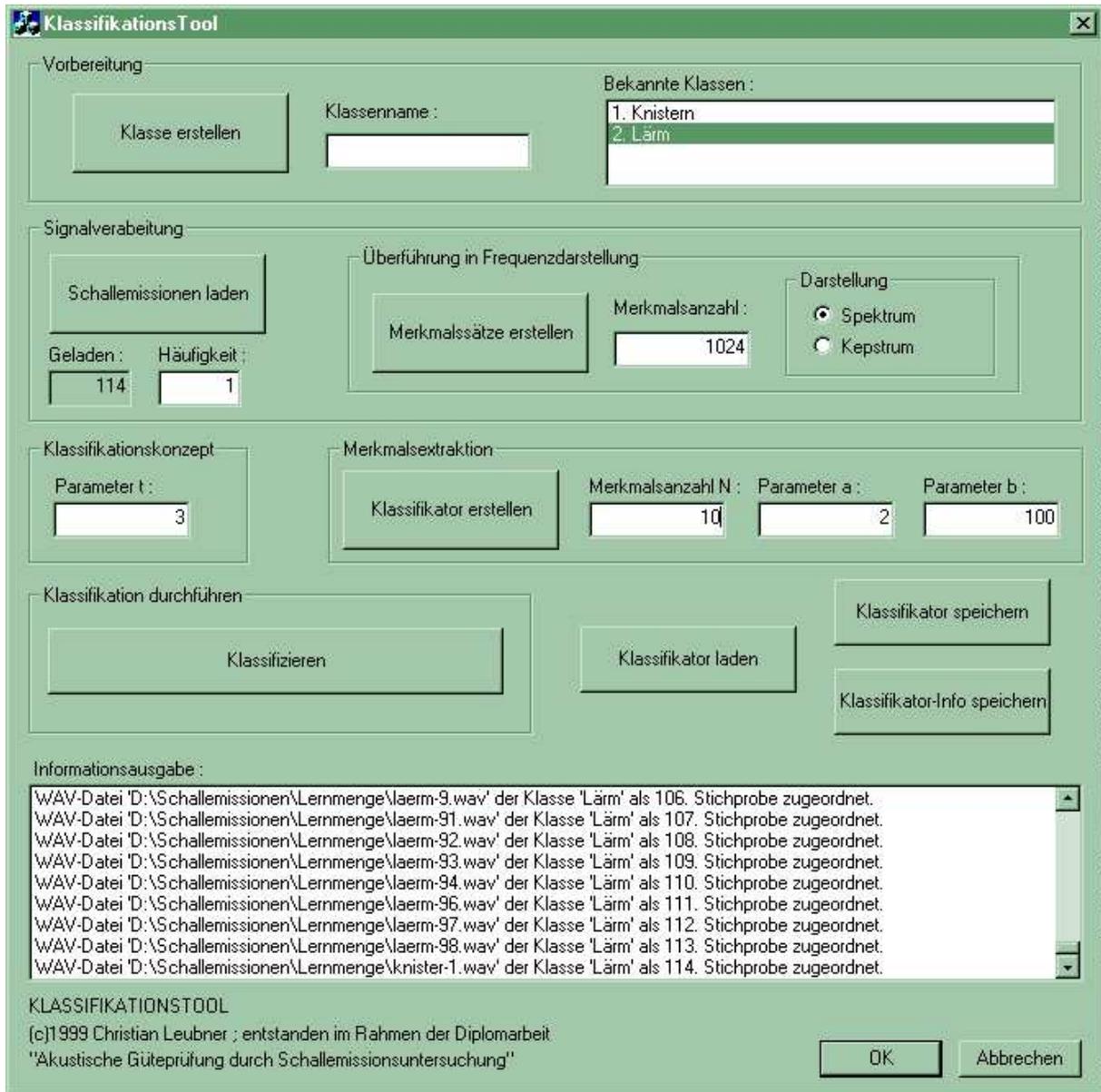


Abbildung 5.3: Oberfläche des Klassifikation Tools

eine Klasse ausgewählt, die farblich unterlegt dargestellt wird. Durch Betätigen des Buttons "Schallemissionen laden" können ein oder mehrere WAV-Dateien angegeben werden, die der zuvor ausgewählten Klasse zugeordnet werden. Das Textfeld "Geladen" zeigt die Anzahl aller bisher geladenen Stichproben an. Über das Textfeld "Häufigkeit" ist ein automatisches mehrmaliges, nach gewünschter Anzahl, Hinzufügen von Stichproben möglich. Wird hier beispielsweise eine "5" eingetragen, werden Stichproben, die über den Button "Schallemissionen laden" ausgewählt werden, jeweils fünfmal der Klasse hinzugefügt. Dadurch ist es möglich, die Stichprobenanzahl von verschiedenen Klassen anzugleichen. Dieses Vorgehen hat keinerlei Auswirkungen auf das Klassifikationskonzept, jedoch auf den Reklassifikationsfehler, anhand dessen die zu findende Merkmalsauswahl optimiert wird. Wird eine Stichprobe falsch reklassifiziert, wiegt dies bei mehrmaligem Vorhandensein der Stichprobe schwerer, als wenn diese nur einmal in der Lernmenge vorhanden ist. Standardmäßig ist das Textfeld "Häufigkeit" auf "1" eingestellt. In der unteren, mit "Informationsausgabe" bezeichneten Listbox werden die einzelnen Schritte, die das Programm ausführt, für den Benutzer transparent protokolliert.

Sind diese Vorbereitungen abgeschlossen, können im nächsten Schritt die Merkmalssätze erstellt werden. Über das Textfeld "Merkmalsanzahl" kann der Benutzer vorgeben, wieviele Merkmale die Merkmalssätze enthalten sollen. Diese Anzahl entspricht dem Wert N_M aus Definition 1 und ist standardmäßig auf 1024 eingestellt. Da für die Berechnung des Leistungsdichtespektrums als Fensterlänge eine Zweierpotenz erwartet wird, wird N_M gegebenenfalls auf die nächste größere Zweierpotenz hochgesetzt. Das Programm versucht, alle im vorherigen Schritt angegebenen und zugeordneten Schallemissionen in Merkmalssätze mit der gewünschten Anzahl N_M an Merkmalen zu überführen. Sollte dies in mindestens einem Fall nicht möglich sein, da beispielsweise die WAV-Datei zu kurz ist bzw. nicht genügend Zeitdaten zur Verfügung stehen, wird für N_M der kleinste Wert genutzt, der für alle Stichproben möglich ist. Dadurch wird sichergestellt, daß Definition 3 entsprochen wird, die unter anderem fordert, daß alle Merkmalssätze einer Lernbasis über die gleiche Anzahl von Merkmalen verfügen müssen. Auch während dieses Schrittes werden alle Handlungen des Programms in der Listbox "Informationsausgabe" für den Anwender angezeigt.

Nach Erstellung der Merkmalssätze ist die Lernbasis für die Merkmalsextraktion vollständig: Eine Menge C von Klassen sowie eine Menge L von Stichproben, in Form von Schallemissionen, wurden dem Programm bekannt gemacht. Da bereits beim Einladen von Schallemissionen eine Klasse angegeben wurde, der die Stichproben zuzuordnen sind, ist durch

den Anwender implizit auch bereits die Abbildung λ elementweise festgelegt worden. Damit ist die Erstellung einer Lernbasis abgeschlossen. Sollen ausgehend von einer Lernbasis mehrere Merkmalsextraktionen ausgeführt werden, ist es sinnvoll, bereits nach Erstellung der Lernbasis die Funktion "Klassifikator speichern" durchzuführen. Für eine weitere Merkmalsextraktion kann die erstellte Lernbasis mit der Funktion "Klassifikator laden" wieder eingeladen werden.

Merkmalsextraktion durchführen

Um die Merkmalsextraktion durchzuführen, ist zunächst das Vorhandensein einer Lernbasis erforderlich, die entweder zuvor, wie im vorhergehenden Abschnitt erläutert, erstellt oder von einem Speichermedium eingeladen wird. Die Parameter der Merkmalsextraktion sind im gleichnamigen Kasten zusammengefaßt. Die Textfelder erlauben die Vorgabe der Parameter N , a und b . Sie entsprechen den in Kapitel 4.5.2 genannten Parametern, wobei N wie gewohnt die Anzahl der Merkmale in der zu findenden Auswahl spezifiziert. Die Zugehörigkeitsfunktion des Klassifikationskonzeptes nach Gleichung 4.16 aus Kapitel 4.3.2 kann über den dort genannten Parameter t modifiziert werden. Der Button "Klassifikator erstellen" startet das Extraktionsverfahren. Aktuelle Zwischenergebnisse werden in der Listbox "Informationsausgabe" angezeigt. Dort wird ebenfalls angezeigt, wenn die Merkmalsextraktion beendet ist. Anschließend kann der erstellte Klassifikator über den Button "Klassifikator speichern" abgespeichert werden.

Bestehende Auswahl erweitern

Um das Merkmalsextraktionsverfahren flexibler zu gestalten, ist die Möglichkeit vorgesehen, eine bereits durchgeführte Merkmalsextraktion fortzusetzen. So können nachträglich Merkmalsauswahlen mit einer größeren Anzahl N von Merkmalen gefunden werden. Um die Konsistenz der Ergebnisse sicherzustellen, werden die Textfelder zu den Parametern a , b und t bei einer Erweiterung nicht berücksichtigt. Für diese Parameter werden die gleichen Werte verwendet, die zuvor auch genutzt wurden.

Um eine Erweiterung durchzuführen, muß zunächst ein Klassifikator vorhanden sein. Dieser kann entweder eingeladen oder durch eine vorhergehende Merkmalsextraktion erzeugt worden sein. Über das Textfeld "Merkmalsanzahl N " wird die neue gewünschte Anzahl von Merkmalen eingegeben, bis zu der die Merkmalsextraktion fortgesetzt werden soll. Durch

den Button "Klassifikator erstellen" wird das Extraktionsverfahren gestartet. Anschließend kann der erstellte Klassifikator wie gewohnt gespeichert werden.

Klassifikationen durchführen

Zum Klassifizieren von beliebigen Schallsignalen, die im WAV-Format vorliegen müssen, sollte das Programm zunächst neu gestartet werden. Über den Button "Klassifizieren" kann die Klassifikation gestartet werden. Der Anwender wird zunächst aufgefordert einen Klassifikator, der zuvor erstellt und abgespeichert wurde, auszuwählen. Danach können ein oder mehrere Schalldateien ausgewählt werden, die anschließend klassifiziert werden. Die Ergebnisse werden in der Listbox "Informationsausgabe" angezeigt.

Das Programm kann über die Buttons "OK" oder "Abbrechen" beendet werden.

5.3 Implementierung

Beiden Programmen "WAV Analyse" und "Klassifikations Tool" liegen zum Teil die gleichen Klassen zugrunde. Daher wird die Implementierung nicht für beide Programme getrennt erläutert. Da die Besprechung des kompletten Quellcodes zudem den Umfang dieser Arbeit übersteigen würde, soll nur ein grober Überblick mit ausgewählten Details verschafft werden.

Zur Realisierung beider Programme wurden folgende Klassen erstellt:

- Klasse `CMerkmalsatz`

- Klasse `CAuswahl`

- Klasse `CKlasse`

- Klasse `CKlassifikator`

- Klasse `CWAVFile`

- Klasse `CBMPFile`

5.3.1 Die Klassen *CMerkmalssatz* und *CAuswahl*

Die Klasse *CMerkmalssatz* bildet die Grundlage zur Darstellung eines Merkmalsatzes entsprechend Definition 1. Des weiteren sind Listenfunktionen integriert, die die Verkettung von Merkmalsätzen in einer Liste unterstützen.

Eine Instanz von *CAuswahl* repräsentiert eine Merkmalsauswahl gemäß Definition 2. Jede Instanz von *CAuswahl* kann die mit ihr erzielten Werte bezüglich der Gütekriterien Reklassifikationsfehler und gemittelte Trennschärfe speichern. Auch hier sind Listenfunktionen integriert.

5.3.2 Klasse *CKlasse*

Eine Instanz der Klasse *CKlasse* dient zur Repräsentation einer Klasse des Mustererkennungssystems. Zusätzlich leistet sie aber auch die elementweise Zuordnung der Stichproben, wie sie in der Definition 3 einer Lernbasis durch die Abbildung λ gefordert wurde, indem sie die zu ihr gehörigen Stichproben selbst verwaltet. Dazu stehen Methoden zur Verfügung, die die Verwaltung und das Hinzufügen von Merkmalsätzen (Instanzen von *CMerkmalssatz*) ermöglichen.

Des weiteren wird in *CKlasse* die Anwendung des Klassifikationskonzeptes nach Kapitel 4.3 durchgeführt. Dies umfaßt im einzelnen die Hauptachsentransformation und die Erstellung der Zugehörigkeitsfunktion. Die zur Berechnung des transformierten Koordinatensystems nötige Bestimmung der Eigenvektoren und -werte wird durch einen Standardalgorithmus nach *Jacobi* durchgeführt, der beispielsweise in [12] gefunden werden kann. Die Laufzeit des *Jacobi-Algorithmus* liegt in $O(N^3)$.

Ist die Hauptachsentransformation berechnet worden, stellt die Klasse *CKlasse* des weiteren eine Methode zur Verfügung, die für beliebige Merkmalsätze den Zugehörigkeitswert zu ihr gemäß der Zugehörigkeitsfunktion nach Gleichung 4.16 ermittelt.

5.3.3 Klasse *CKlassifikator*

Die Klasse *CKlassifikator* faßt alle für das Mustererkennungssystem relevanten Daten und Funktionalitäten zusammen. Dazu gehören Methoden zur Verwaltung der vorhandenen Klassen (Instanzen von *CKlasse*), wodurch entsprechend den Funktionalitäten von *CKlasse*

eine vollständige Lernbasis vorhanden ist. Des Weiteren ist eine Methode zur Merkmalsextraktion vorhanden, die das in Kapitel 4.5.2 vorgestellte suboptimale Auswahlverfahren implementiert. Die gefundenen Auswahlen werden verwaltet und können wahlweise zur Klassifikation eingesetzt werden. Standardmäßig ist stets die beste ermittelte Auswahl aktiviert. Neben der Möglichkeit den Klassifikator komplett abzuspeichern und später wieder einzuladen, kann außerdem auch eine Informationsdatei im Textformat erzeugt werden, die die wesentlichen Daten des Klassifikators ausgibt.

Außerdem existiert eine Methode, die für einen übergebenen Merkmalsatz (Instanz von `CMerkmalssatz`) die Klassifikation durchführt. Dazu wird von jeder Klasse (Instanzen von `CKlasse`) der Zugehörigkeitswert des übergebenen Merkmalsatzes zu ihr ermittelt. Als Ergebnis wird die Klasse mit dem größten Zugehörigkeitswert zurückgeliefert.

5.3.4 Die Klassen *CWAVFile* und *CBMPFile*

Die Klasse `CWAVFile` stellt alle wichtigen Funktionalitäten bezüglich des Umgangs mit Schallsignalen im WAV-Format zur Verfügung. Dazu zählen im einzelnen die Dateioperationen Einladen und Abspeichern, Filter- und Analysefunktionen sowie die Visualisierungsmethoden, die im Programm "WAV Analyse" zur Verfügung stehen. Entsprechend den Anforderungen der Aufgabenstellung verarbeitet die Klasse ausschließlich Dateien mit 44100 Hz, mono, mit 16 Bit Auflösung. Erweiterungsmöglichkeiten zu anderen WAV-Formaten sind prinzipiell vorgesehen und sollten mit verhältnismäßigem Aufwand durchführbar sein.

Die Klasse `CBMPFile` ermöglicht die Erstellung und das Abspeichern von Bilddateien im BMP-Format. Zur Verfügung stehen Zeichenfunktionen wie Punkte setzen, Linien ziehen und Rechtecke zeichnen sowie eine Funktion zur Textausgabe in das Bild hinein. Außerdem ist eine 3D-Funktionalität vorhanden, die das Arbeiten mit dreistelligen Koordinaten erlaubt. Die Klasse nimmt automatisch die Umrechnung in zweistellige Koordinaten vor. Hierfür stehen die gleichen Zeichen- und Textfunktionen zur Verfügung. In der vorliegenden Form können lediglich zweifarbige Bilder (schwarz und weiß) erzeugt werden.

Kapitel 6

Umsetzung des Verfahrens von Mannesmann Dematic

Die auf der Grundlage der vorgestellten Lösungsansätze entwickelte Software, die im vorigen Kapitel vorgestellt wurde, ist allgemein anwendbar. In diesem Kapitel wird das akustische Güteprüfverfahren von Mannesmann Dematic mit dem Programm "KlassifikationsTool" bearbeitet. Im Rahmen der einzelnen Definitionen und Erläuterungen wurden bereits des öfteren Parallelen zu dem akustischen Güteprüfverfahren von Mannesmann Dematic aufgezeigt. Diese sollen hier unter anderem im Zusammenhang dargestellt werden.

6.1 Aufgabenstellung

6.1.1 Ausgangssituation

Das eigentlich Güteprüfverfahren wurde bereits in Kapitel 1.3 beschrieben. Ziel der Prüfung ist die Qualitätsbeurteilung von punktgeschweißten Kranbahnschienen. Anhand der Schallemission, die entsteht, wenn der Prüfer mittels eines Eisens die Schiene nicht zerstörend auseinanderbiegt, kann der erfahrene Prüfer erkennen, ob die Schweißung fehlerfrei oder schadhaft ist. Der menschliche Prüfer charakterisiert die Schallemissionen, die auf schadhafte Schweißungen hindeuten, als "Knistern". Ist die Schweißung fehlerfrei, wird keine Schallemission festgestellt. Allerdings ist hierbei den besonderen Bedingungen des Einsatzortes, einer großen Fabrikhalle, Rechnung zu tragen. Der Prüfplatz ist umgeben von einer Vielzahl von Maschinen, die unterschiedlichste und vor allem laute Geräusche erzeugen. Für

eine Software ist es daher nicht ausreichend, lediglich zwischen Stille (keine Schallemission) und nicht Stille (Knistern) zu unterscheiden, wie trivialerweise angenommen werden könnte. Vielmehr ist es erforderlich, das charakteristische Knistern von beliebigem Umgebungslärm zu unterscheiden.

Das akustische Güteprüfverfahren wird bei Mannesmann Dematic strichprobenartig dreimal täglich durchgeführt. Die Kranbahnschienen werden mittels Widerstandspunktschweißen hergestellt. Zu schadhafte Schweißverbindungen kann es kommen, wenn beispielsweise die Schweißelektrode verschlissen ist. Die Schweißanlage selbst kann durch die Vorgabe verschiedener Parameter konfiguriert werden. Durch unterschiedliche Metalllieferanten kann eine wechselnde Parametrierung der Anlage erforderlich sein. Auch in diesem Zusammenhang treten schadhafte Schweißungen auf, die sich in der akustischen Güteprüfung sowie weiteren Prüfverfahren, die hier nicht von Interesse sind, bemerkbar machen.

Im Rahmen dieser Vorüberlegungen kann für das Mustererkennungssystem bereits die Feststellung getroffen werden, daß es zwei Klassen, $\Omega_{Knistern}$ und Ω_{Laerm} , gibt.

6.1.2 Versuchsaufbau

Bei einem Besuch vor Ort bei Mannesmann Dematic in Wetter/Ruhr werden mit Hilfe eines Laptops, einer handelsüblichen Soundkarte sowie eines handelsüblichen Mikrofons verschiedene Schallemissionen digital aufgezeichnet. Das Mikrophon dient dabei als Körperschallaufnehmer. Das Auseinanderbiegen der Kranbahnschienen sowie die Erkennung des charakteristischen Knistern wird durch den menschlichen Experten, einem Mitarbeiter der Firma Mannesmann Dematic durchgeführt. Da schadhafte Schweißverbindungen in der Praxis eher selten vorkommen, werden durch bewußt falsche Parametrierung der Schweißanlage einige schadhafte Kranbahnschienen hergestellt, die bei der akustischen Güteprüfung, dem Auseinanderbiegen der Schiene mittels einer Eisenstange, erwartungsgemäß das charakteristische Knistern absondern. Des weiteren werden von allen Maschinen in der direkten Umgebung des Prüfplatzes die Schallemissionen aufgezeichnet, die im Rahmen des Mustererkennungssystems zur Klasse "Lärm" zusammengefaßt werden.

Abbildung 6.1 zeigt die unterschiedlichen Einflußgrößen, die bei der Umsetzung des Güteprüfverfahrens von Mannesmann Dematic relevant sind. Über einen Rechner werden Schallemissionen aufgezeichnet, die in die Lernmenge einfließen sollen. Der menschliche Experte führt für diese Stichproben die Klassifikation durch, so daß in der Lernbasis beurteilte Stich-

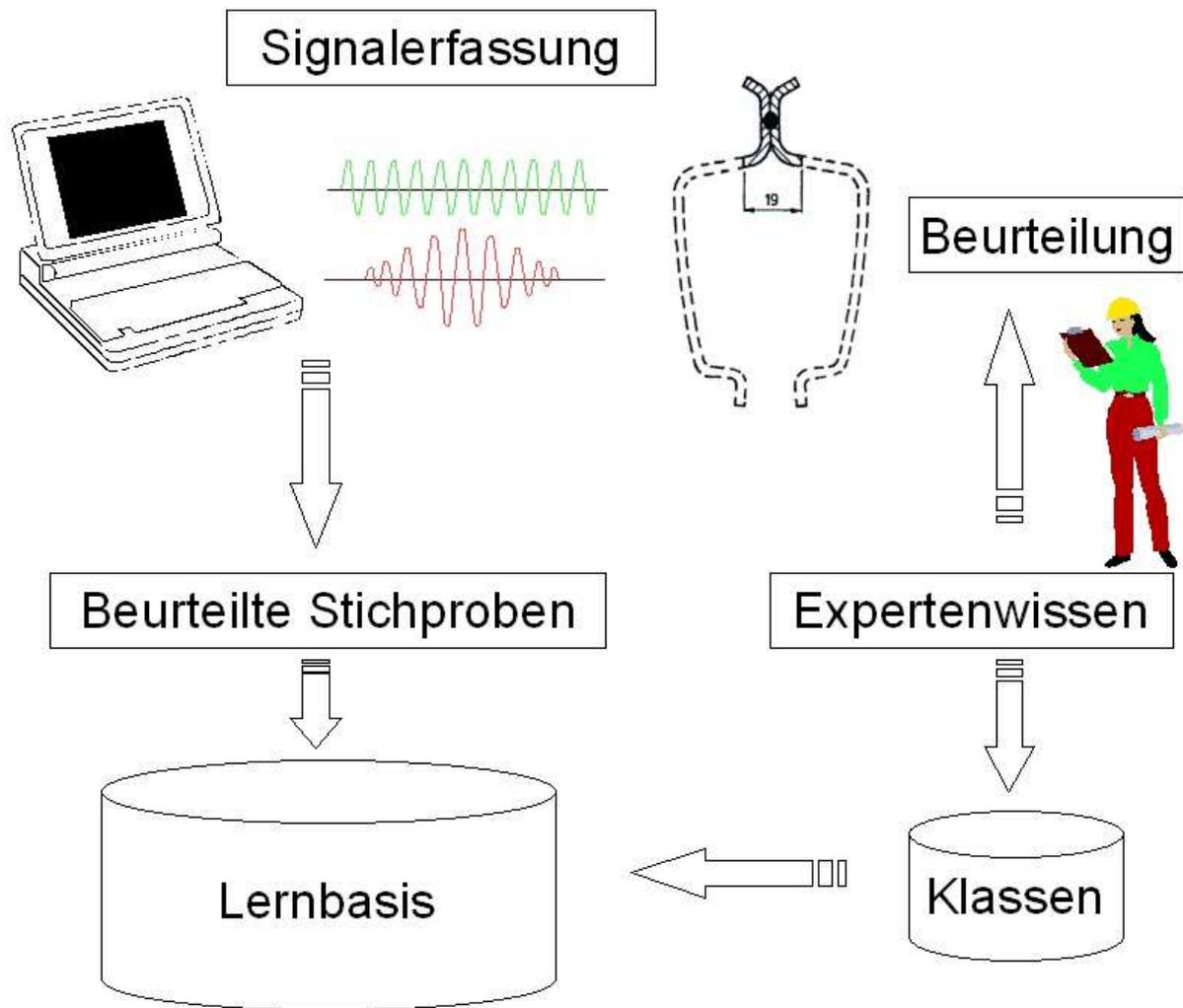


Abbildung 6.1: Erstellung der Lernbasis für Mannesmann Dematic

proben zur Verfügung stehen. Des weiteren werden durch das menschliche Expertenwissen die verschiedenen Klassen definiert. Diese Informationen bilden zusammen eine gültige Lernbasis im Sinne von Definition 3.

6.1.3 Signalverarbeitung

Unzweifelhaft vorteilhaft für das Mustererkennungssystem ist eine möglichst große Anzahl von Stichproben für jede Klasse. Aufgrund des vergleichsweise hohen Aufwands für die Herstellung nur einer einzigen schadhafte Stichprobe stehen nur sieben Stichproben für die Klasse $\Omega_{Knistern}$ zur Verfügung, was als eher wenig anzusehen ist. Da die digitale Aufzeichnung der Schallemissionen allerdings mit einer Samplerate von 44.1 kHz erfolgt, steht trotzdem eine große Menge an Datenmaterial zur Verfügung (jede Sekunde fallen 44100 Werte an). Unterteilt man die Stichproben in jeweils kleinere Abschnitte, entsteht eine größere Anzahl an Stichproben.

Auf der dieser Arbeit beiliegenden CD-ROM können im Verzeichnis "Schallemissionen" die so entstandenen Stichproben im WAV-Format gefunden werden. Die Stichproben sind auf zwei weitere Verzeichnisse, "Lernmenge" und "Prüfmenge", verteilt. Für die Lernbasis werden ausschließlich die Stichproben im Verzeichnis "Lernmenge" herangezogen. Die Stichproben im Verzeichnis "Prüfmenge" werden nicht in den Lernprozeß einbezogen, sondern dienen der unabhängigen Überprüfung von gefundenen Klassifikatoren.

In einem weiteren Schritt, der jedoch nur bei einer größeren Anzahl von Stichproben in Betracht käme, wäre es denkbar, die Überprüfung eines Klassifikators an einer zuvor nicht am Lernprozeß beteiligten Stichprobenmenge in die Merkmalsextraktion als weiteres Gütekriterium einzubeziehen.

Nach der Aufspaltung der "Knister" Stichproben in jeweils 16 kByte große Abschnitte, stehen insgesamt 20 Stichproben für die Klasse $\Omega_{Knistern}$ zur Verfügung. Für die Klasse "Lärm" sind 114 Stichproben vorhanden, die einen Querschnitt über die verschiedenen Umgebungsgeräusche enthalten. Die ersten 16 Stichproben zu "Knistern" finden sich auf der CD-ROM im Verzeichnis "Lernmenge", ebenso wie 97 Stichproben der Klasse Ω_{Laerm} . Die restlichen Stichproben, viermal "Knistern" und 24 mal "Lärm" sind im Verzeichnis "Prüfmenge". Auf der CD-ROM befindet sich des weiteren ein Verzeichnis "Dematic", das die Originalaufnahmen, die bei Mannesmann Dematic in Wetter/Ruhr entstanden sind, thematisch benannt enthält. Die vorhandene Gesamtanzahl von Stichproben wurde im Verhältnis 4:1 auf Lern-

und Prüfmenge verteilt. In beiden Gruppen beginnen die Dateinamen der Stichproben zur Klasse $\Omega_{Knistern}$ mit "Knister-" sowie einer fortlaufenden Nummer. Entsprechend beginnen die Stichproben zur Klasse Ω_{Laerm} mit "Laerm-" und enden mit einer fortlaufenden Nummer. Zur Erstellung einer Lernbasis stehen soweit alle Informationen zur Verfügung. Die Eingabe der Daten in das Programm "KlassifikationsTool" erfolgt wie in Kapitel 5.2.2 beschrieben. Dort wird auch die mögliche Parametrierung der Software erläutert. Über den skizzierten Weg ist es möglich, eine Menge

$$L = \{\vec{M}_1, \vec{M}_2, \dots, \vec{M}_{N_s}\}$$

von Merkmalsätzen zu gewinnen. Die Menge der Klassen

$$C = \{\Omega_{Knistern}, \Omega_{Laerm}\}$$

ist ebenfalls bekannt. Durch die Befragung des menschlichen Experten kann die Zuordnung einer Stichprobe zu einer Klasse erfolgen, so daß die Abbildung

$$\lambda : L \rightarrow C$$

elementweise festgelegt werden kann. Damit ist eine Lernbasis $B = (L, C, \lambda)$ erstellt.

6.1.4 Merkmalsextraktion

Die Merkmalsextraktion erfolgt in dem Programm "KlassifikationsTool" entsprechend der vorgestellten Konzepte und des suboptimalen Auswahlverfahrens. Außerdem wird das unscharfe Klassifikationskonzept nach Kapitel 4.3 genutzt. Als Gütekriterium wird der Reklassifikationsfehler nach Gleichung 4.25 eingesetzt. Die Merkmalsauswahl erfolgt entsprechend des suboptimalen Verfahrens, das in Kapitel 4.5.2 vorgestellt wurde. An den jeweiligen Stellen wurde bereits erläutert, warum die vorgeschlagenen Mittel als besonders geeignet für die Umsetzung des akustischen Güteprüfverfahrens von Mannesmann Dematic erachtet werden. Die Bedienung des Programms "KlassifikationsTool" erfolgt wie in Kapitel 5 erläutert. Durch die Protokollausgaben der Merkmalsextraktion wird offenkundig, daß bereits bei Auswahlen mit $N = 3$ Merkmalen ein Reklassifikationsfehler von 0.0 erzielt wird. Da die Idee des Auswahlverfahrens vorsieht, nur die b besten Merkmalsauswahlen im nächsten Schritt $N = N + 1$ weiterzubetrachten und um ein Merkmal zu erweitern, bleibt es eine Frage der Sortierstrategie, was geschieht, wenn deutlich mehr als b Auswahlen die gleiche und noch

dazu bestmögliche Bewertung haben. Daher wird die bereits in Kapitel 4.5.1 geäußerte Idee aufgegriffen, den Reklassifikationsfehler mit der gemittelten Trennschärfe als Gütekriterium zu kombinieren. Als Sortierstrategie wird nach wie vor primär die Optimierung des Reklassifikationsfehlers verfolgt. Bei Gleichheit des Reklassifikationsfehlers für zwei Auswahlen, wird die Auswahl mit der größeren gemittelten Trennschärfe als besser erachtet, so daß sekundär in Richtung einer möglichst hohen gemittelten Trennschärfe optimiert wird.

6.2 Ergebnisse

Alle Schritte von der Signalverarbeitung bis zur Merkmalsextraktion werden mit der in Kapitel 5 vorgestellten Software "KlassifikationsTool" durchgeführt. Die dort vorgenommene Normierung der Merkmalsätze erfolgt auf der Grundlage von Gleichung 4.6 aus Kapitel 4.2.2.

Das Merkmalsextraktionsverfahren wird mit verschiedenen Parametrierungen durchgeführt. Die daraus resultierenden Klassifikatoren können auf der dieser Arbeit beigelegten CD-ROM im Verzeichnis "Klassifikatoren" gefunden werden. Eine vollständige Auflistung der Ergebnisse liefert die Datei "Ergebnisse.xls"¹. Variationen in den Parametern wurden hauptsächlich bezüglich der Merkmalsanzahl N sowie der Parameter b und t durchgeführt. Zur besseren Übersichtlichkeit wurde für jede untersuchte Parameterkombination ein weiteres Unterverzeichnis auf der CD-ROM angelegt. Das Verzeichnis "a2-b100-t3" enthält beispielsweise die ermittelten Klassifikatoren für $a = 2$, $b = 100$ und $t = 3$. Für einen Parametersatz wurden weiterhin Variationen in der Merkmalsanzahl N durchgeführt, so daß in jedem Unterverzeichnis Klassifikatoren mit 5-10 Merkmalen zu finden sind. Eine Klassifikatordatei endet mit der Kennung ".klf" und führt im Dateinamen neben der Merkmalsanzahl N auch die Parameter a und b , so daß beispielsweise der Dateiname "kl-7-2-100.klf" einen Klassifikator zu den Parametern $N = 7$, $a = 2$ und $b = 100$ beschreibt. Der Parameter t ergibt sich aus dem Verzeichnis, indem sich die Datei befindet. Diese Art der beschreibenden Namensgebung ist lediglich aus Gründen der Übersichtlichkeit erfolgt, prinzipiell kann eine Klassifikatordatei beliebig benannt werden. Des weiteren existiert zu jeder Klassifikatordatei eine zugehörige Informationsdatei, die mit ".info" endet². Diese Datei enthält alle relevanten Informationen

¹Microsoft Excel 97 Dateiformat. Auf den WWW-Seiten von Microsoft kann ein kostenloses Programm zur Betrachtung von Excel-Dateien gefunden werden.

²ASCII Textformat. Darstellbar mit jedem Texteditor (unter Windows 95/98/NT4.0 z.B. mit "NotePad").

eines Klassifikators. Dazu zählen unter anderem die Anzahl der Klassen, die zugehörigen Schallemissionen, die Parameter sowie die ermittelten Merkmalsauswahlen.

Die Leistungsfähigkeit der Klassifikatoren wird an den Schallemissionen im Verzeichnis "Prüfmenge", wie in Abschnitt 6.1.3 beschrieben, durchgeführt. Die nachfolgende Tabelle 6.2 listet die Klassifikationsergebnisse der einzelnen Parameterkombinationen auf. Alle zusammengetragenen Ergebnisse, auch die ermittelten Merkmalsauswahlen, sind in der Datei "Ergebnisse.xls"³ zu finden. Zu jedem Klassifikator wird die Anzahl der richtig reklassifizierten Stichproben der Prüfmenge sowohl für die Klasse "Knistern" (Spalte: "Prüf $\Omega_{Knistern}$ richtig") als auch für die Klasse "Lärm" (Spalte: "Prüf $\Omega_{Lärm}$ richtig") aufgeführt. In der Spalte "Prüf G_R " wird der zugehörige Reklassifikationsfehler nach Gleichung 4.25 bestimmt.

³Auf der beigelegten CD-ROM im Hauptverzeichnis.

N	a	b	t	G_R	G_{V_m}	Prüf $\Omega_{Knistern}$ richtig	Prüf Ω_{Laerm} richtig	Prüf G_R
5	2	100	2	0,000000000	0,8393903	3	23	0,071
6	2	100	2	0,000000000	0,8537624	2	24	0,071
7	2	100	2	0,000000000	0,8662707	1	24	0,107
8	2	100	2	0,000000000	0,871893	2	24	0,071
9	2	100	2	0,000000000	0,8698809	2	24	0,071
10	2	100	2	0,008849560	0,867635	2	24	0,071
5	2	100	3	0,000000000	0,8523686	4	24	0,000
6	2	100	3	0,000000000	0,8815737	3	24	0,036
7	2	100	3	0,000000000	0,9051009	4	24	0,000
8	2	100	3	0,000000000	0,8941729	3	23	0,071
9	2	100	3	0,000000000	0,8916967	4	24	0,000
10	2	100	3	0,000000000	0,8798724	3	24	0,036
5	2	100	4	0,000000000	0,8753569	3	24	0,036
6	2	100	4	0,000000000	0,8949143	2	24	0,071
7	2	100	4	0,000000000	0,9126116	4	24	0,000
8	2	100	4	0,000000000	0,9256204	2	24	0,071
9	2	100	4	0,000000000	0,9394327	2	24	0,071
10	2	100	4	0,000000000	0,9336694	3	24	0,036
5	2	200	2	0,000000000	0,8393904	3	23	0,071
6	2	200	2	0,000000000	0,8537623	2	24	0,071
7	2	200	2	0,000000000	0,8662707	1	24	0,107
8	2	200	2	0,000000000	0,871893	2	24	0,071
9	2	200	2	0,000000000	0,869881	2	24	0,071
10	2	200	2	0,000000000	0,83866	3	24	0,036
5	2	200	3	0,000000000	0,8523688	4	24	0,000
6	2	200	3	0,000000000	0,8815738	3	24	0,036
7	2	200	3	0,000000000	0,9051007	4	24	0,000
8	2	200	3	0,000000000	0,9061368	1	24	0,107
9	2	200	3	0,000000000	0,914783	1	24	0,107
10	2	200	3	0,000000000	0,9185592	2	24	0,071
5	2	200	4	0,000000000	0,8753569	3	24	0,036
6	2	200	4	0,000000000	0,8949143	2	24	0,071
7	2	200	4	0,000000000	0,9130009	4	24	0,000

N	a	b	t	G_R	G_{V_m}	Prüf $\Omega_{Knistern}$ richtig	Prüf Ω_{Laerm} richtig	Prüf G_R
8	2	200	4	0,000000000	0,9256204	2	24	0,071
9	2	200	4	0,000000000	0,9394327	2	24	0,071
10	2	200	4	0,000000000	0,9336694	3	24	0,036
5	2	400	2	0,000000000	0,8393904	3	23	0,071
6	2	400	2	0,000000000	0,8543017	3	24	0,036
7	2	400	2	0,000000000	0,8662707	1	24	0,107
8	2	400	2	0,000000000	0,871893	2	24	0,071
9	2	400	2	0,000000000	0,869881	2	24	0,071
10	2	400	2	0,000000000	0,8475506	1	24	0,107
5	2	400	3	0,000000000	0,8573748	4	24	0,000
6	2	400	3	0,000000000	0,8997133	4	24	0,000
7	2	400	3	0,000000000	0,9150225	4	24	0,000
8	2	400	3	0,000000000	0,9215603	3	24	0,036
9	2	400	3	0,000000000	0,9243987	4	24	0,000
10	2	400	3	0,000000000	0,9180092	3	24	0,036
5	2	400	4	0,000000000	0,8753569	3	24	0,036
6	2	400	4	0,000000000	0,9029375	4	24	0,000
7	2	400	4	0,000000000	0,9239208	3	24	0,036
8	2	400	4	0,000000000	0,9338192	2	24	0,071
9	2	400	4	0,000000000	0,9418555	3	24	0,036
10	2	400	4	0,000000000	0,9452547	2	24	0,071
5	2	50	2	0,000000000	0,824343	3	24	0,036
6	2	50	2	0,000000000	0,8537624	2	24	0,071
7	2	50	2	0,000000000	0,8662708	1	24	0,107
8	2	50	2	0,000000000	0,871893	2	24	0,071
9	2	50	2	0,000000000	0,8698809	2	24	0,071
10	2	50	2	0,008849560	0,867635	2	24	0,071
5	2	50	3	0,000000000	0,8523688	4	24	0,000
6	2	50	3	0,000000000	0,8815738	3	24	0,036
7	2	50	3	0,000000000	0,9051007	4	24	0,000
8	2	50	3	0,000000000	0,8912324	3	24	0,036
9	2	50	3	0,000000000	0,8837098	4	24	0,000
10	2	50	3	0,000000000	0,8798724	3	24	0,036

N	a	b	t	G_R	G_{V_m}	Prüf $\Omega_{Knistern}$ richtig	Prüf Ω_{Laerm} richtig	Prüf G_R
5	2	50	4	0,000000000	0,8753569	3	24	0,036
6	2	50	4	0,000000000	0,8890461	3	24	0,036
7	2	50	4	0,000000000	0,9126115	4	24	0,000
8	2	50	4	0,000000000	0,9252439	3	24	0,036
9	2	50	4	0,000000000	0,9248866	2	24	0,071
10	2	50	4	0,000000000	0,9265305	3	24	0,036
3	3	100	3	0,000000000	0,6607661	4	22	0,071
5	3	100	3	0,000000000	0,8817358	3	24	0,036
6	3	100	3	0,000000000	0,9070604	3	24	0,036
7	3	100	3	0,000000000	0,914321	3	24	0,036
8	3	100	3	0,000000000	0,9207265	3	24	0,036
9	3	100	3	0,000000000	0,9178582	4	24	0,000
10	3	100	3	0,000000000	0,9107648	4	24	0,000
5	1	100	3	0,000000000	0,8523688	4	24	0,000
6	1	100	3	0,000000000	0,8815738	3	24	0,036
7	1	100	3	0,000000000	0,9051007	4	24	0,000
8	1	100	3	0,000000000	0,8941731	3	24	0,036
9	1	100	3	0,000000000	0,8916967	4	24	0,000
10	1	100	3	0,000000000	0,8798724	3	24	0,036

Tabelle 6.2: Ergebnisse der Klassifikatoren

Bevor im folgenden ein Überblick über den Einfluß der einzelnen Parameter auf die Klassifikationsergebnisse verschafft werden soll, sei zunächst angemerkt, daß mit den meisten Parameterkombinationen gute bis sehr gute Ergebnisse bezogen auf die Prüfmenge erzielt werden. Von den 85 untersuchten Parameterkombinationen erreichen 21 den bestmöglichen Reklassifikationsfehler von 0.0, angewendet auf die Prüfmenge. Der schlechteste Reklassifikationsfehler der Prüfmenge beträgt lediglich 0.107 und tritt bei sieben Kombinationen auf. Aufgrund des engen Beieinanderliegens der Ergebnisse ist eine Interpretation der Zusammenhänge nicht einfach, da auch ein gewisser Zufallsfaktor vorhanden sein dürfte.

6.2.1 Merkmalsanzahl N

Bezüglich der Merkmalsanzahl N kann festgestellt werden, daß unabhängig von den restlichen Parametern mit $N = 7$ bezogen auf die Prüfmenge überdurchschnittlich gute Klassifikationsergebnisse erzielt werden. Der Reklassifikationsfehler G_R der Prüfmenge liegt mit $N = 7$ insgesamt achtmal bei 0.0. Die Qualität der Klassifikationsergebnisse scheint nicht proportional zur Anzahl N der Merkmale zu verlaufen, was mit dem Problem der Überadaptation (siehe Abschnitt 4.5.4) erklärbar sein kann. In zwei Fällen verschlechtern sich sogar die Reklassifikationsergebnisse der Lernmenge (bei $a = 2, b = 100, t = 2$ und $a = 2, b = 50, t = 2$), so daß für $N = 10$ der Reklassifikationsfehler der Lernmenge erstmals innerhalb der Versuchsreihe nicht mehr 0.0 beträgt.

6.2.2 Parameter a

Der Parameter a legt fest, bis zu welcher Merkmalsanzahl eine vollständige Überprüfung aller möglichen $\binom{NM}{a}$ Teilmengen im Rahmen der Initialisierungsphase erfolgen soll (siehe Abschnitt 4.5.2). Damit trägt a entscheidend zur Laufzeit der Merkmalsextraktion bei.

Variationen über den Parameter a sind nur in begrenztem Ausmaß durchgeführt worden. Standardmäßig sind die meisten Klassifikatoren mit $a = 2$ erstellt worden, was im Hinblick auf die erforderliche Rechenleistung einen vernünftigen Kompromiß darstellt. Die Nutzung von $a = 1$ wurde in einem Fall durchgeführt ($b = 100, t = 3$) und führte zur selben Merkmalsauswahl wie die Verwendung von $a = 2$. Grundsätzlich sollte die Qualität des Klassifikators proportional zu a besser werden, da eine größere Teilmenge von Merkmalsauswahlen auf ihre Tauglichkeit und Qualität hin überprüft wird. Dies bestätigt sich bei der durchgeführten Versuchsreihe mit $a = 3$ ($b = 100, t = 3$), die zu einer anderen Merkmalsauswahl führt. Die Klassifikationsergebnisse zu diesen Parametern sind sehr gut und liefern für die Prüfmenge einen Reklassifikationsfehler von 0.0 (für $N = 9$ und $N = 10$ Merkmale) und sind damit etwas besser als die Ergebnisse für $a = 2$ (andere Parameter identisch). Die Überprüfung aller $\binom{1024}{3} = 178.433.024$ Teilmengen benötigte allerdings eine Rechenzeit von über 50 Stunden CPU Zeit⁴. Daher seien die Ergebnisse für $a = 3, b = 100$ und $t = 3$ als exemplarisch für die Erhöhung von a betrachtet. Da in der Prüfmenge lediglich 28 Stichproben zur Verfügung stehen, kann eine gewisse Zufälligkeit nicht ausgeschlossen werden, da mit $a = 3$ lediglich

⁴Auf einem System mit AMD K6/2 Prozessor (350 MHz), 128 MB Hauptspeicher unter Windows NT4.0.

eine einzige Stichprobe mehr richtig reklassifiziert wird als mit $a = 2$. Eine Qualitätssteigerung für größere a ist insofern schwierig nachweisbar, als daß bereits mit $a = 2$ sehr gute Ergebnisse erzielt werden.

6.2.3 Parameter b

Über den Parameter b wird die Größe der Bestenliste festgelegt. In jedem Schritt der Erweiterungsschleife des suboptimalen Auswahlverfahrens werden jeweils nur die besten b Merkmalsauswahlen aus dem vorhergehenden Schritt als Grundlage zur Erweiterung um ein weiteres Merkmal genutzt (siehe Abschnitt 4.5.2).

Versuchsreihen wurden für $b = 50$, $b = 100$, $b = 200$ und $b = 400$ durchgeführt. Grundsätzlich sollte aus einem größeren b eine bessere Klassifikationsleistung resultieren, da eine größere Menge von Merkmalsauswahlen untersucht wird. Aus der Auflistung der Ergebnisse kann dies nicht abgelesen werden, da mit allen Werten für b gute bis sehr gute Reklassifikationsergebnisse der Prüfmenge erzielt werden. Der größte Reklassifikationsfehler für $b = 50$ liegt genauso bei 0.107 wie für $b = 400$, auf der anderen Seite existieren sowohl für $b = 50$ als auch für $b = 400$ mehrere Klassifikatoren, die einen Reklassifikationsfehler der Prüfmenge von 0.0 aufweisen.

6.2.4 Parameter t

Über den Parameter t wird die Zugehörigkeitsfunktion des unscharfen Klassifikationskonzeptes nach Gleichung 4.16 beeinflusst (siehe Abschnitt 4.3.2). Anschaulich wächst proportional zur Größe von t der Bereich von hohen Zugehörigkeitswerten, so daß auch Objekte, die eine größere Entfernung zum Cluster aufweisen, mit tendenziell höheren Zugehörigkeitswerten beurteilt werden.

In Kapitel 4.3.2 wurde bereits darauf hingewiesen, daß Strackeljan in [17] ausschließlich mit $t = 3$ arbeitet. Daß dies durchaus eine geeignete Wahl für t darstellt, bestätigen auch die Ergebnisse dieser Arbeit: 17 von 21 Klassifikatoren, die einen Reklassifikationsfehler der Prüfmenge von 0.0 aufweisen, wurden mit dem Parameter $t = 3$ erstellt. Für $t = 4$ wird dieses bestmögliche Ergebnis lediglich viermal erzielt und dies in Fällen, wo für ansonsten identische Parametrierung auch mit $t = 3$ dasselbe Ergebnis erzielt wird. Die schlechtesten Ergebnisse werden mit $t = 2$ erzielt.

6.3 Ergebnis

Wie bereits vorweggenommen wurde, sind die mit dem Programm "KlassifikationsTool" erzielten Ergebnisse sehr zufriedenstellend. Wendet man die gefundenen Klassifikatoren auf die dem Mustererkennungssystem zuvor unbekanntes Stichproben der Prüfmenge an, erzielt jeder vierte Klassifikator den bestmöglichen Reklassifikationsfehler von 0.0. Im schlechtesten Fall werden immerhin noch 25 von 28 Stichproben richtig klassifiziert.

Berücksichtigt man die Beobachtungen bezüglich der verschiedenen Parameter und wählt $t = 3$ und $N = 7$ wird für $a = 2$ zu beliebigem b ein Reklassifikationsfehler der Prüfmenge von 0.0 erzielt. Wird $a = 3$ vorgegeben, wird immerhin für $N = 9$ und $N = 10$ dieses bestmögliche Ergebnis erzielt. Inwieweit diese Beobachtungen für die Anwendung der Software auf andere akustische Prüfverfahren übertragbar sind, bedarf dahingehender Untersuchungen, die nicht mehr Gegenstand dieser Arbeit sind. Insgesamt können die Ergebnisse voll überzeugen.

Kapitel 7

Ausblick

Die akustische Güteprüfung ist ein variables Verfahren, das auch zukünftig neue Einsatzfelder finden wird. Da Schallsignale immer dort entstehen, wo etwas schwingt, sind die Einsatzgebiete hauptsächlich in der produzierenden Industrie zu finden. Aufgrund der fortschreitenden Automatisierung der Produktionsprozesse liegt es nahe, auch die Prüf- und Kontrollvorgänge soweit möglich zu automatisieren. Durch ein objektives Prüfsystem ist eine einheitliche Qualitätsprüfung möglich, die nicht von der Subjektivität eines spezifischen Prüfers abhängt.

Die Ausstellung dieser Arbeit im Rahmen der 6. Dortmunder Fuzzy Tage hat zudem gezeigt, daß in der Industrie nach wie vor Interesse an computerisierten akustischen Prüfsystemen besteht. Durch die im Laufe der Zeit gestiegene Flexibilität der Mustererkennungssysteme bezüglich der spezifischen Aufgabenstellung werden neue Einsatzfelder für akustische Prüfverfahren entdeckt. Bei Verschleißteilen in der Produktion kann der Eintritt des Bauteils in einen kritischen Zustand, der dem Ausfall des Teils vorausgeht, durch genaues Hinhören von einem Menschen erkannt werden. Durch automatisierte Lernverfahren der Mustererkennungssysteme, wie beispielsweise in dieser Arbeit vorgestellt, sollten neue Aufgaben wie diese mit verhältnismäßigem Aufwand gelöst werden können.

Im Bereich der Mustererkennungssysteme sind weitere Forschungsaufgaben im Bereich der Signalverarbeitung und des Klassifikationskonzeptes zu erkennen. Bei der Signalverarbeitung werden zur Zeit häufig nur die Frequenzdaten zur Erzeugung von Merkmalsätzen verwendet. Die zeitliche Darstellung der Schallsignale wird nur selten verwendet. Hier fehlen Konzepte zur aufschlußreichen Handhabung der Daten. Im Bereich des Klassifikationskonzeptes stellt sich das Problem, wie im Fall von nichtdisjunkten Klassen vorgegangen wird. Auch die

Entwicklung alternativer Gütekriterien für die automatisierte Merkmalsextraktion kann zu weiteren Verbesserungen von Klassifikationssystemen führen.

Insgesamt kann die akustische Güteprüfung als interessantes Gebiet betrachtet werden, dessen Weiterentwicklung nicht zuletzt in der Industrie auf reges Interesse stößt.

Literaturverzeichnis

- [1] D. Barschdorff, A. Bothe. *Neuronale Netze als Signal- und Musterklassifikatoren*. Technisches Messen tm. Nr. 57. S. 437-444. 1990.
- [2] Hans Borucki. *Einführung in die Akustik*. BI-Wiss.Verl., 3., erw. Auflage. 1989.
- [3] Hans Goebel, Martin Schrader (Hrsg.). *Datenanalyse, Klassifikation und Informationsverarbeitung*. Physica-Verlag Heidelberg. 1992.
- [4] Joachim Hartung. *Statistik: Lehr- und Handbuch der angewandten Statistik*. R. Oldenbourg Verlag GmbH, München. 1984.
- [5] Frank Höppner. *Fuzzy-Clusteranalyse: Verfahren für die Bildererkennung, Klassifizierung und Datenanalyse*. Vieweg. Wiesbaden. 1997.
- [6] Benno Kotterba. *Eine Methode zur Merkmalsfindung für die Klassifikation von Geräuschen*. Dissertation am Fachbereich 21 (Umwelttechnik) der Technischen Universität Berlin. 1983.
- [7] G.-H. Lietke, L. Schueler, Ch. Weber, H. Wolff. *Klassifikation von Daten und Signalen*. Forschungsbericht aus der Wehrtechnik, Dokumentationszentrum der Bundeswehr im Auftrag des Bundesministeriums der Verteidigung. 1980.
- [8] Mayer, Mechler, Schindwein, Wolke. *Fuzzy Logic. Einführung und Leitfaden zur praktischen Anwendung*. Addison-Wesley. 1993.
- [9] H.E.Meier. *Forschungsbericht T83-005: Objektive akustische Güteprüfung durch Mustererkennung und Signalanalyse*. Bundesministerium für Forschung und Technologie, Fraunhofer-Institut für Informations- und Datenverarbeitung IITB Karlsruhe. Februar 1983.
- [10] Lothar Mensching. *Analyse und Klassifikation fouriertransformierter Wirbelstromsignale bei der Innenrohrprüfung*. Dissertation an der Fakultät für Maschinenwesen der Universität Hannover. 1992.

- [11] H. Niemann. *Klassifikation von Mustern*. Springer-Verlag, Berlin. 1983
- [12] Numerical Recipes Software. *Numerical Recipes in C: The Art of Scientific Computing, Kapitel 12*. Cambridge University Press. 1988-1992.
- [13] Günther Ruske. *Automatische Spracherkennung: Methoden der Klassifikation und Merkmalsextraktion*. R. Oldenbourg Verlag GmbH, München. 1994.
- [14] Thomas Schwarz. *Schallanalyse zur Diagnose von Schäden an Hydraulikpumpen*. Dissertation an der Fakultät für Maschinenwesen der RWTH Aachen. Juni 1990.
- [15] H. Sieber. *Mathematische Formeln, Erweiterte Ausgabe E*. Ernst Klett Verlag, Stuttgart. 1980.
- [16] Wolfgang Stark. *Gütesicherung von Laserpunktschweißungen mittels Schallemissionsanalyse am Beispiel ausgewählter Kupferlegierungen*. Dissertation an TU München. Institut für Werkstoff- und Verarbeitungswissenschaften. Lehrstuhl für Fügetechnik. 1989.
- [17] Jens Strackeljan. *Klassifikation von Schwingungssignalen mit Methoden der unscharfen Mustererkennung*. Dissertation an der Fakultät für Bergbau, Hüttenwesen und Maschinenwesen der Technischen Universität Clausthal. 1993.
- [18] Gerhard Thews, Ernst Mutschler, Peter Vaupel. *Anatomie, Physiologie, Pathophysiologie des Menschen; S. 504-510*. Wissenschaftliche Verlagsgesellschaft mbH, Stuttgart. 1991.
- [19] Helmut Thiele. *Einführung in die Fuzzy-Logik*. Skriptum. Universität Dortmund. 1995.