Fast Fourier Transform (FFT) in Pure Data

Alle komplexen PERIODISCHEN Wellenformen können mit einer Auswahl von harmonisch in Beziehung stehender Sinusschwingungen modelliert werden.

Die FFT ist selbst als ein virtuelles Signal aus vielen periodisch, harmonischen Sinusschwingungen zu sehen, mit denen das Eingangssignal verglichen wird.
Diese Form einer Mustererkennung wird als Multiplikationsprozess durchgeführt.

Wenn 2 Frequenzen ringmoduliert werden, entstehen 2 zwei neue Frequenzen, eine als Summe beider Frequenzen, eine als Differenz der beiden. Wenn beide Frequenzen identisch sind, entsteht lediglich eine unipolare Schwingung mit der doppelten Frequenz.
400 Hz + 401 Hr = 801 Hz + 1 Hz
400 Hz + 400 Hz = 800 Hz (+ 0 Hz), unipolar

Dieser DC-Offset ist direkt proportional zu den Amplituden der Inputsignale. Somit kann mit einem konstanten Referenzsignal die Amplitude eines Eingangssignals gemessen werden.

Das wird rechnerisch vollzogen indem man das Inputsignal durch eine Tiefpassfilterung auf eine Durchschnittsamplitude setzt, die einzelnen Samplepunkte addiert und das Resultat durch die Anzahl der Punkte teilt, der Bereich oberhalb der positiven Hälfte der Welle, wird zu denen der negativen hinzugefügt.
??? RECHNUNG UNKLAR

Jeder Wert > 0 würde somit die Anwesenheit dieser Frequenz mit dieser Amplitude nachweisen.

Die FFT verhält sich wie eine Bank gestimmter analoger Filter. Je exakter die Messung sein soll, desto mehr Filter werden benötigt.
Bei einer Bandbreite von 20 Hz würden zwischen 20 - 20000 Herz 1000 Filter benötigt.
Im Bereich von 440 Herz entspräche 20 Hz ungefähr einem Halbton, wäre aber bei 4000 Hz entsprechend einem Zehntelton deutlich überpräzise d.h. eine lineare Anordnung führt zu einer frequenzabhängigen Genauigkeit der Analyse, da sie die logarithmischer Struktur harmonischer Verhältnisse nicht berücksichtigt. Insofern kann die FFT lediglich die dominanten Tonhöhenregionen, die spektrale Hüllkurve eines Klangs ermitteln.
!!! SIEHE WAVELETS ALS BESSERE LÖSUNG

Wenn eine Frequenz nicht exakt der Zentralfrequenz eines Filters entspricht, werden diese von verschiedenen überlappenden Filtern registriert. Dieses nennt dich SPEKTRAL LEAKAGE.
Darüber hinaus wird dadurch die Gesamtenergie der FFT reduziert, da ein Teil der Energie des Inputsignals auch in angrenzende Filter fließt. Daher wird das Ergebnis in der Regel entsprechend NORMALISIERT.

Es ist möglich jede aktuelle Amplitude eines Klangs festzustellen, nicht jedoch jedoch ohne Weiteres die Frequenz. Es ist z.B. nicht möglich eine Frequenz bei Klängen zu erkennen, die kürzer als 20msec sind. Je nach Klang und Frequenz kann diese Grenze noch deutlich nach oben steigen (bis zu einer halben Sekunde).
Die Größe des FFT Fensters bestimmt daher die tiefste zu analysierende Grundfrequenz. Diese errechnet sich indem die Samplingrate durch die Fenstergröße geteilt wird, z.B.
22050/256 = 86.1 Hz
44100/1024 = 43.1 Hz

Da es eher die Ausnahme darstellt, dass die Wellenform an den Rändern eines FFT Fensters

Nulldurchgänge hat, werden die Fenster mit einer Hüllkurve (meistens einem HAMMING oder HANNING (?) WINDOW) multipliziert.
(Katja Vetter und Richard Dobson präferieren ein Hamming Fenster.)

Ein größeres Fenster hat also eine bessere Frequenzauflösung aber eine schlechtere Zeitauflösung.
512/44100 = 0.011 sec
1024/44100 = 0,023 sec
Das hat Auswirkungen auf die Qualität der Analyse abhängig davon, wie bewegt das Inputsignal ist.

Um diesem Widerspruch entgegen zu wirken, werden die FFT Fenster überlappt, z.B. ein Fenster von 1024 Samples wandert in Schritten von 64 Samples über das Inputsignal. Dadurch wird die Samplefrequenz verachtfacht von 21.53 Hz auf 172.26 Hz entsprechend von 46 msec auf ungefähr 5 msec.

FFT Fenster sollte die Größe einer Zweierpotenz (256, 512, 1024) haben, das diese recheneffektiver sind. Auch Fenster mit vielen Teilern sind möglich z.B. 384 (2+4+6+8).
Entweder man versucht die Grundfrequenz der FFT der des Inputsingals möglichst nahe zu bringen oder man transponiert das Signal entsprechend.

Die Anzahl der "Kanäle" (d.h. Obertöne) der FFT entspricht der Hälfte des FFT Fensters, entsprechend dem Nyquest-Theorem, das besagt, dass die höchste darstellbare Frequenz mindestens 2 Samplepunkte, eine positiven und einen negativen zur Repräsentation benötigt.

In der FFT werden Wellenformen als komplexe Nummern dargestellt. Diese werden als Kombination aus einem realen (Kosinus) und einem imaginären (Sinus) Signal dargestellt mit denen Amplitude und Phase der Schwingung errechnet werden.
Für einfache Analysen ist die Information über die Phase nicht von größerer Bedeutung, keine Informationen geht dadurch verloren, wenn diese ignoriert wird. Für die exakte Resynthese wird sie aber benötigt damit es keine Diskontinuitäten zwischen den FFT Fenstern gibt.
Bei der Berechnung der realen und imaginären Teile des Signale wird dieses aus der Time Domain in die Frequenz Domain überführt.
In der Time Domain hat der Sample Punkte von 0 - N-1. In der Frequenz Domain entstehen daraus die realen und imaginären Anteile mit N/2+1 Punkten. (Die zwei zusätzlichen Werte sind bei 0 Hz und bei N/2, die beide konstant 0 sind). Frequenz Domain und Time Domain beinhalten exakt dieselben Informationen.
Eine reale FFT von 64 Samples ergibt 33 Werte für Sinus- und 33 Cosinuswellen mit den Frequenzen 0,1,2,…32, 33 als Faktoren der Grundfrequenz z.B. 22050/64 = 344,53 Hz
Grundfrequenz:
f0: 0
f1: 1 x 22050/64 = 344,53
f2: 2 x 22050/64 = 689,06
etc.

Da die Wellen alle ganzzahlig sind, passen sie alle genau in das Analysefenster. Der erste Punkt ist dabei identisch mit dem letzten.

Die Amplitude der Frequenz wird folgendermaßen errechnet: amp=sqrt(re^2+im^2)
Durch die Quadrierung wird erreicht, dass alle Magnituden positiv sind.
Die Phase errechnet sich durch: phs=arctan(im/re)

Bei der FFT wird ein zusätzlicher Kanal bei 0 Hz erzeugt, der für alle Frequenzen verwendet wird, die tiefer als die Grundfrequenz des FFT Fensters sind. Solche Frequenzen würden bei einer

Resynthese zu unvorhersehbaren Ergebnissen führen.

Um ein Aliasing bei einer eventuellen Transposition in der Resynthese zu verhindern, können die hohen Frequenzanteile des Eingangssignals vorher durch eine Tiefpassfilterung reduziert werden. Eine Faustregel dabei ist: Tiefpassfilterung des Originals mit der Grenzfrequenz Nyquistgrenze/Transposition.

Pure Data:

[rfft~] Analyse
[rifft~] Resynthese
[rec2pol~] Amplitude in Phase
[pol2rec~] Phase in Amplitude
[atan2~]
[q8_sqrt~] schnelleres sqrt

In PD ist die Block Size Default auf 64 Samples d.h. wenn nichts anderes angegeben ist, wird die FFT mit Fenstern von 64 Samples berechnet. Diese Blocksize kann nur in einem Subpatch mit dem Objekt [block~] geändert werden. [block~ 1024 4] = 4 FFT Fenster mit 1024 Samples und 3 Überlappungen.
Die Normalisierung des Resultats erfolgt nach der Formel: Überlappungen*Blocksize/2 .
??? WARUM

Literatur: INFORMATIONEN ERGÄNZEN

Richard Dobson, The Operation of the Phase Vocoder. A non-mathematical introduction to the Fast Fourier Transform, 1993

Marius Scheballa, FFT & PD, 01.2012

Johannes Kreidler, Fourier Analysis in: Programming Electronic Music in PD

Frank Barknecht, Beginners Guide to the FFT-objects in PD

Katja Vetter, FFT