Zeit-Frequenz-Transformationen für instationäre Geräusche

Robert Liebing[1], Reinhard Weber[2]
[1] BMW AG, 80788 München, Deutschland, Email: Robert.Liebing@BMW.de
[2] Carl von Ossietzky Universität Oldenburg, 26111 Oldenburg, Deutschland, Email: Reinhard.Weber@uni-oldenburg.de

Einleitung

Die beim Soundengineering bearbeiteten Geräusche sind in der Regel instationär. So spielt neben der spektralen Zusammensetzung die zeitliche amplituden- und frequenzmäßige Entwicklung der Komponenten eine entscheidende Rolle für die Wahrnehmung. Zur Analyse der instationären Geräusche können verschiedene Zeit-Frequenz-Algorithmen eingesetzt werden. Der folgende Beitrag stellt Transformationen mit konstanter Bandbreite, wie z.B. die Wigner-Ville-Verteilung und solche mit variabler Bandbreite, zu denen Wavelets, Matching Pursuit und Constant-Q-Transformation gehören, gegenüber, und vergleicht sie am Beispiel eines synthetisierten Sinussweeps und eines instationären, sehr kurzen, realen Geräusches, das z.B. ein Klicken, Klappern oder Knarzen sein könnte, hinsichtlich einer wahrnehmungsnahen Analyse.

Zeit-Frequenz-Analyseverfahren

Transformationen konstanter Bandbreite

Das wohl am Häufigsten verwendete Verfahren zur Zeit-Frequenz-Analyse ist die so genannte Kurzzeit-Fourier-Transformation (STFT für short-time-fourier-transformation). Nach [1] bildet sie die Grundlage für die meisten anderen Zeit-Frequenz-Darstellungen. Die STFT zerlegt ein Signal in seine Kosinus- und Sinusanteile und man berücksichtigt den zeitlichen Bezug dadurch, dass die theoretisch unendlich lange Fourier-Transformation durch Multiplikation mit einer Fensterfunktion im Zeitbereich in kleine Blöcke unterteilt wird. Die Mittelung des Frequenz- und Zeitinhaltes erfolgt dann nur noch im Bereich des Blockes. Daraus resultiert ein festes Verhältnis von Auflösung und Genauigkeit im Zeit- und Frequenzbereich analog der „Heisenberg'schen Unschärferelation“. Damit ist die Tatsache beschrieben, dass die genaue Frequenz eines Schallereignisses zu einer bestimmten Zeit nicht bekannt ist. Zu deren Bestimmung muss das Zeitsignal erst eine gewisse Zeitdauer lang beobachtet werden. Das verringert die Zeitpräzision. Somit gilt es immer, einen Kompromiss zwischen guter Frequenzauflösung und damit einhergehender Verschlechterung der Zeitauflösung durch größere Blockgröße einerseits oder guter Zeitauflösung mit aus der kleineren Blockbreite resultierenden schlechteren Frequenzauflösung andererseits, zu finden.[2]

Bei der Wigner-Ville-Verteilung (WVD für Wigner-Ville-Distribution) handelt es sich um die Fouriertransformierte der zeitvarianten Autokorrelationsfunktion.[3] Man kann sie auch als eine STFT betrachten, deren Fensterfunktion die zeit- und frequenzversetzte Eingangsfunktion an sich ist. Der Vorteil der WVD gegenüber der STFT ist die erheblich höhere Zeit- und Frequenzauflösung, die nicht durch die Unschärferelation beschränkt ist. Allerdings erzeugt sie Kreuzterme (physikalisch unerklärbare Positionen in der Zeit-Frequenz-Ebene) und negative Werte, welche einer negativen Energie entsprechen und somit physikalisch unmöglich sind.

Transformationen variabler Bandbreite

Die Besonderheit dieser Verfahren liegt in der Eigenschaft, hochfrequente Komponenten von Signalen über einen relativ kurzen und niederfrequente über einen relativ langen Zeitraum zu analysieren. Hier zeigt sich auch der größte Vorteil gegenüber den Transformationen mit konstanter Analysebandbreite im ganzen Frequenzbereich. Bei der Wavelet-Analyse handelt es sich im Prinzip um eine Multiplikation des Eingangssignales mit einer Elementarfunktion im Zeitbereich. Deren zeitliche Länge bestimmt dann die spektrale Lage und Ausdehnung der Signalkomposition. So sind zeitlich lange ausgedehnte Elementarfunktionen relativ schmalbandig und tieffrequent, wohingegen eine zeitlich sehr kurze Basisfunktion ein hochfrequentes und breitbandiges Spektrum besitzt. Überlagert man jetzt die Spektren dieser verschiedenen Skalierungen, ergibt sich eine nahezu logarithmische Spektralzusammensetzung, wobei die Bandbreiten der einzelnen Komponenten proportional zu den Mittenfrequenzen sind. Diese Eigenschaft bezeichnet man als Filterung mit konstanter Güte $Q = \frac{Mittenfrequenz}{Bandbreite}$.

Die nahezu logarithmische Frequenzauflösung der Wavelet-Analyse kommt der des Menschen (vgl. Barkskala) sehr entgegen. Gleichzeitig bietet jedoch die STFT durch Faltung mit Sinus- und Kosinus-Funktionen im Frequenzbereich die Möglichkeit, sinusförmige Signalanteile besser aus einem Schall zu extrahieren, als dies mit einem typischen Wavelet möglich ist. Bei der Constant-Q-Transformation (CQT), die man als STFT mit variabler Fensterfunktion ansehen kann, werden beide Methoden miteinander vereint. Die bei der STFT konstante Fensterfunktion wird hier mit steigender Frequenz in der Zeitausdehnung verkürzt, so dass tonale Anteile erkennbar sind und in hohen Frequenzen dennoch eine hohe Zeitauflösung vorhanden ist.[4]

Einen ganz anderen Ansatz verfolgt der Matching-Pursuit Algorithmus (MP) von Mallat und Zhang.[5] Dieses Verfahren modelliert Signale durch iterative Superposition von in einem Wörterbuch abgelegten, variierbaren Prototyp-Funktionen. Der Algorithmus staucht und streckt (warping) die Funktionen aus dem Wörterbuch im Zeitbereich so lange, bis sie einem Teil des gesuchten Signals entsprechen. Im Anschluss wird dieser Teil (Elementarfunktionen) vom Original subtrahiert und mit Hilfe der WVD frei von Kreuztermen in den Zeit-Frequenz-Bereich transformiert. Das Restsignal unterliegt iterativ der gleichen Prozedur. Letztendlich überlagert der MP alle einzelnen Teilsignale in der Zeit-Frequenz-Ebene.

Transformation eines Sinussweeps

Zur exemplarischen Demonstration der Eigenschaften der verschiedenen Transformationen wird zum Einen ein mit 48kHz gesampelter, synthetisierter Sinussweep mit logarithmisch steigender und fallender Frequenz verwendet. Ebenfalls in logarithmischer Darstellung sollte sich eine ansteigende und eine abfallende Gerade für den zeitlichen Frequenzverlauf des Signals ergeben. Abbildung 1 zeigt, dass dies jedoch nur bei den Verfahren mit variabler Bandbreite der Fall ist. Die konstante Bandbreite bei der STFT führt zu einer Verbreiterung des Spektrums im niederen Frequenzbereich. Dies lässt sich jedoch auf Kosten der Zeitauflösung mit einer längeren Fensterfunktion korrigieren. Bei einigen Transformationen treten störende Kreuzterme auf, welche in der Zeit-Frequenz-Ebene nicht vorhandene Signalanteile vortäuschen.

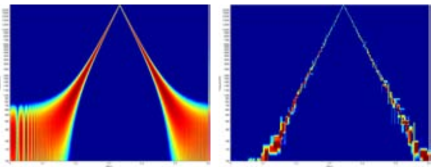

Abbildung 1: Verfahrensvergleich konstante und variable Bandbreite am Bsp. STFT und MP anhand eines Sweeps

Transformation eines transienten, realen Geräusches

Bei den transienten Geräuschen sind sehr kurzzeitige mit tonalen Anteilen besonders interessant. Deshalb besitzt das hier ausgewählte reale Beispielgeräusch neben mehreren, im Abstand von 10ms aufeinander folgenden, breitbandigen Impulsen auch tonale, länger andauernde Komponenten bei 250Hz und 650Hz. Die Algorithmen unterscheiden sich in ihrer Auflösung dadurch, dass die beiden tonalen Anteile unterschiedlich deutlich werden und die Anzahl der Einzelimpulse unterschiedlich hervor tritt. Des Weiteren sollen in der Darstellung die Frequenzschwerpunkte im Hauptgeräusch erkennbar sein. Im Frequenzbereich um 11kHz befinden sich zwei schmalbandige und sehr kurzzeitige tonale Komponenten. Auch diese sind separat voneinander darzustellen.

Wie man Abbildung 2 entnehmen kann, werden bei der STFT bedingt durch die kurze Fensterlänge nur die Impulse zufriedenstellend herausgearbeitet. Die tonalen Anteile gehen auf Grund der in tiefen Frequenzbereichen zu schlechten Frequenzauflösung verloren. Bei 650Hz ist einer bei der WVD gerade noch erkennbar. Allerdings fehlen jegliche Informationen über Frequenzschwerpunkte im Hauptgeräusch.

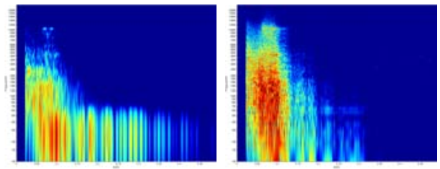

Abbildung 2: Verfahrensvergleich STFT und WVD anhand eines realen Geräusches

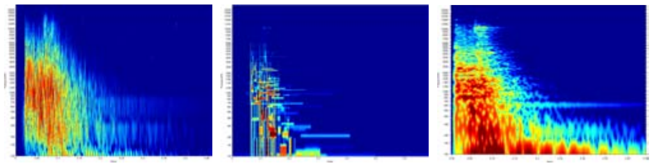

Abbildung 3: Verfahrensvergleich Wavelet, MP und CQT anhand eines realen Geräusches

Bei den Transformationen mit variabler Bandbreite analysiert insbesondere die CQT sowohl die tonalen, als auch die impulshaften Anteile sowie die Frequenzschwerpunkte sehr zufriedenstellend. Aber auch ein Wavelet (hier das Meyer-Wavelet) ist zur Analyse dieses Beispielgeräusches sehr gut geeignet, obgleich eine Aussage über den Frequenzschwerpunkt schwierig erscheint. Der Matching Pursuit Algorithmus ermöglicht die trennschärfste Analyse. Allerdings benötigt dieses Verfahren für eine höhere Genauigkeit als hier noch weitaus mehr Iterationen mit einem enormen Rechenaufwand.

Fazit

Verfahren mit variabler Fensterbreite sind deutlich besser zur Geräuschanalyse geeignet, als solche mit konstanter Fensterlänge. Sie vermeiden die nicht gehöradäquate, konstant hohe Auflösung entweder im Frequenz- oder im Zeitbereich und sind in der Lage, tonale Anteile in den für das Gehör wichtigsten Frequenzbereichen zu identifizieren und dennoch den Impulscharakter in hohen Frequenzbereichen aufzuzeigen. Unter den untersuchten Transformationen erscheint die CQT aufgrund der höchsten Genauigkeit und der relativ kurzen Rechenzeit für die Analyse sehr kurzer Signale am Geeignetsten.

Literatur

[1] Foundations of Time-Frequency Analysis. K. Gröchenig, Birkhäuser, 2001

[2] Time Frequency Signal Analysis and Processing. B. Boashash, Elsevier, 2003

[3] Zur Untersuchung der Frequenzdynamik nichtstationärer Fahrzeugsignale mittels Wavelet-Methoden. S. Ellis, VDI-Verlag, 2003

[4] Calculation of a constant Q spectral transform. J. C. Brown, Journal of the Acoustical Society of America **89** (1990), 425-434

[5] Matching Pursuit With Time-Frequency Dictionaries. S. Mallat and Z. Zhang, IEEE transactions in Signal Processing **26** (1993), 3397-3415