DIE ERKENNUNG VON HÜLLKURVENVERÄNDERUNGEN ALS SYSTEMKOMPONENTE ZUR DATENREDUKTION VON AUDIOSIGNALEN

Friedrich E. Blutner

Mitteilung aus dem Institut für Musikinstrumentenbau, Zwota

1. Allgemeines

Der Apparat der akustischen Perzeption kann als hochkomplexes System ineinandergreifender Mechanismen und spezialisierter Teilfunktionen charakterisiert werden. Dabei bilden die im Verlaufe der biotischen Evolution zu hoher Vollkommenheit entwickelten holoarchisch-hierarchischen Organisationsstrukturen die Grundlage für die technisch noch nicht erreichten Leistungen der Signalselektion, Diskrimination, Mustererkennung und Textverarbeitung.

Auf Grundlage evolutionsbiologischer, neurobiologischer und psychoakustischer Systembetrachtungen kann die Vorstellung entwickelt werden, daß die Schallinformation auf dem Weg zum Gehirn eine Kette selektiver Filter und Entscheidungsinstanzen durchläuft. Dabei werden biologisch bedeutsame akustische Signaleigenschaften in spezifischer Weise hervorgehoben und Störungen wirkungsvoll unterdrückt. Im kortikalen Bereich resultiert somit eine selektive, abstrahierende Abbildung relevanter akustischer Signaleigenschaften: der komplexe akustische Eigenschaftssatz /1/.

Ausgehend von dieser allgemeinen Vorstellung besteht die Aufgabe einer technischen Audiosignalanalyse darin, die akustischen Invarianten des von einem Objekt ausgesandten Schallsignals in einer möglichst kompakten Form darzustellen. Diese Darstellung hat insofern abstrahierenden und idealisierenden Charakter, daß Eigenschaften der Übertragungsstrecke einschließlich der durch Fremdquellen induzierten Störungen zugunsten der informationstragenden Merkmale der Quelle reduziert sind. Zur Lösung dieser Aufgabe ist es hilfreich, grundsätzliche Verarbeitungsprinzipien der Audiosignalverarbeitung des Menschen zu berücksichtigen, ohne jedoch eine detaillierte Nachbildung der sensorischen Vorgänge anzustreben.

Die Suche nach den für uns Menschen informationstragenden Merkmalen der akustischen Welt ist erkenntnistheoretisch und ökonomisch gleichermaßen bedeutsam. Gerade aus der Funktion der akustischen Kommunikation für das Zusammenleben der Menschen ergeben sich wichtige Anwendungsprojekte, z.B. für Aufgaben der akustischen Signalübertragung und -speicherung, der Mensch-Maschine-Kommunikation, der Rehabilitation Behinderter oder der akustischen Diagnostik.

Derart motiviert wurde in den letzten vier Jahrzehnten ein erheblicher Forschungsaufwand zur Lösung des Merkmalproblems getrieben. Im Rahmen dieser Bemühungen wurden wiederholt Analogiebetrachtungen zur Verarbeitung optischer Muster angestellt. In Verfolgung des Ziels einer Reduktion der außerordentlich vielgestaltigen Erscheinungsformen sprachlicher Signale auf einfache psychophysikalische Gesetzmäßigkeiten gelangte beispielsweise TSCHESCHNER bereits in den 60er Jahren zur Auffassung, "daß die Wahrnehmbarkeit von Lauten auf akustischen Erscheinungen beruht, deren optisches Analogon ähnlicher der Berandungsfunktion einer Figur als ihrer Farbe ist." (zit. aus /2/, S. I).

2. Die Rolle des Zeitfaktors bei der Audiosignalverarbeitung

Betrachtet man Raum-Zeit-Metriken unterschiedlichen Organisationsniveaus als Basis der physio-psychologischen Verarbeitung von Sprache und Musik, so sind derartige Analogiebetrachtungen zunächst mit großen Problemen behaftet, weil näm-

lich der Zeitfaktor bei akustischen Signalen grundsätzlich anders als bei optischen Signalen gehandhabt wird. ROEDERER hat am Beispiel der Musik die besondere Rolle des Zeitfaktors tiefgründig analysiert: "Es gibt drei verschiedene Bereiche, in denen sich zeitliche Veränderungen von psychoakustischer Bedeutung einteilen lassen. Erstens die 'mikroskopische Zeitskala', in der die eigentlichen Schwingungen einer Schallwelle auftreten, mit einem Bereich von ca. 0.00007 bis 0.05 sec. Zweitens ein Zwischenbereich, ungefähr bei einer Zehntelsekunde, welcher Schwankungen, wie das Ein- und Abklingen eines Tons, also Variationen der 'mikroskopischen' Vorgänge umfaßt. Drittens der 'makroskopische' Bereich, der sich von ca. 0.1 sec ab aufwärts erstreckt, entsprechend der Dauer gewöhnlicher musikalischer Töne, ihrer Aufeinanderfolge und dem Rhythmus. Bemerkenswert ist, daß jede dieser Zeitstufen ihr eigenes 'Verarbeitungszentrum' im Hörsystem besitzt." (zit. aus /3/, S.6)

Die m i k r o s k o p i s c h e n Schwingungen werden vom Innenohr aufgenommen und kodiert. TERHARDT hat für diese Vorgänge in bemerkenswerter Prägnanz die Analogien zum Sehen aufgezeigt: "Sowohl im Auge als auch im Ohr findet eine Abbildung des Stimulus auf jeweils ein Areal von Sensoren statt. Beim Auge sind dies die Zapfen und Stäbchen der Netzhaut, beim Ohr die Haarzellen des Cortischen Organs im Innenohr. Die bandfilterartigen Eigenschaften der Hydromechanik des Innenohrs bewirken, daß die 'Tief-hoch-Dimension' eines Schalls als örtliche Verteilung der Anregung der Haarzellen abgebildet wird. Dem entspricht die örtliche Helligkeits- beziehungsweise Farbverteilung auf der Retina des Auges. Daß das Ohr in dieser Entsprechung eine Ortsdimension weniger hat als das Auge, ändert an der Analogie nichts, darf vielmehr dann nicht übersehen werden, wenn man sie verstehen will. So verläuft beispielsweise die Kontur einer visuellen Gestalt in der zweidimensionalen Retina und ist selbst e i n dimensional. Daher ist zu erwarten, daß ihr auditives Analogon ein 'Punkt' auf der eindimensionalen Tief-hoch-Skala ist. Es gibt tatsächlich solche 'Punkte': die Spektraltonhöhen der diskreten Teiltöne." (zit. aus /4/)

In konsequenter Verfolgung dieser Analogie konnte deutlich gemacht werden, daß Teiltonhören wahrscheinlich fundamentale Bedeutung für die auditive Informationsverarbeitung im allgemeinen hat. Formal kann zunächst jedes Schallsignal, das von einem Objekt ausgeht, als Summe einer endlichen Anzahl von Teiltönen (Sinusschwingungen) beschrieben werden, deren Frequenzen und Amplituden relativ langsam zeitvariant sind. Vor allem die Teiltonfrequenzen überstehen die Signalübertragung vom Objekt zum Gehör relativ unverändert, sind also weitgehend invariant gegenüber den Einflüssen des akustischen Raums. Die Teiltonfrequenzen spielen demnach in der akustischen Welt eine ähnliche Rolle wie die Gestaltumrisse und Kanten in der optischen. Die (subjektiven) Spektraltonhöhen entsprechen den visuellen Konturen der optischen Gestaltwahrnehmung; kurz: die Teiltöne bilden Primärkonturen des Gehörs /4/ /5/ /6/.

Die Ein- und Ausschwingungsvorgänge sowie die Übergangsvorgänge im "Z w i - s c h e n b e r e i c h der Zeitskala" werden vorzugsweise auf den einzelnen Etappen der neuronalen Hörbahn vom Ohr zum primären Hörzentrum des Gehirns verarbeitet und kodiert. Welche Bedeutung gerade die Ein- und Ausschwingvorgänge für die Signalerkennung besitzen, hat bereits HELMHOLTZ klar erkannt: "Die Arten des Anklingens und Ausklingens sind ja zum Teil so charakteristisch, daß sie für die menschliche Stimme durch eine Reihe verschiedener Buchstaben bezeichnet werden. Es gehören hierher namentlich die explosiven Konsonanten B, D,G und P,T,K." (zit. aus /7/, S. 114). In seiner Systematik der musikalischen Klangfarbe beschränkt sich Helmholtz dann aber mit viel Bedacht nur auf die "Eigentümlichkeiten des gleichmäßig andauernden Klanges".

Die Phase der Erforschung transitorischer Vorgänge im Schallspektrum wurde durch die Arbeiten von Kupfmüller (1924) auf elektroakustischem Gebiet /8/ und die 1932 veröffentlichten Untersuchungen von Backhaus "Über die Bedeutung der Ausgleichsvorgänge in der Akustik" /9/ eingeleitet. Damals bahnte sich die Erkenntnis an, "daß die nicht harmonischen Vorgänge, die in fast jedem andauernden Klange vorkommen und die beim An- und Absetzen des Klanges auftreten, für die Charakterisierung des Klanges von besonderer Bedeutung sind." (zit. aus /8/, S. 32). Diese Feststellung gilt für natürlich erzeugte Sprach- oder Musiksignale in potenzierter Form.

Die Verarbeitung solcher Signale in den einzelnen Hörbahnabschnitten muß mit einer z e i t l i c h e n D i s k r e t i s i e r u n g der zeitvarianten Anregungsverteilung im Innenohr verbunden sein, denn die kortikalen Mechanismen können nur die "m a k r o s k o p i s c h e n" Zeitvorgänge oberhalb von etwa 100 ms verarbeiten. Das bedeutet also, daß E n t s c h e i d u n g e n darüber getroffen werden müssen, ob beispielsweise ein Signalanfang oder ein Signalende vorliegt, welche typischen Formen des Anklingens oder Ausklingens eines Schallereignisses vorhanden sind, ob und in welcher Form ein markanter Signalübergang auftritt oder ob etwa eine charakteristische Signalpause präsent ist. Die Neurophysiologie der zentralen Hörbahn hat eine Reihe von Anhaltspunkten dafür erbracht, daß es einige spezialisierte Neuronengruppen für solche Aufgaben gibt, z.B. auf Intensitätsänderungen phasisch ansprechenden On- bzw. Off-Detektoren, auf Frequenz-Zeitänderungen reagierende Transientdetektoren.

In diesem Beitrag sollen aus technischer Sicht einige Illustrationen für die Auffassung gegeben werden, daß bei den zeitlichen Diskretisierungsvorgängen die Erkennung von Hüllkurven v e r ä n d e r u n g e n eine wichtige Rolle spielt. Dieser Vorgang stellt, wie in Bild 1 illustriert, das Pendant zu dem von TERHARDT

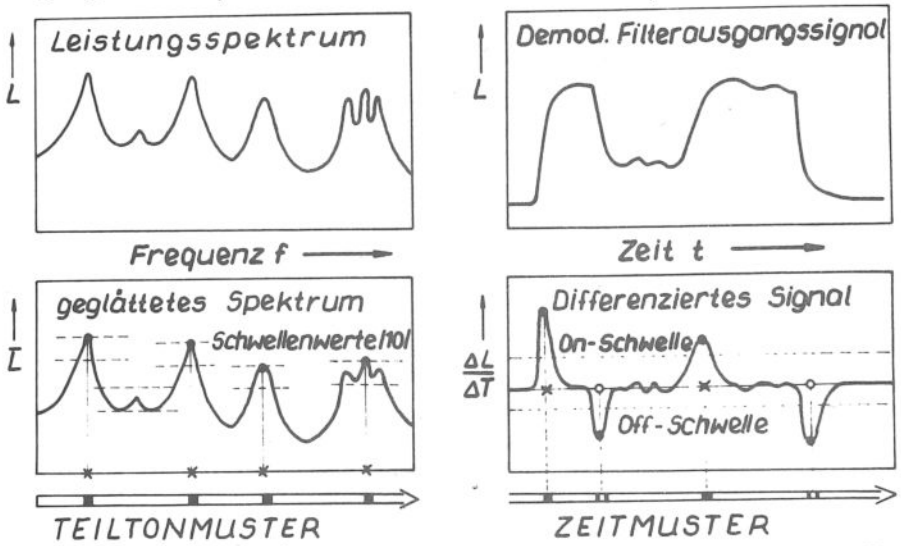


Bild 1: Illustration elementarer Entscheidungsvorgänge zur Datenreduktion von Audiosignalen
links: Erkennung von Teiltonhöhen /4/ /10/
rechts: Erkennung von Hüllkurvenveränderungen

beschriebenen Entscheidungsmechanismus zur Rekrutierung der Teiltonmuster dar. Die Erkennung von Hüllkurvenveränderungen führt dann zu einer Konturisierung in der Zeitebene. Während also zum Beispiel ein stationärer Klang durch eine Reihe singulärer "Punkte" auf der eindimensionalen Tief-hoch-Skala repräsentiert wird /4/, ist eine reguläre Folge kurzer Signalimpulse (etwa ein rhythmisches Muster) durch eine Folge von Zeit p u n k t e n auf der eindimensionalen Zeitskala beschrieben.

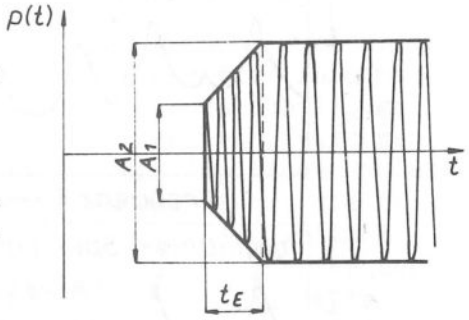
Die übersichtliche Darstellung der Frequenz- und Zeitinformation ist in der allgemein üblichen Weise (z.B. Sonagramm) als Frequenz-Zeitmuster möglich. Diese formale Darstellung hat aber bezüglich der Repräsentation der Zeit für die untere Ebene ("mikroskopische" Zeitskala) und die mittlere Ebene ("Zwischenbereich" um 0.1 sec) einen grundsätzlich verschiedenen Inhalt. Im ersten Falle stellt die Zeit einen "flüchtigen" Parameter dar, im zweiten Falle aber ist sie im Bereich von einigen hundert Millisekunden als räumlicher Parameter fixiert. Für den letzten Fall resultiert also eine z w e i dimensionale x-y-Matrize, die formal völlig analog zu der zweidimensionalen kortikalen Projektion der Retina ist.

Es sei ausdrücklich betont, daß diese Analogie nur für die mittlere Zeitebene gilt, etwa entsprechend dem durch die Informationspsychologie beschriebenen "Zeitmoment" der bewußten Nachrichtenverarbeitung oder dem allgemein als "Flimmerfrequenz" bekannten Zeitauflösungsvermögen des Auges. Der akustische Eigenschaftssatz kann auf dieser kortikalen Ebene als eine "konturisierte Schallrepräsentation" charakterisiert werden, etwa vergleichbar mit einer Strichzeichnung der auf das Wesentliche reduzierten Details einer optischen Szene. Konturen in Richtung der Zeitrepräsentation würden hierbei entsprechend den "Maxigrammen" /10/ die Teilton-Zeitverläufe darstellen, dazu weitgehend orthogonal verlaufende Konturen (also in Richtung der Frequenzrepräsentation) würden prägnante Signaleinsätze, Pausen und andere markante Zeitereignisse abstrahierend abbilden.

3. Beispiele für die Erkennung schneller Hüllkurvenveränderungen

3.1. Einzelereignisse

Untersuchungen zur objektiven Messung der Einschwingvorgänge musikalischer Einzeltöne zeigten am Beispiel von Pianoklängen folgendes Ergebnis /11/: Unterschiede zwischen hart und weich intonierten Pianos können bereits aus der Analyse der Schalldruck-Zeitfunktion erschlossen werden. Hierbei wird die Zeitenveloppe in erster Näherung durch Geradenstücke approximiert (Bild 2). Als Parameter resultieren für den gezeigten einfachen Fall die Einschwingzeit t_E und die prozentuale Sprunghöhe $K=A_1/A_2 x100\%$. Für die Unterscheidung der Härte der Intonation erweist sich der Parameter K als sehr gut geeignet, während die Auswertung der Einschwingzeit t_E keine signifikanten Resultate liefert. In Übereinstimmung mit durchgeführten Perzeptionstests ergab sich, daß offenbar der abrupte Pegelsprung während des Signaleinsatzes das Erkennungskriterium für die Klanghärte darstellt. Das heißt aber, daß Hüllkurvenverläufe hinsichtlich des Zeitpunktes und der Prägnanz von Signalveränderungen ausgewertet werden.



Parameter: Einschwingzeit t_E

Prozentuale Sprunghöhe (Prägnanz) $K = A_1/A_2 \cdot 100\%$

Bild 2: Zeithüllenparameter eines Einzeltons

Daß solche Merkmale ganz analog zu den Teiltonfrequenzen als informationstragend angesehen werden müssen, zeigte sich auch darin, daß sie die Signalübertragung vom Objekt zum Gehör recht unbeschadet überstehen. Die abrupten Hüllkurvensprünge sind im Direktschall stets deutlich auszumachen, insbesondere, wenn man eine grobe Frequenzselektion (Oktavfilterung) vornimmt. Mit überraschender Präzision kann über die Hüllkurvensprünge sogar die Schallquelle geortet werden, wenn man unter Verwendung eines Kunstkopfes die entsprechenden Laufzeitdifferenzen auswertet.

Im Ergebnis solcher Beobachtungen, gezielter psychoakustischer Untersuchungen und technischer Simulationsexperimente wurde ein Funktionsmodell zur Hüllkur-

venverarbeitung entwickelt /12/. Die technische Umsetzung des Modells wurde im Interesse einer möglichst hohen Zeitauflösung wie folgt vorgenommen /13/:

1. Grobe Frequenzselektion (Oktav- bzw. Halboktavfilter)
2. Einweggleichrichtung der erhaltenen Signalfunktionen und Tiefpaßbegrenzung entsprechend der am ersten Neuron gemessenen Grenzfrequenz von 3..4 kHz.
3. Intensitätslogarithmierung
4. Bildung des zeitlichen Gradienten der resultierenden Signalfunktionen für positive und negative Hüllkurven-Zeitveränderungen in separaten Teildetektoren, wobei unterhalb von 1 kHz die Auswertung periodensynchron vorgenommen wird und oberhalb von 1kHz ein festes Zeitfenster von 1 ms benutzt wird.
5. Erkennung der maximalen Hüllkurvenveränderungen separat in den einzelnen Frequenzkanälen mit Hilfe von Maximalwert-Detektoren und Schwellwertvergleichen zur Störunterdrückung.

Das realisierte "Enveloppemeter" bildete die Grundlage für umfangreiche statistische Untersuchungen an Musik- und Sprachsignalen /12/ /14/ /15/ sowie zur akustischen Diagnostik /16/. Als Beispiel sei die Klopftondiagnose von Gitarren-Klangdecken erörtert. Ähnlich wie die versierten Musikinstrumentenbauer schon seit Jahrhunderten aus dem Klang des Klopftons Prognosen über die Resonanzqualität der eingesetzten Hölzer machen konnten, liefert ein für die kommerzielle Anwendung realisiertes Computerdiagnosesystem akustische Gebrauchswertgutachten über die klangbestimmenden Bauteile von Serieninstrumenten. So werden beispielsweise Gitarren - R o h d e c k e n mit Hilfe eines Anschlaghammers angeregt. Der entstehende Klopfton wird mittels breitbandig gehaltener Filter analysiert (8 Frequenzkanäle, Zeitdiskretisierung: 0.5 ms). Die Signal-Onsets bilden sich hierbei sehr markant ab und können in jedem Falle exakt einer 0.5 ms - Zeitscheibe zugeordnet werden. Darauf basierend ist die Erkennung eines Klopfton-Ereignisses auch im akustischen Störfeld eines Industriebetriebes mit Fehlerraten unter 1°/°° möglich.

Im Ergebnis der zeitlichen Konturisierung der Klopftonereignisse resultieren drei Merkmalsätze: die Prägnanzfunktion (Pegeländerungen im Bereich von $\pm$ 0.5 ms, bezogen auf die Onset-Zeit für die einzelnen Frequenzkanäle), das Maximalwertspektrum und die frequenzabhängigen Abklingzeiten. Diese Merkmale tragen Invarianten zur Beschreibung der Resonanzqualität von Klanghölzern. Zur Auffindung der Invarianten wurden im Rahmen einer Lernphase 150 Objekte durch den gesamten Produktionsprozeß bis hin zur fertigen Gitarre meßtechnisch verfolgt. Die Prägnanzfunktion erwies sich bei der Abhebung der Invarianten als besonders wichtig, wobei aber auch das Maximalwertspektrum signifikante Beiträge liefert.

3.2. Ereignisfolgen

Typische Ereignisfolgen liegen bei monophoner (einstimmiger) Musik vor. Es liegt nahe, die Erkennung schneller Hüllkurvenveränderungen zur Ermittlung rhythmischer Muster heranzuziehen. Eine Illustration für die Anwendung des Enveloppemeters zur Analyse einer musikalischen Folgenstruktur wurde in /17/ gegeben. Weitere, mit verbesserter Zeitauflösung (0.5 ms-Raster) durchgeführte Untersuchungen an Gitarren zeigen:

1. Die Reproduzierbarkeit rhythmischer Spielmuster (Schlagdauerschwankungen) liegt bei guten Gitarristen unter 5 ms, der Unterschied von 1 ms ist statistisch hoch signifikant /18/
2. Spielfluß, Prägnanz und Flüssigkeit der Spielmuster liefern Anhaltspunkte für wichtige Gebrauchswerteigenschaften der Instrumente. Diagnostiziert werden können solche Eigenschaften über die Prägnanzfunktionen (s.Kap. 3.2) und die Regularität der Ereigniszeitpunkte /18/.
3. Bei der Beurteilung des Spielflusses werden verschiedene Strategien herangezogen, eine analytische, auf zeitliche Details im ms-Bereich orientierte Strategie und eine synthetische Strategie, die Ereignispunkte auf Basis der Hüllkurven-Zeit-Gestalt (vergl. /19/) erfaßt.

3.3. Sprachliche Silben

Eine recht komplizierte Situation liegt bei sprachlichen Silben vor. Das Problem wird deutlich, wenn man die Zeitfunktion des Gesamtsignals und durch Frequenzselektion ermittelte Signalanteile im logarithmischen Maßstab mittels eines Pegelschreibers darstellt (Bild 3). Die aufgezeichneten Pegeldiagramme für die

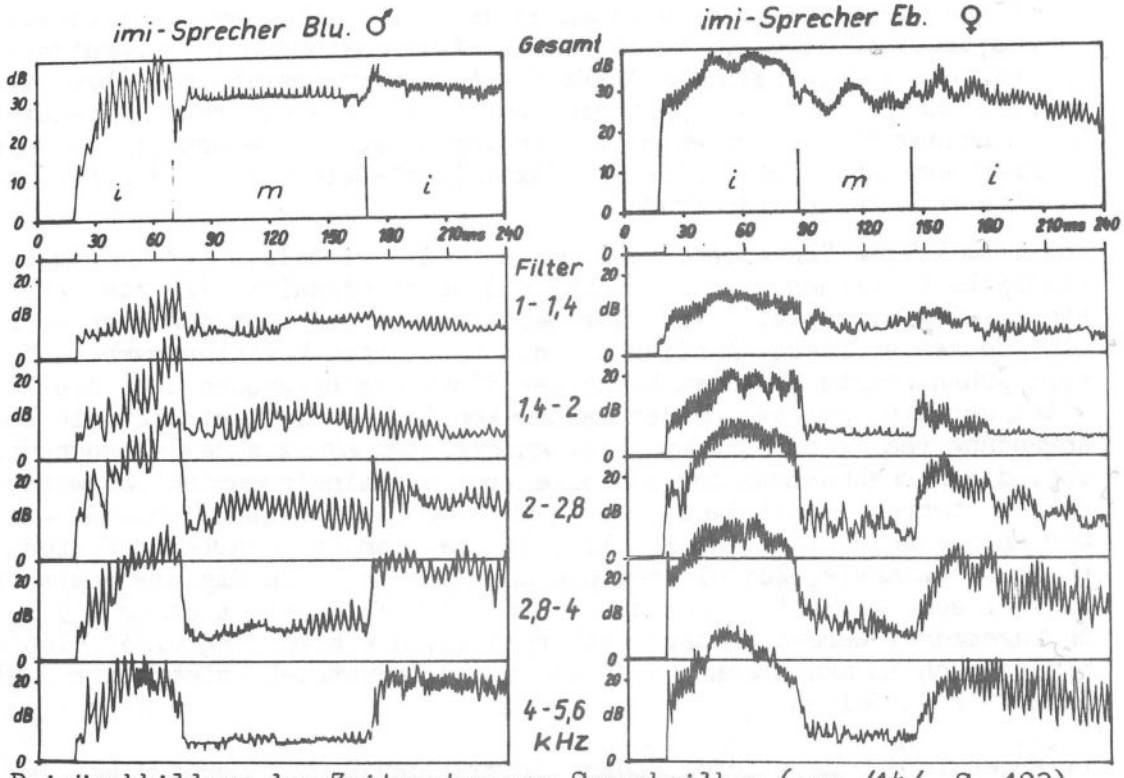


Bild 3: Primärabbildung der Zeitmuster von Sprachsilben (aus /14/, S. 123)

Silbe /imi/ zeigen zwar für die Filterkanäle ab etwa 2 kHz deutlich die Lautübergänge an (im Gesamtsignal sind diese kaum zu erkennen), jedoch weisen sie eine ausgesprochen zerklüftete Zeitstruktur auf. Diese "Rippung" ist auf das stark impulshaltige Anregungssignal der stimmhaften Sprachanteile zurückzuführen. Die Periodendauer der Stimmgrundfrequenz bildet sich naturgemäß bei der Männerstimme (linke Diagramme in Bild 3) deutlicher als bei der Frauenstimme (rechte Diagramme) im "gerippten" Zeitmuster ab.

Die in Abschnitt 2 erörterte Konzeption einer "konturisierten Schallrepräsentation" würde für solche "gerippten" Zeitmuster eine so dichte Folge von Zeitkonturen ergeben, daß eine regelrechte Informationsüberflutung die Folge wäre. Die eigentlich markanten Zeitereignisse (Onsets, Offsets, Transienten) würden in dieser Informationsflut untergehen. Hinzu kommt, daß die Grundfrequenz ja ebenfalls in den Teiltonmustern repräsentiert ist und die Tonhöhenempfindung weitaus adäquater mit Hilfe dieser Teiltonmuster beschrieben werden kann /4/. Es liegt deshalb nahe, nach Möglichkeiten Ausschau zu halten, um die störende "Rippung" zu beseitigen.

Entsprechende Möglichkeiten bieten Demodulator-Anordnungen (Bild 4).
Bei den konventionellen Varianten (obere beiden Illustrationen) besteht aber stets nur eine Kompromißmöglichkeit zwischen Minimierung der Restwelligkeit und Gewährleistung einer maximalen Zeitauflösung. Das Prinzip der Schnellen Hüllkurvendemodulation gewährleistet hingegen die vollständige Unterdrückung der Restwelligkeit bei einer Zeitauflösung, die nur noch eine Frage des Aufwands an Speicher-Grundelementen ist /20/.

Eine andere interessante Möglichkeit ist die Impulsdarstellung der Umhüllenden der Zeitfunktion durch Gaußimpulse /21/. Dieses Verfahren der "Impulsdauerquantisierung" gestattet ebenfalls eine sehr gute Signal-Systemanpassung unter Berücksichtigung von Eigenschaften des menschlichen Gehörs.

Eine Anwendung der Schnellen Hüllkurvendemodulation zur Segmentierung lautsprachlicher Folgenstrukturen ist in Bild 5 illustriert. Dargestellt ist die zeitkonturisierte Abbildung des Wortes "Gerade" sowie die Prägnanzfunktion für den Lauteinsatz /g/ (Kontur ①).

Die Segmentierung des Sprachsignals ist auf Grundlage von konturisierten Abbildungen unter Einbeziehung von Lautdauerstatistiken (Kontrastspeichermodell) mit hoher Zuverlässigkeit möglich, vorausgesetzt, es sind tatsächlich entsprechende Anhaltspunkte im Signal vorhanden. So konnten beispielsweise für die mit anderen Mitteln nur schwer zu strukturierenden Vokal-Nasallaut-Vokal-Verbindungen (vergl. Bild 5) Segmentierungsraten um 95 % erhalten werden /22/.

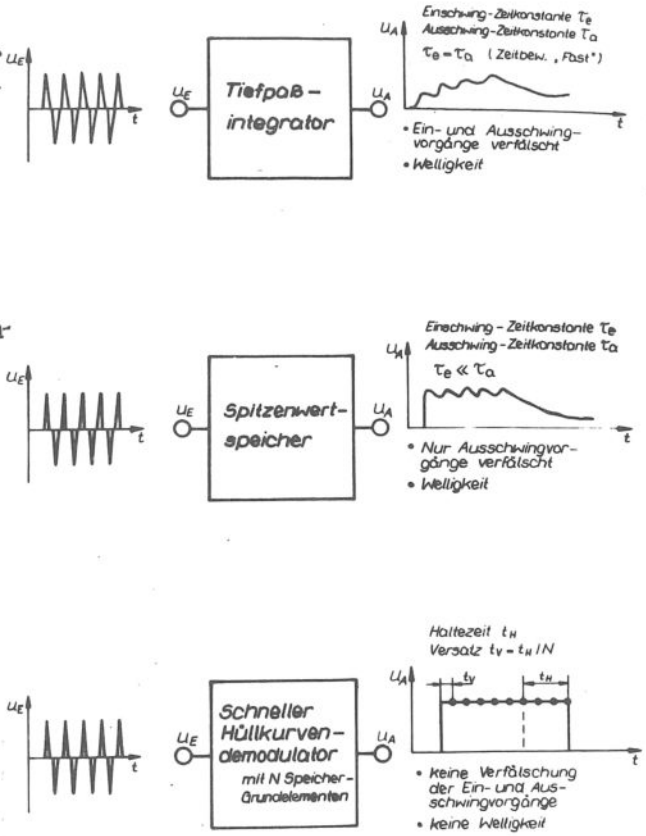


Bild 4: Vergleich verschiedener Demodulator-Anordnungen

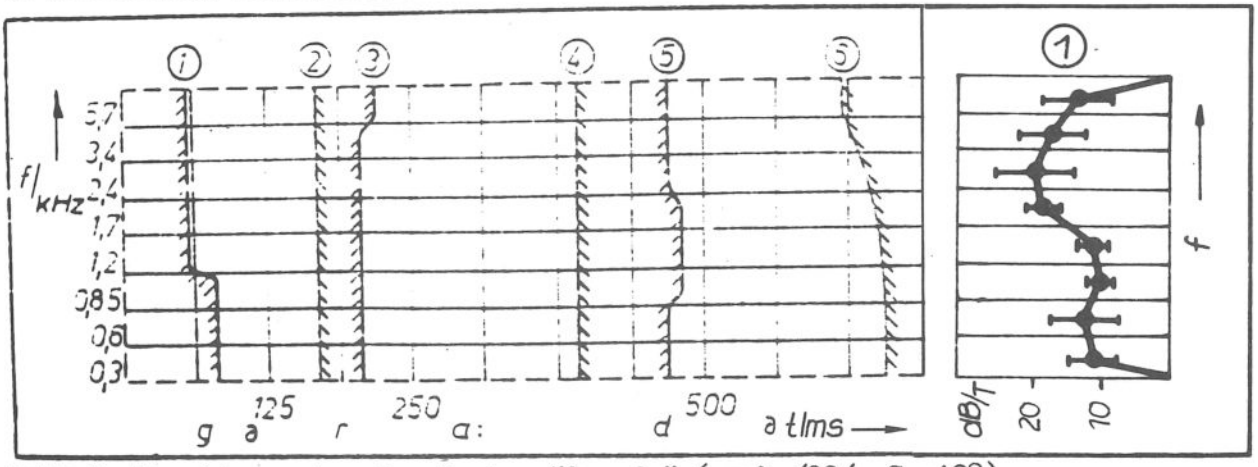


Bild 5: Strukturmuster des Wortes "Gerade" (nach /22/, S. 138)
∣∣ On-Signale ∣∣ Off-Signale, rechts: Prägnanzfunktion ①

4. Zusammenfassung

Es wurde gezeigt, daß die Verarbeitung der schnellen Hüllkurvenveränderungen von Audiosignalen eine Konturisierung im Zeitbereich darstellt. Dieser Vorgang stellt das Pendant zu dem von E. TERHARDT beschriebenen Entscheidungsmechanismus zur Rekrutierung der Teiltonmuster dar. Die für den Zeitbereich erhaltenen Resultate bestätigen die von Terhardt begründete Auffassung, daß es sich bei dem Teiltonhören möglicherweise um ein Schlüsselphänomen des Hörapparats handelt, ähnlich dem der visuellen Konturisierung. Während die Konturisierung in der Frequenzebene im wesentlichen auf dem Niveau der mikroskopischen Zeitskala (Innenohr) vollzogen wird, beruht die Konturisierung im Zeitbereich auf sekundären Prozessen der mittleren Zeitebene (Hörbahn).

Literaturverzeichnis

/1/ Blutner, F.: Entwurf eines psychoakustischen Funktionsmodells und Möglichkeiten der technischen Systemoptimierung. Dissertation B, Techn. Univ. Dresden 1987

/2/ Tscheschner, W.: Die Sprache und ihre Erkennung. Habilitationsschrift, Techn. Univ. Dresden 1967

/3/ Roederer, J.G.: Physikalische und psychoakustische Grundlagen der Musik. Berlin, Heidelberg, New York: Springer-Verlag 1977

/4/ Terhardt, E.: Warum hören wir Sinustöne? Naturwissenschaften 76, 1989

/5/ Terhardt, E.: Psychophysics of audio signal processing and the role of pitch in speech. In: M.E.H. Schouten (ed.): The psychophysics of speech perception. Dordrecht, Boston, Lancaster: Martinus Nijhoff Publishers 1987, S.271-283

/6/ Terhardt, E.: Characteristics of musical tones in relation to auditory acquisition of information. 1st Int. Conference on Music Perc. & Cognition. Kyoto: 17-19 Oct., 1989

/7/ Helmholtz, H.v.: Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik. Braunschweig: Friedr. Vieweg & Sohn, sechste Ausgabe 1913

/8/ Küpfmüller, K.: Einschwingvorgänge in Wellenfiltern. Elektr. Nachr. Techn. 1, 1924, S.141, 152

/9/ Backhaus, H.: Über die Bedeutung der Ausgleichsvorgänge in der Akustik. Z. techn. Phys. 13, 1932, S.31-46

/10/ Heinbach, W.: Gehörgerechte Repräsentation von Audiosignalen durch das Teiltonzeitmuster. Dissertation Techn. Univ. München 1987

/11/ Blutner, F.: Objektive und subjektive Untersuchungen zum Einschwingen von Pianoklängen. Das Musikinstrument, 33, 1982 S.1140-1146

/12/ Blutner, F.: Psychoakustik. Optimierung eines Zeithüllendetektors. Forschungsbericht Inst. für Musikinstr.-bau Zwota 1981

/13/ Blutner, F.: Verfahren zur Erfassung der Zeitenveloppe von Musik- und Sprachklängen WP G10L/210353 vom 5.1.1979

/14/ Blutner, F.: Akustische Gütebewertung mit EDV. Meßsystem zur Erfassung der Klangfarbe unter besonderer Berücksichtigung der Einschwingvorgänge (Enveloppemeter). Forschungsbericht Inst. für Musikinstr.-bau Zwota 1978

/15/ Blutner, F., Haller, K.: Multiples Modell zur Beschreibung subjektiver Klangattribute. Nachrichtentechnik, Elektronik 38, 1983, S.33-35

/16/ Blutner, F., Haller, K., Meinel, E., Dix, R., Schindler, V., Klopfer, F., Kümmel, F.: Klangdiagnose von Gitarren. Ergebnisprotokoll V5/0, Inst. für Musikinstr.-bau Zwota 1990

/17/ Blutner, F.: Psychoakustisches Funktionsmodell und Möglichkeiten der technischen Systemoptimierung. Nachrichtentechnik, Elektronik 38, 1983, S.21-26

/18/ Kluge, R.: Untersuchung über den Zusammenhang ausgewählter Zeithüllenparameter mit der subjektiven Spielflußbeurteilung von Gitarreanspielen. Arbeitsdok. Humboldt-Univ. Berlin 1990

/19/ Schütte, H.: Subjektiv gleichmäßiger Rhythmus: Ein Beitrag zur zeitlichen Wahrnehmung von Schallereignissen. Acustica 41, 1978, S.197-206

/20/ Blutner, F., Starke, H.: Verfahren und Anordnung zur schnellen Hüllkurvendemodulation. WP G10L/284 0096 vom 10.12.1985

/21/ Krause, M.: Impulsdarstellung von Sprachsignalen als Grundlage für ein adaptives Verfahren der Sprachanalyse und Sprachsynthese. Diss. Techn. Univ. Berlin 1971

/22/ Hausfeld, H.: Zur zeitlichen Strukturierung des Sprachsignals auf Grundlage der psychoakustischen Hüllkurvenverarbeitung. Diss. Techn. Univ. Dresden 1983