

Resynthese von Spektrogrammen mit nicht vorhandener oder unzureichender Phaseninformation

Seminararbeit aus Algorithmen in Akustik und Computermusik 02

Christoph Baumgartner
Sebastian Grill

Betreuung: Dr. Franz Zotter, Dr. Matthias Frank

Graz, 16. Juni 2016



institut für elektronische musik und akustik



Zusammenfassung

Die folgenden Seiten beschäftigen sich mit Algorithmen zur Rekonstruktion von Audiosignalen aus Spektrogrammen mit nicht vorhandener Phase. Dabei wird ausschließlich auf die Arbeit von Griffin und Lim, sowie einem wichtigen Bestandteil der Algorithmen, die STFT (Short Time Fourier Transformation), eingegangen. Es geht dabei nicht nur um die Rekonstruktion von Audiosignalen, sondern auch um die Ermittlung eines „Ursignals“ von Mehrkanalaufnahmen. Um Vergleiche anstellen zu können, werden die Algorithmen zusätzlich bezüglich verschiedener Parameter numerisch validiert.

Inhaltsverzeichnis

1	Einleitung	4
2	STFT	5
3	Griffin und Lim Algorithmus	6
3.1	Algorithmus	6
3.2	Bedingungen für die Rekonstruktion	7
3.3	Fast Griffin-Lim Algorithmus	8
4	Numerische Validierung	10
4.1	Grundlegende Validierung der Implementierung	10
4.2	Quadratischer Fehler	11
4.3	Überlappung und Fensterlänge	11
4.4	Laufzeitunterschiede	13
4.5	FGLA	15
5	Untersuchungen mit Aufnahmen eines 64-Kanal-Mikrofonarrays	16
6	Schlussfolgerung und Ausblick	20

1 Einleitung

Diese Arbeit befasst sich damit, wie Spektrogramme mit teilweise zerstörter bzw. überhaupt nicht vorhandener Phaseninformation möglichst schnell rekonstruiert werden können.

Dies wird benötigt, wenn z.B. nur das Amplitudenspektrum eines Signals gegeben ist. Um den Phasenanteil zu schätzen, wird zur Rekonstruktion der Griffin und Lim - Algorithmus verwendet. Kern des Algorithmus ist die Kurzzeit-Fourier-Transformation (Short Term Fourier Transform - STFT). Die STFT ist ein einfaches und effizientes Werkzeug für das Verarbeiten von Audiosignalen.

Die Motivation hinter dieser Arbeit, ist z.B. das Ermitteln des „Ursignals“ einer Mehrkanalaufnahme bzw. das Schätzen der Phase eines monoinkompatiblen Signales.

2 STFT

STFT bedeutet Short Term Fourier Transform. Sie ist eine Methode der Fourier-Analyse, um die zeitliche Änderung des Frequenzspektrums eines Signals darzustellen. Ein Vorteil der STFT ist, dass sie auch für nichtstationäre Signale geeignet ist, d.h., Signale, deren Frequenzeigenschaften sich im Laufe der Zeit verändern.

Bei der Transformation wird das Zeitsignal in einzelne Zeitabschnitte unterteilt. Dies geschieht mit Hilfe einer Fensterfunktion. Diese Abschnitte werden anschließend in einzelne Spektralbereiche transformiert. Die zeitliche Aneinanderreihung der Spektralbereiche ergibt schließlich die STFT [SD11].

Wenn x ein reelles, endliches, zeitdiskretes Signal ist, ist die Kurzzeit-Fouriertransformation davon definiert als [SD11]:

$$X(\omega, m) = \sum_{k=0}^{N-1} e^{-i2\pi \frac{k\omega}{N}} w(k)x(k + Sm). \quad (1)$$

k ... Abtastzeitpunkt (Signal)

ω ... Frequenzindex (STFT)

m ... Zeitindex (STFT)

w ... Analysefenster

N ... Fensterlänge

S ... Analysezeitschritt in Abtastintervallen

3 Griffin und Lim Algorithmus

Der Algorithmus wurde von Daniel Griffin und Jae Lim bereits 1984 präsentiert. Dabei handelt es sich um ein iteratives Gradientensuchverfahren. Der Algorithmus dient dazu, aus Kurzzeit-Spektrogrammen das zugehörige Zeitsignal abzuschätzen, wenn die Phaseninformation verloren oder zerstört ist. D.h., es ist möglich, mit Hilfe einer STFT das Zeitsignal abzuschätzen. Der Schlüssel zur Rekonstruktion ist die Redundanz der Information bei der STFT [Sü11].

Bei einem reellen Signal erhält man $\frac{N}{2} + 1$ komplexe Koeffizienten aus einem Fenster mit der Länge N . Mit der STFT bekommt man also immer eine Überlappung zwischen angrenzenden Fenstern zusammen.

Im Falle einer minimalen Überlappung von 50%, erhält man bei einem realen Signal mit der Länge N entsprechend $2N$ reelle Koeffizienten. Bei einer größeren Überlappung, also $<50\%$, wird die Redundanz der Information immer größer [SD11].

3.1 Algorithmus

Um das geschätzte Signal zu erhalten, kann folgende Formel aus dem quadratischen Abstand zwischen dem Originalspektrum $X(mS, \omega)$ und dem modifizierten Spektrum $Y(mS, \omega)$ hergeleitet werden [SD11]:

$$x(n) = \frac{\sum_{m=-\infty}^{\infty} w(mS - n)y_{\omega}(mS, n)}{\sum_{m=-\infty}^{\infty} w^2(mS - n)} \quad (2)$$

S entspricht dabei der Verschiebung des Fensters und m der fortlaufenden Zahl der verschobenen Fenster. Mit Gleichung 2 kann dann ein iterativer Algorithmus erstellt werden. Dieser vermindert den quadratischen Abstand zwischen den Spektren. Der quadratische Abstand wird auch noch *MSE* (*Mean Square Error*) genannt und ist definiert als [Sü11]:

$$D_M[x(n)|Y_w(mS, \omega)] = \sum_{m=-\infty}^{\infty} \frac{1}{2\pi} \int [|X_w(mS, \omega)| - |Y_w(mS, \omega)|]^2 d\omega. \quad (3)$$

In Abbildung 5 sieht man ein Beispiel für den *MSE* aus Glg. 3. Der quadratische Abstand ist in Abhängigkeit der Anzahl der Iterationen aufgetragen. In Abbildung 1 ist das iterative Verfahren des *G&L – Algorithmus* dargestellt.

Es ist zu beachten, dass bei jeder Iteration das gesamte Zeitsignal berechnet und die STFT durchgeführt werden muss. Es ist also ein erheblicher Rechenaufwand notwendig. Man erhält eine perfekte Rekonstruktion, wenn das ursprüngliche Signal x gleich dem geschätzten Signal \hat{x}

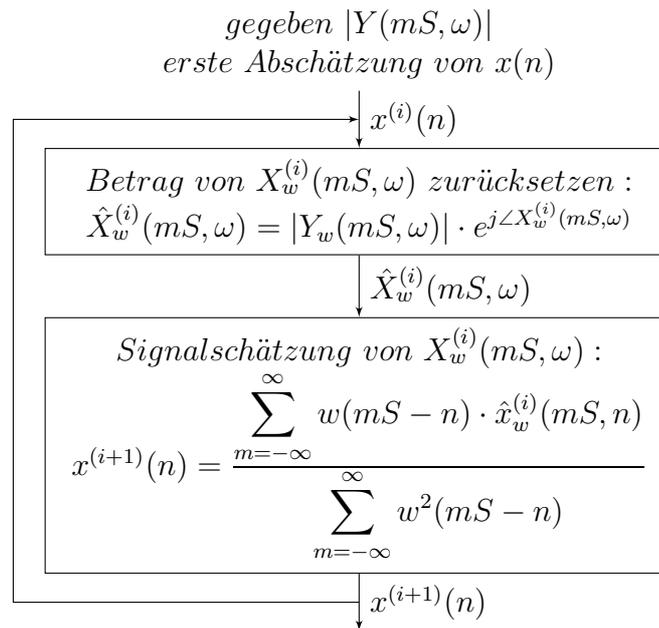


Abbildung 1 – Flussdiagramm des Algorithmus von Griffin und Lim [Sü11] [GL84]

$$\hat{x}(k) = \sum_m s(k - mS) \sum_{\omega} \hat{X}(\omega, m) e^{i2\pi\omega \frac{k-mS}{N}}. \quad (4)$$

- k ... Abtastzeitpunkt (Zeitsignal)
- ω ... Frequenzindex (STFT)
- m ... Zeitindex (STFT)
- s ... Synthesefenster
- N ... Fensterlänge
- S ... Analysezeitschritt in Abtastintervallen

Jedoch gibt es einige Bedingungen für das Abschätzen der Phase bei der Rekonstruktion des Signals.

3.2 Bedingungen für die Rekonstruktion

Um die Phase richtig abschätzen zu können muss zuerst die Fensterfunktion gegeben sein. Um genügend redundante Information zur Verfügung zu haben, muss die Überlappung mindestens 50% betragen.

Um fehlender Information vorzubeugen, darf es keine Nullkoeffizienten der Fensterfunktion geben.

Weiters muss das Signal ein beidseitiges Spektrum besitzen, sprich das Zeitsignal muss kausal sein [SD11].

3.3 Fast Griffin-Lim Algorithmus

Wie bereits angeführt, verursacht die Tatsache, dass beim G&L-Algorithmus in jeder Iteration eine STFT sowie ISTFT des gesamten zu untersuchenden Zeitsignals durchgeführt werden muss, erheblichen Rechenaufwand. Hinzu kommt noch, dass eine große Anzahl von Iterationen notwendig ist, um ein Ergebnis zu erhalten, das perzeptiv dem Originalsignal nahe kommt und keine oder zumindest nur schwer wahrnehmbare Artefakte aufweist. Ein möglicher Ansatz, um die langsame Konvergenz zu verbessern, wird in [PBS13] in Form des Fast Griffin-Lim Algorithmus (FGLA) vorgestellt.

Die dem FGLA zugrundeliegende Idee ist, die Schrittweite des Gradientensuchverfahrens in Abhängigkeit des letzten Schätzfehlers zu vergrößern. Um dies zu bewerkstelligen, wird zu dem in der aktuellen Iteration geschätzten Spektrogramm die mit einem Faktor α_n gewichtete Differenz zwischen dem aktuell geschätzten Spektrogramm sowie der Schätzung der letzten Iteration addiert (siehe Gleichung 5 bzw. Abbildung 2)

$$\tilde{X}_w^{(i)}(mS, \omega) = X_w^{(i)}(mS, \omega) + \alpha(X_w^{(i)}(mS, \omega) - X_w^{(i-1)}(mS, \omega)). \quad (5)$$

Unabhängig von der Nummer der aktuellen Iteration wurde in [PBS13] $\alpha = 0.99$ als optimal angegeben. Auf Basis einer grundlegenden Untersuchung der Abhängigkeit der Konvergenz von α in [PBS13] ist es fraglich, ob eine Adaption dieses Parameters zu einer maßgeblichen Verbesserung führen würde (der Mechanismus, wie dies von den Autoren angedacht ist, bleibt außerdem in dem Artikel unbeantwortet).

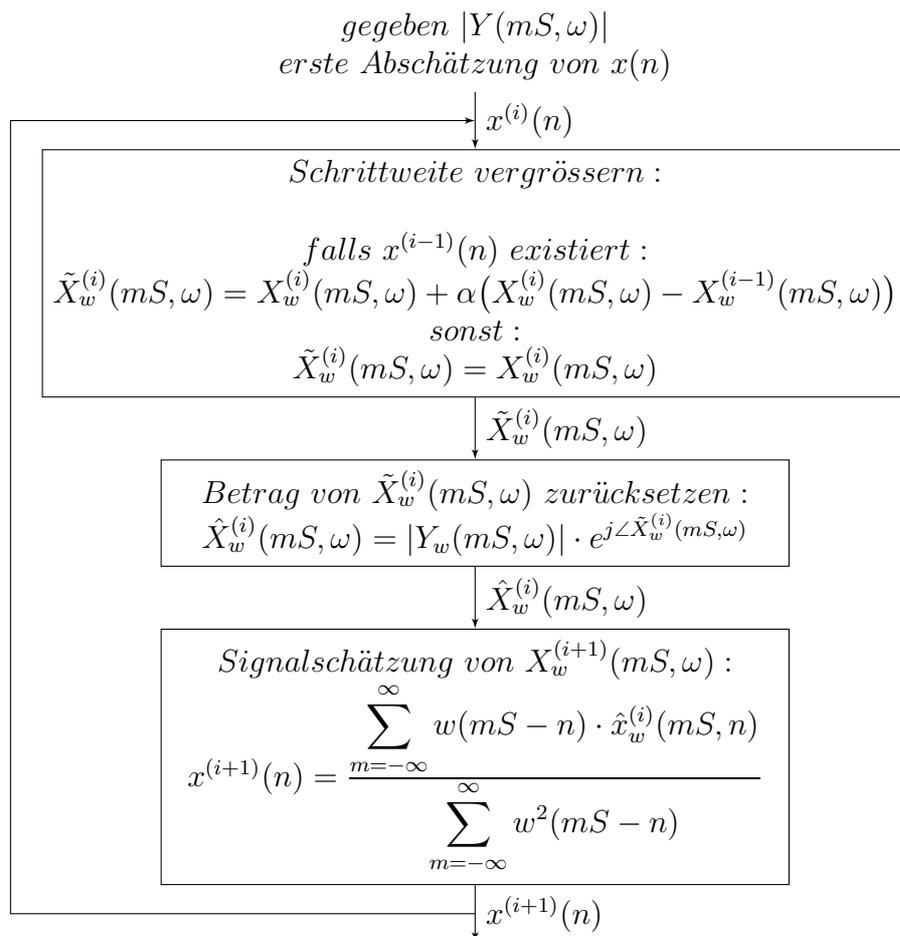


Abbildung 2 – Flussdiagramm des FGLA (abgeleitet aus [PBS13])

4 Numerische Validierung

Nach abgeschlossener Implementierung haben wir eine Reihe von numerischen Experimenten durchgeführt. Diese sollen dazu dienen, den Einfluss verschiedener Parameter auf die Geschwindigkeit der Konvergenz und die benötigte Zeit für die Ausführung der Berechnung zu untersuchen.

Zusätzlich zur numerischen Untersuchung haben wir entwicklungsbegleitend perzeptive Vergleiche angestellt, um die Funktion der Algorithmen subjektiv sicherzustellen.

4.1 Grundlegende Validierung der Implementierung

Eine vergleichsweise einfache Überprüfung der korrekten Funktion des Algorithmus besteht darin, dem G&L-Algorithmus ein völlig intaktes Signal als Ausgangspunkt zu übergeben.

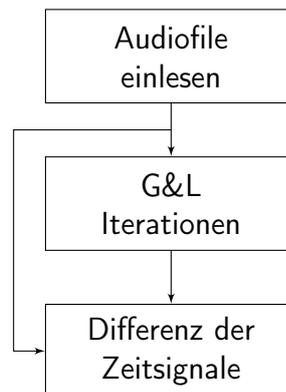


Abbildung 3 – Blockschaltbild des Validierungsversuchs

Da die Phase eines intakten Signals dazu führt, dass die Schätzung des Amplitudenspektrums optimal im Sinne des quadratischen Fehlers ist, sollte bei korrekter Implementierung das Ausgangssignal gleich dem Eingangssignal sein¹.

Als Datenmaterial für alle Versuche in Kapitel 4 haben wir einen kurzen Auszug des Intros von Steely Dan's „Do It Again“ verwendet. Die Parameter wurden, sofern nicht anders angegeben, auf die in der Literatur als optimal angegebenen Werte gesetzt, im Detail bedeutet dies 50% Überlappung bei einer Fensterlänge von 30ms der STFT. In den Versuchen des FGLA wurde α auf 0.99 gesetzt

Das Ergebnis ist in Abbildung 4 ersichtlich. Der Fehler liegt innerhalb eines akzeptablen Ausmaßes, es ist davon auszugehen, dass die Implementierung im Rahmen der Aussagekraft dieses Versuches stimmt. In Verbindung mit den weiteren Versuchen ist die korrekte Funktion damit sichergestellt.

1. Durch die Bearbeitungsschritte im G&L-Algorithmus treten bei endlicher Präzision Rundungsfehler auf, weshalb in der Praxis das Ausgangssignal nicht perfekt dem Eingangssignal gleicht.

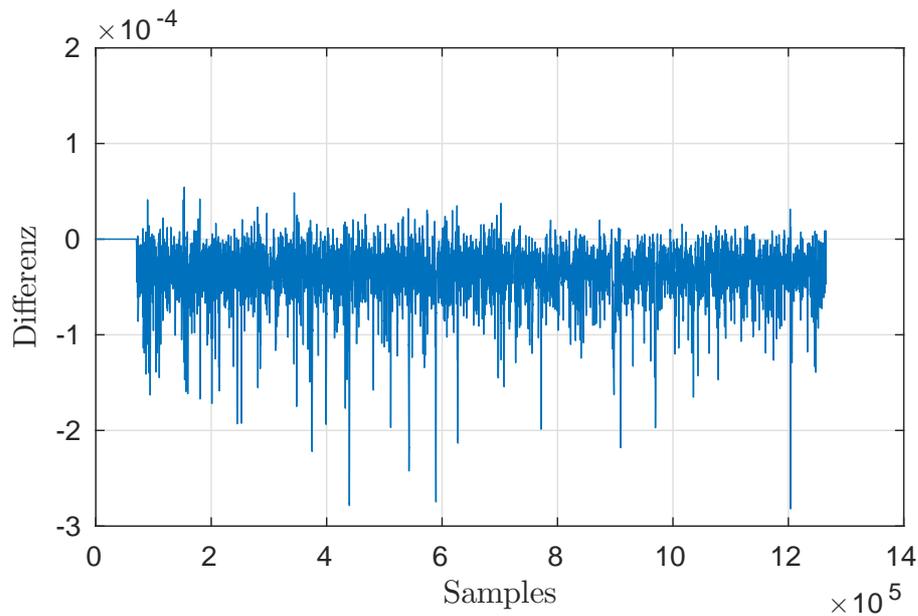


Abbildung 4 – Differenz des ursprünglichen Zeitsignals und des Ausgangs des G&L-Algorithmus nach 10 Iterationen

4.2 Quadratischer Fehler

Als wichtigster Parameter, um die Geschwindigkeit der Konvergenz sowie den residualen Fehler einer Algorithmusiteration zu beurteilen, dient der mittlere quadratische Abstand (MSE - Mean Squared Error) zwischen dem gewünschten und dem aktuell geschätzten Amplitudenspektrum. Mit n als Frequenzindex, m als Zeilenindex, Y als gewünschtes Betragsspektrum und $\hat{X}^{(i)}$ als Schätzer der aktuellen Iteration ergibt sich der Fehler zu:

$$d(Y, X^{(i)}) = \frac{1}{N + M} \sum_{n=1}^N \sum_{m=1}^M \left| |Y(n, m)| - |\hat{X}^{(i)}(n, m)| \right|^2 \quad (6)$$

Abbildung 5 zeigt den Verlauf des MSE über 1000 Iterationen des G&L Algorithmus. Für diesen Versuch haben wir die Ausgangsphase gleich Null gesetzt.

Nach 1000 Iteration beträgt der residuale MSE des geschätzten Amplitudenspektrums $2.5263 \cdot 10^{-9}$.

4.3 Überlappung und Fensterlänge

Wie schon erwähnt, hängt die Performance des Algorithmus stark von der Wahl der Fensterlänge und der Überlappung ab.

Wir haben den Parameterraum in Form eines Rasters abgesucht und für jede Parameterkombination 100 Iterationen des Algorithmus ausführen lassen.

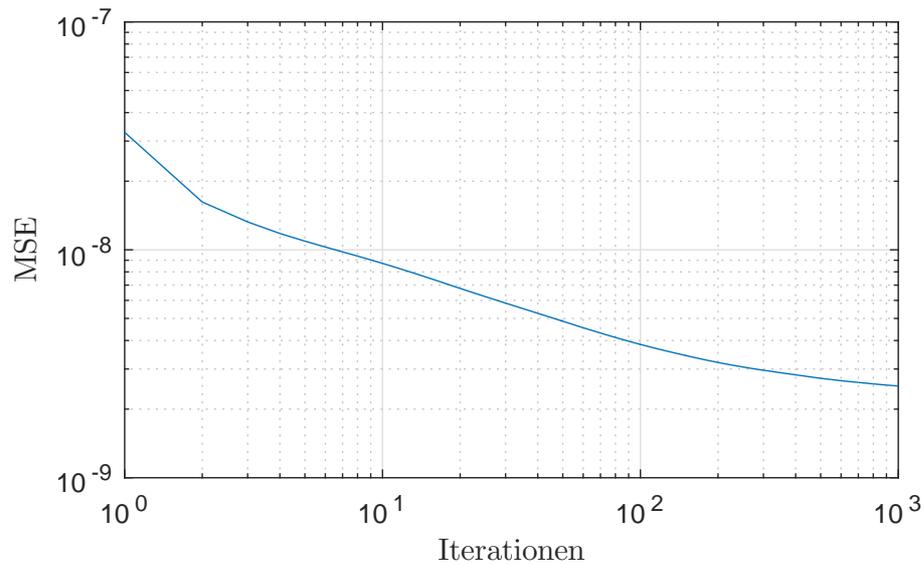


Abbildung 5 – Verlauf des MSE des geschätzten Amplitudenspektrums über 1000 Iterationen

Für den Raster wurde bei der Fensterlänge der Bereich von 16 bis 48ms in 4ms Schritten aufgelöst. Für jede Fensterlänge wurde der Bereich von 50 bis 75% Überlappung in fünf Schritten untersucht. Es ergibt sich damit eine Gesamtanzahl von 45 Rasterpunkten.

Der Versuchsaufbau ist in Abbildung 6 als Blockschaltbild ersichtlich. Das Ergebnis des Versuchs findet sich in Abbildung 7.

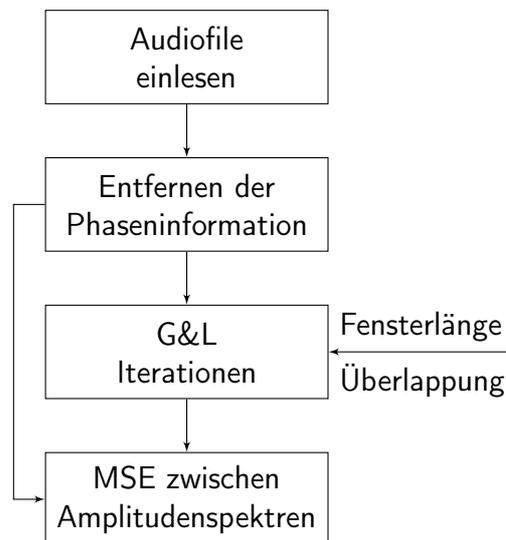


Abbildung 6 – Blockschaltbild der Untersuchungen zu Überlappung und Fensterlänge

Je größer die Überlappung, desto kleiner ist der Residualfehler nach einer bestimmten Anzahl von Iterationen, da die Redundanz der Information des Amplitudensignals erhöht

wird. Ein Nachteil dabei ist jedoch die gleichzeitige Erhöhung der Rechenzeit.

Ein weiterer wichtiger Punkt ist die Wahl der Fensterlänge.

Es hat sich erwartungsgemäß herausgestellt, dass mit größerer Fensterlänge auch der Residualfehler steigt. Dies kann dadurch erklärt werden, dass die zeitliche Auflösung bei steigender Fensterlänge schlechter wird, wodurch es für den G&L-Algorithmus schwieriger wird, die Phase aus der Fortschreitung der Blöcke zu schätzen.

Abbildung 7 zeigt den Residualfehler in Abhängigkeit der besprochenen Parameter nach einer Anzahl von 100 Iterationen. Die Überlappung ist dabei in Prozent bezüglich der Fensterlänge angegeben. Der Residualfehler wird laut Gleichung 6 ermittelt.

Wir haben das Raster nicht danach ausgelegt, die Werte für R und N immer auf Potenzen von 2 zu legen, was für den nichttrivialen Zusammenhang des Fehlers und der Fensterlänge verantwortlich sein dürfte.

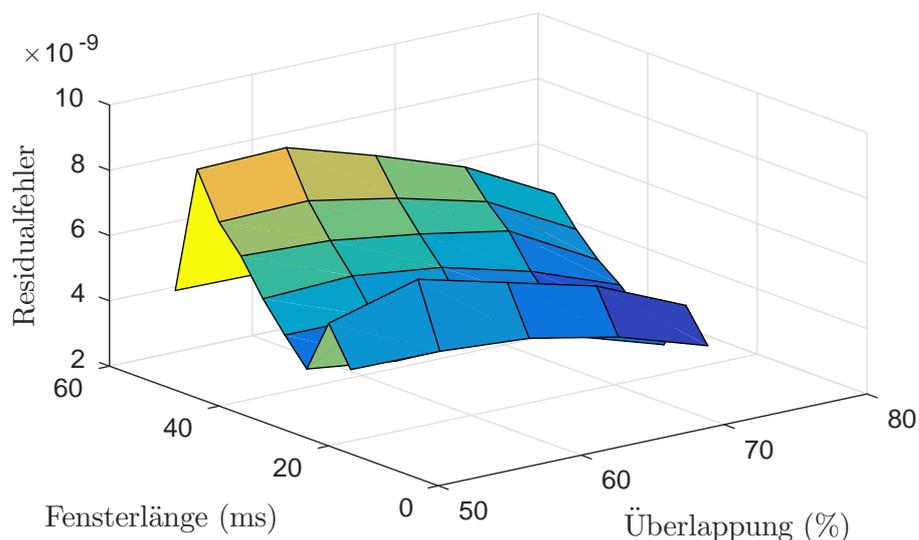


Abbildung 7 – Residualer MSE nach 100 Iterationen in Abhängigkeit von Überlappung und Fensterlänge

4.4 Laufzeitunterschiede

Ein wichtiger Punkt, um den verwendeten Algorithmus zu beurteilen, ist die Laufzeit bei einer gewissen Anzahl an Iterationen. Diese erhöht sich bei größer werdender Überlappung, da sich die Anzahl der zu berechnenden Blöcke erhöht.

Die Laufzeit wurde für verschiedene Parameter bei 100 Iterationen gemessen. Für diesen Versuch wurde der gleiche Raster wie für die Untersuchung der Abhängigkeit des Residualfehlers von Überlappung und Fensterlänge herangezogen.

In Abbildung 9 sieht man, dass die Überlappung einen großen Einfluss auf die Rechenzeit hat. Bei 70% Überlappung erhöht sich die Laufzeit um den Faktor 2 bezogen auf eine

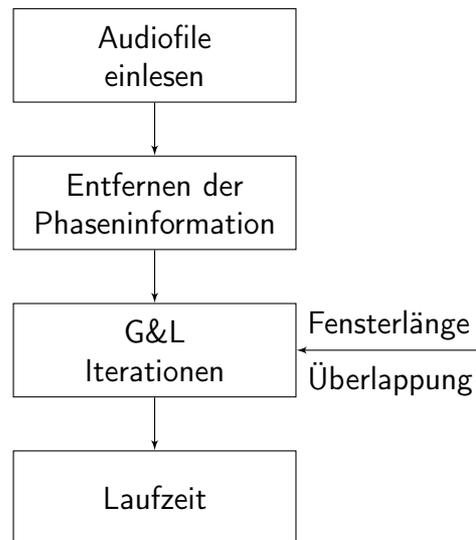


Abbildung 8 – Blockschaltbild der Untersuchungen zu Laufzeit

Überlappung von 50%.

Auch die Fensterlänge hat einen Einfluss, jedoch ist dieser eher klein gegenüber dem der Überlappung. Die Sprünge im Diagramm mit Bezug auf die Fensterlänge könnten dadurch entstehen, dass die Berechnung der FFT mit Fenstern, deren Länge keine Potenz von 2 darstellt, aufwändiger und ungenauer ist.

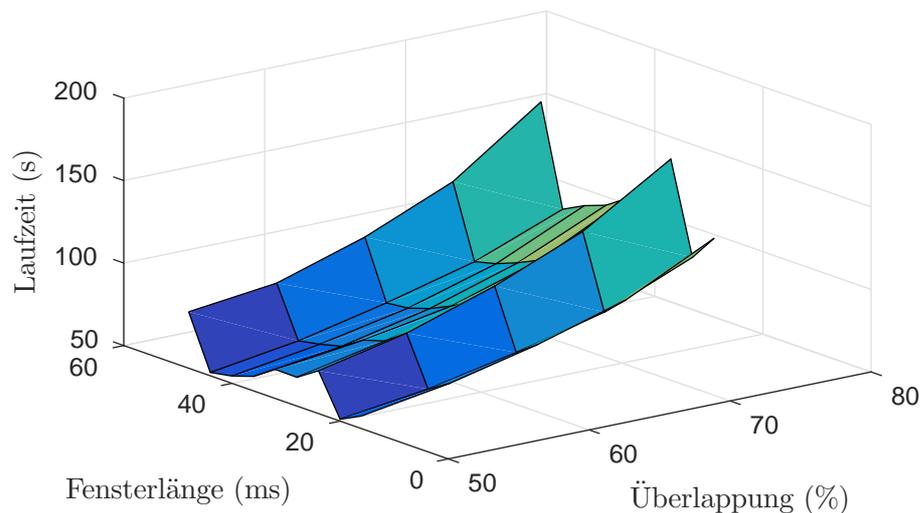


Abbildung 9 – Laufzeit der Berechnung für 100 Iterationen in Abhängigkeit von Überlappung und Fensterlänge

4.5 FGLA

Wie bereits in Kapitel 4.5 beschrieben, wird beim FGLA die Schrittweite adaptiv vergrößert, was dazu führt, dass das gesamte Verfahren schneller konvergieren sollte. Vergleicht man die Abbildungen 10 und 5, so zeigt sich, dass der FGLA zwischen 5 und 30 Iterationen schneller konvergiert, jedoch beginnt auch hier die Konvergenz zu stagnieren. Der Residualfehler dieses Versuchs betrug nach 1000 Iterationen $2.3989 \cdot 10^{-9}$ (vgl. dazu $2.5263 \cdot 10^{-9}$ des herkömmlichen G&L-Algorithmus). Der identische Verlauf der Kurven zwischen erster und zweiter Iteration lässt sich daraus erklären, dass der FGLA bei der ersten Iteration noch nicht über den Fehler der vorhergehenden Iteration verfügt und deshalb in der ersten Iteration dem GLA gleicht.

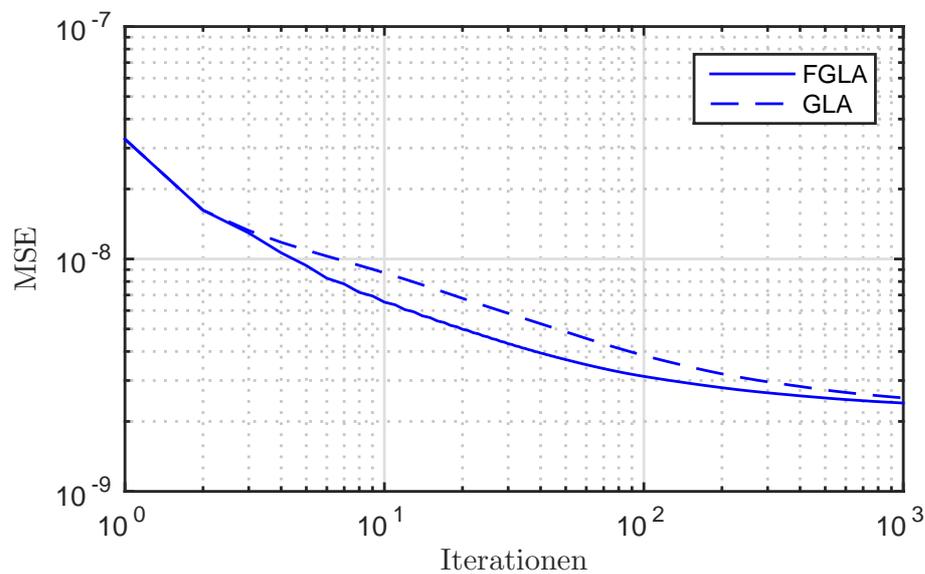


Abbildung 10 – Konvergenzverhalten des FGLA nach 1000 Iterationen

An dieser Stelle sei weiters erwähnt, dass auf Basis einer kurzen perceptiven Beurteilung des resynthetisierten Signals beim FGLA ab etwa 60 Iterationen keine hörbaren Artefakte mehr auftraten, während der GLA circa 120 Iterationen benötigte. Bei fehlender „sinnvoller“ Phaseninformation bietet der FGLA (zumindest im konkreten Fall) einen signifikanten Vorteil in der Berechnungskomplexität.

5 Untersuchungen mit Aufnahmen eines 64-Kanal-Mikrofonarrays

Mehrkanalaufnahmen sind im Normalfall nicht durch einfache Mittelung in ein Monosignal überführbar, da die unterschiedliche Phasenlage zu unerwünschten Auslöschungen bzw. Überhöhungen im Spektrum führt. Hinzu kommt noch, dass wenn die Aufnahmesituation nicht absolut statisch ist (z.B. Musiker bewegt sich), ist auch die Phasenlage der Kanäle variabel. Es kann jedoch gewünscht sein, aus einer solchen Mehrkanalaufnahme ein „Ursignal“ zu gewinnen, sofern die akustischen Laufzeiten zwischen den Kanälen nicht zu große Unterschiede aufweisen.

Verglichen mit anderen Lösungen wie adaptiven Filtern zur variablen Phasenschiebung der einzelnen Kanäle bietet der Algorithmus von Griffin und Lim eine einfache Möglichkeit, die Information der einzelnen Kanäle „konstruktiv“ zu überlagern.

Um dies zu testen haben wir eine 64-Kanal-Aufnahme verwendet, die mit dem vom IEM entwickelten kugelförmigen Mikrofonarray aufgenommen wurde. Die einzelnen Kanäle wurden eingelesen, wiederum wurde die Phaseninformation verworfen und die Betragspektren aller 64 Kanäle wurden gemittelt (Abbildung 11).

Abbildung 12 zeigt den Verlauf des MSE beim vorliegenden Versuch. Der Verlauf des Fehlers ist ähnlich wie in Abbildung 5.

Weiters haben wir den obigen Versuch wiederholt, jedoch wurde dem G&L-Algorithmus die Phase des ersten Kanals als Ausgangsphase übergeben (Abbildung 13). Man sieht in Abbildung 14, dass sich dies sehr stark positiv auf das Konvergenzverhalten und den Residualfehler auswirkt. Der Residualfehler nach 100 Iterationen ist bereits deutlich niedriger als jener nach 1000 Iterationen beim ersten Versuch.

Analog dazu haben wir beide Versuche mit dem FGLA wiederholt. Die Ergebnisse sind in den Abbildungen 15 und 16 ersichtlich. Tabelle 1 enthält zusätzlich eine Übersicht mit numerischen Ergebnissen.

Die Ergebnisse zeigen, dass vor allem bei einer geringen Anzahl an Iterationen das Vorhandensein einer „sinnvollen“ Ausgangsphase als Randbedingung eine signifikant positive Auswirkung auf den Residualfehler hat. Der Wechsel zwischen GL bzw. FGLA hat dagegen vergleichsweise geringe Auswirkungen.

Algorithmus		Residualfehler ohne Ausgangsphase	Residualfehler mit Ausgangsphase
GL	100 Iterationen	$3.568 \cdot 10^{-9}$	$2.3443 \cdot 10^{-9}$
	1000 Iterationen	$2.6638 \cdot 10^{-9}$	$2.2506 \cdot 10^{-9}$
FGLA	100 Iterationen	$3.0334 \cdot 10^{-9}$	$2.2946 \cdot 10^{-9}$
	1000 Iterationen	$2.6031 \cdot 10^{-9}$	$2.2385 \cdot 10^{-9}$

Tabelle 1 – Numerische Ergebnisse des Versuches für den GL sowie den FGLA Algorithmus bei unterschiedlich vielen Iterationen

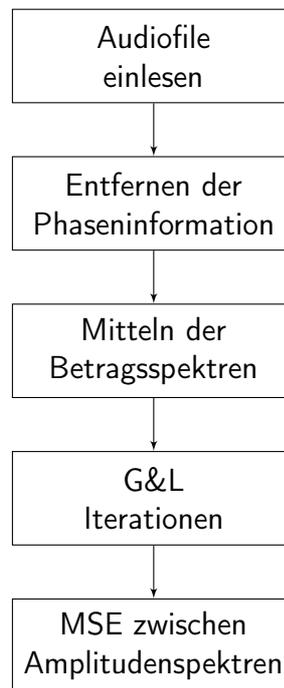


Abbildung 11 – Blockschaltbild der Untersuchungen zu Mehrkanal-Aufnahmen

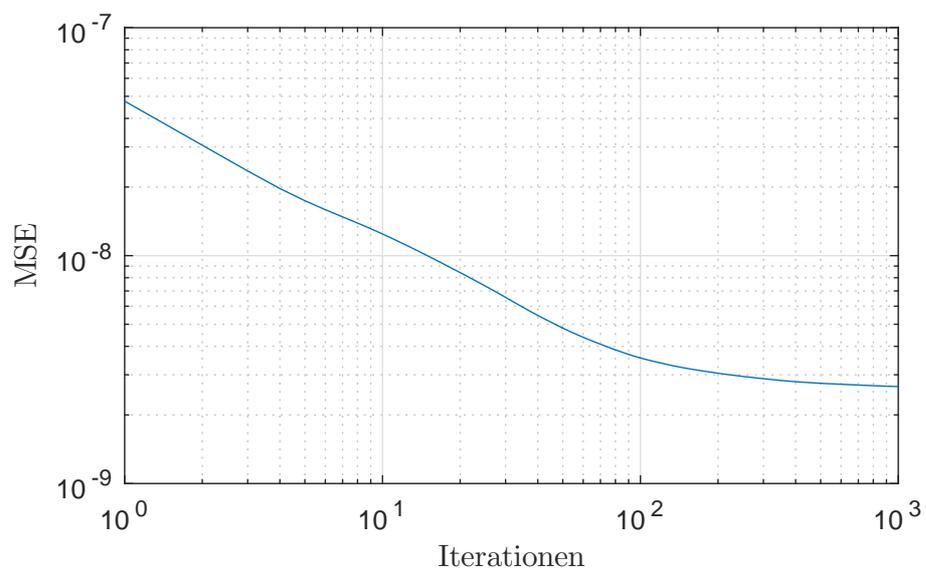


Abbildung 12 – Konvergenzverhalten über 1000 Iterationen

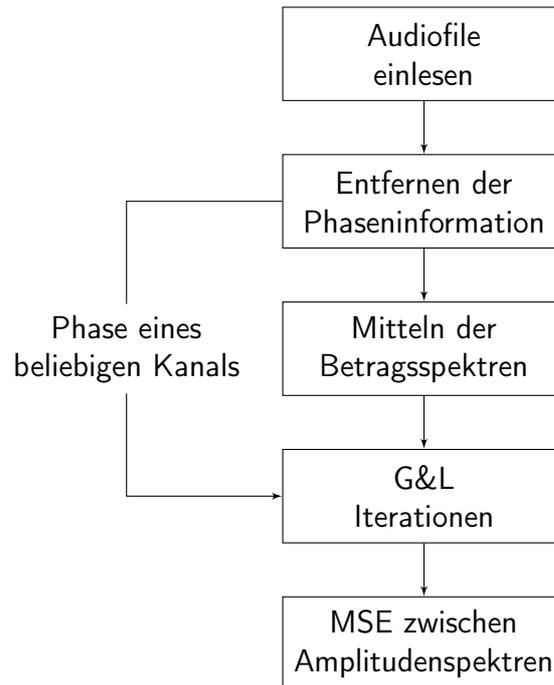


Abbildung 13 – Blockschaltbild der Untersuchungen zu Mehrkanal-Aufnahmen bei vorhandener Ausgangsphase als Randbedingung

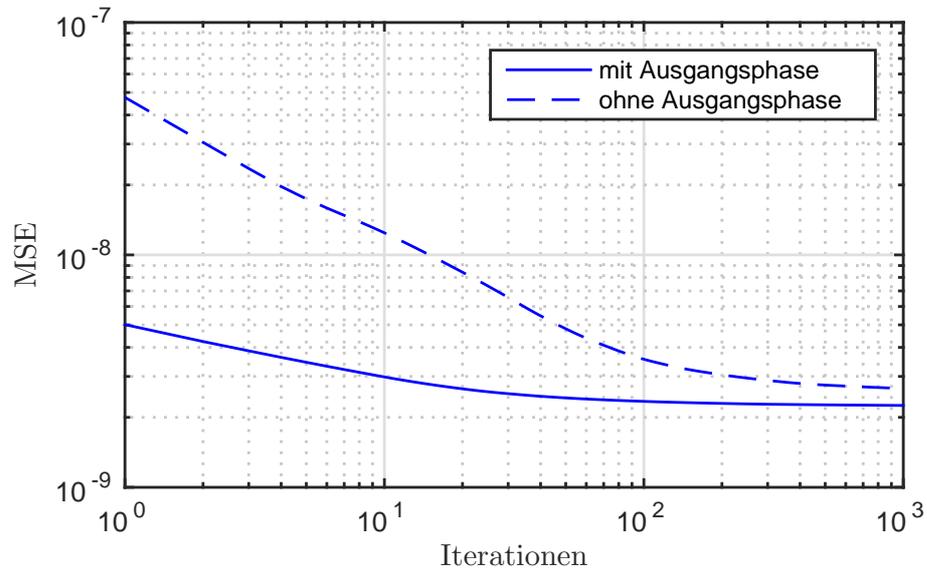


Abbildung 14 – Konvergenzverhalten über 1000 Iterationen bei vorhandener Ausgangsphase

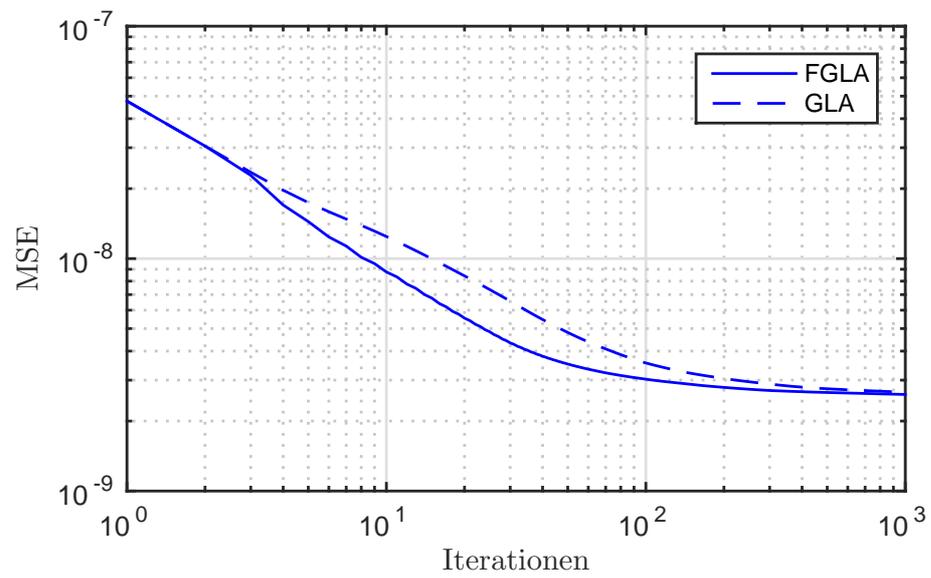


Abbildung 15 – Konvergenzverhalten des FGLA über 1000 Iterationen

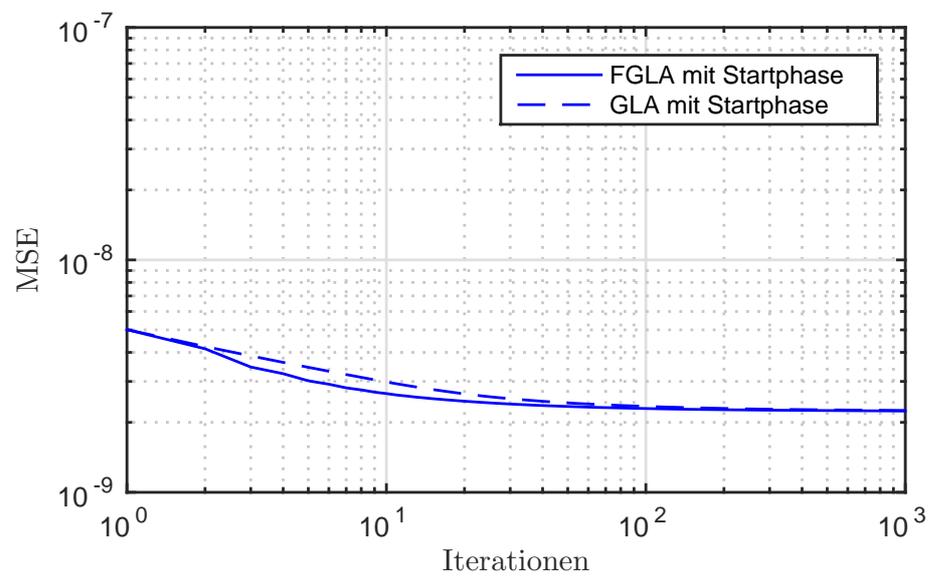


Abbildung 16 – Konvergenzverhalten des FGLA über 1000 Iterationen bei vorhandener Ausgangsphase

6 Schlussfolgerung und Ausblick

In dieser Arbeit konzentrieren wir uns auf den Algorithmus von Griffin und Lim, um eine möglichst gute Rekonstruktion von „Ursignalen“ zu erreichen.

Wir haben den klassischen Griffin&Lim-Algorithmus hinsichtlich Rechenaufwand und Konvergenzzeit untersucht. Dazu haben wir in einer Rastersuche das Analysezeitfenster sowie die Analysefensterlänge der STFT variiert und die Konvergenzzeit sowie den Rechenaufwand protokolliert. Zusammengefasst lässt sich aus diesen Versuchen sagen, dass die Überlappung die Konvergenz verbessert, jedoch den Rechenaufwand deutlich erhöht. Auf Basis dieser Untersuchungen ist es durchaus plausibel, dass der in der Literatur als optimal angegebene Wert von 50% einen guten Kompromiss darstellt.

Hinsichtlich der Analysefensterlänge sollten unbedingt nur Zweierpotenzen zum Einsatz kommen, eine Abweichung davon wirkt sich stark negativ auf den Rechenaufwand aus.

Weiters haben wir Untersuchungen bezüglich des Konvergenzverhaltens des Fast Griffin&Lim-Algorithmus durchgeführt. Im direkten Vergleich mit dem GLA haben wir beobachtet, dass der FGLA ab der zweiten Iteration durchwegs bessere Ergebnisse abliefert. Der Vorteil des FGLA beginnt jedoch bei höheren Iterationszahlen wieder zu schrumpfen, da die Stagnation der Konvergenz früher einzusetzen scheint.

Auch ist der Vorteil von den Ausgangsbedingungen abhängig. Ist Vorab gute (jedoch nicht perfekte) Phaseninformation vorhanden fällt der Vorteil des FGLA sehr gering aus, da die Stagnation sehr bald einsetzt.

Die beiden verwendeten Algorithmen erreichen perzeptiv sehr gute Ergebnisse, der G&L-Algorithmus, sowie auch der schnelle Algorithmus (FGLA). Der G&L-Algorithmus besitzt jedoch im Vergleich zum schnellen Algorithmus eine erheblich langsamere Konvergenz und damit auch einen deutlich höheren Rechenaufwand. Unsere Untersuchung ergab, dass der FGLA im konkreten Fall nur 50% der Iterationen benötigte, bis keine Artefakte mehr im Signal zu hören waren.

Als Ausblick wäre es in weiterer Folge interessant, dieses Ergebnis im Rahmen eines größer angelegten psychoakustischen Versuches mit Versuchspersonen sowie einer größeren Breite an Audiomaterial zu untermauern.

Literatur

- [GL84] D. W. Griffin and J. S. Lim, "Signal estimation from modified short time fourier transform," in *IEEE Transaction on Acoustics, Speech, and Signal Processing*, Vol. ASSP-32, No. 2, 1984.
- [PBS13] N. Perraudin, P. Balazs, and P. L. Søndergaard, "A fast griffin-lim algorithm," in *2013 IEEE Workshop on Application of Signal Processing to Audio and Acoustics*, 2013.
- [SD11] N. Sturmel and L. Daudet, "Signal Reconstruction From STFT Magnitude: A State Of The Art," in *Proc. of the 14th Conference on Digital Audio Effects*, 2011.
- [Sü11] S. Süß, "Auffinden von Ursignalen aus Aufnahmen umgebender kugelförmiger Mikrofonanordnungen," 2011.