

Elektro-akusztikai, szinkronizált kiegyenlítő

Bencsik Gergely

Mérnök Informatikus BSc III.

Pécsi Tudományegyetem

Pollack Mihály Műszaki és Informatikai Kar

Konzulens: Sári Zoltán dr., egyetemi adjunktus

Előszó

Az elektro-akusztika fogalmával egy idős az elektronikus eszközök átviteli karakterisztikájának problémája. Egy zeneszám felvétele, egy koncert, rádióadás vagy televíziós, élő közvetítés alkalmával a hangtechnikus feladata, hogy a rögzített hangokat, egyfelől az eredeti hangjához hűen, érthetően és a felvétel különböző, egyéb hanghatásaitól elkülöníthetően, összefoglalva: jól definiálható jelként továbbítsa, rögzítse.

Számos egyéb probléma mellett, (melyektől a dolgozat eltekint) a rögzítésre, továbbításra és lesugárzásra használt eszközök átviteli karakterisztikája mind torzítja az eredeti hangképet. Jellemzően egy hangtechnikai eszköz árát elsősorban az határozza meg, hogy mennyire „lineáris” átvitelű, azaz milyen mértékben képes továbbítani, feldolgozni a jelet annak eredeti formájában.

Ezen probléma kiküszöbölésére használható egy u.n. equalizer, azaz kiegyenlítő, amellyel nem csak az eredeti hangkép jó megközelítésű formája definiálható, hanem több, egymás mellett szóló jel szeparálható, a hangkép tovább optimalizálható.

A dolgozat egy olyan eszközzel foglalkozik, amely ezen a ponton használható.

Tartalomjegyzék

A hangról.....	3
A felharmonikusokról.....	5
A hang feldolgozásának problémáiról.....	8
Elektro-akusztikai, szinkronizált kiegyenlítő.....	11
A szinkronizált kiegyenlítő elvi háttere.....	12
A prototípusról.....	15
Felhasználási lehetőségek.....	26
Fejlesztési tervezet.....	29
A fejlesztési projekt mérföldkövei.....	30
Irodalomjegyzék.....	31

A hangról

A hang nem más, mint egy tetszőleges közegben terjedő, longitudinális hullámok összessége. Hogy ezek között a hullámok között mit és meddig nevezünk önálló hangnak, az szubjektív értelmezés kérdése. Gondoljunk a „koktélparty-jelenségre”, ami nem más, mint amikor egy nagyobb társaságot összefoglaló eseményen, a különböző, egy időben beszélgető emberek közül, a hozzánk közelebb állók hangjait egyenként felismerjük és akár érthetjük is, annak ellenére, hogy mellette több azonos párbeszédet is hallunk. Hangként definiálhatjuk a helység összes hangját egyben, az emberek beszédhangjait egyenként, egy-egy ember beszédében a magánhangzókat és így tovább.

Fontos megjegyezni, hogy ehhez az elkülönítéshez ismerni kell a fent felsorolt, köthető fogalmakat, így ezek szeparálásának képességét ez a tény rögtön erősen szűkíti és (ha minden egyéb példától eltekintünk) kijelenthetjük, hogy erre csak az ember képes.

A hang hullám, a hullámnak pedig alapvető paramétere a frekvenciája. Minél összetettebb egy hang, annál sűrűbben és annál szélesebb spektrumon változik ez az értéke. Az előző példában említett szeparáció ezekből a kevert frekvenciaértékekből és azok változásából, ritmikájából engedi nekünk embereknek, hogy a köthető foglmaink, felismert mintáink és kognitív sémáink segítségével azonosítsuk a hordozott információt. Ennek részletes pszichológiai és fizikai tárgyalása nem tartozik a dolgozat témájához.¹

Ez az összetettség egészen a fehér zajhoz és a hozzá hasonló, teljes hallható (és további) spektrumot lefedő jelig bonyolíthatja a hangot.

Fontos tény egy hullám esetében, hogy bármilyen rezgés (annak formájától függően) további felharmonikusokat generál. Ezek a felharmonikusok adják az általunk felismerhető hang karakterét. Gondoljunk egy azonos hangmagasságon, egy hangon megszólaló női és férfi énekhangra, egy hegedűre, egy trombitára, vagy egy egyszerű négyszög jelű oszcillátorra. Ezen hangok (ideális esetben) karakterükből fakadóan egyértelműen és konkrétan elkülöníthetők, ez pedig mind a sajátos felhangrendszerüknek köszönhető.

Ezt a felhangrendszert a keltett alaprezgés formája, (a periodikusan ismétlődő jel alakja) illetve a (tekintsünk el a további példáktól) hangszer testének felépítése befolyásolja. Amennyiben

egy akusztikus hangszer testének hangját teljes mértékig le tudnánk választani a keltett alaphangtól, azt tapasztalnánk, hogy az alaphang részben eltérő, de továbbra is felismerhető karakterrel szól. Jó példa erre egy elektromos gitár és egy akusztikus gitár hangjainak összehasonlítása, mivel igen hasonló, bár az említett jelenséggel nem azonos eltérés mutatkozik a két eszköz hangja között.

Ez annak köszönhető, hogy a felhangok szinguláris velejárói egy hullámnak, így a hangszer teste nem tesz mást ezekkel, pusztán egy, a rá jellemző átviteli karakterisztikán viszi át az alaphangot és a keletkezett felhangokat. (Tekintsünk el a hangszer testében létrejövő, további rezgésektől.)

Az említett átviteli karakterisztika bizonyos pontokon (hangmagasság) erősíti, máshol elnyomja a jelet, ezzel torzítva azt. Ugyanez a jelenség egy hang feldolgozható, elektromos jellé való alakítása esetén minden egyes eszköznél megtörténik, amit a művelet folyamán használunk, illetve visszaalakításkor szintén. Mindez befolyásolja a hang karakterét, bármit is definiáljunk hangként az adott jelfolyam egészén belül.

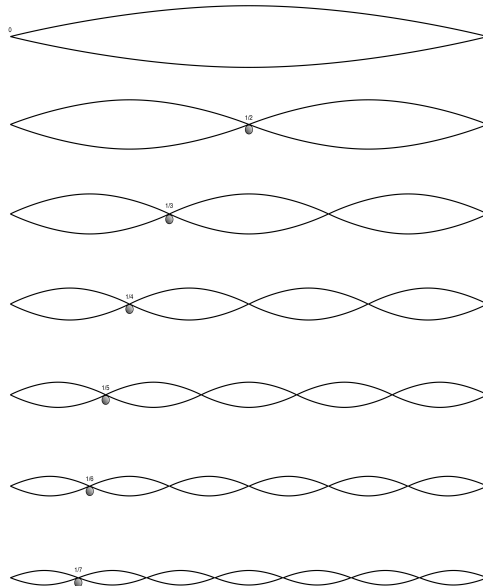
A hétköznapi gyakorlatban így definiálhatjuk a hangot, mint fogalmat, illetve így gondolhatunk rá, ha a felépítését szeretnénk megismerni.

A felharmonikusokról

Fontos szót ejteni a felharmonikusokról, mert felépítésük és működésük alapvető tétele a későbbiekben tárgyalt eszköz működésének. Egy adott hangmagasságú rezgés alapvető eleme egy periodikusan ismétlődő jel. Ennek a jelnek a periódusideje határozza meg a rezgés gyorsaságát, (esetünkben a hangmagasságot) formája pedig a ráépülő felharmonikusok eloszlását, illetve a különböző harmonikus pontokon jelentkező jelerősséget. Jellemzően egy felharmonikus rendszer első tagja az alap körfrekvencia után, annak oktávja, tehát periódusidejének a felénél található. Ezt követi a második oktávja, majd egy kvint és így tovább. Ennek a rendszernek az igazolását Joseph Fourier tézisében találjuk. A Fourier-análízis ezen fizikai rendszer alkalmazása szerint minden periodikus rezgés felírható adott számú szinuszos rezgések összegeként.

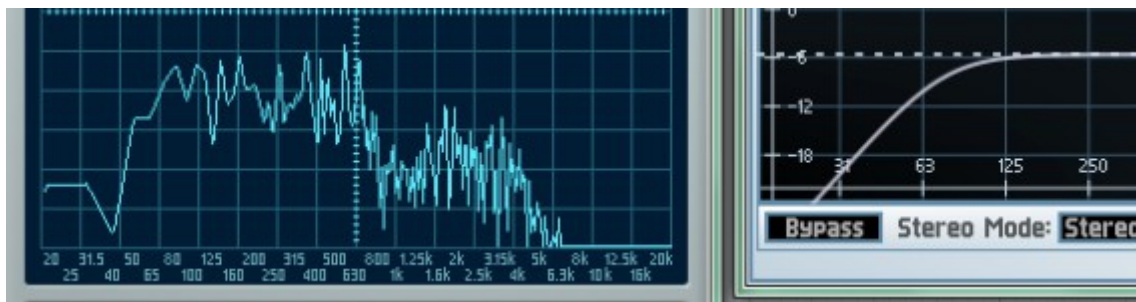
$$y(t) = a_0 + \sum_{n=1}^{\infty} (a_n \cdot \sin(n \cdot \omega_0 \cdot t - \varphi_n))$$

Ezek az úgynevezett felharmonikusok – mint azt korábban tárgyaltuk – határozzák meg a megszólaltatott hang karakterét. A valóságban lehetetlen olyan jelet előállítani, amelyik nem tartalmaz felharmonikusokat, elméletben azonban a szinuszos rezgések ilyenek. Ebből következik, hogy a gyakorlatban jelenlevő rezgések mind rendelkeznek ezekkel az alapvető építőelemekkel, amelyek sorra gyengülnek, ám – fontos megjegyezni – a számuk végtelen. Mi azonban az elhanyagolható mértékűektől, illetve a feldolgozásban részt nem vevő spektrumon jelentkezőktől eltekintünk.

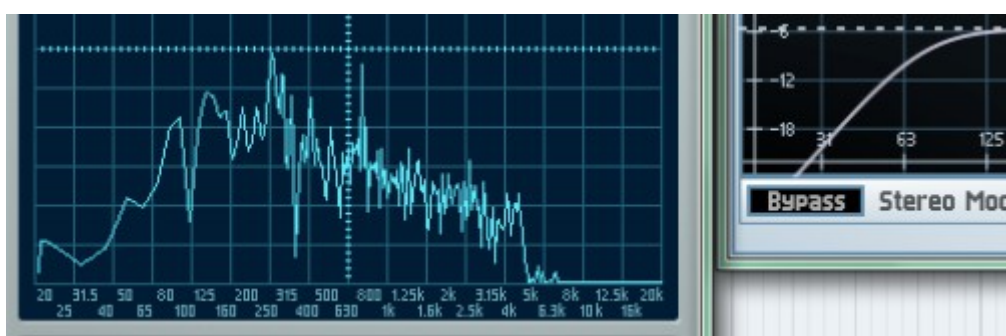


1. A felhangrendszer egyszerű reprezentációja.

Természetesen a Fourier állításában szereplő, tiszta szinuszos rezgések a gyakorlati életben „tisztá” formában nem léteznek, azonban egy Fourier-transzformáció segítségével kimutathatók, a periodikus jel alakjától függő formában jelen vannak. Ha ez nem így lenne, egy, az alap körfrekvenciára helyezett „high-pass”, azaz felüláteresztő szűrővel csillapított hangot el is némíthatnánk. Ha ezt a kísérletet elvégezzük, csak annyit tapasztalunk, hogy a hangkarakter megváltozik, a felharmonikusok átrendeződnek, de továbbra is jelen vannak.



2. Egy jel és felharmonikusai, érintetlenül. (A jobbra látható HPF kikapcsolt állapotban van.)



3. Az előző jel, most bekapcsolt HPF-el való szűrést követően.

A fenti állításból (Fourier-analízis) egyértelműen következik, hogy ideális esetben a felharmonikusok minél magasabb frekvencián szólnak, annál alacsonyabb az amplitúdójuk, ezen kívül egyik felharmonikus amplitúdója sem haladhatja meg az alap körfrekvencián jelentkező amplitúdót. A gyakorlatban azonban a különböző eszközök, effektek átviteli karakterisztikájával is számolnunk kell. Egy akusztikus hangszer esetén a legszembetűnőbb magának a hangszer testének a karaktere, amely sok esetben színezhethet annyit a leadott rezgésen, hogy az előző állítás gyakorlati szempontból értelmét veszti. Példaként gondoljunk csak a fent bemutatott HPF hatására, vagy egy hangszeren megszólaltatott üveghangra.

Ennek a ténynek később, ahol a szinkronizált kiegyenlítő működését tárgyaljuk még fontos szerepe lesz.

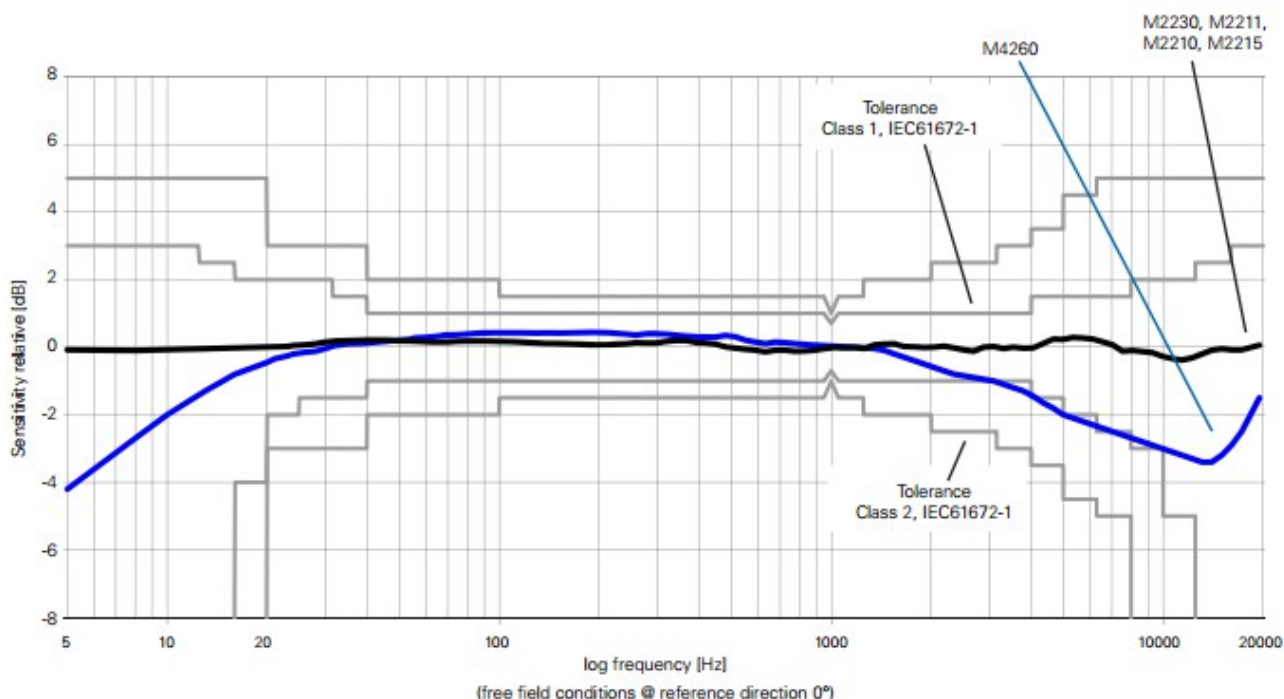
A hang feldolgozásának problémáiról

A hang feldolgozásához azt olyan jellé kell alakítani, amelyet értelmezni és kezelni tud az erre szánt eszköz. A célravezető technológiai lehetőségek és azok határai (esetünkben egy PCM jel) szabványok közé foglalt, konkrét értékekkel segítik a munkát.²

Az átalakításhoz a több fázison, több eszközön halad keresztül a hang, amely bizonyos mértékben el fog térni a rögzített jeltől. Példaként vegyünk egy mikrofont, legyen ez az NTI Audio M2211 modellje. Ez egy csúcskategóriás mérőmikrofon, melynek jellemzője, hogy igen pontos, lineáris átvitelű, a forrás hangjához igen hű jelet továbbító eszköz. A technikai leírásából azonban kiderül, hogy ez az eszköz sem biztosít teljes pontosságot.

Az említett mikrofon átviteli karakterisztikája:

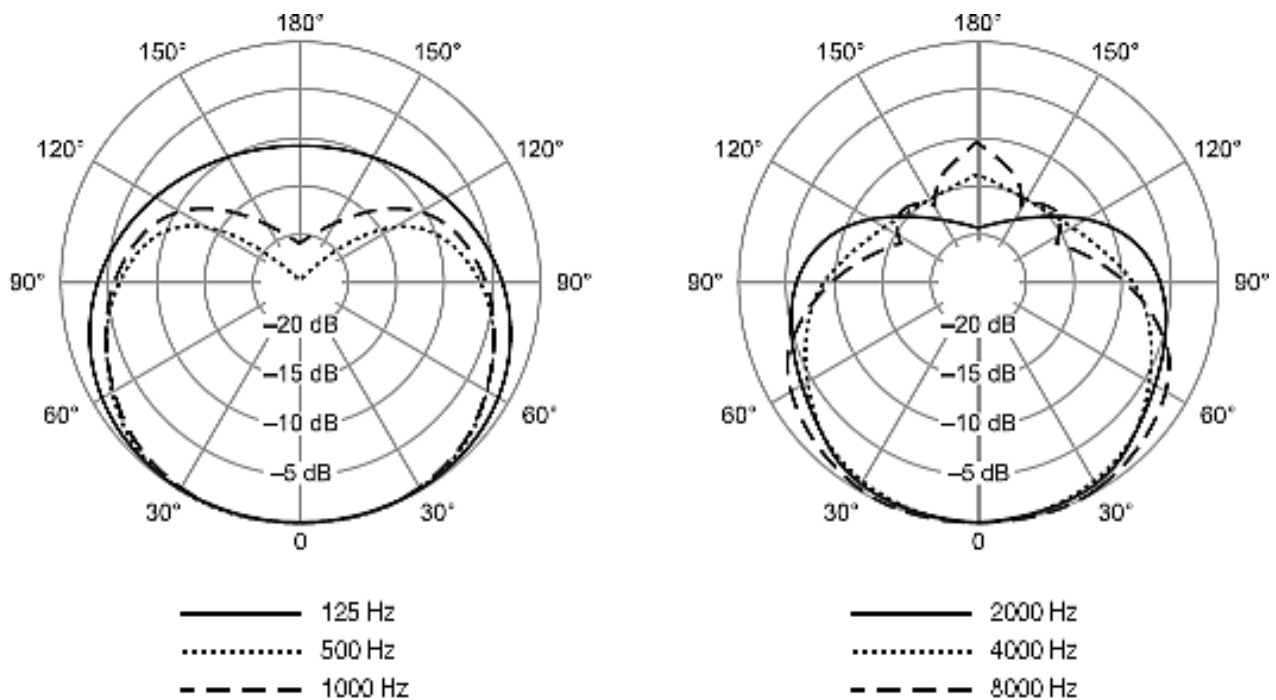
Typical Frequency Response of Measurement Microphones



4. Az említett mikrofon és a gyártó további mikrofonjainak általános átviteli karakterisztikája.

A gyártó eltekint a nem hallható tartományhoz rendelt átvitel ábrázolásától.

A jelölések itt egy sor mérésből kapott átlagot mutatnak, az adott modellszám általános átviteli karakterisztikáját reprezentálандó, továbbá a grafikon jelöl egy-egy megengedhető tűrést is. A legkritikusabb spektrumon, 100Hz és 1kHz között is a pontosabb, „Class 1” osztályba sorolt modellek is $\pm 1\text{dB}$ -t tévedhetnek. Ennek kiküszöbölésére a mikrofon felvételét kezelő technikus egy kiegyenlítő (equalizer) fog használni, amihez jellemzően a mikrofonok forgalmazói egy segéd táblázatot is biztosítanak, ezzel is segítve a technikus munkáját.



TYPICAL POLAR PATTERNS

5. Kardioid mikrofonok tipikus átviteli ábrája, a membránhoz adott irányból érkező jelek esetén.
Az ábra feltételezi, hogy a jelforrásunk a membrán horizontján helyezkedik el. A görbék a hozzájuk rendelt frekvenciák hangnyomásának csillapítását hivatottak jelezni.

Nominal Frequency [Hz]	M2230 Microphone [dB]	M2211 Microphone [dB]	M2215 Microphone [dB]
50	0.0	0.0	0.0
63	0.0	0.0	0.0
80	0.0	0.0	0.0
100	0.0	0.0	0.0
125	0.0	0.0	0.0
160	0.0	0.0	0.0
200	0.0	0.0	0.0
250	0.0	0.0	0.0
315	0.0	0.0	0.0
400	0.0	0.0	0.0
500	0.0	0.1	0.0
630	0.0	0.1	0.0
800	0.1	0.1	0.0
1000	0.2	0.1	0.0
1250	0.3	0.3	0.0
1600	0.4	0.5	0.1
2000	0.5	0.6	0.2
2500	0.6	0.7	0.6
3150	0.8	1.1	0.7
4000	1.0	1.7	1.2
5000	1.5	2.2	2.1
6300	2.2	3.3	3.0
8000	3.3	4.2	3.9
10000	4.6	5.4	3.7
12500	6.5	7.3	6.7
16000	8.0	9.2	9.0
20000	9.0	11.0	10.6

6. Az NTI Audio különböző mikrofonmodelljeihez ajánlott beállítások, a felvételen alkalmazott kiegyenlítőn.

A 4. képen látható átviteli karakterisztika linearizációjához ezeket a beállításokat javasolja a mikrofon gyártója.

Ezen felül további, egyéb módon is torzul a felvételünk, ezek tárgyalására viszont itt most nem kerül sor. Az eltérések viszonylagos kiküszöbölésére több lehetőség adott:

- lineáris átvitelű eszközök használata
- hangszínező, kiegyenlítő használata
- kompresszor használata
- pszicho-akusztikai effektek használata
- az emberi fül természetes torzításának³ modellezése
- panorámázás
- minél nagyobb mintavételi frekvencia használata (48kHz - 96kHz)

Egyik megoldás sem ígér teljes mértékű megoldást a problémára, függetlenül az eszközöket beállító személyzet szakképzettségétől vagy a használt algoritmus pontosságától, optimalizáltságától. Ez fizikailag lehetetlen, azonban az eltérés mértékét elhanyagolhatóra csökkenthetjük.

Bizonyos esetekben nem a linearizáció és a felvétel eredeti dinamikájának visszaállítása a cél, hanem egy, az eredetitől eltérő hangzás kialakítása. Ez az történhet akár a hangzás optimalizálásának, esetleg egy zenei mix érthetőségének kialakítása, stb. céljából.

Elektro-akusztikai, szinkronizált kiegyenlítő

A dolgozat által tárgyalt eszköz elsődleges célja, hogy kiegyenlítőként működtetve, a megadott paraméterek mentén alkalmazza a szűrési feltételeket, a bemenő jel meghatározott értékeinek függvényében. Természetesen ez az eszköz sem tud hiánytalan megoldást biztosítani a feljebb tárgyalt problémára, de azonosan a kiegyenlítőhöz, a felvétel optimalizációjában használatos.

A következő egy igen speciális, ám a zenei-produceri gyakorlatban igen gyakori példa. Feltételezzük, hogy a kezelendő jel a rögzítés folyamán elhanyagolható mértékben torzult, a felvétel pedig egyetlen hangszer hangja, néhány eltérő hangot lejátszva. Legyen a célunk a felvétel hangzásának optimalizálása, azonban a hangszínezésen kívül tekintsünk el minden egyéb optimalizálástól. Erre egy kiegyenlítőt használhatunk.

A kiegyenlítő különböző típusú sávszűrőkből álló egység, működésének és paramétereinek részletes tárgyalása nem tartozik a dolgozathoz. A felvételen szóló hangszer sajátos hangterjedelmét ismerve a kiegyenlítő első paramétereit megadhatjuk. Ezt követi a hangszer testének karakterisztikája, amely (helyzettől függően) további beállításokat igényel a kiegyenlítőn. Végül elhelyeztük és felparamétereztük a kiegyenlítő szűrőit. Eredményül egy olyan jelet kaptunk, amely a hangszer testének sajátosságait kihasználva optimalizált, azaz a kellő tartományokban erősebben és érthetőbben szól. Azonban a különböző hangmagasságokon játszott hangok, amelyek a kiegyenlítőn áthaladó felvételen vannak, a kiegyenlítőben stabilan lehelyezett szűrőkön

áthaladva mind-mind különböző arányú torzítást kapnak. Ez annak köszönhető, hogy az alap körfrekvenciájuk és az arra épülő felharmonikusaik minden hangváltásnál más magasságra esnek, így a felvétel teljes ideje alatt a felvétel alapvető jelleggörbéje folyamatosan változik, a kiegyenlítőben található szűrőkkel és a hangszer testével ellentétben.

Amennyiben a linearizáció a célunk, itt eleget is tettünk a feladatnak, hiszen az eredeti jelforrás karakteréhez közel hű átviteli karakterisztikát hoztunk létre. Azonban ha az optimalizáció a célunk, akkor meg kell elégednünk azzal, hogy a keletkező hangokat egy stabil kiegyenlítő azonosan torzítja.

A szinkronizált kiegyenlítő lényege, hogy az előzőleg leírt kiegyenlítő mellett dolgozzon a hangszínezésen. A következőkben a feldolgozandó hangunkat új kontextusba helyezzük, eltekintünk a hangszer testétől és a húrokon keletkezett rezgéstől független karakterétől. Ahhoz, hogy ezt megtegyük, minden időpillanatban (gyakorlatban a nullához közelítő időintervallumban) fizikailag is definiálnunk kell a hangot, hogy azt a paraméterei segítségével elemezhessük és feldolgozhassuk. A szinkronizált kiegyenlítő elvégzi ezt a műveletet és a benne lévő szűrőket ennek függvényében, a létrejövő felharmonikusokat figyelembe véve engedi elhelyezni.

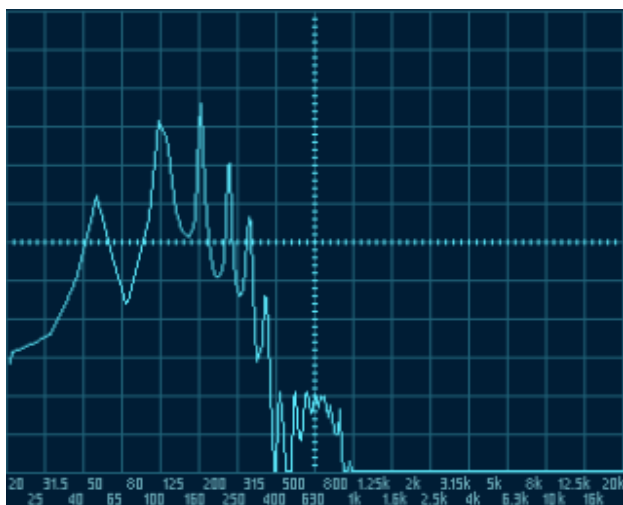
Eredményképpen egy olyan dinamikusan változó karakterisztikát kapunk, amely a megszólaló hang függvényében, annak valódi, hangmagasságtól függetlenül, általános karakterét engedi színezni, ezzel további lehetőségeket biztosítva a hangzás optimalizálására.

Az előző példán kívül természetesen több további funkciót láthat el az eszköz, ezek részletezésére később visszatérünk.

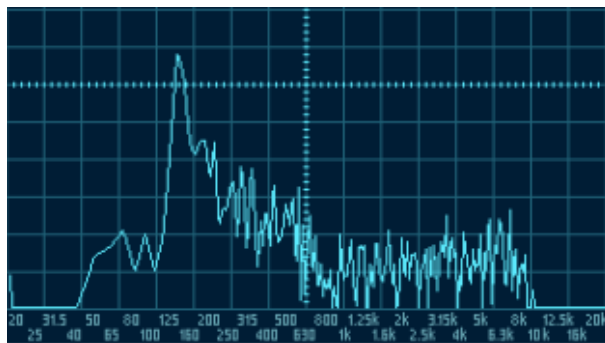
A szinkronizált kiegyenlítő elvi háttere

Az egyszerűség kedvéért az előző példán haladunk tovább. Mint láthattuk, a tárgyalt eszköznek ismernie kell az aktuális időpillanatban szóló hang körfrekvenciáját. Ennek megállapításához a korábban említett Fourier-transzformáció mellett, néhány kísérlet eredményéből kapott optimalizáció használható. Mint azt korábban tárgyaltuk, a gyakorlatban

bizonyos esetekben nem az alaphang magasságára esik a legnagyobb amplitúdó. Alább a néhány vizsgált hang spektrumán határozottan megfigyelhető a probléma.



7. Natúr, hangszedőn át rögzített gitárhang. Az játszott hang H1, 61.74Hz frekvenciával. Jól láthatóan az első oktáv 123.48Hz-nél nagyobb amplitúdójú, illetve ennek kvintje 185.22Hz-nél a jel legmagasabb amplitúdójával rendelkezik.



8. Egy pergődob hangja. A hangkarakteréhez ajánlott mikrofonnal rögzítve, pusztán egy vákuumcsöves előerősítő kompressziójával. A pergődob alaphangja jól kivehetően 130Hz környékén a legnagyobb amplitúdóval szól, azonban ez alatt is vélhetünk felfedezni csúcsokat. Ezek a felvétel körülményeiből adódhatnak. Például a dobtesthez rögzített mikrofon a dobőr megütésével, a dobtestről mozgási energiát vesz át, ami a benne lévő membránon offset-hibát okoz. Ez ugyancsak rezonanciaként jelentkezik.

A probléma megoldására több lehetőség is adott:

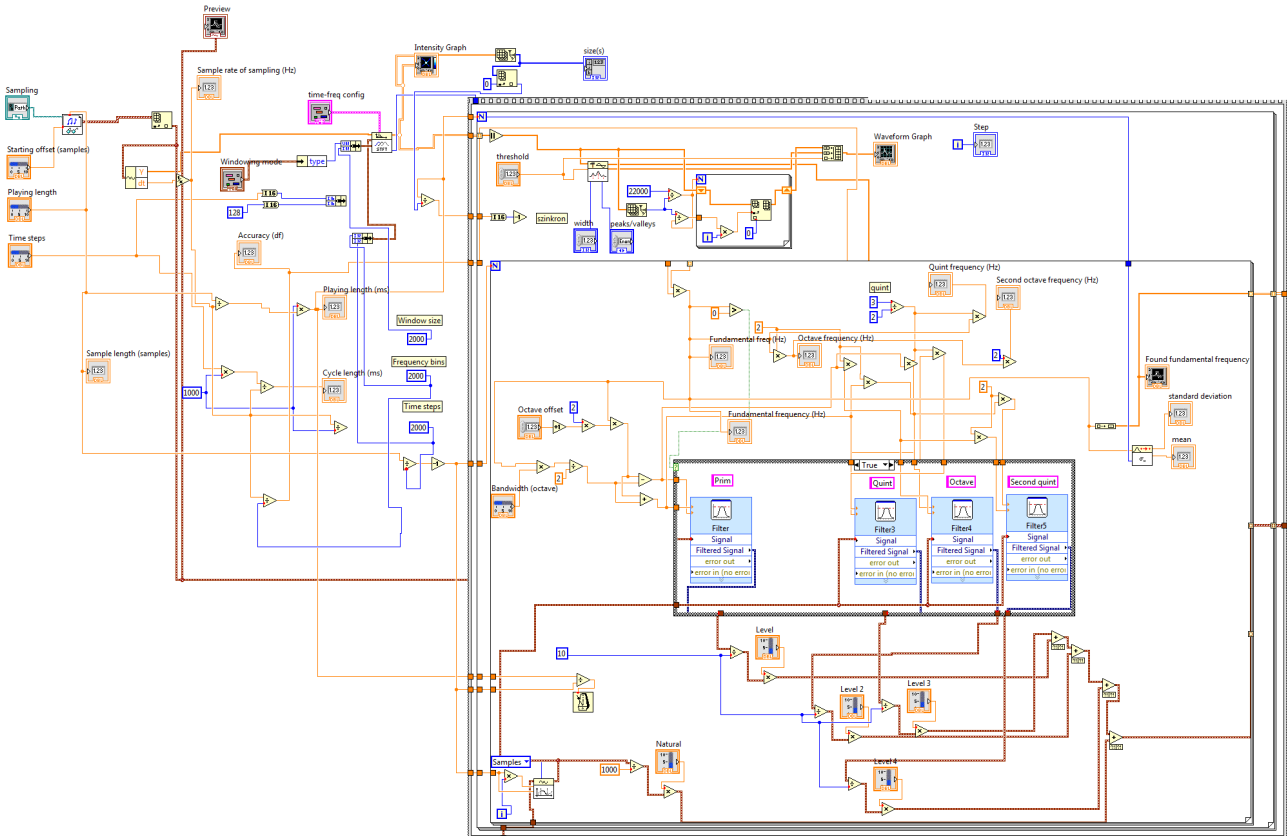
- threshold érték fölötti csúcsponatok keresése
- a kapott csúcsponatok közül kiválasztjuk a legmagasabbat (például a pergődob esetében)
- a kapott csúcsponatok közül az elsőt vesszük figyelembe
- az összes csúcsponatok közül a két egymást követő, de egymástól legnagyobb távolságra lévő csúcsponatok közül az elsőt vesszük figyelembe, feltételezve, hogy ez az alaphang és az oktáv párja közötti „völgy”
- előzőleg, vagy alkalmazás közben definiálható profil segítségével történő kiválasztás
- stb.

A jel körfrekvenciájának megállapítását követően paraméterezhetővé válnak a kiegyenlítő szűrői. A szűrők típusán és sávszélességén túl, a szűrők lehelyezésének lehetősége, az alaphangtól való harmonikus távolságokra is adott a lehetőség. A későbbiekben a körfrekvencia változásakor a szűrők az itt definiált viszonylagos távolságot követik. Eredményképpen egy, a jel alapfrekvenciájától függő kiegyenlítőt kapunk.

A körfrekvencia változása további problémákat vet fel, többek közt a jelben pillanatnyi szünet esetén eltűnő csúcsérték, illetve gyors hangmagasság-váltás esetén a kiegyenlítő a Fourier-transzformációhoz szükséges idő miatt a kiegyenlítő pontatlan működését eredményezi. Ezek a problémák a kiegyenlítő működési elvének optimalizálásával kezelhetők. Erre ismét több lehetőség adott, például:

- a kiegyenlítő által előállított jel rákeverésének mértékét a Fourier-transzformációhoz szükséges idő függvényében növeljük
- egy alsó frekvenciahatár alá eső alaphang esetén a kiegyenlítő jelét nem keverjük a jelre, a jelet érintetlenül adjuk tovább
- változó alapfrekvencia esetén a definiált szűrők, idő függvényében való, adott sebességű és gyorsulású mozgatása

A prototípusról



9. A prototípus blokkdiagramja National Instruments LabVIEW-ban.

A különböző felhasználási területek más-más paraméterezési lehetőségeket és bizonyos esetekben eltérő konstans értékekkel rendelkeznek. A felhasználási lehetőségeket tárgyaló fejezetben előrekül egy LFO-s (low frequency oscillator) és egy ADSR (attack, decay, sustain, release) vezérlés, azonban ezeken túl még számtalan egyéb paraméterrel közelíthető a megvalósítandó részfeladat. Fontos kihangsúlyozni, hogy a paraméterezési lehetőségek természetesen csak kezelési módszerek meghatározását igénylik. A konstans értékek megadása azonban egy-egy célra készített koncepció vagy disztribúció szigorúbb optimalizálásának okán igényelték, jellemzően az eszköz működésének alapvető matematikai hátterét érintik, így a felhasználó közvetlenül nem kap lehetőséget ezek szabad módosítására. Ilyen például a Fourier-transzformáció során előkerülő érzékenység, amely megszabja az eszköz működésének pontosságát, illetve az egy-egy szűrőhöz tartozó erősítési érték maximuma.

Az elkészült prototípus esetében a legfontosabb az volt, hogy gyorsan, egyszerűen és látványosan, az eltéréseket jól hallhatóan reprezentálja a bemutatók alkalmával. A fent említett

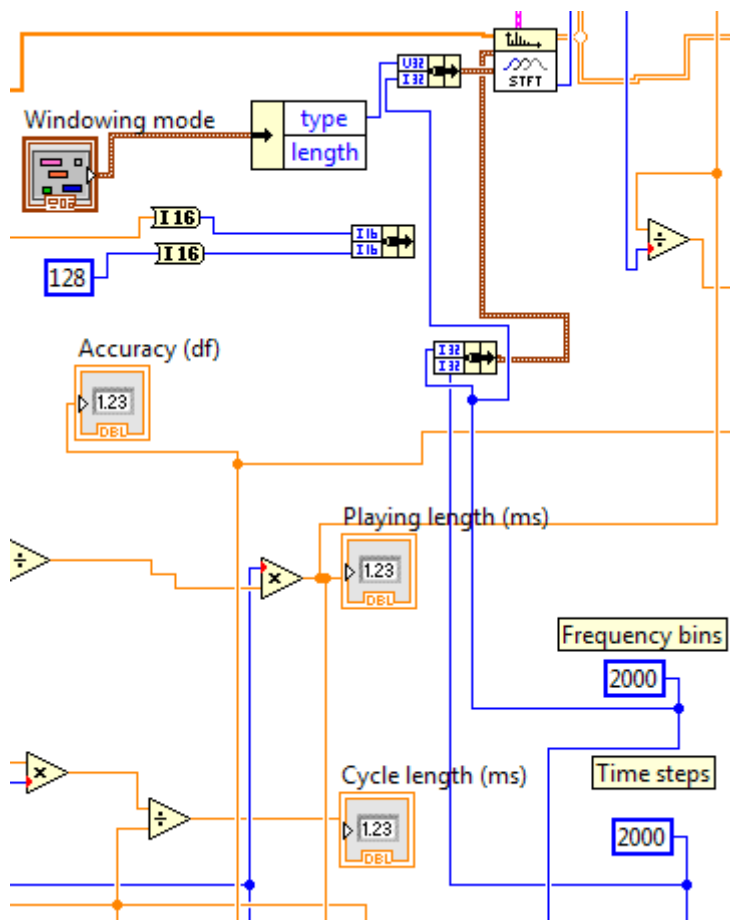
példák a teljes felhasználási palettából csak intuitív sorrendben és módszerrel kiválasztott elemek, sem kategóriákat sem a lehetőségek bezárt halmazának elemeit nem hivatottak körvonalazni. A prototípus egyik felsorolt példa funkciójához sem lett alkalmazva, hogy megmaradjon irányadó bemutatónak.

A prototípus a National Instruments LabVIEW rendszertervező software segítségével készült. Jelen formájában nem implementálható semmilyen egyéb platformra. A későbbiekben az előzőleg meghatározott felhasználási koncepcióknak általánosan megfelelő alapmodell, majd az egyéni specifikációknak megfelelő héjprogramok kerülnek fejlesztésre. Ezeket részletesebben a dolgozat későbbi fejezetei tárgyalják.

A beállított konstans értékek és a megadott paraméterezési lehetőségek, azok kijelölt intervallumaival jelen esetben ezt a célt szolgálják. Fontos hangsúlyozni, hogy a további alkalmazási területeknek megfelelően ezek természetesen bizonyos mértékig változtathatók, a jelenlegi értékeik nem szükséges feltételei a működésnek.

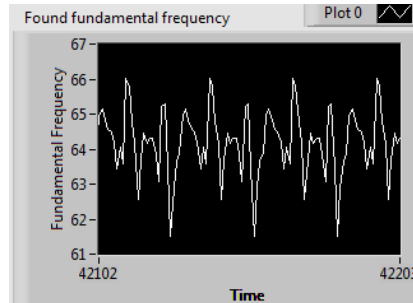
Az alábbi táblázat összefoglalja és magyarázza a konstansnak vett értékeket a prototípus programjában:

Paraméter megnevezése	Felvett érték	Magyarázat
Frequency bins [N]	2000	A Frouier transzformációhoz szükséges minták (samples) mennyisége időosztásonként.
Time steps	2000	Az előző értékhez tartozó időosztás hossza. Ezzel az értékkel csúszik az ablak a mintán ciklusonként. (samples)
Window size	2000	Az ablakoláshoz megadott ablak hossz. (samples)
Ablakolás típusa	Négyszög	-
Threshold	0,001	A peak-detector modul ez alapján az amplitúdó érték alapján választja ki az első csúcsot a spektrumból.
Width (Peak)	3	A keresendő csúcspontok szélessége. (samples)



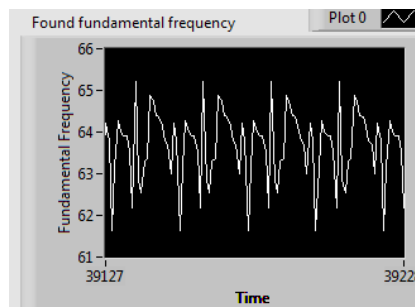
A korábban ismertetett, spektrumgörbén kiválasztandó, alap frekvenciát jelölő csúcs kiválasztásának problémája és az STFT modul miatt szükséges volt az ablakolás módszerével is foglalkozni. A különböző ablakolási eljárások közül a minél alacsonyabb ciklusidő használatának érdekében a négyszöges ablakolás bizonyult a legpontosabbnak. Alább látható egy összehasonlítás a különböző ablakolási módszerek és fenti paraméterek használatával megállapított frekvenciák idő-frekvencia diagramja. A felvételen egy elektromos gitáron, torzítás nélkül, folyamatosan szól egy C₂ hang, 65.41Hz-en.

Blackman-Harris



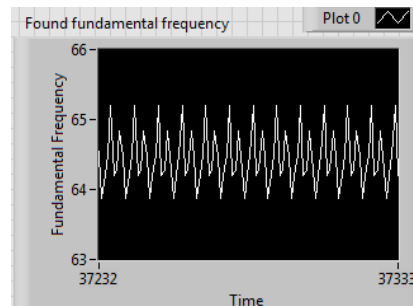
$$\mu = 64,1 \text{ Hz}$$
$$\sigma = 1,07 \text{ Hz}$$

Gaussian



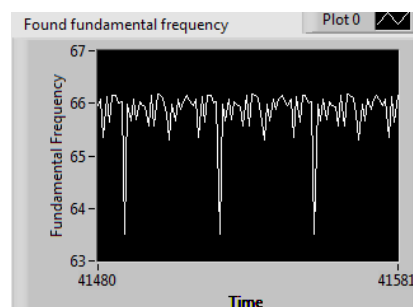
$$\mu = 63,71 \text{ Hz}$$
$$\sigma = 0,88 \text{ Hz}$$

Hamming



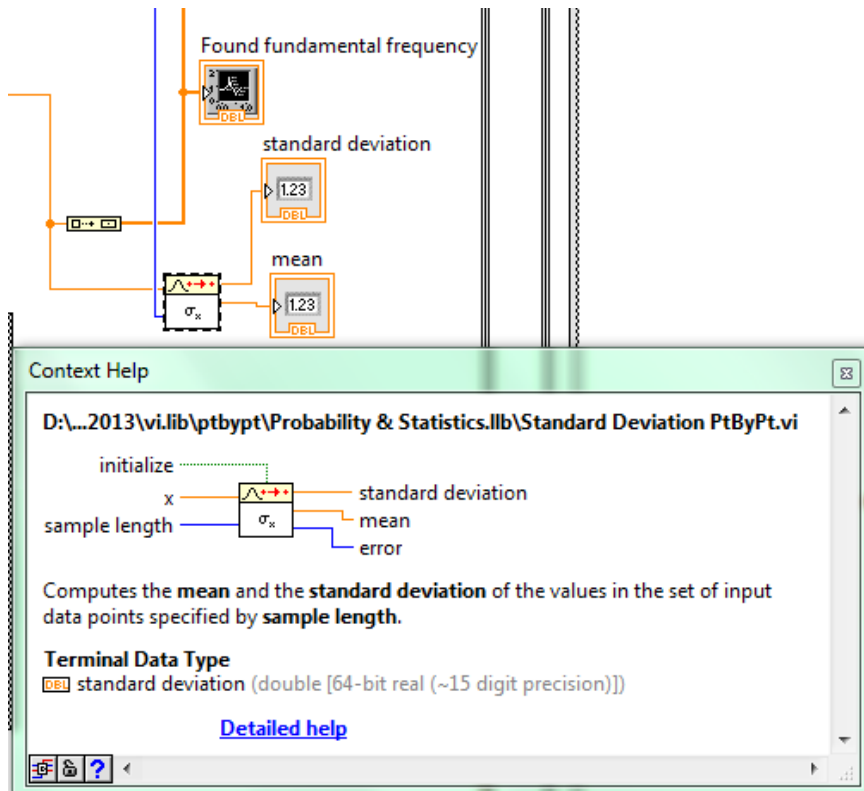
$$\mu = 64,55 \text{ Hz}$$
$$\sigma = 0,84 \text{ Hz}$$

Négyszög

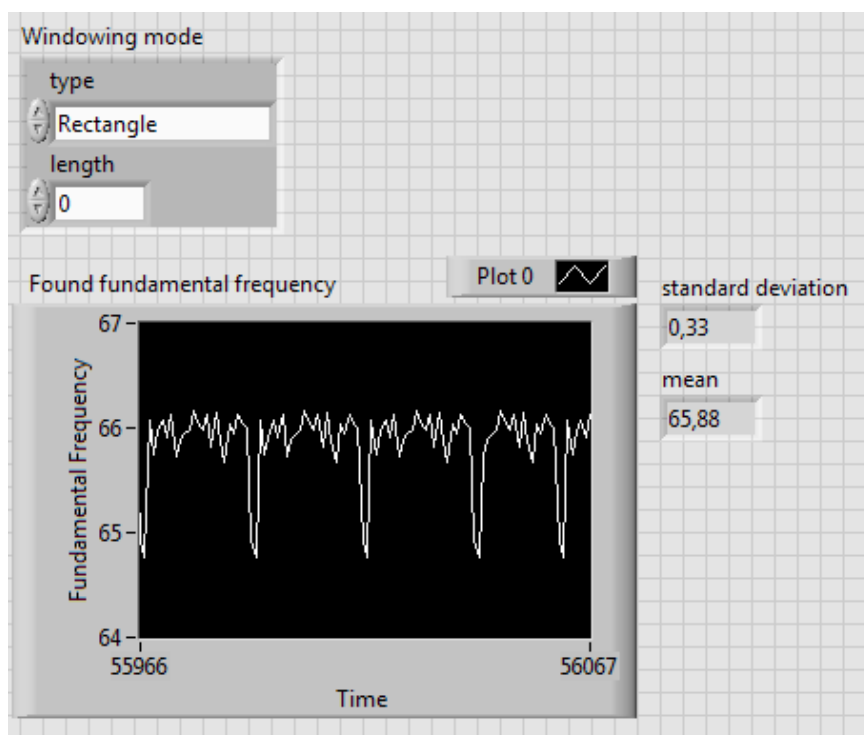


$$\mu = 65,88 \text{ Hz}$$
$$\sigma = 0,33 \text{ Hz}$$

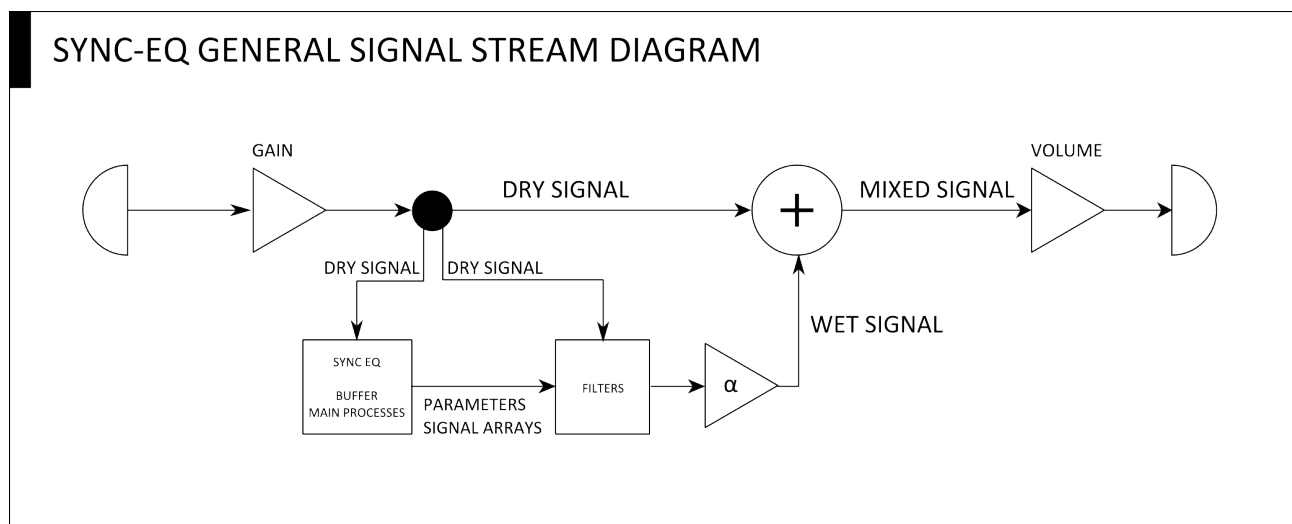
A Hamming ablakolás kivételével mindegyiken jól látható egy tranziens, amely a minta indulásakor keletkezik. Az ismétlődés a sampling folyamatosan ismétlődő lejátszásának köszönhető. Ennek ellenére az elvégzett számítás alapján is a legpontosabb ablakolás a legkisebb szórással a négyszög ablakkal sikerült, ahogy azt az előző táblázat mutatja.



11., 12. Az alábbi képeken a statisztikai számításhoz használt modul és bekötése látható, illetve a vonatkozó indikátorok a kezelőfelületen.



A lehető legalacsonyabb ciklusidő elérése, és a négyszöges ablakolási mód mellett szól a vett minták számának, az ablak méretének és az ablak eltolásának azonos értéke. A ciklusidő pedig meghatározó a valós idejű feldolgozás lehetőségének igénye mellett. Sajnos a valós idejű feldolgozás nem része a jelenlegi prototípusnak, azonban a következőkben egy, a hardwares integrálás esetén jelentkező probléma és optimalizációs lehetőség alátámasztja az ablakolási mód megválasztását és a különböző paraméterek kapcsolatát.



13. A szinkron-eq általános jelfolyam diagramja. Ez már az implementálható rendszer alapja.

A prototípus ezt nem követi.

Az alábbi táblázatban a különböző mintavételezési ráták mellett látható egy adott ciklus feldolgozásához szükséges idő, illetve ennek torlódása egy 10 másodperc hosszú minta feldolgozása esetén.

A torlódó feldolgozási idő az alábbi képlet szerint alakul:

$$\text{torlódó feldolgozási idő [ms]} = \frac{1 \cdot 10^3}{CL_H} \cdot CL_t$$

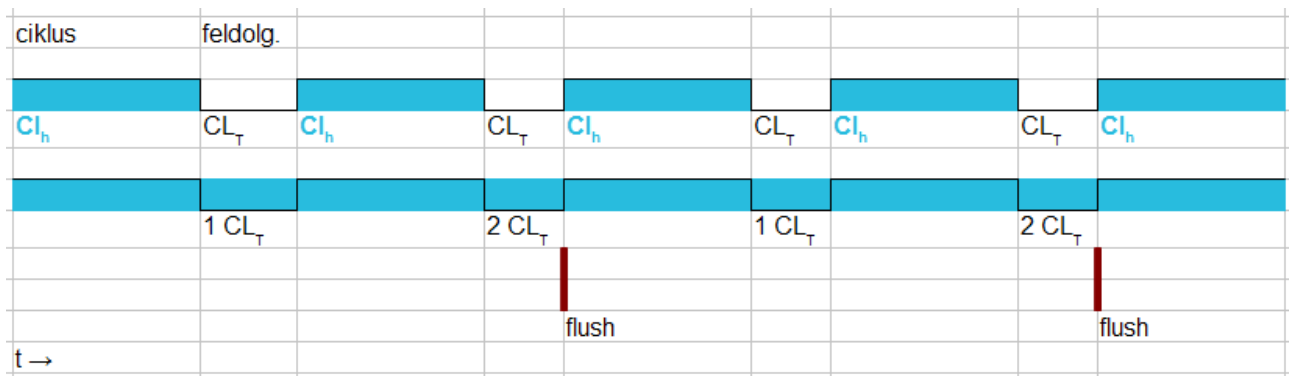
Ahol CL_t az egy ciklus feldolgozásához szükséges idő hossza. Ez az érték az eljárást végrehajtó eszköz konfigurációjától függ. CL_H a ciklus időbeli hosszúsága.

Mintavételezési ráta (Hz)	1 minta hossza (ms)	CL _h (ms)	ΔF mérési pontosság (Hz)	Torlódó ciklusidő 10 sec esetén (ms)
44100	0,02276	45,52	22,05	219,6837 · CL _t
48000	0,02083	41,66	24	240,0384 · CL _t
96000	0,01042	20,84	48	479,8464 · CL _t

Hogy a torlódás ne lehetetlenítse el a valós idejűnek tűnő jelfeldolgozást, érdemes adott időszeletenként eldobni a még fel nem dolgozott ciklusokat, vagy a feldolgozás ideje alatt eltekinteni a mintavételezéstől. Abban az esetben, ha hardwares integrálás folyamán a műveletet végző processzor megszakításvezérlésére bízunk, hogy a ciklusokat felgyűjtse, párhuzamosan feldolgozza és keverje a nyers jellel, további műveleti időt veszünk. Ha azonban egy valós idejű műveleteket támogató architektúrára (például PRU társprocesszorral szerelt processzorral⁴) készítjük a programot, kihasználhatjuk a hierarchikus megszakításvezérlés nyújtotta lehetőséget, hogy adott, pontosan hangolt időközönként eldobjuk a feldolgozandó ciklusokat, ürítsük az előzőleg töltött puffert.

Az alábbi két példában négy ciklus mintavételezése és feldolgozása látható. Az első alkalmával (eltekintve a párhuzamos feldolgozás lehetőségétől) nem ürítjük a feldolgozandó ciklusokat, viszont mintát csak a feldolgozás előtt és után gyűjtünk, ezzel is időt spórolva a rendszernek. Ebben az esetben csak a pufferek, processzusváltás, feldolgozás, processzusváltás folyamatait ismétljük. Ezzel a megoldással elkerüljük a feldolgozási idő torlódását, azonban a rövid megszakítások révén nem kapunk pontos eredményt.

A második példában a ciklusidő párhuzamosan fedi a feldolgozási időt. Itt a processzorra bízunk a két folyamat közötti processzusváltás időzítését, azonban a példa szerint minden második ciklusidő letelte után megszakítjuk a feldolgozást és ürítjük (flush) a feldolgozandó tömböket.



14. A valósidejű jelfeldolgozás két példájának idő-funkció diagramja.

Mindkét esetben a lehető legkisebb műveleti időre van szükség, hiszen a valósidejűnek tűnő feldolgozás feltétele, hogy 8ms alá essen a feldolgozott jel késése. ⁵

Eszerint az alábbi egyenlőtlenségnek minden beállítás mellett igaznak kell lennie, a valósidejű hatás igénye mellett:

$$8ms \geq \sum_{i=k}^{k+n} \left(\frac{1 \cdot 10^3}{CL_{Hi}} \cdot CL_{ti} \right)$$

Ahol k az adott sorszámú flush, n pedig a két flush közötti ciklusok száma.

Tovább haladva a prototípus tesztelése folyamán előkerült egy másik probléma, mégpedig a STFT modul viszonylagos pontatlansága. Korábban az ablakolási módszerek szemléltetésénél láthattuk, hogy pontos értéket már ott sem kaptunk, némi tűréssel számolnunk kellett. Az alábbi táblázat, a fent korábban beállítások mellett Fourier-transzformáció pontosságát adja meg, amely az alapvető körfrekvencia megállapításánál ad egy tűrési értéket, különböző mintavételezési rátával készült jelfolyamok esetén.

A tűrés meghatározása az alábbi képlet segítségével történik:

$$\Delta F = \frac{f_s}{N}$$

Ahol ΔF az a lépték, amelyen az STFT műveletünk végül értéket állapíthat meg. Értéke a jelfolyam mintavételi rátájának és az STFT-hez időosztásonként felvett minták számának hányadosa.

A ΔF értékéből következtethetünk egy általános hibatartományra:

$$H_f = \frac{\Delta F}{2}$$

Sample rate of sampling (Hz) 44100

Sample length (samples) 103544

Accuracy (df) 22,05

Windowing mode

type Rectangle

length 0

Cycle length (ms) 45,52

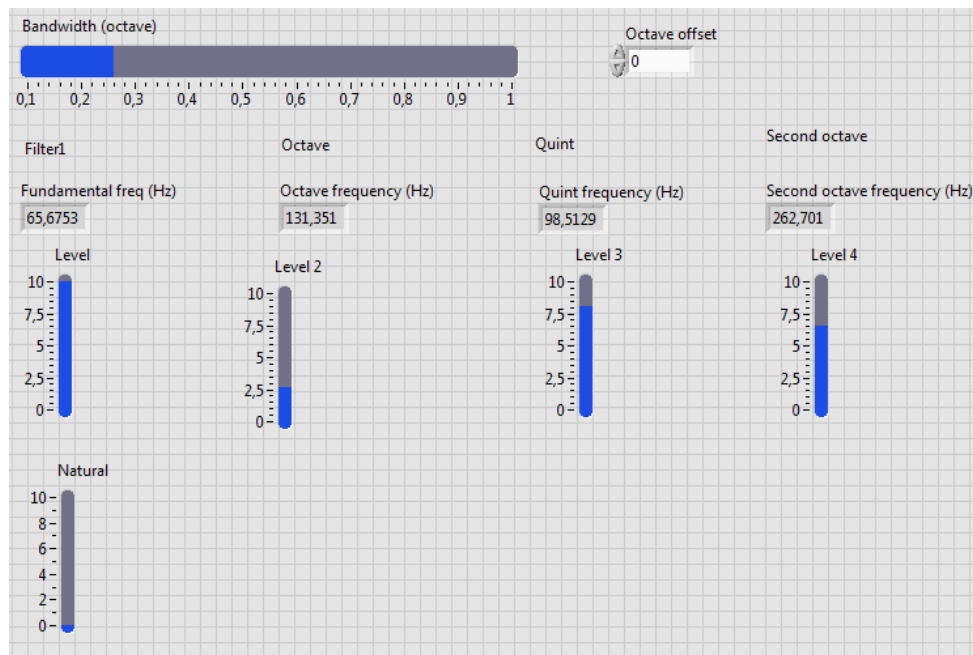
Playing length (ms) 2347,93

threshold

15. A prototípus által megadott értékek a samplingről, illetve a művelethez tartozó egyéb, számított adatok.

Mint azt a korábbi táblázatban is láttuk, ez az érték a különböző mintavételezési ráták és az előzőleg felsorolt konstans értékek mellett az alábbi módon alakul:

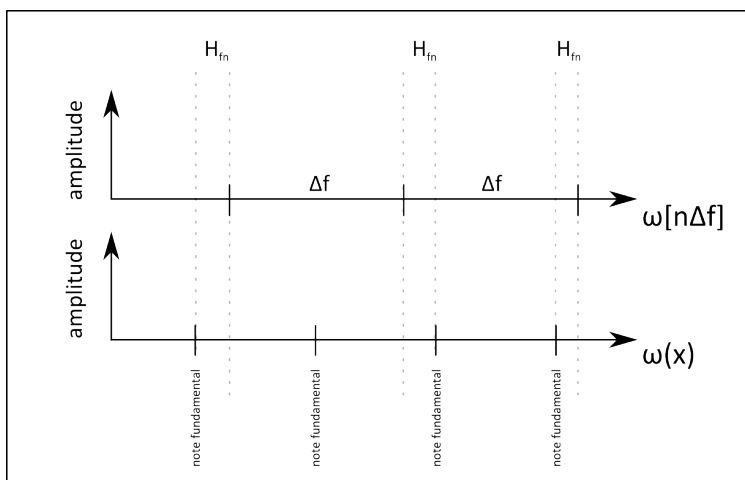
f_s (Hz)	44100	48000	96000
ΔF (Hz)	22,05	24	48
H_f (Hz)	11,025	12	24
Legalacsonyabb, hiba nélkül megkülönböztethető zenei hang.	G_4 392.00 Hz	G_4 392.00 Hz	A_4 440.00 Hz



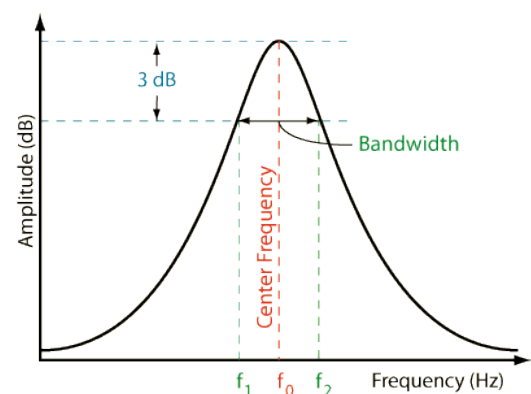
16. A prototípus paraméterezési lehetőségei csúszkák segítségével.

Itt állítható be a különböző szűrők erősítése, illetve a szűrők sávszélessége.

A táblázat utolsó sorában láthatóak azok a zenei hangok, amik tévedés nélkül megállapíthatók az adott beállításokkal. Ez azt jelenti, hogy ha minden szűrő sávszélessége 0-hoz tart, jó közelítéssel a zeneileg temperált rendszerben megadott hangon és annak meghatározott felharmonikusain, zeneileg pontos szűrés lesz az eredmény. Azonban a szűrők sávszélessége változtatható, így a tévedés bizonyos fokig korrigálható, lefedik a valóban játszott hangmagasságot és a meghatározott felharmonikusait.



17. A tűrésből adódó hibák.



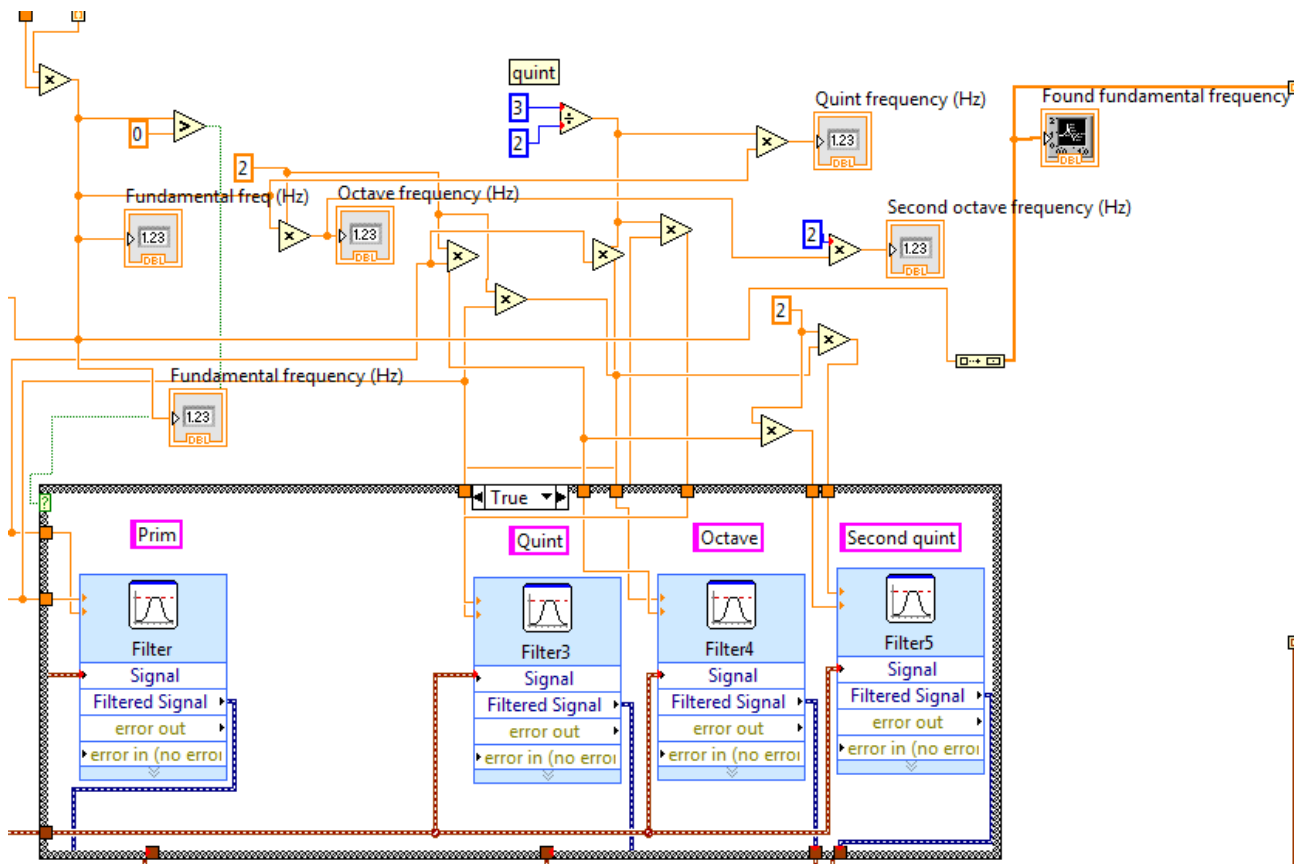
18. A band-pass filterek karakterisztikája.

Kép forrása:

<http://www.rane.com/note170.html>

A prototípus jelenleg négy band-pass filterrel van ellátva, melyek az alábbi táblázat szerinti felhangokra vannak beállítva:

Harmonikus	Alaphang	Kvint	Oktáv	Második kvint
Szorzó érték	1	$3/2$	2	3



19. A prototípus filterei és bekötött paramétereik blokkdiagramon.

Felhasználási lehetőségek

A szinkronizált-kiegyenlítő adott paraméterezés mellett természetesen nem csak a fent levezetett példa esetében és a példában betöltött funkciójára használható. Pár példa a megfelelő paraméterezés esetén további elektro-akusztikai (egyéb jelfeldolgozási területektől a dolgozat eltekint) felhasználási lehetőségekhez:

Adott egy felvétel, amelyen egy rövid dallam hallható. A felvétel egy DAW (digital audio workstation) rendszerben rögzített jelfolyam. Tegyük fel, hogy a kiegyenlítést elvégeztük, jelenleg egy tiszta, arányos hangot hallunk. A keverésnél azonban előkerül egy probléma: egy másik hangszer uniszónóban, azaz párhuzamosan vezeti a dallamot, amit az imént a felvételen kezeltünk. Sajnos a hangkarakterük igen hasonló, így a fáziskülönbségek miatt irreális, esztétikátlan keveréket hallunk. Ha úgy tetszik, a két hang súrlódik, ráadásul nem telíti ki a hangképet, hanem erősen lefojtja a felhangtorlódás miatt. Romlik a hangosságérzet, a dinamika, végeredményben kapunk egy olyan hibát, amit azonnal meg kell oldani az egyik sáv kitörlésével. A megoldás egy szinkron-eq segítségével is előállhat: különbözővé kell tenni a két hangszer felhang-struktúráját, hogy érthetőek legyenek és mégse súrlódjanak. Előállhat egy olyan probléma is, hogy ettől idegen lesz a hangszerek hangja, így ezt is figyelembe véve a következő konstrukciót használhatjuk:

A két hangszer azonos felhangjaira illesztünk egy-egy sáváteresztő szűrőt. A célunk az, hogy a hangszerek hangjait különbözővé tegyük egymástól, de lehetőség szerint hűek maradjunk azok eredetijéhez. Így az előzőleg definiált szűrők erősítési és csillapítási értékeit egy-egy LFO-ra bízuk melyeket adott frekvencián mozgatunk. Ott tartunk, hogy a két hangszer karaktere továbbra is azonos, de már folyamatosan változó és mintegy átlagként járják körbe az eredeti hangkarakterüket. A két hangszer hangjának elkülönítéséhez már csak fázisban kell eltolni az első és második hangszer szinkron-eq szűrőit vezérlő LFO-kat egymástól. (Konkrétan az első hangszerhez tartozókat a másik hangszerhez tartozóktól, vagy fordítva.)

Az eredmény az, hogy a két hangszer hangja elkülöníthető, felhangjaik nem torzulnak és a súrlódás érzetét is a minimumra csökkentettük, így nincsenek a dinamikát és a hangosságérzetet romboló csúcspontok és megtarthattuk az eredeti elképzelésnek megfelelő zenei koncepciót.

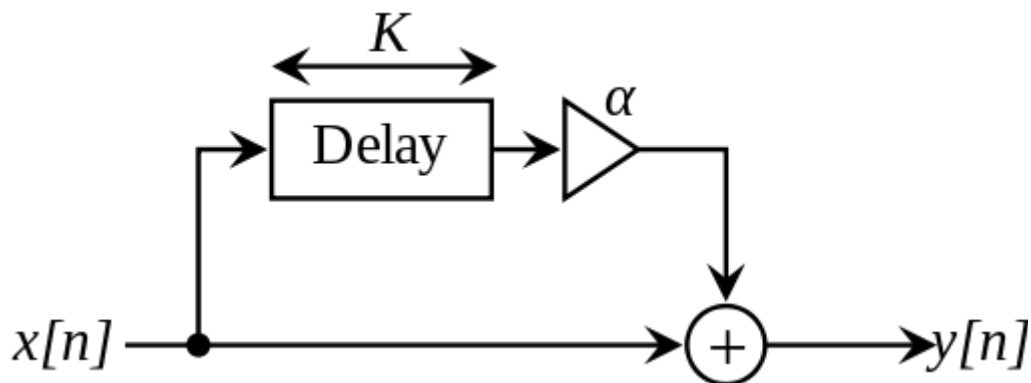
Többsávós-kompresszor – Amennyiben az adott szűrők erősítését vagy csillapítását egy kompresszor működéséhez hasonlóan automatizáljuk, úgy a kiemelt felharmonikusok mentén akár egy többsávós kompresszor képes működni. A különbség egy ilyen eszköztől többek között, hogy a többsávós-kompresszor esetén a kompresszor működéséhez különböző intervallumokat adhatunk meg adott frekvenciahatárokkal. Itt a különböző kompresszorok adott felharmonikusokhoz igazodnak és működnek. A szinkron-eq szűrőinek erősítését vagy csillapítását ebben a példában nem LFO-k szabályozzák, hanem az adott szűrőhöz vagy a – koncepcióból adódóan – szűrőcsoporthoz rendelt ADSR, határérték és kompresszási arány paraméterek.

Az előző példából következik még egy linearizációs lehetőség. Ezzel az eljárással az előzőleg definiált, a hangszer saját hangkarakteréhez igazodó kiegyenlítő statikus átviteli karakterisztikája által felvetett problémát kezelhetjük. Az adott időpillanatban szóló hang esetében elérhető, hogy a fix pontokra helyezett szűrőkkel ellentétben az adott időpillanatban szóló hang ne szenvedjen eltérő torzítást a játszott hang magasságának függvényében, mégis megőrizzük a statikus kiegyenlítővel definiált hangkaraktert.

Egy keverésnél jellemzően a mélyebb tartomány hangjai torlódnak fel először, ezzel rengeteget rontva a dinamikán. Ez annak köszönhető, hogy sok felvételen nem vágják le az alaphang alatti sávot. Ezek jellemzően nem szükséges elemei a felvételnek, sőt, bizonyos zajok is ebben a sávban jelentkezhetnek, például a korábban említett mikrofont érő ütés, vagy a mikrofon hirtelen elmozdulásából adódó nullpont-eltolás okozta zaj. Ezeket a helyzeteket könnyen elképzelhetjük, gondoljunk csak egy kimikrofonozott dobszerelésre. Előfordul, hogy a dobos a mikrofont vagy annak rögzítőjét, illetve a kábelt üti vagy mozdítja meg játék közben.

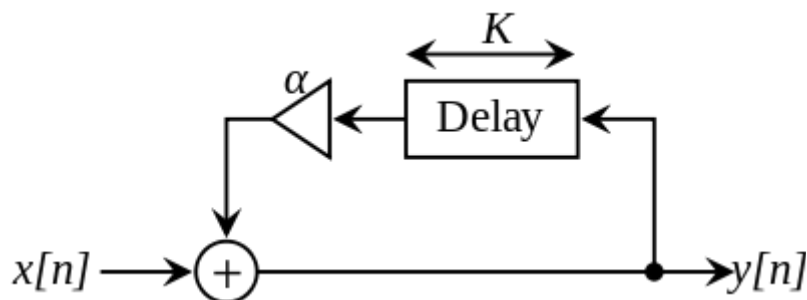
A szinkron-eq segítségével a megállapított alap körfrekvencia alá eső tartomány egyszerűen eldobható egy high-pass filter segítségével, amely természetesen mindig az adott időben megállapított alaphang után mozdul a frekvenciaskálán. Több sáv keverése esetén itt sok energiát spórolhatunk meg egy erősítőnek és a lesugárzó membránoknak, ezzel is optimalizálva a hangzást, növelve a hangosságérzetet és tisztábbá, érthetőbbé téve a keverést.

Fésű filter – A fésű filter effektust késleltetéssel lehet előállítani úgy, hogy a késleltetett jelet visszakeverjük az eredeti jelre, így a keletkező interferenciával „roncsoljuk” a jelet. A fésű filter egy igen idegen, „harapós” hangkaraktert állít elő, azonban a késleltetés miatt a gerjesztő jel megszüntetésével a kevert jel egy plusz lecsengéssel ér véget. Gyakorlatilag a hatás ugyanaz, mintha egyenlő távolságokra sáváteresztő szűrőket helyeznénk le a jelen, azonos sáv szélességgel és erősítéssel.



20. Feedforward visszacsatolással megvalósított fésű filter.

Kép forrása: http://en.wikipedia.org/wiki/Comb_filter



21. Feedback visszacsatolással megvalósított fésű filter.

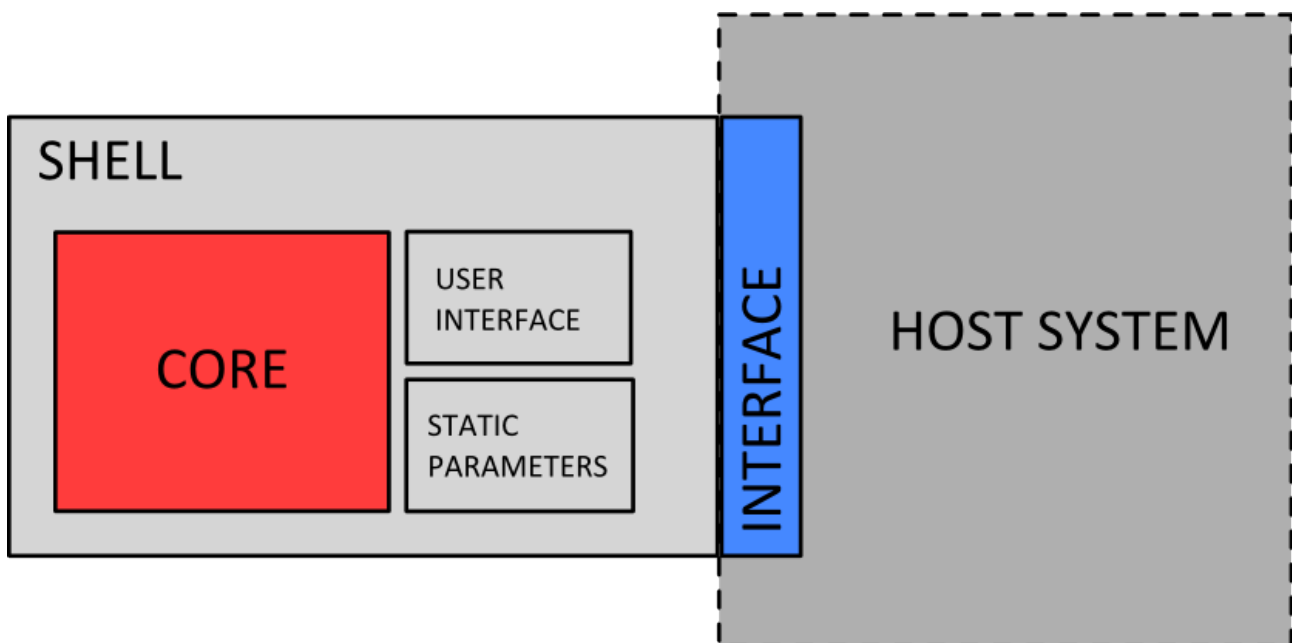
Kép forrása: http://en.wikipedia.org/wiki/Comb_filter

A szinkron-eq segítségével a virtuálisan definiált szűrők száma adott mértékig változtatható, így akár egy fésű filter is létrehozható a segítségével. A fésű filter hatás emellett harmonizálható és paraméterezhető. Bár ez az eljárás nem garantálja a visszacsatolás pontosságát, azonban a lecsengés ilyen módú elhagyása és a fésű filter egyéb beállítási lehetőségei (Például csak az oktáv-kvart vagy kvart-kvint, esetleg csak kvint stb. felhangok bevonása.) egy új fésű filter eljárás megvalósítását teszik lehetővé.

Az adott hangkarakter színesítése – Ez az eszköz szubjektív, művészi értékű felhasználásától függő, igen széles kört felölelő lehetőségeket jelent.

Fejlesztési tervezet

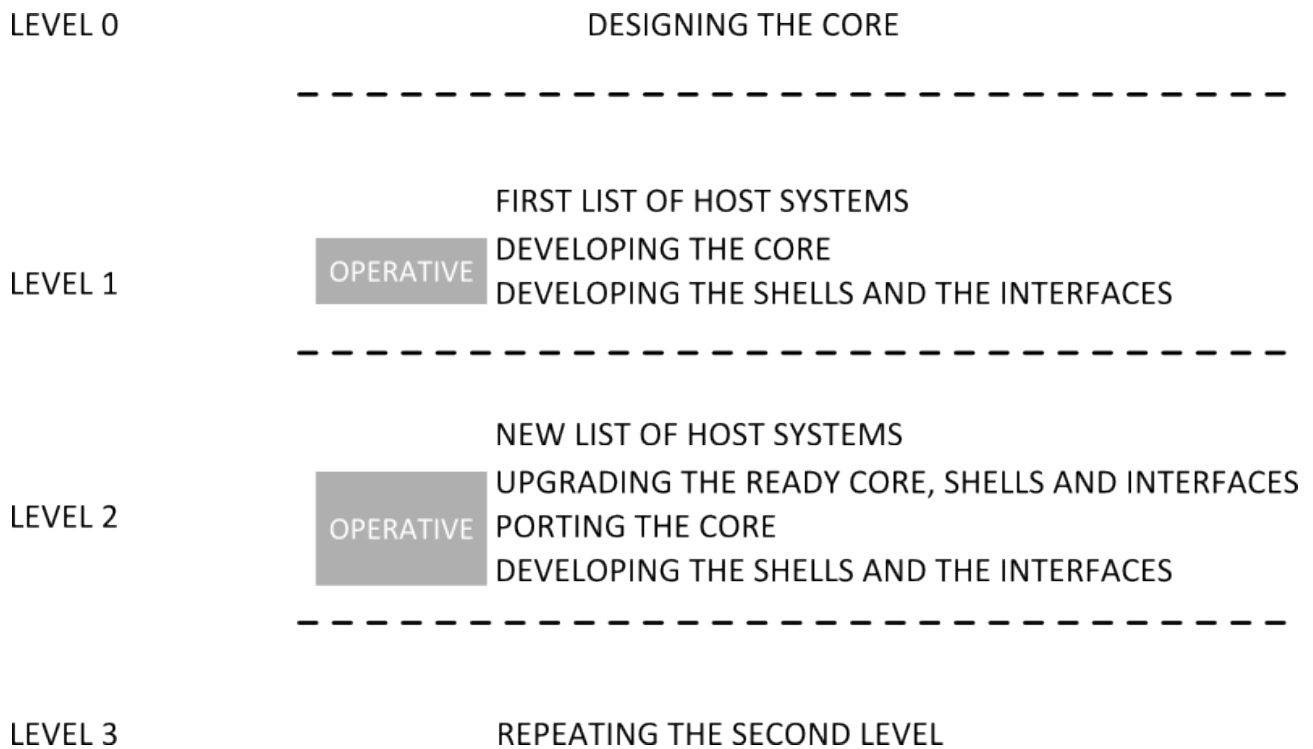
A szinkron-eq mögötti projekt jelenlegi modellje lépcsőzetesen épül fel. A végcél egy VST software és egy hardware-be integrálható rendszer, amely akár softwares, akár hardwares beépülő modulként, illetve önálló egységként is realizálható. Ezek három jellemző részből állnak: a mag, amely a különböző felhasználási területeknek megfelelő héjaknak azonosan megfelelő alap, illetve az említett héjak, amik az adott feladathoz optimalizált beállításokat és paraméterezési lehetőségeket tartalmazzák, végül pedig egy interface rendszer, amely a megfelelő, befogadó rendszerhez csatolja az eszközt.



22. A fejlesztési terv szerinti általános rendszerszerkezet.

Ennek elérése érdekében a projekt egyik vetülete a mag fejlesztésétől az interface fejlesztéséig haladó idő-folyamat diagram, másik vetülete pedig egy verziókövető, befogadó eszközöket skálázó modell.

Végeredményben a második vetületben az első modell operatív műveletekként szerepel, így a fejlesztés folyamata a befogadó rendszerek skálájának bővülését a modell második szintje után folyamatosan követheti.



23. A ciklikus fejlesztési tervezet.

A fejlesztési projekt mérföldkövei

1. Valósídejű prototípus létrehozása a vonatkozó fejezetben bemutatott jelfolyam diagram alapján.
2. Első implementáció VSTfx-ként, legfeljebb három felhasználási területre optimalizálva.
3. Ingyenesen hozzáférhető 32, illetve 64 bites rendszerre optimalizált VSTfx programcsomag előállítása.
4. Legfeljebb tíz felhasználási területre, adott mértékben szigorított tűrésekkel optimalizált programcsomag előállítása és kereskedelmi célú terjesztése.
5. Első hardwares implementáció prototípusa.
6. Autonóm működésű hardwares implementáció kereskedelmi célú terjesztése.
7. Softwares beépülő modulok előállítása különböző hardwares platformokra. (Több alprogramot tartalmazó és párhuzamosan működtető DSP rendszerek, multieffektek.)

Irodalomjegyzék

1. John, Chowning, "Az észlelés egysége és a hallásperspektíva", ford. Kovács Balázs 2009.
2. Microsoft (2014) <http://www-msp.ece.mcgill.ca/Documents/AudioFormats/WAVE/WAVE.html>

E. Fleischman, Microsoft Corporation, <http://tools.ietf.org/html/rfc2361>
Microsoft (2014) [http://msdn.microsoft.com/en-us/library/windows/hardware/dn653308\(v=vs.85\).aspx](http://msdn.microsoft.com/en-us/library/windows/hardware/dn653308(v=vs.85).aspx)
3. Gardánfalvi András, A fül természetes torzításának kutatása
4. http://www.ti.com/lscs/ti/arm/sitara_arm_cortex_a_processor/arm_cortex_a9_core/am437x_arm_cortex_a9/getting_started.page
5. Al Keltz, The "L" word - latency and digital audio systems