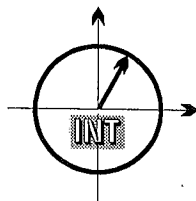


■ *Forschungsberichte aus dem
Institut für Nachrichtentechnik
der Universität Karlsruhe (T.H.)*



Christoph Delfs

■ **Zeit-Frequenz-Signalanalyse:
Lineare und quadratische
Verfahren sowie vergleichende
Untersuchungen zur
Klassifikation von Klaviertönen**

■ Band 2

**Forschungsberichte aus dem Institut für Nachrichtentechnik
der Universität Karlsruhe (T.H.)**

Herausgeber: Prof. Dr. rer. nat. Friedrich Jondral

Band 1 Marcel Kohl
**Simulationsmodelle für die Bewertung von Satelliten-
übertragungsstrecken im 20/30 GHz Bereich**

Band 2 Christoph Delfs
**Zeit-Frequenz-Signalanalyse: Lineare und quadratische
Verfahren sowie vergleichende Untersuchungen zur Klas-
sifikation von Klaviertönen**

Vorwort des Herausgebers

Der französische Physiker und Mathematiker Jean Baptiste Joseph de Fourier (1768-1830) hat den Naturwissenschaften und den technischen Wissenschaften, insbesondere der Nachrichtentechnik, mit der nach ihm benannten Transformation ein analytisches Werkzeug zur Verfügung gestellt, mit dem Einsichten in das Frequenzverhalten von Signalen zu gewinnen sind. Nun ist der analytische Umgang mit Signalen und Systemen eine, ihre praktische Anwendung aber eine andere Sache: Für reale Signale liegt in aller Regel keine Beschreibung in Form eines Funktionszusammenhangs vor. Darüber hinaus sind die zu untersuchenden Signale häufig Pfade nichtstationärer (und damit nichtergodischer) stochastischer Prozesse und neben anderen einschränkenden Randbedingungen können sie nur über ein endliches Zeitintervall beobachtet werden. Daraus folgt die Erkenntnis, daß die Fouriertransformierte eines Signals normalerweise nicht berechnet werden kann, sie muß vielmehr geschätzt werden. Damit hängen die gewonnenen Erkenntnisse über den Frequenzgehalt eines Signals nicht nur von diesem selbst sondern auch von der verwendeten Schätzmethode ab.

Ist das zu untersuchende Signal nichtstationär, ändert sich natürlich sein Frequenzgehalt mit der Zeit. Damit ergeben sich unbefriedigende Einblicke in die Signalstruktur, wenn nur Zeitfunktion oder Fouriertransformierte betrachtet werden. Es ist vielmehr notwendig, das Signal über der Zeit-Frequenz-Ebene darzustellen. Seitdem die digitale Signalverarbeitung Bedeutung erlangt hat, werden Verfahren zur Zeit-Frequenz-Signalanalyse ausführlich diskutiert. Insbesondere die Kurzzeit-Fourier-Transformation hat (z.B. in Realzeit-Spektralanalysegeräten oder in Vielkanalempfängern und -peilern) praktische Anwendungen gefunden.

Die vorliegende Dissertation stellt verschiedene Zeit-Frequenz-Signalanalyseverfahren vergleichend am Beispiel der Wiedererkennung von Klaviertönen dar. Ein wichtiger Aspekt der hier vorgestellten Ergebnisse ist, daß sie auf andere Signalanalyseaufgaben, wie z.B. die Klassifikation von RADARimpulsen oder die Burstdetektion, übertragen werden können. Insbesondere wird deutlich, daß mathematisch elegante, aber vergleichsweise aufwendige Algorithmen (wie die Wavelet-Packet-Transformation) in der praktischen Signalanalyse nicht unbedingt einfacheren Verfahren (wie der Kurzzeit-Fourier-Transformation) vorzuziehen sind.

Karlsruhe im Mai 1998

Friedrich Jondral

Copyright: Institut für Nachrichtentechnik
Universität Karlsruhe, 1998

Druck: Druckerei Ernst Grässer, Humboldtstr. 1,
76131 Karlsruhe, Tel. 0721/615050

ISSN: 1433-3821

**Zeit-Frequenz-Signalanalyse:
Lineare und quadratische Verfahren sowie vergleichende
Untersuchungen zur Klassifikation von Klaviertönen**

Zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

der Fakultät für
Elektrotechnik
der Universität Fridericiana Karlsruhe

genehmigte

DISSERTATION

von

Dipl.-Ing. Christoph Delfs
aus Hamburg

Tag der mündlichen Prüfung:
Hauptreferent:
Korreferent:

5. Mai 1998
Prof. Dr. rer. nat. Friedrich Jondral
Prof. Dr.-Ing. Johann F. Böhme

Danksagung

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Angestellter am Institut für Nachrichtentechnik der Universität Karlsruhe.

Mein besonderer Dank gilt an erster Stelle Herrn Prof. Dr. rer. nat. F. Jondral, Leiter des Instituts für Nachrichtentechnik der Universität Karlsruhe, für die kritische wie auch wohlwollende Begleitung und Förderung meiner Arbeit und die Übernahme des Hauptreferates.

Ebenso möchte ich mich recht herzlich bei Herrn Prof. Dr.-Ing. J. F. Böhme, Leiter des Lehrstuhles für Signaltheorie an der Ruhr-Universität Bochum, für sein Interesse an der Arbeit und die Übernahme des Korreferates bedanken.

Frau A. Olbrich möchte ich für das verlässliche Umsetzen meiner Skizzen in ansprechende Bilder und Vortragsfolien danken.

Weiterhin danke ich meiner Kollegin Anne Wiesler und meinen Kollegen Marc Ihle, Ulrich Kaage, Marcel Kohl, Ralph Machauer und Gunnar Wetzker für Ihre Unterstützung und Hilfe. Allen Studenten, die im Rahmen von Diplomarbeiten und Studienarbeiten zur Anfertigung dieser Arbeit beigetragen haben, sei herzlich gedankt.

Für viele persönliche Gespräche und Hinweise möchte ich mich bei unserer Sekretärin Frau G. Kuntermann bedanken.

Schließlich gilt mein besonderer Dank meiner Frau und meinen Eltern für ihre Unterstützung und ihr Verständnis.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Aufgabenstellung	2
1.2	Signalverarbeitung: Mathematisches Modell und physikalische Deutung	4
2	Merkmalsextraktion	7
2.1	Lineare Zeit-Frequenz-Verfahren	8
2.1.1	Reellwertige, paraunitäre FIR-Filterbänke	8
2.1.1.1	Theorie paraunitärer Filterbänke	10
2.1.1.2	Filterentwurf paraunitärer Filterbänke	17
2.1.1.3	Deutung paraunitärer Filterbänke	21
2.1.2	Wavelet-Transformation	23
2.1.2.1	Signalanalyse durch Mehrfachauflösung	25
2.1.2.2	Implementierung	28
2.1.2.3	Eigenschaften der Projektionsfilter	31
2.1.2.4	Entwurf der Projektionsfilter als paraunitäre Filterbank	35
2.1.3	Allgemeine translationsinvariante Analysefilterbänke	37
2.1.3.1	Filterbänke mit Baumstruktur	38
2.1.3.2	Translationsinvariante Filterbänke	42
2.1.3.3	Interpretation der Merkmale in der Zeit-Frequenz-Ebene	51
2.1.4	Kurzzeit-Fourier-Transformation	55
2.1.4.1	Implementierung durch eine Filterbank	56
2.1.4.2	Interpretation der Merkmale in der Zeit-Frequenz-Ebene	60
2.2	Quadratische Zeit-Frequenz-Verfahren	62
2.2.1	Die Cohen-Klasse	62
2.2.2	Ausgewählte Transformationen der Cohen-Klasse	66
2.2.3	Die Wigner-Verteilung (WV)	66
2.2.4	Die Pseudo-Wigner-Verteilung (PWV)	67
2.2.5	Die Smoothed Pseudo-Wigner-Verteilung (SPWV)	68
2.2.6	Die Choi-Williams-Verteilung (CWV)	68
2.2.7	Die diskrete Cohen-Klasse	68
2.2.8	Interpretation der Merkmale in der Zeit-Frequenz-Ebene	72
2.3	Modellbasierte Spektralschätzverfahren	73
2.3.1	Das autoregressive Modell	74
2.3.2	Schätzung der Modellparameter	80
2.3.3	Wahl der Modellordnung	81
2.3.4	Interpretation der Merkmale in der Zeit-Frequenz-Ebene	82

3	Merkmalsselektion und Klassifikation	83
3.1	Klassenunabhängige Merkmalsselektion und Klassifikation . . .	85
3.1.1	Der Algorithmus der besten Basis	86
3.1.2	Musterbildung und Klassifikation	89
3.1.2.1	Musterbildung bei Adaption und Klassifikation	89
3.1.2.2	Klassifikation	90
3.2	Klassenabhängige Merkmalsselektion und Klassifikation	91
3.2.1	Die beste Basis für klassenabhängige Merkmalsselektion .	92
3.2.2	Musterbildung und Klassifikation	94
3.2.2.1	Musterbildung bei Adaption und Klassifikation	94
3.2.2.2	Klassifikation	95
3.3	Auswertung der Klassifikationsergebnisse	96
4	Signalquelle und Signalvorverarbeitung	98
4.1	Klavier und Flügel als Signalquelle	98
4.2	Aufnahme und Signalaufbereitung	100
4.2.1	Aufnahme	100
4.2.2	Signalerkennung und Signalvorverarbeitung	101
5	Ergebnisse	104
5.1	Klassifikationsergebnisse für die vollständigen Klaviertöne . . .	104
5.1.1	Kurzzeit-Fourier-Transformation	104
5.1.2	Wavelet-Transformation	106
5.1.3	Filterbänke mit Baumstruktur	108
5.1.4	Quadratische Zeit-Frequenz-Verfahren	115
5.1.5	Modellbasierte Spektralschätzverfahren	122
5.2	Klassifikationsergebnisse für die Anschlagphase	126
5.2.1	Kurzzeit-Fourier-Transformation und modellbasierte Spek- tralschätzverfahren	126
5.2.2	Allgemeine translationsinvariante Analysefilterbänke . . .	128
5.3	Vergleich der Klassifikationsergebnisse	133
6	Anhang	137
6.1	Eigenschaften eines Hilbertraums	137
6.1.1	Innenprodukt	138
6.1.2	Norm und Metrik	138
6.1.3	Orthogonalität und orthogonale Basen	139
6.1.4	Orthogonale Unterräume	139
6.2	Elemente und Eigenschaften von Multiratenystemen	140
6.2.1	Aufwärtstaster	141
6.2.2	Abwärtstaster	142

6.2.3	Äquivalenzen	145
6.2.4	Polyphasendarstellung eines Signals	145
6.2.5	Innenprodukt und Faltung bei Dezimation	147
6.3	Quadrature-Mirror-Filterbank (QMF)	148
6.4	Evolutionäre Programme	150
6.5	Rechenaufwand der eingesetzten Verfahren zur Merkmalsex- traktion	151
6.5.1	Kurzzeit-Fourier-Transformation	151
6.5.2	Allgemeine translationsinvariante Filterbänke mit Baum- struktur	152
6.5.3	Quadratische Zeit-Frequenz-Verfahren	153
6.5.4	Modellbasierte Spektralschätzverfahren	154
6.6	Optimierte Filter	155
6.7	Aufgezeichnete Instrumente	157
6.8	Formelzeichen	159
6.9	Abkürzungen	166
7	Literaturverzeichnis	167
7.1	Diplom- und Studienarbeiten	176
8	Lebenslauf	177

Zusammenfassung

Die Aufgabe, kurze, zeitveränderliche Signale zu klassifizieren, ist für Industrie, Medizin und im Bereich der Funksignalanalyse von großer Bedeutung. Die damit verbundenen Teilaufgaben werden in dieser Arbeit mit der Erkennung eines Klaviers bzw. Flügels aus einem Pool von 34 Instrumenten anhand eines angeschlagenen Akkords an einem realen Beispiel behandelt.

In einem Schneidvorgang werden zuerst Anfang und Ende des Signals bestimmt. Anschließend folgen die Merkmalsextraktion und die Selektion der Merkmale. Durch den Klassifikator werden die ausgewählten Merkmale einer Klasse zugeordnet.

Den Schwerpunkt der Arbeit bilden die Verfahren zur Merkmalsextraktion. Zeit-Frequenz-Verfahren eignen sich besonders für die betrachtete Signalklasse, da sie die Änderungen eines Signals bezüglich der Zeit und der Frequenz erfassen. Mit ihrem Einsatz wird die Erwartung verbunden, die für eine Erkennung notwendigen Charakteristika eines Signals aus dessen Darstellung in der Zeit-Frequenz-Ebene ermitteln zu können. Jedes Verfahren ermittelt für ein Signal eine andere Repräsentation. Durch die Klassifikation wird festgestellt, welche Methode die für die Aufgabe besten Merkmale erzeugt.

Es werden lineare und quadratische Verfahren sowie alternativ modellbasierte Spektralschätzmethoden eingesetzt. Die betrachteten quadratischen Verfahren, die Wigner-Verteilung, Pseudo-Wigner-Verteilung, Smoothed Pseudo-Wigner-Verteilung und die Choi-Williams-Verteilung, gehören zur Cohen-Klasse. Bei den linearen Verfahren wird zum einen die Kurzzeit-Fourier-Transformation verwendet. Zum anderen werden orthogonale Verfahren den klassischen Ansätzen gegenübergestellt. Orthogonaltransformationen bieten den Vorteil, den erzeugten Merkmalen bei voller Signalinformation keine Redundanz hinzuzufügen. Ihr bekanntester Vertreter, die Wavelet-Transformation, wird seit mehreren Jahren für viele Anwendungen, darunter auch für die in dieser Arbeit vorliegende, diskutiert.

Die Wavelet-Transformation ist ein Sonderfall orthogonaler Filterbänke. Die dargelegte Erweiterung auf Filterbänke mit Baumstruktur stellt eine Menge an Orthogonaltransformationen zur Verfügung und ist daher die Grundlage für eine aufgabenangepasste Zeit-Frequenz-Darstellung des Signals. Die Einführung translationsinvarianter Filterbänke erlaubt, die Zeit-Frequenz-Repräsentation für mehrere, zeitverschobene Signalversionen zugleich zu berechnen. Die von einem Verfahren erzeugten Merkmale können in der Zeit-Frequenz-Ebene visualisiert werden und erlauben, die Klassifikationsergebnisse zu interpretieren.

Teil der Merkmalsselektion ist im Fall der Filterbänke mit Baumstruktur die Wahl einer geeigneten Orthogonaltransformation durch den Algorithmus der besten Basis. Dadurch gelingt eine besonders kompakte Beschreibung des Signals durch wenige Merkmale. Bei allen Verfahren werden diejenigen Merkmale für die Musterbildung ausgewählt, die den Maxima der Zeit-Frequenz-Darstellung entsprechen. Neben dieser klassenunabhängigen Selektion wird bei Filterbänken mit Baumstruktur auch eine klassenabhängige Variante vorgestellt.

Die Klassifikation erfolgt durch den Nächster-Nachbar-Klassifikator. Klassenabhängige bzw. -unabhängige Selektion führen zu unterschiedlichen Klassifikatorkonzepten.

Während des Schneidvorgangs wird alternativ das Ende eines angeschlagenden Akkords durch dessen Verklingen festgestellt oder durch das Ende der Anschlagphase bestimmt. Für beide Varianten werden jeweils Klassifikationsläufe durchgeführt.

Bei der Klassifikation der vollständigen Klaviertöne erweist sich die Kurzzeit-Fourier-Transformation gegenüber allen Verfahren in Hinblick auf Rechenaufwand und Ergebnis als überlegen. Die aus Orthogonaltransformationen abgeleiteten Merkmale weisen eine große Abhängigkeit vom Schneidvorgang auf. Die geringe spektrale Selektivität der zugehörigen Filterbänke erschwert nicht nur die visuelle Analyse, sondern führt auch zu weniger effizienten Merkmalen. Quadratische Verfahren und modellbasierte Spektralschätzverfahren erreichen die Ergebnisse der Orthogonaltransformationen.

Für die Klassifikation der Anschlagphase eignen sich die aus einer modellbasierten Spektralschätzung gewonnenen Merkmale am besten. Durch translationsinvariante Filterbänke erreichen die Orthogonaltransformationen vergleichbare Klassifikationsresultate. Die klassenabhängige Merkmalsselektion führt zu keiner Verbesserung der Erkennungsergebnisse. Der hohe Rechenaufwand, der mit einer translationsinvarianten Filterung und Klassifikation verbunden ist, macht die Verwendung der Filterbänke kaum attraktiv.

Der Einsatz der Wavelet-Transformation bzw. der Orthogonaltransformationen zur Merkmalsextraktion in der Klassifikation wird in der Literatur empfohlen. Die vorliegende Arbeit zeigt, daß in der vorgestellten, realen Anwendung ein Vorteil dieser Verfahren nicht zu erkennen ist. Vielmehr darf auf einen Vergleich mit traditionellen Verfahren nicht verzichtet werden.

1 Einleitung

Werkzeuge und Maschinen haben den Menschen befähigt, die durch die eigene Geschicklichkeit und Muskelkraft vorgegebenen Beschränkungen zu lockern. Fortgeschrittene Sensorik und digitale Signalverarbeitung werden entwickelt, um auch die menschliche Reaktion auf Ereignisse zu ersetzen.

Ebenso wie der Entwurf mechanischer Hilfen durch die Erfahrung des Entwicklers stets „menschlich“ geprägt ist, bleibt die automatische Entscheidungsfindung durch sensorische Auswertung auf die Fälle begrenzt, die der Mensch auch selbst bewältigen kann. Man kann von einem Automaten nicht mehr an Ergebnis erwarten, als man zu seinem Entwurf an Wissen hineinsteckt hat. Die dennoch auch in der Wissenschaft anzutreffende Bereitschaft, neuen Verfahren unerklärbare Eigenschaften zuzuordnen [Dea97], mag tiefere Ursachen haben.

Die automatische Klassifikation von Ereignissen durch einen Rechner hat zum Ziel, die menschliche Fähigkeit nachzuahmen, zu abstrahieren und eine Entscheidung auf der Grundlage weniger, aber wichtiger Merkmale zu treffen. Der Einsatz des Rechners soll dabei Routineentscheidungen des Menschen ersetzen. Wenn der Computer eine Klassifikationsaufgabe lösen soll, muß das Abstraktionsvermögen des Menschen auf Algorithmen abgebildet werden. Dafür gibt es zwei Ansätze:

- Der Rechner erhält fertige Merkmale, zum Beispiel als Ergebnis einer Modellbildung oder durch Expertenwissen. Diese Lösung setzt voraus, daß dem Menschen die Merkmale bereits bekannt sind.
- Der Rechner muß wie der Mensch am Beispiel lernen, Wichtiges von Unwichtigem zu unterscheiden. Dieser Weg verlagert den Aufwand in den Rechner. Die Anpassung an die Klassifikationsaufgabe erfolgt durch Trainingsalgorithmen. Die Trainingsalgorithmen bestehen aus einer Phase der Merkmalsfindung (Merkmalsextraktion) und der Abstraktion (Merkmalsselektion) auf einem Satz an Beispielsignalen.

In dieser Arbeit wird der zweite Weg beschritten. Dieser Ansatz birgt die Schwierigkeit, zunächst geeignete Merkmale zu erzeugen.

Die Zeit-Frequenz-Analyse versucht zu erfassen, wie sich der spektrale Gehalt eines Signals mit der Zeit ändert. Da sich jedes physikalische Signal mit der Zeit ändert, sind Zeit-Frequenz-Beschreibungen eines Signals als universelle Merkmalslieferanten geeignet und dienen in der vorliegenden Arbeit zur Merkmalsextraktion.

Die Erkennung eines Klaviers oder Flügels anhand eines angeschlagenen Akkords dient als Beispiel für die Aufgabenstellung, transiente Signale zu er-

kennen. Dem Klavierkenner fällt diese Unterscheidung leicht, für den Laien erscheint die Aufgabe kaum lösbar. Daher ist die Automatisierung der Erkennung dieser Signalquelle vergleichbar mit der Anforderung, den Experten für Routineaufgaben durch den Rechner zu entlasten.

Die Spracherkennung ist ein populärer Vertreter für das Gebiet der Erkennung transients Signale. Weitere Aufgabenbereiche sind:

- Im industriellen Bereich werden Verfahren zur Überwachung, Test und Diagnose entwickelt. In [Hei95] werden die Geräusche eines Motors für einen Qualitätstest herangezogen. Eine Vibrationsanalyse dient zur Bestimmung von Fehlerzuständen eines Helikopterrotors [BPJ92]. Das Aufspüren von Kontaktgeräuschen zwischen Metallteilen einer Anlage, die sich nicht berühren dürfen, dient der Früherkennung eines drohenden Anlagenausfalls [PW92]. Störungen in Hochspannungssystemen z.B. durch Kurzschlüsse oder Blitzeinschläge, werden für eine post-mortem-Analyse zur Fehlerdiagnose verwendet [RCM94].
- In medizinischen Anwendungen werden Anomalien in EEG-Signalen zum Aufspüren bzw. der Diagnose von Herzfehlern [LZT95] verwendet. ECG-Signale lassen sich zur Erkennung von Gehirnerkrankungen [HCTS97] nutzen.
- Im Bereich der Funksignalanalyse eignet sich die Erkennung transients Signale für die Klassifikation von Unterwassergeräuschen [BPK92, LGR97, LAJ94], die durch Schiffe oder U-Boote erzeugt werden. Auch die Erkennung von Walgeräuschen [LKW92] mag diesem Aufgabengebiet zugeordnet werden. Flugzeuge können durch ihr Start- und Landegeräusch identifiziert werden [DK94]. Die Unterscheidung von Geschütztypen anhand ihres Knalls wird in [Har96] untersucht.

1.1 Aufgabenstellung

Bisherige Untersuchungen an realen Signalen stellen im allgemeinen ein Verfahren zur Zeit-Frequenz-Analyse vor und prüfen die Eignung der gewonnenen Merkmale für die Klassifikation, z.B. [Hei95, LKW92, HCTS97, DK94]. Dagegen werden bei Einsatz mehrerer Verfahren alle erzeugten Merkmale gemeinsam verarbeitet, z.B. [BPJ92], oder beschränken sich auf eine visuelle Analyse, z.B. [BPK92].

Jedoch kann nur eine vergleichende Untersuchung von Zeit-Frequenz-Verfahren die für die vorliegende Anwendung optimale Transformation hinsicht-

lich des Aufwands und des Erfolgs ermitteln. Bisher ist keine Arbeit bekannt, die diese Aufgabenstellung für reale Signale behandelt.

Ende der 80er Jahre [CGT89] wurde mit der Wavelet-Transformation ein neues Verfahren in die Menge der Zeit-Frequenz-Analyseverfahren aufgenommen, das durch seine neuen Gestaltungsmöglichkeiten Interesse erregte. Durch die Einführung der orthogonalen Wavelet-Transformation wurde ein Boom in der digitalen Signalverarbeitung ausgelöst: Ausgehend von Anwendungen in der Datenkompression, z.B. [ST93], wurden Wavelets für eine visuelle Datenanalyse und in der Klassifikation eingesetzt. Seit kurzem wird auch die Impulsformung durch Wavelets bei digitalen Modulationsverfahren erprobt [LT96, Lin97, WWDJ97].

Die unabhängig entwickelte Theorie orthogonaler Filterbänke [Vai87], als deren Sonderfall sich die orthogonale Wavelet-Transformation herausstellte, hat in der Anwendung dagegen neben der Datenkompression bislang kaum Resonanz gefunden. Der Vorteil orthogonaler Filterbänke besteht darin, daß das Signal ohne Informationsverlust und ohne zusätzliche Redundanz transformiert werden kann. Die Ergänzung zu Filterbänken mit Baumstruktur eröffnet die Möglichkeit, aus einer Menge von Orthogonaltransformationen wählen zu können. Der Schwerpunkt der vorliegenden Arbeit liegt in dem Vergleich orthogonaler Transformationen durch Filterbänke mit herkömmlichen Zeit-Frequenz-Verfahren. Der Vergleich schließt die Erweiterung auf translationsinvariante Filterbänke und klassenabhängige Merkmalsextraktion ein. Bisherige Arbeiten zu orthogonalen Filterbänken verzichten auf einen Vergleich bzw. verwenden synthetisch erzeugte Signale [Sai94, MWJ94].

Die Aufgabenstellung gliedert sich in vier Teilaufgaben:

1. In der Signalvorverarbeitung werden Anfang und Ende des Eingangssignals ermittelt und eine Filterung durchgeführt.
2. Die Merkmalsextraktion bildet das Eingangssignal durch eine Transformation auf eine Menge von Merkmalen ab.
3. Die Merkmalsselektion reduziert die Menge der erzeugten Merkmale durch eine Auswahl und bildet sie auf Muster ab.
4. In der abschließenden Klassifikation wird das Muster einer Klasse zugeordnet.

Der Ablauf der Signalverarbeitung und die Zuordnung zu den Kapiteln dieser Arbeit ist in Bild 1.1 skizziert.

In Kapitel 5 werden die Klassifikationsergebnisse vorgestellt und diskutiert.

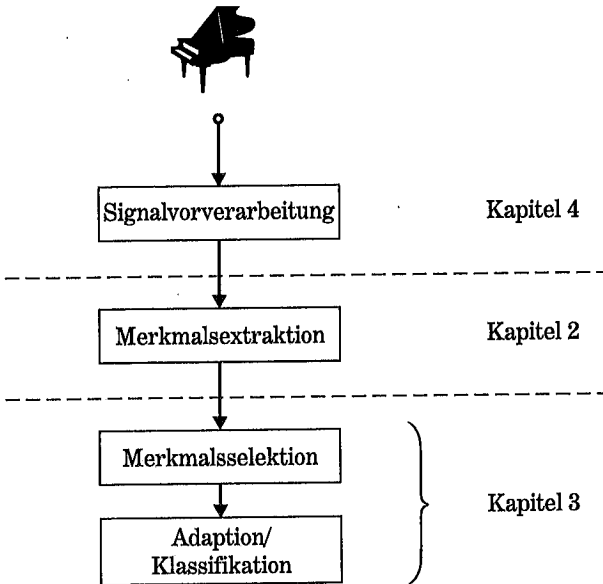


Bild 1.1: Aufbau der Signalverarbeitung zur Klassifikation und ihre Zuordnung zu den Kapiteln

1.2 Signalverarbeitung: Mathematisches Modell und physikalische Deutung

Das zu klassifizierende Signal liegt zunächst als physikalisches Ereignis vor, das durch einen Sensor wahrgenommen und durch eine analoge Signalverarbeitung verstärkt wird.

Durch Zeit- und Wertquantisierung (Analog-Digital-Wandlung) am Eingang des Digitalrechners wird das Signal auf eine Folge von Zahlen abgebildet und liegt in einer rechnerspezifischen Repräsentation vor.

In einer Idealisierung mag das analoge Sensorsignal und die digitale Form als mathematisches Signalmodell verstanden werden. Diese Idealisierung berücksichtigt z.B. nicht die Eigenschaften des Sensors, des Analogverstärkers, die abschließende Wertquantisierung, die zeitlichen Ungenauigkeiten bei der Abtastung (Jitter) oder die endliche Beobachtungsdauer. Sie erlaubt jedoch die Zuordnung des zu verarbeitenden Signals zu einem Element eines Signalraumes. Die Eigenschaften des verwendeten Signalraumes erlauben die für

die Merkmalsextraktion verwendeten Transformationen. Die vom Rechner durchgeführten Transformationen werden durch Algorithmen beschrieben. Die ermittelten Merkmale können wiederum physikalisch interpretiert werden.

Im folgenden soll zwischen dem physikalischen und dem mathematischen Modell unterschieden werden, da physikalische Größen durch einen Zahlenwert und eine Einheit, mathematische Größen hingegen nur durch einen Zahlenwert beschrieben werden.

Dem mathematischen Modell kann eine physikalische Deutung wie in Bild 1.2 gegenüber gestellt werden. Das Symbol $\circ \bullet$ repräsentiert die Fourier-Transformation, die für Funktionen und Folgen aus den betrachteten Signalmräumen definiert ist [Chu92, Fli91]. Bild 1.2 stellt das kontinuierliche Zeitsignal $s(t)$ im Frequenzbereich durch seine Fourier-Transformierte $S(j\omega)$ dar. Die Fourier-Transformierte des (zeit)diskreten Signals $S(e^{j\omega})$ bzw. $S(e^{j\omega\Delta t})$ ist zusätzlich in ihrer frequenzdiskretisierten Form durch $S(e^{j2\pi n/N})$ angegeben. Die Größen $\Delta\omega$ und Δf stellen die resultierende Rasterung bezüglich der Kreisfrequenz ω und der Frequenz $f = \omega/(2\pi)$ dar.

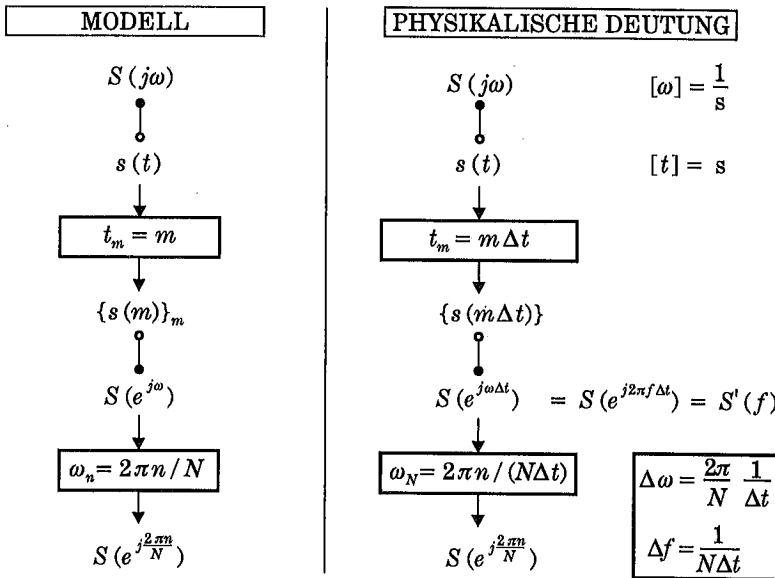


Bild 1.2: Modell und physikalische Deutung für das abgetastete und kontinuierliche Signal $s(t)$ im Zeit- und Frequenzbereich

In Bild 1.3 ist der Ablauf der Signalverarbeitung in mathematischem Modell und physikalischer Ebene bis zur Merkmalsextraktion abgebildet. Die physikalische Interpretation der Merkmale $F(m, n)$ in der Zeit-Frequenz-Ebene mit m als Zeitparameter und n als Frequenzparameter wird in Kapitel 2 für jedes Verfahren der Merkmalsextraktion gegeben.

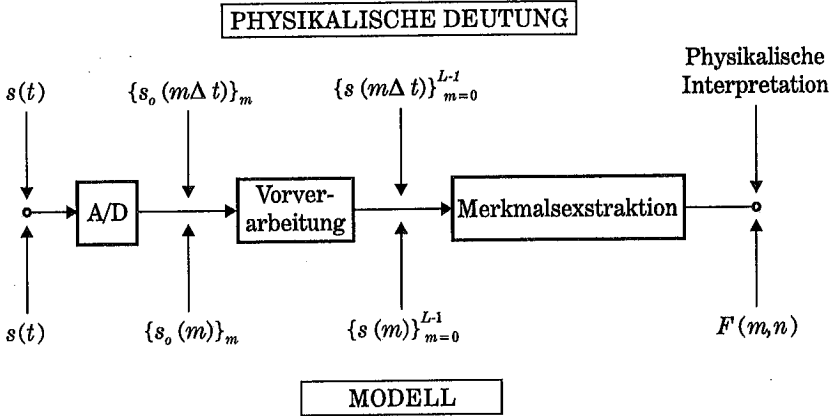


Bild 1.3: Mathematisches Modell und physikalische Deutung von dem kontinuierlichen Signal bis zur Merkmalsextraktion

In der Literatur wird die Verbindung von mathematischem Modell und physikalischer Deutung uneinheitlich behandelt und führt zu verfahrensabhängigen Darstellungen [Fli93]. Die Trennung beider Ebenen, wie z.B. in [Vai93], erlaubt eine einheitliche Beschreibung der Merkmale unabhängig von ihrer Zuordnung zu einem Frequenzband oder einem Zeittakt. Dadurch wird eine einfache Beschreibung des Übergangs von der Merkmals- auf die Musterdarstellung möglich. Diese Darstellung hat allerdings ebenso zur Folge, daß der Takt einer zeitdiskreten Folge $\{s(m)\}_m$ stets aus dem Zusammenhang erkannt werden muß.

2 Merkmalsextraktion

Für die Merkmalsextraktion wird das zu untersuchende Signal als Element eines Signalraumes interpretiert. Die Eigenschaften des betrachteten Signalraumes erlauben die Anwendung der für die Merkmalsbildung benötigten Transformationen.

In der Merkmalsextraktion werden die Hilberträume $l^2(\mathbb{Z})$ bzw. $L^2(\mathbb{R})$ betrachtet. Zwischen zwei Elementen eines Hilbertraums ist das Innenprodukt definiert, das die zentrale Grundlage aller verwendeten Analyseverfahren bildet. Im Anhang 6.1 sind die für die Merkmalsextraktion wesentlichen Eigenschaften eines Hilbertraums, d.h. Norm, Metrik und Innenprodukt, und die im folgenden verwendete Notation für Vektoren erläutert.

In Abschnitt 2.1 werden lineare Verfahren vorgestellt, bei denen mit Hilfe des Innenproduktes das Signal auf eine Merkmalsmenge abgebildet wird. Erfolgt diese Innenproduktbildung mit einer Menge von Funktionen, die eine orthonormale Basis des betrachteten Funktionenraums bilden, spricht man von orthogonalen Verfahren. Orthogonale Verfahren, insbesondere die Wavelet-Transformation, bilden den Schwerpunkt der hier verwendeten linearen Merkmalsextraktion.

Die quadratischen Zeit-Frequenz-Verfahren in Abschnitt 2.2 werden über eine Analogie aus der Quantenmechanik hergeleitet. Die das quantenmechanische System beschreibende Wellenfunktion besitzt zufälligen Charakter. Durch Erwartungswertbildung über den Orts- und den Impulsoperator wird das System bezüglich des Ortes und des Impulses charakterisiert. Die Formulierung der Erwartungswertbildung erfolgt durch eine Innenproduktbildung im Hilbertraum. Für die Signalanalyse wird die Wellenfunktion durch das zu untersuchende Signal ersetzt. An die Stelle des Orts- bzw. Impulsoperators treten Zeit- bzw. Frequenzoperator.

Das Signal wird in Abschnitt 2.3 als Realisierung eines durch ein autoregressives Modell beschreibbaren, zeitdiskreten stochastischen Prozesses interpretiert. Die Interpretation der Zufallsvariablenfolge als Menge von Vektoren erlaubt eine geometrische Deutung des Prozeßmodells über das Innenprodukt.

In Anhang 6.5 ist der für die Implementierung der Verfahren benötigte Rechenaufwand zusammengestellt.

2.1 Lineare Zeit-Frequenz-Verfahren

Als lineare Zeit-Frequenz-Verfahren werden Multiratenysteme [Fli93] eingesetzt, d.h. der Takt der Ausgangsgrößen unterscheidet sich um einen ganzzahligen Faktor vom Takt des Eingangssignals. Eine Darstellung der Elemente von Multiratenystemen und ihrer Eigenschaften ist in Anhang 6.2 angegeben.

Zunächst wird in Abschnitt 2.1.1 die paraunitäre Filterbank als grundlegende orthogonale Abbildung im Folgenraum $l^2(\mathbf{Z})$ vorgestellt. Anschließend wird in Abschnitt 2.1.2 die Wavelet-Transformation diskutiert. Die orthogonale Wavelet-Transformation wird als kaskadierte paraunitäre Filterbank mit zusätzlichen Anforderungen an die Filterkoeffizienten implementiert.

Die paraunitäre Filterbank als Grundbaustein einer orthogonalen Abbildung wird in Abschnitt 2.1.3 durch eine Filterbank mit Baumstruktur und durch translationsinvariante Filterung verallgemeinert. Die Merkmale der orthogonalen Verfahren werden gemeinsam in Abschnitt 2.1.3.3 in der Zeit-Frequenz-Ebene interpretiert.

Die Kurzzeit-Fourier-Transformation als klassisches Verfahren der Zeit-Frequenz-Analyse bildet in Abschnitt 2.1.4 das letzte und einzige, hier diskutierte, nichtorthogonale, lineare Zeit-Frequenz-Verfahren.

2.1.1 Reellwertige, paraunitäre FIR-Filterbänke

Paraunitäre FIR-Filterbänke gehören zu der Klasse der perfekt rekonstruierenden Filterbänke. In Bild 2.1 ist die Filterbank dargestellt. Paraunitäre Filterbänke vermitteln eine orthogonale Abbildung der Eingangsfolge $\{s(m)\}_m$ auf die Folgen $\{y_n(m)\}_m$.

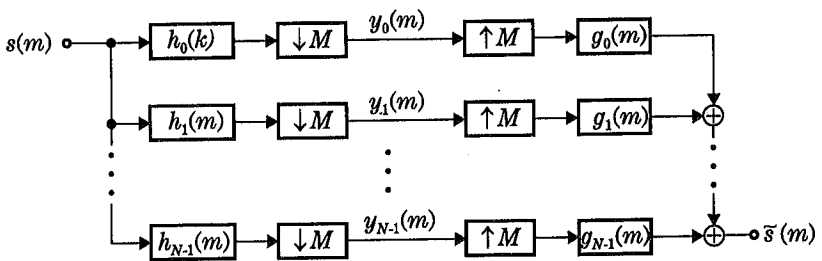


Bild 2.1: Signalfluß mit paraunitärer Analyse- und Synthesefilterbank in einem perfekt rekonstruierenden System

Die Eingangsfolge $\{s(m)\}_m$ wird als unendliche Folge beschrieben, wobei in der Anwendung nur eine beliebige endliche Anzahl von Elementen von Null verschieden ist. Die Folge $\{s(m)\}_m$ wird durch eine N -kanalige Filterbank mit den Filtern $\{h_n(k)\}_{k=0}^{K-1}$, die eine endliche Impulsdauer aufweisen, mit $n = 0, 1, \dots, N-1$, bei einer Dezimation um den Faktor M auf die Folgen $\{y_n(m)\}_m$ abgebildet:

$$y_n(m) = \sum_{k=-\infty}^{\infty} s(k)h_n(mM - k), \quad m \in \mathbb{Z}.$$

Man beachte, daß physikalisch den Folgen $\{y_n(m)\}_m$ ein um den Faktor M kleinerer Takt gegenüber der Eingangsfolge $\{s(m)\}_m$ zugeordnet wird.

Durch die Angabe $\{h_n(k)\}_{k=0}^{K-1}$ werden hier unendliche Folgen beschrieben, deren Elemente außerhalb des Intervalls $[0, K-1]$ Null sind.

Durch die Synthesefilterbank $\{g_n(k)\}_{k=0}^{K-1}$ werden die Folgen $\{y_n(m)\}_m$ zur Folge $\{\tilde{s}(m)\}_m$ zusammengesetzt:

$$\tilde{s}(m) = \sum_{n=0}^{N-1} \sum_{k=-\infty}^{\infty} y_n(k)g_n(m - kM), \quad m \in \mathbb{Z}.$$

Perfekte Rekonstruktion liegt vor, wenn die rekonstruierte Folge $\{\tilde{s}(m)\}_m$ bis auf eine Laufzeit L' mit der Eingangsfolge $\{s(m)\}_m$ übereinstimmt:

$$\{\tilde{s}(m)\}_m = \{s(m - L')\}_m, \quad m \in \mathbb{Z}. \quad (2.1)$$

Im folgenden werden nur reellwertige, paraunitäre Systeme betrachtet. Wie in Abschnitt 2.1.1.1 gezeigt wird, besteht die Filterentwurfsaufgabe bei kausalen, paraunitären FIR-Systemen darin, die Bedingung (2.1) unter der weiteren Bedingung

$$\{g_n(k)\}_{k=0}^{K-1} = \{h_n(K-1-k)\}_{k=0}^{K-1} \quad (2.2)$$

zu erfüllen.

In Abschnitt 2.1.1.2 wird beschrieben, wie ein paraunitäres System als Produkt frei parametrisierbarer, paraunitärer Teilsysteme entworfen wird. Auf diese Weise wird die Entwurfsaufgabe mit den Randbedingungen (2.1) und (2.2) durch eine Strukturvorgabe in eine Entwurfsaufgabe ohne Beschränkungen überführt. Die Interpretation der Filterbank als orthogonale Abbildung ist Gegenstand des Abschnitts 2.1.1.3.

Gilt $N = M$, heißt die Filterbank kritisch abgetastet, für $N > M$ spricht man von einer überabgetasteten Filterbank. Im folgenden wird der Filterentwurf für den Fall $N = M$ vorgestellt.

2.1.1.1 Theorie paraunitärer Filterbänke

Die Darstellung der Analyse- und Synthesefilter durch ihre Polyphasenkomponenten bezüglich der Dezimation M vereinfacht die Formulierung der Entwurfsaufgabe. Die Definitionen der Polyphasenkomponenten erster und zweiter Art sind in Anhang 6.2.4 angegeben.

Ausgangspunkt der folgenden Umformungen sind die z -Transformierten der Analyse- und Synthesefilter sowie der Teilbandfolgen $\{y_n(m)\}_m$:

$$\begin{aligned} H_n(z) &= \sum_{k=0}^{K-1} h_n(k) z^{-k}, \\ G_n(z) &= \sum_{k=0}^{K-1} g_n(k) z^{-k} \quad \text{und} \\ Y_n(z) &= \sum_{m=-\infty}^{\infty} y_n(m) z^{-m}, \quad n = 0, 1, \dots, M-1. \end{aligned}$$

Für eine einfache Darstellung wird die später in diesem Abschnitt hergeleitete Eigenschaft, daß die Länge der Impulsantwort K ein ganzzahliges Vielfaches des Dezimationsfaktors M ist, benutzt:

$$K = M\tilde{G}, \quad \tilde{G} \in \mathbb{N}. \quad (2.3)$$

Die Polyphasenkomponenten erster Art $H_{n,\lambda}^{(1)}(z)$ mit $\lambda = 0, 1, \dots, M-1$ der Analysefilter $H_n(z)$ lauten mit der Größe $\tilde{G}$:

$$H_{n,\lambda}^{(1)}(z) = \sum_{k=0}^{\tilde{G}} h_n(kM + \lambda) z^{-k}. \quad (2.4)$$

Umgekehrt folgt dann ebenso:

$$H_n(z) = \sum_{\lambda=0}^{M-1} z^{-\lambda} H_{n,\lambda}^{(1)}(z^M). \quad (2.5)$$

Für die Synthesefilterbank werden die Polyphasenkomponenten zweiter Art verwendet:

$$G_{\lambda,n}^{(2)}(z) = \sum_{k=0}^{\tilde{G}} G_n((k+1)M-1-\lambda) z^{-k}. \quad (2.6)$$

Die Synthesefilter lassen sich umgekehrt durch

$$G_n(z) = \sum_{\lambda=0}^{M-1} z^{-(M-1-\lambda)} G_{\lambda,n}^{(2)}(z^M) \quad (2.7)$$

ausdrücken. Mit den z -Transformierten $S(z)$ und $\tilde{S}(z)$ der Eingangsfolge bzw. der rekonstruierten Folge lassen sich Analyse- und Synthesefilterbank aus Bild 2.1 im z -Bereich durch Verwendung der Polyphasenfilter $H_{n,\lambda}^{(1)}(z)$ und $G_{\lambda,n}^{(2)}(z)$ darstellen wie in Bild 2.2 und Bild 2.3.

Unter Ausnutzung der in Anhang 6.2.3 dargestellten Äquivalenzen läßt sich für das System die Darstellung in Bild 2.4 angeben, bei der alle Filter aufwandsgünstig mit einem um den Faktor M reduzierten Takt arbeiten.

In Bild 2.4 werden die Zwischengrößen $\tilde{S}_{\lambda}^{(2)}(z)$ und $U_{\lambda}^{(2)}(z)$ verwendet. Es sind $\tilde{S}_{\lambda}^{(2)}(z)$ die Polyphasenkomponenten zweiter Art von $\tilde{S}(z)$. Man kann mit Anhang 6.2.4 zeigen, daß $U_{\lambda}^{(2)}(z)$ die Polyphasenkomponenten zweiter Art der Größe $U(z) = z^{-(M-1)}S(z)$ sind.

Sie stehen mit den Polyphasenkomponenten zweiter Art $S_{\lambda}^{(2)}(z)$ des Eingangssignals $S(z)$ auf folgende Weise in Beziehung:

$$U_{\lambda}^{(2)}(z^M) = z^{-(M-1)} S_{\lambda}^{(2)}(z^M). \quad (2.8)$$

Aus Bild 2.4 läßt sich für die Polyphasensignale $\tilde{S}_{\lambda}^{(2)}(z)$ mit Hilfe der Polyphasenmatrix

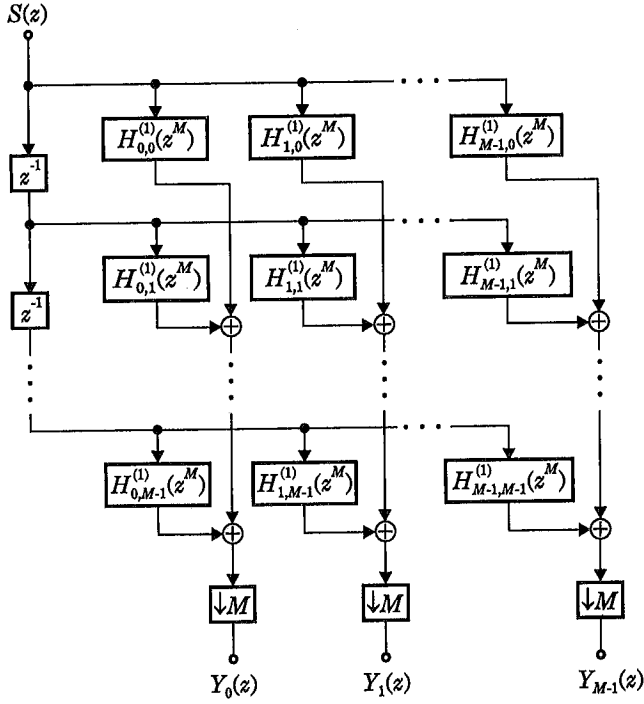


Bild 2.2: Darstellung der Analysefilterbank durch die Polyphasenfilter

$$\mathbf{G}^{(2)}(z) = \begin{pmatrix} G_{0,0}^{(2)}(z) & G_{0,1}^{(2)}(z) & G_{0,2}^{(2)}(z) & \dots & G_{0,M-1}^{(2)}(z) \\ G_{1,0}^{(2)}(z) & G_{1,1}^{(2)}(z) & G_{1,2}^{(2)}(z) & \dots & G_{1,M-1}^{(2)}(z) \\ \dots & \dots & \dots & \dots & \dots \\ G_{M-1,0}^{(2)}(z) & G_{M-1,1}^{(2)}(z) & G_{M-1,2}^{(2)}(z) & \dots & G_{M-1,M-1}^{(2)}(z) \end{pmatrix}.$$

folgender Zusammenhang angeben:

$$\tilde{S}_{\lambda}^{(2)}(z) = \sum_{n=0}^{M-1} G_{\lambda,n}^{(2)} Y_n(z),$$

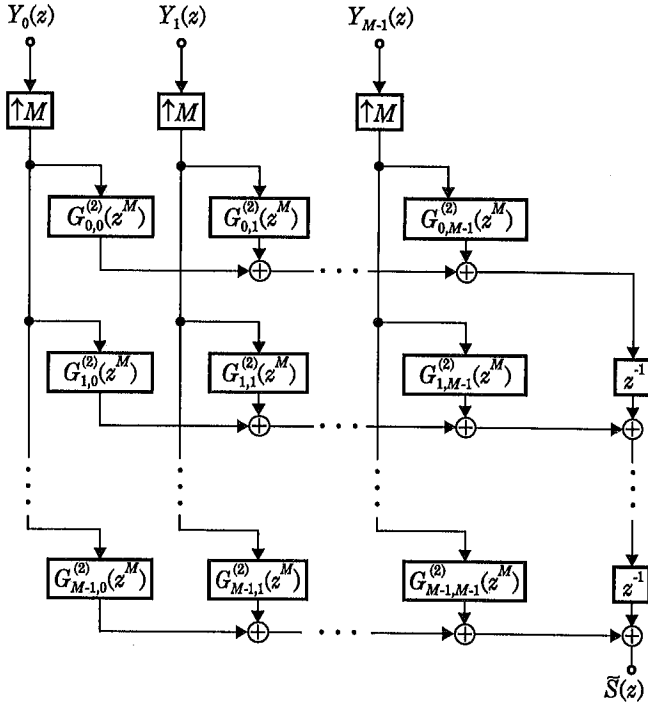


Bild 2.3: Darstellung der Synthesefilterbank durch die Polyphasenfilter

$$\begin{pmatrix} \tilde{S}_0^{(2)}(z) \\ \tilde{S}_1^{(2)}(z) \\ \vdots \\ \tilde{S}_{M-1}^{(2)}(z) \end{pmatrix} = \mathbf{G}^{(2)}(z) \begin{pmatrix} Y_0(z) \\ Y_1(z) \\ \vdots \\ Y_{M-1}(z) \end{pmatrix}. \quad (2.9)$$

Ebenso folgt für die Teilbandsignale $Y_n(z)$ mit Hilfe der Polyphasenmatrix

$$\mathbf{H}^{(1)}(z) = \begin{pmatrix} H_{0,0}^{(1)}(z) & H_{0,1}^{(1)}(z) & H_{0,2}^{(1)}(z) & \cdots & H_{0,M-1}^{(1)}(z) \\ H_{1,0}^{(1)}(z) & H_{1,1}^{(1)}(z) & H_{1,2}^{(1)}(z) & \cdots & H_{1,M-1}^{(1)}(z) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{M-1,0}^{(1)}(z) & H_{M-1,1}^{(1)}(z) & H_{M-1,2}^{(1)}(z) & \cdots & H_{M-1,M-1}^{(1)}(z) \end{pmatrix} \quad (2.10)$$

die Darstellung:

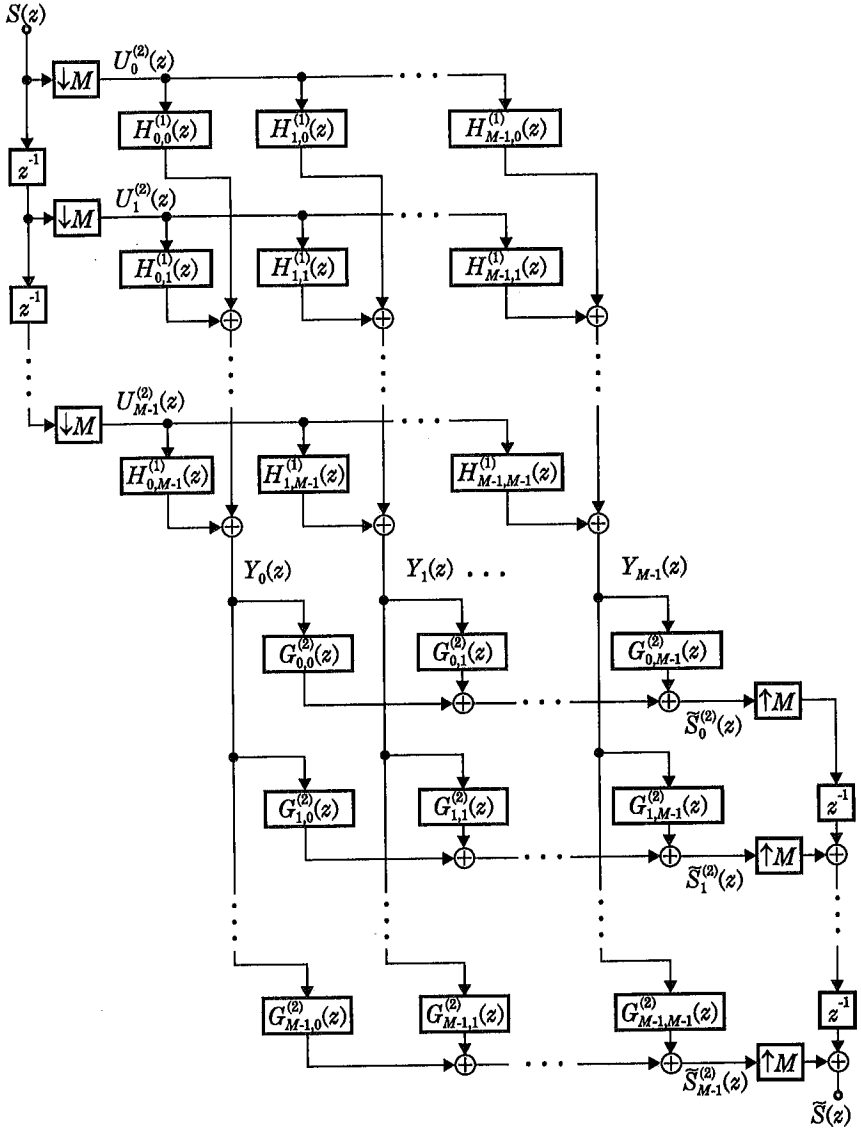


Bild 2.4: Darstellung des perfekt rekonstruierenden Systems aus Bild 2.1 durch Polyphasenfilterbänke auf Analyse- und Syntheseseite

$$Y_n(z) = \sum_{\lambda=0}^{M-1} H_{n,\lambda}^{(1)} U_{\lambda}^{(2)}(z), \quad n = 0, 1, \dots, M-1,$$

$$\begin{pmatrix} Y_0(z) \\ Y_1(z) \\ \dots \\ Y_{M-1}(z) \end{pmatrix} = \mathbf{H}^{(1)}(z) \begin{pmatrix} U_0^{(2)}(z) \\ U_1^{(2)}(z) \\ \dots \\ U_{M-1}^{(2)}(z) \end{pmatrix}. \quad (2.11)$$

Über (2.9) und (2.11) folgt für den Zusammenhang der Signale $U_{\lambda}^{(2)}(z)$ und $\tilde{S}_{\lambda}^{(2)}(z)$:

$$\begin{pmatrix} \tilde{S}_0^{(2)}(z) \\ \tilde{S}_1^{(2)}(z) \\ \dots \\ \tilde{S}_{M-1}^{(2)}(z) \end{pmatrix} = \mathbf{G}^{(2)}(z) \mathbf{H}^{(1)}(z) \begin{pmatrix} U_0^{(2)}(z) \\ U_1^{(2)}(z) \\ \dots \\ U_{M-1}^{(2)}(z) \end{pmatrix}. \quad (2.12)$$

Perfekte Rekonstruktion liegt vor, wenn:

$$\mathbf{G}^{(2)}(z) = z^{-G'} [\mathbf{H}^{(1)}(z)]^{-1} = z^{-G'} \frac{\mathbf{H}^{A,(1)}(z)}{\det \mathbf{H}^{(1)}(z)}, \quad G' \in \mathbf{Z} \quad (2.13)$$

mit $\det \mathbf{H}^{(1)}(z) \neq 0$. Die Matrix $\mathbf{H}^{A,(1)}(z)$ ist die adjungierte Matrix zu der Polyphasenmatrix $\mathbf{H}^{(1)}(z)$, wobei jedes Element $H_{n,\lambda}^{A,(1)}(z)$ der Matrix $\mathbf{H}^{A,(1)}(z)$ die zum Element $H_{n,\lambda}^{(1)}(z)$ gehörige Adjunkte ist.

Unter der Bedingung (2.13) ergibt sich die Eigenschaft

$$\det \mathbf{G}^{(2)}(z) \det \mathbf{H}^{(1)}(z) = z^{-G'} \quad (2.14)$$

für die Determinanten der Polyphasenmatrizen.

Die aufwendige und ggf. auf instabile Synthesefilter führende Matrixinversion läßt sich durch den Ansatz einer paraunitären Polyphasenmatrix $\mathbf{H}^{(1)}(z)$ umgehen:

Die Polyphasenmatrix $\mathbf{H}^{(1)}(z)$, deren Elemente $H_{n,\lambda}^{(1)}(z)$ reelle Koeffizienten besitzen, heißt paraunitär, wenn gilt [VK95, Fli93, Vai93]:

$$[\mathbf{H}^{(1)}(z)]^{-1} = c[\mathbf{H}^{(1)}(z^{-1})]^T, \quad c \in \mathbb{R} \setminus \{0\}. \quad (2.15)$$

Im folgenden wird stets $c = 1$ gewählt. In diesem Fall heißt das System normiert paraunitär [Vai93]. Ist jedes Filter $H_{n,\lambda}^{(1)}(z)$ der Matrix $\mathbf{H}^{(1)}(z)$ stabil, heißt das System verlustlos.

Ist nun $\mathbf{H}^{(1)}(z)$ paraunitär und wird die Polyphasenmatrix

$$\mathbf{G}^{(2)}(z) = z^{-G'} [\mathbf{H}^{(1)}(z^{-1})]^T \quad (2.16)$$

gewählt, so folgt aufgrund (2.15) für das Produkt beider Polyphasenmatrizen:

$$\mathbf{G}^{(2)}(z) \mathbf{H}^{(1)}(z) = \mathbf{I} z^{-G'}, \quad (2.17)$$

mit $\mathbf{I}$ als (M, M) -Einheitsmatrix.

Die Paraunitarität von $\mathbf{H}^{(1)}(z)$ sei im folgenden stets vorausgesetzt.

Der Parameter G' legt die Laufzeit des Systems fest. Für realisierbare Systeme gilt daher stets $G' \geq 0$. Mit Bild 2.4 oder über eine kleine Nebenrechnung unter Beachtung von (2.8) kann man zeigen, daß das System in Bild 2.1 mit der Laufzeit $L' = (G' + 1)M - 1$ die Bedingung (2.1) erfüllt:

$$\tilde{S}(z) = z^{-((G'+1)M-1)} S(z). \quad (2.18)$$

Die Wahl von G' legt außerdem die Länge der Synthese- und Analysefilter fest. Denn beide Filterbänke lassen sich, ausgehend von (2.5) und (2.6), mit den zugeordneten Polyphasenmatrizen und dem Verzögerungsvektor

$$\mathbf{e}^T(z) = (1 \quad z^{-1} \quad \dots \quad z^{-(M-1)}) \quad (2.19)$$

ausdrücken durch:

$$\begin{pmatrix} H_0(z) \\ H_1(z) \\ \dots \\ H_{M-1}(z) \end{pmatrix} = \mathbf{H}^{(1)} z^M \mathbf{e}(z) \quad \text{und} \quad (2.20)$$

$$\begin{pmatrix} G_0(z) \\ G_1(z) \\ \dots \\ G_{M-1}(z) \end{pmatrix}^T = z^{-(M-1)} \mathbf{e}^T(z^{-1}) \mathbf{G}^{(2)} z^M \quad (2.21)$$

und bilden die paraunitäre Analyse- und Synthesefilterbank. Durch Einsetzen von (2.16) in (2.21) und mit (2.20) folgt:

$$\begin{pmatrix} G_0(z) \\ G_1(z) \\ \dots \\ G_{M-1}(z) \end{pmatrix} = z^{-((G'+1)M-1)} \begin{pmatrix} H_0(z^{-1}) \\ H_1(z^{-1}) \\ \dots \\ H_{M-1}(z^{-1}) \end{pmatrix}$$

und durch Rücktransformation in den Zeitbereich die Darstellung in (2.2) mit

$$K - 1 = (G' + 1)M - 1 = L', \quad (2.22)$$

wenn kausale Analysefilter vorausgesetzt werden. Für die Filterlänge K gilt daher

$$K = (G' + 1)M \quad (2.23)$$

bzw. für die in (2.3) eingeführte Größe $\tilde{G} = G' + 1$.

2.1.1.2 Filterentwurf paraunitärer Filterbänke

Die Elemente der Polyphasenmatrix $\mathbf{H}^{(1)}(z)$ seien kausale FIR-Filter. Dann besitzt auch die Determinante $\det \mathbf{H}^{(1)}(z)$ die FIR-Eigenschaft. Da dies aufgrund von (2.13) auch für die Determinante $\det \mathbf{G}^{(2)}(z)$ zutrifft, muß mit (2.14) gelten:

$$\det \mathbf{H}^{(1)}(z) = z^{-G}, \quad G \in \mathbb{N}. \quad (2.24)$$

Die Größe G heißt Grad des Systems (McMillan-Grad) [Vai93] und entspricht der Anzahl der für die Implementierung benötigten Verzögerungselemente.

Die im folgenden eingeführten Matrizen besitzen die Größe (M, M) , wenn nichts anderes angegeben ist. Jedes kausale, reellwertige und paraunitäre System mit $N = M$ Kanälen vom Grad G kann durch folgende Faktorisierung beschrieben werden [Vai93]:

$$\mathbf{H}^{(1)}(z) = \left(\prod_{i=0}^{G-1} \mathbf{V}_i(z) \right) \mathbf{U}_M \quad (2.25)$$

mit $\mathbf{U}_M$ als orthogonaler Matrix und den Matrizen

$$\mathbf{V}_i(z) = \mathbf{I} + (z^{-1} - 1)\mathbf{v}_i\mathbf{v}_i^T.$$

Dabei sind die $\mathbf{v}_i$ reellwertige $(1, M)$ Spaltenvektoren der Norm 1, d.h. $\mathbf{v}_i^T \mathbf{v}_i = \|\mathbf{v}_i\|^2 = 1$. Vektoren endlicher Dimension werden auch im folgenden durch Fettschrift gekennzeichnet. Der Signalfluß in dem Element $\mathbf{V}_i(z)$, das den Grad $G = 1$ besitzt, ist in Bild 2.5 dargestellt.

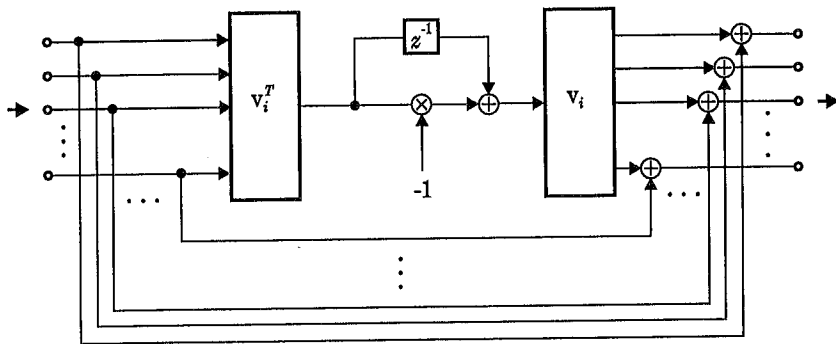


Bild 2.5: Darstellung des allgemeinen paraunitären Systems $\mathbf{V}_i(z)$

Wegen (2.16) und $\mathbf{V}_i^T(z) = \mathbf{V}_i(z)$ folgt für die Polyphasenmatrix der Synthesefilterbank:

$$\mathbf{G}^{(2)}(z) = z^{-G'} \mathbf{U}_M^T \left(\prod_{i=0}^{G-1} \mathbf{V}_{G-1-i}(z^{-1}) \right). \quad (2.26)$$

Um kausale Synthesefilter wie in (2.2) zu erhalten, muß

$$G' = G \quad (2.27)$$

gewählt werden.

Die Faktorisierung ist vollständig und minimal, d.h. sie kommt mit einer minimalen Anzahl von Verzögerungsgliedern aus [Vai93].

Jeder der Vektoren $\mathbf{v}_i$ mit den Elementen $\{v_{i,l}\}_{l=0}^{M-1}$ besitzt durch die Normvorgabe $M - 1$ Freiheitsgrade und läßt sich durch die Winkel $\{\theta_{i,l}\}_{l=0}^{M-2}$ parametrisieren:

$$\begin{aligned}
v_{i,0} &= \cos \theta_{i,0}, \\
v_{i,l} &= \left(\prod_{k=0}^{l-1} \sin \theta_{i,k} \right) \cos \theta_{i,l}, \quad l = 1, 2, \dots, M-2, \\
v_{i,M-1} &= \prod_{k=0}^{M-2} \sin \theta_{i,k}.
\end{aligned}$$

Jede orthogonale (M, M) Matrix $\mathbf{U}_M$ kann durch folgende Rekursion faktorisiert werden [VK95, Vai93]:

$$\mathbf{U}_M = \left(\prod_{k=1}^{M-1} \mathbf{R}_{M,M-k} \right) \begin{pmatrix} \mathbf{U}_{M-1} & \mathbf{0}_{M-1} \\ (\mathbf{0}_{M-1})^T & \mu_M \end{pmatrix}. \quad (2.28)$$

Dabei sind $\mu_l = \pm 1$ mit $l = 1, 2, \dots, M$, $\mathbf{0}_{M-1}$ ein $(1, M-1)$ Nullvektor und $\mathbf{R}_{M,M-k}$ eine (M, M) Matrix, die zwischen den Zeilen $M-k$ und M eine Givens-Rotation [Vai93] durchföhrt. Die Rekursion endet mit der $(1, 1)$ Matrix $\mathbf{U}_1 = \mu_1$.

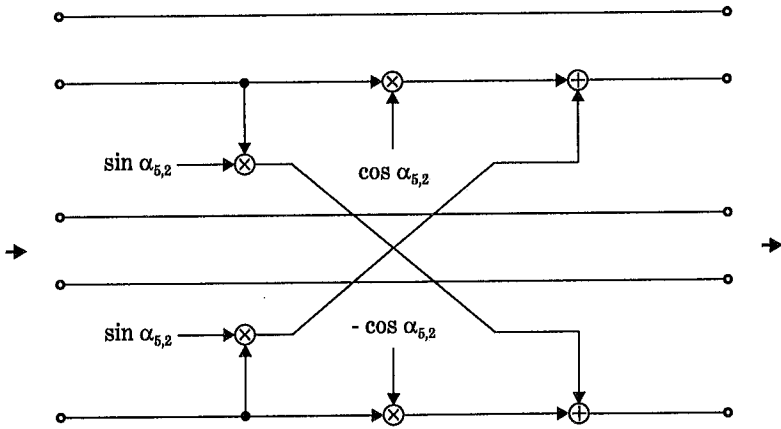
Als Beispiel sei die Matrix $\mathbf{R}_{5,2}$, im Signalfußdiagramm in Bild 2.6 dargestellt, angegeben:

$$\mathbf{R}_{5,2} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos \alpha_{5,2} & 0 & 0 & \sin \alpha_{5,2} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & \sin \alpha_{5,2} & 0 & 0 & -\cos \alpha_{5,2} \end{pmatrix}.$$

Zur eindeutigen Faktorisierung der Matrix $\mathbf{U}_M$ nach (2.28) werden $(M-1) + (M-2) + \dots + 1 = M(M-1)/2$ Parameter benötigt.

Die Faktorisierung läßt sich geometrisch so deuten:

Ausgehend von der (M, M) Einheitsmatrix, lassen sich die Zeilen i mit $i = 1, 2, \dots, M$ als Basisvektoren i eines kartesischen Koordinatensystems in einem M -dimensionalen Raum ansehen. Man kann nun die Vektoren l mit $l = 1, 2, \dots, M-1$ und M in der zwischen den Vektoren l und M aufgespannten Ebene rotieren, ohne die Orthogonalität der Basis aufzugeben. Durch die Multiplikation mit ± 1 wird die Orientierung des Vektors M ggf. umgedreht.

Bild 2.6: Signalflußdarstellung der Matrix $\mathbf{R}_{5,2}$

Anschließend wird der Raum, der durch die Vektoren $1, 2, \dots, M-1$ aufgespannt wird, betrachtet und der Vorgang für Vektor $M-1$ wiederholt. Der Prozeß endet, wenn der Raum nur eine Dimension umfaßt und lediglich die Orientierung des verbliebenen Basisvektors gewählt werden kann.

Da die Vorzeichenwahl der Elemente μ_k nicht als Freiheitsgrad gezählt wird, ergeben sich für ein System der Ordnung G insgesamt $f_M^G = (M-1)G + M(M-1)/2$ freie Parameter. Für eine zweikanalige Filterbank gilt

$$f_2^G = G + 1. \quad (2.29)$$

Da jedes Filter der Analysefilterbank für $M=2$ wegen (2.23) gerade die Länge $2(G+1)$ besitzt, wird die Hälfte der Freiheitsgrade beim Filterentwurf durch die Bedingung der Paraunitarität festgelegt.

Die Analysefilter $H_n(z)$ werden nun so entworfen, daß sie für $H_n(z)|_{z=e^{j\omega}}$ eine Filterbank wie in Bild 2.7 bilden. In Bild 2.7 sind auch die Stoppbandintervalle $B_n(\varepsilon)$ für jedes Filter $H_n(z)$ eingetragen:

$$\begin{aligned} B_0(\varepsilon) &= \left[\frac{\pi}{M} + \varepsilon, \pi \right] \\ B_n(\varepsilon) &= \left[0, \frac{n}{M}\pi - \varepsilon \right] \cup \left[\frac{n+1}{M}\pi + \varepsilon, \pi \right], \quad n = 1, 2, \dots, M-2 \\ B_{M-1}(\varepsilon) &= \left[0, \frac{M-1}{M}\pi - \varepsilon \right]. \end{aligned}$$

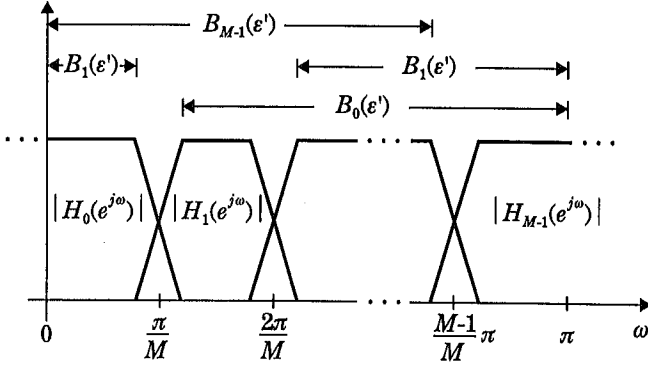


Bild 2.7: Gewünschte Amplitudengänge mit eingetragenen Stoppbandintervallen $B_n(\varepsilon)$ für gegebenes $\varepsilon = \varepsilon'$. Die Übergangsbereiche zwischen Durchlaß- und Sperrbereich deuten den im Entwurf erreichbaren Verlauf an.

Die Entwurfsaufgabe ist ein nichtlineares Optimierungsproblem. Für den Filterentwurf wird der $(1, f_M^G)$ Vektor $\mathbf{o}$ gebildet, der alle freien Parameter enthält und die Analysefilter $H_k(z)$ durch (2.5) eindeutig festlegt. Mit Hilfe des Verfahrens von Nelder-Mead [NM65], das das zum Entwurf verwendete Programm *Matlab* bereitstellt, sowie dem Vektor $\mathbf{o}$ als zu optimierenden Parametersatz und der zu minimierenden Funktion

$$K(\mathbf{o}, \varepsilon) = \sum_{n=0}^{M-1} \int_{\omega \in B_n(\varepsilon)} |H_n(e^{j\omega})|^2 d\omega \quad (2.30)$$

werden die Filter entworfen.

2.1.1.3 Deutung paraunitärer Filterbänke

Aus der Paraunitaritätseigenschaft (2.15) mit $c = 1$ folgt direkt die Darstellung:

$$\sum_{\lambda=0}^{M-1} H_{n,\lambda}^{(1)}(z) H_{k,\lambda}^{(1)}(z^{-1}) = \delta_{n,k}$$

und durch Rücktransformation in den Zeitbereich:

$$\sum_{l=0}^{K-1} h_n(l) h_k(l + mM) = \delta_{n,k} \delta_m, \quad n, k = 0, 1, \dots, M-1, m \in \mathbb{Z}. \quad (2.31)$$

Aus (2.31) folgt mit $z = e^{j\omega}$ unter Nutzung von (6.10) im Anhang 6.2.5 die Eigenschaft:

$$\sum_{l=0}^{M-1} |H_n(e^{j(\omega + \frac{2\pi}{M}l)})|^2 = M, \quad n = 0, 1, 2, \dots, M-1. \quad (2.32)$$

Für den Fall $M = 2$ kann man außerdem zeigen, daß die Filter $H_0(z)$ und $H_1(z)$ durch

$$\begin{aligned} H_1(z) &= \mu z^{-(K-1)} H_0(-z^{-1}), \quad \mu = \pm 1, \text{ bzw.} \\ \{h_1(k)\}_{k=0}^{K-1} &= \{(-1)^{k_0-k} h_0(K-1-k)\}_{k=0}^{K-1}, \quad k_0 \in \mathbb{Z} \end{aligned}$$

in Beziehung stehen (Anhang 6.3) und daher auf dem Einheitskreis leistungskomplementäre Amplitudengänge haben:

$$|H_0(e^{j\omega})|^2 + |H_1(e^{j\omega})|^2 = 2. \quad (2.33)$$

(2.31) erlaubt außerdem die Interpretation der paraunitären Filterbank als Entwicklung der Eingangsfolge nach einem orthonormalen Basissystem im Raum $l^2(\mathbb{Z})$, dessen Basisvektoren $\bar{b}_{n,mM}$ als unendliche Folge mit K reellwertigen, von Null verschiedenen Elementen beschrieben werden (vergl. Anhang 6.2.5):

$$\bar{b}_{n,mM} = \{h_n(mM - k)\}_k, \quad m \in \mathbb{Z}.$$

Für die Entwicklungskoeffizienten $\alpha_{n,m}$ des Vektors $\bar{s} = \{s(k)\}_k$ bezüglich des Basissystems gilt dann:

$$\begin{aligned} \alpha_{n,m} &= \langle \bar{s}, \bar{b}_{n,mM} \rangle \\ &= \sum_{k=-\infty}^{\infty} s(k) h_n(mM - k) \stackrel{!}{=} y_n(m). \end{aligned}$$

Wegen dieser Interpretation wird eine paraunitäre Filterbank im folgenden auch orthogonale Filterbank genannt. Durch die Orthogonalität der Filterbank gilt für die Normen der Vektoren $\bar{y}_n = \{y_n(m)\}_m$:

$$\|\bar{s}\|^2 = \sum_{n=0}^{M-1} \|\bar{y}_n\|^2.$$

2.1.2 Wavelet-Transformation

In [GM84] wurde die Wavelet-Transformation (WT) im Rahmen der Signalanalyse für die Untersuchung seismischer Schwingungen vorgestellt. Für das Signal $s(t)$ ergibt sich durch:

$$WT(a, b) = \int_{-\infty}^{\infty} s(t) \tilde{\psi}_{a,b}^*(t) dt = \langle s(t), \tilde{\psi}_{a,b}(t) \rangle, \quad a, b \in \mathbb{R}, a \neq 0 \quad (2.34)$$

die Transformationsvorschrift mit

$$\tilde{\psi}_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad (2.35)$$

als Wavelet-Funktion, die durch das Mutterwavelet $\psi(t)$ sowie den Skalierungsparameter a und die Verschiebung b beschrieben wird. Der Faktor $1/\sqrt{|a|}$ sichert die Normierungseigenschaft $\|\tilde{\psi}_{a,b}(t)\| = \|\psi(t)\|$.

Damit die inverse Wavelet-Transformation

$$s(t) = C_{\psi}^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} WT(a, b) \tilde{\psi}_{a,b}(t) \frac{da db}{a^2}$$

existiert, muß die Zulässigkeitsbedingung

$$C_{\psi} = \int_{-\infty}^{\infty} \frac{|\Psi(j\omega)|^2}{|\omega|} d\omega < \infty \quad (2.36)$$

mit der Fourier-Transformierten $\Psi(j\omega) \bullet \circ \psi(t)$ erfüllt sein [Chu92, Dau92].

In der Praxis verwendet man nur Funktionen $\psi(t)$ endlicher Dauer, d.h. es ist $\psi(t) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Wegen $\psi(t) \in L^1(\mathbb{R})$ ist $\Psi(j\omega)$ stetig [Chu92] und das Integral (2.36) daher für $\tilde{\Psi}(0) = 0$ beschränkt.

Da die Fourier-Transformierte $\Psi(j\omega)$ dann auch die Eigenschaft $\Psi(j\omega) \rightarrow 0$ für $\omega \rightarrow \pm\infty$ besitzt, hat $\psi(t)$ Bandpaßcharakter. Ihrem zeitbegrenzten, oszillierenden Verlauf verdankt $\psi(t)$ den Namen Wavelet-Funktion.

Die Skalierung der Funktion $\tilde{\psi}_{a,b}(t)$ durch den Parameter a führt zu einer frequenzabhängigen Auflösung.

Für den in der Praxis wichtigen Fall

$$a_n = 2^n \quad \text{und} \quad b_{n,m} = a_n m = 2^n m, \quad n, m \in \mathbb{Z} \quad (2.37)$$

wird über (2.35) definiert:

$$\tilde{\psi}_{2^n, 2^n m}(t) = \psi_{n,m}(t) = 2^{-n/2} \psi(2^{-n}t - m). \quad (2.38)$$

In der Zeit-Frequenz-Ebene erhält man das in Abbildung 2.8 dargestellte symbolische Raster. Die Wavelet-Transformation heißt für diese Diskretisierung der Parameter a und b dyadische Wavelet-Transformation (DWT).

Für große n wird die Wavelet-Funktion $\psi_{n,m}(t)$ gedehnt bzw. ihre Fourier-Transformierte $\Psi_{n,m}(j\omega)$ gestaucht und umgekehrt.

In der Literatur werden mehrere Varianten der Wavelet-Transformation diskutiert:

- Die nichtorthogonale Wavelet-Transformation [CGT89] wurde bislang vor allem für die visuelle Signalanalyse eingesetzt. Ihr Einsatz in der Klassifikation zeigte bisher keine erkennbaren Vorteile, z.B. [BPJ92]. Daher wird sie in der vorliegenden Arbeit nicht betrachtet.
- Der Matching-Pursuit-Algorithmus [MZ93, MC97] ermittelt für das Eingangssignal aus einer Menge nichtorthogonaler Wavelet-Funktionen durch einen Suchalgorithmus eine signalangepaßte Zerlegung. Die fehlende Orthogonalität führt zu einer Instabilität der Zerlegung gegenüber kleinen Signaländerungen bzw. Signalverschiebungen. Obwohl

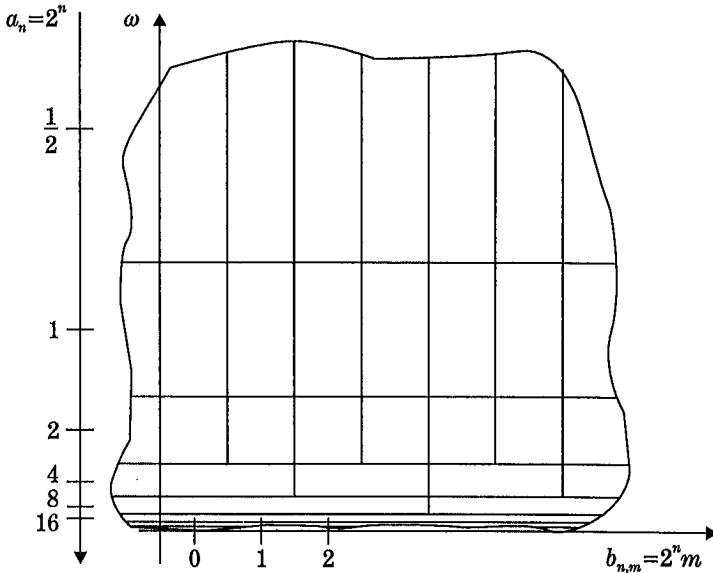


Bild 2.8: Ausschnitt aus der Zeit-Frequenz-Belegung der Wavelet-Transformation $WT(a_n, b_{n,m})$.

eine Variante des Algorithmus [Rod97] zur Charakterisierung von Klaviertönen eingesetzt wurde, verspricht die Anwendung in einer Klassifikationsaufgabe nur geringen Erfolg.

- Die orthogonale Wavelet-Transformation, die auf dem Konzept der Mehrfachauflösung beruht, wurde bislang in nur geringem Umfang für die Klassifikation realer Signale verwendet. Sie wird daher im folgenden betrachtet.

2.1.2.1 Signalanalyse durch Mehrfachauflösung

Erst durch das Konzept der Mehrfachauflösung [Mal89b] erlangte die Wavelet-Transformation praktische Bedeutung. Ausgangspunkt der Mehrfachauflösung ist die Zerlegung des Hilbertraums $L^2(\mathbb{R})$ in zueinander orthogonale, lineare, abgeschlossene Unterräume W_n mit $n \in \mathbb{Z}$ (vergl. Anhänge 6.1 und 6.1.4). Jeder der Unterräume W_n möge durch die orthogonalen Basisfunktionen $\{\psi_{n,m}(t)\}$ aufgespannt werden:

$$W_n = \text{span}_m \{\psi_{n,m}(t)\}. \quad (2.39)$$

Die aus der direkten Summe abgeleiteten Räume $\{V_n\}_{n \in \mathbb{Z}}$ mit

$$V_n = W_{n+1} \oplus W_{n+2} \oplus \dots = W_{n+1} \oplus V_{n+1} \quad (2.40)$$

sind ebenfalls abgeschlossen [Chu92, S. 120]. Das Symbol $\oplus$ in (2.40) drückt die direkte Summe aus. Die Räume $\{V_n\}_{n \in \mathbb{Z}}$ bilden eine Folge von verschachtelten Unterräumen:

$$\{0\} \dots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \dots \subset L^2(\mathbb{R}), \quad (2.41)$$

deren Vereinigung dicht in $L^2(\mathbb{R})$ und deren Schnittmenge der Nullraum $\{0\}$ ist:

$$\bigcap_{n \in \mathbb{Z}} V_n = \{0\}, \quad \bigcup_{n \in \mathbb{Z}} V_n = L^2(\mathbb{R}). \quad (2.42)$$

Die Räume $\{V_n\}_{n \in \mathbb{Z}}$ mögen die Eigenschaft haben:

$$s(t) \in V_n \Leftrightarrow s(2t) \in V_{n-1}. \quad (2.43)$$

Weiterhin möge eine Skalierungsfunktion $\phi(t)$ existieren, die mit ihren verschobenen Versionen $\{\phi(t-m)\}_{m \in \mathbb{Z}}$ eine orthonormale Basis des Raumes V_0 aufspannt:

$$V_0 = \text{span}_m \{\phi(t-m)\}. \quad (2.44)$$

Aus (2.41) und (2.43) folgt, daß die skalierten und verschobenen Skalierungsfunktionen

$$\phi_{n,m}(t) = 2^{-n/2} \phi(2^{-n}t - m) \quad (2.45)$$

ebenso eine orthonormale Basis für die Räume $\{V_n\}_{n \in \mathbb{Z}}$ darstellen:

$$V_n = \text{span}_m \{\phi_{n,m}(t)\}. \quad (2.46)$$

Aus (2.44) folgt auch:

$$s(t) \in V_0 \iff s(t - m) \in V_0, \quad m \in \mathbb{Z}$$

bzw. allgemein wegen (2.45):

$$s(t) \in V_n \iff s(t - 2^n m) \in V_n \quad m \in \mathbb{Z}. \quad (2.47)$$

Die Eigenschaften (2.40), (2.41), (2.42), (2.43), (2.44) und (2.47) seien bei jeder Signalanalyse durch Mehrfachauflösung erfüllt. In diesem Fall bilden die Funktionen

$$\psi_{n,m}(t) = 2^{-n/2} \psi(2^{-n}t - m), \quad n, m \in \mathbb{Z} \quad (2.48)$$

eine orthonormale Basis des $L^2(\mathbb{R})$ [VK95].

Zur Bestimmung einer zulässigen Skalierungsfunktion kann man ausnutzen [Dau88], daß sich die Basis des Raumes V_0 durch die Basisfunktionen des Raumes V_{-1} darstellen läßt:

$$\phi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} c(k) \phi(2t - k)$$

bzw. mit (2.47):

$$\phi_{n,m}(t) = \sum_{k=-\infty}^{\infty} c(k) \phi_{n-1,k}(t - 2^n m). \quad (2.49)$$

Ebenso lassen sich die Basisfunktionen des Raumes W_n durch eine Linearkombination von Basisfunktionen des Raumes V_{n-1} darstellen:

$$\psi_{n,m}(t) = \sum_{k=-\infty}^{\infty} d(k) \phi_{n-1,k}(t - 2^n m). \quad (2.50)$$

Für die Praxis sind Wavelet-Funktionen endlicher Länge für eine einfache Implementierung der Wavelet-Transformation von besonderem Interesse. In diesem Fall ist nur eine endliche Anzahl von Koeffizienten $\{c(k)\}_{k \in \mathbb{Z}}$ bzw. $\{d(k)\}_{k \in \mathbb{Z}}$ von Null verschieden [Ste92]. Aus (2.49) sowie (2.50) folgen mit einem kausalen Ansatz für die Koeffizientenfolgen:

$$\phi_{n,m}(t) = \sum_{k=0}^{K-1} c(k) \phi_{n-1,k}(t - 2^n m), \quad (2.51)$$

$$\psi_{n,m}(t) = \sum_{k=0}^{K-1} d(k) \phi_{n-1,k}(t - 2^n m). \quad (2.52)$$

Durch die Koeffizienten $\{c(k)\}_{k=0}^{K-1}$ und $\{d(k)\}_{k=0}^{K-1}$ werden die Wavelet- und Skalierungsfunktionen vollständig beschrieben. Die Bestimmung der Koeffizientensätze sind Gegenstand des Abschnitts 2.1.2.4.

2.1.2.2 Implementierung

Ist nun die Projektion $\{\gamma_n(m)\}$ des Signals $s(t)$ auf die Basisfunktionen des Raumes V_n bekannt, d.h.:

$$\gamma_n(m) = \langle s(t), \phi_{n,m}(t) \rangle, \quad m \in \mathbb{Z},$$

lassen sich die Projektionen auf die Räume $\{W_{n'}\}_{n' > n}$ sukzessive mit (2.51) und (2.52) bestimmen.

Zunächst wird gezeigt, wie die Entwicklungskoeffizienten $\{\gamma_{n+1}(m)\}$ des Signals $s(t)$ bezüglich der Basisvektoren des Raumes V_{n+1} aus den Koeffizienten $\{\gamma_n(m)\}$ folgen:

$$\begin{aligned} \gamma_{n+1}(m) &= \langle s(t), \phi_{n+1,m}(t) \rangle \\ &= \langle s(t), \sum_{k=0}^{K-1} c(k) \phi_{n,k}(t - 2^{n+1} m) \rangle \\ &= \langle s(t), \sum_{k=0}^{K-1} c(k) \phi_{n,2m+k}(t) \rangle \\ &= \sum_{k=0}^{K-1} c(k) \langle s(t), \phi_{n,2m+k}(t) \rangle \\ &= \sum_{k=0}^{K-1} c(k) \gamma_n(2m + k). \end{aligned}$$

Mit $\{h_0(k) = c(K-1-k)\}_{k=0}^{K-1}$ folgt schließlich:

$$\gamma_{n+1}(m) = \sum_{k=0}^{K-1} h_0(k) \gamma_n(2m + K - 1 - k). \quad (2.53)$$

Ebenso folgt für die Projektion auf die Basis des Raumes W_{n+1} mit $\{h_1(k) = d(K-1-k)\}_{k=0}^{K-1}$:

$$\begin{aligned} WT(n+1, m) &= \langle s(t), \psi_{n+1, m}(t) \rangle \\ &= \sum_{k=0}^{K-1} h_1(k) \gamma_n(2m + K - 1 - k). \end{aligned} \quad (2.54)$$

Die Filter $\{h_0(k)\}$ und $\{h_1(k)\}$ in Verbindung mit einer Dezimation um den Faktor 2 dienen als Orthogonalprojektoren bzw. ideale Filter (Anhang 6.1.4). Die Implementierung einer Zerlegungsstufe als Filterbank ist in Bild 2.9 dargestellt.

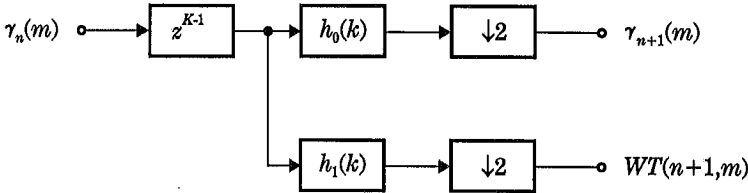


Bild 2.9: Implementierung der Orthogonalprojektoren mit der akausalen Verschiebung der Eingangsfolge um $K-1$ Elemente

Man beachte, daß bei Wahl des akausalen Ansatzes

$$\begin{aligned} \phi_{n, m}(t) &= \sum_{k=0}^{K-1} c(k) \phi_{n-1, -k}(t - 2^n m) \text{ und} \\ \psi_{n, m}(t) &= \sum_{k=0}^{K-1} d(k) \phi_{n-1, -k}(t - 2^n m) \end{aligned}$$

statt (2.51) bzw. (2.52) das Verschiebungselement in Bild 2.9 entfällt und die Koeffizienten $\{c(k)\}_{k=0}^{K-1}$ bzw. $\{d(k)\}_{k=0}^{K-1}$ unmittelbar als Filter eingesetzt werden können.

Die wiederholte Anwendung der Projektion auf die Entwicklungskoeffizienten bezüglich der Basis des Raumes $V_{n'}$ zur Bestimmung der Koeffizienten zur Basis des Raumes $V_{n'+1}$ erlaubt stets auch die Bestimmung der Wavelet-Transformierten zur Skale $a_{n'+1} = 2^{n'+1}$ durch Orthogonalprojektion.

Die Implementierung erfolgt nun unter folgenden Randbedingungen:

1. Die Eingangsfolge $\{s(m)\}_m$ wird als Projektion auf die Basis des Raumes $V_{n'}$ interpretiert. Die Zuordnung von $n' \in \mathbb{Z}$ zu einer bestimmten ganzen Zahl ist für die Signalanalyse und die Interpretation der Abbildung in der Zeit-Frequenz-Ebene ohne Bedeutung. In der Literatur hat sich der Wert $n' = 0$ durchgesetzt und soll daher im folgenden verwendet werden.
2. Es wird, ausgehend von der gegebenen Projektion in den Raum V_0 , eine endliche Zahl T von Projektionsschritten vorgenommen.
Die Zerlegung $V_0 = W_1 \oplus W_2 \dots \oplus W_T \oplus V_T$ beschreibt die Wavelet-Transformation mit einer Zerlegungstiefe T .
3. Der Zeitversatz in (2.53) und (2.54), der sich aus der Verwendung eines kausalen Filters für die Projektion ergibt und in Abbildung 2.9 durch ein Verschiebungsglied repräsentiert wird, ist für die Merkmalsextraktion ohne Bedeutung und wird weggelassen.

Es wird daher folgende Implementierung für die Erzeugung der Merkmale der DWT verwendet:

$$F^{DWT}(m, n-1) = \sum_{k=0}^{K-1} h_1(k) \gamma'_{n-1}(2m-k), \quad (2.55)$$

$$F^{DWT}(m, T) = \gamma'_T(m) \quad (2.56)$$

mit

$$\gamma'_{n+1}(m) = \sum_{k=0}^{K-1} h_0(k) \gamma'_n(2m-k),$$

$$\gamma'_0(m) = s(m)$$

für $n = 1, 2, \dots, T$ und $m \in \mathbb{Z}$.

Die zugehörige Filterbank ist in Bild 2.10 dargestellt.

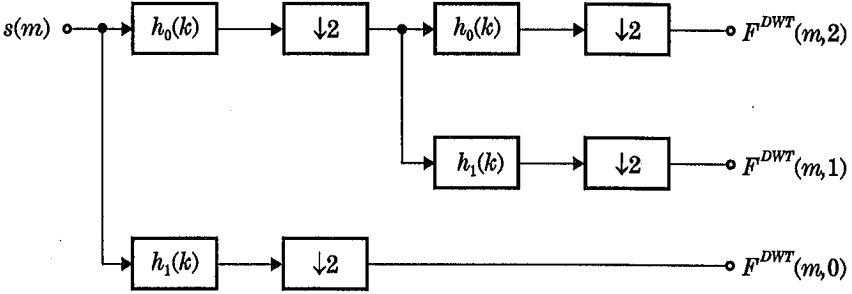


Bild 2.10: Implementierung der Wavelet-Transformation ohne Berücksichtigung akausaler Verschiebungen für die Zerlegungstiefe $T = 2$

2.1.2.3 Eigenschaften der Projektionsfilter

Die Forderung, daß die Skalierungsfunktionen $\{\phi_{n,m}(t)\}_m$ eine orthonormale Basis des Raumes V_n bilden (2.46), entspricht im Zeitbereich der Darstellung:

$$\langle \phi_{n,m}, \phi_{n,m+l} \rangle = \delta_l. \quad (2.57)$$

Ebenso gilt für die den Raum W_n aufspannenden Wavelet-Funktionen $\{\psi_{n,m}(t)\}_m$:

$$\langle \psi_{n,m}, \psi_{n,m+l} \rangle = \delta_l. \quad (2.58)$$

Schließlich folgt aus der Orthogonalität der Wavelet- und Skalierungsfunktionen gleicher Skale n :

$$\langle \phi_{n,m_1}, \psi_{n,m_2} \rangle = 0, \quad m_1, m_2 \in \mathbb{Z}. \quad (2.59)$$

Man kann durch eine Nebenrechnung zeigen, daß die Koeffizienten $\{c(k)\}_{k=0}^{K-1}$ und $\{d(k)\}_{k=0}^{K-1}$ die Eigenschaften

$$\sum_{k=0}^{K-1} c(k)c(k+2l) = \delta_l, \quad (2.60)$$

$$\sum_{k=0}^{K-1} d(k)d(k+2l) = \delta_l, \quad (2.61)$$

$$\sum_{k=0}^{K-1} c(k)d(k+2l) = 0, \quad l = 0, 1, 2, \dots, \frac{K}{2} - 1 \quad (2.62)$$

besitzen. Man beachte, daß die Koeffizienteneigenschaften (2.60, 2.61, 2.62) mit denen einer orthogonalen Filterbank übereinstimmen (2.31). Dieser Zusammenhang zu paraunitären Filterbänken dient in diesem Abschnitt zur Herleitung von Eigenschaften der Koeffizienten und wird für deren Bestimmung in Abschnitt 2.1.2.4 genutzt. Die Gleichungen (2.60, 2.61, 2.62) können nur erfüllt werden, wenn die Filterlänge K eine gerade Zahl ist. Dies folgt durch eine Nebenrechnung oder für $M = 2$ direkt aus (2.23).

Weiterhin folgt aus (2.51) durch Fourier-Transformation für die Skalierungsfunktion:

$$\Phi(j2^n\omega) = \frac{1}{\sqrt{2}} C(e^{j2^{n-1}\omega}) \Phi(j2^{n-1}\omega) \quad (2.63)$$

mit

$$C(e^{j\omega}) = \sum_{k=0}^{K-1} c(k)e^{-j\omega k}. \quad (2.64)$$

Aus (2.63) folgt durch I -faches iteratives Einsetzen:

$$\Phi(j2^n\omega) = \left(\prod_{i=0}^I \left(\frac{1}{\sqrt{2}} \right)^i C(e^{j2^{n-1-i}\omega}) \right) \Phi(j2^{n-1-I}\omega). \quad (2.65)$$

Für den Grenzübergang $I \rightarrow \infty$ folgt unter der Voraussetzung, daß das unendliche Produkt konvergiert:

$$\Phi(j2^n\omega) = \left(\prod_{i=0}^{\infty} \left(\frac{1}{\sqrt{2}} \right)^i C(e^{j2^{n-1-i}\omega}) \right) \Phi(0). \quad (2.66)$$

Es muß also $\Phi(0) \neq 0$ erfüllt sein, damit $\Phi(j\omega) \neq 0$ ist. Aus der Orthogonalitätseigenschaft (2.57) in Verbindung mit (2.65) läßt sich diese Eigenschaft präzisieren zu [Dau92]:

$$\Phi(0) = 1 \quad \text{bzw.} \quad \int_{-\infty}^{\infty} \phi(t) dt = 1. \quad (2.67)$$

Durch Einsetzen von (2.67) in (2.63) läßt sich sofort ablesen, daß

$$C(e^{j0}) = \sqrt{2} \quad \text{bzw.} \quad \sum_{k=0}^{K-1} c(k) = \sqrt{2} \quad (2.68)$$

gilt. Aus der Orthogonalitätsbedingung (2.60) läßt sich ebenfalls ableiten ([VK95] oder Abschnitt 2.1.1.3), daß für das Polynom $C(e^{j\omega})$ entsprechend zu (2.33) gilt:

$$|C(e^{j\omega})|^2 + |C(e^{j(\omega+\pi)})|^2 = 2. \quad (2.69)$$

Aus (2.69) und (2.68) folgt, daß $C(e^{j\pi}) = 0$, bzw.

$$C(e^{j\omega}) = (1 + e^{-j\omega})Q(e^{j\omega}) \quad (2.70)$$

mit $Q(e^{j\omega})$ als trigonometrischem Polynom gilt. Aus dieser Eigenschaft folgt durch Einsetzen in (2.64):

$$\sum_{k=0}^{K-1} (-1)^k c(k) = 0. \quad (2.71)$$

Man beachte, daß (2.71) und (2.70) keine neuen Bedingungen an die Folge $\{c(k)\}_{k=0}^{K-1}$ stellen, sondern sich über (2.68) aus der Orthogonalitätsbedingung (2.60) ergeben.

Über (2.62) kann gezeigt werden [VK95], daß die Folge $\{d(k)\}_{k=0}^{K-1}$ durch

$$D(e^{j\omega}) = \rho(e^{j\omega})C^*(e^{j(\omega+\pi)})$$

mit

$$D(e^{j\omega}) = \sum_{k=0}^{K-1} d(k)e^{-j\omega k} \quad (2.72)$$

und $\rho(e^{j\omega})$ als mit 2π periodischer Funktion von ω mit der Eigenschaft

$$\rho(e^{j\omega}) = -\rho(e^{j(\omega+\pi)})$$

mit der Folge $\{c(k)\}_{k=0}^{K-1}$ in Beziehung steht. Mit

$$\rho(e^{j\omega}) = -e^{-j(K-1)\omega}$$

folgt die Koeffizientenfolge:

$$d(k) = (-1)^k c(K-1-k), \quad k = 0, 1, 2, \dots, K-1. \quad (2.73)$$

Man beachte, daß mit dieser Wahl auch (2.61) erfüllt ist. In der Literatur wird mit der Ausnahme [Fli93] $\rho(e^{j\omega}) = -e^{-j\omega}$ und damit die Folge $\{d(k)\}_k$ zyklisch verschoben bzw. akausal gewählt [Dau88, VK95, Mer96].

Die Orthonormalität der Basisfunktionen (2.57, 2.58, 2.59) läßt sich also durch (2.60, 2.61, 2.62) sowie durch eine der Gleichungen (2.68, 2.70, 2.71) als Forderung an die Koeffizienten $\{c(k)\}_{k=0}^{K-1}$ und $\{d(k)\}_{k=0}^{K-1}$ formulieren. Man beachte, daß gegenüber paraunitären Filterbänken eine zusätzliche Bedingung erfüllt sein muß.

Die Konvergenz des unendlichen Produkts

$$\lim_{I \rightarrow \infty} \prod_{i=0}^I \left(\frac{1}{\sqrt{2}} \right)^i C(e^{j2^{n-1-i}\omega}) \quad (2.74)$$

läßt sich durch weitere Anforderungen an den Koeffizientensatz $\{c(k)\}_{k=0}^{K-1}$ sichern. Man kann zeigen, daß das unendliche Produkt existiert und punktweise konvergiert, wenn folgende hinreichenden Bedingungen erfüllt sind [Dau88, VK95]:

- Das Polynom $C(e^{j\omega})$ besitzt eine N' -fache Nullstelle bei $\omega = \pi$:

$$C(e^{j\omega}) = \sqrt{2} \left(\frac{1 + e^{-j\omega}}{2} \right)^{N'} P(e^{j\omega}), \quad N' \geq 1. \quad (2.75)$$

- Das Polynom $P(e^{j\omega})$ ist für $\omega \in [0, 2\pi]$ beschränkt:

$$P(e^{j\omega}) \leq 2^{N'-1}. \quad (2.76)$$

Es wird darauf hingewiesen, daß in der Literatur eine alternative hinreichende Bedingung in [Mal89a] sowie eine notwendige und hinreichende Bedingung in [LMR94] angegeben werden.

Die punktweise Konvergenz sichert allerdings nicht die Konvergenz bezüglich der Norm des $L^2(\mathbb{R})$. Für diese Frage und die Prüfung der Vollständigkeit der Waveletfunktionen (2.48) als Basis für den $L^2(\mathbb{R})$ wird auf die Literatur verwiesen [LMR94].

Durch mehrfaches Differenzieren von (2.64) nach ω erhält man wegen (2.71) N' weitere Bedingungen an die Folge $\{c(k)\}_{k=0}^{K-1}$:

$$\sum_{k=0}^{K-1} (-1)^k c(k) k^n = 0, \quad n = 0, 1, 2, \dots, N'. \quad (2.77)$$

Die Größe N' kann, wie es bei den Daubechies-Wavelets der Fall ist, wegen (2.60) maximal zu

$$N' = K/2 - 1 \quad (2.78)$$

gewählt werden und legt die Filter vollständig fest.

Man kann zeigen, daß die Bedingungen (2.60), (2.68) und (2.77) zur Bestimmung der Wavelets nach Daubechies für das Filter $\{h_0(m)\}_{m=0}^{K-1}$ ebenso gelten. In [Lai95] wird gezeigt, daß für $N' \rightarrow \infty$ das Filter $\{h_0(m)\}_{m=0}^{K-1}$ gegen einen idealen Halbbandtiefpas bzw. $\{h_1(m)\}_{m=0}^{K-1}$ gegen einen idealen Halbbandhochpas strebt.

Die Motivation hinter der Faktorisierung (2.75) besteht in dem Ziel, für die Waveletfunktion $\psi(t)$ mehrfach stetig differenzierbare Funktionen zu erhalten. Man kann zeigen, daß eine N' -fache Nullstelle des Polynoms $C(e^{j\omega})$ in π eine notwendige Bedingung für die Existenz $N' - 1$ stetiger Ableitungen von $\psi(t)$ ist [Dau88].

2.1.2.4 Entwurf der Projektionsfilter als paraunitäre Filterbank

Die Wavelet- und die Skalierungsfunktion werden durch die Folgen $\{c(k)\}_{k=0}^{K-1}$ und $\{d(k)\}_{k=0}^{K-1}$ beschrieben. Ihre z -Transformierten $C(z)$ und $D(z)$ stehen zu den Projektionsfiltern $H_0(z)$ und $H_1(z)$ durch

$$\begin{pmatrix} H_0(z) \\ H_1(z) \end{pmatrix} = z^{-(K-1)} \begin{pmatrix} C(z^{-1}) \\ D(z^{-1}) \end{pmatrix} \quad (2.79)$$

in Beziehung.

$C(z)$ bzw. $D(z)$ besitzen nach Abschnitt 2.1.2.3 die Eigenschaften paraunitärer Filterbänke, so daß die Entwurfsverfahren des Abschnitts 2.1.1.2 für

die Konstruktion der Projektionsfilter verwendet werden können. Die im folgenden aus Abschnitt 2.1.1.1 verwendeten Zusammenhänge werden stets für $M = 2$ betrachtet.

Die Orthogonalität der Wavelet- und Skalierungsfunktionen wird aber erst durch (2.70) gesichert. Durch analytische Fortsetzung erhält man aus (2.70)

$$C(z) = (1 + z^{-1})Q(z). \quad (2.80)$$

Mit dem Ansatz aus (2.20)

$$\begin{pmatrix} C(z) \\ D(z) \end{pmatrix} = \mathbf{H}^{(1)}(z^2) \begin{pmatrix} 1 \\ z^{-1} \end{pmatrix} \quad (2.81)$$

wird in [ZT93] untersucht, ob durch eine geeignete Parametrisierung von $\mathbf{H}^{(1)}(z)$ über die Faktorisierung (2.25):

$$\mathbf{H}^{(1)}(z) = \left(\prod_{i=0}^{G-1} \mathbf{V}_i(z) \right) \mathbf{U}_2 \quad (2.82)$$

die Eigenschaft (2.80) erfüllt werden kann.

Da $\mathbf{V}_i(z^2)|_{z=-1} = \mathbf{I}_2$, $i = 1, 2, \dots, G-1$, und $D(-1) = \sqrt{2}$, muß die Faktorisierung an der Stelle $z = -1$ die Bedingung

$$\begin{pmatrix} C(z) \\ D(z) \end{pmatrix} \Big|_{z=-1} = \begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix} \stackrel{!}{=} \mathbf{U}_2 \begin{pmatrix} 1 \\ z^{-1} \end{pmatrix} \Big|_{z=-1}$$

erfüllen. Dies ist der Fall für:

$$\mathbf{U}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Für den Entwurf von Projektionsfiltern ist es daher notwendig, eine Faktorisierung nach

$$\begin{pmatrix} C(z) \\ D(z) \end{pmatrix} = \frac{1}{\sqrt{2}} \left(\prod_{i=0}^{G-1} \mathbf{V}_i(z^2) \right) \begin{pmatrix} 1 + z^{-1} \\ 1 - z^{-1} \end{pmatrix} \quad (2.83)$$

zu verwenden. Die Freiheitsgrade der Vektoren $\mathbf{v}_i$ der Teilsysteme $\mathbf{V}_i(z)$ können genutzt werden, um die Amplitudengänge nach (2.30) zu optimieren.

Für den Entwurf der Projektionsfilter nach Daubechies muß das Polynom $C(z)$ wegen (2.75) eine N' -fache Nullstelle in $z = 1$ mit $N' \geq 1$ aufweisen. Für das Polynom $C(z^{-1})$ folgt der Ansatz:

$$C(z^{-1}) = \sqrt{2} \left(\frac{1+z}{2} \right)^{N'} Q(z^{-1}) \quad (2.84)$$

und gegenüber (2.83) N' neue Bedingungen:

$$\left(\begin{array}{c} \frac{d^m C(z^{-1})}{dz^m} \\ \frac{d^m D(z^{-1})}{dz^m} \end{array} \right) \bigg|_{z=-1} = \begin{pmatrix} 0 \\ q \end{pmatrix}, \quad q \in \mathbb{R}, \quad m = 1, 2, \dots, N'. \quad (2.85)$$

Wegen (2.29) kann $N' = G$ in Übereinstimmung mit (2.78) zu $G = K/2 - 1$ gewählt werden.

In [ZT93] wird gezeigt, wie aus diesen G Gleichungen (2.85) G Bedingungen an die Vektoren $\mathbf{v}_i$ abgeleitet werden können. Es wird darauf hingewiesen, daß alternativ die Filter nach Daubechies durch spektrale Faktorisierung eines Lagrange-Halbbandfilters bestimmt werden können [AGK91].

2.1.3 Allgemeine translationsinvariante Analysefilterbänke

Die Merkmalsextraktion durch Zeit-Frequenz-Analyse liefert für die Klassifikationsaufgabe geeignete Merkmale, wenn

- die Transformation an die Lokalisierung des Signals in der Zeit-Frequenz-Ebene angepaßt ist, um mit möglichst wenigen Merkmalen ein Signal zu erfassen, und
- die Merkmale gegenüber Verschiebungen des Eingangssignals robust sind.

In der Literatur werden Möglichkeiten beschrieben, die diskrete orthogonale Wavelet-Transformation so zu verallgemeinern, daß die Transformation diese Eigenschaften besitzt [LP96, PKLH96, CW92]. Diese Erweiterungen nutzen die Orthogonalität der Transformation und lassen sich daher ebenso für paraunitäre Filterbänke verwenden.

Der Grundbaustein einer orthogonalen Filterbank sei durch eine M -kanalige Filterbank wie in Bild 2.11 gegeben. Die Ein- und Ausgabegrößen werden in Abschnitt 2.1.3.1 erläutert. Auch in den folgenden Bildern wird für eine einfache und einheitliche Darstellung der z -Bereich gewählt.

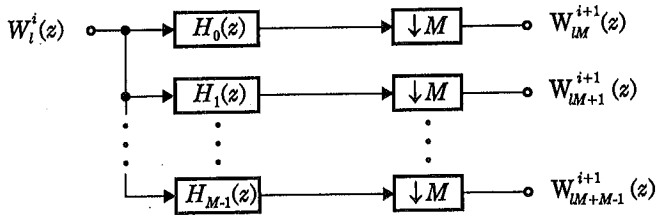


Bild 2.11: Grundbaustein einer M -kanaligen Filterbank mit Baumstruktur

Eine Anpassung der Frequenz-Auflösung lässt sich durch Kaskadierung von Grundbausteinen erreichen und führt auf Filterbänke mit Baumstruktur, die in Abschnitt 2.1.3.1 vorgestellt werden. In der Literatur werden auch zeitveränderliche Filterbänke mit Baumstruktur diskutiert, z.B. [XRHO97]. Wegen des deutlich höheren Rechenaufwands, der mit diesem Ansatz verbunden ist, wird auf eine Betrachtung verzichtet.

Durch die Dezimationsblöcke im Grundbaustein ist die Filterbank periodisch translationsinvariant [Fli93]. Völlige Translationsinvarianz wird durch taktverschoben arbeitende Parallelbausteine erreicht und ist Gegenstand von Abschnitt 2.1.3.2.

2.1.3.1 Filterbänke mit Baumstruktur

Jeder Ausgang des Grundbausteins kann wiederum Gegenstand einer Orthogonalzerlegung sein. Wird dieser Vorgang $(T - 1)$ -mal wiederholt, wird eine Filterbank mit Baumstruktur der Zerlegungstiefe T aufgebaut. Für eine M -kanalige Filterbank mit den Analysefiltern $\{h_b(k)\}_{k=0}^{K-1}$ mit $b = 0, 1, 2, \dots, M - 1$ werden die Ausgabefolgen $\{w_l^i(m)\}_m$ auf folgende Weise bestimmt:

1. Initialisierung:

$$w_0^0(m) = s(m),$$

2. Iterationsschritt i :

für $i = 0, 1, \dots, T - 1$

(2.86)

$$\begin{aligned} w_{Ml+b}^{i+1}(m) &= \sum_{k=0}^{K-1} h_b(k) w_l^i(Mm - k), \\ b &= 0, 1, 2 \dots M - 1, \\ l &= 0, 1, 2 \dots M^i - 1. \end{aligned}$$

Mit jeder Zerlegung wird die Anzahl der Kanäle um den Faktor M größer. In der Zerlegungsstufe i beträgt die Anzahl der Kanäle M^i . Der Parameter l indiziert den gewählten Kanal.

Über die Abbildungsvorschrift

$$n(i, l) = \begin{cases} 0 & \text{wenn } i = 0, l = 0 \\ \sum_{k=0}^{i-1} M^k + l & \text{wenn } i \geq 1 \end{cases} \quad (2.87)$$

können die Folgen $\{w_l^i(m)\}_m$ auf die Merkmale $F^B(m, n)$ abgebildet werden:

$$F^B(m, n(i, l)) = w_l^i(m). \quad (2.88)$$

Bild 2.12 zeigt ein Beispiel für den Fall einer Filterbank mit $M = 2$ Kanälen und einer Zerlegungstiefe $T = 2$ mit den Größen $W_l^i(z) \xrightarrow{z} \{w_l^i(m)\}_m$.

Umgekehrt läßt sich für ein gegebenes l und $i \geq 1$ die eindeutige Zerlegung

$$l = \sum_{\tilde{l}=0}^{i-1} b_{\tilde{l}} M^{\tilde{l}}, \quad b_{\tilde{l}} \in \{0, 1, 2, \dots, M - 1\}$$

angeben, wobei die Koeffizienten $b_{\tilde{l}}$ für $\tilde{l} = 0, 1, 2, \dots, i - 1$ die Abfolge der verwendeten Filter $\{h_{b_{\tilde{l}}}(k)\}_{k=0}^{K-1}$ beschreiben. Bild 2.13 verdeutlicht den Zusammenhang.

Mit Hilfe der in Anhang 6.2.3 angegebenen Äquivalenzen läßt sich der Signalfluß wie in Bild 2.14 umformen.

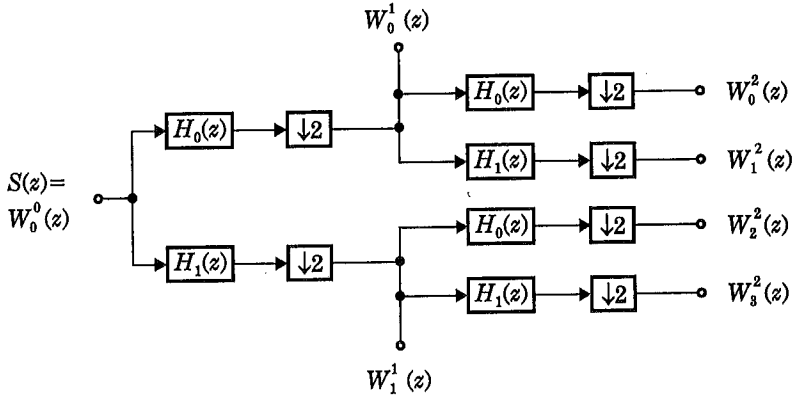


Bild 2.12: Filterbank mit Baumstruktur der Tiefe $T = 2$ mit $M = 2$ Kanälen

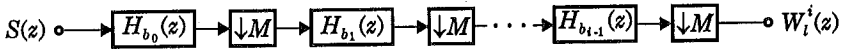


Bild 2.13: Signalfluß zwischen dem Eingangssignal $S(z)$ und dem Ausgang $W_i^i(z)$ der Filterbank

Die Impulsantwort des resultierenden Filters $H_{i,l}^{res}(z)$

$$H_{i,l}^{res}(z) = \prod_{i=0}^{i-1} H_{b_i}(z^{M^i}) \quad (2.89)$$

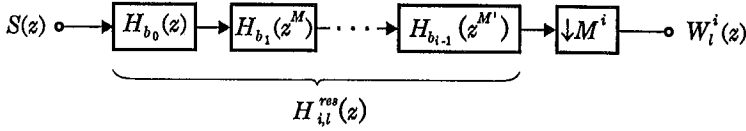
besitzt die rekursiv definierte Länge K_i :

$$\begin{aligned} K_1 &= K, \\ K_i &= (K_{i-1} - 1)M + K, \quad i > 1. \end{aligned} \quad (2.90)$$

Aus Bild 2.14 läßt sich für die Ausgangsfolgen $\{w_i^i(m)\}_m$ mit

$$H_{i,l}^{res}(z) \stackrel{z}{\bullet} \{h_{i,l}^{res}(k)\}_{k=0}^{K_i}$$

auch die Darstellung

Bild 2.14: Umgeformter Signalfluß aus Bild 2.13 mit $M' = M^{i-1}$

$$w_l^i(m) = \sum_{k=0}^{K_i-1} h_{i,l}^{res}(k) s(M^i m - k) \quad (2.91)$$

angeben.

Der Grundbaustein vermittelt nach Abschnitt 2.1.1.3 eine Abbildung auf ein System orthonormaler Basisfolgen. Eine Filterbank mit Baumstruktur stellt eine Abbildung nach einer Menge von Folgenbasen dar. Filterbänke mit Baumstruktur können daher als Orthogonalzerlegungen des Raumes $l^2(\mathbb{Z})$ interpretiert werden. Die Folgen $\{w_l^i(m)\}_m$ lassen sich nach Abschnitt 2.1.1.3 als Entwicklungskoeffizienten des Vektors $\bar{s}$ bezüglich der Basisvektoren

$$\bar{b}_{l,mM}^i = \{h_{i,l}^{res}(mM^i - k)\}_k$$

deuten, da aus

$$w_l^i(m) = \langle \bar{s}, \bar{b}_{l,mM}^i \rangle \quad (2.92)$$

unmittelbar (2.91) folgt. Wenn der Raum, den die Vektoren $\{\bar{b}_{l,mM}^i\}_m$ aufspannen, mit R_l^i bezeichnet wird, führt der Grundbaustein die Zerlegung

$$R_l^i = R_{lM+1}^{i+1} \oplus R_{lM+1}^{i+1} \oplus \dots \oplus R_{lM+M-1}^{i+1} \quad (2.93)$$

durch. Dabei ist der Raum $R_0^0 = l^2(\mathbb{Z})$.

Der Zusammenhang sei graphisch veranschaulicht. Repräsentiert man den Grundbaustein durch ein graphisches Element wie in Bild 2.15, läßt sich die Filterbank für den Fall $M = 2$ und der Zerlegungstiefe $T = 3$ auf die in Bild 2.16 abgebildete Form bringen.

Jeder Knoten repräsentiert den Folgenraum R_l^i , bezüglich dessen Basisvektoren die Entwicklungskoeffizienten nach (2.92) als Filterausgabe bestimmt

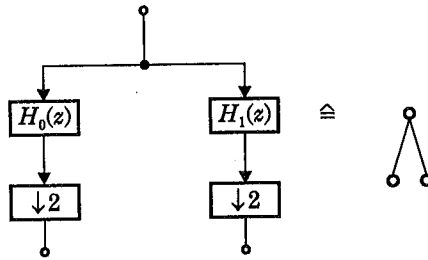


Bild 2.15: Repräsentation des Grundbausteins durch einen Graphen für $M = 2$ Kanäle

werden. Die Wahl einer geeigneten Basis bzw. Repräsentation der Eingangsfolge wird auf den Schritt der Merkmalsselektion verlagert.

Die Auswahl einer Folgenbasis erfolgt in dem Graphen durch Knotenmarkierung. Die Markierung eines Knotens impliziert, daß alle Nachfolgerknoten nicht markiert sind. Bild 2.17 (a) stellt die Wahl entsprechend der diskreten Wavelet-Transformation dar, Bild 2.17 (b) liefert ein weiteres Beispiel.

Für den Fall, daß in der Filterbank bei $M = 2$ die in Abschnitt 2.1.2.1 vorgestellten Filter verwendet werden, spricht man von einer dyadischen Wavelet-Packet-Transformation (DWPT) [Chu92, Dau92].

2.1.3.2 Translationsinvariante Filterbänke

Die Abbildung A :

$$s(m) \circ \overset{A}{\longrightarrow} \odot F(m, n), \quad n, m \in \mathbb{Z}$$

heißt translationsinvariant, wenn die Verschiebung der Eingangsfolge um m_0 einer gleich großen Verschiebung im Merkmalsbereich entspricht:

$$s(m + m_0) \circ \overset{A}{\longrightarrow} \odot F(m + m_0, n), \quad m_0 \in \mathbb{Z}. \quad (2.94)$$

Für orthogonale Abbildungen spielt Translationsinvarianz eine große Rolle. Definiert man die Autokorrelationsfolge (AKF) $\{r_{ss}(m_0)\}_{m_0}$ reellwertiger Folgen $\{s(m)\}_m$ zu [Mer96]

$$r_{ss}(m_0) = \sum_{k=-\infty}^{\infty} s(k)s(k + m_0)$$

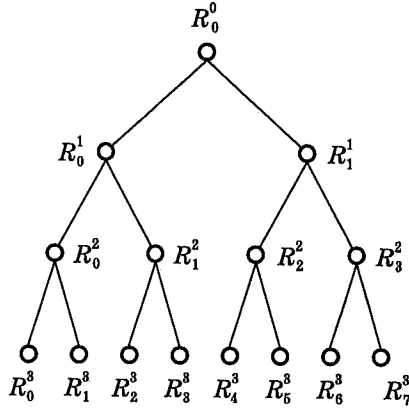


Bild 2.16: Repräsentation einer Filterbank mit Baumstruktur durch einen Graphen für $M = 2$ mit der Zerlegungstiefe $T = 3$

und ist $F^0(m, n)$ die orthogonale Abbildung für das unverschobene Signal sowie $F^{m_0}(m, n)$ die des verschobenen Signals, gilt:

$$\frac{1}{2} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} |F^0(m, n) - F^{m_0}(m, n)|^2 = \left[\sum_{m=-\infty}^{\infty} s^2(m) \right] - r_{ss}(m_0). \quad (2.95)$$

Dies folgt aus folgender Überlegung: Der Summenterm auf der linken Seite ist wegen der Orthogonalität der Abbildung A bis auf den Faktor $\frac{1}{2}$ die Norm des Differenzsignals $\{\Delta s(m)\}_m = \{s(m) - s(m + m_0)\}_m$. Die rechte Seite ergibt sich über

$$\sum_{m=-\infty}^{\infty} \Delta s^2(m) = \sum_{m=-\infty}^{\infty} [s^2(m) + s^2(m + m_0) - 2s(m)s(m + m_0)].$$

Fällt die AKF gegenüber ihrem Maximum bei $m_0 = 0$ rasch ab, wächst die Differenz in (2.95) in gleichem Maß. Die Abbildung eines verschobenen Signals unterscheidet sich dann sehr von seiner unverschobenen Version.

Lediglich für die Wavelet-Transformation ist die Abhängigkeit des Transformationsergebnisses von der Verschiebung des Eingangssignals bereits diskutiert worden [LP96], obgleich es eine gemeinsame Eigenschaft aller ortho-

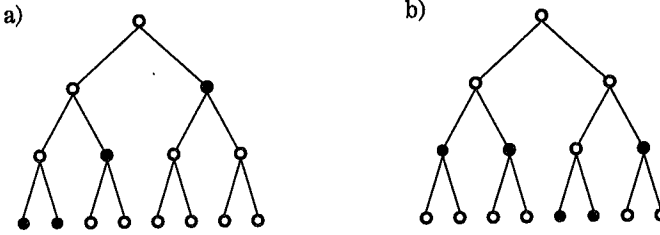


Bild 2.17: Unterraumzerlegung bei (a) einer Wavelet-Transformation und (b) einer anderen möglichen Basis

gonaler Abbildungen ist. Für die Klassifikation ist diese Eigenschaft offensichtlich von Nachteil, soll doch das Signal und nicht dessen Verschiebung erkannt werden. Abhilfe ist möglich, wenn die Abbildung translationsinvariant ist und die Verschiebung im Merkmalsbereich rückgängig gemacht werden kann. Auf diese Weise läßt sich die gewünschte Robustheit erzielen. In [DJ97b] wurde erstmals eine translationsinvariante Filterbank mit Baumstruktur für die Klassifikation eingesetzt.

Eine Filterbank mit Baumstruktur mit dem Grundbaustein wie in Bild 2.11 kann durch die Verwendung der Dezimatoren nicht translationsinvariant im Sinne von (2.94) sein. Da das Weglassen der Dezimatoren die Aufgabe der Paraunitarität und damit der Orthogonalität des Grundbausteins bedeutet, muß der Begriff der Translationsinvarianz erweitert werden, um die Verschiebung im Merkmalsbereich dennoch zu ermöglichen.

Eine Abbildung A durch eine Filterbank mit Baumstruktur soll im folgenden translationsinvariant bezüglich einer Menge von Verschiebungen $\mathbb{W}$ heißen, wenn ein Algorithmus existiert, der zu einer gegebenen Verschiebung $m_0 \in \mathbb{W}$ eine Abbildung $A(m_0)$ auf eindeutige Weise bestimmt, so daß gilt:

$$s(m + m_0) \circ \xrightarrow{A(m_0)} \odot = \tilde{F}_{m_0}^B(m, n), \quad (2.96)$$

mit $m_0 \in \mathbb{W}$. Dabei ist $\tilde{F}_{m_0}^B(m, n)$ die durch die Abbildung $A(m_0)$ vermittelte Merkmalsrepräsentation mit der Eigenschaft

$$\tilde{F}_{m_0}^B(m - m'(n, m_0), n) = F_{m_0}^B(m, n) = F_0^B(m, n), \quad (2.97)$$

wobei $F_0^B(n, m)$ die Merkmale des unverschobenen Eingangssignals und $F_{m_0}^B(n, m)$ die Merkmale darstellen, die bei Ausgleich der Verschiebung m_0 des Eingangssignals erzeugt werden. Das Transformationsergebnis $\tilde{F}_{m_0}^B(m, n)$

aus (2.96) muß einer abschließenden Verschiebung unterzogen werden, um das Ergebnis des unverschobenen Signals $F_{m_0}^B(m, n)$ zu erhalten (2.97).

Man beachte, daß die Verschiebung $m'(n, m_0)$ nicht nur von dem Parameter m_0 , sondern auch von dem Parameter n abhängen darf.

Im folgenden wird zunächst untersucht, wie bei gegebener Verschiebung $m_0 \in \mathbb{W}$ aus dem Grundbaustein geeignete Abbildungen abgeleitet werden können. Anschließend wird der dabei gefundene Ansatz auf eine Filterbank mit Baumstruktur übertragen. Den Abschluß bildet ein Algorithmus, der die Baumstruktur so auslegt, daß für alle Verschiebungen aus $\mathbb{W}$ die notwendigen Abbildungen erzeugt werden können und Bedingung (2.97) erfüllt wird.

Das Signal $\{s(m)\}_m$ sei um m_0 entsprechend (2.96) verschoben. Dann muß die Abbildung eine Rückverschiebung nach (2.97) ermöglichen und führt zu dem Ansatz $\{s(m + m_0)\}_m \xrightarrow{z} S(z)z^{m_0} = W_0^0(z)z^{m_0}$, wie in Bild 2.18 (a) dargestellt.

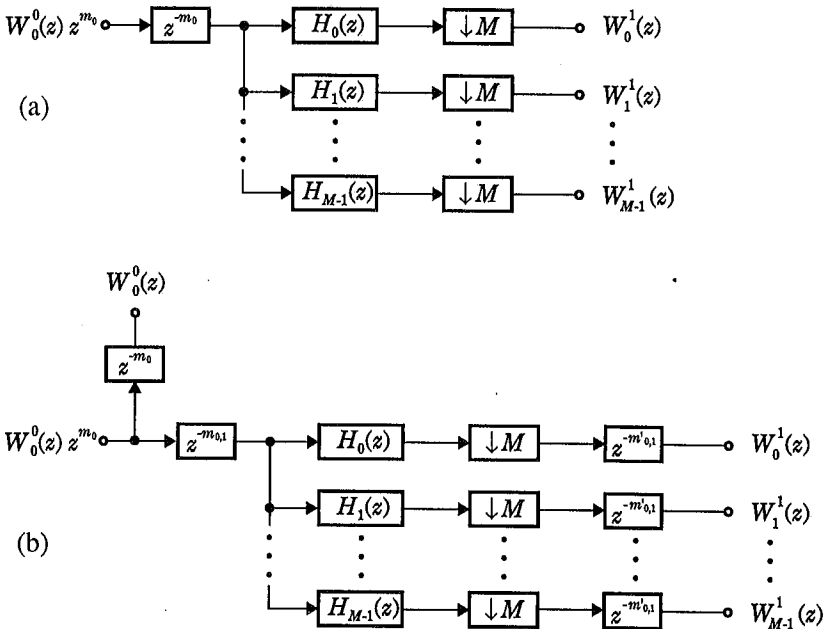


Bild 2.18: Transformation einer um m_0 verschobenen Eingangsfolge durch eine Filterbank mit Baumstruktur (a) vor und (b) nach der Umformung

Durch die Äquivalenzen nach Anhang 6.2.3 läßt sich der Ansatz in die Form in Bild 2.18 (b) überführen, wobei mit mod als modulo-Operation gilt:

$$\begin{aligned} m_{0,1} &= m_0 \bmod M, \\ m'_{0,1} &= (m_0 - m_{0,1})/M. \end{aligned} \quad (2.98)$$

Die Verzögerung m_0 wird auf zwei Verzögerungsblöcke verteilt, so daß gilt

$$m_0 = m_{0,1} + Mm'_{0,1}, \quad m_{0,1} \in \{0, 1, 2, \dots, M-1\}, \quad m'_{0,1} \in \mathbb{Z}. \quad (2.99)$$

Soll die Abbildung aus 2.18 (b) zu einer Filterbank mit Baumstruktur erweitert werden, führt dies mit der in Bild 2.19 vereinbarten Kurzform für den Fall $M = 2$ auf die Form in Bild 2.20.

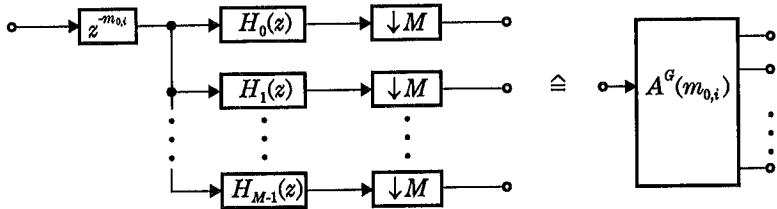


Bild 2.19: Grundbaustein einer Filterbank mit Baumstruktur für eine Verschiebung um m_0 und sein Symbol

Die Größen $m_{0,i}$ und $m'_{0,i}$ werden für allgemeine M auf rekursive Weise bestimmt:

$$\begin{aligned} m'_{0,0} &= m_0 \\ m_{0,i+1} &= m'_{0,i} \bmod M \\ m'_{0,i+1} &= (m'_{0,i} - m_{0,i+1})/M \\ i &= 0, 1, 2, \dots, T-1. \end{aligned} \quad (2.100)$$

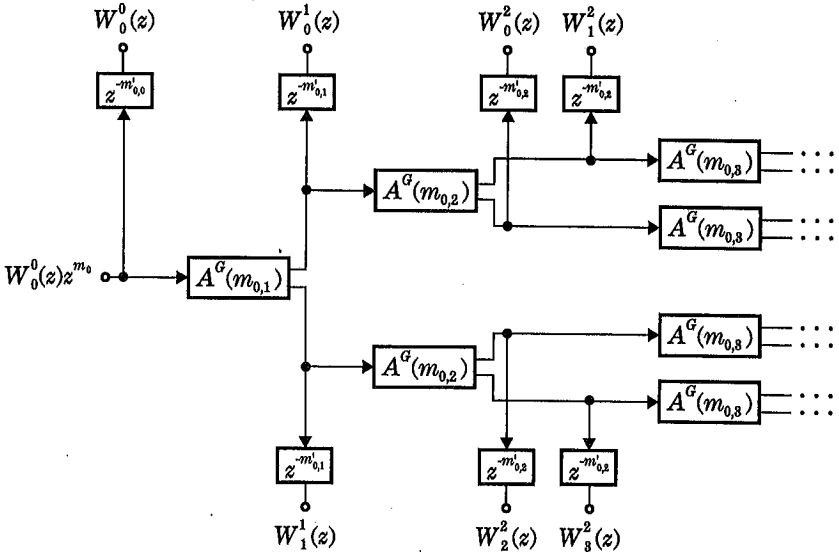


Bild 2.20: Filterbank mit Baumstruktur der Zerlegungstiefe $T = 3$ mit $M = 2$, die für eine Verschiebung m_0 ausgelegt ist

Auf diese Weise wird die Verschiebung bei Zerlegungstiefe i_0 zerlegt in die Form:

$$m_0 = m'_{0,i_0} + \tilde{m}_0(i_0) \quad (2.101)$$

mit

$$\tilde{m}_0(i_0) = \sum_{i=1}^{i_0} m_{0,i} M^{i-1}. \quad (2.102)$$

Man kann die Summe $\tilde{m}_0(i_0)$ so interpretieren, daß sie den Teil der Verzögerung repräsentiert, der durch Anpassung der Filterbank aufgefangen wird. Der Term m'_{0,i_0} entspricht der Verschiebung, die abschließend ausgeführt werden muß, um das entsprechende Ausgangssignal zum unverschobenen Signal zu erhalten.

Soll die Abbildung für eine Menge von Verschiebungen $\mathbb{W}$ durchgeführt werden, ist es recheneffizient, die Filterbank für alle vorgesehenen Verschiebun-

gen auszulegen. Zur Bestimmung der Abbildung $F_{m_0}^B(m, n)$ für eine bestimmte Verschiebung $m_0 \in \mathbb{W}$ ist dann eine Untermenge der Ausgaben der ergänzten Filterbank relevant.

Ausgangspunkt sei das Signal $W_l^{i,p}(z)$, das bezüglich aller Verschiebungen der Menge $\mathbb{W}^{i,p}$ zerlegt werden soll. Der Index i beschreibt wie in Abschnitt 2.1.3.1 die aktuelle Zerlegungstiefe, l repräsentiert den gewählten Kanal und $p \in \mathbb{Z}$ ist der Parameter der Verschiebung.

Aus der Menge $\mathbb{W}^{i,p}$ mit der Mächtigkeit $|\mathbb{W}^{i,p}|$ muß die Entscheidung abgeleitet werden, ob die Berechnung für den Baustein $A^G(q)$ mit $q = 0, 1, 2, \dots, M-1$ durchgeführt werden soll. Dazu werden die Steuergrößen $d^{i+1,p+qM^i}$ mit der Bedeutung

$$d^{i+1,p+qM^i} = \begin{cases} 1 & \text{Berechnung erfolgt} \\ 0 & \text{Berechnung erfolgt nicht} \end{cases}$$

vereinbart. Die Menge der Verschiebungen $\mathbb{W}^{i,p}$

$$\mathbb{W}^{i,p} = \{v_1^{i,p}, v_2^{i,p}, \dots, v_{|\mathbb{W}^{i,p}|}^{i,p}\}$$

wird zunächst in disjunkte Teilmengen $\mathbb{W}_T^{i+1,p+qM^i}$ zerlegt, die von den Grundbausteinen $A^G(q)$ verarbeitet werden können:

$$\mathbb{W}_T^{i+1,p+qM^i} = \{v \in \mathbb{W}_T^{i,p} | (v - q) \bmod M = 0\}. \quad (2.103)$$

Aus den Teilmengen $\mathbb{W}_T^{i+1,p+qM^i}$ werden die Nachfolgermengen $\mathbb{W}^{i+1,p+qM^i}$ abgeleitet:

$$\mathbb{W}^{i+1,p+qM^i} = \{(v_T - q)/M | v_T \in \mathbb{W}_T^{i+1,p+qM^i}\}. \quad (2.104)$$

Die Entscheidungsgrößen $d^{i+1,p+qM^i}$ werden schließlich bestimmt durch:

$$d^{i+1,p+qM^i} = \begin{cases} 1 & \text{wenn } |\mathbb{W}^{i+1,p+qM^i}| \neq 0 \\ 0 & \text{wenn } |\mathbb{W}^{i+1,p+qM^i}| = 0 \end{cases}. \quad (2.105)$$

Eine Filterbank mit Baumstruktur der Zerlegungstiefe T wird für eine Menge von Verschiebungen $\mathbb{W}$ rekursiv aufgebaut. Ausgehend von

$$\begin{aligned} W_0^{0,0}(z) &= S(z) \\ \mathbb{W}^{0,0} &= \mathbb{W} \end{aligned}$$

werden die Nachfolgermengen $\mathbb{W}^{i+1,p+qM^i}$ und daraus die Entscheidungsgrößen $d^{i+1,p+qM^i}$ bestimmt. Die Filterausgänge $W_{Ml+b}^{i+1,p+qM^i}$ werden für alle $b = 0, 1, 2, \dots, M-1$ bzw. q berechnet, für die $d^{i+1,p+qM^i} = 1$ gilt. Der Zusammenhang ist im Signalfluß in Bild 2.21 dargestellt.

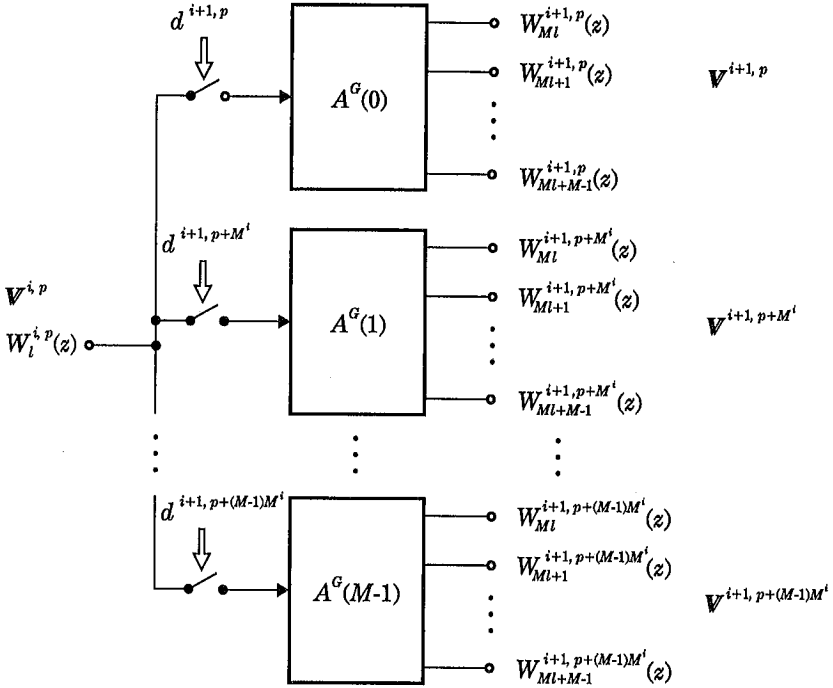


Bild 2.21: Zerlegung des Eingangssignals $W_l^{i,p}(z)$ bezüglich einer Menge von Verschiebungen $\mathbb{V}^{i,p}$

Jedes Ausgangssignal $W_{Ml+b}^{i+1,p+qM^i}$ wird mit der zugehörigen Nachfolgermenge $\mathbb{W}^{i+1,p+qM^i}$ der gleichen Zerlegung unterzogen, bis die gewünschte Zerlegungstiefe T erreicht ist.

Entsprechend den Bildern 2.13 und 2.14 läßt sich ein ähnlicher Zusammenhang zwischen dem Eingangssignal $S(z)$ und $W_l^{i,p}(z)$ angeben. Ausgehend von der Darstellung der Verschiebung $m_0 \in \mathbb{W}$ durch (2.101) folgt die Darstellung in Bild 2.22. Die Verzögerungen lassen sich wegen (2.102) zu einem Verzögerungsglied zusammenfassen. Man beachte, daß durch die Bildung der Koeffizienten p , wie in Bild 2.21 verdeutlicht, folgender Zusammenhang zwi-

schen den Koeffizienten i und p bei dem Ausgangssignal $W_l^{i,p}(z)$ für eine Verschiebung m_0 gilt:

$$W_l^{i,p}(z), m_0 \rightarrow p = \tilde{m}_0(i). \quad (2.106)$$

Daher läßt sich Bild 2.22 in die Form des Bilds 2.23 überführen.



Bild 2.22: Zusammenhang zwischen dem Eingangssignal $S(z)$ und dem Ausgangssignal $W_l^{i,p}(z)$ bei Betrachtung der Verschiebung m_0

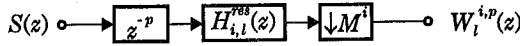


Bild 2.23: Umformung von Bild 2.22

Für die Ausgangsfolge $\{w_l^{i,p}(m)\}_m \stackrel{z}{\circ} W_l^{i,p}(z)$ ergibt sich:

$$w_l^{i,p}(m) = \sum_{k=0}^{K_i-1} h_{i,l}^{res}(k) s(M^i m - p - k). \quad (2.107)$$

Aus (2.107) folgt die Interpretation für die Größe p : Bezüglich der nichtdezimierten, um $M^i - 1$ verzögerten Folge

$$\tilde{w}_l^i(m) = \sum_{k=0}^{K_i-1} h_{i,l}^{res}(k) s(m - (M^i - 1) - k)$$

ist die Folge $\{w_l^{i,p}(m)\}_m$ nach Anhang 6.2.4 die Polyphasenkomponente p zweiter Art bezüglich der Dezimation um M^i :

$$\{(\tilde{w}_l^i)_p^{(2)}(m)\}_m = \{w_l^{i,p}(m)\}_m.$$

In Analogie zu Abschnitt 2.1.3.1 spannen die orthonormalen Vektoren

$$\bar{v}_{l,mM}^{i,p} = \{h_{i,l}^{res}(mM^i - p - k)\}_k, \quad m \in \mathbb{Z}$$

den Raum $R_l^{i,p}$ auf mit $w_l^{i,p}(m)$ als Entwicklungskoeffizienten des Vektors $\bar{s}$ zu den Vektoren $\bar{b}_{l,mM}^{i,p}$:

$$w_l^{i,p}(m) = \langle \bar{s}, \bar{b}_{l,mM}^{i,p} \rangle.$$

Der Filtervorgang einer translationsinvariant ausgelegten Filterbank mit Baumstruktur läßt sich als Orthogonalraumzerlegung interpretieren. Das Signal $\{w_l^{i,p}(m)\}_m$ in dem Raum $R_l^{i,p}$ wird verschiebungsabhängig in die folgenden Orthogonalräume zerlegt, wobei $R_0^{0,0} = l^2(\mathbb{Z})$ gilt.

Gilt $d^{i+1,p+qM^i} = 1$ mit $q \in \{0, 1, 2, \dots, M-1\}$ wird die Zerlegung

$$R_l^{i,p} = R_{lM}^{i+1,p+qM^i} \oplus R_{lM+1}^{i+1,p+qM^i} \oplus \dots \oplus R_{lM+M-1}^{i+1,p+qM^i} \quad (2.108)$$

durchgeführt.

Abschließend sei die Merkmalsrepräsentation des unverschobenen Eingangssignals hergeleitet. Da sich die Verschiebung m_0 durch (2.101) darstellen läßt, folgt mit (2.106) und der Abbildung (2.87):

$$F_{m_0}^B(m, n(i, l)) = w_l^{i, \tilde{m}_0(i)}(m - m'_{0,i}) \quad (2.109)$$

und erfüllt damit die in (2.97) geforderte Eigenschaft. In Bild 2.24 ist ein Beispielgraph für $M = 2$ und $T = 3$ dargestellt, der die Unterraumzerlegung für m_0 nach (2.108) wie in Bild 2.16 repräsentiert.

Für den Fall $M = 2$ ist der Zusammenhang zwischen den Ausgangssignalen $W_l^i(z)$ des unverschobenen Signals und der Filterausgänge $W_l^{i,p}(z)$ in Bild 2.25 verdeutlicht.

2.1.3.3 Interpretation der Merkmale in der Zeit-Frequenz-Ebene

Die diskrete Wavelet-Transformation aus Abschnitt 2.1.2.1 und die para-unitäre Filterbank aus Abschnitt 2.1.1 sind als Sonderfälle einer Filterbank mit Baumstruktur anzusehen.

Für die Interpretation wird davon ausgegangen, daß die verwendeten Filter $\{h_b(k)\}_k$ mit $b = 0, 1, 2, \dots, M-1$ und damit auch die resultierenden Filter $\{h_{i,l}^{res}(k)\}_k$ orthogonal sind.

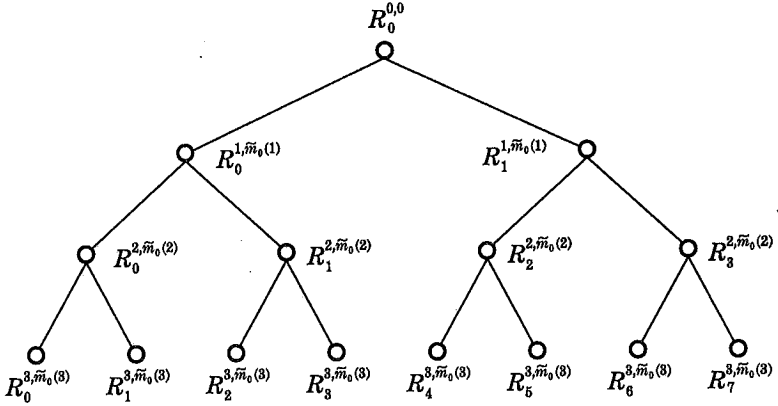


Bild 2.24: Graphische Darstellung der Unterraumzerlegung für die Verschiebung m_0 durch einen Graphen für $M = 2$ und $T = 3$

Für die resultierenden Filter gilt bei einem Takt Δt der Eingangsfolge die Zuordnung:

$$H'_{i,l}{}^{res}(f) = \sum_{k=0}^{K_i} h'_{i,l}{}^{res}(k\Delta t) e^{j2\pi k f \Delta t}.$$

In Bild 2.26 sind die entsprechenden idealen Amplitudengänge der Filter $\{h_b(k)\}_k$ für das Beispiel $\{h_{2,7}^{res}(k)\}_k$ in Bild 2.27 ($M = 3$) dargestellt, so daß $H_{2,7}^{res}(z) = H_1(z^3)H_2(z)$ gilt.

Durch Umkehrung der Abbildung (2.87) lassen sich zu gegebenem n die Parameter i und l bestimmen. Den Merkmalen $F^B(m, n)$ lassen sich in Abhängigkeit von n in der Zeit-Frequenz-Ebene die Bandbreiten, die dem idealen Amplitudengang des Filters $H'_{i,l}{}^{res}(f)$ entsprechen, und wegen (2.91) der Takt $M^i \Delta t$ zuordnen. In den Bildern 2.28 und 2.29 sind die Zeit-Frequenz-Raster der Merkmale der Filterbänke, die durch die Graphen in Bild 2.17 (a) und (b) gegeben sind, dargestellt.

Für die Darstellung eines Signals, das mit einer translationsinvarianten Filterbank transformiert worden ist, werden die Merkmale (2.109) unter Vernachlässigung der Verschiebung genauso dargestellt, wie die einer gewöhnlicher Filterbank mit Baumstruktur.

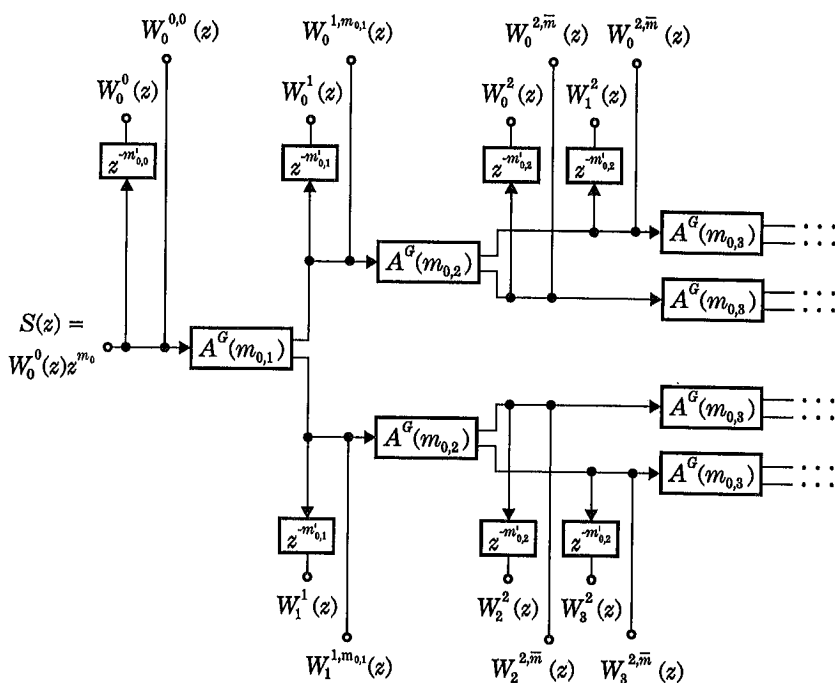


Bild 2.25: Filterbank mit Baumstruktur ohne Verschiebung der Filterausgangssignale mit $M = 2$ für eine Verschiebung m_0 ($\bar{m} = m_{0,1} + Mm_{0,2}$)

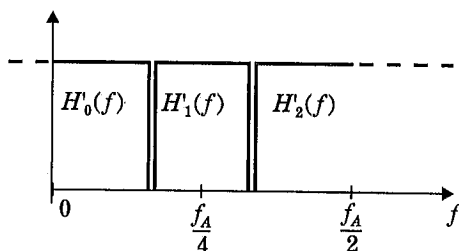


Bild 2.26: Ideale Amplitudengänge der paraunitären Filterbank für $M = 3$

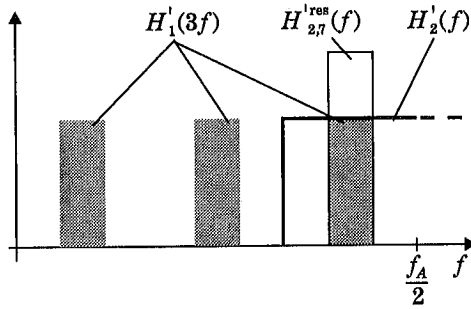


Bild 2.27: Idealer Amplitudengang des Filters $H_{2,7}'^{res}(f)$

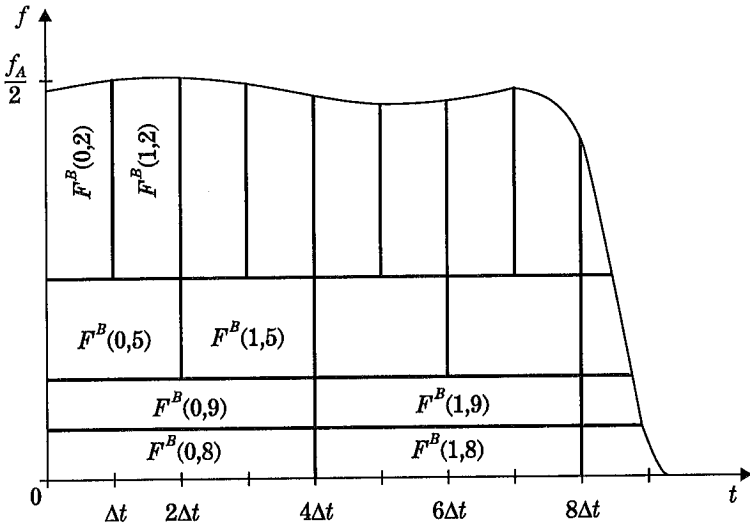


Bild 2.28: Zeit-Frequenz-Belegung der Wavelet-Transformation nach Bild 2.17 (a)

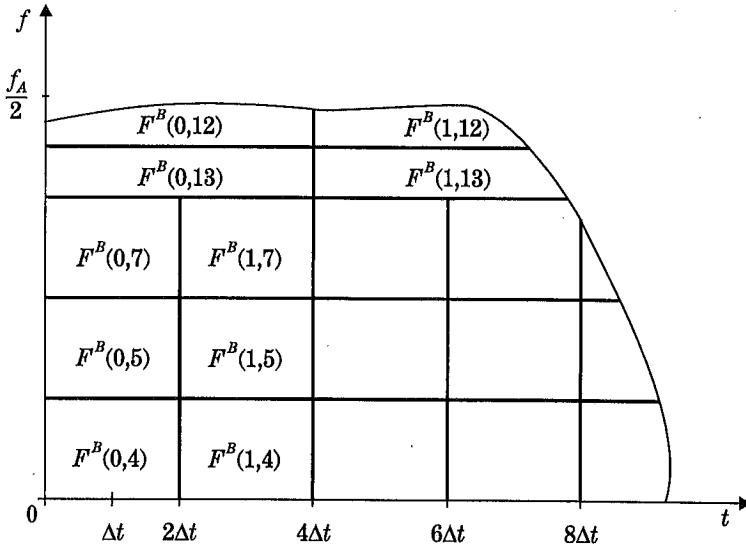


Bild 2.29: Zeit-Frequenz-Belegung für Unterraumzerlegung nach Bild 2.17 (b)

2.1.4 Kurzzeit-Fourier-Transformation

Die Kurzzeit-Fourier-Transformation (KFT) hat sich aus der Gabor-Transformation [Gab46] entwickelt und wird durch folgende Vorschrift beschrieben:

$$\begin{aligned}
 KFT(\tau, \omega) &= \int_{-\infty}^{\infty} s(t) w^*(t - \tau) e^{-j\omega t} dt \\
 &= \langle s(t), w(t - \tau) e^{j\omega t} \rangle \quad \omega, \tau \in \mathbb{R}.
 \end{aligned}$$

Die Eingangsfunktion wird durch eine um τ verschobene, im folgenden stets als reell angenommene, Fensterfunktion $w(t)$ gewichtet und einer Fourier-Transformation unterzogen.

Der Fensterungsprozeß läßt sich auch als Faltung mit der zeitgespiegelten Funktion $v(t) = w(-t)$ interpretieren:

$$KFT(\tau, \omega) = \int_{-\infty}^{\infty} s(t)v(\tau - t)e^{-j\omega t} dt.$$

Die zeitdiskrete Form der KFT ist mit der diskreten Verschiebung $\tau = m$ definiert durch [Vai93]:

$$KFT(m, e^{j\omega}) = \sum_{k=-\infty}^{\infty} s(k)v(m - k)e^{-j\omega k}, \quad m \in \mathbb{Z}.$$

Für die Implementierung wird die Größe ω äquidistant an den Stellen

$$\omega_n = 2\pi n/N \quad (2.110)$$

abgetastet. Mit $W_N = e^{-j2\pi/N}$ ergibt sich die Darstellung:

$$KFT(m, n) = \sum_{k=-\infty}^{\infty} s(k)v(m - k)W_N^{nk}, \quad n = 0, 1, 2, \dots, N - 1. \quad (2.111)$$

2.1.4.1 Implementierung durch eine Filterbank

Durch Umformen von (2.111) in

$$KFT(m, n) = W_N^{mn} \sum_{k=-\infty}^{\infty} v(k)W_N^{-nk}s(m - k)$$

läßt sich die Kurzzeit-Fourier-Transformation als modulierte Filterbank interpretieren:

$$KFT(m, n) = W_N^{mn} \sum_{k=-\infty}^{\infty} h_n(k)s(m - k), \quad (2.112)$$

$$h_n(k) = v(k)W_N^{-nk}, k \in \mathbb{Z}. \quad (2.113)$$

Für die Merkmalsextraktion werden die Filter $\{h_n(k)\}_k$ aus einem reellwertigen Tiefpaßfilter $\{v(k)\}_{k=0}^{K-1}$ mit endlicher Impulsantwort (FIR) entworfen. Als Entwurfsverfahren wird hier die Fourierapproximation mit Kaiser-Fensterung bei Vorgabe des Amplitudengangs über die Grenzfrequenz $\omega_c(N) = \pi/N$ verwendet [MK93, Seite 174]:

$$|V(e^{j\omega})| = \begin{cases} 1 & |\omega| \leq \omega_c(N) \\ 0 & \omega_c(N) < |\omega| \leq \pi \end{cases} \quad (2.114)$$

Das Ergebnis der KFT kann als Menge von Signalfolgen $\{y'_n(m)\}_m$ interpretiert werden:

$$y'_n(m) = W_N^{-mn} KFT(m, n), \quad m \in \mathbb{Z}.$$

Unter der vereinfachenden Annahme, daß die entworfenen Filter ideal bandbegrenzt sind, sind die Ausgangssignale $\{y'_n(m)\}_m$ ebenfalls bandbegrenzt und können um den Faktor $M = N$ dezimiert werden.

$$y_n(m) = KFT(Mm, n) = y'_n(mM) = \sum_{k=0}^{K-1} h_n(k) s(mM - k). \quad (2.115)$$

Man beachte, daß der Takt der Folge $\{y_n(m)\}_m$ um den Faktor M langsamer ist als der der Folge $\{y'_n(m)\}_m$.

Die durch (2.115) gegebene, in Zeitrichtung dezimierte Form der Kurzzeit-Fourier-Transformation wird für die Merkmalsextraktion eingesetzt und ist in Bild 2.30 für reelle Folgen $\{s(m)\}_m$ dargestellt. Die doppelt ausgeführten Linien bezeichnen komplexwertige Signale.

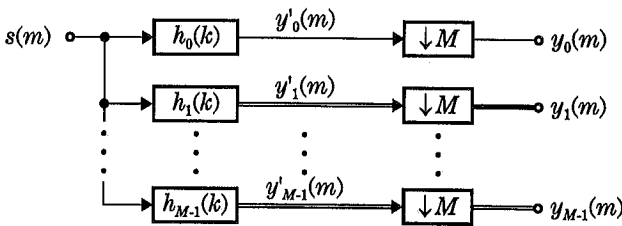


Bild 2.30: Filterbankdarstellung der Kurzzeit-Fourier-Transformation bei einer Dezimation um M

Als Merkmale werden die Betragsquadrate der Ausgangsfolgen $\{y_n(m)\}_m$ verwendet:

$$F^{KFT}(m, n) = |y_n(m)|^2, \quad n = 0, 1, 2, \dots, N-1. \quad (2.116)$$

Eine effiziente Implementierung der Filterbank mit Hilfe der schnellen Fourier-Transformation (FFT) wurde zuerst in [Sch74] angegeben und ist in [CR83, Jon91, Fli93] beschrieben. Die Idee sei im folgenden skizziert.

Für die Filter $\{h_n(m)\}_{m=0}^{K-1}$ läßt sich im z -Bereich mit $N = M$ und mit den Polyphasenkomponenten $H_{0,\lambda}^{(1)}(z)$ wie in (2.4) die Darstellung

$$H_n(z) = H_0(zW_M^n) = \sum_{\lambda=0}^{K-1} W_M^{-n\lambda} H_{0,\lambda}^{(1)}(z^M) z^{-\lambda} \quad (2.117)$$

angeben. Ausgehend von (2.112) folgt für die nicht dezimierten Folgen $\{y'_n(m)\}_m$ im z -Bereich:

$$Y'_n(z) = \sum_{\lambda=0}^{K-1} W_M^{-n\lambda} H_{0,\lambda}^{(1)}(z^M) S(z) z^{-\lambda}. \quad (2.118)$$

Über die Hilfsgröße $\hat{S}(z) = z^{-1}S(z)$ läßt sich der in Bild 2.31 dargestellte Signalfluß unter Nutzung der Polyphasenkomponenten zweiter Art von $\hat{S}(z)$ herleiten.

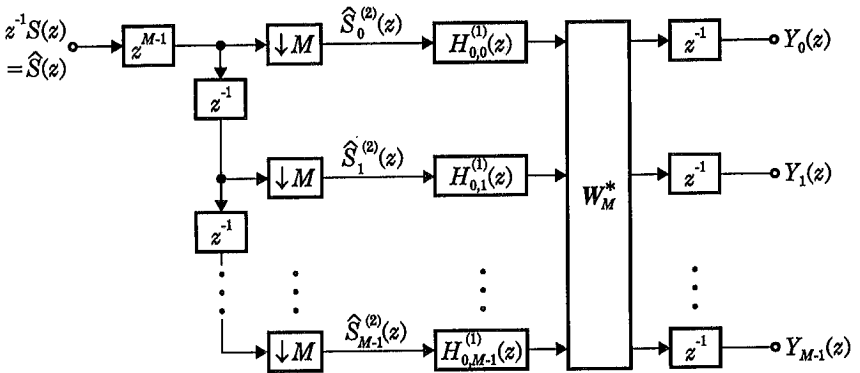


Bild 2.31: Herleitung der DFT-Filterbank

Dabei ist $\mathbf{W}_M^*$ bis auf den Vorfaktor $1/M$ die IDFT-Matrix [Sch94].

Wenn die ausgangsseitigen Verzögerungsglieder durch den Matrixblock, die Filter und die Dezimatoren geschoben werden, erhält man den Signallauf in Bild 2.32.

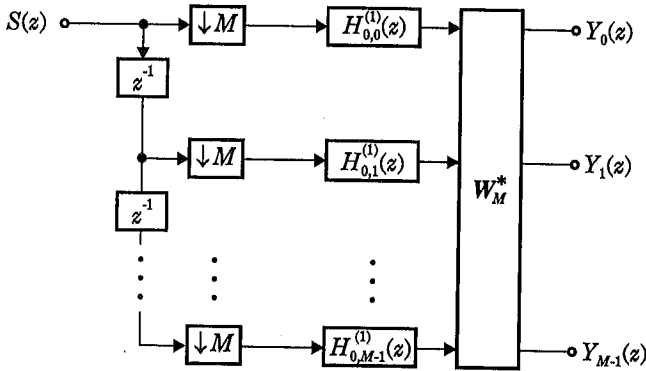


Bild 2.32: Implementierung der dezimierten Kurzzeit-Fourier-Transformation durch eine DFT-Filterbank

Aus Bild 2.32 läßt sich durch Vergleich mit Bild 2.4 und (2.10) für die Analysefilterbank die Polyphasenmatrix erster Art $\mathbf{H}^{(1)}(z)$ zu

$$\mathbf{H}^{(1)}(z) = \mathbf{W}_M^* \mathbf{H}_0^{(1)}(z)$$

mit

$$\mathbf{H}_0^{(1)}(z) = \begin{pmatrix} H_{0,0}^{(1)}(z) & 0 & 0 & \dots & 0 \\ 0 & H_{0,1}^{(1)}(z) & 0 & \dots & 0 \\ 0 & 0 & H_{0,2}^{(1)}(z) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & H_{0,M-1}^{(1)}(z) \end{pmatrix}$$

ablesen.

Die Matrix $\mathbf{W}_M^*$ ist unitär. Soll die Polyphasenmatrix $\mathbf{H}^{(1)}(z)$ paraunitär nach (2.15) sein, muß dies auch für die Matrix $\mathbf{H}_0^{(1)}(z)$ gelten.

Man kann leicht zeigen, daß dies nur für $\mathbf{H}_0^{(1)}(z) = c\mathbf{I}_M$ mit $c \neq 0$ möglich ist, wenn die Analysefilterbank aus kausalen Filtern mit endlicher Impulsantwort

bestehen soll und Lösungen, die nur Verzögerungsglieder enthalten, ausgeschlossen werden. Dies erzwingt die Filterlänge $K = M$ und damit für den Tiefpaß $\{v(k)\}_{k=0}^{M-1}$ die Eigenschaft

$$v(k) = c, \quad k = 0, 1, 2, \dots, M-1.$$

Die diskrete Kurzzeit-Fourier-Transformation ist bei Beschränkung auf den in der Praxis wichtigen Fall einer FIR-Analysefilterbank und Dezimation um M also nur für den trivialen Fall, daß als Analysefenster bzw. Analysefilter Rechteckfolgen der Länge M gewählt werden, orthogonal. Durch den in dieser Arbeit verwendeten Filterentwurf wird keine orthogonale Transformation durchgeführt, d.h.

$$\sum_{m=-\infty}^{\infty} |s(m)|^2 \neq \sum_{n=0}^{M-1} \sum_{m=-\infty}^{\infty} |y_n(m)|^2.$$

2.1.4.2 Interpretation der Merkmale in der Zeit-Frequenz-Ebene

Die Merkmale $F^{KFT}(m, n)$, die aus der Form der diskreten Kurzzeit-Fourier-Transformation aus (2.116) abgeleitet werden, lassen sich in der Zeit-Frequenz-Ebene physikalisch interpretieren.

Durch die in Bild 1.2 angegebenen Zuordnungen läßt sich für die N -kanalige Filterbank aus (2.113) die Darstellung

$$H'_n(f) = \sum_{k=0}^{K-1} v(k\Delta t) e^{j2\pi k(f - \frac{f_A}{N})\Delta t}$$

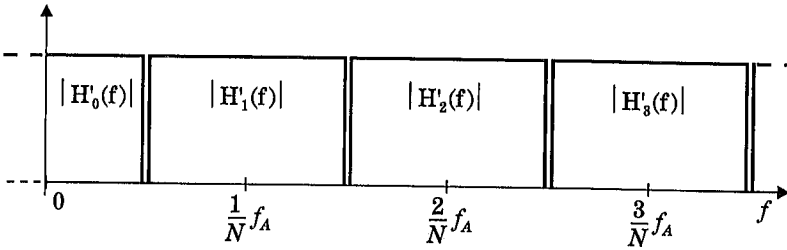
ableiten, die in Bild 2.33 durch die idealen Amplitudengänge einer Filterbank mit N Kanälen veranschaulicht ist.

Die Merkmale $F^{KFT}(m, n)$ lassen sich unter Berücksichtigung der Dezimation einem diskreten Zeit-Frequenz-Raster zuordnen:

$$F^{KFT}(m, n) \triangleleft \triangleright F^{KFT}(mM\Delta t, n\Delta f)$$

mit

$$\Delta f = \frac{1}{N\Delta t} = \frac{f_A}{N}.$$

Bild 2.33: Ideale Amplitudengänge einer KFT-Filterbank mit N Kanälen

Jedem Element $F^{KFT}(mM\Delta t, n\Delta f)$ ist eine Fläche in der Zeit-Frequenz-Ebene zugeordnet, die dem Takt der Elemente und der Diskretisierung der Frequenzvariablen entspricht. Alternativ kann man Δf als Bandbreite der Analysekanäle nach Bild 2.33 interpretieren. Die Zuordnung in der Zeit-Frequenz-Ebene ist in Bild 2.34 veranschaulicht.

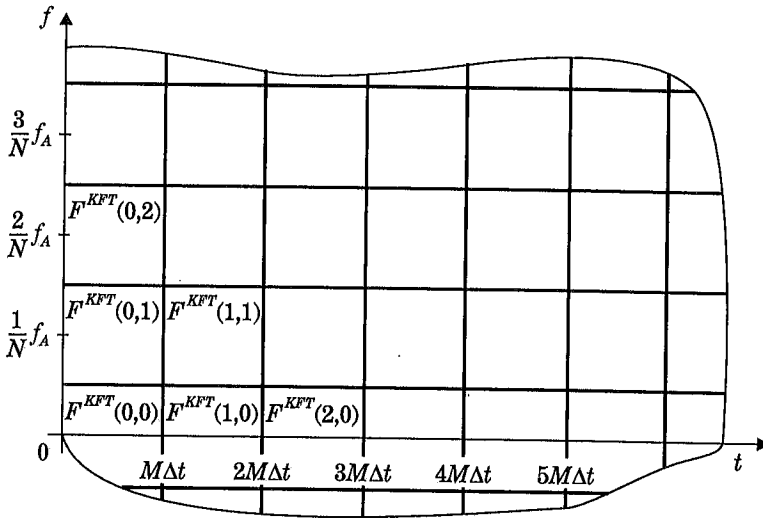


Bild 2.34: Belegung der Zeit-Frequenz-Ebene durch die Merkmale der KFT

Eine Zeit-Frequenz-Darstellung der Elemente $F^{KFT}(m, n)$ zeigt näherungsweise, wo die Signalenergie des transformierten Signales bezüglich der Zeit und der Frequenz lokalisiert ist. Die Ungenauigkeit der Näherung ist durch die einfache Zuordnung der Merkmale zu Ersatzflächen in der Zeit-Frequenz-

Ebene und durch die fehlende Orthogonalität der Transformation bedingt.

2.2 Quadratische Zeit-Frequenz-Verfahren

Quadratische Verfahren zur Zeit-Frequenz-Analyse wurden erstmals im Rahmen der Quantenmechanik in [Wig32, Moy49] angesprochen, nach [Vil48] allerdings erst 1980 [CM80a, CM80b, CM80c] für die Signalanalyse erschlossen. Ebenso wie die Unschärferelation der Nachrichtentechnik [Gab46] lassen sich quadratische Signalanalyseverfahren durch eine formale Analogie aus der Quantenmechanik ableiten. Alle im folgenden beschriebenen Verfahren werden durch Parametrisierung einer allgemeinen Klasse von Zeit-Frequenz-Verfahren, der Cohen-Klasse, ermittelt. Im Abschnitt 2.2.1 werden zunächst die für die Herleitung der Cohen-Klasse notwendigen quantenmechanischen Grundlagen erläutert.

2.2.1 Die Cohen-Klasse

Nach den Axiomen der Quantentheorie [Fic84] wird der Zustand eines Objekts durch eine Wellenfunktion ξ beschrieben. Physikalisch meßbare Größen (Observable) wie z.B. Ort, Impuls oder Energie, werden durch hermitesche Operatoren ersetzt. Der Vorgang der Zuordnung klassischer Größen zu Operatoren heißt Quantisierung und die ermittelte Zuordnung Korrespondenzregel. Im folgenden werden nur eindimensionale Objekte betrachtet, die vollständig durch ihren Ort oder ihren Impuls beschrieben werden können. Die Zeitabhängigkeit wird nicht betrachtet. Die Operatoren beziehen sich auf ein bestimmtes vorgegebenes Koordinatensystem und werden durch ihre Eigenvektoren bzw. Eigenwerte beschrieben. Die Eigenwerte eines Operators sind die Meßwerte, die die zugehörige physikalische Größe annehmen kann. So gilt für den Ortsoperator $\mathcal{X}^q$ in der Ortsdarstellung die Korrespondenzregel $x \rightarrow \mathcal{X}_x^q$ mit x als Ortskoordinate. Allgemein liefert die Projektion des Zustandsvektors ξ auf die Eigenfunktionen des Ortsoperators $\mathcal{X}_x^q$ die Ortsdarstellung des Zustands $\xi(x)$. Das Argument der Funktion $\xi(x)$, die Ortskoordinate x , ist dabei als Eigenwert des Ortsoperators $\mathcal{X}_x^q$ zu interpretieren.

Für komplizierte Funktionen von physikalischen Größen gibt es verschiedene Ansichten, ob und wie die zugehörigen Operatorfunktionen darstellbar sind [Fic84].

Der Mittelwert einer Observablen mit zugehörigem Operator $\mathcal{A}^q$ ist durch

das Skalarprodukt

$$\langle A^q \xi, \xi \rangle \quad (2.119)$$

gegeben.

In der klassischen Physik ist der Impuls p bei Verschiebungen eine Erhaltungsgröße. Man kann damit zeigen [Fic84], daß der Verschiebungsoperator $\mathcal{V}^q(\tau)$ und der Impulsoperator $\mathcal{P}^q$ über folgende Beziehung zusammenhängen:

$$\mathcal{V}^q(\tau) = e^{-\frac{i}{\hbar} \tau \mathcal{P}^q}, \quad (2.120)$$

wobei $\hbar$ das Plancksche Wirkungsquantum ist und $\hbar = h/2\pi$ gilt. Mit Hilfe des Verschiebungsoperators $\mathcal{V}^q(\tau)$ folgt ebenso, daß für den Impulsoperator $\mathcal{P}^q$ in der Koordinatendarstellung des Ortes die Darstellung

$$\mathcal{P}_X^q = \frac{\hbar}{j} \frac{d}{dx} \quad (2.121)$$

gilt.

Für eine gemeinsame, dynamische Darstellung des Objekts als Funktion des Ortes x und des Impulses p wird der Erwartungswert bezüglich der Funktion $e^{j\theta x + j\tau p}$, d.h. die charakteristische Funktion, ermittelt [Moy49]. Eine eindeutige Korrespondenzregel für diese Funktion existiert nicht. Verschiedene Korrespondenzen führen zu verschiedenen Darstellungen. Cohen hat in [Coh66] gezeigt, daß die Korrespondenzregel

$$e^{j\theta x + j\tau p} \longrightarrow G(\theta, \tau) e^{j\theta \mathcal{X}^q + j\tau \mathcal{P}^q} = \mathcal{M}^q(\theta, \tau, \mathcal{X}^q, \mathcal{P}^q) \quad (2.122)$$

mit dem Mittelwert

$$\mathcal{M}^q(\theta, \tau) = \langle G(\theta, \tau) e^{j\theta \mathcal{X}^q + j\tau \mathcal{P}^q} \xi, \xi \rangle \quad (2.123)$$

zu plausiblen Darstellungen führt, wenn

$$\mathcal{M}_q(\theta, 0, \mathcal{X}^q, \mathcal{P}^q) = e^{j\theta \mathcal{X}^q}$$

und

$$\mathcal{M}^q(0, \tau, \mathcal{X}^q, \mathcal{P}^q) = e^{j\tau \mathcal{P}^q}$$

gilt und damit die **Randeigenschaften**

$$M^q(\theta, 0) = \langle e^{j\theta \mathcal{X}^q} \xi, \xi \rangle \quad \text{und} \quad (2.124)$$

$$M^q(0, \tau) = \langle e^{j\tau \mathcal{P}^q} \xi, \xi \rangle \quad (2.125)$$

erfüllt sind. Für (2.124) und (2.125) müssen ebenso die Bedingungen an die Gewichtsfunktion $G(\theta, \tau)$, die auch Kern genannt wird,

$$G(\theta, 0) = 1 \quad \text{und} \quad (2.126)$$

$$G(0, \tau) = 1 \quad (2.127)$$

eingehalten werden.

Nach Einsetzen der Operatoren, entsprechend ihrer Darstellung in dem gewählten Koordinatensystem, wird durch zweifache inverse Fourier-Transformation aus (2.123) die gemeinsame Darstellung bezüglich x und p ermittelt:

$$C^q(x, p) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M^q(\theta, \tau) e^{-j\theta x - j\tau p} d\theta d\tau. \quad (2.128)$$

Die Interpretation der Funktion $C^q(x, p)$ als klassische Wahrscheinlichkeitsdichte ist nicht möglich, da sie abhängig von ξ auch negative Werte annehmen kann. Lediglich für Funktionen ξ , die Exponentialfunktionen von quadratischen Polynomen sind, und der Wahl $G(\theta, \tau) = 1$, konnte die Nichtnegativität der Funktion $C^q(x, p)$ nachgewiesen werden [Hud74].

Man kann zeigen, daß die Projektionen des Zustandsvektors ξ auf die Eigenfunktionen des Ortsoperators $\mathcal{X}^q$ und des Impulsoperators $\mathcal{P}^q$ über die Fourier-Transformation in Beziehung stehen [Fic84]

$$\xi(x) \circ \bullet \Xi(p).$$

Für den Übergang zur Zeit-Frequenz-Analyse ersetzt man nun den Zustandsvektor $\xi(x)$ im Ortsbereich durch das Zeitsignal $s(t)$ und den Zustandsvektor $\Xi(p)$ im Impulsbereich durch die Fourier-Transformierte $S(\omega)$. Im Zeitbereich werden Zeitoperator $\mathcal{T}_T$, Frequenzoperator $\mathcal{W}_T$ und der Verschiebungoperator $\mathcal{V}_T$ definiert [Coh95]:

$$\begin{aligned}
\mathcal{T}_T &= t, \\
\mathcal{W}_T &= \frac{1}{j} \frac{d}{dt}, \\
\mathcal{V}_T(\tau) &= e^{j\tau\mathcal{W}_T}.
\end{aligned}$$

Aus (2.123) und unter Nutzung des Verschiebungsoperators $\mathcal{V}_T(\tau)$ ergibt sich die verallgemeinerte Ambiguitätsfunktion:

$$\begin{aligned}
M(\theta, \tau) &= \langle G(\theta, \tau) e^{j\theta\mathcal{T}_T + j\tau\mathcal{W}_T} s(t), s(t) \rangle \\
&= G(\theta, \tau) \int_{-\infty}^{\infty} s(t + \frac{\tau}{2}) s^*(t - \frac{\tau}{2}) e^{j\theta t} dt. \quad (2.129)
\end{aligned}$$

Dabei wurde berücksichtigt, daß die Operatoren $\mathcal{T}_T$ und $\mathcal{W}_T$ nicht vertauschbar sind und

$$\begin{aligned}
e^{j\theta\mathcal{T}_T + j\tau\mathcal{W}_T} s(t) &= e^{j\theta\frac{\tau}{2}} e^{j\theta\mathcal{T}_T} e^{j\tau\mathcal{W}_T} s(t) \\
&= e^{j\theta\frac{\tau}{2}} e^{j\theta\mathcal{T}_T} \mathcal{V}_T(\tau) s(t) \\
&= e^{j\theta\frac{\tau}{2}} e^{j\theta t} s(t + \tau)
\end{aligned}$$

verwendet [Coh66][Fic84, Seite 125]. Die Cohen-Klasse der Zeit-Frequenz-Analyse ergibt sich schließlich in Analogie zu (2.128) mit (2.129) zu:

$$\begin{aligned}
C(t, \omega) &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M(\theta, \tau) e^{-j\theta t - j\tau\omega} d\theta d\tau \\
&= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\theta, \tau) s(u + \frac{\tau}{2}) s^*(u - \frac{\tau}{2}) e^{-j\theta t - j\tau\omega + j\theta u} du d\theta d\tau. \quad (2.130)
\end{aligned}$$

Die Interpretation der Zeit-Frequenz-Darstellung $C(t, \omega)$ wird in der Praxis durch zwei Eigenschaften erschwert:

- Die Cohen-Klasse ist eine nichtlineare Transformation des Signals $s(t)$. Die Summe der Transformationen verschiedener Signale (Autoterme) unterscheidet sich von der Transformation des Summensignals durch die Kreuzterme. Kreuz- und Autoterme lassen sich oft nur durch Vorkenntnisse über das Signal $s(t)$ unterscheiden.

- Die Funktion $C(t, \omega)$ kann negative Werte annehmen.

Durch geeignete Wahl des Kerns $G(\theta, \tau)$ kann man beide Eigenschaften beeinflussen.

Die verallgemeinerte Ambiguitätsfunktion (2.129) erlaubt in der Schreibweise

$$\begin{aligned} M(\theta, \tau) &= G(\theta, \tau) \int_{-\infty}^{\infty} s(t) s^*(t + \tau) e^{j\theta t} dt \\ &= G(\theta, \tau) \langle s(t), s(t + \tau) e^{-j\theta t} \rangle \end{aligned}$$

eine einfache Deutung des Kerns $G(\theta, \tau)$:

Es wird die Ähnlichkeit (Korrelation) des Signals $s(t)$ und seiner um τ in der Zeit und um θ in der Frequenz verschobenen Version bestimmt. Mit Hilfe der Funktion $G(\theta, \tau)$ lassen sich die Korrelation zwischen Signalkomponenten, die den Zeit-Frequenz-Versatz (θ_0, τ_0) aufweisen, durch den Wert $G(\theta_0, \tau_0)$ gewichten. Daher bestimmt die Wahl der Funktion $G(\theta, \tau)$ die Eigenschaften der Kreuztermunterdrückung [HMUJ95].

Der Wunsch, alle Kreuzterme zu eliminieren, führt dazu, den Kern $G(\theta, \tau)$ nur um den Ebenenursprung von Null verschieden zu gewichten. Dies birgt den Nachteil, daß durch diese Fensterung im Ambiguitätsbereich im Zeit-Frequenz-Bereich eine Verschmierung der Autoterme bewirkt und damit die Auflösung verschlechtert wird.

2.2.2 Ausgewählte Transformationen der Cohen-Klasse

Im folgenden Abschnitt werden verschiedene, in der Merkmalsextraktion eingesetzte Zeit-Frequenz-Verfahren, die aus einer bestimmten Wahl für den Kern $G(\theta, \tau)$ hervorgehen, in ihren Eigenschaften diskutiert. In Bild 2.35 sind die Verfahren zusammengestellt.

2.2.3 Die Wigner-Verteilung (WV)

Für die Wigner-Verteilung wird $G_{WV}(\theta, \tau) = 1$ gewählt. Diese Transformation unterdrückt daher keine Kreuzterme. Durch Einsetzen in die Cohen-Klasse (2.130) erhält man die Transformationsvorschrift [Coh95]:

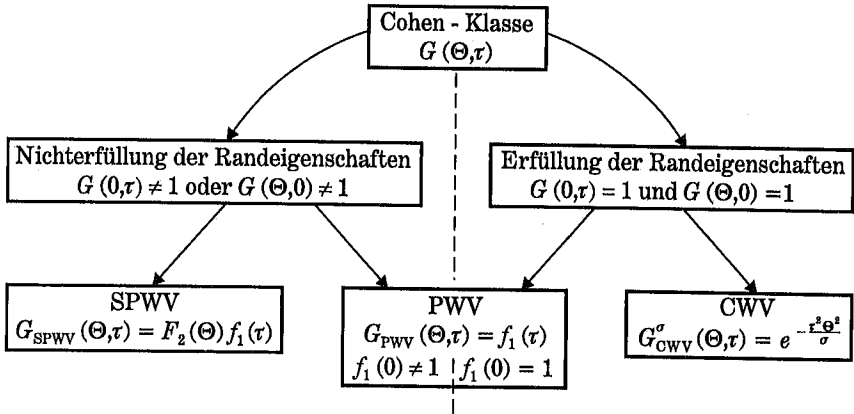


Bild 2.35: Übersicht über die für die Merkmalsextraktion eingesetzten Verfahren der Cohen-Klasse

$$C_{WV}(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} s\left(t + \frac{\tau}{2}\right) s^*\left(t - \frac{\tau}{2}\right) e^{-j\omega\tau} d\tau.$$

Die Wigner-Verteilung wurde als erste quadratische Transformation eingeführt [Wig32], blieb aber ohne praktische Bedeutung. Allein die im folgenden angesprochenen Varianten Pseudo-Wigner-Verteilung und Smoothed Pseudo-Wigner-Verteilung eignen sich auch für eine Implementierung.

2.2.4 Die Pseudo-Wigner-Verteilung (PWV)

Bei der Pseudo-Wigner-Verteilung wird mit $G_{PWV}(\theta, \tau) = f_1(\tau)$ die Ambiguitätsfunktion in Zeitrichtung gefensterter:

$$C_{PWV}(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f_1(\tau) s\left(t + \frac{\tau}{2}\right) s^*\left(t - \frac{\tau}{2}\right) e^{-j\omega\tau} d\tau.$$

Kreuzterme, die in Frequenzrichtung oszillieren, werden durch diese Kernwahl geglättet. Die PWV erfüllt für $f_1(0) = 1$ die Randeigenschaft (2.125).

2.2.5 Die Smoothed Pseudo-Wigner-Verteilung (SPWV)

Die Smoothed Pseudo-Wigner-Verteilung verwendet einen separablen Kern $G_{SPWV}(\theta, \tau) = F_2(\theta) f_1(\tau)$, der eine unabhängige Fensterung in Zeit- und Frequenzrichtung erlaubt. In der Praxis werden für $f_2(t) \circ \bullet F_2(\theta)$ und $f_1(\tau)$ zwei Tiefpaßfilter eingesetzt:

$$C_{SPWV}(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_1(\tau) f_2(u - t) s(u + \frac{\tau}{2}) s^*(u - \frac{\tau}{2}) e^{-j\omega\tau} du d\tau.$$

Kreuzterme, die in Zeit- oder in Frequenzrichtung oszillieren, werden durch diese Kernwahl geglättet. Die SPWV erfüllt keine der Randeigenschaften (2.124) und (2.125).

2.2.6 Die Choi-Williams-Verteilung (CWV)

Der Kern der Choi-Williams-Verteilung [CW89] ist eine von dem Parameter σ abhängige exponentielle Funktion $G_{CWV}^\sigma(\theta, \tau) = e^{-\frac{\tau^2 \theta^2}{\sigma}}$:

$$C_{CWV}(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\sqrt{\sigma}}{2|\tau|\sqrt{\pi}} e^{-\frac{\sigma(u-t)^2}{4\tau^2}} s(u + \frac{\tau}{2}) s^*(u - \frac{\tau}{2}) e^{-j\omega\tau} du d\tau.$$

Je größer der Parameter σ gewählt wird, desto rascher fällt die Kernfunktion $G_{CWV}^\sigma(\theta, \tau)$ für steigende τ und θ ab bzw. desto stärker werden die Kreuzterme, die im Ambiguitätsbereich einen Zeit- und Frequenzversatz aufweisen, unterdrückt. Die CWV erfüllt beide Randeigenschaften (2.124) und (2.125).

2.2.7 Die diskrete Cohen-Klasse

Für die Implementierung werden die Zeit-Frequenz-Verteilungen der Cohen-Klasse auf einem zeit- und frequenzdiskreten Raster berechnet. Im folgenden wird die zeit- und frequenzdiskrete Cohen-Klasse kurz diskrete Cohen-Klasse genannt.

Die diskrete Cohen-Klasse wird entsprechend der kontinuierlichen Cohen-Klasse definiert.

Zunächst wird (2.130) mit der bezüglich θ inversen Fourier-Transformierten $G(\theta, \tau) \xrightarrow{\theta} g(u, \tau)$ umgeformt:

$$\begin{aligned} C(t, \omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(u, \tau) s(u + t + \frac{\tau}{2}) s^*(u + t - \frac{\tau}{2}) e^{-j\omega\tau} du d\tau \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(u, 2\tau) s(u + t + \tau) s^*(u + t - \tau) e^{-j2\omega\tau} du d\tau. \end{aligned} \quad (2.131)$$

(2.131) dient zur Definition der zeitdiskreten Cohen-Klasse [Mer96]:

$$C(m, e^{j\omega}) = 2 \sum_{k=-K_1}^{K_1} \sum_{l=-K_2}^{K_2} g(l, k) s(l + m + k) s^*(l + m - k) e^{-j2k\omega}. \quad (2.132)$$

Zu der Diskretisierung der Funktion $g(u, \tau)$ sind zwei Bemerkungen notwendig:

- In der Literatur [Mer96, CW89] bleibt die Skalierung der Funktion $g(u, 2\tau)$ bezüglich τ bei der Diskretisierung unbeachtet. Dies mag dadurch begründet sein, daß die Matrix $g(l, k)$ in der Praxis direkt im Zeitbereich entworfen wird.
- Durch die endliche Summation in (2.132), die durch die Parameter K_1 und K_2 ausgedrückt wird, sind die Eigenschaften von $g(u, \tau)$ und $g(l, k)$ z.B. für die Choi-Williams-Verteilung verschieden. Die Folgen für die Berechnung der Zeit-Frequenz-Darstellung werden im praktischen Einsatz als vernachlässigbar angenommen [CW89].

Für die Zeit-Frequenz-Verteilungen WV, PWV und SPVW werden folgende diskrete Kernfunktionen definiert:

$$g_{WV}(l, k) = \delta_l, \quad (2.133)$$

$$g_{PWV}(l, k) = f_1(k) \delta_l, \quad (2.133)$$

$$g_{SPVW}(l, k) = f_1(k) f_2(l). \quad (2.134)$$

Bei der Choi-Williams-Verteilung wird wegen der endlichen Summation eine Normierung der Koeffizienten vorgenommen [Mer96]:

$$g_{CWV}^{\sigma}(l, k) = \begin{cases} \frac{1}{\beta_k} e^{-\frac{\sigma l^2}{4k^2}}, & k \neq 0 \\ \delta(l), & k = 0 \end{cases} \quad (2.135)$$

mit

$$\beta_k = \sum_{l=-K_2}^{K_2} e^{-\frac{\sigma l^2}{4k^2}}, \quad k = -K_1, \dots, -1, 1, \dots, K_1.$$

Bei der Implementierung der diskreten Cohen-Klasse wird nicht die reelle Folge $\{s(m)\}_m$, sondern die komplexwertige Folge $\{s_a(m)\}_m$ verwendet [Vil48, CM83], deren Spektrum $S_a(e^{j\omega})$ nur im positiven Bereich des Intervalls $-\pi < \omega < \pi$ von Null verschieden ist. Der Hintergrund ist am Beispiel der diskreten Wigner-Verteilung in Bild 2.36 erläutert [Boa88].

Zunächst wird die zeitdiskrete Wigner-Verteilung der Folge $\{s(m)\}_m$ berechnet, deren Spektrumsbelegung in Bild 2.36 (a) dargestellt ist. Neben Kreuztermen zwischen den symmetrisch zu $\omega = 0$ liegenden Spektralanteilen (Spiegelspektren) wird durch die Zeitdiskretisierung eine in π periodische Transformation erzeugt, die ggf. auch zu einem Überfalten (Aliasing) der Autoterme führt (Bild 2.36 (b)). Das Aliasing kann durch Wahl der Abtastrate $f_A \geq 4f_{max}$ vermieden werden, wenn f_{max} die höchste in $s(t)$ vorkommende Frequenzkomponente ist. Durch Verwendung der komplexwertigen Folge ist eine Erhöhung der Abtastrate nicht nötig. Darüberhinaus treten auch keine Kreuzterme zwischen den Spiegelspektren auf (Bild 2.36 (c)).

Für die Implementierung wird die Frequenzvariable ω durch

$$\omega_n = \pi n / N \quad (2.136)$$

diskretisiert mit $n = 0, 1, 2, \dots, N-1$, so daß nur das Intervall $[0, \pi)$ abgetastet wird:

$$C(m, n) = 2 \sum_{k=-K_1}^{K_1} \sum_{l=-K_2}^{K_2} g(l, k) s_a(l+m+k) s_a^*(l+m-k) W_N^{nk}. \quad (2.137)$$

Für die Merkmalsextraktion wird (2.137) durch einen Dezimationsfaktor M in Zeitrichtung ergänzt. Die Dezimation dient der Reduktion des Rechenaufwands und der Anzahl der berechneten Merkmale:

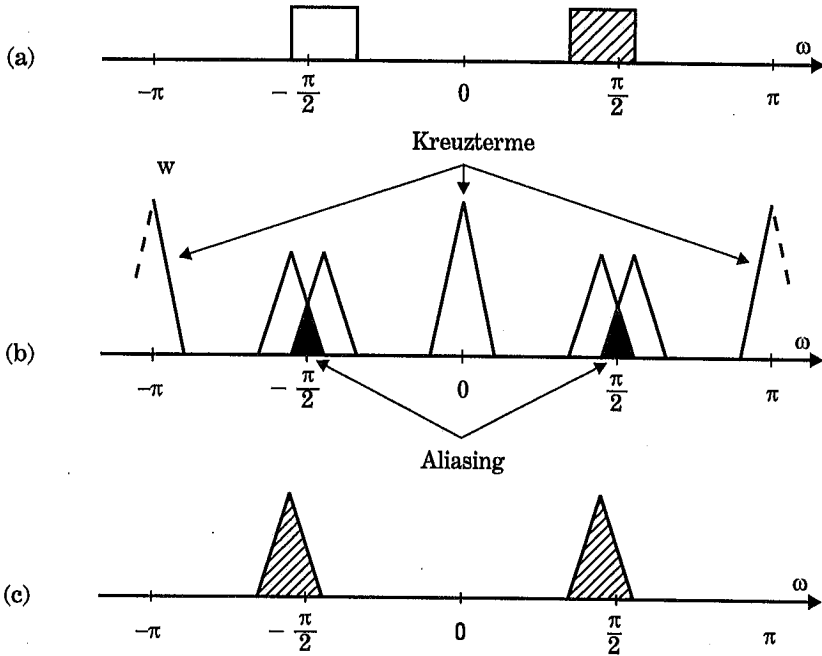


Bild 2.36: Reellwertiges und analytisches (schraffiert) Bandpaßsignal (a), Frequenzbelegung der WV des reellwertigen (b) und des analytischen (c) Signals

$$F^{Cohen}(m, n) = 2 \sum_{k=-K_1}^{K_1} \sum_{l=-K_2}^{K_2} g(l, k) s_a(l + mM + k) s_a^*(l + mM - k) W_N^{nk}. \quad (2.138)$$

Ist die diskrete Kernfunktion $g(l, k)$ bezüglich k konjugiert gerade, d.h.

$$g(l, k) = g^*(l, -k), \quad (2.139)$$

ist das Transformationsergebnis $F^{Cohen}(m, n)$ reell. Denn mit dieser Voraussetzung ist die Größe

$$\tilde{g}(m, k) = \sum_{l=-K_2}^{K_2} g(l, k) s_a(l + mM + k) s_a^*(l + mM - k) \quad (2.140)$$

bezüglich k konjugiert gerade und es folgt für die Darstellung (2.138):

$$F^{Cohen}(m, n) = 2 \sum_{k=-K_1}^{K_1} \tilde{g}(m, k) W_N^{nk}.$$

Mit

$$g'(m, k) = \begin{cases} \tilde{g}(m, k) & \text{für } 0 \leq k < K_1 \\ 0 & \text{für } k = K_1 \\ \tilde{g}(m, k - 2K_1 - 1) & \text{für } K_1 < k \leq 2K_1 + 1 \end{cases}$$

folgt für die Implementierung von (2.138) durch eine diskrete Fourier-Transformation (DFT):

$$F^{Cohen}(m, n) = 2 \sum_{k=0}^{2K_1+1} g'(m, k) W_N^{nk}. \quad (2.141)$$

2.2.8 Interpretation der Merkmale in der Zeit-Frequenz-Ebene

Die Merkmale $F^{Cohen}(m, n)$ lassen sich unter Berücksichtigung der Dezimation einem diskreten Zeit-Frequenz-Raster zuordnen:

$$F^{Cohen}(m, n) \leftrightarrow F(mM\Delta t, n\Delta f).$$

Jedem Element $F(mM\Delta t, n\Delta f)$ ist eine Fläche in der Zeit-Frequenzebene zugeordnet, die den Takt der Elemente $M\Delta t$ und die Diskretisierung der Frequenz-Variablen berücksichtigt.

$$\Delta f = \frac{1}{2N\Delta t} = \frac{f_a}{2N}$$

Die Zuordnung in der Zeit-Frequenz-Ebene ist in Bild 2.37 veranschaulicht.

Die Zeit-Frequenz-Darstellung der Elemente $F^{Cohen}(m, n)$ zeigt in einer groben Näherung, wo die Signalenergie des transformierten Signales bezüglich der Zeit und der Frequenz lokalisiert ist. Die Probleme der Näherung bestehen darin, daß die Merkmale $F^{Cohen}(m, n)$ negativ sein können und daß ggf. die Implementierung durch die endliche Summation in (2.138) unpräzise ist.

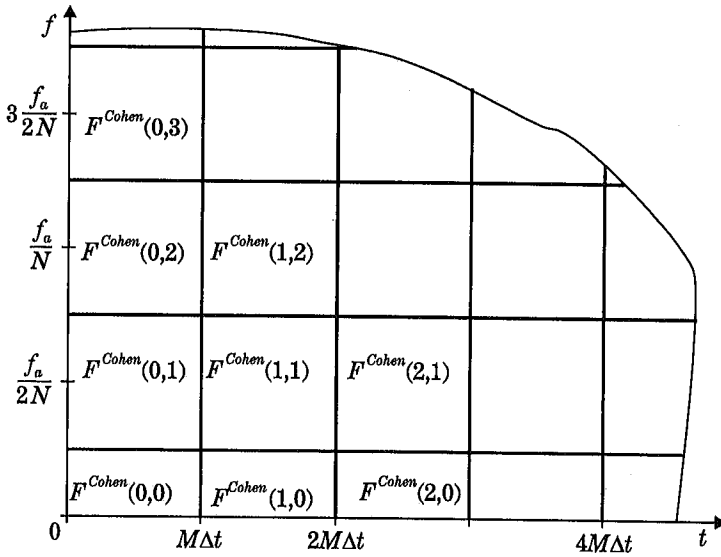


Bild 2.37: Zeit-Frequenz-Belegung durch eine Transformation der Cohen-Klasse

2.3 Modellbasierte Spektralschätzverfahren

Spektralschätzverfahren interpretieren die reellwertige Eingangsfolge $\{s(m)\}_m$ als Realisierung eines stationären Zufallsprozesses $\{s(m)\}_m$ [Kay88]. Ziel der Algorithmen ist die Schätzung der spektralen Leistungsdichte $P_{ss}(e^{j\omega})$, die sich durch Fouriertransformation der Autokorrelationsfunktion $p_{ss}(m)$ ergibt:

$$p_{ss}(l) = E\{s(m)s(m+l)\} \circ \bullet P_{ss}(e^{j\omega}).$$

Modellbasierte Verfahren gehen von der Annahme aus, daß sich der Prozeß durch ein Modell beschreiben läßt. In der Literatur sind zahlreiche Prozeßmodelle beschrieben [Mar87, Kay88]. Es ist zu erwarten, daß sich die betrachteten Signalquellen (Klaviertöne) durch scharfe spektrale Maxima auszeichnen. Für die Merkmalsextraktion wird daher das autoregressive Prozeßmodell (AR-Modell) verwendet. In Abschnitt 2.3.1 werden das Modell und seine Implementierung durch die Levinson-Durbin-Rekursion beschrieben. Aus der Prozeßrealisierung $\{s(k)\}_k$ werden die Parameter des Modells sowie das Leistungsdichtespektrum $P_{ss}(e^{j\omega})$ geschätzt. Voraussetzung für

eine sinnvolle Schätzung ist die Vorgabe einer geeigneten Modellordnung, die in Abschnitt 2.3.3 diskutiert wird. Die Einbindung der geschätzten Modellparameter und des Leistungsdichtespektrums in die Merkmalsbildung ist Gegenstand des Abschnitts 2.3.2.

2.3.1 Das autoregressive Modell

Bei einem reellwertigen autoregressiven Modell geht man von der Annahme aus, der Eingangsprozeß $s(m)$ sei durch die Differenzengleichung

$$s(m) = - \sum_{k=1}^P a(k)s(m-k) + u(m), \quad m \in \mathbb{Z}, a(k) \in \mathbb{R}, \quad (2.142)$$

beschrieben [Kay88], wobei $\{u(m)\}_m$ einen mittelwertfreien, weißen Prozeß mit der Varianz σ_{uu}^2 darstellt. Der zugehörige Signalfuß ist in Bild 2.38 dargestellt.

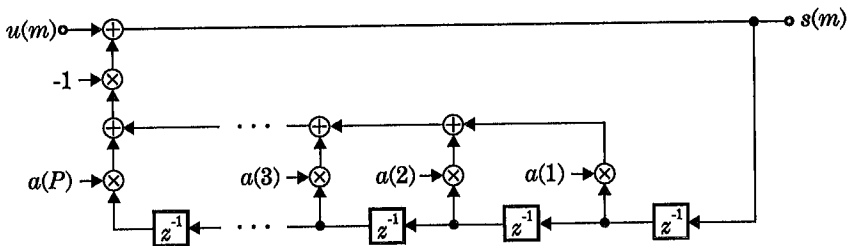


Bild 2.38: Erzeugung des AR-Prozesses nach (2.142) durch ein FIR-Filter

Der Prozeß $\{u(m)\}_m$ dient zur Erregung des linearen Systems und erzeugt den Ausgangsprozeß $\{s(m)\}_m$.

Aus (2.142) läßt sich über die Folge $\{a(k)\}_{k=0}^P$ mit $a(0) = 1$ für die z -Transformierte der Autokorrelationsfolgen

$$P_{ss}(z) \stackrel{z}{\bullet \circ} p_{ss}(k) \quad \text{und} \quad P_{uu}(z) \stackrel{z}{\bullet \circ} p_{uu}(k)$$

wegen $P_{uu}(z) = \sigma_{uu}^2$ der Zusammenhang

$$P_{ss}(z) = \frac{\sigma_{uu}^2}{A(z)A^*(1/z)} \quad (2.143)$$

herleiten [Kay88], wobei $A(z) \stackrel{z}{\bullet} \{a(k)\}_{k=0}^P$ ist. Mit $B(z) = A^*(1/z)$ folgt $B^*(1/z) = A(z)$, d.h. (2.143) folgt auch aus dem Ansatz:

$$s(m) = - \sum_{k=1}^P a(k)s(m+k) + u(m), \quad m \in \mathbb{Z}. \quad (2.144)$$

Beide Ansätze liefern die gleiche spektrale Leistungsdichte. Wertet man (2.143) auf dem Einheitskreis mit $z = e^{j\omega}$ aus, folgt für das Signalleistungsdichtespektrum:

$$P_{ss}(e^{j\omega}) = \frac{\sigma_{uu}^2}{|A(e^{j\omega})|^2}. \quad (2.145)$$

(2.142) kann auch als lineare Prädiktionsaufgabe interpretiert werden, bei der die Zufallsvariable $s(m)$ als zu bestimmende Linearkombination der mittelwertfreien Zufallsvariablen $\{s(m-k)\}_{k=1}^P$ bei minimaler Signalfehlerleistung $P_{uu}(e^{j\omega}) = \sigma_{uu}^2$ ausgedrückt wird [Mar87]. Der Signalfluß ist in Bild 2.39 dargestellt.

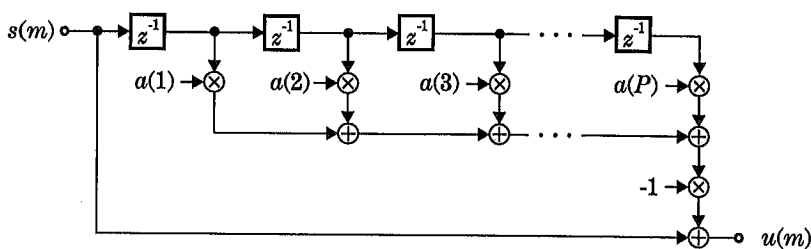


Bild 2.39: Interpretation des AR-Modells nach (2.142) als lineare Prädiktionsaufgabe

Man nennt die in (2.142) formulierte Aufgabe Vorwärtsprädiktion, (2.144) entsprechend Rückwärtsprädiktion.

Die Vorwärts- und Rückwärtsprädiktion führen auf die Levinson-Durbin-Rekursion zur Bestimmung der AR-Koeffizienten $\{a(k)\}_{k=1}^P$ und auf die Interpretation der Zufallsvariablen $\{s(m-k)\}_{k=0}^P$ als Vektoren $\{\bar{s}_{m-k}\}_{k=0}^P$ in einem $P+1$ -dimensionalen Hilbertraum mit dem Innenprodukt [Orf85]:

$$\langle \bar{s}_{m-i}, \bar{s}_{m-k} \rangle = E\{s(m-i)s(m-k)\} = p_{ss}(i-k), \text{ bzw.}$$

$$\langle \bar{\mathbf{s}}_{m-k}, \bar{\mathbf{s}}_{m-k} \rangle = ||\bar{\mathbf{s}}_{m-k}||^2 = p_{ss}(0), \quad k, i = 0, 1, 2, \dots, P.$$

Gesucht sind die Koeffizienten $\{a(k)\}_{k=1}^P$, die die Projektion $\bar{\mathbf{s}}_m^P$ des Vektors $\bar{\mathbf{s}}_m$ auf den von den Vektoren $\{\bar{\mathbf{s}}_{m-k}\}_{k=1}^P$ aufgespannten P -dimensionalen Unterraum durch

$$\bar{\mathbf{s}}_m^P = - \sum_{k=1}^P a(k) \bar{\mathbf{s}}_{m-k} \quad (2.146)$$

mit minimalem quadratischen Fehler bestimmen:

$$||\bar{\mathbf{s}}_m - \bar{\mathbf{s}}_m^P||^2 = ||\bar{\mathbf{u}}_m||^2 = \sigma_{uu}^2. \quad (2.147)$$

Die AR-Koeffizienten sind genau dann optimal gewählt, wenn der Fehlervektor $\bar{\mathbf{u}}_m$ senkrecht auf den Vektoren $\{\bar{\mathbf{s}}_{m-k}\}_{k=1}^P$ steht [Kay88].

Die Levinson-Durbin-Rekursion sei nun anhand der Gram-Schmidt-Orthogonalisierung [Orf85, Dav74] erläutert. Ausgehend von den Vektoren $\{\bar{\mathbf{s}}_{m-k}\}_{k=1}^P$ werden die orthonormalen Basisvektoren $\{\bar{\mathbf{e}}_m^k\}_{k=1}^P$ über die Hilfsvektoren $\{\bar{\mathbf{c}}_m^k\}_{k=1}^P$ durch

$$\bar{\mathbf{c}}_m^1 = \bar{\mathbf{s}}_{m-1}, \quad (2.148)$$

$$\bar{\mathbf{e}}_m^1 = \bar{\mathbf{c}}_m^1 / ||\bar{\mathbf{c}}_m^1||,$$

$$\bar{\mathbf{c}}_m^2 = \bar{\mathbf{s}}_{m-2} - \langle \bar{\mathbf{s}}_{m-2}, \bar{\mathbf{e}}_m^1 \rangle \bar{\mathbf{e}}_m^1, \quad (2.149)$$

$$\bar{\mathbf{e}}_m^2 = \bar{\mathbf{c}}_m^2 / ||\bar{\mathbf{c}}_m^2||,$$

⋮

$$\bar{\mathbf{c}}_m^P = \bar{\mathbf{s}}_{m-P} - \sum_{k=1}^{P-1} \langle \bar{\mathbf{s}}_{m-P}, \bar{\mathbf{e}}_m^k \rangle \bar{\mathbf{e}}_m^k, \quad (2.150)$$

$$\bar{\mathbf{e}}_m^P = \bar{\mathbf{c}}_m^P / ||\bar{\mathbf{c}}_m^P||$$

bestimmt. Es folgt dann für den Vektor $\bar{\mathbf{s}}_m^P$:

$$\begin{aligned} \bar{\mathbf{s}}_m^P &= \sum_{k=1}^P \langle \bar{\mathbf{s}}_m, \bar{\mathbf{e}}_m^k \rangle \bar{\mathbf{e}}_m^k \\ &= \sum_{k=1}^P \frac{\langle \bar{\mathbf{s}}_m, \bar{\mathbf{c}}_m^k \rangle}{||\bar{\mathbf{c}}_m^k||^2} \bar{\mathbf{c}}_m^k. \end{aligned} \quad (2.151)$$

Mit den als Reflexionskoeffizienten bezeichneten Größen

$$r(k) = -\frac{\langle \bar{\mathbf{s}}_m, \bar{\mathbf{c}}_m^k \rangle}{\|\bar{\mathbf{c}}_m^k\|^2} \quad (2.152)$$

folgt aus (2.151):

$$\bar{\mathbf{s}}_m^P = -\sum_{k=1}^P r(k) \bar{\mathbf{c}}_m^k. \quad (2.153)$$

Die Levinson-Durbin-Rekursion berechnet in der Iteration l aus dem AR-Modell der Ordnung $l-1$:

$$\bar{\mathbf{s}}_m^{l-1} = -\sum_{k=1}^{l-1} a_{l-1}(k) \bar{\mathbf{s}}_{m-k}, \quad (2.154)$$

$$\bar{\mathbf{s}}_m^{l-1} = -\sum_{k=1}^{l-1} r(k) \bar{\mathbf{c}}_m^k \quad (2.155)$$

das Modell der Ordnung l , wie im folgenden ausgeführt wird.

In (2.148)-(2.150) sind die Hilfsvektoren $\{\bar{\mathbf{c}}_m^k\}_{k=1}^{l-1}$ aufgeführt, die den $l-1$ -dimensionalen Raum der Vektoren $\{\bar{\mathbf{s}}_{m-k}\}_{k=1}^{l-1}$ aufspannen.

Die Prädiktionsaufgabe (2.144) schreibt sich jetzt in der Form

$$\bar{\mathbf{s}}_{m-l} = -\sum_{k=1}^P a(l-k) \bar{\mathbf{s}}_{m-k} + \bar{\mathbf{u}}_{m-l}.$$

Das Modell der Ordnung $l-1$ wird entsprechend zu (2.154) definiert durch

$$\bar{\mathbf{s}}_{m-l}^{l-1} = -\sum_{k=1}^{l-1} a_{l-1}(l-k) \bar{\mathbf{s}}_{m-k}. \quad (2.156)$$

Die Vektoren $\{\bar{\mathbf{u}}_{m-l}^k\}_{k=1}^{l-1}$ seien entsprechend zu (2.148)-(2.150) mit $i = 2, 3, \dots, l-1$ definiert:

$$\bar{u}_{m-l}^1 = \bar{s}_{m-l+1}, \quad (2.157)$$

$$\bar{u}_{m-l}^i = \bar{s}_{m-l+i} - \sum_{k=1}^{i-1} \frac{\langle \bar{s}_{m-l+k}, \bar{u}_{m-l}^k \rangle}{\|\bar{u}_{m-l}^k\|} \bar{u}_{m-l}^k. \quad (2.158)$$

$$(2.159)$$

Für den Iterationschritt l ergibt sich der Vektor $\bar{u}_{m-l}^l$ zu

$$\bar{u}_{m-l}^l = \bar{s}_{m-l} - \sum_{k=1}^{l-1} \frac{\langle \bar{s}_{m-l}, \bar{u}_{m-l}^k \rangle}{\|\bar{u}_{m-l}^k\|^2} \bar{u}_{m-l}^k. \quad (2.160)$$

Der Summenterm in (2.160) stellt die beste Approximation an den Vektor $\bar{s}_m$ aus den Vektoren $\{\bar{s}_{m-k}\}_{k=1}^{l-1}$ dar. Da dies auch auf den Summenterm in (2.154) zutrifft, entspricht der Fehler in der Rekursion $l-1$ der Vorwärtsprädiktion gerade dem Vektor $\bar{u}_{m-l}^l$:

$$\bar{u}_{m-l}^l = \bar{s}_m - \bar{s}_{m-l}^{l-1}.$$

Umgekehrt ist der Vektor $\bar{e}_m^l$ gerade der Fehler der Rückwärtsprädiktion des Modells der Ordnung $l-1$:

$$\bar{e}_m^l = \bar{s}_{m-l} - \bar{s}_{m-l}^{l-1}. \quad (2.161)$$

Dieser Zusammenhang zwischen Vorwärts- und Rückwärtsprädiktion wird in der Levinson-Durbin-Rekursion ausgenutzt.

Ausgehend von (2.155) werden ein von den Koeffizienten $\{a_{l-1}(k)\}_{k=1}^{l-1}$ abhängiger Ausdruck und durch Koeffizientenvergleich die AR-Koeffizienten $\{a_l(k)\}_{k=1}^l$ ermittelt:

$$\begin{aligned} \bar{s}_m^l &= - \sum_{k=1}^l r(k) \bar{e}_m^k \\ &= -r(l) \bar{e}_m^l - \sum_{k=1}^{l-1} r(k) \bar{e}_m^k \\ &= -r(l) \bar{e}_m^l - \sum_{k=1}^{l-1} a_{l-1}(k) \bar{s}_{m-k} \end{aligned}$$

$$= -r(l) \left[\bar{s}_{m-l} - \sum_{k=1}^{l-1} \frac{\langle \bar{s}_{m-l}, \bar{c}_m^k \rangle}{\|\bar{c}_m^k\|^2} \bar{c}_m^k \right] - \sum_{k=1}^{l-1} a_{l-1}(k) \bar{s}_{m-k}.$$

Wegen (2.161) und (2.156) folgt jetzt:

$$\begin{aligned} \bar{s}_m^l &= -r(l) \left[\bar{s}_{m-l} + \sum_{k=1}^{l-1} a_{l-1}(l-k) \bar{s}_{m-k} \right] - \sum_{k=1}^{l-1} a_{l-1}(k) \bar{s}_{m-k} \\ &= -\sum_{k=1}^{l-1} [a_{l-1}(k) + r(l)a_{l-1}(l-k)] \bar{s}_{m-k} - r(l) \bar{s}_{m-l}. \end{aligned} \quad (2.162)$$

Es folgt daher für die rekursive Bestimmung der Koeffizienten $\{a_l(k)\}_{k=1}^l$ über die Koeffizienten $\{a_{l-1}(k)\}_{k=1}^{l-1}$ bzw. $r(l)$:

$$a_l(k) = a_{l-1}(k) + r(l)a_{l-1}(l-k), \quad k = 1, 2, \dots, l-1, \quad (2.163)$$

$$a_l(l) = r(l). \quad (2.164)$$

Für die Rekursion l muß lediglich der Reflexionskoeffizient $r(l)$ aus (2.152) bestimmt werden. Es ist [Kay88]:

$$r(l) = -\frac{\langle \bar{s}_m, \bar{c}_m^l \rangle}{\|\bar{c}_m^l\|^2} = -\frac{p_{ss}(l) + \sum_{k=1}^{l-1} a_{l-1}(k)p_{ss}(l-k)}{\|\bar{c}_m^l\|^2}.$$

Für die Norm des Vektors $\bar{u}_{m-l}^l$ kann man zeigen [Kay88], daß gilt

$$\begin{aligned} \|\bar{u}_{m-l}^l\|^2 &= \|\bar{c}_m^l\|^2 \\ &= \|\bar{u}_{m-l}^{l-1}\|^2 (1 - |r(l-1)|^2). \end{aligned} \quad (2.165)$$

Also folgt für den Reflexionskoeffizienten $r(l)$:

$$r(l) = -\frac{p_{ss}(l) + \sum_{k=1}^{l-1} a_{l-1}(k)p_{ss}(l-k)}{\|\bar{u}_{m-l}^l\|^2}. \quad (2.166)$$

Die Gleichungen (2.163), (2.164), (2.165) und (2.166) erlauben den Aufbau eines AR-Modells der Ordnung l aus dem der Ordnung $l - 1$. Zusammen mit den Startwerten

$$a_1(1) = r(1) = -\frac{p_{ss}(1)}{p_{ss}(0)} \quad \text{und} \quad (2.167)$$

$$\|\bar{u}_{m-l}^2\|^2 = p_{ss}(0)(1 - |r(1)|^2) \quad (2.168)$$

bilden sie die Levinson-Durbin-Rekursion, die mit dem Aufbau des Modells der Ordnung P endet.

2.3.2 Schätzung der Modellparameter

Zur Schätzung der AR-Koeffizienten der Ordnung P stehe die Realisierung $\{s'(k)\}_{k=0}^{L-1}$ mit $L > P$ zur Verfügung. In [Bur67] wird ein Verfahren (Burg-Algorithmus) vorgeschlagen, das mit Hilfe der Levinson-Durbin-Rekursion die AR-Parameter schätzt.

Man beachte, daß in (2.167), (2.168) und (2.166) Schätzungen für die Autokorrelationsfolge $\{p_{ss}(l)\}_l$ eingesetzt werden müssen. Der Burg-Algorithmus umgeht die Schätzung der Autokorrelationsfolge und schätzt die Reflexionskoeffizienten und damit die AR-Koeffizienten direkt. Für eine Beschreibung des Burg-Algorithmus wird auf die Literatur verwiesen [Kay88, Böh93].

Die Eingangsfolge $\{s(k)\}_k$ wird für die Merkmalsextraktion in Teilfolgen $\{s'_m(k)\}_{k=0}^{L-1}$ durch

$$\{s'_m(k)\}_{k=0}^{L-1} = \{s(k + Mm)\}_{k=0}^{L-1} \quad (2.169)$$

zerlegt. Der Faktor M bestimmt den Versatz der durch den Parameter m gekennzeichneten Abschnitte $\{s'_m(k)\}_{k=0}^{L-1}$. Der Abschnitt $\{s'_m(k)\}_{k=0}^{L-1}$ wird einer Schätzung durch den Burg-Algorithmus unterzogen. Der Algorithmus liefert die Schätzwerte für die AR-Koeffizienten $\{\hat{a}_m(k)\}_{k=0}^P$, die Reflexionskoeffizienten $\{\hat{r}_m(k)\}_{k=0}^P$ sowie $\hat{\sigma}_m^2$ für die Varianz des weißen Rauschens (vergl. [Mar87, Seite 208]). Mit $\{\hat{a}_m(k)\}_{k=0}^P$ und (2.145) kann das Leistungsdichtespektrum $P_{ss}(e^{j\omega})$ geschätzt werden. Wird das geschätzte Spektrum diskret abgetastet mit $\omega_n = 2\pi n/N$, folgt:

$$\hat{P}_{ss}^m(e^{j2\pi n/N}) = \frac{\hat{\sigma}_m^2}{\sum_{k=0}^{P-1} |\hat{a}_m(k)e^{-j2\pi n/N}|^2}, \quad n = 0, 1, 2, \dots, N-1, \quad (2.170)$$

wobei der Index m die Zugehörigkeit zu der Folge $\{s'_m(k)\}_{k=0}^{K-1}$ ausdrückt.

In der Merkmalsextraktion werden zwei Merkmalsmengen aus dem Schätzvorgang abgeleitet. Zum einen wird das diskret abgetastete Leistungsdichtespektrum als Merkmalssatz ermittelt:

$$F^{AR,S}(m, n) = \hat{P}_{ss}^m(e^{j2\pi n/N}), \quad n = 0, 1, 2, \dots, N/2. \quad (2.171)$$

Zum anderen besteht die Möglichkeit, die geschätzten Reflexionskoeffizienten als Merkmale zu verwenden:

$$F^{AR,R}(m, n) = \hat{r}_m(n+1), \quad n = 0, 1, 2, \dots, P-1. \quad (2.172)$$

2.3.3 Wahl der Modellordnung

Die Wahl der richtigen Modellordnung P ist ein kritischer Punkt bei der Modellierung. Wird die Modellordnung zu klein angesetzt, ist die Schätzung des Leistungsdichtespektrums zu grob. Bei zu großer Modellordnung neigen die Modelle zur Erzeugung real nicht vorhandener spektraler Spitzen [Kay88, Mar87].

Liegt der Messung ein echter AR-Prozeß der Ordnung P zugrunde, sind die Reflexionskoeffizienten $r(l)$ für $l > P$ Null und die Leistung des Vorhersagefehlers σ_{uu} ändert sich nicht mehr. Daher beruhen fast alle in der Literatur angegeben Verfahren zur Schätzung der Modellordnung auf der Auswertung des Vorhersagefehlers.

Bei der Merkmalsextraktion wird für alle Abschnitte $\{s'_m(k)\}_{k=0}^{L-1}$ die gleiche Modellordnung verwendet. Ausgehend von einer Realisierung $\{s'_m(k)\}_{k=0}^{L-1}$ der Länge L hat Akaike zwei Verfahren vorgeschlagen [Mar87]. Für jedes Kriterium wird die Modellordnung P_{min} gewählt, die die folgenden Entscheidungsgrößen minimiert:

- FPE-Kriterium (*Finale Prediction Error*):

$$FPE(P_{min}) = \min_P \frac{L+P}{L-P} \hat{\sigma}_m^2, \quad (2.173)$$

- AIC (*Akaike Information Criterion*):

$$AIC(P_{min}) = \min_P \{ L \ln(\hat{\sigma}_m^2) + 2P \}. \quad (2.174)$$

Die Anwendbarkeit der Verfahren zur Schätzung der Modellordnung muß bei realen Daten stets überprüft werden. Erfahrung und subjektives Urteil mögen ebenso als Kriterium dienen [Mar87]. Kay schlägt als Faustformel daher alternativ

$$L/3 < P < L/2 \quad (2.175)$$

als für den praktischen Einsatz sinnvolle Wahl vor [Kay88].

Eine dritte Möglichkeit ergibt sich aus dem Umstand, daß die Zahl der Nullstellen in dem Polynom $A(z)$ in guter Näherung die Anzahl spektraler Spitzen in dem Leistungsdichtespektrum beschreibt. Vaseghi schlägt in [Vas96] deshalb für ein reellwertiges Signal vor, die Modellordnung doppelt so groß zu wählen wie die Anzahl erwarteter spektraler Maxima im Intervall $[0, \pi)$.

Die Wahl der Modellordnung beeinflusst auch die Varianz der Schätzung des Leistungsdichtespektrums. In [Kay88, S. 194] wird darauf hingewiesen, daß die Varianz linear von dem Quotienten P/L abhängt.

2.3.4 Interpretation der Merkmale in der Zeit-Frequenz-Ebene

Die Merkmale $F^{AR,S}(m, n)$ beschreiben die Änderung des geschätzten Leistungsdichtespektrums über der Zeit. Für die Visualisierung wird die gleiche Rasterung der Zeit-Frequenz-Ebene wie bei der Kurzzeit-Fourier-Transformation verwendet (Bild 2.34).

Für die Merkmale $F^{AR,R}(m, n)$ gibt es keine Interpretation in der Zeit-Frequenz-Ebene.

3 Merkmalsselektion und Klassifikation

Nach der Merkmalsextraktion enthält die Merkmalsmatrix $F(m, n)$ eine solche Fülle von Daten, daß zunächst durch eine Merkmalsselektion eine Auswahl vorgenommen werden muß. Ziel der Selektion kann die Entfernung von Merkmalen sein, die statistisch von anderen abhängig sind oder keinen Beitrag zur Unterscheidung der Klassen leisten. Für eine Diskussion der aus diesen Forderungen abgeleiteten Verfahren wird auf [Nie74] verwiesen. Bei der vorliegenden Klassifikationsaufgabe erlaubt die geringe Menge an Signalen keine relevanten Schlüsse auf die statistischen Merkmalseigenschaften.

Deshalb wird ein heuristischer Ansatz gewählt. In Abschnitt 3.1 wird ein klassenunabhängiges Kriterium vorgestellt, das die Merkmale nach der Größe ihrer Beträge auswählt. Für Filterbänke mit Baumstruktur wird eine klassenabhängige Merkmalsselektion in Abschnitt 3.2 betrachtet. Die Merkmalsselektion für orthogonale Filterbänke mit Baumstruktur setzt nach den Ausführungen in Abschnitt 2.1.3.1 zunächst die Bestimmung einer geeigneten Basis voraus. In den folgenden Abschnitten werden für beide Selektionsansätze Verfahren zur Bestimmung geeigneter Basen vorgestellt.

Für den Klassifikator besteht die Aufgabe, bei einem empfangenen Muster $\mathbf{P}$, das die bereits einer Selektion unterzogene Merkmalsmenge $F(m, n)$ repräsentiert, auf die richtige Klassenzugehörigkeit c zu schließen.

Bei einer Klassifikationsaufgabe mit C Klassen soll c einem Klassenindex $c \in \{1, 2, 3, \dots, C\}$ entsprechen. Es ist sinnvoll, zusätzlich eine Rückweisklasse mit dem Index $C + 1$ einzuführen, um unsichere Entscheidungen zu vermeiden.

Diese Aufgabe läßt sich als statistisches Entscheidungsproblem formulieren [Mer96, Kro96, Sch77, Sch96, DH73], wobei drei Fälle unterschieden werden können [FH51]:

1. Die Verbunddichte $p(\mathbf{P}, c)$ zwischen dem Muster $\mathbf{P}$ und der tatsächlichen Klassenzugehörigkeit c ist bekannt. Die optimale Lösung läßt sich auf analytischem Weg allgemein durch das Bayes-Kriterium angeben [Mer96].
2. Die Verbunddichte $p(\mathbf{P}, c)$ liegt als parametrierbares Modell vor. Anhand von Beispielmustern, deren korrekte Klassenzuordnung bekannt ist, lassen sich die Parameter der Modelldichte schätzen [Kro96].
3. Über die Verbunddichte $p(\mathbf{P}, c)$ liegt keine Information vor. Eine auf-

wendige Möglichkeit besteht darin, mit korrekt klassifizierten Beispielmustern durch Histogrammbildung die Dichtefunktion zu approximieren. Alternativ wird in [FH51] und [CH67] die Nächster-Nachbar-Regel als Klassifikationsverfahren vorgeschlagen. Dabei wird ein Muster $\mathbf{P}$ derjenigen Klasse zugeordnet, zu deren Repräsentanten es den geringsten Abstand aufweist: Dazu liegen in einer Datenbank N_M klassifizierte Tupel der Form $\{\mathbf{P}_i, c_i\}_{i=1}^{N_M}$ vor. Der nächste Nachbar $\mathbf{P}_n \in \{\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_{N_M}\}$ wird über die Metrik d bestimmt:

$$\min_{1 \leq i \leq N_M} d(\mathbf{P}, \mathbf{P}_i) = d(\mathbf{P}, \mathbf{P}_n).$$

Das Muster $\mathbf{P}$ wird der Klasse c_n des nächsten Nachbarn $\mathbf{P}_n$ zugeordnet. Dieses Verfahren ist praktikabel, wenn der Umfang der Datenbank nicht zu umfangreich ist. Für eine Beschreibung der Leistungsfähigkeit des Nächster-Nachbar-Klassifikator und seinen Zusammenhang zum Bayes-Klassifikator wird auf die Literatur verwiesen [CH67].

Für die vorliegende Klassifikationsaufgabe wird der Nächster-Nachbar-Klassifikator verwendet. Auf Vergleiche mit anderen Klassifikatorkonzepten wird verzichtet, da in dieser Arbeit vor allem die Güten der durch die verschiedenen Verfahren ermittelten Merkmale verglichen werden sollen.

Folgende Vereinbarungen werden getroffen:

- Die zur Erzeugung der Muster für die Datenbank erforderlichen Signale werden mit $\{s^{c,l}(m)\}_m$ bezeichnet, wobei c den Klassenindex und l einen Zählindex darstellen.
- Für die c -te Klasse stehen L_c Signale für die Erstellung der Datenbank zur Verfügung, d.h. für Klasse c ist $l \in \{1, 2, \dots, L_c\}$. Damit gilt für die Gesamtzahl der Muster der Datenbank

$$N_M = \sum_{c=1}^C L_c.$$

- In der Klasse c stehen insgesamt N_c Signale für Datenbank und Durchführung der Klassifikation zur Verfügung.

3.1 Klassenunabhängige Merkmalsselektion und Klassifikation

Lediglich bei Filterbänken mit Baumstruktur wird eine Verschiebung der Eingangsfolge durchgeführt. Um eine gemeinsame Formulierung der Musterbildung und Klassifikation unabhängig vom Merkmalsextraktionsverfahren zu erreichen, wird für alle anderen Verfahren die Verschiebungsmenge $\mathbf{V} = \{0\}$ gewählt. Für die klassenunabhängige Merkmalsselektion und Klassifikation wird ein Verarbeitungskonzept verwendet, dessen Signalfluß in Bild 3.1 dargestellt ist.

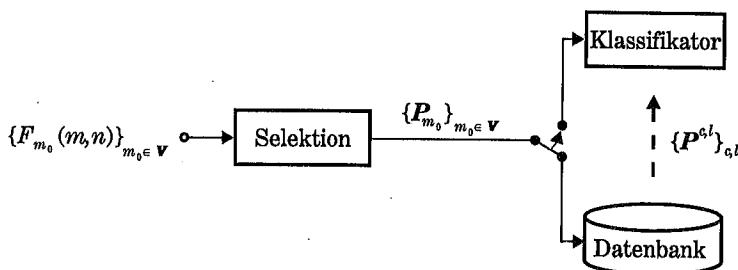


Bild 3.1: Klassenunabhängige Merkmalsselektion und Klassifikation

Das aktuelle Muster wird einer Merkmalsselektion unterzogen. Während der Adaptionsphase werden die Muster zu den Signalen $\{s^{c,l}(m)\}_m$ in der Datenbank abgelegt und bilden den Musterpool $\{\mathbf{P}^{c,l}\}_{c,l}$ mit $c = 1, 2, \dots, C$ und $l = 1, 2, \dots, L_c$. In der Anwendungsphase werden die Muster einer Nächster-Nachbar-Klassifikation unterzogen.

Abhängig vom Merkmalsextraktionsverfahren werden nicht notwendig alle von Merkmalsextraktionsmethoden ermittelten Merkmale $F(m, n)$ verwendet. Die Menge N beschreibt die Menge der Indizes n , die bei der Musterbildung berücksichtigt werden.

Für die Merkmale der Kurzzeit-Fourier-Transformation (2.116) gilt bei einer geradzahlgigen Frequenzdiskretisierung N (2.110):

$$N^{KFT} = \{0, 1, 2, \dots, N/2\},$$

da reellwertige Signale verarbeitet werden.

Für die Wavelet-Transformation mit der Zerlegungstiefe T gilt (2.55, 2.56):

$$N^{DWT} = \{0, 1, 2, \dots, T\}.$$

Für alle Verfahren der Cohen-Klasse gilt gemeinsam (2.138) mit dem Parameter der Frequenzdiskretisierung N (2.136):

$$N^{Cohen} = \{0, 1, 2, \dots, N\}.$$

Bei den modellbasierten Spektralschätzverfahren gilt für die Parameter nach (2.171)

$$N^{AR,S} = \{0, 1, 2, \dots, N/2\}$$

und nach (2.172)

$$N^{AR,R} = \{0, 1, 2, \dots, P\}$$

mit P als Modellordnung.

Für eine Filterbank mit Baumstruktur wird die Menge N^B durch den Algorithmus der besten Basis, dargestellt in Abschnitt 3.1.1, bestimmt. Musterbildung und Klassifikation sind Gegenstand des Abschnitts 3.1.2.

3.1.1 Der Algorithmus der besten Basis

Als Kostenfunktion zur Bewertung der Optimalität der Basis wird das in der Informationstheorie eingeführte Maß der Entropie einer Ereignismenge auf die Entwicklungskoeffizienten bezüglich einer Basis übertragen [CW92]. Dieses Maß sei anhand der Entwicklungskoeffizienten bezüglich der Basis des Raumes R_i^j nach Abschnitt 2.1.3.1. eingeführt (vergleiche auch (2.92)):

$$M(\{w_i^j(m)\}_m) = - \sum_{m=-\infty}^{\infty} |w_i^j(m)|^2 \ln |w_i^j(m)|^2 \quad (3.1)$$

und

$$M(w_i^j(m)) = \begin{cases} -|w_i^j(m)|^2 \ln |w_i^j(m)|^2 & \text{für } |w_i^j(m)| \neq 0 \\ 0 & \text{sonst} \end{cases} \quad (3.2)$$

Das Maß hat die Eigenschaft:

$$M(\{w_l^i(m)\}_m) = \sum_{m=-\infty}^{\infty} M(w_l^i(m)). \quad (3.3)$$

Unter der Annahme, daß

$$\sum_{m=-\infty}^{\infty} |w_l^i(m)|^2 = 1$$

gilt, kann man die Größe $|w_l^i(m)|^2$ als Wahrscheinlichkeit dafür interpretieren, daß das Signal $\{w_l^i(m)\}_m$ in dem durch den Vektor $\bar{b}_{l,m}^i$ aufgespannten Raum liegt. Für die Unterraumzerlegung nach (2.93):

$$R_l^i = R_{lM}^{i+1} \oplus R_{lM+1}^{i+1} \oplus \dots \oplus R_{lM+M-1}^{i+1}$$

bezeichnet die Summe

$$\sum_{m=-\infty}^{\infty} |w_{lM+b}^{i+1}(m)|^2$$

die Wahrscheinlichkeit, daß das Signal in dem Unterraum R_{lM+b}^{i+1} mit $b \in \{0, 1, 2, \dots, M-1\}$ liegt.

Der Algorithmus der besten Basis ermittelt unter allen durch die Filterbank vermittelten Basen diejenige, bezüglich der das Entropiekriterium in (3.1) für eine Basis des Raumes R_0^0 minimal wird. Der Algorithmus läßt sich durch die in Abschnitt 2.1.3.1 im Bild 2.15 verwendete Graphendarstellung erläutern. Jeder Knoten repräsentiert einen Unterraum R_l^i . Zu Beginn werden alle Endknoten eines Baumes der Zerlegungstiefe T markiert, die Entropie der Entwicklungskoeffizienten berechnet und den Größen

$$K_l^i = M(\{w_l^i\}_m) \quad (3.4)$$

mit

$$i = 0, 1, 2, \dots, T-1, \quad l(i) = 0, 1, 2, \dots, M^i - 1$$

zugeordnet. Anschließend werden, beginnend mit $i = T-1$, die Entropien der Signalentwicklungen in den Räumen R_l^i mit $l = 0, 1, 2, \dots, M^i - 1$ mit

denen der eigenen Nachfolgerknoten R_{Ml+b}^{i+1} mit $b = 0, 1, 2, \dots, M-1$ verglichen. Der Knoten zu dem Raum R_l^i wird markiert und die Markierung der Nachfolgerknoten werden gelöscht, wenn gilt:

$$K_l^i < \sum_{b=0}^{M-1} K_{lM+b}^{i+1}. \quad (3.5)$$

Im anderen Fall unterbleibt eine Markierung und die Kosten der Nachfolger werden übernommen:

$$K_l^i = \sum_{b=0}^{M-1} K_{lM+b}^{i+1}.$$

Für jeden Durchgang wird i um eins verringert, bis $i = 0$ erreicht und jeder Knoten des Graphen einer Bewertung unterzogen worden ist. Man beachte, daß der Vergleich nach (3.5) erlaubt ist, da die Unterraumzerlegung orthogonal ist und das Bewertungsmaß (3.1,3.2) die Eigenschaft (3.3) erfüllt.

Die ermittelte Markierung des Graphen stellt die beste Basis für den Raum R_0^0 dar. In der Zeit-Frequenz-Ebene kann die beste Basis als zu derjenigen orthogonalen Transformation gehörig interpretiert werden, die die Zeit-Frequenz-Lokalisierung des Signals optimiert. Der Menge $\mathcal{N}^B$ werden alle Indizes $n(i, l)$ (2.87) zugeordnet, die die Entwicklungskoeffizienten $\{w_l^i(m)\}_m$ bezüglich der gewählten Basis beschreiben.

Wird die Filterbank für eine Menge von Verschiebungen $\mathbb{W}$ nach Abschnitt 2.1.3.2 ausgelegt, so kann die beste Basis für jede Verschiebung $m_0 \in \mathbb{W}$ bestimmt werden. In Bild 2.24 ist der zugehörige Graph dargestellt. Nun werden die Kosten K_l^i entsprechend (3.4) und (2.109) durch

$$K_l^i = M(\{w_l^{i, \tilde{m}_0(i)}\}_m)$$

bestimmt. Der Ablauf des Algorithmus entspricht dem im unverschobenen Fall und erlaubt nach erfolgtem Durchlauf die Bestimmung der Menge $\mathcal{N}_{m_0}^B$, die die Indizes $n(i, l)$ der markierten Knoten für die Verschiebung m_0 enthält.

Es wird auf die Möglichkeit hingewiesen, für eine bestimmte Basis auch signalangepaßte Filter zu verwenden, z.B. [Gre96], d.h. durch eine geeignete Wahl der Kostenfunktion (2.30). In der vorliegenden Arbeit wird darauf unter der Annahme verzichtet, daß die Wahl der Basis einen größeren Einfluß auf die Klassifikationsergebnisse hat als die verwendeten Filter. Im Abschnitt 5.1.3 wird bei der Diskussion der Ergebnisse auf die Zulässigkeit dieser Annahme eingegangen.

3.1.2 Musterbildung und Klassifikation

Im folgenden wird bei den Indexmengen und bei den Merkmalen die Kennzeichnung des Zeit-Frequenz-Analyseverfahrens weggelassen, da auf dieser Ebene der Merkmalsextraktionsalgorithmus unwesentlich ist.

Die Abhängigkeit der Indexmenge M von der Verschiebung m_0 ist nur bei Filterbänken mit Baumstruktur gegeben. Für eine gemeinsame Darstellung werden in der Musterbildung allgemein die Merkmale $F_{m_0}(m, n)$ in Verbindung mit der Menge M_{m_0} auf Muster $P_{m_0}(m, n)$ abgebildet.

In Abschnitt 3.1.2.1 wird die Mustererzeugung erläutert und deren Einbindung in den Klassifikationsprozeß folgt in Abschnitt 3.1.2.2.

3.1.2.1 Musterbildung bei Adaption und Klassifikation

Eine einfache Merkmalsselektion besteht darin, die betragsgrößten Merkmale der vorliegenden Merkmalsmenge $F_{m_0}(m, n)$ zu ermitteln und als Muster weiterzuverarbeiten. Es werden damit die Elemente verwendet, durch die das Signal in der Zeit-Frequenz-Ebene lokalisiert wird. Die Musterbildung erfolgt segmentweise mit der Segmentlänge V und der Schwelle S :

$$P'_{m_0}(mV + k, n) = \begin{cases} F_{m_0}(mV + k, n) & \text{wenn } \begin{cases} E_{m,V} > S \\ \text{und} \\ n \in M_{m_0} \end{cases} \\ 0 & \text{sonst} \end{cases} \quad (3.6)$$

mit

$$E_{m,V} = \sum_{v=0}^{V-1} |F_{m_0}(mV + v, n)|^2$$

als Segmentenergie für $k = 0, 1, 2, \dots, V - 1$, $m_0 \in \mathbb{W}$ und $n, m \in \mathbb{Z}$. Damit für jedes Verfahren dieselbe Anzahl ausgewählter Segmente B vorliegt, wird die Schwelle S adaptiv gewählt. In der Implementierung ist dafür eine Sortierung der Segmentenergien nach der Größe notwendig. Je größer V gewählt wird, um so schneller ist die Sortierung bzw. größer die Maximasuche. Abschließend werden die Muster normiert durch

$$P_{m_0}(m, n) = \frac{P'_{m_0}(m, n)}{\|P'_{m_0}\|} \quad (3.7)$$

mit der Musternorm:

$$\|\mathbf{P}'\| = \sqrt{\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} |P'(m, n)|^2}.$$

Alternativ zu (3.7) wird bei einzelnen Verfahren auch der Betrag der Merkmale zu Musterbildung herangezogen, d.h.:

$$P_{m_0}(m, n) = \frac{|P'_{m_0}(m, n)|}{\|\mathbf{P}'_{m_0}\|}. \quad (3.8)$$

3.1.2.2 Klassifikation

Während der Adaptionsphase werden die Trainingssignale $\{s^{c,l}(m)\}_m$ auf Merkmale $P_{m_0}^{c,l}(m, n)$ abgebildet. Die Muster der Datenbank $P_0^{c,l}(m, n)$ werden über (3.6) für die Verschiebung $m_0 = 0$ gebildet.

In der Anwendungsphase wird die aktuelle Mustermenge $P_{m_0}(m, n)$ für $m_0 \in \mathbb{W}$ mit den Mustern der Datenbank verglichen. Anstelle der Darstellung $P_{m_0}(m, n)$ wird die Matrixschreibweise $\mathbf{P}_{m_0}$, bzw. $\mathbf{P}^{c,l}$ für $P_0^{c,l}(m, n)$, verwendet. Der Klassifikator entscheidet sich für die Klasse c_1 , die einen Repräsentanten l_1 mit minimalem euklidischen Abstand d^{c_1, l_1} zu einem der Muster $\{P_{m_0}(m, n)\}_{m_0 \in \mathbb{V}}$ aufweist. Es ist:

$$d^{c_1, l_1} = \min_{1 \leq c \leq C} \min_{1 \leq l \leq L_c} \min_{m_0 \in \mathbb{V}} \|\mathbf{P}_{m_0} - \mathbf{P}^{c, l}\|^2.$$

Ein Muster wird zurückgewiesen, wenn der Abstand d^{c_2, l_2} zu dem nächstgelegenen Repräsentanten einer anderen Klasse:

$$d^{c_2, l_2} = \min_{1 \leq c \leq C, c \neq c_1} \min_{1 \leq l \leq L_c} \min_{m_0 \in \mathbb{V}} \|\mathbf{P}_{m_0} - \mathbf{P}^{c, l}\|^2$$

sich um weniger als Δd_{min} von d^{c_1, l_1} unterscheidet, d.h.

$$\text{Rückweisung, wenn: } d^{c_2, l_2} - d^{c_1, l_1} \leq \Delta d_{min}.$$

Dieses Rückweisungskriterium wird in [Sch96] als Differenz-Kriterium (DIF-Kriterium) bezeichnet.

3.2 Klassenabhängige Merkmalsselektion und Klassifikation

Das Konzept der klassenabhängigen Merkmalsselektion wird nur für die Merkmale aus einer Filterbank mit Baumstruktur betrachtet. In [Sai94, SC94] wird ein Ansatz vorgeschlagen, um den Algorithmus der besten Basis aus Abschnitt 3.2.1 zur Ermittlung der Unterschiede zwischen den Merkmalen zu ermitteln. In Bild 3.2 ist das Konzept aus Merkmalsselektion und Klassifikation dargestellt. Das aktuelle Muster wird einer klas-

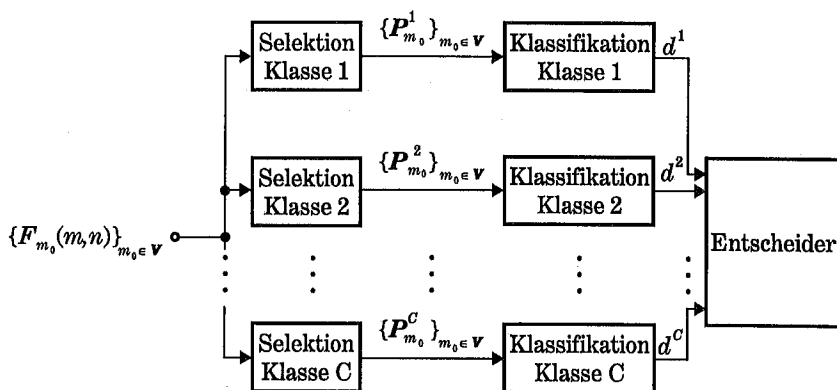
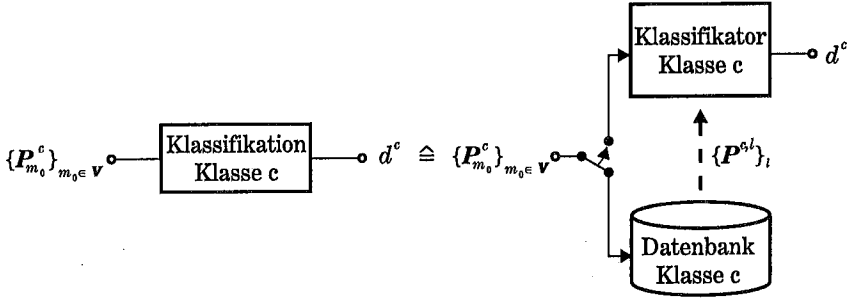


Bild 3.2: Klassenabhängige Merkmalsselektion und Klassifikation

senabhängigen Merkmalsselektion und einer klassenspezifischen Nächster-Nachbar-Klassifikation unterzogen. Jeder Klassifikator, im Signalfluß dargestellt in Bild 3.3, gibt den ermittelten minimalen Abstand an einen Entscheiderblock weiter, der die eigentliche Klassifikation durchführt. Dabei enthält die Datenbank zum Klassifikator der Klasse c' nur die Muster $P_0^{c',l}(m, n)$ mit $l = 1, 2, \dots, L_{c'}$.

In Abschnitt 3.2.1 wird der Algorithmus zur besten Basis für klassenabhängige Merkmalsselektion erläutert. Seine Einbindung in die Musterbildung und Klassifikation ist Gegenstand des Abschnitts 3.2.2.

Bild 3.3: Merkmalsselektion und Klassifikation für Klasse c

3.2.1 Die beste Basis für klassenabhängige Merkmalsselektion

Die Idee des Algorithmus sei an einem Beispiel erläutert. Es seien $\{s_1(m)\}_m$ und $\{s_2(m)\}_m$ zwei Folgen sowie $F^1(m, n)$ und $F^2(m, n)$ ihre Merkmalsrepräsentationen. Dabei sei vorausgesetzt, daß die Merkmale beider Signale einem Filterbaum mit derselben Zerlegungstiefe T entstammen. Die Aufgabe bestehe darin, ihren Abstand bezüglich der $l^2(\mathbf{Z})$ -Norm durch eine Abbildung auf wenige Merkmale zu konzentrieren.

Die Lösung besteht darin, die Differenzfolge $\{s_1(m) - s_2(m)\}_m$ mit der Merkmalsrepräsentation:

$$\Delta F(m, n) = F^1(m, n) - F^2(m, n)$$

dem Algorithmus der besten Basis zu unterwerfen.

Dieser Ansatz kann auf zwei Klassen von Signalen erweitert werden, indem die Differenzsignale zwischen jedem Signal der einen Klasse und jedem der anderen Klasse ermittelt und durch geeignete Mittelung der Differenzrepräsentationen eine beste Basis zur Unterscheidung der beiden Klassen errechnet wird. Im folgenden wird die beste Basis für die Unterscheidung der Klasse c von der Restklasse, die alle anderen Klassen enthält, bestimmt. Diese Restklasse wird durch $\bar{c}$ gekennzeichnet. Es sind $F_{m_0}^{\bar{c}, l}(m, n)$ die zugehörigen Merkmale mit $l \in \{1, 2, \dots, L_{\bar{c}}\}$ und $L_{\bar{c}} = N_M - L_c$. Auch hier erfolgt die Merkmalsbildung mit der gemeinsamen Zerlegungstiefe T .

Die Differenzbildung soll im folgenden im Merkmalsbereich vorgenommen werden. Die Größe

$$\Delta F^c(m, n) = \sum_{l_1=1}^{L_c} \sum_{l_2}^{L_{\bar{c}}} |F_{m_{l_1}}^{c, l_1}(m, n) - F_{m_{l_2}}^{\bar{c}, l_2}(m, n)|^2$$

dient zur Ermittlung der im Mittel besten Basis der beiden Klassen c und $\bar{c}$. Für die Wahl der Verschiebungen m_{l_1} und m_{l_2} werden zwei Strategien verwendet [DJ97b]:

1. Bei Strategie A werden die Verschiebungen zu Null gesetzt:

$$\begin{aligned} m_{l_1}^c &= 0, & l_1 &= 1, 2, \dots, L_c, \\ m_{l_2} &= 0, & l_2 &= 1, 2, \dots, L_{\bar{c}}. \end{aligned}$$

2. Bei Strategie B werden im Unterschied zu Strategie A die Verschiebungen $\{m_{l_1}^c\}_{l_1=1}^{L_c}$ der Klasse c so gewählt, daß die Signale minimalen euklidischen Abstand untereinander aufweisen:

$$E = \min_{\{m_{l_1}^c\}_{l_1=1}^{L_c}} \sum_{i=1}^{L_c-1} \sum_{l=i+1}^{L_c} \sum_{m=-\infty}^{\infty} |s^{c,i}(m+m_i) - s^{c,l}(m+m_l)|^2. \quad (3.9)$$

Dahinter steckt die Erwartung, daß die aussagekräftigeren Merkmale gewonnen werden, wenn durch die Verschiebung die Ähnlichkeit der Signale der Klasse c vergrößert wird. Für die Verschiebungen der Klasse $\bar{c}$ wird wiederum

$$m_{l_2} = 0, \quad l_2 = 1, 2, \dots, L_{\bar{c}}$$

gewählt. Betrachtet wird im folgenden stets eine Verschiebung im Intervall $[-R, R]$. Die Optimierung der Abstände wird mit Hilfe evolutionärer Programme (Anhang 6.4) durchgeführt, da ein erschöpfender Vergleich aller Möglichkeiten praktisch nicht durchführbar ist.

Die Elemente der Matrix $\Delta F^c(m, n)$ können als quadrierte Merkmale eines Differenzsignals $\{\Delta s(m)\}_m$ interpretiert werden. Durch Umkehr der Abbildung (2.87) lassen sich aus n die Größen $i(n)$ und $l(n)$ bestimmen. So ist $i(n) \in \mathbb{N}$ die größte Zahl, für die die Aussage

$$n - \sum_{k=0}^{i(n)} M^k > 0$$

wahr ist. Es folgt dann sofort

$$l(n) = n - i(n).$$

Mit Hilfe der Umkehrabbildung lassen sich den Merkmalen $\Delta F^c(m, n)$ die Beträge der Ausgabefolgen $\{(\Delta w_{l(n)}^{i(n)})^c(m)\}_m$ einer Filterbank mit Baumstruktur ableiten:

$$|(\Delta w_{l(n)}^{i(n)})^c(m)| = \sqrt{\Delta F^c(m, n)}.$$

Auf die Folgen $\{|\Delta w_{l(n)}^{i(n)}(m)|\}_m$ läßt sich der Algorithmus der besten Basis mit den Kosten

$$K_l^i = M(\{|\Delta w_{l(n)}^{i(n)}(m)|^2\}_m)$$

anwenden. Die bei Abschluß des Algorithmus gewählten Knoten dienen zur Bildung der Indexmenge N_c^B .

3.2.2 Musterbildung und Klassifikation

Entsprechend zu den Abschnitten 3.1.2.1 und 3.1.2.2 werden in den folgenden beiden Abschnitten die Musterbildung und deren Einbindung in Adaption und Klassifikation dargestellt.

3.2.2.1 Musterbildung bei Adaption und Klassifikation

Die Muster werden entsprechend zu Abschnitt 3.1.2.1 erzeugt. Sollen die aktuellen Muster bezüglich der Merkmalsselektion für die Klasse c gebildet werden, ergibt sich:

$$P_{m_0}'^c(mV + k, n) = \begin{cases} F_{m_0}(mV + k, n) & \text{wenn } \begin{cases} \Delta E_{m,V}^c > S \\ \text{und} \\ n \in N_c^B \end{cases} \\ 0 & \text{sonst} \end{cases} \quad (3.10)$$

mit

$$\Delta E_{m,V}^c = \sum_{v=0}^{V-1} \Delta F_{m_0}^c(mV + v, n).$$

Abschließend werden die Muster normiert durch

$$P_{m_0}^c(m, n) = \frac{P_{m_0}^{'c}(m, n)}{\|P_{m_0}^{'c}\|}.$$

3.2.2.2 Klassifikation

Während der Adaptionphase werden die Trainingssignale $\{s^{c,l}(m)\}_m$ klassenspezifisch auf Merkmale $F_{m_0}^{c,l}(m, n)$ abgebildet. Die Muster der Datenbank $\mathbf{P}^{c,l} = P^{c,l}(m, n)$ werden über (3.6) für die gemäß Abschnitt 3.2.1 ermittelten signalabhängigen Verschiebungen $\{m_{l_1}^c\}_{l_1=1}^{L_c}$ gebildet, d.h.

$$P^{c,l}(m, n) = P_{m_{l_1}}^{c,l}(m, n)$$

mit $P_{m_{l_1}}^{c,l}(m, n)$ als Muster zu dem Signal $\{s^{c,l}(m)\}_m$ für alle $l = 1, 2, \dots, L_c$. In der Anwendungsphase wird das aktuelle Signal $\{s(m)\}_m$ unter der Hypothese, daß das Signal zu Klasse c gehört, jeweils einer Musterbildung für Klasse c unterworfen. Für jede Hypothese ergibt sich die Mustermenge $\{P_{m_0}^c\}_{m_0 \in V}$ wie in Bild 3.2 dargestellt.

Der Klassifikator der Klasse c ermittelt den Repräsentanten aus der Datenbank, der einen minimalen euklidischen Abstand d^c zu einem der Muster $\{P_{m_0}^c\}_{m_0 \in V}$ aufweist. Für den minimalen Abstand gilt:

$$d^c = \min_{1 \leq l \leq L_c} \min_{m_0 \in V} \|\mathbf{P}_{m_0} - \mathbf{P}^{c,l}\|^2.$$

Die ermittelten Abstände werden dem Entscheider (Bild 3.2) zugeführt, der das Minimum d^{c_1} der ermittelten Abstände und die zugehörige Klasse c_1 wählt:

$$d^{c_1} = \min_{1 \leq c \leq C} d^c.$$

Die Klassifikation wird zurückgewiesen, wenn der Abstand d^{c_1} zu dem nächstgelegenen Repräsentanten einer anderen Klasse :

$$d^{c_2} = \min_{1 \leq c \leq C, c \neq c_1} d^c$$

sich um weniger als $\Delta d_{\min}$ von d^{c_1} unterscheidet, d.h.

$$\text{Rückweisung, wenn: } d^{c_2} - d^{c_1} \leq \Delta d_{\min}.$$

3.3 Auswertung der Klassifikationsergebnisse

Die Menge der zur Verfügung stehenden Signale wird durch eine Variante des *Leaving-One-Out*-Verfahrens [DH73] in disjunkte Mengen der Trainings-signale und Anwendungssignale aufgeteilt. Dabei wird jeweils ein Teil der Signale zur Adaption verwendet und der Rest zur Anwendung. Jedes Signal wird entweder für das Training oder für die Anwendung eingesetzt. Das Verfahren wird in zwei Varianten durchgeführt:

- Zur Klassifikation des Signals l_c der Klasse c mit $1 \leq l_c \leq N_c$ besteht die Menge der Trainingssignale aus

$$N_M = \left(\sum_{c=1}^C N_c \right) - 1$$

Elementen, d.h. Es gibt $N_M + 1$ Klassifikationsdurchläufe, da jeweils ein Signal pro Durchlauf klassifiziert wird (*Leaving-One-Out*).

- Für die klassenabhängige Merkmalsselektion werden zur Reduktion des Rechenaufwands $D = N_c/5$ Durchläufe durchgeführt. Beginnend mit Durchlauf $d = 1$ werden jeweils die Signale $\{s^{c,l}(m)\}_m$ mit den Indizes $l = 5(d-1) + 1, 5(d-1) + 2, \dots, 5d$ für die Klassifikation verwendet. Die übrigen

$$N_M = \sum_{c=1}^C \frac{3}{4} N_c$$

Signale werden für die Adaption in jedem Durchlauf eingesetzt. Pro Klasse und Durchlauf werden D Signale klassifiziert. Es wird vorausgesetzt, daß der Stichprobenumfang N_c für jede Klasse c gleich groß ist.

Die Anzahl der Muster $a^{i,k}$ der Klasse i , die der Klassifikator der Klasse k zuordnet, werden durch

$$r^{i,k} = \frac{a^{i,k}}{N_T} 100 \quad [\%], \quad i = 1, 2, \dots, C, \quad k = 1, 2, \dots, C + 1$$

in Klassifikationshäufigkeiten umgerechnet, die wiederum die $(C, C + 1)$ Vertauschungsmatrix $\mathbf{R}$ bilden. Die Summe

$$r_K = \sum_{i=1}^C r^{i,i} \quad (3.11)$$

heißt Klassifikationsrate, die Summe

$$r_F = \sum_{i=1}^C \sum_{\substack{k=1 \\ k \neq i}}^C r^{i,k} \quad (3.12)$$

wird Fehlerkennungsrate genannt. Der Umfang an Rückweisungen wird durch die Rückweisungsrate

$$r_R = \sum_{i=1}^C r^{i,C+1} \quad (3.13)$$

ausgedrückt.

4 Signalquelle und Signalvorverarbeitung

Die Ergebnisse der Zeit-Frequenz-Analyse lassen sich durch Vorkenntnisse über die verwendeten Signalquellen, Klavier und Flügel, leichter interpretieren. Daher werden in Abschnitt 4.1 die Eigenschaften eines auf dem Klavier angeschlagenen Tons diskutiert. Die Durchführung der Aufnahmen und die Aufbereitung der Signale für die Klassifikationsaufgabe werden in Abschnitt 4.2 vorgestellt.

4.1 Klavier und Flügel als Signalquelle

Ein Klavier bzw. ein Flügel besteht aus Tastatur, Mechanik zur Umsetzung des Anschlags, Saiten und dem Instrumentgehäuse, das als Resonanzkörper wirkt. Für die Tonerzeugung wird durch Anschlagen der Taste ein filzübzogener Hammer auf die Saiten geschlagen. Damit nicht durch einen Nachschlag ein Doppelton erzeugt wird und nach Loslassen der Taste das System sofort wieder in die Ausgangslage versetzt wird, wird der Aufbau durch zusätzliche Mechanik unterstützt.

Flügel werden durch stabile Stahlrahmen gehalten, die die Verwendung von dicken, steifen Saiten bei hoher Spannung ermöglichen. Jede Komponente beeinflusst die Eigenschaften eines angeschlagenen Tones.

Die Tonhöhe wird durch die Länge der angeschlagenen Saite bestimmt. Das Verhalten einer ausgelenkten, ungedämpften Saite kann durch eine Fourierreihe beschrieben werden [Heu86]. Der Anschlag einer Taste erzeugt daher einen Grundton und Obertöne, deren Frequenz ganzzahlige Vielfache der Grundfrequenz sind. Je schneller der Anschlag erfolgt, desto stärker werden die Obertöne angeregt. Gegenüber diesem Modell gibt es in der Praxis Abweichungen:

- Die Saiten sind nicht exakt harmonisch. Diese Abweichung hängt von der Saitensteifigkeit, der Unnachgiebigkeit der Saitenhalterung und anderen Faktoren ab: Je größer die Steifigkeit, desto mehr weicht die Frequenz der Obertöne von den entsprechenden harmonischen Vielfachen ab [Tay92, Dic95]. Auch die Intonation (Perforierung des Hammerfilzes) beeinflusst den Klang (Obertonanteil) bedeutend.
- Jedem Ton ist ein Geräuschanteil überlagert, z.B. bei dem Klavier durch das Anschlaggeräusch.

- Der Anschlag wird in der Musik durch eine Reihe von Eigenschaften geprägt. Die Geschwindigkeit des Anschlags (Anschlagstärke), der Einsatz der Pedalen und Bindungen beim Klavierspiel erlauben dem Musiker Gestaltungsmöglichkeiten.
- Die Schwingung des Klaviers besitzt einen Einschwingvorgang, einen quasistationären Zustand und einen Ausklingvorgang.

Die Zerlegung eines angeschlagenen Tones in Einschwingphase, quasistationäre Phase und Ausklingvorgang soll im folgenden vertieft werden.

Die Einschwingphase umfaßt den Zeitraum von völliger Ruhe bis die Schwingung ihre maximale Amplitude erreicht hat. Die Dauer der Einschwingphase ist durch Materialeigenschaften, wie z.B. die Härte des Hammerfilzes bestimmt, nicht aber durch die Geschwindigkeit des Anschlags. Die Dauer der quasistationären Phase bei einem Klavierton wird in der Literatur unterschiedlich behandelt. Nach Meyer folgt der Einschwingphase eine kurze, ca. 150 – 200 ms dauernde quasistationäre Phase [Mey81]. Nach Dickreiter gibt es keinen quasistationären Zustand, vielmehr folgt direkt der Ausklingvorgang [Dic95]. Der Ausklingvorgang erfolgt langsam, für hohe Töne rascher als für niedrige, und dauert 20 – 40 Sekunden [Dic95].

Ein Modell zur Beschreibung eines durch ein Musikinstrument erzeugten Tones ist in [Rod97] angegeben. Das empfangene Signal $s(t)$ wird durch eine Summe zeitlich veränderlicher sinusförmiger Schwingungen und einen Rauschanteil $r(t)$ ausgedrückt:

$$s(t) = \sum_{i=1}^I a_i(t) \cos(\beta(t)) + r(t),$$

wobei $a_i(t)$ die zeitveränderliche Amplitude und $\beta(t)$ die Phase der Schwingung beschreiben. Durch Anwendung von Zeit-Frequenz-Analyseverfahren auf Beispielsignale werden die Parameter des Modells geschätzt [Rod97]. In dieser Arbeit werden als Merkmale nicht die geschätzten Parameter des Klaviertones verwendet, sondern direkt die Ausgabe der Zeit-Frequenz-Verfahren. Dies birgt den möglichen Nachteil, zu viele Merkmale für die Beschreibung eines Tones zu berücksichtigen.

Andererseits müssen die Funktionen $a_i(t)$, $\beta(t)$ und $r(t)$ für eine präzise Schätzung wiederum modelliert werden. Die Güte der heuristischen Modellierung, die z.B. aus den visuellen Ergebnissen der Zeit-Frequenz-Analyse abgeleitet wird, bestimmt den Erfolg der Schätzung. Ein ungenaues Modell

führt dazu, daß durch die Parameterschätzung Signalinformation verloren geht.

In [Rod97] werden aktuelle Verfahren zur Schätzung der Modellparameter diskutiert. Neue Verfahren, wie das in [Rod97] vorgestellte HRMP (High Resolution Matching Pursuit), verwenden direkt die Merkmale aus der Zeit-Frequenz-Ebene zur Modellierung des Signals $s(t)$ (vergl. Abschnitt 2.1.2). Der hier verfolgte Ansatz, durch Zeit-Frequenz-Analyse Merkmale für die Klassifikation abzuleiten, ist daher im Einklang mit aktuellen Entwicklungen zur Signalmodellierung von Klaviertönen.

4.2 Aufnahme und Signalaufbereitung

In Abschnitt 4.2.1 werden die Randbedingungen der Klavier- und Flügelaufnahmen diskutiert. Die in Abschnitt 4.2.2 beschriebene Signalaufbereitung umfaßt die Signalerkennung, d.h. das Schneiden des Signals, und die Umsetzung auf eine niedrigere Abtastrate sowie die Erzeugung des komplexwertigen Signals für die quadratischen Zeit-Frequenz-Verfahren.

4.2.1 Aufnahme

Es wurden Aufnahmen von 33 verschiedenen Klavieren und Flügeln sowie von einem digitalen Piano vorgenommen. Unter den Flügeln sind zwei modellgleiche Instrumente. Im Anhang 6.7 ist eine Liste der Instrumente aufgeführt.

Die Aufnahmen wurden im Musikhaus Schlaile, Karlsruhe, durchgeführt und erstreckten sich über einen Zeitraum von eineinhalb Jahren. Für die Aufnahmen wurde das Mikrofon *MK 5g* der Fa. Schöppts GmbH eingesetzt, das eine umschaltbare Kugel- und Nieren-Charakteristik aufweist. Als Aufzeichnungsgerät stand der DAT-Recorder *D-07* (Pioneer) zur Verfügung, der das Eingangssignal mit 16 Bit Auflösung bei einer Abtastrate $f_a = 48\text{kHz}$ digitalisiert. Jede Aufnahme umfaßte den Anschlag des C-Dur-Dreiklangs $c^1 - e^1 - g^1$ [Tay92]. Die Aufnahme des Anschlags endete mit Verklingen des Tones, d.h. nach etwa 10-30 Sekunden. Der Anschlag erfolgte bei jedem Klavier mit etwa gleicher Geschwindigkeit ohne den Einsatz der Pedale.

In [Dic95] wird vorgeschlagen, für einkanalige Flügelaufnahmen das Mikrofon vor dem Flügel zu positionieren. Für Klavieraufnahmen wurden keine Literaturhinweise gefunden.

Da nur ein Akkord angeschlagen wurde, ist das Mikrofon bei jedem Flügel bzw. Klavier so plaziert worden, daß eine dem eigenen akustischen Eindruck gleichende Aufnahme durchgeführt wurde. Dafür wurde das Mikrofon mit Kugelcharakteristik betrieben. Für eine Klavieraufnahme wurde das Mikrofon mittig im Klavierkasten versenkt, bei Flügelaufnahmen im Abstand von 20 – 30 cm über den angeschlagenen Saiten bei halbgeöffnetem Flügeldeckel positioniert.

Von jedem Instrument wurden zahlreiche Aufnahmen gemacht, von denen pro Instrument jeweils 20 für die Klassifikationsaufgabe herangezogen wurden. Bei dem verwendeten digitalen Piano wurden ebenfalls 20 Aufnahmen bei direktem Anschluß an den DAT-Recorder erzeugt. Für die Klassenzahl C und die Anzahl der Aufnahmen N_c pro Klasse gilt also:

$$\begin{aligned} C &= 34, \\ N_c &= 20, \quad c = 1, 2, \dots, 34. \end{aligned}$$

4.2.2 Signalerkennung und Signalvorverarbeitung

Beginn und Ende des Klaviertons werden durch einen Detektionsalgorithmus ermittelt. Zunächst wird jedes Signal so normiert, daß das betragsgrößte Folgeelement des zeitdiskreten Eingangssignals den Wert 1 hat. Der Anfang des Klaviertons wird durch Verschieben eines Fensters über der normierten Folge $\{s_o(m)\}_m$ ermittelt. Der Algorithmus erkennt den Start des Tons durch Messen der durchschnittlichen Leistung P_M^{LW} innerhalb eines Fensters der Länge L_W beginnend mit dem Abtastwert M :

$$P_M^{LW} = 10 \log_{10} \left\{ \frac{1}{L_W} \sum_{l=M}^{M+L_W-1} |s_o(l)|^2 \right\} .$$

Der erste Abtastwert des Tons M_{Start} ist der erste Wert, für den die auf 1 bezogene Leistung

$$P_{M_{Start}}^{80} \geq P_{start} = -25\text{dB}$$

überschritten wird.

Das Ende jedes Tons wird durch zwei Detektionsmodi, die durch den gemeinsamen Signalfuß in Bild 4.1 beschrieben werden, ermittelt:

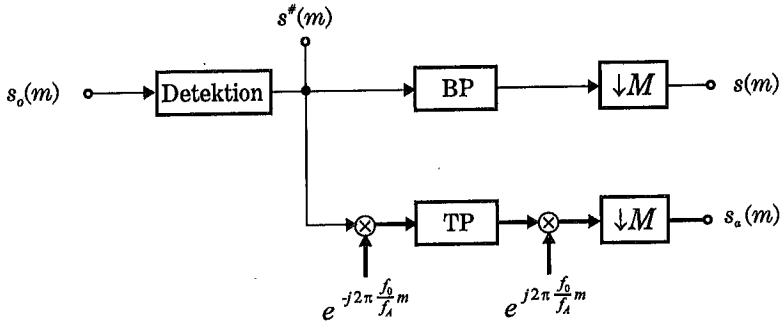


Bild 4.1: Signalfluß zur Ermittlung und Vorverarbeitung des Klaviertons

1. In diesem Fall wird der gesamte Klavierton analysiert. Daher ist der letzte Abtastwert des Tons M_{Stop} der erste Wert, für den die Leistung:

$$P_{M_{Stop}}^{7200} \geq P_{start} = -50\text{dB} \quad (4.1)$$

unterschritten wird. Das geschnittene Signal $\{s^\#(m)\}_m$ wird in dem oberen Zweig in Bild 4.1 einer Bandpaßfilterung unterzogen und um den Faktor $M = 8$ dezimiert. Die Grenzfrequenzen des Bandpasses werden zu $f_u = 200\text{Hz}$ (untere Grenzfrequenz) und $f_o = 3\text{ kHz}$ gewählt. In den Klassifikationsläufen hat sich gezeigt, daß die Merkmalsselektion nur Merkmale verwendet, die in der Zeit-Frequenz-Ebene dem Frequenzbereich bis 3 kHz entsprechen. Zur Erzeugung des komplexwertigen Signals wird die Folge $\{s^\#(m)\}_m$ mit $f_0 = 1600\text{ Hz}$ gemischt, einer Tiefpaßfilterung mit der Grenzfrequenz $f_G = 1400\text{ Hz}$ unterzogen und anschließend mit $-f_0$ zurückgemischt. Der untere Zweig erzeugt durch diese Operationen die komplexwertige Folge $\{s_a(m)\}_m$. Jeder so geschnittene Ton umfaßt etwa 48000 Abtastwerte bei einer Abtastrate von $f_A = 6\text{ kHz}$.

2. Im zweiten Fall soll nur der Einschwingvorgang verarbeitet werden. In der Literatur [Dic87] wird die Einschwingphase mit ca. 25 ms angegeben. Dieser Wert dient als Kriterium für das Ende des Einschwingvorgangs und führt bei einer Abtastrate von 48 kHz auf

$$M_{Stop} = M_{Start} + 1200.$$

Eine Analyse durch gemittelte Periodogramme [KK89] der Einschwingphase ergab, daß es ausreichend ist, den Frequenzbereich bis 2 kHz zu berücksichtigen.

Daher wird die Signalverarbeitung gegenüber dem ersten Fall auf folgende Weise geändert. Die Grenzfrequenzen des Bandpasses werden zu $f_u = 200$ Hz und $f_o = 2$ kHz gewählt. Die Mischfrequenz wird zu $f_0 = 1100$ Hz gewählt, die Grenzfrequenz des Tiefpasses zu $f_G = 900$ Hz und die Dezimation zu $M = 12$, so daß die Abtastrate der Ausgangsfolgen 4 kHz beträgt.

Bandpaßfilter und Tiefpaß werden durch Fourierapproximation und Kaiserfensterung entworfen. Die Filterlänge beträgt $K = 481$. Die Laufzeit der Filter muß bei Ermittlung der Einschwingphase mit berücksichtigt werden. Aus diesem Grund werden die ersten $(K - 1)/2$ Abtastwerte aus der Ausgangsfolge entfernt. Für den Fall, daß nur der Einschwingvorgang analysiert werden soll, werden die nachfolgenden 100 Abtastwerte für reelles und komplexwertiges Ausgangssignal verwendet, d.h. $\{s(m)\}_{m=0}^{L-1}$ bzw. $\{s_a(m)\}_{m=0}^{L-1}$ mit $L = 100$.

5 Ergebnisse

In Abschnitt 5.1 werden die Klassifikationsergebnisse mit klassenunabhängiger Merkmalsselektion für die vollständigen Klaviertöne diskutiert und zu deren Interpretation Beispiele für Visualisierungen der Merkmale in der Zeit-Frequenz-Ebene angegeben. Auf eine klassenabhängige Merkmalsselektion bzw. Klassifikation wurde wegen des hohen Rechenaufwands für die vollständigen Klaviertöne verzichtet.

In Abschnitt 5.2 folgen mit klassenabhängiger und -unabhängiger Merkmalsselektion die Erkennungsergebnisse für die Anschlagphase.

Den Abschluß bildet in Abschnitt 5.3 ein Vergleich der für jedes Merkmalsextraktionsverfahren optimalen Klassifikationsergebnisse.

5.1 Klassifikationsergebnisse für die vollständigen Klaviertöne

Für die vollständigen Klaviertöne wurde der Klassifikator mit den in Tabelle 5.1 eingetragenen Einstellungen betrieben.

B	:	5	10	15	20	30	50	75	100	150	200
V	:	5									
Δd_{min}	:	0,1									

Tabelle 5.1: Wahl von Segmentanzahl B und -länge V sowie der Rückweisungsschwelle Δd_{min} für den Klassifikator

Zur Wahl der Rückweisungsschwelle Δd_{min} wird folgendes bemerkt: Wird für die Musterbildung der Betrag der Merkmale nach (3.8) verwendet, beträgt der maximale Abstand zweier Muster $\sqrt{2}$, bei der Bildung nach (3.7) entsprechend 2. Die Wahl $\Delta d_{min} = 0,1$ führte in den Durchläufen zu einer geringen Fehlerkennungsrate r_F (3.12), ohne die Klassifikationsrate r_K (3.11) deutlich zu mindern. In [DJ97a, DJ98b] wurden Klassifikationsläufe mit einem anderen Rückweisungskriterium durchgeführt.

5.1.1 Kurzzeit-Fourier-Transformation

Für die KFT werden die Visualisierung der Merkmale in der Zeit-Frequenz-Ebene und die Auswertung der Klassifikationsergebnisse ausführlich erläutert.

Der angeschlagene Klavierakkord $c^1 - e^1 - g^1$ besitzt im Frequenzbereich bis 2 kHz die in Tabelle 5.2 angegebenen Grundswingungen und Oberwellen. Mit Hilfe der Tabelle 5.2 lassen sich die in der Zeit-Frequenz-Ebene

Grundswingung c^1 : 262 Hz	2.Oberwelle zu g^1 : 1176 Hz
Grundswingung e^1 : 330 Hz	4.Oberwelle zu c^1 : 1308 Hz
Grundswingung g^1 : 392 Hz	3.Oberwelle zu e^1 : 1319 Hz
1.Oberwelle zu c^1 : 523 Hz	3.Oberwelle zu g^1 : 1568 Hz
1.Oberwelle zu g^1 : 659 Hz	5.Oberwelle zu c^1 : 1570 Hz
1.Oberwelle zu e^1 : 784 Hz	4.Oberwelle zu e^1 : 1658 Hz
2.Oberwelle zu c^1 : 785 Hz	6.Oberwelle zu c^1 : 1831 Hz
2.Oberwelle zu e^1 : 989 Hz	4.Oberwelle zu g^1 : 1960 Hz
3.Oberwelle zu c^1 : 1047 Hz	5.Oberwelle zu e^1 : 1978 Hz

Tabelle 5.2: Liste der Grundswingungen und Oberwellen bis 2 kHz, die durch den Anschlag des Akkords $c^1 - e^1 - g^1$ angeregt werden (gerundete Frequenzwerte)

visualisierten Merkmale der KFT interpretieren. Als Beispiel ist der Ton 1 des Yamaha C2-Flügels, der im folgenden als Referenzton dient, in Bild 5.1 dargestellt, wobei die KFT mit der Parameterwahl $N = 512$ und $K = 1024$ betrieben wurde. Die Merkmale wurden für die Visualisierung auf das betragsgrößte Merkmal normiert und im Betrag durch einen Grauwert in dB dargestellt. Weiße Flächen in der Darstellung bedeuten, daß die zugehörigen Merkmale unter der Darstellungsgrenze von -70 dB liegen.

In Bild 5.1 sind deutlich die drei Grundswingungen und die Oberwellen aus Tabelle 5.2 zu erkennen. Durch die Breite der Analysefilter, die größer als der ideale Wert $6000\text{Hz}/512 \approx 12$ Hz ist, können dicht benachbarte Schwingungen nicht aufgelöst werden.

Die KFT wurde so wie in Tabelle 5.3 parametrisiert, so daß Kanalzahl N und Filterlänge K in dem festen Verhältnis $K = 2N$ standen.

Kanalzahl N :	64	128	256	512	1024
Filtertyp :	Fourier-Approximation/ Kaiserfensterung				
Filterlänge K :	128	256	512	1024	2048

Tabelle 5.3: Parameterwahl bei der KFT

In Tabelle 5.4 sind die Ergebnisse eines Klassifikationslaufs für das Beispiel

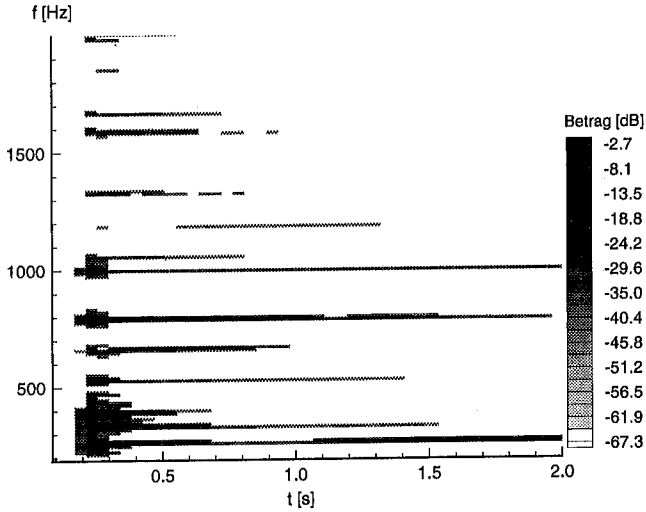


Bild 5.1: Zeit-Frequenz-Darstellung des Tons 1 (Yamaha C2-Flügel) durch die KFT

$N = 128$ und $K = 256$ aufgeführt. Die Klassifikation erfolgte für Segmentzahlen $B \geq 30$ ohne Fehlerkennung und erreicht dennoch sehr hohe Erkennungsraten r_K . Klassifikationen, die für $\Delta d_{\min} = 0,2$ durchgeführt wurden, lieferten bei der KFT unverändert hohe Klassifikationsraten. Die hohe Leistungsfähigkeit dieses Zeit-Frequenz-Verfahrens wird auch durch die gemeinsame Darstellung der Klassifikationsergebnisse für diese Parameterwahl und die übrigen Einstellungen in Bild 5.2 deutlich. Für $N \geq 256$ Kanäle erreichte der Klassifikator mit $B \geq 20$ Segmenten eine Klassifikationsrate von 100%.

5.1.2 Wavelet-Transformation

Die dyadische Wavelet-Transformation (DWT) wurde mit den in Tabelle 5.5 aufgeführten Parametereinstellungen durchgeführt.

Es wurden die Merkmale nach (3.8) und (3.7) bestimmt. Die Klassifikationsrate war von der Filterlänge und der Merkmalsbildung praktisch unabhängig und erreichte ab einer Zerlegungstiefe von $T = 5$ bezüglich T konstante Werte, wie für $K = 10$ in Bild 5.3 zu erkennen ist.

Die geringen Klassifikationsraten der DWT können dadurch erklärt werden,

Segmentzahl B :	r_K [%]	r_F [%]	r_R [%]
5	: 65,15	2,50	32,35
10	: 82,06	0,88	17,06
20	: 92,50	0	7,50
30	: 95,29	0	4,71
50	: 97,21	0	2,79
75	: 98,08	0	1,92
100	: 98,38	0	1,62
150	: 99,41	0	0,59
200	: 99,41	0	0,59

Tabelle 5.4: Klassifikations-, Fehlerkennungs- und Rückweisungsdaten für $N = 128$ und $K = 256$

Filtertyp	:	Daubechies					
Filterlänge K	:	10	10	10	20	20	20
Zerlegungstiefe T	:	2	5	8	2	5	8

Tabelle 5.5: Parameterwahl bei der DWT

daß die verwendete Filterbank die Charakteristik des Klavierakkords nur ungenügend beschreibt und die feinere Aufteilung des Frequenzbereichs für niedrige Frequenzen keine neuen Informationen beiträgt.

Zum einen werden durch die Signalvorverarbeitung Frequenzanteile des Signals unter 200 Hz weggefiltert. Zum anderen zerlegt die DWT nach 5 Stufen den Frequenzbereich unter $6000\text{Hz}/2^5 = 187,5\text{ Hz}$, in dem der Klavierton ohnehin keinen Beitrag mehr aufweist.

Dies wird auch durch die Visualisierung der Merkmale des Klaviertons Nr. 1 des Yamaha C2-Flügels in der Zeit-Frequenz-Ebene in dem Frequenz-Bereich 187,5 – 2000 Hz deutlich (Bild 5.4), bei dem die DWT mit Daubechies Filtern der Länge $K = 20$ und der Zerlegungstiefe $T = 5$ durchgeführt wurde. Zum Verständnis vergleiche man die zugehörige Filterbank, die für $T = 2$ in Bild 2.10 dargestellt ist und die in Bild 2.28 angegebene Rasterung der Zeit-Frequenz-Ebene. Den Merkmalen $F(m, 0)$ wird der Frequenzbereich 1500 – 3000 Hz zugeordnet, von dem der Ausschnitt 1500 – 2000 Hz dargestellt ist. Entsprechend werden die Merkmale $F(m, 1)$ im Bereich 750 – 1500 Hz visualisiert. Der Klavierton wird durch die geringe Frequenzauflösung bezüglich der Frequenz verschmiert. Die Eigenschaft der DWT, das Signal

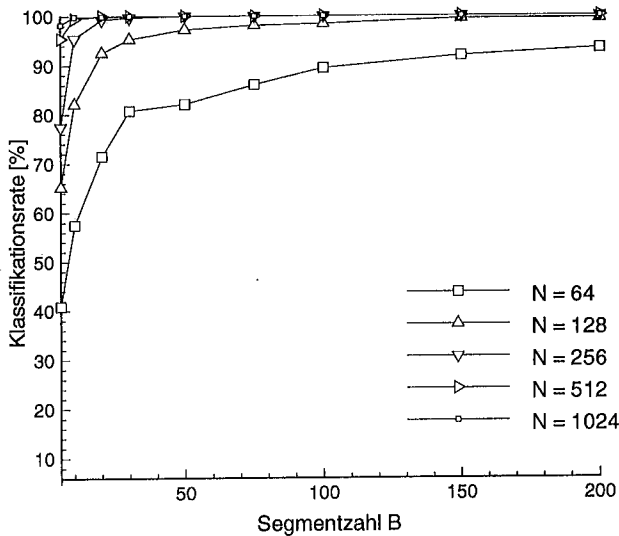


Bild 5.2: Klassifikationsraten für die KFT in Abhängigkeit von der Kanalzahl N

mit hoher Zeitaufösung zu analysieren, stellt hier keinen Vorteil dar.

Ein weiterer Aspekt ist die in Kapitel 2.1.3.2 angesprochene Abhängigkeit des Transformationsergebnisses von der Verschiebung des Signals. Der Startzeitpunkt, der durch den Schneidalgorithmus ermittelt wird, beeinflusst damit auch das Transformationsergebnis. Ein Klavierton wird nicht erkannt, wenn das zugehörige Muster zu den Referenzmustern der Datenbank durch den Schneidversatz einen zu großen Abstand aufweist.

5.1.3 Filterbänke mit Baumstruktur

Die Filterbank mit Baumstruktur wurde wie in Tabelle 5.6 parametrisiert. Bei Einsatz der Daubechies Filter, d.h. bei einer diskreten Wavelet-Packet-Transformation (DWPT), ergaben sich bei vorzeichenbehafteten Merkmalen nach (3.7) und Filterlänge $K = 20$ die in Bild 5.5 aufgeführten Klassifikationsraten.

Durch den Algorithmus der besten Basis wird gegenüber der DWT eine bessere Anpassung der Transformation an das Signal erzielt. In Bild 5.6 ist die Zeit-Frequenz-Darstellung des Referenztons für die Zerlegungstiefe $T = 10$

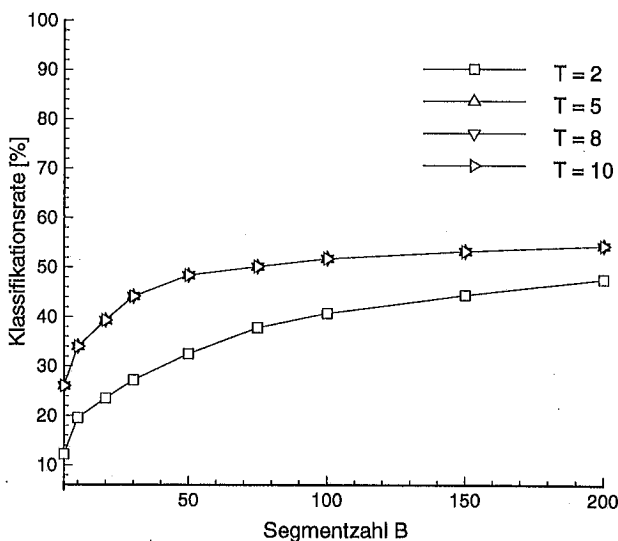


Bild 5.3: Klassifikationsergebnisse der DWT bei Merkmalsbildung nach (3.8) für verschiedene Zerlegungstiefen T , Filterlänge $K = 10$

und der Filterlänge $K = 20$ abgebildet. Es ist zu erkennen, daß der Algorithmus frequenzabhängig unterschiedliche zeitliche Auslösungen gewählt hat. Die Repräsentation des Klaviertons ist im Vergleich zur KFT in Bild 5.1 schwierig zu interpretieren:

1. Je feiner die Frequenzauflösung ist, umso länger ist die damit verknüpfte Filterkaskade nach Bild 2.13 bzw. die Signallaufzeit durch die Filter. Zudem sind praktisch sinnvolle FIR-Filter einer zweikanaligen, orthogonalen Filterbank nicht linearphasig [Vai85, HV91]. Diese Eigenschaften führen dazu, daß die Zeit-Frequenz-Darstellung des Signals bezüglich der Zeit „zerissen“ scheint. Beispiel: Der Algorithmus der besten Basis hat um $f = 1500$ Hz eine grobe, um $f = 300$ Hz eine feine Frequenzauflösung gewählt. Das Signal erscheint bei $f = 300$ Hz um etwa eine halbe Sekunde verzögert.
2. Die Kaskadierung der Filter führt zu resultierenden Filtern (2.89), die nicht frequenzselektiv sind. Beispiel: Das resultierende Filter der Kaskade

$$H_{5,13}^{res}(z) = H_1(z)H_0(z^2)H_1(z^4)H_0(z^8)H_1(z^{16})$$

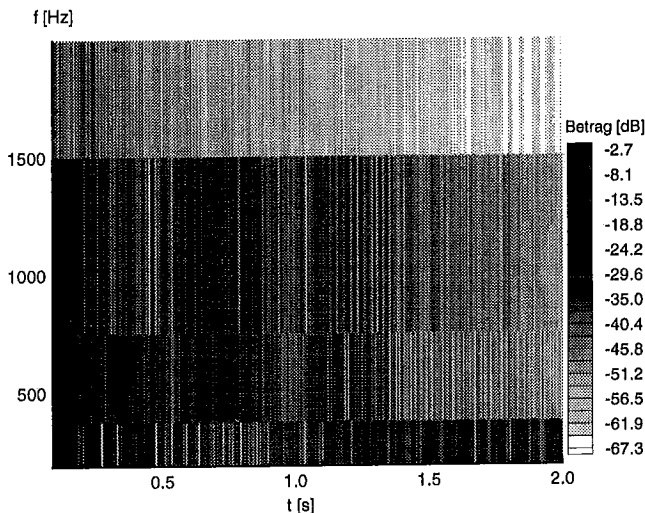


Bild 5.4: Zeit-Frequenz-Darstellung des Tons 1 (Yamaha C2-Flügel) durch die DWT

besitzt auf dem Einheitskreis $z = e^{j\omega}$ den in Bild 5.7 dargestellten Amplitudengang. Die geringe Frequenzselektivität führt zu einer Verschmierung bezüglich der Frequenz und erlaubt keine Zuordnung der Zeit-Frequenz-Repräsentation zu den Teilschwingungen des Klaviertons.

Die Abhängigkeit des Transformationsergebnisses vom Schneidvorgang ist ebenso bei der DWPT gegeben. Durch die Fähigkeit der DWPT, sich der Signalstruktur durch den Best-Basis-Algorithmus anzupassen, scheint sich diese Abhängigkeit für $T \geq 5$ auf die Vorzeichen der berechneten Merkmale zu konzentrieren. Denn durch Verwendung der Betragsbildung nach (3.8) wurden die Erkennungsergebnisse deutlich besser und erreichten für die Zerlegungstiefe $T = 8$ die besten Ergebnisse von über 93% bei $B \geq 10$ (Bild 5.8). Für $T = 2$ ist die ausgewählte Transformation noch nicht genügend an das Signal adaptiert und die durch Betragsbildung erzeugten Merkmale liefern schlechtere Resultate.

Wird die Rückweisungsschwelle jedoch auf den Abstand $\Delta d_{min} = 0,2$ angehoben, sinkt die Klassifikationsrate um 15% Prozent, d.h. die Entscheidungen der KFT im Vergleich zur DWPT sind besser, als der Vergleich für

Filtertyp	:	Daubechies							
Filterlänge L	:	10	10	10	10	20	20	20	20
Zerlegungstiefe T	:	2	5	8	10	2	5	8	10
Filtertyp	:	$M = 2$, Entwurf nach (2.30)							
Filterlänge K	:	10	10	10	10	20	20	20	20
Zerlegungstiefe T	:	2	5	8	10	2	5	8	10
Filtertyp	:	$M = 3$, Entwurf nach (2.30)							
Filterlänge K	:	30	30	30	30	30			
Zerlegungstiefe T	:	2	3	4	5	6			

Tabelle 5.6: Parameterwahl bei der Filterbank mit Baumstruktur

$d_{min} = 0,1$ anzeigt. Die Schwelle $d_{min} = 0,1$ für die DWPT ist dennoch geeignet gewählt, denn Fehlerkennungen gibt es nicht für $T = 8$ und $B \geq 30$.

Bei Einsatz der nach (2.30) optimierten Filter ($M = 2$) ergaben sich die in Bild 5.9 dargestellten Klassifikationsresultate. Die Impulsantworten der entworfenen Filter sind in Anhang 6.6 aufgeführt. Die Ergebnisse für $T = 5$ sind um über 10% besser als bei Verwendung der Daubechies-Filter, bei $T = 8$ erreichen die Filter um 1 – 2% bessere Resultate. Dies ist darauf zurückzuführen, daß die optimierten Filter frequenzselektiver sind. Bild 5.10 zeigt im Vergleich das nach (2.30) optimierte Tiefpaßfilter und das Daubechies-Filter für $K = 20$. Bei Verwendung der Filterlänge $K = 10$ bei den Daubechies und den selbstentworfenen Filtern wurden um etwa 3 – 5% schlechtere Ergebnisse gegenüber dem Fall $K = 20$ erzielt.

Die höhere spektrale Selektivität der optimierten Filter kann als Anpassung an die Signaleigenschaften der Klaviertöne interpretiert werden. Die in Abschnitt 3.1.1 getroffene Annahme, daß die Filtereigenschaften eine geringe Wirkung auf die Klassifikationsergebnisse haben, bestätigt sich. Vielmehr bestimmt die Zerlegungstiefe T und damit auch die Wahl der Basis die Resultate.

Bei $M = 3$ Kanälen wurden die Klassifikationsläufe für Zerlegungstiefen von $T = 2, 3, 4, 5, 6$ durchgeführt. Dabei wurden die in Bild 5.11 dargestellten Klassifikationsraten erreicht. Die besten Ergebnisse wurden für $T = 5, 6$ mit über 95% für $B \geq 30$ Segmente erzielt. Die Ergebnisse für die Zerlegungstiefe $T = 5$ entsprechen denen der Tiefe $T = 8$ im Fall $M = 2$. Im direkten Vergleich stellt der Fall $M = 3$ durch die geringere Zerlegungstiefe nach Anhang 6.5 das aufwandsgünstigere Verfahren mit etwa 443 Operationen pro Takt dar.

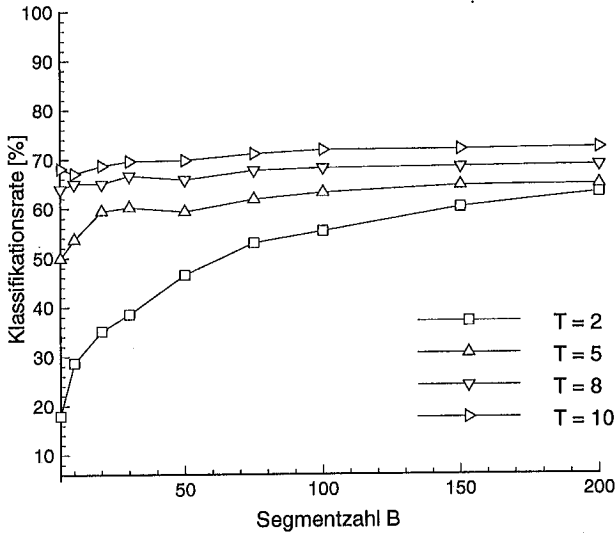


Bild 5.5: Klassifikationsergebnisse der DWPT bei Merkmalsbildung nach (3.7) und $K = 20$ in Abhängigkeit von der Zerlegungstiefe T

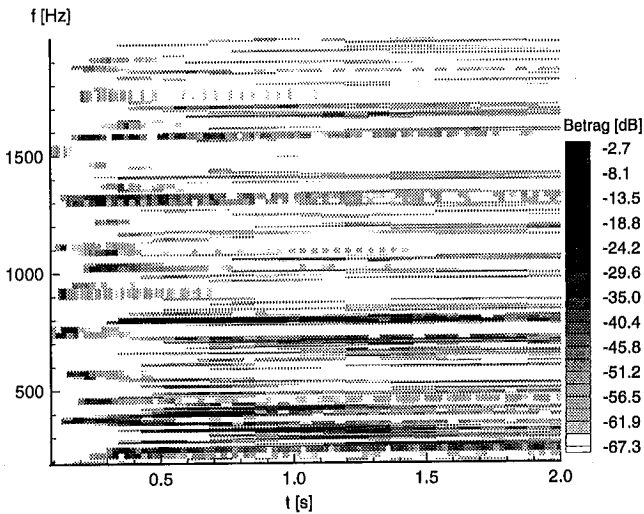


Bild 5.6: Zeit-Frequenz-Darstellung des Tons 1 (Yamaha C2-Flügel) durch die DWPT

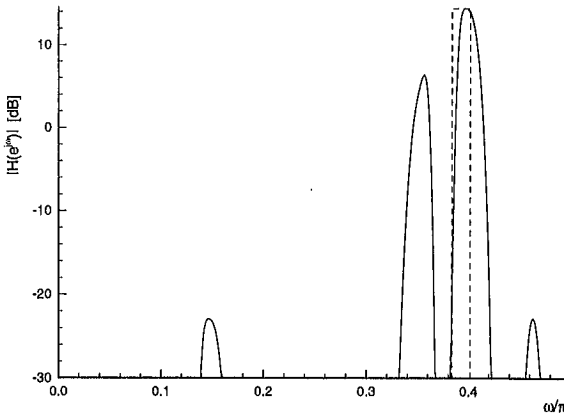


Bild 5.7: Amplitudengang des Filters $H_{5,13}^{res}(e^{j\omega})$ und idealer Verlauf (unterbrochen gezeichneter Graph)

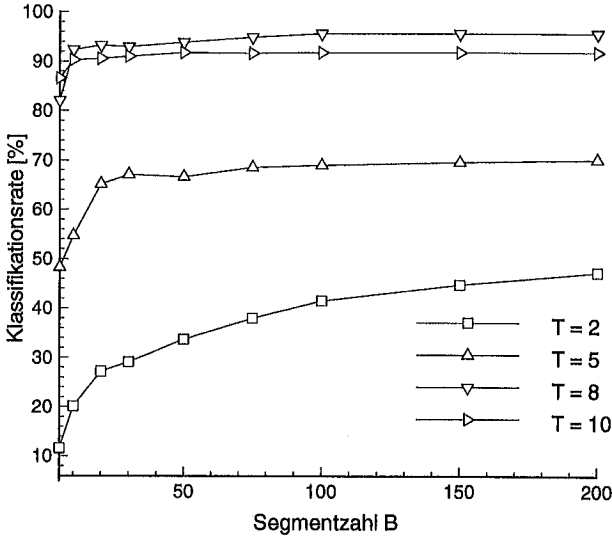


Bild 5.8: Klassifikationsergebnisse der DWPT bei Merkmalsbildung nach (3.8)

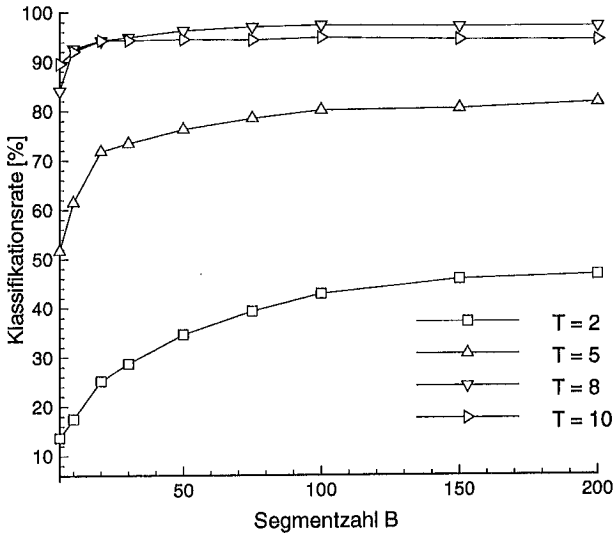


Bild 5.9: Klassifikationsergebnisse für die nach (2.30) optimierten Filter ($M = 2$, $K = 20$)

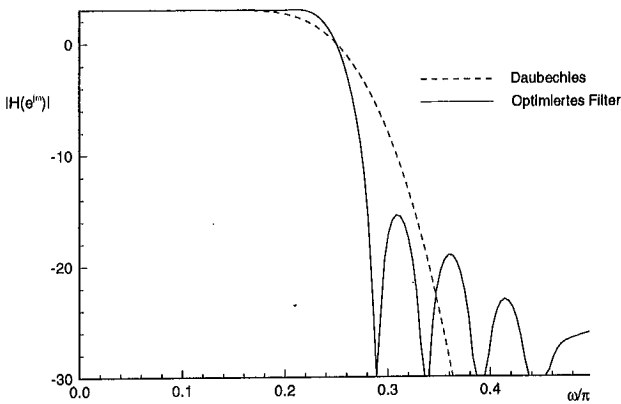


Bild 5.10: Vergleich der Amplitudengänge des Daubechies-Filters und dem nach (2.30) optimierten Filter

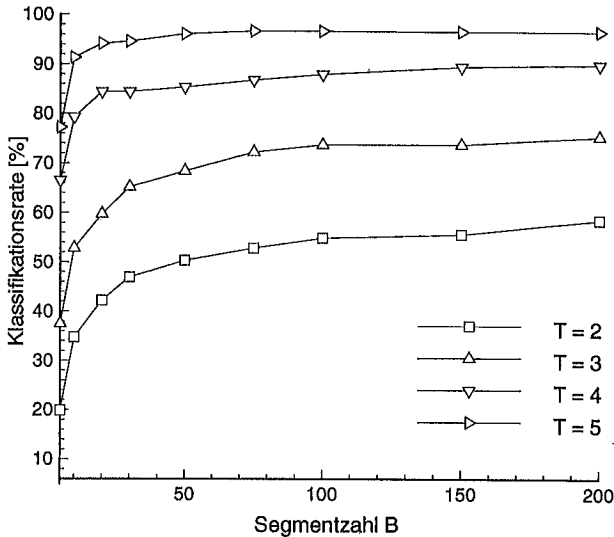


Bild 5.11: Klassifikationsergebnisse für die nach (2.30) optimierten Filter ($M = 3$, $K = 20$)

5.1.4 Quadratische Zeit-Frequenz-Verfahren

Für die Verfahren der Cohen-Klasse wurde die Implementierung nach (2.141) verwendet. Die Kernfunktion für die betrachteten Verfahren sind durch (2.133)-(2.135) gegeben.

Die Ergebnisse für die Pseudo-Wigner-Verteilung (PWV) wurden durch eine Parametrisierung des Verfahrens nach Tabelle 5.7 ermittelt. Die Definition und Eigenschaften des Hamming-Fensters sind in [Har78] angegeben.

Fenstertyp $f_1(k)$:	Hamming			
Fensterlänge K_1 :	63	127	255	511
Dezimation M :	32			

Tabelle 5.7: Parameterwahl bei der PWV

Die Resultate der Klassifikation sind bei einer Merkmalsbildung nach (3.7) in Bild 5.12, bzw. für (3.8) in Bild 5.13 dargestellt.

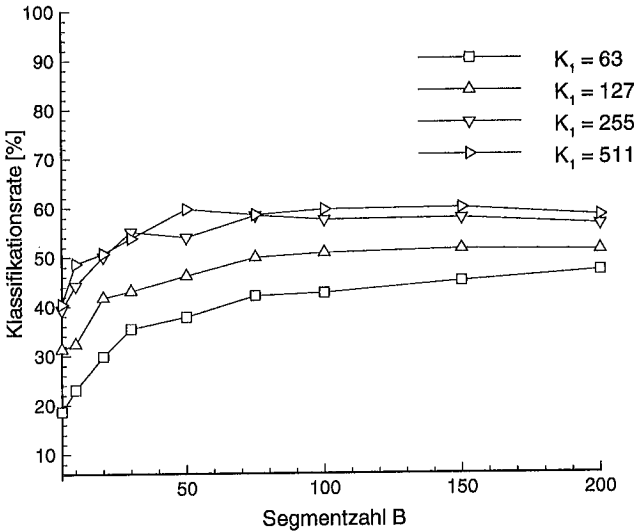


Bild 5.12: Klassifikationsergebnisse der PWV bei Merkmalsbildung nach (3.7)

Mit größerer Fensterlänge K_1 nimmt stets die Erkennungsrate zu, wobei der Vergleich beider Ergebnisse zeigt, daß die Betragsbildung für lange Fenster eine deutliche Verbesserung der Erkennungsleistung bringt, für kurze eine Verschlechterung. Mit größerer Fensterlänge nimmt die Frequenzauflösung zu und die Beschreibung des Signals in der Zeit-Frequenz-Ebene liefert bessere Merkmale.

Zwischen jeweils zwei Frequenzen f_1 und f_2 der im Signal enthaltenen Schwingungen (vergl. Tabelle 5.2) tritt bei der Wigner-Verteilung ein oszillierender Kreuzterm der Frequenz $(f_1 + f_2)/2$ als zusätzliches Merkmal auf. Die Geschwindigkeit, mit der die einzelnen Teilsignale einschwingen, beeinflusst die Kreuztermbildung, denn die Oszillation der Kreuzterme hängt von der Phase der beteiligten Schwingungen ab [Hla84]. Die PWV reduziert diese Kreuzterme nicht, da sie nur Kreuzterme, die durch zeitlich getrennte Signale erzeugt werden, unterdrückt.

Durch die Betragsbildung wird für $L \geq 127$ der Einfluß der Phase und damit des Einschwingvorgangs reduziert, da Kreuz- und Autoterme durch die hohe Frequenzauflösung getrennt werden. Für kleine Fensterlängen dominiert der Einfluß der geringen Frequenzauflösung. Durch die Betragsbildung der

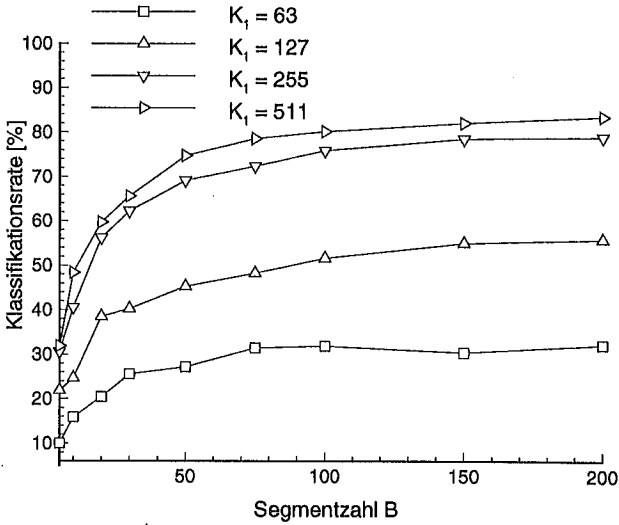


Bild 5.13: Klassifikationsergebnisse der PWV bei Merkmalsbildung nach (3.8)

bezüglich der Frequenz verschmierten Auto- und Kreuzterme werden für die Klassifikation schlechtere Merkmale erzeugt.

Für die Smoothed Pseudo-Wigner-Verteilung (SPWV) wurden die Parameter der Merkmalsextraktion wie in Tabelle 5.8 gewählt. Die Klassifikations-

Fenstertyp $f_1(k)$:	Hamming				
Fenstertyp $f_2(l)$:	Hamming				
Fensterlänge K_1 : 127	127	127	255	255	255
Fensterlänge K_2 : 63	127	255	63	127	255
Dezimation M :	32				

Tabelle 5.8: Parameterwahl bei der SPWV

ergebnisse sind für $K_1 = 127$ und für $K_2 = 255$ in Bild 5.14 dargestellt.

Da das Hammingfenster ein Tiefpaß der Grenzfrequenz $\omega_c(K_2) \approx 4\pi/K_2$ ist, werden in dem Klaviersignal Kreuzterme zu den Teilschwingungen unterdrückt, die sich um mindestens $\Delta f(K_2) \approx (2/K_2)f_a = 12000/K_2\text{Hz}$ in

der Frequenz unterscheiden. Es sind $\Delta f(63) = 190$ Hz und $\Delta f(255) = 47$ Hz. In der Visualisierung (Bild 5.15) läßt sich diese Überlegung bestätigen. Die Kreuzterme für $K_2 = 63$ werden kaum unterdrückt, während für $K_2 = 255$ die Kreuzterme weitgehend eliminiert sind. Die Kreuztermunterdrückung führt damit zu einer deutlich besseren Erkennungsrate. Wie an den Ergebnissen in Bild 5.14 (oben) zu erkennen ist, spielt die Kreuztermunterdrückung gerade für geringe Frequenzauflösungen, d.h. kleines K_1 , eine große Rolle.

Tabelle 5.9 enthält die Parametersätze, für die die Choi-Williams-Verteilung (CWV) berechnet wurden.

Parameter σ	: 0.01	0.1	1	10	0.01	0.1	1	10
Fensterlänge K_1	: 127	127	127	127	255	255	255	255
Fensterlänge K_2	: 127	127	127	127	255	255	255	255
Dezimation M	:				32			

Tabelle 5.9: Parameterwahl bei der CWV

Die erreichten Klassifikationsraten sind in den Bildern 5.16 und 5.17 für verschiedene σ dokumentiert.

Je kleiner σ gewählt wird, um so stärker werden Kreuzterme von Signalen, die einen Zeit- und einen Frequenzversatz aufweisen, unterdrückt. Zur Deutung der Kreuztermunterdrückung kann man sich die Teilschwingungen aus Tabelle 5.2 in Zeitrichtung segmentiert vorstellen. Es werden dann genau die Kreuzterme aus der Darstellung entfernt, die aus zwei zeitlich versetzten Segmenten unterschiedlicher Teilschwingungen resultieren. Kreuzterme aus zeitlich übereinstimmenden Signalsegmenten bleiben bestehen. Ein Beispiel für die Wirkung der Kreuztermunterdrückung durch die CWV liefert Bild 5.18. Die verbliebenen Kreuzterme bilden einen „Teppich“, der charakteristisch für die CWV bzw. für Zeit-Frequenz-Verfahren ist [HMUJ95], die die Bedingungen (2.126) und (2.127) erfüllen.

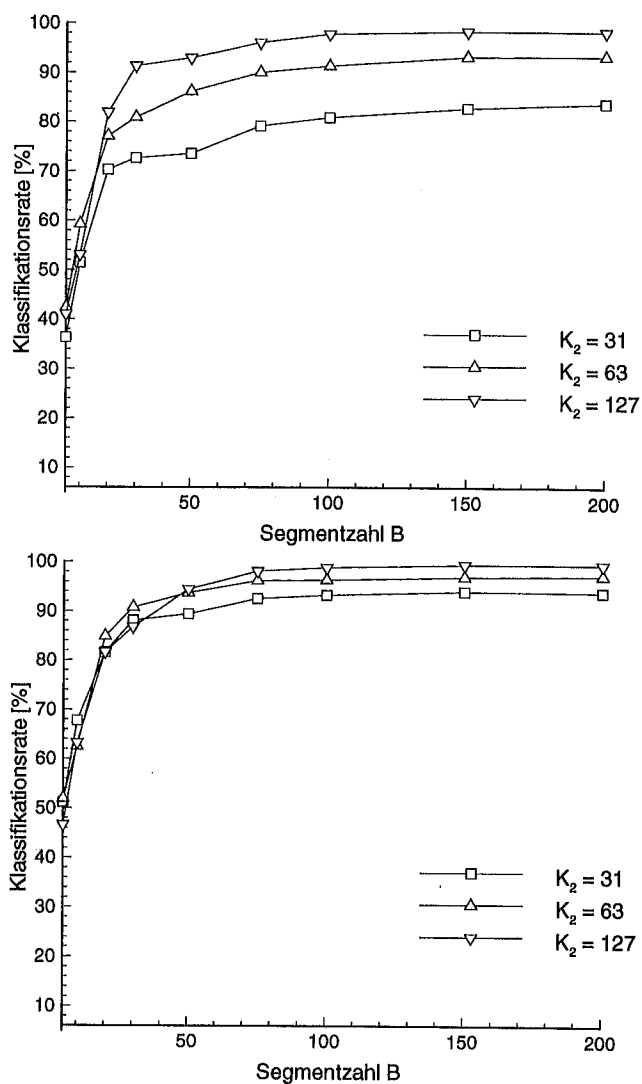


Bild 5.14: Klassifikationsergebnisse der SPWV mit $K_1 = 127$ (oben) bzw. $K_1 = 255$ (unten) bei der Merkmalsbildung nach (3.8)

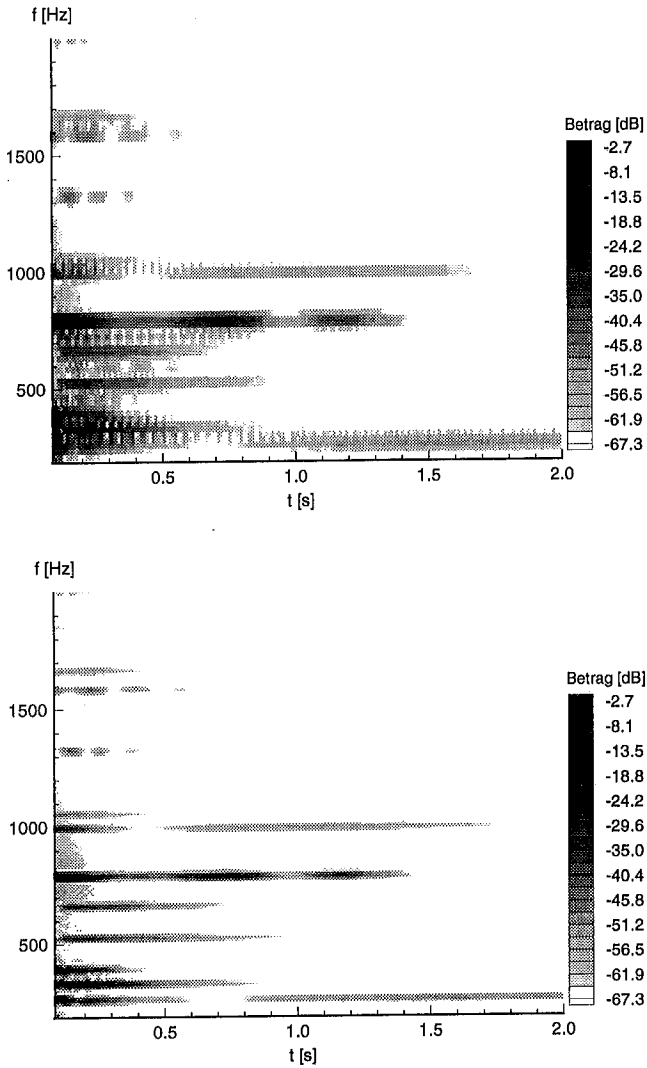
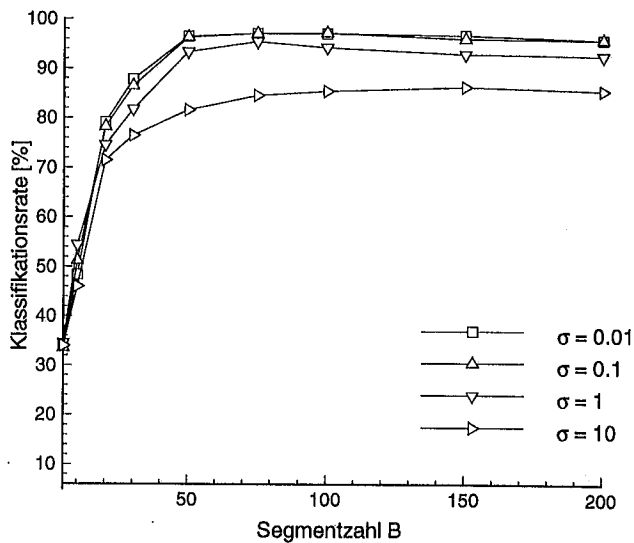
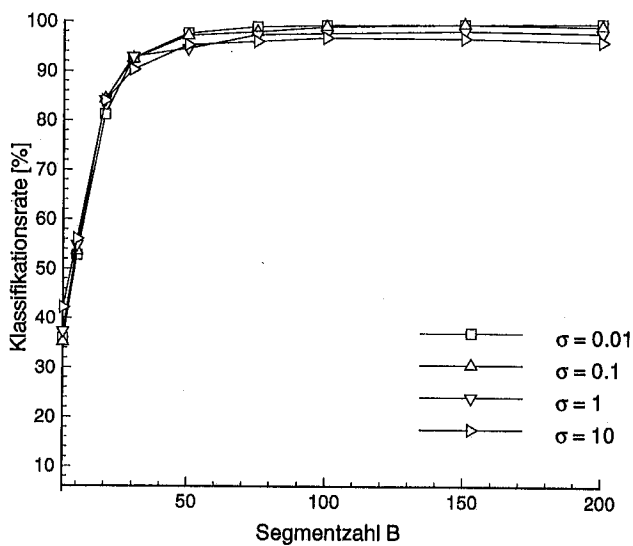


Bild 5.15: Zeit-Frequenz-Darstellungen des Tons 1 (Yamaha C2-Flügel) durch die SPWV mit $K_1 = 127$ sowie $K_2 = 63$ (oben) und $K_2 = 255$ (unten)

Bild 5.16: Klassifikationsergebnisse der CWV für $K_1 = K_2 = 127$ Bild 5.17: Klassifikationsergebnisse der CWV für $K_1 = K_2 = 255$

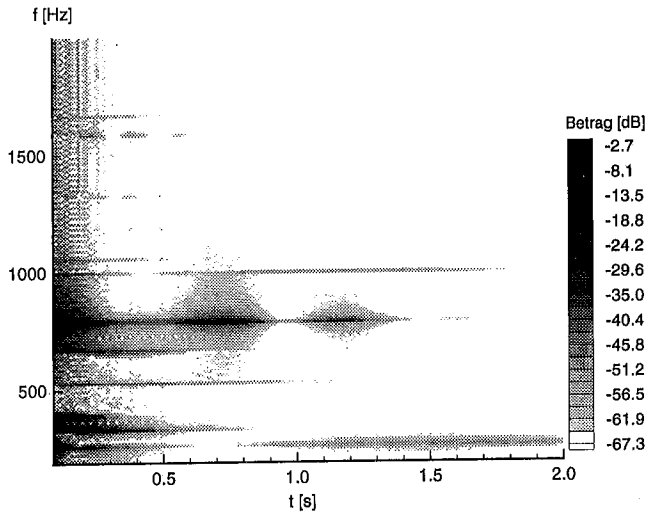


Bild 5.18: Zeit-Frequenz-Darstellungen des Tons 1 (Yamaha C2-Flügel) durch die CWV mit $K_1 = K_2 = 255$ sowie $\sigma = 0.01$

5.1.5 Modellbasierte Spektralschätzverfahren

Für die Modellierung der Eingangsfolge durch eine Folge von AR-Prozessen wurden die Parameter nach Tabelle (5.10) verwendet. Die Wahl der Modellordnung durch das AIC oder FPE-Kriterium war nicht möglich, da für alle zu einem Test herangezogenen Signalausschnitte die Funktionen (2.173) und (2.174) kein Minimum aufwiesen. Daher wurden die Modellordnungen so ausgewählt, daß das Modell 3 ($P = 6$) bis 24 ($P = 48$) spektrale Maxima erzeugen kann.

Modellordnung P :	6	12	18	24	48
Abschnittslänge L :	512				
DFT-Länge N :	512				
Dezimation M :	512				

Tabelle 5.10: Parameterwahl bei der Modellierung durch AR-Prozesse

Die Resultate der Klassifikation sind in Bild 5.19 bei Nutzung der geschätz-

ten Leistungsdichtespektren (LDS) und in Bild 5.20 bei Verwendung der Reflexionskoeffizienten als Merkmale dargestellt.

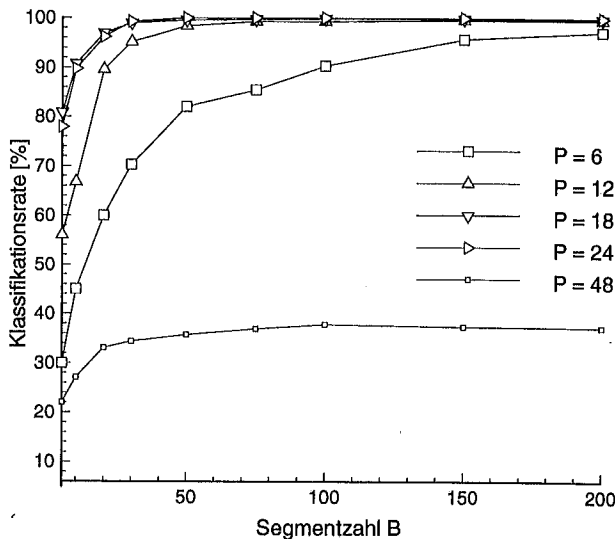


Bild 5.19: Klassifikationsergebnisse der AR-Modelle bei Verwendung der geschätzten Leistungsdichtespektren als Merkmale

Die Ergebnisse weisen signifikante Unterschiede auf. Bei der Wahl des LDS als Merkmalsquelle ist eine deutliche Abhängigkeit von der gewählten Modellordnung zu erkennen. Für eine Ordnung $12 \leq P \leq 24$ und Segmentzahlen $B \geq 30$ werden hohe Erkennungsleistungen erzielt.

Für $P = 6$ muß eine hohe Segmentzahl gewählt werden, um vergleichbare Ergebnisse zu erreichen. Die Modellordnung $P = 48$ dagegen liefert mit Klassifikationsraten unter 40% unbefriedigende Resultate. Die in Bild 5.21 abgebildeten Visualisierungen der Merkmale zeigen, daß für $P = 12$ der Klavierton in seinem spektralen Gehalt nur unzureichend beschrieben wird, während für $P = 48$ Grundschwingungen und Oberwellen klar zu erkennen sind. Real nicht vorhandene spektrale Spitzen sind nicht festzustellen.

Der Grund für die geringen Erkennungsraten bei $P = 48$ mag darin liegen, daß die Varianz der Schätzung bei gleicher Abschnittslänge L größer ist als bei $P = 12$. Wegen des ausgeprägten sinusförmigen Charakters der Signale liegen die Pole des Modells dicht am Einheitskreis. Geringe Verschiebungen erzeugen bei der Schätzung des Leistungsdichtespektrums große

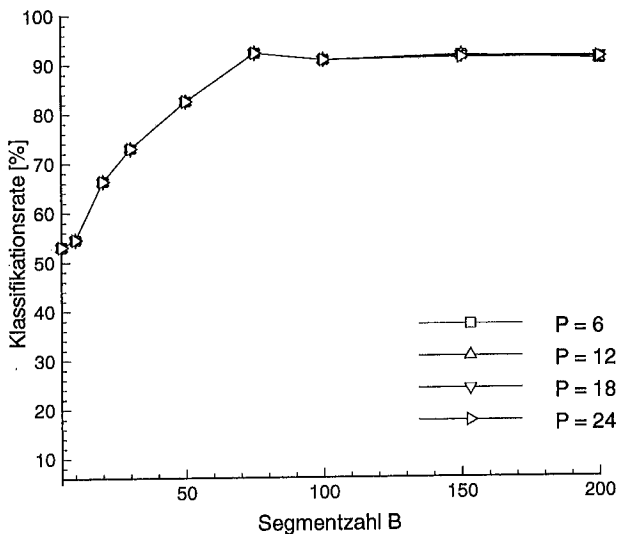


Bild 5.20: Klassifikationsergebnisse der AR-Modelle bei Verwendung der geschätzten Reflexionskoeffizienten als Merkmale

Änderungen. Die für $P = 48$ erreichte genauere Beschreibung der spektralen Eigenschaften des Eingangssignals, wie in Bild 5.21 erkennbar, ist für die Unterscheidung der Klaviere ohne Belang.

Die Reflexionskoeffizienten sind robuster gegenüber Änderungen der AR-Koeffizienten. Dies zeigt sich in den Resultaten: Für die Reflexionskoeffizienten ist keine Abhängigkeit von der Modellordnung erkennbar.

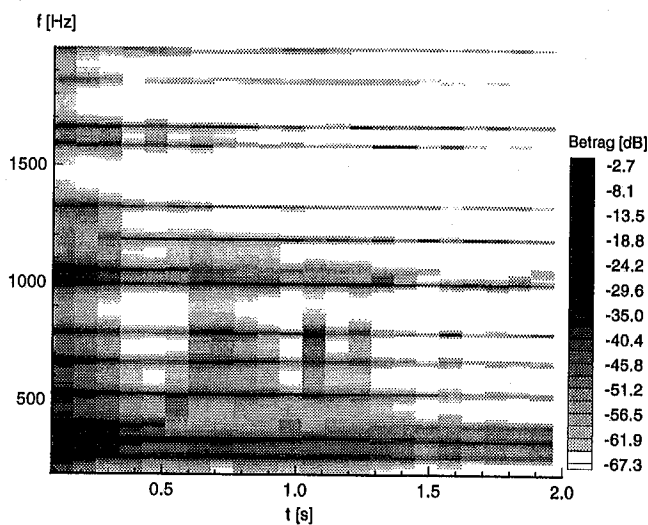
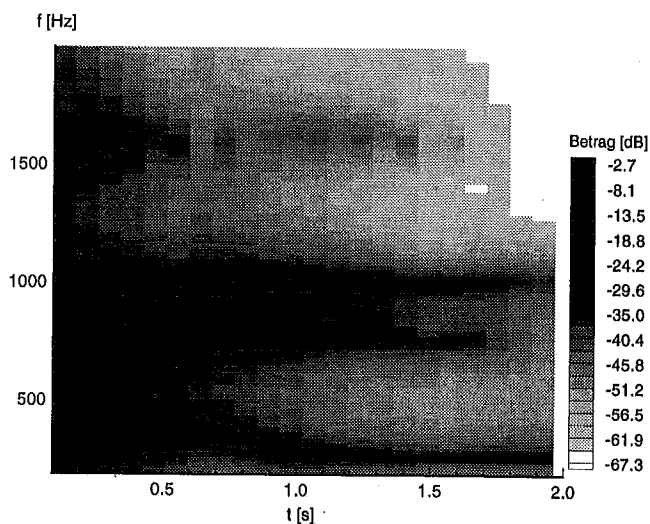


Bild 5.21: Zeit-Frequenz-Darstellung des Tons 1 (Yamaha C2-Flügel) durch eine Folge von AR-Modellen mit $P = 12$ (oben) und $P = 48$ (unten)

5.2 Klassifikationsergebnisse für die Anschlagphase

Für die Erkennung der Klaviere durch die Anschlagphase wurde der Klassifikator nach Tabelle 5.11 parametrisiert.

B	:	5	10	15	20	30	50	60
V	:				1			
Δd_{min}	:				0,1			

Tabelle 5.11: Wahl von Segmentanzahl B und -länge V sowie der Rückweisungsschwelle Δd_{min} für den Klassifikator

Die Diskussion der Ergebnisse beschränkt sich auf den Fall allgemeiner, translationsinvarianter Filterbänke sowie der Kurzzeit-Fourier-Transformation und der modellbasierten Spektralschätzverfahren. Die Verfahren der Cohen-Klasse erwiesen sich als völlig ungeeignet und erreichten nahezu unabhängig von den gewählten Parametern Klassifikationsraten von 5 – 15%.

5.2.1 Kurzzeit-Fourier-Transformation und modellbasierte Spektralschätzverfahren

Die Wahl der AR-Modellordnung wurde wie in Abschnitt 5.1.5 vorgenommen. Die übrigen Parameter wurden der Länge des Signalausschnitts so angepaßt, daß die Eingangsfolge ohne Aufteilung bearbeitet wurde und nur ein AR-Modell errechnet wird.

Modellordnung P	:	6	12	18	24	48
Abschnittslänge L	:				100	
DFT-Länge N	:				128	
Dezimation M	:				100	

Tabelle 5.12: Parameterwahl bei der Modellierung durch AR-Prozesse

Die Kurzzeit-Fourier-Transformation wurde so parametrisiert, daß die Eingangsfolge einer DFT unterzogen wurde, d.h. als Fensterfolge $\{v(k)\}_{k=0}^{K-1}$ wurde eine Rechteckfolge verwendet. Es wird darauf hingewiesen, daß auch

andere Fensterfolgen in der Klassifikation eingesetzt wurden. Die Verwendung einer Rechteckfolge als Fenster bei der KFT lieferte gegenüber anderen Fensterfolgen (Hanning, Hamming, Kaiser-Fenster [Har78]) die besten Resultate. Der Abfall der übrigen Fensterfolgen an den Rändern ist offenbar für die Klassifikation kritisch. Dies mag insbesondere daran liegen, daß dadurch das Ende der Einschwingphase ausgeblendet wird. Die übrigen Parameter wurden wie in Tabelle 5.13 gewählt.

Kanalzahl N	: 128
Filtertyp	: Rechteckfolge
Filterlänge K	: 100

Tabelle 5.13: Parameterwahl bei der KFT (Parametrisierung einer DFT)

Die Resultate der Klassifikation sind bei Nutzung der geschätzten Leistungsdichtespektren als Merkmale mit denen der KFT in Bild 5.22 dargestellt.

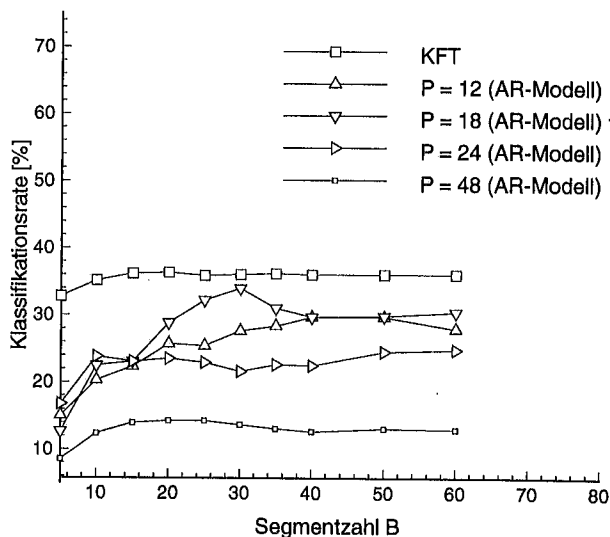


Bild 5.22: Klassifikationsergebnisse der KFT sowie des AR-Modell bei Verwendung des geschätzten Leistungsdichtespektrums als Merkmal

Im Vergleich erreicht die KFT bessere Klassifikationsraten als das AR-Modell mit 36%.

Bei dem AR-Modell werden wie in Abschnitt 5.1.5 die besten Resultate für die Ordnungen $12 \leq B \leq 24$ erreicht.

Die Robustheit der Reflexionskoeffizienten zeigt sich auch bei der Klassifikation der Anschlagphase: Werden die geschätzten Reflexionskoeffizienten als Merkmale verwendet, steigen die Erkennungsraten auf über 60%. In Bild 5.23 sind die Resultate für $P \geq 12$ dargestellt. Die unterschiedlichen Längen der Graphen resultieren aus der mit der Ordnung wachsenden Anzahl verfügbarer Merkmale.

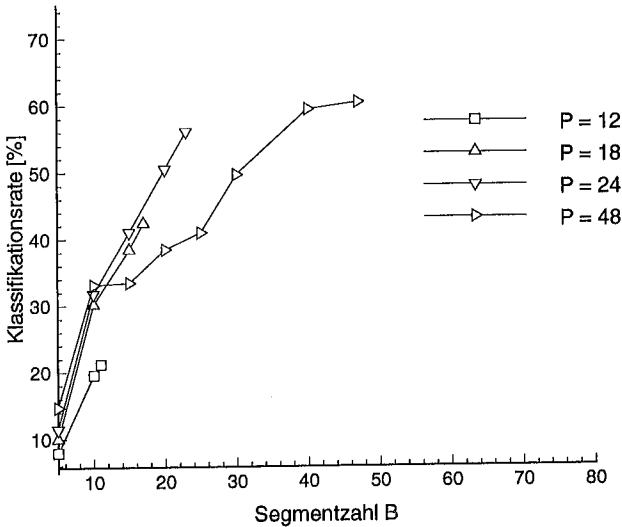


Bild 5.23: Klassifikationsergebnisse der AR-Modelle bei Verwendung der Reflexionskoeffizienten als Merkmal

5.2.2 Allgemeine translationsinvariante Analysefilterbänke

Die Filterbank mit Baumstruktur wurde zuerst für den Zerlegungsbereich $\mathcal{W} = \{0\}$ und für eine klassenunabhängige Merkmalss Selektion bzw. Klassifikation verwendet. Die in Tabelle 5.14 angegebenen Parameter wurden für die Merkmalsextraktion eingesetzt. Die Musterbildung wurde mit den Merkmalen nach (3.7) durchgeführt.

In den Klassifikationsläufen zeigten sich kaum Unterschiede für die be-

Filtertyp	: Daubechies bzw. optimierte Filter ($M = 2$)							
Zerlegungstiefe T	1	2	3	4	1	2	3	4
Filterlänge K	10	10	10	10	20	20	20	20

Tabelle 5.14: Parameterwahl für die Filterbank mit Baumstruktur

trachteten Filterlängen bzw. Filtertypen [DJ98a]. Stellvertretend sind die Resultate für Daubechies-Filter der Länge $K = 10$ in Bild 5.24 dargestellt.

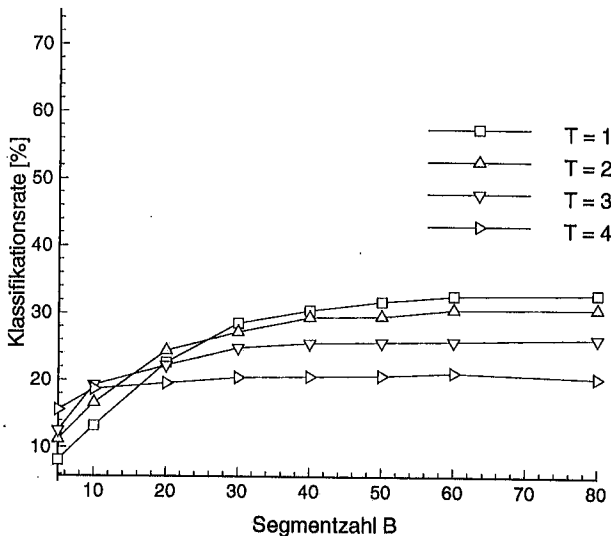


Bild 5.24: Klassifikationsergebnisse der Filterbank mit Baumstruktur (Daubechies-Filter, $K = 10$) in Abhängigkeit von der Zerlegungstiefe T

Es ist zu erkennen, daß bei besonders kleiner Zerlegungstiefe T die besten Klassifikationsraten bei etwa $B \geq 50$ Segmenten bzw. Merkmalen die besten Ergebnisse erzielen. Im Gegensatz zu der Analyse der vollständigen Klaviertöne ist es nicht möglich, durch Betragsbildung bei der Musterbildung nach (3.8) die Verschiebungsabhängigkeit zu kompensieren, da wegen des geringen Umfang an Abtastwerten des Eingangssignals eine höhere Zerlegungstiefe als $T = 4$ nicht sinnvoll ist. Man beachte, daß für $K = 10$ und $T = 4$ die Impulsantwort des resultierenden Filters der letzten Zerlegungsstufe nach (2.90) bereits 136 Elemente umfaßt. Vielmehr wird durch die

Betragsbildung, wie Klassifikationsläufe gezeigt haben, die Unterscheidbarkeit der Klaviere verringert. Die so erzielten Klassifikationsraten lagen bei lediglich 10 – 15%.

Weitere Klassifikationsläufe wurden für eine Filterbank mit Baumstruktur bei Wahl des Verschiebungsbereichs $\mathbb{W} = \{-R, -(R-1), \dots, 0, \dots, R-1, R\}$ durchgeführt. Die Auslegung von R ist in Tabelle 5.15 angegeben. In diesem Fall wurden die Parameter R mit allen Einstellungen für die Zerlegungstiefe T permutiert, d.h. es wurden 18 Läufe durchgeführt.

Filtertyp	:	Daubechies					
Parameter R	:	2	5	8	10	15	20
Zerlegungstiefe T	:	1	2			3	
Filterlänge K	:	10					

Tabelle 5.15: Parameterwahl für die translationsinvariante Filterbank mit Baumstruktur

Die Erweiterung der Filterbank mit Baumstruktur durch translationsinvariante Filterung und Klassifikation führt zu einem sprunghaften Anstieg der Erkennungsergebnisse. Die Klassifikationsraten für die Zerlegungstiefen $T = 1$ und $T = 2$ sowie Verschiebungsparameter $R \leq 10$ sind in den Bildern 5.25 bzw. 5.26 dargestellt.

Bereits für $R = 2$ werden um den Faktor vier höhere Erkennungsleistungen erreicht. Die Erweiterung des Verschiebungsbereichs führt zu geringfügig besseren Resultaten, die für $R \geq 10$ nahezu unverändert bleiben. Bei einem Vergleich beider Bilder ist zu erkennen, daß die Ergebnisse für die Zerlegung $T = 1$ besser sind. Da in dem Klassifikationslauf mit $T = 3$ schlechtere Klassifikationsergebnisse als mit $T = 2$ erzielt wurden, muß auch auf die Frage eingegangen werden, ob die Eingangsfolge selbst die besten Merkmale liefert. Denn je größer die zeitliche Auflösung der Merkmale bei der Zerlegung ist, um so höher sind die erreichten Klassifikationsraten. Bild 5.27 zeigt, daß für große Segmentzahlen $B \geq 50$ die Klassifikationsergebnisse mit denen der Filterbank mit Baumstruktur für $T = 1$ übereinstimmen, für einen kleineren Merkmalsumfang die Merkmalsextraktion durch die Filterbank jedoch durchaus einen Vorteil bietet. Der hohe Rechenaufwand (vergl. Abschnitt 6.5) wird die Einsatzmöglichkeiten translationsinvarianter Filterung dennoch beschränken.

Die klassenabhängige Merkmalsselektion und Klasifikation nach Abschnitt 3.2 wird für die Parametrisierung nach Tabelle 5.16 betrachtet.

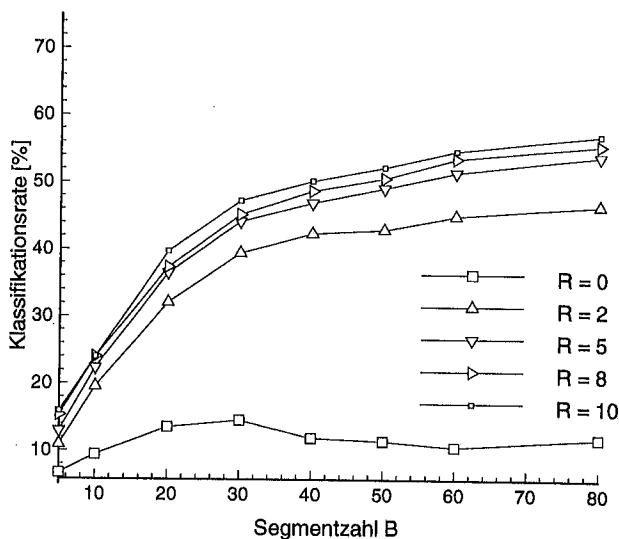


Bild 5.25: Klassifikationsergebnisse der Filterbank mit Baumstruktur (Daubechies-Filter, $K = 10$, $T = 1$)

Filtertyp	:	Daubechies					
Parameter R	:	2	5	8	10	15	20
Zerlegungstiefe T	:	1			2		
Filterlänge K	:	10					

Tabelle 5.16: Parameterwahl für die translationsinvariante Filterbank mit Baumstruktur bei klassenweiser Merkmalsselektion und Klassifikation

In den Klassifikationsläufen hat sich gezeigt, daß unabhängig von der Zerlegungstiefe T und der gewählten Strategie zur Signalverschiebung in der Trainingsphase nach Abschnitt 3.2.1 die Ergebnisse der klassenunabhängigen Merkmalsselektion und Klassifikation nicht erreicht werden konnten. Für $T = 2$ und Strategie B sind die Ergebnisse der Klassifikation in Bild 5.28 abgebildet.

Die Robustheit der klassenunabhängigen Merkmalsselektion mag folgenden Grund haben: Die klassenweise erfolgende Merkmalsselektion ermittelt die Merkmale, die die Abstände zwischen den Mustern zweier Klassen im Mittel

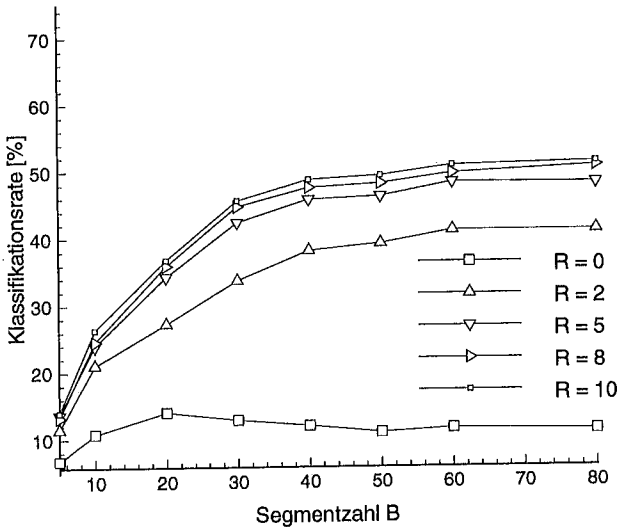


Bild 5.26: Klassifikationsergebnisse der Filterbank mit Baumstruktur (Daubechies-Filter, $K = 10$, $T = 2$)

maximieren. Dies setzt voraus, daß die Muster einer Klasse bereits in guter Näherung eine Ballung im Musterraum bilden. In den Trainingsläufen hat sich gezeigt, daß der mittlere euklidische Abstand zweier Signale einer Klasse nach (3.9) fast unabhängig von der Klasse einen Wert um 1 annahm. Durch die Optimierung der Signalverschiebungen nach Strategie B konnte dieser Wert in Abhängigkeit von der Klasse zwischen 15% und 50% gesenkt werden. Dadurch ergab sich eine stärkere Ballung auch im Musterraum, die aber für die Klassifikation nicht ausreichend war:

1. Durch die geringe Ballung der Muster in der Trainingsphase wurden Merkmale ausgewählt, die nicht zur Charakterisierung der Klassenunterschiede geeignet waren.
2. Der Einfluß der translationsinvarianten Filterung und Klassifikation überwog den der Ballungsverbesserung: Bei klassenweiser wie bei klassenunabhängiger Klassifikation fielen die Entscheidungen zugunsten des Repräsentanten einer Klasse bei Abständen von etwa 0,3 - 0,6.

Es ist anzunehmen, daß der Einsatz einer klassenweisen Merkmalsselektion und Klassifikation dann sinnvoll ist, wenn die Signale einer Klasse bereits eine

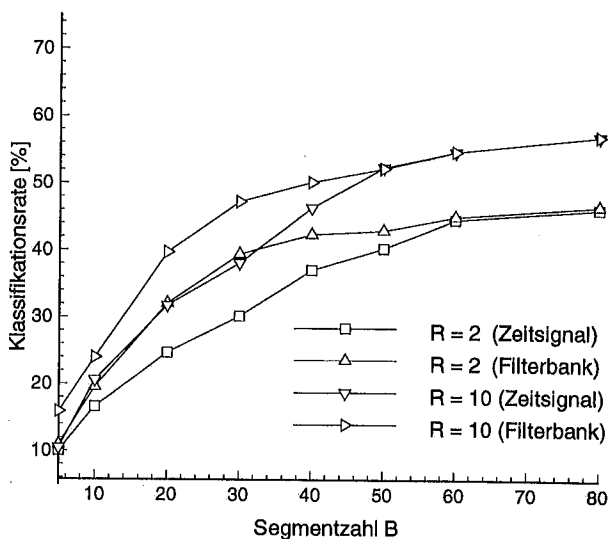


Bild 5.27: Vergleich der Klassifikationsergebnisse bei Nutzung der Eingangsfolge als Merkmalsmenge mit denen einer Filterbank mit Baumstruktur (Daubechies-Filter, $T = 1$, $K = 10$)

enge Ballung bilden oder diese z.B. durch die Wahl der Signalverschiebung nach (3.9) hergestellt werden kann.

5.3 Vergleich der Klassifikationsergebnisse

Für einen Vergleich der Klassifikationsergebnisse in Abhängigkeit von den Merkmalsextraktionsverfahren werden die Resultate für die für jedes Verfahren optimalen Parameter zusammen dargestellt.

Zuerst wird der Vergleich für die Klassifikationsraten, die mit den vollständigen Klaviertönen ermittelt wurden, durchgeführt. Bild 5.29 zeigt die Klassifikationsresultate für die Kurzzeit-Fourier-Transformation (KFT), die dyadische Wavelet-Transformation (DWT), die Filterbank mit Baumstruktur (B), die Choi-Williams-Verteilung (CWV) als Vertreter der Cohen-Klasse und für aus den geschätzten Leistungsdichtespektren abgeleiteten Merkmale der AR-Modellierung (AR).

Bei KFT, CWV und AR-Modellen werden die besten Ergebnisse erzielt.

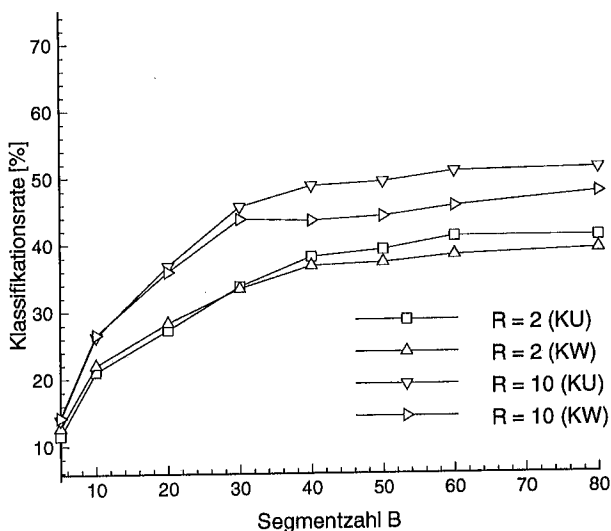


Bild 5.28: Klassifikationsergebnisse bei klassenweiser (KW) und klassenunabhängiger (KU) Merkmalsselektion und Klassifikation (Daubechies-Filter, $T = 2$, $K = 10$)

Geringfügig schlechtere Resultate werden mit einer Filterbank mit Baumstruktur erreicht. Lediglich die Erkennungsraten der DWT lagen deutlich unter denen der übrigen Verfahren. Bei der Entscheidung für ein Verfahren muß der benötigte Rechenaufwand einbezogen werden. In Tabelle 5.17 sind die aus Anhang 6.5 ermittelten Rechenoperationen pro Eingangstakt für alle Verfahren aufgeführt, wobei für die AR-Modelle der zusätzliche Aufwand für Berechnung der FFT berücksichtigt wurde.

KFT	DWT	B	CWV	AR
43	45	443	352	260

Tabelle 5.17: Vergleich des Rechenaufwand der Merkmalsextraktionsverfahren für die optimalen Parametereinstellungen (Klassifikation der vollständigen Klaviertöne)

Die KFT liefert das beste Resultat bei sehr geringem Aufwand und ist für die vorliegende Klassifikationsaufgabe die beste Wahl. Im Vergleich dazu liegt der Aufwand für eine Filterbank mit Baumstruktur bzw. für die CWV um

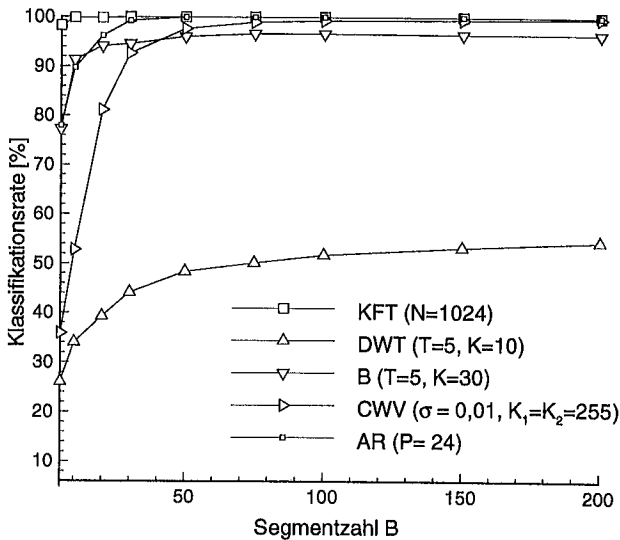


Bild 5.29: Vergleich der Merkmalsextraktionsverfahren bei Klassifikation der vollständigen Klaviertöne

eine Größenordnung darüber.

Ein Vergleich der Klassifikationsresultate für die Anschlagphase ist in Bild 5.30 dargestellt, wobei bei den AR-Modellen die Reflexionskoeffizienten als Merkmale verwendet wurden. Die Merkmale aus AR-Modellen erreichen schon für einen Merkmalsumfang von $B = 48$ die besten Ergebnisse, während die dyadische Wavelet-Packet-Transformation (DWPT) bei klassenunabhängiger Merkmalsselektion (KU) vergleichbare Resultate erst für $B = 80$ Merkmale erzielt. Die KFT erreicht deutlich geringere Klassifikationsraten.

In Tabelle 5.18 ist der Rechenaufwand für jedes eingesetzte Merkmalsextraktionsverfahren eingetragen.

Bei der DWPT ist zu beachten, daß durch die verschiebungsabhängige Mustererzeugung pro Signal nicht nur ein Muster klassifiziert werden muß, sondern $2R + 1$. Durch diesen zusätzlichen Aufwand im Klassifikationsablauf ist die Merkmalerzeugung durch die AR-Modelle für die Klassifikation der Anschlagphase am besten geeignet.

Sowohl für die vollständigen Klaviertöne als auch für die Anschlagphase ist

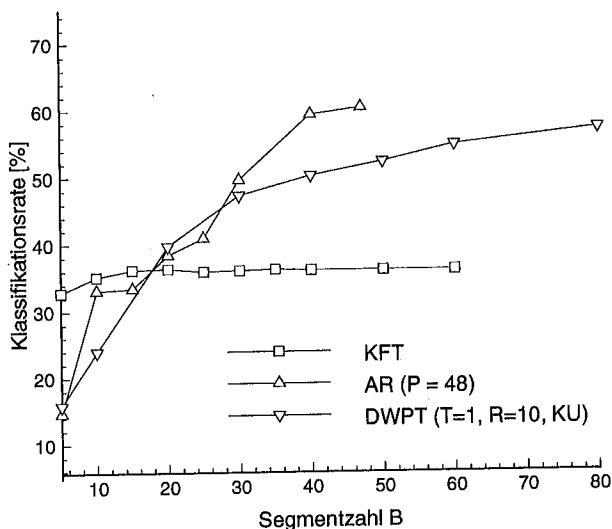


Bild 5.30: Vergleich der Merkmalsextraktionsverfahren bei Klassifikation der Anschlagphase

KFT	AR	DWPT
36	74	45

Tabelle 5.18: Vergleich des Rechenaufwand der Merkmalsextraktionsverfahren für die optimalen Parametereinstellungen (Klassifikation der Anschlagphase)

ein Vorteil orthogonaler Filterbänke als Merkmalsextraktionsverfahren nicht zu erkennen. Eine offene Frage ist, ob reale Aufgabenstellungen existieren, in denen die zu klassifizierenden Signale vorteilhaft mit diesen Verfahren verarbeitet werden können. Diese Arbeit hat gezeigt, daß auch in zukünftigen Arbeiten ein umfassender Vergleich mit anderen Methoden der Zeit-Frequenz-Signalanalyse unumgänglich ist.

6 Anhang

6.1 Eigenschaften eines Hilbertraums

Die Funktion $s(t)$ bzw. die Folge $\{s(m)\}_m$ werden als Elemente des Signalraumes $L^2(\mathbb{R})$ bzw. $l^2(\mathbb{Z})$ angenommen:

$$s(t) \in L^2(\mathbb{R}) \leftrightarrow \int_{-\infty}^{\infty} |s(t)|^2 dt < \infty$$

$$\{s(m)\}_m \in l^2(\mathbb{Z}) \leftrightarrow \sum_{m=-\infty}^{\infty} |s(m)|^2 < \infty.$$

Beide Räume sind Vektorräume, ihre Elemente lassen sich als Vektoren in einem abzählbar unendlich dimensionalen Vektorraum deuten und erfüllen Eigenschaften, die in [VK95] diskutiert werden. Diese Interpretation erlaubt die gemeinsame Beschreibung durch einen Vektor $\bar{s}$, der im Raum $L^2(\mathbb{R})$ der Funktion $s(t)$ und im Folgenraum $l^2(\mathbb{Z})$ der Folge $\{s(m)\}_m$ entspricht. Jedes Element eines abzählbar unendlich dimensional Vektorraumes kann durch eine Linearkombination von Basisvektoren $\bar{b}_m$, die den Raum aufspannen, beschrieben werden:

$$\bar{s} = \sum_{m=-\infty}^{\infty} \alpha_m \bar{b}_m. \quad (6.1)$$

Dabei heißen die Größen α_m Entwicklungskoeffizienten bezüglich der Basisvektoren.

Ein Raum heißt abgeschlossen, wenn er alle seine Häufungspunkte ¹ enthält [Hen56]. Ein Hilbertraum ist ein vollständiger ² Innenproduktraum. Über das Innenprodukt sind die Norm und die Metrik des Raumes definiert. Das Innenprodukt erlaubt, die Eigenschaft der Orthogonalität zweier Vektoren zu beschreiben. Über das Innenprodukt wird der Begriff der Orthogonalität eingebracht und führt auf orthogonale Basen und Räume, deren Eigenschaften in der Merkmalsextraktion (Kapitel 2) ausgenutzt werden.

¹Man nennt einen Punkt Häufungspunkt eines metrischen Raumes, wenn in jeder Umgebung des Punktes im Sinne der Metrik unendlich viele Punkte des Raumes liegen.

²D.h. jede Cauchy-Folge besitzt einen Grenzwert, der in dem Raum enthalten ist [Heu86].

6.1.1 Innenprodukt

Das Innenprodukt

$$\langle \bar{s}_1, \bar{s}_2 \rangle$$

ist die Abbildung zweier Vektoren $\bar{s}_1$ und $\bar{s}_2$ auf eine komplexe Zahl mit den Eigenschaften [Fli93]:

$$\begin{aligned} \langle \bar{s}_1, \bar{s}_2 \rangle &= \langle \bar{s}_2, \bar{s}_1 \rangle^*, \\ \langle \alpha_1 \bar{s}_1 + \alpha_2 \bar{s}_2, \bar{s}_3 \rangle &= \alpha_1 \langle \bar{s}_1, \bar{s}_3 \rangle + \alpha_2 \langle \bar{s}_2, \bar{s}_3 \rangle, \quad \alpha_1, \alpha_2 \in \mathbb{C} \\ \langle \bar{s}_1, \bar{s}_1 \rangle &\geq 0, \quad \langle \bar{s}_1, \bar{s}_1 \rangle = 0 \leftrightarrow \bar{s}_1 = \bar{0}, \end{aligned}$$

wobei $\bar{0}$ dem Nullvektor entspricht. In den betrachteten Signalräumen ist das Innenprodukt definiert durch:

$$\begin{aligned} L^2(\mathbb{R}) : \quad \langle \bar{s}_1, \bar{s}_2 \rangle &= \int_{-\infty}^{\infty} s_1(t) s_2^*(t) dt, \\ l^2(\mathbb{Z}) : \quad \langle \bar{s}_1, \bar{s}_2 \rangle &= \sum_{m=-\infty}^{\infty} s_1(m) s_2^*(m). \end{aligned}$$

6.1.2 Norm und Metrik

Norm und Metrik sind über das Innenprodukt definiert. Die Norm eines Vektors $\bar{s}$ ist gegeben durch:

$$\|\bar{s}\| = \sqrt{\langle \bar{s}, \bar{s} \rangle}.$$

Die Metrik im Vektorraum entspricht dem Abstand zweier Vektoren $\bar{s}_1$ und $\bar{s}_2$. Er ist über die Norm des Differenzvektors definiert:

$$d(\bar{s}_1, \bar{s}_2) = \|\bar{s}_1 - \bar{s}_2\|.$$

6.1.3 Orthogonalität und orthogonale Basen

Gilt für zwei Vektoren $\bar{s}_1$ und $\bar{s}_2$

$$\langle \bar{s}_1, \bar{s}_2 \rangle = 0,$$

heißen die Vektoren orthogonal. Mit Hilfe des Innenproduktes läßt sich aus jeder Basis des Hilbertraumes eine orthonormale Basis $\bar{b}_m$ ableiten, d.h.

$$\langle \bar{b}_i, \bar{b}_j \rangle = \delta_{i,j}, \quad i, j \in \mathbb{Z}$$

mit $\delta_{i,j}$ als Kronecker-Symbol [BS87]. Nur in diesem Fall sind die Entwicklungskoeffizienten (6.1) des Vektors $\bar{s}$ bezüglich dieser Basis gegeben durch:

$$\alpha_m = \langle \bar{s}, \bar{b}_m \rangle. \quad (6.2)$$

Ein Hilbertraum heißt separabel, wenn die Menge der Basisvektoren abzählbar ist, wie in (6.1) und (6.2) bereits vorausgesetzt wird. Die Signalmräume $L^2(\mathbb{R})$ und $l^2(\mathbb{Z})$ sind separabel [Heu86, S. 179, S.156].

6.1.4 Orthogonale Unterräume

Es sei V ein Hilbertraum und U ein vollständiger Unterraum von V . U liegt dicht in V , wenn der Abschluß von U gleich V ist [Heu86, S. 62].

Im folgenden sei U ein echter Unterraum von V . Der Raum W heißt orthogonales Komplement von U bezüglich V , wenn alle Elemente aus U orthogonal zu den Elementen von W sind und V die direkte Summe der Unterräume U und W ist [Hen56]:

$$V = U \oplus W$$

mit $\oplus$ als Symbol der direkten Summe. Jedes Element $\bar{s} \in V$ läßt sich als Summe zweier, eindeutig bestimmter Elemente der zueinander orthogonalen Unterräume U und W darstellen:

$$\bar{s} = \bar{s}_u + \bar{s}_w, \quad \bar{s}_u \in U, \bar{s}_w \in W.$$

Die Vektoren $\bar{s}_u$ und $\bar{s}_w$ lassen sich durch Projektion des Vektors $\bar{s}$ auf die Räume U und W , d.h. durch Projektionsoperatoren P_u und P_w , bestimmen, also:

$$\begin{aligned}\bar{s}_u &= P_u \bar{s}, \\ \bar{s}_w &= P_w \bar{s}.\end{aligned}$$

Projektionsoperatoren im Signalraum $L^2(\mathbb{R})$ heißen ideale Filter [Hen56].

6.2 Elemente und Eigenschaften von Multiraten-Systemen

Ein zeitdiskretes System ist ein Multiraten-System, wenn innerhalb des Systems mit verschiedenen Abtastraten gearbeitet wird.

Die z -Transformierte einer Folge $\{x(m)\}_m$ ist definiert durch:

$$X(z) = \sum_{m=-\infty}^{\infty} x(m)z^{-m}, \quad z \in \mathbb{C}.$$

Dieser Zusammenhang wird auch kurz durch das Symbol

$$\{x(m)\}_m \xrightarrow{z} X(z)$$

ausgedrückt.

Die z -Transformierte der Faltung zweier Folgen $\{x_1(m)\}_m$ und $\{x_2(m)\}_m$

$$y(m) = \sum_{k=-\infty}^{\infty} x_1(k)x_2(m-k)$$

ist das Produkt ihrer z -Transformierten:

$$Y(z) = X_1(z)X_2(z).$$

Weitere Eigenschaften der z -Transformation sind in [Jon91] erläutert. Viele Operationen in Multiraten-Systemen lassen sich vorteilhaft im z -Bereich interpretieren.

Zu den grundlegenden Elementen zählen Verzögerungsglieder sowie Auf- und Abwärtstaster, die zunächst beschrieben werden. Ein Verzögerungsglied um L Takte ist in Bild 6.1 dargestellt.

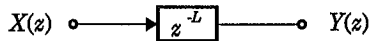


Bild 6.1: Symbol für ein Verzögerungsglied

Diese Elemente werden benötigt, um die Polyphasenkomponenten eines Signals in Abschnitt 6.2.4 einzuführen. Die Verknüpfung der Faltung mit dem Innenprodukt des Folgenraumes $l^2(\mathbb{Z})$ bei einer Dezimation wird abschließend in Abschnitt 6.2.5 diskutiert.

6.2.1 Aufwärtstaster

Bei einem Aufwärtstaster um den Faktor M werden zwischen zwei Elementen der Eingangsfolge $\{x(m)\}_m$ $M - 1$ Nullen eingefügt. Die Ausgangsfolge $\{y(m)\}_m$ ist gegeben durch:

$$y(m) = \begin{cases} x(m/M), & \text{für } m = nM, \quad m, n \in \mathbb{Z} \\ 0, & \text{sonst.} \end{cases}$$

Das Symbol für den Aufwärtstaster ist in Bild 6.2 dargestellt.

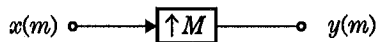


Bild 6.2: Symbol des Aufwärtstasters

Bild 6.3 zeigt die Auswirkungen im Zeitbereich für den Fall $M = 2$. Für die z -Transformierte des aufwärtsgetasteten Signals $Y(z)$ gilt:

$$\begin{aligned} Y(z) &= \sum_{m=-\infty}^{\infty} y(m) z^{-m} \\ &= \sum_{k=0}^{M-1} \sum_{m=-\infty}^{\infty} y(mM + k) z^{-mM - k}. \end{aligned} \quad (6.3)$$

Alle Terme in (6.3) mit $k \neq 0$ sind Null. Daher folgt weiter:

$$Y(z) = \sum_{m=-\infty}^{\infty} y(mM) z^{-mM}$$

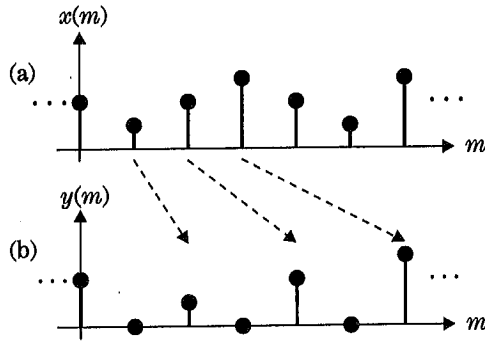


Bild 6.3: Beispiel für eine Aufwärtstastung um den Faktor $M = 2$

$$\begin{aligned}
 &= \sum_{m=-\infty}^{\infty} x(m)(z^M)^{-m} \\
 &= X(z^M)
 \end{aligned} \tag{6.4}$$

6.2.2 Abwärtstaster

Bei einem Abwärtstaster um den Faktor M wird nur jeder M -te Wert der Eingangsfolge in der Ausgangsfolge weiterverwendet:

$$y(m) = x(Mm), \quad k \in \mathbb{Z}.$$

Das Symbol für den Abwärtstaster ist in Bild 6.4 dargestellt.

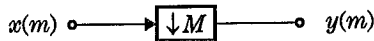
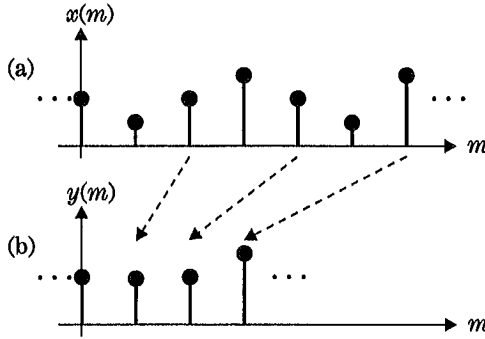


Bild 6.4: Symbol des Abwärtstasters

Ein Beispiel für $M = 2$ liefert Bild 6.5.

Zur Bestimmung der z -Transformierten der Ausgangsfolge $Y(z)$ wird zunächst die Hilfsfolge

$$C_M(m) = \frac{1}{M} \sum_{k=0}^{M-1} W_M^{-km} = \begin{cases} 1, & \text{für } m = Mn, \quad n, m \in \mathbb{Z} \\ 0, & \text{sonst} \end{cases}$$

Bild 6.5: Beispiel für eine Abwärtstastung um den Faktor $M = 2$

mit $W_M = e^{-j2\pi/M}$ definiert. Die Zwischenfolge

$$\{x_c(m)\}_m = \{C_M(m)x(m)\}_m$$

blendet alle von der Dezimation betroffenen Elemente aus. Sie steht mit der Ausgangsfolge in folgendem Zusammenhang:

$$y(m) = x_c(Mm) \text{ bzw.} \quad (6.5)$$

$$Y(z) = \sum_{m=-\infty}^{\infty} x_c(Mm)z^{-m} = \sum_{m=-\infty}^{\infty} x_c(m)z^{-m/M}. \quad (6.6)$$

Die Umformung in (6.6) ist nur zulässig, da die Elemente der Folge $\{x_c(m)\}_m$ Null sind mit Ausnahme der Indizes m , die ein Vielfaches von M sind. Für die z -Transformierte $X_c(z)$ gilt:

$$\begin{aligned} X_c(z) &= \sum_{m=-\infty}^{\infty} \left[\frac{1}{M} \sum_{k=0}^{M-1} W_M^{-km} x(m) \right] z^{-m} \\ &= \frac{1}{M} \sum_{k=0}^{M-1} \sum_{m=-\infty}^{\infty} x(m) (z W_M^k)^{-m} \\ &= \frac{1}{M} \sum_{k=0}^{M-1} X(z W_M^k). \end{aligned} \quad (6.7)$$

Aufgrund (6.6) und (6.7) folgt für $Y(z)$:

$$Y(z) = \frac{1}{M} \sum_{k=0}^{M-1} X(z^{1/M} W_M^k). \quad (6.8)$$

Die physikalische Interpretation von Abwärts- und Aufwärts-
tastung ist für den Signalfluß in Bild 6.6 und in Bild 6.7 angegeben. Die Abwärts-

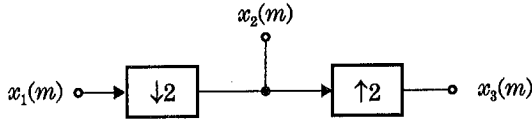


Bild 6.6: Signalfluß mit Abwärts -und Aufwärtstaster

stastung um den Faktor 2 entspricht einer Halbierung der Abtastrate auf $2\Delta t$. Durch die Aufwärts-
tastung wird der ursprüngliche Takt Δt wiederher-
gestellt.

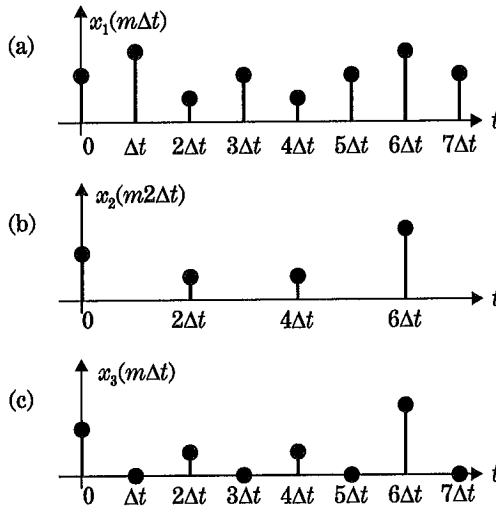


Bild 6.7: Physikalische Interpretation von Abwärts- und Aufwärtstaster

6.2.3 Äquivalenzen

In einem Multiratsensystem gelten folgende Äquivalenzen [Vai93]:

- Ein Verzögerungsglied kann unter den in Bild 6.8 dargestellten Bedingungen mit einem Ab- bzw. Aufwärtstaster in der Reihenfolge vertauscht werden.

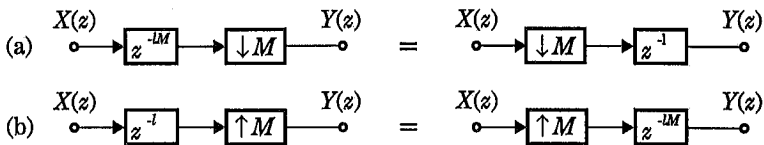


Bild 6.8: Äquivalenzen bei der Vertauschung von Verzögerungsgliedern und Abwärtstastern (a) bzw. Aufwärtstastern (b)

- Ein Filter kann unter den in Bild 6.9(a) und (b) dargestellten Bedingungen mit einem Ab- bzw. Aufwärtstaster in der Reihenfolge vertauscht werden.

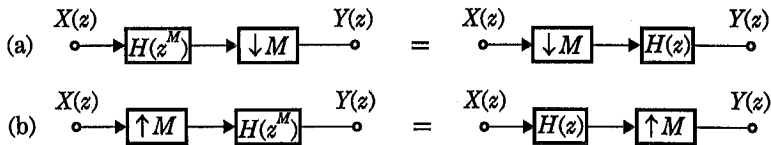


Bild 6.9: Äquivalenzen bei der Vertauschung von Filtern und Abwärtstastern (a) bzw. Aufwärtstastern (b)

6.2.4 Polyphasendarstellung eines Signals

Von zentraler Bedeutung sind in einem Multiratsensystem die Polyphasendarstellungen eines Signals bezüglich einer Dezimation um den Faktor M . Man unterscheidet Polyphasendarstellungen erster und zweiter Art [Vai93].

- Die Polyphasendarstellung erster Art der Folge $\{x(m)\}_m$ erfolgt durch die Polyphasenkomponenten:

$$x_k^{(1)}(m) = x(mM + k), \quad m \in \mathbb{Z}, k = 0, 1, \dots, M-1$$

mit der z -Transformierten $X_k^{(1)}(z)$, die zu $X(z)$ durch

$$X(z) = \sum_{k=0}^{M-1} z^{-k} X_k^{(1)}(z^M)$$

in Beziehung steht. In Bild 6.5 (b) stellt die Folge $\{y(m)\}_m$ die Polyphasenkomponente $\{x_0^{(1)}(m)\}_m$ der Folge aus Bild 6.5 (a) dar. Man kann die Erzeugung der Polyphasenkomponenten erster Art und ihre Synthese zum Originalsignal $X(z)$ wie in Bild 6.10 darstellen.

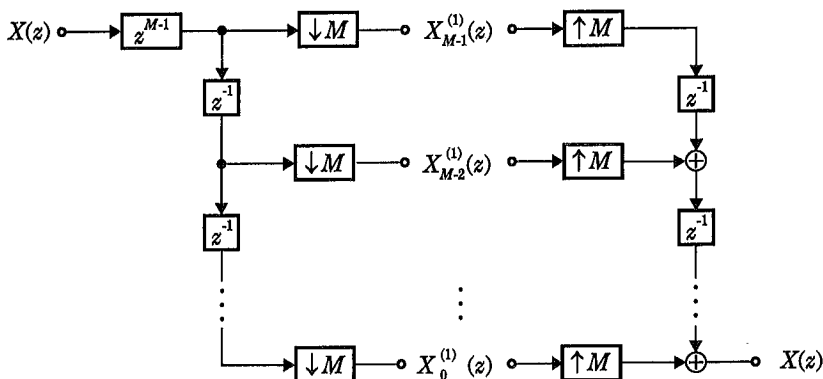


Bild 6.10: Erzeugung der Polyphasenkomponenten erster Art

- Die Polyphasendarstellung zweiter Art ist definiert durch

$$x_k^{(2)}(m) = x((m+1)M - 1 - k), \quad m \in \mathbb{Z}, \quad k = 0, 1, \dots, M-1$$

mit der z -Transformierten $X_k^{(2)}(z)$, die zu $X(z)$ durch

$$X(z) = \sum_{k=0}^{M-1} z^{-(M-1-k)} X_k^{(2)}(z^M)$$

in Beziehung steht. Man kann die Erzeugung der Polyphasenkomponenten zweiter Art und ihre Synthese zum Originalsignal $X(z)$ wie in Bild 6.11 darstellen.

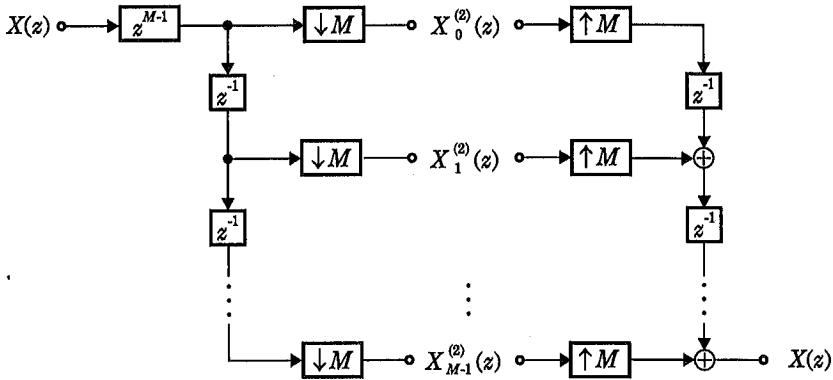


Bild 6.11: Erzeugung der Polyphasenkomponenten zweiter Art

6.2.5 Innenprodukt und Faltung bei Dezimation

Die Faltung zweier Folgen $\{x_1(k)\}_k$ und $\{x_2(k)\}_k$ lässt sich als Innenprodukt der Vektoren $\bar{x}_1 = \{x_1(k)\}_k$ und $\bar{x}_{2,m} = \{x_2(m-k)\}_m$ interpretieren:

$$\begin{aligned} y(m) &= \sum_{m=-\infty}^{\infty} x_1(k)x_2(m-k), \quad k \in \mathbb{Z}, \\ &= \langle \bar{x}_1, \bar{x}_{2,m} \rangle. \end{aligned}$$

Die Folgeelemente $y(m)$ können als Entwicklungskoeffizienten des Vektors $\bar{x}_1$ bezüglich der Vektoren $\bar{x}_{2,m}$ interpretiert werden. Im z -Bereich folgt die Darstellung:

$$Y(z) = X_1(z)X_2(z). \quad (6.9)$$

Wird die Ausgangsfolge $\{y(m)\}_m$ um den Faktor M zur Folge $\{y_d(k)\}_k$ dezimiert, gilt im Zeitbereich:

$$y_d(m) = y(mM) = \sum_{m=-\infty}^{\infty} x_1(k)x_2(Mm-k)$$

und im z -Bereich durch Einsetzen von (6.9) in (6.8):

$$Y_d(k) = \frac{1}{M} \sum_{m=0}^{M-1} X_1(zW_M^m) X_2(zW_M^m). \quad (6.10)$$

6.3 Quadrature-Mirror-Filterbank (QMF)

Für den Fall $M = 2$ erhält bei einer paraunitären Filterbank den in Bild 6.12 dargestellten Signalfluß und für das rekonstruierte Signal über (6.10) den Ausdruck :

$$\tilde{S}(z) = \frac{1}{2} \begin{bmatrix} G_0(z) & G_1(z) \end{bmatrix} \begin{bmatrix} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{bmatrix} \begin{bmatrix} S(z) \\ S(-z) \end{bmatrix}.$$

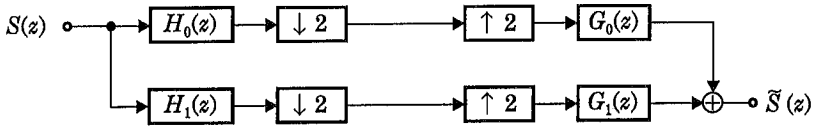


Bild 6.12: Signalfluß bei einer zweikanaligen Analyse- und Synthesefilterbank

Mit Einsetzen von $G_i(z) = z^{K-1}H_i(z^{-1})$, $i = 0, 1$, folgt ausgeschrieben:

$$\begin{aligned} \tilde{S}(z) = & \frac{1}{2} z^{-(K-1)} [H_0(z^{-1})H_0(z) + H_1(z^{-1})H_1(z)] S(z) + \\ & \frac{1}{2} z^{-(K-1)} [H_0(z^{-1})H_0(-z) + H_1(z^{-1})H_1(-z)] S(-z). \end{aligned}$$

Damit perfekte Rekonstruktion vorliegt, müssen die Bedingungen:

$$H_0(z^{-1})H_0(z) + H_1(z^{-1})H_1(z) = 1 \quad \text{und} \quad (6.11)$$

$$H_0(z^{-1})H_0(-z) + H_1(z^{-1})H_1(-z) = 0 \quad (6.12)$$

erfüllt sein. Aus (6.11) folgt, daß die Polynome $H_0(z)$ und $H_1(z)$ teilerfremd sein müssen. Andernfalls folgt mit $H_0(z) = L(z)M_0(z)$ und $H_1(z) = L(z)M_1(z)$, wobei $M_0(z)$ und $M_1(z)$ teilerfremd sind, so daß

$$L(z)L(z^{-1})(M_0(z^{-1})M_0(z) + M_1(z^{-1})M_1(z)) = 1$$

ist. Da das Polynom $L(z)L(z^{-1})$ Nullstellen aufweist, mit der Ausnahme $L(z) = 1, \forall z$, ist (6.11) für diese Nullstellen nicht erfüllt. Damit sind auch $H_0(z^{-1})$ und $H_1(z^{-1})$ bzw. $H_0(-z)$ und $H_1(-z)$ teilerfremd.

Damit scheidet die Lösung $H_1(z) = -H_0(z)$ für (6.12) aus. Die einzige Alternative zur Lösung von (6.12) besteht darin, daß die Polynome $H_1(z^{-1})$ und $H_0(-z)$ die gleichen Nullstellen, deren Betrag ungleich 0 ist, aufweisen oder

$$\frac{H_0(z^{-1})}{H_1(z^{-1})} = -\frac{H_1(-z)}{H_0(-z)}$$

und damit für $H_1(z^{-1}) = \pm z^{K'-1} H_0(-z)$, $K' \in \mathbb{Z}$ bzw.

$$H_1(z) = \mu z^{-(K-1)} H_0(-z^{-1}), \quad \mu = \pm 1.$$

Damit $H_1(z)$ kausal ist, wird $K' = K$ gewählt und für die Impulsantwort folgt:

$$\{h_1(k)\}_{k=0}^{K-1} = \{(-1)^{k_0-k} h_0(K-1-k)\}_{k=0}^{K-1}, \quad k_0 \in \mathbb{Z}.$$

Daraus folgt der Zusammenhang

$$|H_1(e^{j\omega})|^2 = |H_0(e^{j(\omega+\pi)})|^2.$$

Durch Einsetzen in (2.32) erhält man die Aussage:

$$|H_0(e^{j\omega})|^2 + |H_1(e^{j\omega})|^2 = 2. \quad (6.13)$$

Für reellwertige Filter folgt, daß die quadrierten Amplitudengänge beider Filter gerade spiegelsymmetrisch zu $\omega = \pi/2$ sind und hat für die Analysefilterbank im Bild 6.12 zu dem Namen QMF (Quadrature-Mirror-Filterbank) geführt.

6.4 Evolutionäre Programme

Evolutionäre Programme [Mic96, SHK94] interpretieren eine Menge von Parametervektoren $\mathbf{O}^i = \{\mathbf{o}_1^i, \mathbf{o}_2^i, \dots, \mathbf{o}_N^i\}$ als Population der Generation i mit N Individuen $\mathbf{o}_n^i$. Jedes Individuum $\mathbf{o}_n^i$ sei im folgenden ein Vektor mit L Elementen. Dabei stellt jedes Individuum $\mathbf{o}_n^i$ eine mögliche Lösung für ein zu lösendes Optimierungsproblem dar. Über eine Kostenfunktion $K(\mathbf{o})$ wird die Eignung eines Individuums, d.h. die damit bezeichnete Lösung des Problems, bewertet.

In jedem Schritt wird, ausgehend von einer zufällig ausgewählten Anfangspopulation $\mathbf{O}^1$, der Übergang zur nächsten Generation durch zwei heuristisch geprägte genetische Operatoren vorbereitet. Man unterscheidet Mutation und Kreuzung. Nach Durchführung dieser Operationen werden von den alten, veränderten bzw. neu gebildeten Individuen die bezüglich der Kostenfunktion besten N für die nächste Generation ausgewählt. Der Vorgang wird nach einer gegebenen Anzahl von Schritten oder durch ein anderes Abbruchkriterium beendet. Zu Beginn jeder Iteration wird geprüft, ob ein Individuum erzeugt wurde, dessen Eignung die des bisher besten Individuums $\mathbf{o}_{opt}$ übertrifft und gegebenenfalls als neues bestes abgelegt. Nach Ablauf des Programms wird $\mathbf{o}_{opt}$ als Lösung zurückgegeben.

Alle Entscheidungen in einem evolutionären Programm werden durch Realisierungen von Zufallsvariablen beschrieben. So seien X_1 im Intervall $[0, 1]$, X_L im Intervall $[1, L]$, X_R im Intervall $[-R, R]$ und X_N im Intervall $[1, N]$ gleichverteilte Zufallsvariable, deren Realisierungen ganzzahlig sind. In der Praxis werden die Zufallsvariablen durch Pseudozufallsgeneratoren [PTVF92] approximiert. Für die in Abschnitt 3.2.1 formulierte Optimierungsaufgabe mit der Kostenfunktion (3.9) wird zur Initialisierung für jedes Element $\mathbf{o}_n^1$ eine Realisierung der Zufallsvariablen X_R gewählt, und das Optimum $\mathbf{o}_{opt}$ bestimmt.

In der vorliegenden Arbeit werden die Operationen Mutation und Kreuzung bei gegebenen Parametern p_M (Mutationswahrscheinlichkeit) und p_K (Kreuzungswahrscheinlichkeit) wie folgt definiert. Man beachte, daß in der Implementierung die Kreuzung vor der Mutation vorgenommen wird.

- Für jedes Element $\mathbf{o}_n^i$ der Generation i wird eine Realisierung der Zufallsvariablen X_1 ermittelt. Überschreitet die Realisierung den Wert des Parameters p_M , wird das Element einer Mutation unterzogen. Dafür wird ein durch eine Realisierung der Variablen X_L zufällig ausgewähltes Element des Vektors $\mathbf{o}_n^i$ auf den Wert einer neuen Realisierung der Variablen X_R gesetzt.

- Bei der Kreuzung werden die Elemente der Population $\mathbf{O}^i$ so ausgewählt wie bei der Mutation unter Verwendung des Parameters p_K , wobei die Anzahl A der ausgewählten Individuen geradzahlig sein muß. Gegebenenfalls wird die Auswahl so lange wiederholt, bis die Bedingung erfüllt ist. Die ausgewählten Individuen $\{\mathbf{o}'_n\}_{n=1}^A$ werden in einer Menge der Mächtigkeit A zusammengestellt. Mit der Zufallsvariablen X_A , die im Intervall $[1, A]$ gleichverteilte ganzzahlige Realisierungen erzeugt, wird ein Element $\mathbf{o}'_l$ mit $l \in \{1, 2, \dots, A\}$ der Menge ausgewählt. Aus der nunmehr um ein Element kleineren Menge wird wiederum ein Element $\mathbf{o}'_k$ mit $k \in \{1, 2, \dots, A\}$ ausgewählt und mit dem vorherigen gekreuzt. Beide ausgewählten Elemente bilden ein Kreuzungspaar $(\mathbf{o}'_l, \mathbf{o}'_k)$. Diese Auswahl endet, wenn $A = 0$ gilt, d.h. die Liste leer ist und $A/2$ Kreuzungspaare erzeugt worden sind. Bei der Kreuzung wird mit X_L für jedes Kreuzungspaar eine Komponente ausgewählt, bei der der zweite Partner den Komponentenwert des ersten übernimmt.

6.5 Rechenaufwand der eingesetzten Verfahren zur Merkmalsextraktion

Der Rechenaufwand wird für jedes Verfahren in Hinblick auf die erforderlichen reellen Multiplikationen O_{Mul} und reellen Additionen O_{Add} pro Abtastwert der Eingangsfolge, d.h. bezogen auf den Eingangstakt, im Mittel angegeben.

6.5.1 Kurzzeit-Fourier-Transformation

Die diskrete Kurzzeit-Fourier-Transformation wird durch die Struktur in Bild 2.32 implementiert. Die Kanalzahl $N = M$ wird zu einer Potenz von 2 gewählt. Für die eingesetzte schnelle Fourier-Transformation (FFT) der Länge N werden dann (vergl. [Jon91]) $\frac{N}{2} \lg N$ komplexe Multiplikationen und $N \lg N$ komplexe Additionen eingesetzt. Da jede komplexe Multiplikation vier reellen Multiplikationen und ebenso eine komplexe Addition zwei reellen Additionen entspricht, werden

$$O_{Mul}^{FFT}(N) = 2N \lg N \quad (6.14)$$

Multiplikationen und

$$O_{Add}^{FFT}(N) = 2N \lg N \quad (6.15)$$

Additionen benötigt.

Die Länge des Tiefpaßfilters $\{v(k)\}_{k=0}^{K-1}$ sei $K = \tilde{G}M$ mit $\tilde{G} \geq 0$. In diesem Fall haben die Polyphasenfilter $H_{0,n}(z)$ (vergl. Bild 2.32) mit $n = 0, 1, 2, \dots, M$ die Länge $\tilde{G} = K/M$. Für jeden Ausgabewert eines der M Polyphasenfilter fallen genau $\tilde{G}$ Multiplikationen und $\tilde{G} - 1$ Additionen an. Da Polyphasenfilter und FFT-Baustein mit einem gegenüber dem der Eingangsfolge um den Faktor M niedrigeren Takt arbeiten, gilt daher für die Anzahl an Multiplikationen und Additionen bei einer Implementierung der Kurzzeit-Fourier-Transformation:

$$\begin{aligned} O_{Mul}^{KFT}(M, K) &= \frac{M \frac{K}{M} + 2M \lg M}{M} = \frac{K}{M} + 2 \lg M, \\ O_{Add}^{KFT}(M, K) &= \frac{M(\frac{K}{M} - 1) + 2M \lg M}{M} = \frac{K}{M} - 1 + 2 \lg M. \end{aligned}$$

6.5.2 Allgemeine translationsinvariante Filterbänke mit Baumstruktur

Die für eine paraunitäre Filterbank mit M Kanälen mit Filtern der Länge $K = (G + 1)M$ angegebene Faktorisierung (2.25) kann direkt für eine Implementierung verwendet werden. Der Signallauf der für die Faktorisierung benötigten Elemente ist in den Bildern 2.5 bzw. 2.6 dargestellt. Die Anzahl benötigter Multiplikationen $O_{Mul}^{Para}(M, K)$ bzw. Additionen $O_{Add}^{Para}(M, K)$ beträgt [Vai93, S. 329]:

$$\begin{aligned} O_{Mul}^{Para}(M, K) &= \frac{2K}{M} + M, \\ O_{Add}^{Para}(M, K) &= \frac{2K}{M} + \frac{(M - 1)^2}{M}. \end{aligned}$$

Für eine Kaskadierung der Filterbank wie bei der Wavelet-Transformation ergibt sich mit $M = 2$ und der Zerlegungstiefe T der Rechenaufwand [VK95]:

$$\begin{aligned} O_{Mul}^{DWT}(T, K) &= O_{Mul}^{Para}(2, K) \left(1 + \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^T}\right) \\ &= O_{Mul}^{Para}(2, K) \left(2 - \frac{1}{2^{T+1}}\right), \\ O_{Add}^{DWT}(T, K) &= O_{Add}^{Para}(2, K) \left(2 - \frac{1}{2^{T+1}}\right). \end{aligned}$$

Der Rechenaufwand ist also stets kleiner als Doppelte einer einfachen zweikanaligen orthogonalen Filterbank.

Für eine Filterbank mit Baumstruktur gilt für M Kanäle allgemein:

$$\begin{aligned} O_{Mul}^B(T, M, K) &= O_{Mul}^{Para}(M, K)(1 + M \frac{1}{M} + \dots + M^{T-1} \frac{1}{M^{T-1}}) \\ &= O_{Mul}^{Para}(M, K)T, \\ O_{Add}^B(T, M, K) &= O_{Add}^{Para}(M, K)T, \end{aligned}$$

d.h. der Rechenaufwand steigt linear mit der Zerlegungstiefe. Wird die Filterbank translationsinvariant ausgelegt, so hängt der zusätzliche Rechenaufwand von der betrachteten Menge der Translationen $\mathbb{W}$ ab. Eine Filterbank mit Baumstruktur der Tiefe T , die vollständig translationsinvariant ist, z.B. $\mathbb{W} = \{0, 1, 2, \dots, M^T - 1\}$, erfordert

$$\begin{aligned} O_{Mul}^{BT}(T, M, K) &= O_{Mul}^{Para}(M, K)(M + M^2 M \frac{1}{M} + M^3 M^2 \frac{1}{M^2} + \dots \\ &\quad \dots + M^T M^{T-1} \frac{1}{M^{T-1}}) \\ &= O_{Mul}^{Para}(M, K) \sum_{i=1}^T M^i \quad \text{und} \\ O_{Add}^{BT}(T, M, K) &= O_{Add}^{Para}(M, K) \sum_{i=1}^T M^i \end{aligned}$$

Multiplikationen und Additionen.

6.5.3 Quadratische Zeit-Frequenz-Verfahren

Für quadratische Verfahren der Cohen-Klasse ist (2.140) Ausgangspunkt der Implementierung. Der Kern $g(l, k)$ sei für $l \in [-K_1, K_1]$ und $k \in [-K_2, K_2]$ definiert sowie reell und bezüglich k gerade (2.139).

Zur Bestimmung jedes Summanden in (2.140) werden 4 Multiplikationen und 2 Additionen benötigt, da die Folge $\{s_a(m)\}_m$ komplexwertig ist. Da $\tilde{g}(m, k)$ bezüglich k konjugiert gerade ist, ist das Ergebnis der nachfolgenden DFT (2.141) reell. In [BB87] wird daher vorgeschlagen, jeweils die Elemente $\tilde{g}(m, k)$ für $m = 2m'$ und $m = 2m' + 1$ mit $m' \in \mathbb{Z}$ gemeinsam durch eine DFT bestimmen zu lassen.

Es sei die Länge der DFT zu $N = 2K_1 + 2$ zu einer Potenz von 2 gewählt. Dann gilt für die Anzahl der Multiplikationen und Additionen bei einer Dezimation um den Faktor M :

$$\begin{aligned} O_{Mul}^{Cohen}(M, K_2, K_1) &= \frac{6(K_2 + 1) + \frac{1}{2}O_{Mul}^{FFT}(2K_1 + 2)}{M} \quad \text{und} \\ O_{Add}^{Cohen}(M, K_2, K_1) &= \frac{2K_2 + \frac{1}{2}O_{Add}^{FFT}(2K_1 + 2)}{M}. \end{aligned}$$

6.5.4 Modellbasierte Spektralschätzverfahren

Zur Ermittlung des Rechenaufwands des Burg-Algorithmus wird dessen Darstellung in [Böh93, S. 181] verwendet. Wird die Eingangsfolge in Segmente der Länge L mit einem Versatz der Segmente von M zerlegt (2.169), wobei für jedes Segment das AR-Modell der Ordnung P ermittelt wird, folgt für den Rechenaufwand:

$$\begin{aligned} O_{Mul}^{AR}(L, P, M) &= \frac{5PL - 2P^2 + 5P - 6L - 5}{M} \quad \text{und} \\ O_{Add}^{AR}(L, P, M) &= \frac{5PL - 2P^2 + 3P - 6L - 8}{M} \end{aligned}$$

sowie $(P + 1)/M$ Divisionen.

6.6 Optimierte Filter

Für $M = 2$ wurden mit $\varepsilon = 0$ nach (2.30) die folgenden Filter der Länge $K = 10$ und $K = 20$ entworfen. Es ist jeweils der Tiefpaß angegeben:

$$\{h_0(k)\}_{k=0}^9 = \{ -0.05493548704259, -0.04526488720061, 0.00750756275607 \\ -0.27818886685200, -0.72714644271372, -0.57429488051066 \\ 0.03172526057076, 0.23402033331432, 0.03574232069865 \\ -0.04337847539332 \},$$

$$\{h_0(k)\}_{k=0}^{19} = \{ -0.07348256737690, -0.00919701603906, 0.07810331343385 \\ 0.07766118449058, 0.05772788263789, 0.07452373123945 \\ -0.07832216952074, -0.46010219371474, -0.68774105296929 \\ -0.40168231120433, 0.07314470004784, 0.16749345557588 \\ -0.10550230035982, -0.21922995426330, -0.03087150971778 \\ 0.13707689946169, 0.03937736400556, -0.08011003456178 \\ -0.00397638637126, 0.03177063932494 \}.$$

Für $M = 3$ wurde mit $\varepsilon = 0$ nach (2.30) die Filterbank mit Filtern der Länge $K = 30$ entworfen. Es sind $\{h_0(k)\}_{k=0}^{29}$ ein Tiefpaß, $\{h_1(k)\}_{k=0}^{29}$ ein Bandpaß und $\{h_2(k)\}_{k=0}^{29}$ ein Hochpaß:

$$\{h_0(k)\}_{k=0}^{29} = \{ -0.0632946199946, -0.0405228055493, 0.0894908294763 \\ 0.160009442871, 0.0448957564718, -0.200952959283 \\ -0.471665875273, -0.599624204665, -0.498849783211 \\ -0.24088075652, 0.00204860605676, 0.0990956070187 \\ 0.0608639176952, -0.0459280961345, -0.0863320684762 \\ -0.0134614907795, 0.0706014972627, 0.0266316776963 \\ -0.00354666303037, -0.013613202613, -0.0121953639138 \\ 0.00896997459881, -0.00690012892628, 0.00292514985231 \\ 0.00048815063079, -0.00022116994735, 0.00026619029275 \\ -3.971682795e-05, 2.55224322e-05, -1.653377205e-05 \},$$

$$\{h_1(k)\}_{k=0}^{29} = \{ \begin{array}{lll} -0.00458947763789, & -0.00348285359733, & 0.00564835408563 \\ -0.109995055898, & -0.0695227050511, & 0.158932235584 \\ -0.0404052176092, & -0.016491517794, & 0.27574965765 \\ -0.174614874235, & -0.480005138662, & 0.289074569306 \\ 0.523616861563, & -0.228191139033, & -0.372631847185 \\ 0.0765778506755, & 0.155427407828, & 0.12425578871 \\ 0.0506201159896, & -0.126676328999, & -0.0429346626315 \\ 0.0510290553541, & -0.0373403102098, & 0.0179702211075 \\ 0.00195515074651, & -0.00079944117697, & 0.00112495730468 \\ -0.00019616112691, & 0.00012605510865, & -8.166018096e-05 \end{array} \} ,$$

$$\{h_2(k)\}_{k=0}^{29} = \{ \begin{array}{lll} -0.0412353995345, & -0.0276900984348, & 0.0563102361587 \\ -0.0335083903622, & -0.0444315717599, & 0.0714425456272 \\ -0.0377418055223, & -0.0131059856885, & -0.027082231195 \\ 0.206517522779, & -0.432101915816, & 0.57487816394 \\ -0.504927180273, & 0.254768675165, & 0.0500861416382 \\ -0.21487720456, & 0.189761997159, & -0.0332044003053 \\ -0.089316379333, & 0.0974071385643, & -0.00983477345156 \\ -0.0216580311119, & 0.0154643426731, & -0.00788870517189 \\ -0.00066503814325, & 0.00023448578808, & -0.00040813645768 \\ 8.279630308e-05, & -5.320573523e-05, & 3.446738505e-05 \end{array} \} .$$

6.7 Aufgezeichnete Instrumente

Flügel

Nummer	Name	Fabrikat-Nr.	Zusatz
1	Bechstein Modell A (Berlin)	184089	
2	Bechstein Modell B (Berlin) B-88	184077	
3	Bechstein Modell M (Berlin)	183046	
4	Boston GP 184	105101	
5	Bösendorfer	43668-3853	
6	Bösendorfer 170cm	43789-1546	
7	Grotrian Steinweg Modell 192	unbekannt	
8	Sauter	99497	185 cm
9	Schimmel	317604	174 cm
10	Steinway A	476733	
11	Wilhelm Steinberg IQ-77	21.6.96	
12	Yamaha C2	5458794	173 cm
13	Yamaha C3	5470851	186cm
14	Yamaha C5	5460023	
15	Yamaha C6	5406811	
16	Yamaha C7	5380441	
17	Yamaha GH1	B5449489	

Klaviere

Nummer	Name	Fabrikat-Nr.	Zusatz
18	Bechner	923427	
19	C. Bechstein Modell Zimmermann	unbekannt	
20	C. Bechstein Modell 8a	183442	
21	Eterna-Piano	ER30R5243686	121 cm
22	Furstein	unbekannt	
23	Gratiae	5447 PU120E	121 cm
24	Hohner Modell HP-116	1455865	
25	Kemble	266322	116cm, Kirschbaum
26	Rameau Paris: Christian Adam	57457	
27	Sauter Modell Carus	102598	

Klaviere

Nummer	Name	Fabrikat-Nr.	Zusatz
28	Sauter	unbekannt	112 cm
29	Schimmel	318.976	116cm
30	Schimmel	316999	Birn- baum
31	Schimmel	318.048	
32	Wilhelm Steinberg	unbekannt	
33	Steinway & Sons Modell V	500 844	

Keyboards

Nummer	Name	Fabrikat-Nr.	Zusatz
34	Yamaha PSR 400	-	

6.8 Formelzeichen

A	Symbol für Abbildung
$A(m_0)$	Symbol für eine verschiebungsabhängige Abbildung
$A(z)$	Autoregressives Filter
$A^G(q)$	Symbol für eine von der Verschiebung abhängige Abbildung
a	Skalierungsparameter bei der WT
$a(k)$	AR-Modell-Parameter
$\hat{a}_m(k)$	Geschätzte AR-Modell-Parameter mit Zeitindex m
$a_l(k)$	AR-Modell-Parameter bei Modellordnung l
$B_n(\epsilon)$	Stopbandintervall für Filter n
b	Verschiebungsparameter der WT
$\bar{b}_{n,mM}$	Vektorinterpretation einer paraunitären Filterbank
$\bar{b}_{n,mM}^i$	Vektorinterpretation einer Filterbank mit Baumstruktur
$\bar{b}_{n,mM}^{i,p}$	Vektorinterpretation einer translationsinvarianten Filterbank mit Baumstruktur
C	Anzahl der Klassen
C_ψ	Konstante der inversen WT
$C(z)$	z -Transformierte der Folge $\{c(k)\}_{k=0}^{K-1}$
$C(e^{j\omega})$	Fourier-Transformierte der Folge $\{c(k)\}_{k=0}^{K-1}$
$C(m, n)$	Zeit-Frequenz-Darstellung der diskreten Cohen-Klasse
$C(t, \omega)$	Zeit-Frequenz-Darstellung der Cohen-Klasse
$C_{WV}(t, \omega)$	Zeit-Frequenz-Darstellung der WV
$C_{PWV}(t, \omega)$	Zeit-Frequenz-Darstellung der PWV
$C_{SPWV}(t, \omega)$	Zeit-Frequenz-Darstellung der SPWV
$C_{CWV}(t, \omega)$	Zeit-Frequenz-Darstellung der CWV
$C(t, \omega)$	Zeit-Frequenz-Darstellung der Cohen-Klasse
$C^q(x, p)$	Ort-Impuls-Darstellung (Quantentheorie)
$c(k)$	Koeffizientensatz zur Beschreibung der Skalierungsfunktion
$\bar{c}_m^l$	Hilfsvektor
$D(z)$	z -Transformierte der Folge $\{d(k)\}_{k=0}^{K-1}$
$D(e^{j\omega})$	Fourier-Transformierte der Folge $\{d(k)\}_{k=0}^{K-1}$
$d(k)$	Koeffizientensatz zur Beschreibung der Waveletfunktion
d^i, p	Entscheidungsparameter
d^c	Abstand des Nächster-Nachbar-Klassifikators c
d^{c_1, l_1}	Minimaler Abstand
d^{c_2, l_2}	Nächstgrößere Abstand im Vergleich zu d^{c_1, l_1}

d^{c1}	Minimaler Abstand bei klassenweiser Merkmalsselektion
d^{c2}	Nächstgrößerer Abstand im Vergleich zu d^{c1}
$d(\mathbf{P}, \mathbf{P}_n)$	Abstand zweier Muster $\mathbf{P}$ und $\mathbf{P}_n$
Δd_{min}	Minimale Abstandsdifferenz
$E_{m,V}$	Segmentenergie
$\Delta E_{m,V}^c$	Segmentenergie bei klassenweiser Merkmalsselektion
$\mathbf{e}(z)$	Verzögerungsvektor
$F(m, n)$	Merkmal mit Zeitindex m und Frequenzindex n
$F^{AR}(m, n)$	Merkmale bei einem AR-Modell
$F^B(m, n)$	Merkmale bei einer Filterbank mit Baumstruktur
$F_{m_0}^B(m, n)$	Merkmale $F^B(n, m)$ bei verschobener Eingangsfolge
$F_0^B(m, n)$	Merkmale $F^B(n, m)$ bei unverschobener Eingangsfolge
$\tilde{F}_{m_0}^B(m, n)$	Merkmalsdarstellung der Abbildung $A(m_0)$
$F^{CWV}(m, n)$	Merkmale der CVW
$F^{Cohen}(m, n)$	Merkmale der Cohen-Klasse
$F^{DWT}(m, n)$	Merkmale der Wavelet-Transformation
$F^{PWV}(m, n)$	Merkmale der PWV
$F^{SPWV}(m, n)$	Merkmale der SPWV
$F^{WV}(m, n)$	Merkmale der WV
$F_2(\theta)$	Frequenzfenster bei der SPWV
$\Delta F(m, n)$	Merkmalrepräsentation einer Signaldifferenz
$\Delta F^c(m, n)$	Merkmalrepräsentation der Signaldifferenzen einer Klasse
f_A	Abtastrate nach der Signalvorverarbeitung
f_a	Abtastrate des DAT-Recorders
f_{max}	Höchste in einem Signal vorkommende Frequenzkomponente
f_M^G	Freiheitsgrade eines paraunitären Systems
$f_1(\tau)$	Zeitfenster bei der PWV und SPWV
$f_2(\tau)$	Frequenzfenster bei der SPWV
G	Grad eines paraunitären Systems (McMillan-Grad)
G'	Laufzeitparameter eines paraunitären Systems
$G_n(z)$	Filter der Synthesefilterbank
$G_{\lambda,n}^{(2)}(z)$	Polyphasenfilter zweiter Art der Synthesefilterbank
$G(\theta, \tau)$	Kern oder Gewichtsfunktion der Cohen-Klasse
$G_{WV}(\theta, \tau)$	Kern oder Gewichtsfunktion der WV
$G_{PWV}(\theta, \tau)$	Kern oder Gewichtsfunktion der PWV
$G_{SPWV}(\theta, \tau)$	Kern oder Gewichtsfunktion der SPWV
$G_{CWV}^\sigma(\theta, \tau)$	Kern oder Gewichtsfunktion der CWV

$\mathbf{G}^{(2)}(z)$	Polyphasenmatrix zweiter Art der paraunitären Synthesefilterbank
$g_n(k)$	Filter der Synthesefilterbank
$g_{WV}(u, \tau)$	Inverse Fourier-Transformierte von $G_{WV}(\theta, \tau)$ bezüglich θ
$g_{PWV}(u, \tau)$	Inverse Fourier-Transformierte von $G_{PWV}(\theta, \tau)$ bezüglich θ
$g_{SPWV}(u, \tau)$	Inverse Fourier-Transformierte von $G_{SPWV}(\theta, \tau)$ bezüglich θ
$g_{CWV}(u, \tau)$	Inverse Fourier-Transformierte von $G_{CWV}^{\sigma}(\theta, \tau)$ bezüglich θ
$g(l, k)$	Allgemeine diskrete Kernfunktion
$g_{WV}(l, k)$	Diskrete Kernfunktion der WV
$g_{PWV}(l, k)$	Diskrete Kernfunktion der PWV
$g_{SWWV}(l, k)$	Diskrete Kernfunktion der SPWV
$g_{CWV}^{\sigma}(l, k)$	Diskrete Kernfunktion der CWV
$H_n(z)$	Filter der Analysefilterbank
$H_{n,\lambda}^{(1)}(z)$	Polyphasenfilter erster Art der Analysefilterbank
$H_{i,l}^{res}(z)$	Resultierendes Filter
$H_{i,l}^{res}(f)$	Physikalische Interpretation des Filters $H_{i,l}^{res}(z)$
$\mathbf{H}^{(1)}(z)$	Polyphasenmatrix erster Art der paraunitären Analysefilterbank
h	Plancksches Wirkungsquantum $h = (6,62620 \pm 0.00005)10^{-34}$ Js
$\hbar$	$h/2\pi$
$h_n(k)$	Filter der Analysefilterbank
$h_{i,l}^{res}(k)$	Resultierendes Filter
$h_{i,l}^{res}(k)$	Physikalische Interpretation des Filters $h_{i,l}^{res}(k)$
i	Laufparameter
i	Parameter der Zerlegungstiefe
K	Länge der Filterimpulsantwort
K_1	Ausdehnung der Kernfunktionen $g(l, k)$ bezüglich k
K_2	Ausdehnung der Kernfunktionen $g(l, k)$ bezüglich l
K_l^i	Kosten der Koeffizienten $w_l^i(m)$
$K(\mathbf{o}, \epsilon)$	Kostenfunktion der Optimierung
$KFT(m, n)$	Diskrete Kurzzeit-Fourier-Transformation
k	Laufparameter
k	Index der Impulsantwort
L	Länge der Signalabschnitte bei Segmentierung
L	Länge des geschnittenen Signals

L_c	Anzahl Muster der Klasse c
$L_{\bar{c}}$	Anzahl Muster der Klasse $\bar{c}$
L_W	Länge des Fensters beim Schneidevorgang
$L^2(\mathbb{R})$	Raum der quadratintegrablen Funktionen
$L^1(\mathbb{R})$	Raum der integrablen Funktionen
l	Laufparameter
l	Kanalparameter bei Filterbank mit Baumstruktur
$\tilde{l}$	Laufparameter
$l^2(\mathbb{Z})$	Raum der quadratsummierbaren Folgen
M	Dezimationsfaktor
$M(\theta, \tau)$	Verallgemeinerte Ambiguitätsfunktion
$M^q(\theta, \tau)$	Verallgemeinerte Ambiguitätsfunktion (Quantentheorie)
$M(\{w_l^i(m)\}_m)$	Entropiemaß zur Bewertung der Folge $\{w_l^i(m)\}_m$
m	Verschiebungsparameter
m_0	Verschiebungsparameter
$m_{0,i}$	Verschiebung in Zerlegungsstufe i
m'_{0,i_0}	Verbleibende Verschiebung nach der Filterbank
$\tilde{m}_{0,i_0}$	Verschiebung durch die Filterbank
$m_{l_1}^c$	Signalverschiebungen der Klasse c
$m_{l_2}^c$	Signalverschiebungen der Klasse $\bar{c}$
N	Kanalanzahl einer Filterbank
N	Parameter der Frequenzdiskretisierung
N'	Anzahl der Nullstellen des Polynoms $C(e^{j\omega})$ in $\omega = \pi$
N_M	Gesamtzahl der Muster in der Datenbank
$N^{AR,R}$	Indexmenge bei AR-Modellen (Reflexionskoeffizienten)
$N^{AR,S}$	Indexmenge bei AR-Modellen (Leistungsdichtespektren)
N^{KFT}	Indexmenge bei der KFT
N^{DWT}	Indexmenge bei der DWT
N^{Cohen}	Indexmenge bei Verfahren der Cohen-Klasse
$N_{m_0}^B$	Indexmenge bei translationsinvarianten Filterbänken
n	Kanalindex
n	Skalierungsparameter
n'	Fester Skalierungsparameter
$n(i, l)$	Abbildungsvorschrift bei Filterbänken mit Baumstruktur
O_{Mul}	Anzahl Multiplikationen
O_{Add}	Anzahl Additionen
$\mathbf{o}$	Parametervektor der Optimierung
$\mathbf{o}_n^i$	Parametervektor n der Optimierung der Generation i

P	Ordnung eines AR-Prozesses
P_M^{LW}	Signalleistung beim Schneidvorgang
$P_{m_0}(m, n)$	Muster mit Verschiebung m_0
$P_{ss}(z)$	z -Transformierte von $p_{ss}(m)$
$P_{ss}(e^{j\omega})$	Fourier-Transformierte von $p_{ss}(m)$, Leistungsdichtespektrum
$\hat{P}_{ss}^m(e^{j\omega})$	Geschätztes Leistungsdichtespektrum $P_{ss}(e^{j\omega})$ mit Zeitindex m
$P(e^{j\omega})$	Trigonometrisches Polynom
$P_{m_0}^c$	Muster mit Verschiebung m_0 und Klassenindex c
P_{m_0}	Muster mit Verschiebung m_0
$P^{c,l}$	Muster der Datenbank
p	Verschiebungsparameter bei translationsinvarianter Filterbank
$p_{ss}(m)$	Autokorrelationsfolge
$p(\mathbf{P}, c)$	Verbunddichte zwischen Muster $\mathbf{P}$ und Klasse c
$Q(z)$	Polynom in z
q	Verschiebungsparameter
R_l^i	Unterraum des $l^2(\mathbb{Z})$
$R_l^{i,p}$	Unterraum des $l^2(\mathbb{Z})$
$\mathbf{R}_{M,M-j}$	Rotationsmatrix
$r(k)$	Reflexionskoeffizient
r_K	Klassifikationsrate
r_F	Fehlerkennungsrate
r_R	Rückweisungsrate
$r_{ss}(m_0)$	Autokorrelationsfolge
$r^{i,k}$	Elemente der Vertauschungsmatrix
$\hat{r}_m(k)$	Geschätzte Reflexionskoeffizienten mit Zeitindex m
S	Schwellwert
$S(z)$	z -Transformierte des Eingangssignals $s(t)$
$\hat{S}(z)$	Verzögertes Eingangssignal $S(z)$
$\tilde{S}(z)$	z -Transformierte des rekonstruierten Signals
$s(t)$	Eingangssignal
$s(m)$	Eingangsfolge
$\tilde{s}(m)$	Rekonstruiertes Signal
$s'_m(k)$	Segmentierte Eingangsfolge
$s_o(l)$	Ungeschnittene Eingangsfolge
$s^\#(l)$	Geschnittene Eingangsfolge vor der Filterung
$\Delta s(t)$	Differenzsignal
$s(m)$	Interpretation der Eingangsfolge $s(m)$ als Zufallsprozeß

$\bar{s}_m$	Interpretation des Prozesses $s(m)$ als Vektor
$\bar{s}_m^l$	Approximation von $\bar{s}_m$ durch ein AR-Modell der Ordnung l
T	Zerlegungstiefe
$U(z)$	Hilfssignal
U_M	Orthogonale Matrix der Dimension (M, M)
$u(m)$	Weißer Rauschprozeß
$\bar{u}_m^l$	Hilfsvektor
V	Segmentlänge
V_n	Unterraum des $L^2(\mathbb{R})$
$\mathbb{W}$	Verschiebungsmenge
$\mathbb{W}^{i,p}$	Verschiebungsmenge in Zerlegungsstufe i und Verzögerung p
$\mathbb{W}_T^{i,p}$	Hilfsmenge zur Bestimmung der Verschiebungsmenge $\mathbb{W}^{i,p}$
V_i	Paraunitäre Matrix vom Grad 1
$v(t)$	Zeitgespiegelte Fensterfunktion/Filter der KFT
$v(k)$	Filter der zeit- und frequenzdiskreten KFT
W_M	Komplexer Drehfaktor $e^{j2\pi/M}$
W_N	Komplexer Drehfaktor $e^{j2\pi/N}$
W_n	Unterraum des $L^2(\mathbb{R})$
$W(n, m)$	Entwicklungskoeffizienten bezüglich der Waveletfunktionen
$W_l^{i,p}(z)$	Ausgabesignal einer translationsinvarianten Filterbank mit Baumstruktur
$W_l^i(z)$	Ausgabesignal einer Filterbank mit Baumstruktur
$w(t)$	Fensterfunktion der KFT
$w_l^i(m)$	Ausgabefolgen einer Filterbank mit Baumstruktur
$w_l^{i,p}(m)$	Ausgabefolgen einer translationsinvarianten Filterbank mit Baumstruktur
$\tilde{w}_l^{i,p}(m)$	Nichtdezimierte Variante der Folge $w_l^{i,p}(m)$
$y_n(m)$	Ausgangsfolgen der Analysefilterbank
$y_n^l(m)$	Modulierte, nichtdezimierte Ausgangsfolgen der KFT
$\mathcal{A}^q$	Allgemeiner Operator der Quantentheorie
$\mathcal{M}^q$	Operator der charakterischen Funktion (Quantentheorie)
$\mathcal{M}$	Operator der charakterischen Funktion

P^q	Impulsoperator (Quantentheorie)
P_X^q	Impulsoperator (Quantentheorie) in der Ortsdarstellung
$\mathcal{T}_T$	Zeitoperator in der Zeitdarstellung
$\mathcal{V}$	Verschiebungsoperator
$\mathcal{V}^q$	Verschiebungsoperator (Quantentheorie)
$\mathcal{V}_T$	Verschiebungsoperator in der Zeitdarstellung
$\mathcal{W}_T$	Frequenzoperator in der Zeitdarstellung
$\mathcal{X}^q$	Ortsoperator (Quantentheorie)
β_k	Normierungskoeffizienten der zeitdiskreten CWV
$\gamma_n(m)$	Entwicklungskoeffizienten bezüglich der Skalierungsfunktionen
$\gamma'_n(m)$	Filterausgabe bei Implementierung der DWT
ε	Entwurfsparameter der Filteroptimierung
θ	Frequenzversatz
$\theta_{i,k}$	Winkel der Givens-Rotation
λ	Polyphasenparameter
ξ	Wellenfunktion/Zustandsvektor (Quantentheorie)
$\xi(x)$	Zustandsvektor in der Ortsdarstellung
$\xi(p)$	Zustandsvektor in der Impulsdarstellung
$\rho(e^{j\omega})$	Hilfsfunktion
σ	Parameter der CWV
σ_{uu}	Varianz des Rauschprozesses $u(m)$
τ	Zeitversatz
$\Phi(j\omega)$	Fourier-Transformierte der Skalierungsfunktion
$\phi(t)$	Skalierungsfunktion
$\Psi(j\omega)$	Fourier-Transformierte der Wavelet-Funktion
$\psi(t)$	Wavelet-Funktion
$\tilde{\psi}_{a,b}(t)$	Parametrisierte Wavelet-Funktion
$\psi_{n,m}(t)$	Parametrisierte Wavelet-Funktion

6.9 Abkürzungen

CWV	Choi-Williams-Verteilung
DFT	Diskrete Fourier-Transformation
DWT	Dyadische Wavelet-Transformation
DWPT	Dyadische Wavelet-Packet-Transformation
FFT	Fast Fourier Transform
PWV	Pseudo-Wigner-Verteilung
QMF	Quadrature-Mirror-Filterbank
SPWV	Smoothed Pseudo-Wigner-Verteilung
WV	Wigner-Verteilung

7 Literaturverzeichnis

- [AGK91] ANSARI, R., C. GUILLEMOT und J.F. KAISER: *Wavelet Construction Using Lagrange Halfband Filters*. IEEE Transactions on Circuits and Systems, 38:1116–1118, 1991.
- [BB87] BOASHASH, B. und P.J. BLACK: *An Efficient Real-Time Implementation of the Wigner-Ville-Distribution*. IEEE Trans. Acoust. Speech, Sig. Proc., ASSP-35(9):1518–1521, 1987.
- [Böh93] BÖHME, J.F.: *Stochastische Signale*. Stuttgart: B.G. Teubner, 1993.
- [Boa88] BOASHASH, B.: *Note On the Use of the Wigner Transform*. IEEE Trans. Acoust. Speech, Sig. Proc, 36(9):1518–1521, 1988.
- [BPBK92] BROTHERTON, T., T. POLLARD, R. BARTON und R. KRIEGER: *Application of Time-Frequency and Time-Scale Analysis to Underwater Acoustic Transients*. In: *IEEE Internat. Symposium on Time-Frequency & Time-Scale Analysis*, Seiten 513–517, 1992.
- [BPJ92] BROTHERTON, T., T. POLLARD und D. JONES: *Applications of Time-Frequency and Time-Scale Representations to Fault Detection and Classification*. In: *IEEE Internat. Symposium on Time-Frequency & Time-Scale Analysis*, Seiten 95–98, 1992.
- [BS87] BRONSTEIN, I.N. und K.A. SEMENDJAJEW: *Taschenbuch der Mathematik*. Frankfurt: Harri Deutsch, 1987.
- [Bur67] BURG, J.P.: *Maximum Entropy Spectral Analysis*. Proc. of the 37th Meeting of the Society of Exploration Geophysicists, 1967.
- [CGT89] COMBES, J.M., A. GROSSMANN und PH. TCHAMITCHIAN: *Wavelets: Time-Frequency Methods and Phase Space*. Berlin/Heidelberg: Springer, 1989.
- [CH67] COVER, T.M. und P.E. HART: *Nearest Neighbor Pattern Classification*. IEEE Trans. on Information Theory, IT-13:21–27, 1967.
- [Chu92] CHUI, C. K.: *An Introduction to Wavelets*. San Diego: Academic Press, Inc., 1992.
- [CM80a] CLAASEN, T. und W. MECKLENBRÄUKER: *The Wigner Distribution: A Tool for Time-Frequency Signal Analysis: Continuous-Time Signals*. Philips Journal of Research, 35(3):217–250, 1980.

- [CM80b] CLAASEN, T. und W. MECKLENBRÄUKER: *The Wigner Distribution: A Tool for Time-Frequency Signal Analysis: Discrete-Time Signals*. Philips Journal of Research, 35(3):276–300, 1980.
- [CM80c] CLAASEN, T. und W. MECKLENBRÄUKER: *The Wigner Distribution: A Tool for Time-Frequency Signal Analysis: Relations with other Transformations*. Philips Journal of Research, 35(3):372–389, 1980.
- [CM83] CLAASEN, T.A.C.M. und W.F.G. MECKLENBRÄUKER: *The Aliasing Problem in Discrete-Time Wigner Distributions*. IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-31(5):1067–1072, 1983.
- [Coh66] COHEN, L.: *Generalized Phase-Space Distribution Functions*. J. Math. Physics, 7:781–786, 1966.
- [Coh95] COHEN, L.: *Time-Frequency Analysis*. Englewood Cliffs, New Jersey: Prentice Hall, 1995.
- [CR83] CROCHIERE, R. E. und L. R. RABINER: *Multirate Digital Signal Processing*. Englewood Cliffs, New Jersey: Prentice Hall, 1983.
- [CW89] CHOI, H.-I. und W.J. WILLIAMS: *Improved Time-Frequency Representations of Multicomponent Signals Using Exponential Kernels*. IEEE Transactions on Acoustics, Speech, and Signal Processing, 37(6):862–871, 1989.
- [CW92] COIFMAN, R.R. und M.V. WICKERHAUSER: *Entropy-Based Algorithms for Best Basis Selection*. IEEE Trans. on Information Theory, 38(2):713–718, 1992.
- [Dau88] DAUBECHIES, I.: *Orthonormal Bases of Compactly Supported Wavelets*. Comm. Pure Appl. Math., 41:909–996, 1988.
- [Dau92] DAUBECHIES, I.: *Ten Lectures on Wavelets*. SIAM, 1992.
- [Dav74] DAVIS, P.J.: *Interpolation & Approximation*. New York: Dover Publications, 1974.
- [Dea97] DEAN, J.S.: *Artificial Intelligence in Test*. IEEE AES Systems Magazine, Seiten 16–20, July 1997.

- [DH73] DUDA, R. O. und P. E. HART: *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
- [Dic87] DICKREITER, M.: *Musikinstrumente: Moderne Instrumente, historische Instrumente, Klangakustik*. Kassel: Bärenreiter Verlag, 1987.
- [Dic95] DICKREITER, M.: *Mikrofon-Aufnahme-Technik*. Stuttgart/Leipzig: Hirzel, 1995.
- [DJ97a] DELFS, C. und F. JONDRAL: *Classification of Piano Sounds Using Time-Frequency Signal Analysis*. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*, Band 3, Seiten 2093–2096, 1997.
- [DJ97b] DELFS, C. und F. JONDRAL: *New Aspects on Classification Using Translation-Invariant Wavelet Packet Transforms*. In: *IEEE UK Symposium on Applications of Time-Frequency and Time-Scale Methods*, Warwick, Seiten 73–76, 1997.
- [DJ98a] DELFS, C. und F. JONDRAL: *Classification of Piano Sounds Using DFT and Wavelet Packet Based Methods*. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*, Band III, Seiten 1569–1572, 1998.
- [DJ98b] DELFS, C. und F. JONDRAL: *Zeit-Frequenz-Verfahren für die Klassifikation transienter Signale*. Frequenz, Seiten 60–65, März/April 1998.
- [DK94] DRESS, W. B. und S.W. KERCEL: *Wavelet-based Acoustic Recognition of Aircraft*. In: SZU, HAROLD S. (Herausgeber): *Wavelet Applications*, Band 2242, Seiten 778–791. SPIE - The International Society for Optical Engineering, 1994.
- [FH51] FIX, E. und J.L. HODGES: *Discriminatory Analysis: Non-Parametric Discrimination: Consistency Properties*. Technischer Bericht 4, USAF School of Aviation Medicine, Randolph Field, Texas, February 1951.
- [Fic84] FICK, E.: *Einführung in die Grundlagen der Quantentheorie*. Wiesbaden: Aula-Verlag, 1984.
- [Fli91] FLIEGE, N.: *Systemtheorie*. Informationstechnik. Stuttgart: B.G. Teubner, 1991.

- [Fli93] FLIEGE, N.: *Multiraten-Signalverarbeitung*. Informationstechnik. Stuttgart: B.G. Teubner, 1993.
- [Gab46] GABOR, D.: *Theory of communications*. Journal Inst. Elec. Eng., 93:429–457, 1946.
- [GM84] GROSSMANN, A. und J. MORLET: *Decomposition of Hardy Functions into square integrable wavelets of constant shape*. SIAM J. Math., 15:723–736, 1984.
- [Gre96] GREINER, T.: *Orthogonal and Biorthogonal Texture-Matched Wavelet Filterbanks for Hierarchical Texture Analysis*. Signal Processing, 54:1–22, 1996.
- [Har78] HARRIS, F. J.: *On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform*. Proceedings of the IEEE, 66:51–83, 1978.
- [Har96] HARDER, S.: *Beiträge zur Erkennung und Zuordnung transien-ter Schallsignale*. Doktorarbeit, Universität der Bundeswehr, Hamburg, 1996.
- [HCTS97] HAZARIKA, N., J.Z. CHEN, A.C. TSOI und A. SERGEJEV: *Classification of EEG Signals Using the Wavelet Transform*. Signal Processing, 59:61–72, 1997.
- [Hei95] HEITZ, C.: *Optimum Time-Frequency Representations for the Classification and Detection of Signals*. Applied Signal Processing, 3:124–143, 1995.
- [Hen56] HENZE, E.: *Abriß einer verallgemeinerten Filtertheorie*. Arch. elektr. Übertr., Band 10:541–551, 1956.
- [Heu86] HEUSER, H.: *Funktionalanalysis*. Stuttgart: B.G. Teubner, 1986.
- [Hla84] HLAWATSCH, F.: *Interference Terms in the Wigner Distribution*. Proc. Int. Conf. on Dig. Sig. Proc., Seiten 363–367, 1984.
- [HMUJ95] HLAWATSCH, F., T.G. MANICKHAM, R.L. URBANKE und W. JONES: *Smoothed Pseudo-Wigner distribution, Choi-Williams Distribution, and Cone-Kernel Representation: Ambiguity-Domain Analysis and Experimental Comparison*. Signal Processing, 43:149–168, 1995.

- [Hud74] HUDSON, R. L.: *When is the Wigner Quasi-Probability Density Non-Negative?* Reports on Mathematical Physics, 6(2):249–252, 1974.
- [HV91] HERLEY, C. und M. VETTERLI: *Linear Phase Wavelets: Theory and Design*. Proceedings IEEE ICASSP 1991, Seiten 2017–2020, 1991.
- [Jon91] JONDRAL, F.: *Funksignalanalyse*. Stuttgart: B.G. Teubner, 1991.
- [Kay88] KAY, STEVEN M.: *Modern Spectral Estimation: Theory and Application*. Signal Processing Series. New York: Prentice Hall, Englewood Cliffs, 1988.
- [KK89] KAMMEYER, K. D. und K. KROSCHER: *Digitale Signalverarbeitung*. Stuttgart: B.G. Teubner, 1989.
- [Kro96] KROSCHER, K.: *Statistische Nachrichtentheorie*. Berlin: Springer, 1996.
- [Lai95] LAI, M.-J.: *On the Digital Filter Associated with Daubechies' Wavelets*. IEEE Trans. on Signal Processing, 43:2203–2205, 1995.
- [LAJ94] LOE, R. S., K. ANDERSON und K. JUNG: *Comparative Analysis Results for Underwater Transient Analysis*. In: SZU, HAROLD S. (Herausgeber): *Wavelet Applications*, Band 2242, Seiten 815–823. SPIE - The International Society for Optical Engineering, 1994.
- [LGR97] LOUGHLIN, P., D. GROUTAGE und R. ROHRBAUGH: *Time-Frequency Analysis of Acoustic Transients*. In: *ICASSP 97-International Conference on Acoustics, Speech and Signal Processing*, Band 3, Seiten 2125–2128, 1997.
- [Lin97] LINDSEY, A.R.: *Wavelet Packet Modulation for Orthogonally Multiplexed Communication*. IEEE Trans. on Signal Processing, 45:1336–1339, 1997.
- [LKW92] LEARNED, R.E., W.C. KARL und A.S. WILLSKY: *Wavelet Packet Based Transient Signal Classification*. In: *IEEE Internat. Symposium on Time-Frequency & Time-Scale Analysis*, Seiten 109–113, 1992.

- [LMR94] LOUIS, A.K., P. MAASS und A. RIEDER: *Wavelets*. Stuttgart: B.G. Teubner, 1994.
- [LP96] LIANG, J. und T.W. PARKS: *A Translation-Invariant Representation Algorithm with Applications*. IEEE Trans. Signal Processing, 44(2):225–232, 1996.
- [LT96] LIVINGSTON, J.N. und C.-C. TUNG: *Bandwidth Efficient PAM Signaling Using Wavelets*. IEEE Trans. on Communications, 44:1629–1631, 1996.
- [LZT95] LI, C., C. ZHENG und C. TAI: *Detection of ECG Characteristic Points Using Wavelet Transforms*. IEEE Trans. Biomedical Engineering, 42:21–28, 1995.
- [Mal89a] MALLAT, S.: *Multiresolution Approximation and Wavelets*. Trans. Amer. Math. Soc., 315:69–88, 1989.
- [Mal89b] MALLAT, S.: *A Theory for Multiresolution Signal Decomposition*. IEEE Trans. Pattern Analysis and Machine Intelligence, 2(7):674–693, 1989.
- [Mar87] MARPLE, S. L.: *Digital Spectral Analysis*. Englewood Cliffs: Prentice Hall, 1987.
- [MC97] MCCLURE, M.R. und L. CARIN: *Matching Pursuits with a Wave-Based Dictionary*. IEEE Trans. on Signal Processing, 45:2912–2927, 1997.
- [Mer96] MERTINS, A.: *Signaltheorie*. Informationstechnik. Stuttgart: B.G. Teubner, 1996.
- [Mey81] MEYER, J.: *Akustik und musikalische Aufführungspraxis*. Frankfurt am Main: Verlag Das Musikinstrument, 1981.
- [Mic96] MICHALWICZ, Z.: *Genetic Algorithms + Data Structures = Evolution Programs*. New York: Springer Verlag, 1996.
- [MK93] MITRA, S.K. und J.F. KAISER: *Handbook for digital signal processing*. New York: John Wiley & Sons, 1993.
- [Moy49] MOYAL, J.E.: *Quantum Mechanics as a Statistical Theory*. Proc. Camb. Phil. Soc., 45:99–124, 1949.

- [MWJ94] MARCO, S. DEL, J. WEISS und K. JAGLER: *Wavelet-based Transient Signal Detector Using a Translation Invariant Wavelet Transform*. In: SZU, H. S. (Herausgeber): *Wavelet Applications*, Band 2242, Seiten 792–797. SPIE - The International Society for Optical Engineering, 1994.
- [MZ93] MALLAT, S. und Z. ZHANG: *Matching Pursuit with Time-Frequency Dictionary*. IEEE Trans. Signal Processing, 41(12):3397–3415, 1993.
- [Nie74] NIEMANN, H.: *Methoden der Mustererkennung*. Frankfurt am Main/Akademische Verlagsgesellschaft, 1974.
- [NM65] NELDER, J.A. und R. MEAD: *A Simplex Method for Function Minimization*. Computer Journal, 7:308–313, 1965.
- [Orf85] ORFANIDIS, S. J.: *Optimum Signal Processing: An Introduction*. New York: MacMillan Publishing Company, 1985.
- [PKLH96] PESQUET, J.-C., H. KRIM, D. LEPORINI und E. HAMMAN: *Bayesian Approach to Best Basis Selection*. In: *Proceedings ICASSP 96*, Band 5, Seiten 2634–2637, 1996.
- [PTVF92] PRESS, W.H., S.A. TEUKOLSKY, W.T. VETTERLING und B.P. FLANNERY: *Numerical Recipes in C*. Cambridge: Cambridge University Press, 1992.
- [PW92] PRIEBE, R. und G. WILSON: *Application of "Matched" Wavelets to Identification of Metallic Transients*. In: *IEEE Internat. Symposium on Time-Frequency & Time-Scale Analysis*, Seiten 349–353, 1992.
- [RCM94] ROBERTSON, D., O.I. CAMPS und J. MAYER: *Wavelets and Power Transients: Feature Detection and Classification*. In: SZU, H. S. (Herausgeber): *Wavelet Applications*, Band 2242, Seiten 474 – 487. SPIE - The International Society for Optical Engineering, 1994.
- [Rod97] RODET, X.: *Musical Sound Signal Analysis/Synthesis: Sinusoidal+Residual and Elementary Waveform Models*. In: *IEEE UK Symposium on Applications of Time-Frequency and Time-Scale Methods, Warwick*, Seiten 111–120, 1997.

- [Sai94] SAITO, N.: *Local Feature Extraction and Its Applications Using a Library of Bases*. Doktorarbeit, Yale University, Dezember 1994.
- [SC94] SAITO, N. und R.R. COIFMAN: *Local Discriminant Bases*. Proceedings SPIE, 2303, 1994.
- [Sch74] SCHALLER, W.: *Verwendung der schnellen Fouriertransformation in digitalen Filtern*. ntz, 11:425–431, 1974.
- [Sch77] SCHÜRMANN, J.: *Polynomklassifikatoren für die Zeichenerkennung*. München/Wien: R. Oldenbourg Verlag, 1977.
- [Sch94] SCHÜSSLER, H. W.: *Digitale Signalverarbeitung 1*. Berlin: Springer, 1994.
- [Sch96] SCHÜRMANN, J.: *Pattern Classification*. New York: Wiley, 1996.
- [SHK94] STENDER, J., E. HILLEBRAND und J. KINGDON: *Genetic Algorithms in Optimisation, Simulation and Modelling*. Amsterdam: IOS Press, 1994.
- [ST93] SINHA, D. und A.H. TEWFIK: *Low Bit Rate Transparent Audio Compression Using Adapted Wavelets*. IEEE Trans. on Signal Processing, 41:3463–3478, 1993.
- [Ste92] STEFFEN, P.: *Closed Form Derivation of Orthogonal Wavelet Bases for Arbitrary Integer Scales*. Technischer Bericht Lehrstuhl für Nachrichtentechnik, Universität Erlangen-Nürnberg, 1992.
- [Tay92] TAYLOR, C.: *Der Ton macht die Musik*. Braunschweig/Wiesbaden: Vieweg, 1992.
- [Vai85] VAIDYANATHAN, P.P.: *On Power-Complementary FIR Filters*. IEEE Trans. on Circuits and Systems, CAS-32:1308–1310, 1985.
- [Vai87] VAIDYANATHAN, P.P.: *Theory and Design of M-Channel Maximally Decimated Quadrature Mirror Filters with Arbitrary M, Having the Perfect-Reconstruction Property*. IEEE Trans. on Acoustics, Speech and Signal Processing, ASSP-35:476–492, 1987.
- [Vai93] VAIDYANATHAN, P.P.: *Multirate Systems and Filter Banks*. Englewood Cliffs: Prentice Hall, 1993.

- [Vas96] VASEGHI, S.V.: *Advanced Signal Processing and Digital Noise Reduction*. Stuttgart: Teubner und Chichester: John Wiley, 1996.
- [Vil48] VILLE, J.: *Theorie et Applications de la Notion de Signal Analytique*. Cable et Transmission, 2A:61–74, 1948.
- [VK95] VETTERLI, M. und J. KOVAČEVIĆ: *Wavelets and Subband Coding*. Englewood Cliffs: Prentice Hall, 1995.
- [Wig32] WIGNER, E. P.: *On the Quantum Correction for Thermodynamic Equilibrium*. Physical Review, 40:749–759, 1932.
- [WWDJ97] WONG, K.M., J. WU, T.N. DAVIDSON und Q. JIN: *Wavelet Packet Division Multiplexing and Wavelet Packet Design Under Timing Error Effects*. IEEE Trans. on Signal Processing, 45:2877–2890, 1997.
- [XRHO97] XIONG, Z., K. RAMCHANDRAN, C. HERLEY und M.T. ORCHAD: *Flexible Tree-Structured Signal Expansions Using Time-Varying Wavelet Packets*. IEEE Trans. on Signal Processing, 45:333–345, 1997.
- [ZT93] ZOU, H. und A.H. TEWFIK: *Parametrization of Compactly Supported Orthonormal Wavelets*. IEEE Trans. on Signal Processing, 41(3):1428–1431, 1993.

7.1 Diplom- und Studienarbeiten

BÄUERLE, Detlef: Aufbau einer Filterbank für die Wavelet-Transformation, November 1994

KOCH, Rene: Signalanalyse mit einer FFT-Filterbank und einer Multiratenfilterbank, August 1995

SCHNIERINGER, Martin: Nichtorthogonale Wavelets in der Signalanalyse, August 1995

DECKER, Andreas: Klassifikationsverfahren in der Signalanalyse, Dezember 1995

KUNZ, Dietmar: Signalanalyse mit der Wigner-Ville-Verteilung, April 1996

BAYER, Gerd: Visualisierung von Multiraten-Systemen, Juni 1996

GEIßLER, Christian: Konzeption einer Ergebnisvisualisierung mit dem Grafik-Tool Tecplot unter besonderer Berücksichtigung von Multiraten-Systemen, August 1996

SCHWEIZER, Martin: Vergleichende Untersuchung der Eignung moderner Spektralschätzverfahren für die Signalklassifikation, August 1996

ROTHSCHILD, Steffen: Signalanalyse und Klassifikation mit der Choi-Williams-Verteilung, Oktober 1996

JAKI, Stephan: Theorie und Anwendung perfekt rekonstruierender Filterbanken in der Klassifikation, Juli 1997

8 Lebenslauf

Persönliche Daten

Name	Christoph Delfs
Geburtsdatum/-ort	5.2.1967 in Hamburg

Schulausbildung

1973–1977	Grundschule in Kaltenkirchen
1977–1986	Gymnasium Kaltenkirchen (Abitur)

Grundwehrdienst

1986–1987

Studium und Berufsweg

1987–1993	Studium der Elektrotechnik an der Universität Karlsruhe (Diplom) Studienmodell: Technik der Informationsverarbeitung
1993–1998	Wissenschaftlicher Angestellter am Institut für Nachrichtentechnik der Universität Karlsruhe

Druck und Bindung: DRUCKEREI ERNST GRÄSSER
76131 Karlsruhe · Humboldtstraße 1 · Tel. 07 21/61 50 50

