

Anwendung von Rahmen-Multiplikatoren für die Extraktion von Kurvenquietschen von Zugsaufnahmen

P. Balazs , C. Kasess , W. Kreuzer , T. Maly , Z. Průša , F. Jaillet 

Für viele Anwendungen in der Akustik ist es notwendig, Signale und Funktionen mithilfe von zeitvarianten Filtern zu bearbeiten, z. B. um Komponenten aus einem Signal zu entfernen, deren Frequenzverlauf sich über die Zeit ändert. Es wird eine Methode vorgestellt, die auf einer Darstellung des Signals durch Rahmen (engl. Frames) basiert, und mit deren Hilfe Filter auf der Zeit-Frequenz-Ebene definiert werden können. Nach einer kurzen Beschreibung des theoretischen Hintergrunds von Rahmen wird ihre Anwendung anhand eines Beispiels aus der Lärmforschung erläutert. Mithilfe einer einfachen grafischen Oberfläche wird aus einer Aufnahme einer Kurvenfahrt eines Zugs eine durch den Dopplereffekt zeitvariante Komponente (Kurvenquietschen) herausgeschnitten und in ein zweites Signal eingefügt. Auf diese Art und Weise lassen sich kontrollierte Signale generieren, die dann zur Lärmbewertung eingesetzt werden können.

Schlüsselwörter: Rahmen; Lärm; numerische Methoden; Zeit-Frequenz; open science

Application of frame multipliers for the extraction of curve squeals from train signals.

For many applications in acoustics it is necessary to process signals and functions with the help of time-variant filters, e.g., to remove components from a signal whose frequency profile changes over time. A method is presented that is based on a representation of the signal in terms of frames, where such time-varying filters can be defined in the time frequency plane. After a brief description of the theoretical background of frames, their application is explained using an example from noise research. With the help of a simple graphical user interface, a component (curve squeal) that is not time-invariant due to the Doppler effect is cut out of a recording of a train and inserted into a second signal. This way, controlled signals can be generated, which can then be used for noise assessment.

Keywords: frames; noise; numerics; time frequency; open science

Eingegangen am 15. Jänner 2021, angenommen am 11. März 2021, online publiziert am 13. April 2021
© The Author(s) 2021



1. Einleitung

In vielen Anwendungen in der Akustik sind Methoden zur Darstellung und Manipulation von Signalen notwendig, zum Beispiel um (menschliche) Sprache zu analysieren, oder um kontrolliert Testsignale für psychoakustische Tests zu generieren, die unter anderem zur Lärmbewertung benutzt werden. Oft ist es nicht möglich, Signal-komponenten mit einem einfachen linearen Filter zu isolieren, weil sich verschiedene Frequenzkomponenten über die Zeit ändern. Ein Beispiel dafür tritt bei der Bewertung und Modellierung von Kurvenquietschen auf. Kurvenquietschen ist ein tonaler, also schmalbandiger Klang, der durch periodisches Rutschen des Rades auf der Schiene verursacht wird [26]. Das auf diese Art zur Schwingung angeregte Rad strahlt in einem oder mehreren schmalbandigen Moden ab. An sich ändert sich das daraus resultierende Quellsignal am Rad nicht über die Zeit, aber durch die Bewegung des Zugs und des Rades relativ zur Beobachtungsstelle ändert sich die Frequenz des Tons durch den Dopplereffekt an der stationären Messstelle.

In Abb. 1 ist das Spektrogramm eines Kurvenquietschens, das in 25 m Entfernung zur Gleisachse gemessen wurde, dargestellt. Entlang der x-Achse kann der zeitliche Verlauf des Signals abgelesen werden, entlang der y-Achse sind die Frequenzen aufgetragen und helle Stellen bedeuten eine hohe Energie des Signals an der jeweiligen Zeit-Frequenz-Position. Das Kurvenquietschen ist leicht im Frequenzbereich um 4000 Hz zu erkennen.

Um nun aus der ganzen Aufnahme das Kurvenquietschen für weitere psychoakustische Experimente zu isolieren, muss ein *zeit-variant*er Filter verwendet werden. Eine Möglichkeit, so einen Filter zu generieren, ist, eine geeignete Zeit-Frequenz-Darstellung des Signals zu finden, diese Darstellung zu manipulieren und anschließend aus der neuen Zeit-Frequenz-Darstellung das gefilterte Signal zu re-synthetisieren. Gute Zeit-Frequenz-Darstellungen bedingen jedoch Redundanz, also etwa durch eine Überlappung der Fenster in einer Kurzzeit-Fourier-Transformierten, die durch eine klassische Darstellung mittels Basisfunktionen nicht gegeben ist. Wir verwenden daher spezielle Methoden, die auf dem Konzept von Rahmen und Rahmen-Multiplikatoren basieren und das „Ausschneiden“ und „Kombinieren“ von Signalteilen aus Spektrogrammen ermöglichen, bzw. ganz vereinfacht ausgedrückt: Methoden, die uns „Photoshop für Signale“ ermöglichen. Dafür verwenden wir MULACLAB, eine

Balazs, Peter, Österreichische Akademie der Wissenschaften, Institut für Schallforschung, Wohllebengasse 12-14, 1040 Wien, Österreich (E-Mail: peter.balazs@oeaw.ac.at); **Kasess, Christian**, Österreichische Akademie der Wissenschaften, Institut für Schallforschung, Wien, Österreich; **Kreuzer, Wolfgang**, Österreichische Akademie der Wissenschaften, Institut für Schallforschung, Wien, Österreich; **Maly, Thomas**, Technische Universität Wien, Institut für Verkehrswissenschaften, Wien, Österreich; **Průša, Zdeněk**, LEWITT GmbH, Wien, Österreich; **Jaillet, Florent**, Université Côte d'Azur, CNRS, I3S, Sophia Antipolis, France

open-source Implementierung, die auf der Large Time-Frequency Analysis Toolbox¹ (LTFAT) basiert.

In Abschn. 2 wird der theoretische Hintergrund dieser Zeit-Frequenz-Darstellung kurz beschrieben, wir definieren die Begriffe „Rahmen“ und „Rahmen-Multiplikatoren“ und zeigen, wie basierend darauf auf einfache Weise zeitabhängige Filter definiert werden können. In Abschn. 3 stellen wir MULACLAB vor, ein Matlab/Octave-Programm mit grafischer Oberfläche, mit dessen Hilfe Signale in der Zeit-Frequenz-Ebene manipuliert werden können. Als Beispiel wird MULACLAB in Abschn. 4 verwendet, um Testsignale aus Aufnahmen von Zugvorbeifahrten in Kurven zu generieren, die für psychoakustische Untersuchungen von Kurvenquietschen verwendet werden. Zum Abschluss werden die Ergebnisse und Methoden in Abschn. 5 zusammengefasst.

2. Rahmen

Eine wichtige Eigenschaft zahlreicher numerischer Methoden ist die Darstellung von kontinuierlichen Funktionen (z. B. Signalen) durch diskrete Bausteine (Atome) und deren Koeffizienten:

$$f(t) = \sum_{k=0}^{\infty} f_k \psi_k(t). \quad (1)$$

Ein bekanntes Beispiel aus der Signalverarbeitung sind Fourierreihen, mit denen periodische Funktionen als Summe von Sinus- und Kosinusfunktionen dargestellt werden können, und das Abtasttheorem [23], das eine Brücke zwischen kontinuierlichen zeitbeschränkten Signalen und den abgetasteten Werten des Signals an diskreten Stützstellen schlägt. Ganz klassisch werden häufig orthonormale Basen als Atome verwendet, es ist jedoch in vielen Anwendungen von Vorteil, redundante, also nicht eindeutige Darstellungen zu verwenden. Solche Darstellungen erlauben es zum Beispiel Spektrogramme „glatter“ erscheinen zu lassen oder die Darstellung effizienter und robuster gegenüber Störungen und Rauschen zu machen. Für die Kurzzeit-Fourier-Transformation (STFT) [12] (auch Gabor Transformation [10] genannt) kann sogar gezeigt werden, dass es ohne Redundanz unmöglich ist, eine gut lokalisierte Darstellung zu haben, im Sinne, dass nur eine Zeit-Frequenz-Region nahe des jeweiligen Punkts diesen beeinflusst. Durch die Zeit-Frequenz-Unschärfe ist es jedoch nie möglich, hier eine beliebige Genauigkeit zu bekommen. Für „vernünftige“ Funktionen/Signale zum Beispiel mit endlicher Energie (d.h. $f(t) \in L^2(\mathbb{R})$) kann so ein darstellendes Erzeugendensystem durch Rahmen beschrieben werden. Ein Rahmen [7, 8] in einem (Hilbert)-Raum \mathcal{H} ist eine Familie von Funktionen $\Psi = (\psi_k)$, wofür Konstanten $A, B > 0$ existieren, sodass

$$A\|f\|^2 \leq \sum_{k \in K} |\langle f, \psi_k \rangle|^2 \leq B\|f\|^2 \quad (2)$$

für alle $f \in \mathcal{H}$ gilt. Hier ist $k \in K$ aus einer (unendlich großen) Indexmenge K , und $\langle f, g \rangle$ bezeichnet das innere Produkt im Hilbertraum. Z. B. für $f, g \in L^2(\mathbb{R})$ gilt: $\langle f, g \rangle = \int_{-\infty}^{\infty} f(t)g^*(t)dt$, wobei $g^*(t)$ die konjugiert komplexe Funktion zu $g(t)$ ist. Die auf den ersten Blick abstrakte Bedingung Gl. (2) hat einige wichtige Konsequenzen [4]:

- **Stabilität:** Gl. (2) stellt eine Beziehung zwischen einem Signal f und den Koeffizienten $f_k = \langle f, \psi_k \rangle$ einer diskreten Darstellung her. Für moderate Rahmenparameterquotienten $\frac{B}{A}$ ist die Darstellung stabil, d.h. ähnliche Signale besitzen ähnliche Koeffizienten, und ähnliche Koeffizienten erzeugen umgekehrt wieder ähnliche Signale. Für numerische Verfahren bedeutet dies, dass die Konditionszahl der Transformation durch $\sqrt{\frac{B}{A}}$ bestimmt ist.

¹<http://lftfat.github.io/>.

- **Verlustfreie Wiederherstellung:** Eine Konsequenz der Bedingung ist, dass jedes Signal mit Hilfe eines Rahmens Ψ zerlegt werden kann (Analyse), und dass es mindestens einen dualen Rahmen Φ gibt, mit dem das Signal aus den Komponenten wieder *verlustfrei* zusammengesetzt werden kann (Synthese), d.h.

$$f(t) = \sum_{k \in K} \langle f, \psi_k \rangle \phi_k = \sum_{k \in K} \langle f, \phi_k \rangle \psi_k. \quad (3)$$

- **Redundanz:** Im Allgemeinen ist die Darstellung mittels Rahmen nicht eindeutig, was wie oben beschrieben oft erst eine gute Darstellung erlaubt. Gegenüber Basen gibt es mehr Freiheiten, Rahmen mit besonderen Eigenschaften zu konstruieren, und falls bei einer möglichen Datenübertragung eine Komponente verrauscht ist oder verloren geht, ist ein Teil der Information immer noch in anderen Koeffizienten enthalten [7]. Darüber hinaus erlaubt die Redundanz mehr als einen dualen Rahmen, der eine verlustfreie Wiederherstellung garantiert, was etwa in der Resynthese auch wieder mehr Optionen ermöglicht [18].
- **Linearität:** Das Zusammenspiel von Rahmen und dualen Rahmen erlaubt eine signalunabhängige, lineare Invertierung und eine einfache Manipulation von Signalen oder Signalteilen. Während viele akustische Phänomene inhärent nicht-linear sind, und daher ein Rahmen-Ansatz nur eine Approximation der Realität ist, sei darauf hingewiesen, dass die lineare Darstellung neben der Einfachheit auch unschlagbare Vorteile in der Interpretation bietet. Selbst wenn auf eine Darstellung mittels Rahmen eine nicht-lineare Manipulation folgt, ist es etwa immer noch klar, was mit Begriffen wie Bandbreite gemeint ist. In einer linearen Darstellung ist auch eine Trennung im Zeit-Frequenz-Bereich äquivalent zu einer Trennung in der Signaldomäne.

Die Rahmen-Theorie ist nicht auf unendlich dimensionale, kontinuierliche Räume, wie zum Beispiel Funktionenräume, beschränkt. Man kann die Theorie auch auf diskrete Räume (z.B. den Raum ℓ^2 der (unendlichen) Folgen mit endlicher Energie) oder auch auf endlich-dimensionale Räume (z. B. \mathbb{R}^N oder \mathbb{C}^N) anwenden. Im letzteren Fall entspricht ein Rahmen wie oben schon erwähnt einem Erzeugendensystem, wie es aus der linearen Algebra bekannt ist. Jedoch erlaubt uns ein Rahmen in diesem Fall *immer*, eine explizite Rekonstruktionsformel anzugeben und die Robustheit der Darstellung durch die Konditionszahl $\sqrt{\frac{B}{A}}$ zu beschreiben.

2.1 Rahmen-Multiplikatoren

Basierend auf der Darstellung mittels Rahmen ist es möglich, durch die Manipulation der Rahmenkoeffizienten einen zeitabhängigen Filter direkt in der Zeit-Frequenz-Ebene zu definieren. Durch das Zusammenspiel aus Analyse, Manipulation der Rahmenkoeffizienten mittels des Symbols (oder der Maske) m_k und Resynthese mit Hilfe eines dualen Rahmens lässt sich somit ein Rahmenmultiplikator [3, 9] durch

$$M_{m, \Phi, \Psi} f = \sum_{k \in K} m_k \langle f, \phi_k \rangle \psi_k \quad (4)$$

definieren. Diese Multiplikatoren treten in vielen verschiedenen wissenschaftlichen Disziplinen auf. In der *Mathematik* werden sie für die Diagonalisierung von Operatoren [22] verwendet. In der *Physik* stellen sie eine mögliche Verbindung zwischen klassischer Mechanik und Quantenmechanik dar, sogenannte Quantisierungsoperatoren [2]. In der *Signalverarbeitung* ermöglichen sie es, zeitvariante Filter zu implementieren [15], welche dann in der *Akustik* eingesetzt werden, etwa für die Signaltrennung [27]. Rahmen-Multiplikatoren

treten daher nicht nur als mathematische Objekte auf [6, 25], es werden auch deren Anwendungen studiert, wie zum Beispiel für die Entfernung von perceptiv irrelevanten Zeit-Frequenz-Punkten [5], oder auch als Möglichkeit, die Ähnlichkeit von Instrumenten [17] oder Sprechern [14] zu beurteilen.

Ein einfaches Beispiel, das etwa in der computergestützten auditorischen Szenenanalyse (CASA [28]) aber auch für das Beispiel in Abschn. 4 verwendet wird, ist eine binäre Maske: $m_k = 1$, wobei k in einer Untermenge $K' \subset K$ liegt. Für alle anderen k gilt: $m_k = 0$. Mit dieser einfachen Maske ist es möglich, Komponenten aus Signalen zu schneiden.

2.2 Gabor-Rahmen

Gabor-Rahmen sind eine spezielle Art von Rahmen, die in der Signalverarbeitung sehr häufig vorkommen, und die den diskreten Unterbau zur Kurzzeit-Fourier-Transformation (STFT)

$$S_g f(\tau_k, \nu_\ell) = \int_{-\infty}^{\infty} f(t)g(t - \tau_k)e^{-2\pi i \nu_\ell t} dt$$

liefern. Die STFT wird u.a. verwendet, um Zeit-Frequenz-Darstellungen von Signalen zu generieren; Spektrogramme sind der (quadrierte) Absolutbetrag der STFT, oft in einer logarithmischen Darstellung. Dabei wird das Signal mit einem Fenster multipliziert, das (normalerweise) um den Zeitpunkt $\tau_0 = 0$ konzentriert ist. Durch eine Fourier-Transformation kann dann diesem Zeitpunkt ein Spektrum mit Frequenzvariable ν_ℓ zugeordnet werden. Das Fenster wird auf den Zeitpunkt $\tau_{k+1} = \tau_k + \Delta t$ verschoben und die ganze Prozedur wiederholt. Interpretieren wir die einzelnen Funktionen $g_{k,\ell}(t) = g(t - \tau_k)e^{-2\pi i \nu_\ell t}$ als Atome eines Erzeugendensystems, kann gezeigt werden, dass unter bestimmten Bedingungen an das Fenster $g(t)$ und an die Parameter τ_k und ν_ℓ das Gaborssystem $\mathcal{G}(\mathcal{G}, \tau_k, \nu_\ell) = \{g_{k,\ell}\}_{k \in K, \ell \in L}$ einen s.g. Gabor-Rahmen bildet und somit die oben erwähnten Eigenschaften besitzt. Es kann auch gezeigt werden [12], dass zu jedem Gabor-Rahmen mindestens ein dualer Rahmen existiert, der ebenfalls eine Gaborstruktur besitzt, d.h. der duale Rahmen ergibt sich aus (Zeit-)Translation und (Frequenz-)Modulation eines dualen Fensters \tilde{g} , und es gilt:

$$f(t) = \sum_{k,\ell} f_{k,\ell} g_{k,\ell} \text{ mit } f_{k,\ell} = \langle f, \tilde{g}_{k,\ell} \rangle. \quad (5)$$

In einem Spektrogramm ergibt sich dann jeder Punkt als $|f_{k,\ell}|^2$. Da hier die Phaseninformation verloren geht, ist eine direkte Invertierung nicht trivial und führt zum Thema der Phasen-Rekonstruktion [1, 19]. Daher wird normalerweise ein Multiplikator nicht auf das Spektrogramm selbst, sondern auf die dahinter liegenden komplexen Koeffizienten $f_{k,\ell}$ der Zeit-Frequenz-Darstellung mittels Rahmen angewandt.

Die Toolbox LTFAT (Large Time-Frequency Analysis Toolbox), die am ISF entwickelt und betreut wird [20, 24], stellt in diesem Zusammenhang eine Vielzahl von nützlichen Funktionen zur Verfügung. Sie wird für Matlab/Octave entwickelt, beinhaltet C/C++ Implementierungen und eine Anbindung an Python ist gerade in Ausarbeitung. Sie ist open source, in vielen bekannten Linux-Distributionen bereits enthalten und fördert durch ihre freie Verfügbarkeit auch reproduzierbare Forschung.

So sollen möglichst viele Zeit-Frequenz- bzw. Rahmen-Implementierungen in diese Toolbox eingebunden, gut dokumentiert, gewartet und auch verwendet werden. Eine Bestimmung der Gabor-Koeffizienten lässt sich in LTFAT zum Beispiel durch den einfachen Befehl `dgtrreal` bewerkstelligen, die Anwendung eines Rahmen-Multiplikators mit Hilfe von `framemul`.

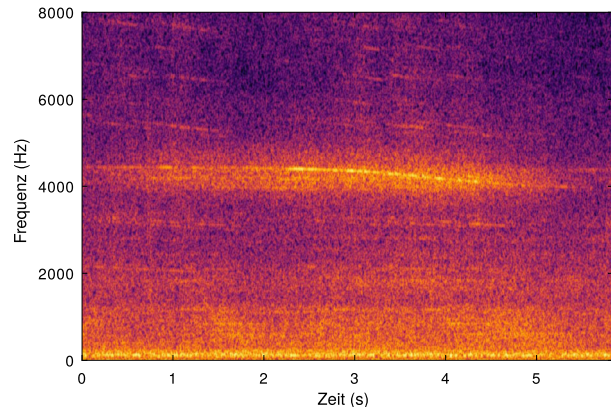


Abb. 1. Spektrogramm eines Kurvenquietschens, das in 25 m Abstand zur Gleisachse aufgenommen wurde

3. MULA CLAB

Speziell zur Bearbeitung von Zeit-Frequenz-Darstellungen bietet LTFAT auch eine Routine mit graphischer Oberfläche (Abb. 2) an, in der auf einfache Weise Regionen in der Zeit-Frequenz-Ebene definiert und manipuliert werden können. Nach dem Aufruf von MULA CLAB kann eine wav-Datei importiert werden, aus der ein Spektrogramm generiert wird. In diesem Spektrogramm können Abschnitte markiert und bearbeitet werden. Die Auswahl dieser Abschnitte erfolgt entweder durch das manuelle Markieren des Randes der Region mittels Maus oder durch ein „Zauberstab“-Tool, das rund um einen Punkt mit Hilfe eines „Regiongrowing“-Algorithmus eine Region innerhalb eines definierten dynamischen Bereichs festlegt.

Verschiedene Regionen können nun ausgeschnitten, kombiniert, und miteinander geschnitten werden. Wichtig dabei ist, dass der Multiplikator auf die vollen komplexen STFT-Koeffizienten angewandt wird. Durch ihre Interpretation und Verallgemeinerung im Konzept von Rahmen können in der LTFAT auch andere Zeit-Frequenz-Darstellungen verwendet werden, etwa Wavelets [11], eine konstant-Q-Transformation [13] oder auch perzeptive Filterbänke [16].

4. Anwendung

Als Beispiel für zeitabhängige Filterung betrachten wir die psychoakustische Untersuchung von Kurvenquietschen [26]. Beim Kurvenquietschen entsteht ein schmalbandiger Klang, dessen Frequenz sich durch den Dopplereffekt beim ortsfesten Empfänger über die Zeit ändert.

Um den Effekt eines solchen Signals auf die Wahrnehmung systematisch zu untersuchen, muss das Quietschen aus Aufnahmen isoliert, möglicherweise manipuliert und dann wieder mit dem Geräusch eines vorbeifahrenden Zugs kombiniert werden. So ist es möglich, *kontrolliert* vergleichbare Testsignale mit und ohne Quietschen zu generieren, die in psychoakustischen Experimenten verwendet werden können.

Als Testsignale wurden Aufnahmen von Zugs vorbeifahrenden in einer Kurve in 25 m bzw. 50 m Entfernung verwendet. Zur Manipulation der Signale wurde ein Gabor-Rahmen (Hanning Fenster mit einer Länge von 1024 Samples) verwendet, wobei die Fenster jeweils um 256 Samples verschoben wurden. Für jede Zeitinstanz wurden 2048 Frequenz-bins berechnet.

In Abb. 1 ist ein Signal mit Quietschen dargestellt, das deutlich im Bereich um 4 kHz erkennbar ist. Nachdem sich die Frequenz des tonalen Signals durch den Dopplereffekt über die Zeit verändert, wurde die Signalkomponente mit Hilfe des Spektrogramms einerseits

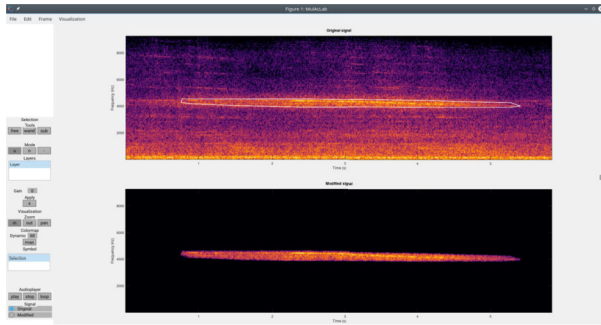


Abb. 2. Benutzeroberfläche von MULACLAB in Octave. Links befinden sich die verschiedenen Werkzeuge, um Regionen auszuwählen und in verschiedenen Ebenen zu kombinieren. Das obere Spektrogramm zeigt wieder das Kurvenquietschen mit einer manuell ausgewählten Region, das untere Spektrogramm zeigt die Zeit-Frequenz-Darstellung, nachdem diese Region isoliert wurde

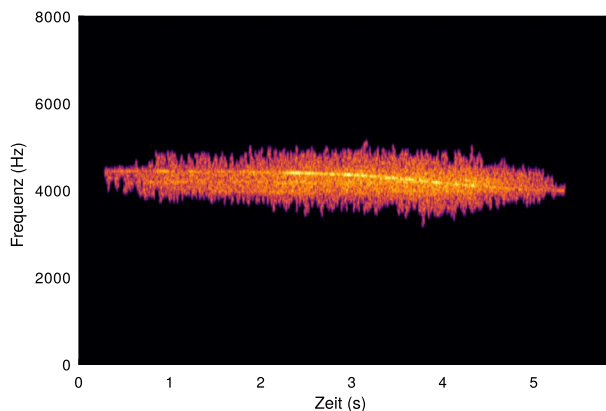


Abb. 3. Spektrogramm des Quietschens, nachdem es isoliert wurde. Die Maske wurde mittels des Magic-Wand-Tools erzeugt

durch händisches Markieren (Abb. 2) und andererseits durch das Magic-Wand-Tool mit 25 dB Range ausgewählt (siehe Abb. 3) und dann mit Hilfe eines Multiplikators der Rest des Signales entfernt.

Die isolierte Komponente wurde dann mit dem Signal einer Vorbeifahrt ohne Quietschen kombiniert, um somit ein kontrolliertes Testsignal zu generieren. Nachdem das Quietschen in der Amplitude groß genug ist, reicht ein einfaches Überlagern der Signale, um ein realistisches Signal zu erzeugen. Alternativ kann die Maske für die tonale Komponente invertiert und vor dem Kombinieren auf das Testsignal angewandt werden, um Effekte, die durch das Überlagern entstehen könnten, zu verringern (Abb. 4). Im Versuch mit einem, in Bezug auf Zuglärm erfahrenen Hörer (einer der Koautoren), in dem das originale Signal mit Quietschen und kombinierten Testsignalen anonymisiert dargeboten wurden, wurden die modifizierten Signale zwar wie zu erwarten als unterschiedlich, aber alle für realistisch befunden. Die Untersuchung der subjektiven Beurteilung des Quietschens wird Teil zukünftiger psychoakustischer Experimente sein.

5. Zusammenfassung

Das obige Beispiel zeigt die Flexibilität, die Rahmen und Rahmen-Multiplikatoren im Bezug auf Manipulation von Signalen erlauben. MULACLAB erlaubt eine einfache Modifikation von Signalen über ihre Spektrogramme ähnlich zu grafischen Werkzeugen (Ausschneiden, Verschieben, Kombinieren).

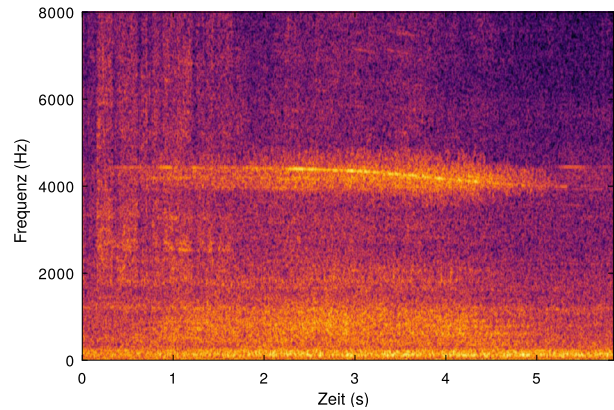


Abb. 4. Spektrogramm des Quietschens, nachdem es mit dem Testsignal kombiniert wurde. Zuvor wurde die Maske verwendet, um Signalkomponenten des Testsignals im Bereich des Quietschens zu entfernen

Durch diese Methode ist es möglich, kontrollierte Testsignale aus der Kombination verschiedener Quellen zu generieren. Diese Signale können danach auch manipuliert werden. Im Falle des Quietschens kann z. B. die Grundfrequenz verändert [21] oder ein räumlicher Eindruck durch Modifikation mittels einer Kopfübertragungsfunktion (HRTF, [29]) erzeugt werden. Dadurch ist es möglich, psychoakustische Tests durchzuführen, um zum Beispiel Vorher-Nachher Effekte zu simulieren, oder um Testsignale von Schallquellen zu generieren, die nicht oder nur unter großem messtechnischen Aufwand gemessen werden können.

Danksagung

Diese Arbeit wurde zum Teil durch die Österreichische Forschungsförderungsgesellschaft (FFG Projekt 860523), das Bundesministerium für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie, die ÖBB-Infrastruktur AG und dem FWF (START Projekt FLAME Y551-N13) unterstützt.

Hinweis des Verlags Der Verlag bleibt in Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutsadressen neutral.

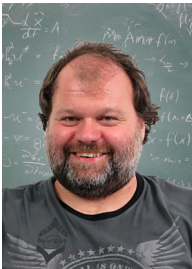
Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Alaifari, R., Grohs, P. (2019): Gabor phase retrieval is severely ill-posed. *Appl. Comput. Harmon. Anal.*, 50, 401–419.
- Ali, S. T., Antoine, J. P., Gazeau, J. P. (2000): Coherent states, wavelets and their generalization. *Graduate texts in contemporary physics*. New York: Springer.
- Balazs, P. (2007): Basic definition and properties of Bessel multipliers. *J. Math. Anal. Appl.*, 325(1), 571–585.

4. Balazs, P., Holighaus, N., Necciarì, T., Stoeva, D. (2017): Frame theory for signal processing in psychoacoustics. In R. Balan, J. J. Benedetto, W. Czaja, K. Okoudjou (Hrsg.), *Excursions in harmonic analysis* (Vol. 5, S. 225–268). Berlin: Springer.
5. Balazs, P., Laback, B., Eckel, G., Deutsch, W. A. (2010): Time-frequency sparsity by removing perceptually irrelevant components using a simple model of simultaneous masking. *IEEE Trans. Audio Speech Lang. Process.*, 18(1), 34–49.
6. Benedetto, J., Pfander, G. (2006): Frame expansions for Gabor multipliers. *Appl. Comput. Harmon. Anal.*, 20(1), 26–40.
7. Christensen, O. (2016): *An introduction to frames and Riesz bases*. Basel: Birkhäuser.
8. Duffin, R. J., Schaeffer, A. C. (1952): A class of nonharmonic Fourier series. *Trans. Am. Math. Soc.*, 72, 341–366.
9. Feichtinger, H. G., Nowak, K. (2003): *A first survey of Gabor multipliers* (S. 99–128). Boston: Birkhäuser. Chap. 5.
10. Feichtinger, H. G., Strohmer, T. (1998): *Gabor analysis and algorithms - theory and applications*. Boston: Birkhäuser.
11. Flandrin, P. (1999): *Time-frequency/time-scale analysis*. San Diego: Academic Press.
12. Gröchenig, K. (2001): *Foundations of time-frequency analysis*. Boston: Birkhäuser.
13. Holighaus, N., Dörfler, M., Velasco, G. A., Grill, T. (2013): A framework for invertible, real-time constant-Q transforms. *IEEE Trans. Audio Speech Lang. Process.*, 21(4), 775–785.
14. Huang, F., Balazs, P. (2019): Harmonic-aligned frame mask based on non-stationary Gabor transform with application to content-dependent speaker comparison. In *Inter-speech* (Vol. 2019).
15. Matz, G., Hlawatsch, F. (2002): Linear time-frequency filters: on-line algorithms and applications. In A. Papandreou-Suppappola (Hrsg.), *Application in time-frequency signal processing* (S. 205–271). Boca Raton (FL): CRC Press. Chap. 6.
16. Necciarì, T., Holighaus, N., Balazs, P., Průša, Z., Majdak, P., Derrien, O. (2018): Audlet filter banks: a versatile analysis/synthesis framework using auditory frequency scales. *Appl. Sci.*, 8(1), 96.
17. Olivero, A., Torrèsani, B., Kronland-Martinet, R. (2013): A class of algorithms for time-frequency multiplier estimation. *IEEE Trans. Audio Speech*, 21(8), 1550–1559.
18. Perraudin, N., Holighaus, N., Søndergaard, P., Balazs, P. (2018): Designing Gabor windows using convex optimization. *Appl. Math. Comput.*, 330, 266–287.
19. Průša, Z., Balazs, P., Søndergaard, P. L. (2017): A non-iterative method for (re)construction of phase from stft magnitude. *IEEE Trans. Audio Speech Lang. Process.*, 25(5), 1154–1164.
20. Průša, Z., Søndergaard, P. L., Holighaus, N., Wiesmeyr, C., Balazs, P. (2014): The large time-frequency analysis Toolbox 2.0. In M. Aramaki, O. Derrien, R. Kronland-Martinet, S. Ystad (Hrsg.), *Sound, music, and motion, lecture notes in computer science* (S. 419–442). Berlin: Springer.
21. Průša, Z., Holighaus, N. (2017): Phase vocoder done right. In *Proceedings of 25th European signal processing conference (EUSIPCO-2017)*, Kos (S. 1006–1010).
22. Schatten, R. (1960): *Norm ideals of completely continuous operators*. Berlin: Springer.
23. Shannon, C. E. (1949): Communication in the presence of noise. *Proc. IRE*, 37(1), 10–21.
24. Søndergaard, P., Torrèsani, B., Balazs, P. (2012): The linear time frequency analysis toolbox. *Int. J. Wavelets Multiresolut.*, 10(4), 1250, 032.
25. Stoeva, D. T., Balazs, P. (2012): Invertibility of multipliers. *Appl. Comput. Harmon. Anal.*, 33(2), 292–299.
26. Thompson, D. (2009): Curve squeal noise. In D. Thompson (Hrsg.), *Railway noise and vibration* (S. 315–342). Amsterdam: Elsevier. Chapter 9.
27. Vincent, E., Gibbonval, R., Fevotte, C. (2006): Performance measurement in blind audio source separation. *IEEE Trans. Audio Speech Lang. Process.*, 14(4), 1462–1469.
28. Wang, D., Brown, G. J. (2006): *Computational auditory scene analysis: principles, algorithms, and applications*. New York: Wiley-IEEE Press.
29. Ziegelwanger, H., Majdak, P., Kreuzer, W. (2015): Numerical calculation of head-related transfer functions and sound localization: microphone model and mesh discretization. *J. Acoust. Soc. Am.*, 138(1), 208–222.

Autoren

**Peter Balazs**

studierte Mathematik und Mathematik/Physik Lehramt an der Universität Wien. 2001 schloss er sein Studium mit einer Diplomarbeit zum Thema Polynome über Gruppen mit Auszeichnung ab. Im Juni 2005 beendete er sein Dissertationsstudium mit einer Arbeit über Regular and Irregular Gabor Multiplier with Application to Psychoacoustic Masking mit Auszeichnung.

Peter Balazs ist seit 1999 Teil des Instituts für Schallforschung (ISF) der österreichischen Akademie der Wissenschaften. Er hat 2003–2004 am CNRS Marseille und 2005 an der UCL, Louvain-La-Neuve gearbeitet. Im Jahr 2011 hat er den Start-Preis gewonnen, im selben Jahr schrieb er seine Habilitation mit dem Titel „New Concepts in Frame Theory Motivated by Acoustical Applications“. 2012 wurde er zum Direktor des ISFs ernannt.

Seine Vision ist es, (theoretische) Mathematik mit angewandteren wissenschaftlichen Themen zu verbinden. Er engagiert sich leidenschaftlich für anwendungsorientierte Mathematik für Akustik und Signalverarbeitung. Seine Forschungsgebiete sind die Rahmentheorie, die Operatortheorie und die Zeit-Frequenz-Analyse. Diese Ansätze sind verbunden mit numerischer linearer Algebra, Signalverarbeitung, numerischer Akustik, mathematischer Physik, Psychoakustik, akustischen Messungen und Simulationen, Bioakustik und maschinellem Lernen.

**Christian Kasess**

studierte Elektrotechnik an der Technischen Universität Graz (Diplomthema: Estimation of Time-Variant Multivariate Autoregressive Models using Kalman Filtering). Danach studierte er Physik an der Universität North Carolina und beendete seine Arbeit an einem Modell für Richtungsselektivität bei Sweeps im primären auditiven Kortex.

2009 beendete er seine Dissertation in medizinischer Physik an der medizinischen Universität Wien mit seiner Arbeit über Konnektivitätsmodelle für Neuroimaging-Daten (Dissertation: Investigations into Effective Connectivity using Dynamic Causal Modeling based on Functional MRI). Im September 2010 begann er am Institut für Schallforschung zu arbeiten.

**Wolfgang Kreuzer**

studierte Technische Mathematik an der Technischen Universität in Wien. Nach dem Abschluss des Studiums (Diplom 1996: Modellierung des Cortischen Organs mit Hilfe der finiten Elemente Methode) und eines anschließenden Doktoratsstudiums (2001: Computergestützte Analyse verschiedener Varianten der iterierten Defektkorrektur unter besonderer Berücksichtigung steifer Differentialgleichungen) war Wolfgang Kreuzer bis Ende 2004 wissenschaftlicher Mitarbeiter am Institut für Angewandte und Numerische Mathematik der TU Wien.

Von 2004 bis 2005 war er als wissenschaftlicher Mitarbeiter an einem FWF Projekt beteiligt, das in Kooperation zwischen dem Institut for Distributed and Multimedia Systems der Universität Wien und

dem Institut für Theoretische Chemie durchgeführt wurde. Seit 2004 ist Wolfgang Kreuzer Mitarbeiter am Institut für Schallforschung der Österreichischen Akademie der Wissenschaft.



Thomas Maly

hat Nachrichtentechnik an der TU Wien mit dem Fächerschwerpunkt Hochfrequenztechnik studiert und sein Doktorat 2012 zum Themenschwerpunkt Zuverlässigkeit der Zustandserkennung infrastrukturgebundener Sensorsysteme für die Zugüberwachung abgeschlossen. Seit 2003 ist er an der TU Wien beschäftigt mit dem Forschungsschwerpunkt Multisensorsysteme im Bahnbereich.



Zdeněk Průša

erhielt den Bc. (B.Sc.), Ing. (M.Sc.) und Ph.D. Abschluss an der Technischen Universität Brno, Tschechische Republik, 2006, 2008 bzw. 2012. Von 2012 bis 2018 war er Forscher am Institut für Schallforschung der Österreichischen Akademie der Wissenschaften in Wien. Er war der Betreuer eines Open-Source-Projekts, der Large Time-Frequency Toolbox, und der Gründer der Phase Retrieval

Toolbox. Er forschte auf dem Gebiet der Zeit-Frequenz-Analyse und Signalverarbeitung. Seit 2018 arbeitet er als Embedded Software Developer bei der Lewitt GmbH in Wien und entwickelt Firmware für analoge und digitale Mikrofone und Audio-Interfaces.



Florent Jaillet

erhielt im Jahr 2000 ein Ingenieurdiplom in Elektrotechnik und Informatik. Danach spezialisierte er sich auf Akustik, Signalverarbeitung und Informatik für Musik und erhielt 2001 einen MSc und 2005 einen Dokortitel. Anschließend arbeitete er für verschiedene Unternehmen und akademische Forschungslabore in der Forschung und Entwicklung im Bereich Audiosignalverarbeitung und Akustik. Im Jahr 2010 wechselte er als Ingenieur zum CNRS (Frankreich) und unterstützte die Bereiche Signalverarbeitung, Softwareentwicklung und wissenschaftliches Rechnen. In dieser Position arbeitete er nacheinander in verschiedenen Forschungslabore in den Bereichen Neurowissenschaften und Informatik. Derzeit arbeitet er für das Laboratoire d'Informatique, Signaux et Systèmes de Sophia Antipolis (Frankreich), als Softwareentwickler und Berater für wissenschaftliches Rechnen für künstliche Intelligenz.