

Széchenyi István Egyetem
Műszaki Tudományi Kar

DIPLOMAMUNKA

Laczkovits Áron
Villamosmérnöki MSc szak

2013



SZÉCHENYI ISTVÁN EGYETEM
MŰSZAKI TUDOMÁNYI KAR
Távközlési Tanszék



DIPLOMAMUNKA

Hangszíntézisek

Laczkovits Áron

Villamosmérnöki MSc szak
Távközlési rendszerek és szolgáltatások szakirány

Győr, 2013

Feladatkiírás

Diplomamunka címe: Hangszintézis módszerek

Hallgató neve: Laczkovits Áron

Szak: Villamosmérnöki

Képzési szint: MSc

Típus: nyilvános

A diplomaterv keretében a meglévő szintézis módszerek csoportosítását, bemutatását, valamint e módszerek objektív összehasonlítását kell elvégezni. Minden csoportba reprezentatív hang szintézis technikákat választva. A kiválasztott szintézis módszerek a hozzájuk szorosan kapcsolódó szempontok szerint kerüljenek osztályozásra.

- Mutassa be a hang és szintéziseinek elméleti alapjait.
- Ismertesse a különböző hangszintézis csoportokat.
- Valósítsa meg a kiválasztott hangszintéziseket MATLAB programmal.
- Vizsgálja meg a szintézisek hatásfokát a paraméterek változtatásával.
- A számítási kapacitások és a hatásfok eredmények alapján határozza meg az optimális a paraméterek értékeit.
- Értékelje a kapott eredményeket és hasonlítsa össze különböző szintéziseket.

Győr, 2013-02-18

belső konzulens

tanszékvezető

Értékelő lap

Diplomamunka címe: Hangszíntézis módszerek

Hallgató neve: Laczkovits Áron

Szak: Villamosmérnöki

Képzési szint: MSc

Típus: nyilvános

A diplomamunka beadható / nem adható be: (a nem kívánt rész törlendő)

dátum

aláírás (belső konzulens)

A diplomamunka bírálatra bocsátható / nem bocsátható: (a nem kívánt rész törlendő)

dátum

aláírás (tanszékvezető)

A bíráló:

- Neve:
- Munkahelye:
- Beosztása:

A bíráló javaslata:

dátum

érdemjegy

aláírás (bíráló vagy tszv.)

A belső konzulens javaslata:

dátum

érdemjegy

aláírás (belső konzulens)

A ZVB döntése:

dátum

érdemjegy

aláírás (ZVB elnök)

Nyilatkozat

Alulírott, Laczkovits Áron, Villamosmérnöki, MSc szakos hallgató kijelentem, hogy a Hangszíntézisek című diplomamunka feladat kidolgozása a saját munkám, abban csak a megjelölt forrásokat, és a megjelölt mértékben használtam fel, az idézés szabályainak megfelelően, a hivatkozások pontos megjelölésével.

Eredményeim saját munkán, számításokon, kutatáson, valós méréseken alapulnak, és a legjobb tudásom szerint hitelesek.

dátum

hallgató aláírása

Tartalomjegyzék

1. Tartalom

FELADATKIÍRÁS	3
ÉRTÉKELŐ LAP	4
NYILATKOZAT.....	5
TARTALOMJEGYZÉK.....	6
BEVEZETŐ.....	9
1. A HANG STRUKTÚRÁJÁNAK ISMERTETÉSE	10
1.1. A hangszintézisről általánosan	12
2. ABSZTRAKT ALGORITMUSOK.....	16
2.1. Frequency Modulation Synthesis (Frekvencia modulációs szintézis)	18
2.2. Waveshaping (Hullámalakítás)	22
2.3. Karplus-Strong Algorithm Basics (Karplus-Strong algoritmus alapjai).....	25
3. MINTAVÉTELEZÉS ÉS FELDOLGOZOTT FELVÉTELEK	25
3.1. Sampling Synthesis (Minta-alapú szintézis).....	25
3.2. Wavetable (Hullámtábla), Vector (Vektor) Synthesis.....	26
3.2.1. Hullámtábla szintézis megvalósítása MATLAB programmal:	28
3.3. Granular Synthesis (szemcsés szintézis).....	33
3.3.1. Granulár szintézis megvalósítása MATLAB programmal:	36
4. SPEKTRÁLIS MODELLEK:	41
4.1. Spectral Modelling Synthesis (Spektrális modellező szintézis).....	42
4.2. Transient Modeling Sythesis (Tranziens modellező szintézis).....	42
4.3. Additive synthesis (Összegzéses szintézis)	43

4.3.1.	Additív szintézis megvalósítása MATLAB programmal:.....	44
4.4.	Subtractive Synthesis (Kivonó szintézis) – Analóg szintézis	49
4.4.1.	Szubtraktív szintézis megvalósítása MATLAB programmal:	50
4.4.2.	Szubtraktív szintézis egy emberi példán keresztül:	54
4.5.	Formant Synthesis.....	56
5.	FIZIKAI MODELLEZÉS:	58
5.1.	Finite Difference Methods (Parciális Differenciálegyenletek Numerikus Megoldása).....	59
5.2.	Modal Synthesis (Alaki szintézis).....	60
5.3.	Mass-Spring Networks – the CORDIS System (Tömeg-rugó hálózatok).....	61
5.4.	Numerikus módszerek összehasonlítása	61
5.5.	Digital Waveguides (Digitális hullámvezetés)	63
5.5.1.	Waveguide	64
5.5.2.	Waveguide alkalmazásai:	65
5.6.	Waveguide Meshes („Waveguide hálók”)	66
5.6.1.	2D Waveguide mesh:.....	67
5.6.2.	Banded waveguide:.....	68
5.7.	Single Delay Loop Models (Egyszerű késleltető húrok modellek).....	69
6.	ÉRTÉKELÉSI RENDSZER:.....	70
6.1.	A paraméterek használhatósága:.....	72
6.2.	A létrehozott hangok minősége és sokfélesége	73
6.3.	Végrehajtással kapcsolatos kérdések.....	74
7.	NÉHÁNY HANGSZÍNTÉZIS MÓDSZER ÉRTÉKELÉSE	76
7.1.	Absztrakt algoritmusok vizsgálata	76
7.1.1.	FM szintézis.....	76
7.1.2.	Waveshaping szintézis.....	77
7.1.3.	Karplus-Strong szintézis	78
7.2.	Mintavételezés és feldolgozott felvételek vizsgálata.....	78
7.2.1.	Mintavételezés	78
7.2.2.	Többszörös hullámtábla szintézis	79

7.2.3.	Granulár szintézis	79
7.3.	A Spektrális modellek vizsgálata	80
7.3.1.	Additív szintézis	80
7.3.2.	Szubtraktív szintézis	81
7.3.3.	Spektrális modellező szintézis	81
7.3.4.	Tranziens modellező szintézis	82
7.3.5.	Formant szintézis	82
7.4.	Fizikai modellek vizsgálata	83
7.4.1.	Véges differencia módszerek	83
7.4.2.	Modal szintézis	84
7.4.3.	CORDIS	84
7.4.4.	Digital Waveguides szintézis	85
7.4.5.	Waveguide Meshes	86
7.5.	Értékelések összefoglalása	86
7.6.	Megvalósított szintézisek összefoglalása CPU és RAM felhasználtságuk alapján	88
8.	ÖSSZEFOGLALÁS	89
9.	SUMMARY	90
10.	IRODALOMJEGYZÉK	92
11.	ÁBRAJEGYZÉK	97
12.	MELLÉKLETEK	99

Bevezető

A digitális hang szintézis módszerek olyan numerikus algoritmusok, amelyeknek célja, hogy zeneileg érdekes és lehetőleg valós idejű zenei hangokat állítsunk elő. A hang szintézisek bemenetei kizárólag vezérlő utasításokból állnak. Ezzel kapcsolatban számos különböző megközelítés érhető el.

A dolgozat célja nem az, hogy minden módszert részletesen bemutassak, hanem, hogy áttekintést adjak a számos különféle szintézis módszerekről, illetve részletesen összehasonlítsak néhány általam választott módszert. Ezen kívül célja az, hogy megállapítsam a felmerülő szintézis problémákat, hogy melyek a legalkalmasabbak egy adott módszerhez. Ezt a szintézis algoritmusok összehasonlításával végeztem.

A szintézis algoritmusok reprezentatív példák szerint lettek kiválasztva minden osztályban. Az egyes algoritmusokkal megpróbálok áttekintést adni a létező módszerekről.

Az összehasonlítási megközelítések Smith (1991) taxonómiáján alapulnak. Smith négy csoportra osztotta a digitális hang szintézis módszereket: absztrakt algoritmusok, feldolgozott felvételek, spektrális modellek és fizikai modellek. A dolgozat követi Smith taxonómiáját egy kissé módosított formában és minden kategóriában megvitatja a reprezentatív módszereket.

Minden módszer a következők szerint lett kategorizálva: absztrakt algoritmusok, mintavételezett és feldolgozott felvételek, spektrális modellek és fizikai modellek. Nagyobb hangsúlyt fektettem a mintavételezett és a jelmodell alapú szintézis módszerekre. A választott szintézis módszerek a 1 – 4. fejezetekben vannak bemutatva. Ezt követően az összehasonlítás kiértékelési szempontjai Jaffe (1995) által javasoltak. Egy további kritérium szerepel az eljárások alkalmasságára vonatkozóan: osztott és párhuzamos feldolgozás. Az összehasonlítási eredmények egy táblázatban vannak összegyűjtve, melyben az egyes módszerek minősítését lehet összehasonlítani.

1. A HANG STRUKTÚRÁJÁNAK ISMERTETÉSE

A hang a levegő érzékelt rezgése (oszcilláció), ami a hangforrás rezgéséből ered (pl.: gitár hang, hangszóró, hajszáritó, stb.) Az egyszerűbb rezgések összegének szempontjából periodikus rezgésnek tekinthető. Más szóval minden periodikus rezgés és az így keletkező hullámforma leírható harmónikusok összegével. Minden harmonikus egyszerű szinusz hullám a saját adott frekvenciájával és amplitúdójával (gyakran tiszta hangnak is nevezik) írható le. Az alábbi grafikon azt mutatja, hogy egy egyszerű, tiszta hang hogyan változik az időben:



1. ábra Tiszta hang (szinusz hullám)

Az alábbi grafikon egy bonyolultabb hullámformát mutat, amely magába foglal három tiszta hangot, különböző adott frekvenciával és amplitúdóval:



2. ábra Összetett hullámforma (tiszta hangok összege)

A harmónikusok relatív frekvenciája és amplitúdója adott időben meghatározza a hangszínt, azaz, hogy milyen hangot hallunk. A következő grafikon a vizuális értelmezését mutatja a harmonikus tartalomnak, mégpedig úgy, hogy az ábrázolt amplitúdót és egy adott hullámformát a frekvencia függvényében jeleníti meg:



3. ábra Harmónikusok a frekvencia tartományban

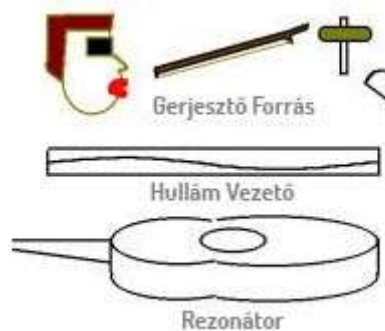
Figyelembe veszi, hogy a harmonikus tartalom időben változni fog, ha a hullámforma nem ismétlődik, vagyis ha nem periodikus. A szintézis módot ad a hangszínek szerkesztéséhez, hogy felülmúlja az eddigi hangszereket és hogy új hangokat állítson elő.

A hangszínek szintetizálásának különböző megközelítései:

1. Önmagának a hangnak az elemzése és közvetlenül próbált utánzása
2. A hangszer fizikai működésének az elemzése és modellezése

Akusztikus hang:

A legtöbb hangszer három részből áll:



4. ábra Gerjesztés, hullámvezető, rezonátor

1, *Excitation Source* (Gerjesztő forrás): Ez adja a rendszer energiáját, pl.: egy hegedűvonó, kalapács, vagy akár a zenész lélegzete.

2, *Wave-guide* (hullámvezető): ez a fő része a hangszernek, amely oszcillál (pl.: egy gitár húrja, vagy egy levegő oszlop a fuvalában). Mikor a rendszer folyamatosan periódikusan rezeg, olyan összetett hullámalakot alakít ki, melyek alaprezgésszám és felhang tekintetében fejezhetőek ki.

3, *Resonator* (rezonátor): Tipikusan egy hang doboz, harang vagy egy hangszer teste. Ez elsősorban felveszi az energiát a hullám-vezetőtől, ahhoz hogy hangot állítson elő. A rezonátor a saját maga módján fog rezonanciában oszcillálni a hullámvezetővel, így megváltozik a hullám-vezető rezgése és ez módosítja a hangszínt. Az összerezgő oszcilláló rezonátor tulajdonságait általában formánsnak nevezik. Ezek a formánsok fontos hatást gyakorolnak a hang minőségére.



5. ábra Folyamatábra a hullámvezetőtől a hangszer kimenetéig

A hang szintetizálásakor figyelembe kell venni a gerjesztés szintjét, a rezgést/harmonikust és azt, hogy a rendszer hogyan szabadítja fel az energiát, hogy hallható hangként halljuk.

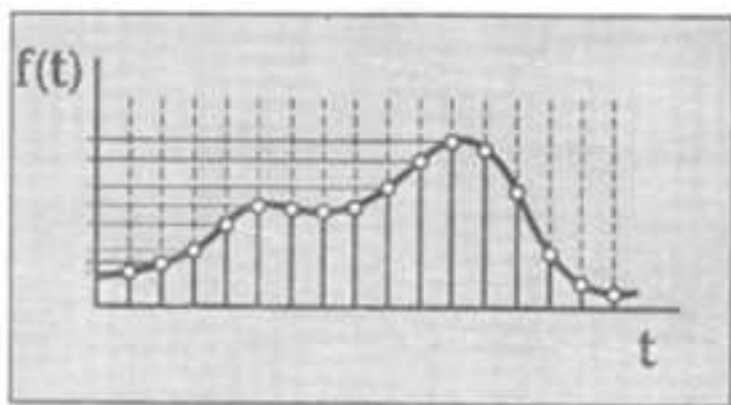
1.1. A hangszintézisről általánosan

Elektronikus módszerekkel alapvetően két eljárással hozhatunk létre hangot: Vagy rezgéskeltők (oszcillátorok), illetve a hozzájuk kapcsolódó elektronikus áramkörök alkalmazásával állítunk elő hangokat vagy a már meglévő hangeseményeket dolgozzuk fel. A gyakorlatban mindkettőt használjuk.

A rezgéskeltők és az elektronikus áramkörök alkalmazása a hangszintézis, amelyek során az oszcillátorokat úgy vezérlik, hogy megfelelő hangokat állítsunk elő. A hangzás paramétereinek a beállításával előállítható egy hangprogram. A hangszergyártók gondolva az egyszerű felhasználókra, akik nem képesek több száz paramétert beállítani a hangszerekbe írnak be ilyen hangprogramokat. Ez azt jelenti, hogy egy hangon sok minden megváltoztatható, új hangok is létrehozhatóak, de sajnos hátránya, hogy mivel nincs beadva több száz paraméter, így nem hasonlítanak a klasszikus hangszerek hangjaira. Ez a módszer

tehát arra jó, hogy a műkedvelők számára megfelelő szintézist biztosítson, de nem felel meg azok számára akik hangszerek szerinti akusztikus jellegű megoldást alkalmaznak szívesen.

Az akusztikus jellegű hangszerek számára, a másik módszer alkalmas, ami digitalizált hangokat használ. Ezeket a hangszer hangokat hangminta könyvtárakból lehet kiválasztani. Amennyiben a hangminta könyvtárakban nincs olyan hang ami számunkra szükséges, akkor mikrofonnal mi is rögzíthetjük az adott zenei hangot. A kész hangmintát belehet tenni a hangminta lejátszó rendszerbe és ott különböző effektek, szűrők segítségével feldolgozhatóak.



6. ábra Mintavételezés

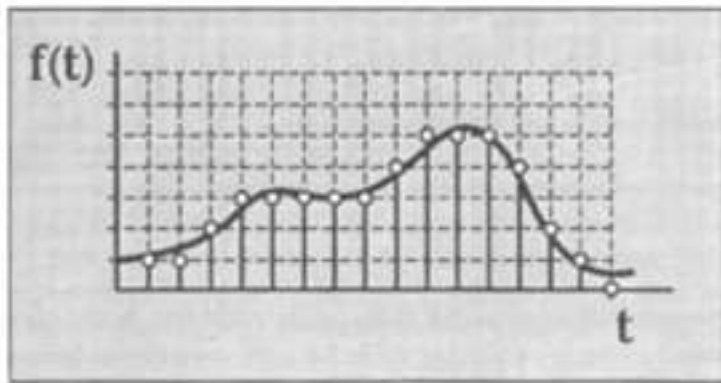
A hangdigitalizálás során a hangot számokká alakítjuk és tároljuk. Ez úgy történik, hogy egy mikrofonnal felvesszük a hangszer jeleit pl.: egy egyszerű digitalizálóval. A digitalizáló

tulajdonképpen egy forgó tárcsa nélküli közönséges magnóként dolgozik. A mikrofon a levegő rezgéseit elektromos feszültséggel változásaivá alakítja, majd adott időnként mintát vesz, azaz rögzíti az elektromos rezgés pillanatnyi értékét. Ez a mintavett hangszelet olyan feszültségérték, ami egy számmal írható le. A számokat rögzítjük és kész is a digitalizált hang. Minél gyakrabban vesszük a mintákat annál élethűbb a hang.

A CD-nél ez az intervallum 22,6757 mikro szekundum, ami 44100 Hz-nek felel meg. Ez a másodpercenkénti hangminták száma. Miután 20 és 20000 Hz közötti frekvencia tartományban érzékeli az emberi fül a hangokat, Mr. Shannon rájött, hogy a 44100 Hz-es mintavételi frekvencia a megfelelő, mert egy hangból két mintát kell venni egy teljes rezgés alatt, hogy élethűen és veszteségmentesen vissza tudjuk játszani. A 44100 Hz-nek történeti okai vannak: az U-matic szabványos videomagnók képrögzítő képességét erre a frekvenciára alkalmazták.

A hangkártyáknál és különböző stúdió eszközökkel 96 vagy 192 kHz-es a mintavételi frekvenciát is alkalmaznak, így az ember élethűbbnek érzi a hangzást, mivel a 20000 Hz

feletti felharmónikusok is megmaradnak, annak ellenére, hogy nem hallhatóak, de mégis érezhetőek.

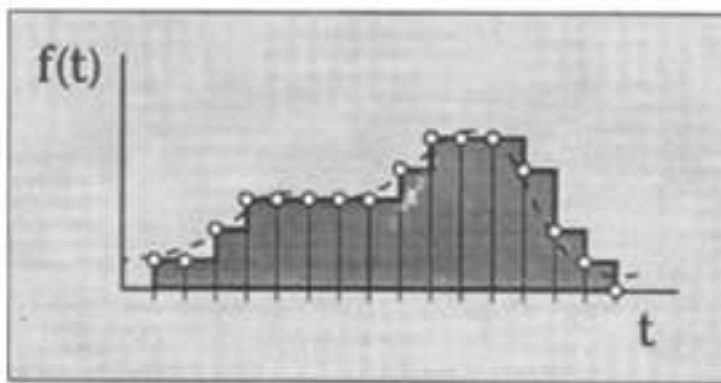


7. ábra Kvantálás

A mintavett jelek nagyságát úgy alakítjuk számokká, hogy megadjuk az adott elektromos jelhez tartozó feszültséget pl.: milivoltban. A kvantálás az a folyamat, amikor különböző

jelszintekhez különböző számokat rendel hozzá a mintavevő berendezés. A kvantálás (kerekítés) pontosságát meghatározza a függőleges tengelyen lévő négyzetháló sűrűsége. A függőleges tengely felosztásának az egysége a bit. A bitszám növelése a hangminőség fokozódását mutatja.

A CD-ben 16 bites a felbontás, a stúdió eszközökben már akár 24 bites és egyes eszközöknél ennél is nagyobb felbontást alkalmaznak. A ma alkalmazott 16 bit 65536 vonalbeosztást, a 24 bit 16 milliónál is több vonalbeosztást tartalmaz. Feladatunk, hogy hogyan játszunk le két egymást követő mintavett jelet.



8. ábra Lépcsők

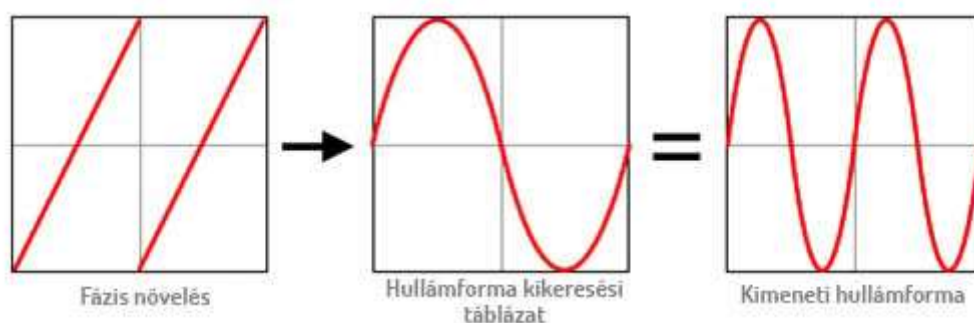
1. Megoldás: a jelet az előző szintet tartjuk, mindaddig amíg nem jön a következő jel. Ettől kissé lépcsős lesz a jel. Ez a lépcsős jel még mindig jobb, mintha éles tüskés jelet kapnánk,

hiszen a lépcsős jel még mindig jobban hasonlít az eredetihez. A lépcsők élei pedig szűrővel lekerekíthetőek.

2. Megoldás: a meglévő jelből extrapoláltjuk a következő jelet. Amennyiben helyes az extrapoláció, akkor kedvező lesz az eredmény, ha nem túl szerencsés az extrapolálás, akkor kedvezőtlenebben alakul az eredmény, mint az előző esetben.

Vagy az egyes vagy a kettes eset szerint digitalizáltuk a hangot és a számsorozatot eltároltuk a mintavevő vagy a számítógép memóriájában. Lejátszáskor visszafelé kell csinálnunk vagy egyes vagy a kettes módszert. Mivel digitális technikát alkalmaztunk, a felvételtől adódó minőségromlásra kívül soha nem lesz minőségvesztés akárhányszor játszunk le a hangot.

Mielőtt elkezdeném a hangszintézisek tárgyalását elsősorban meg kell, hogy értsük a *Direct Digital Synthesis* oszcillátor alapelvét. Ez a legalapvetőbb típusú digitális oszcillátor. Egy nagy számlálót használ, amit *Phase Akkumlátor*-nak nevezünk. Ez a számláló minden mintánál adott értékkel növekszik (*Frequency Increment*). A számláló kimenete egy index, ami a hullámforma kikeresési táblázatán alapul.



9. ábra Folyamatára a kimeneti hullámformáig

CPU és RAM mérő program bemutatása:

Gyakran felmerül a kérdés, hogy hogyan mérhetjük az aktuális MATLAB folyamatunk CPU és RAM felhasználását. Ezt valójában nagyon egyszerűen és egyértelműen meg is tehetjük ezt MATLAB-on belül. Ezzel minden alapvető információt kinyerhetünk a Windows-os gépünkről, amely teljes körű ismeretet ad a teljesítményéről. A részletes információk a következők:

- Gép neve, teljes fizikai memória, processzor (CPU) információ
- CPU használat százalékban
- Memória használat MB-ban

Ezt a mérő programot úgy használhatjuk, hogy minden alkalommal, amikor meghívjuk megméri és tárolja a CPU, illetve a memória felhasználást. A mérő programot úgy használhatjuk, hogy legelőször meghívjuk a 'startRecordPerformance' függvényt, majd azt csinálunk MATLAB alatt amit mérni szeretnénk. Ez jellemzően egy folyamat vagy

szimuláció. Ha ezzel végeztünk nincs más dolgunk, mint a 'stopRecordandDisplay' függvényt meghívni a programunk után és megkapjuk a CPU és RAM felhasználtságát egy grafikonon. Természetesen a folyamatunk során ne használjuk a *clear all* parancsot. Az eltárolt adatok törléséhez használjuk a 'delete(perfData), delete(t)' parancsot. Az alábbi rendszert leíró információkhoz a következő függvényt készítettem el MATLAB-ban:

```
classdef SysInfoData < handle
    % tulajdonságok, amelyek meg lesznek jelenítve
    properties (SetAccess = 'private') []

    % tulajdonságok, amelyek nem lesznek megjelölve
    properties (SetAccess = 'private', GetAccess = 'private') []

    methods
        function obj = SysInfoData() % szerkesztés [] % szerkesztés

        function reset(obj) [] % alaphelyzetbe állítás

        function measure(obj) [] % mérések

        function data = getTimeArray(obj) [] % megkapjuk az idő tömböt

        function data = getUsedCPUArray(obj) [] % megkapjuk a felhasznált CPU tömböt

        function data = getUsedMemoryArray(obj) [] % megkapjuk a felhasznált memória tömböt

    end
end % osztályok definiálása

% funkciók
function cpuSpeedStr = getCpuSpeed()
    % a CPU sebességét MHz-ben kapjuk meg
    cpuSpeedMHz = winqueryreg('HKEY_LOCAL_MACHINE', 'HARDWARE\DESCRIPTION\System\CentralProcessor\0', 'MHz');
    % GHz-re konvertálás
    cpuSpeedGHz = double(cpuSpeedMHz)/1000;
    cpuSpeedStr = sprintf('%2.2fGHz', cpuSpeedGHz);
end
```

10. ábra CPU és RAM mérőprogram

2. ABSZTRAKT ALGORITMUSOK

Az 1920-as években néhány zeneszerző változtatható sebességű lemezlejátszóval kísérletezett a koncertjeiken, mint például *Milhaud* és *Toch*, melyek a számítógépes zene elődeiként is értelmezhetőek. 1950-ben *Pierre Schaffer* megalapította a *Studio de Musique Concrète*-t Párizsban. A *Musique Concrète*-ben a zeneszerző hang elemek kinyerésével foglalkozott felvételekből, illetve valódi hangokból.

A bemutatott módszerek ebben a fejezetben az absztrakt algoritmusok, illetve a valódi hangok felvételein alapulnak. Smith szerint ezek a módszerek kevésbé válhatnak gyakorivá a kereskedelmi szintetizátorokban a nagyobb teljesítményű és a felmerülő kifejező technikák miatt. Ezek azonban, még mindig hasznos háttérrel szolgálnak a bonyolultabb hang szintézis

módszerekhez, különösen, hogy még mindig kiemelkedőnek bizonyulhatnak néhány hang szintézis problémák esetén.

Distortin Synthesis (Torzításos szintézis)

A torzításos szintézis, egy olyan hangszintézis technika, mely a már meglévő hangokat módosítja, hogy összetettebb hangokat (vagy hangszíneket) állítsanak elő. Ezt leggyakrabban nem lineáris áramkörökkel vagy matematikailag valósíthatunk meg. Ide tartozik az *FM* szintézis, a *hullámalakító* szintézis és az *összegzéses formulát* használó szintézis.

A korai digitális időkben a legfontosabb fejlesztések közé tartozott számos torzítás alapú módszer. A módszer nagy teret hódított magának a kutatásokban és a digitális zene gyakorlatában. Néhány kiemelkedő példa:

John Chowning, aki nagy úttörőnek számított a Frekvencia Moduláció (FM) szintézisben;

Godfrey Windham, Ken Steiglitz és Andy Moorer együttesen dolgoztak a *Diszkrét Összegzéses Formulán* (DSF);

Daniel Arfib és Marc Leburn közreműködött a *Digitális Hullámformánál*.

Ezek a különböző technikák ugyanabból az alapelvből erednek. Ezen a területen a *Phase-Align Format* (PAF) szintézis fejlesztése az 1990-es években volt.

A torzításos szintézis választ adhat a legfontosabb fennálló problémára; hogy hogyan lehet létrehozni egy összetett, időben változó spektrumot, ami diszkrét komponensekből áll. Az alapvető megoldási módszerek:

A *brute force* módszer: egy szinusz hullám oszcillátort és néhány burkoló (amp., frek.) funkciót használ részenként, majd összekeveri az összes forrást (*Additív szintézis*).

Az *elegant solution* módszer: megtalálja a módját, hogy néhány egyszerű forrást kombinálva (pl.: szinusz hullám, oszcillátor) sok komponens létrehozunk.

A fent említett utóbbi módszer, a különféle torzításos algoritmusokkal biztosítja a komponensek létrehozását.

Matematikailag egy $s(t)$ jelet szeretnénk létrehozni, ahol adott az a_k amplitúdó, a $w_k (=2\pi f)$ frekvencia és az Φ_k fázis-eltolás:

$$s(t) = \sum_{k=0}^{N-1} a_k \cos(\omega_k t + \phi_k)$$

A torzításos szintézis előnyét, leginkább az ebből kifejlesztett új szintézis algoritmusokban és hallható hatásokban élvezhetjük. Hátrányát itt is a módszer matematikai hátterében, és összetett, bonyolult számításokban találjuk.

2.1. Frequency Modulation Synthesis (Frekvencia modulációs szintézis)

Dr. Chowning volt az első, aki felfedezte és használta a frekvencia modulációs szintézist. Ennek alapja, hogy a vivő hullámot legalább egy operátorral moduláljuk. Ezáltal, olyan zenei hangokat állítunk elő, hogy folyamatosan gyorsan változtatjuk a hang alaphékvenciáját.

A frekvencia modulációs technika alapvetően már rég óta használatos a rádiózásban. A frekvencia modulációhoz alapvetően két oszcillátorra van szükségünk. Az egyik oszcillátor segítségével a második frekvenciáját moduláljuk, a modulált második oszcillátor kimenetét pedig a lehallgatásra használjuk. Leggyakrabban szinus hullámot használunk, de a Modulátor nem csak ezekre a hullámokra van korlátozva. Bármilyen hullámformát használhatunk, sőt még hangokat is küldhetünk az audio bemenetre. Minél összetettebb a modulációs hullámforma, annál komplexebb lesz a végeredmény is, emiatt a sokrétű kapott hullámformák zajnak tekinthetők. A fentiekből következik, hogy a frekvencia moduláció nem végződhet csak zajjal.

A frekvencia modulációnak két típusát különböztetjük meg. Az egyik típusba tartoznak azok, ahol a frekvencia modulációban az aktuális frekvenciát változtatjuk meg. A másik fő típus a *Phase Modulation*, amiben a frekvencia paraméter fix, de a hullámforma fázisa változtatható. Gyakran találkozunk az *FM* betűösszetétellel, bár a technika, ami mögötte áll, az valójában a *PM*. Két előnye van a *PM*-nek az *FM*-mel ellentétben: az első, hogy lehetővé teszi egy oszcillátor ön maga modulációját anélkül, hogy a kapott hullámforma alaphékvenciája megváltozna; a másik, hogy a kapott hullámforma alaphangja nem mozdul el, ha a moduláló hullámforma DC komponenszt tartalmaz, ami azt jelenti, hogy a hullámforma átlagos energiája enyhén pozitív vagy negatív, a nulla helyett.

John Chowning a 1960-as évek közepén egy nem tervezett esemény következtében váratlanul fedezett fel egy teljesen új hang szintézis formát. Az új hangszintézis felfedezése közben a hang rezgéseivel kísérletezett a Stanford Egyetemen. Chowning rájött, hogy amikor a modulált jel frekvenciája egy bizonyos ponton túl nő, az úgynevezett *vibrato* hatás eltűnik a modulált jelből, és egy új komplex hangra cserélődik, ekkor az eredeti jel lényegében eltűnik. Ma már látjuk, hogy a tudós rábukkant a mai legelterjedtebb kódolási technikára a nyilvános rádiós átvitelben (FM rádió). Hamar rájött, hogy az FM nagyon hatékony szintézis módszer. 1966-ban Ő volt az első, aki egy zenéből egy szeletet az új szintézis módszerével és a hang generálás kizárólagos használatával állított össze és rögzített zenei részletet.

A következők állíthatók a modulátor és a kimeneti jel kapcsolata között:

A Modulációs Index a spektrális komponensek lényeges számát és amplitúdóit határozza meg, ami arányos a modulátor amplitúdójával, de fordítottan arányos a modulátor frekvenciájával:

$$\beta = \frac{\Delta W_c}{W_m}$$

A kifejezés számlálója a változó vivő frekvenciát mutatja, ami azt jelenti, hogy a β közvetlenül kapcsolódik a modulátor amplitúdójához. Ez az a pont, ahol bármely adott modulátor frekvencia értéke a modulációs index (a modulátor amplitúdója), amely minden egyes komponens amplitúdóját meghatározza, a kimeneti jel spektrumában.

Bármely adott vivőfrekvencián, egyedül a modulátor frekvenciája határozza meg a spektrális komponensek pozícióját.

A módszer alkalmazásának rövid leírása:

Egy egyszerű szinuszos jel harmonikus eloszlását, egy másik szinuszos jellel modulált, Bessel-függvényekkel reprezentálhatjuk – ez adja az *FM* szintézis matematikai megértésének egyszerű alapját.

Az *FM* szintézis értelmezhető egyfajta *torzítás*- vagy *nemlineáris* szintézisként. Ugyanis egy oszcillátor, egy hangfrekvenciás hullámformájú vivőt generál, majd a kapott frekvenciát változtatja meg (vagy modulálja) egy moduláló frekvencia segítségével.

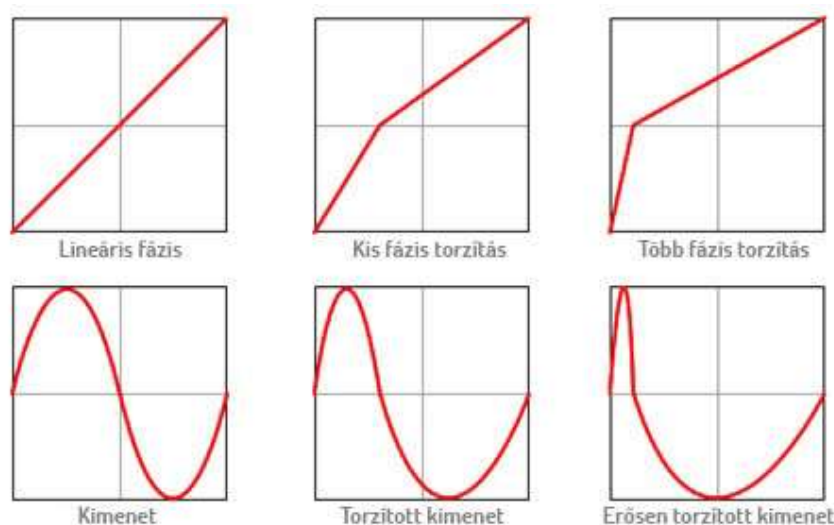
Előnyét elsősorban az adja, hogy frekvencia modulált értékéből megállapítható lesz a jel kimeneti pozíciója. Ugyanis a frekvencia moduláció igen hatékony szintézis módszer (digitális- és analóg szintetizátorokhoz is használható), képes hangok előállítására. Más módszerekkel ez elérhetetlen.

Hatalmas hátránya a többi módszerrel szemben, főleg a zenészek számára, hogy az összetett matematikai háttérének megoldása (feltételezés és bizonyítás kalkulációkkal) bonyolult feladat is lehet, habár az alapelv könnyen értelmezhető marad. További hátránya, hogy a spektrális komponensek pozíciója kizárólagosan a modulátor frekvenciájával határozható meg.

Phase Distortion Synthesis (Fázistorzításos szintézis)

Fázis Torzítás (PD), olyan szintézis technika, amely szinusz függvény *kikeresési táblázatán* alapul, egy nem-lineáris leképezési *modulo* számlálót használva. Igazolható, hogy a *PD* valójában egy részhalmaza a Fázis modulációs (*PM*) szintézisnek; szokásos végrehajtási módszere a Frekvencia modulációnak (*FM*) a hardver szintetizátorokban, mint például a Yamaha DX sorozat A Virtuális Analóg (*VA*) oszcillátorok kialakításához is felhasználták ezt a szintézist.

A fázistorzításos szintézis működésének alapja, hogy megváltoztatja a fázis akkumulátor függvény alakját. Ha a fázis akkumulátor függvény nem lineáris, a kimeneti hullámforma is torzul. A nem lineáris hullám harmonikus változásokat fog eredményezni a hangban. A következő példa egy egyszerű fázis torzulási függvény és egy szinusz hullám folyamatát mutatja.



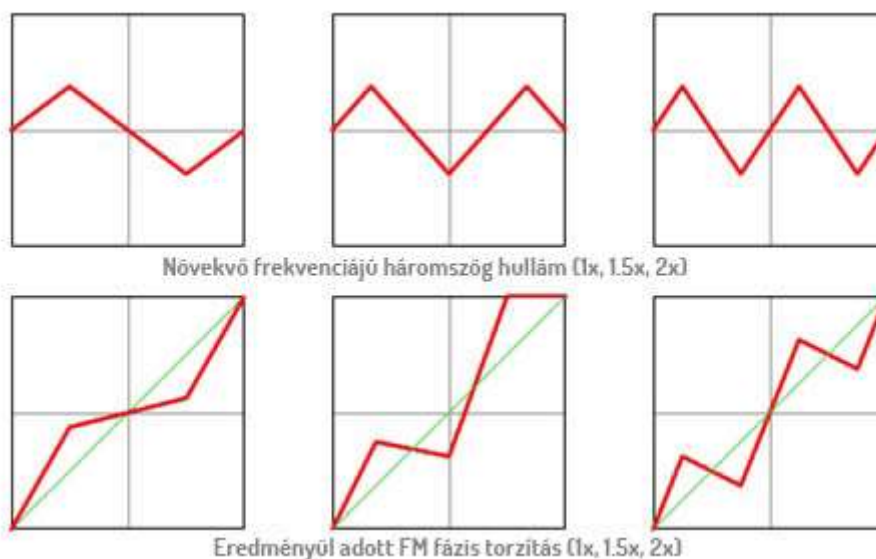
11. ábra Fázis torzulási függvény és szinusz hullám működése

Felmerül a kérdés, hogy mi a különbség a *PD* és az *FM/PM* között. Lényegében nem sok. Ugyanis a fázis torzításos szintézis egy speciális esete az általánosabb fázis modulációs szintézisnek, ahogyan például a Yamaha használta az *FM* szintetizátoraiban. A fázistorzításos és a modulációs szintézis tehát egyenértékű, ami az alábbi ábrán látható, melyen a fázistorzítási függvényt tekinthetjük úgy, mint háromszög hullám hozzáadását az alap fázis akkumulátorhoz:



12. ábra Fázistorzítási függvény

A *PD* sokkal korlátozottabb, mint az *FM*, mert mindig modulátorokat használ. A modulátorok ugyanazon a frekvencián vannak, mint a vivők, vagy csupán olykor a frekvencia egyszerű többszörösei. Növelve a háromszög hullám frekvenciáját egy összetett hatást hoz létre a fázistorzításon (ez lehetséges *FM* szintetizátoron).



13. ábra FM Fázis torzítás

A PD szintézis érdekessége, összehasonlítva a Yamaha FM-jével, hogy a modulátor hullámalakja nem egy egyszerű szinusz görbe, és tovább módosítható.

A Casio CZ sorozatú szintetizátorokon mutatták be először 1984-ben, ahol a tipikus szubtraktív szintézis jelfolyammal versengett, ami forrás-szűrő vezérlőkből állt, így egy szabályozottabb gondolati eljárás-mód is elérhető lett. A PD nem lett alaposan kidolgozva a jelfeldolgozási témakörben, azonban továbbra is használatos maradt.

Az alapelv szerint, szinuszos függvény és adott fázis modulo-val $\varphi(t)$, a következő:

$$x(t) = -\cos(2\pi\varphi(t))$$

Komplex harmonikus spektrumot hoz létre, ha a fázisa egy nem-lineáris függvénnyel $x(t)$ van kialakítva:

$$x(t) = -\cos(2\pi f(\varphi(t)))$$

Az $x(t)$ alakjától függően, különböző spektrumok állíthatóak elő.

Ha $x(t)$ lineáris, akkor nincs torzítás és tiszta szinuszos hangot kapunk. Így a PD-re tekinthetünk úgy, mint egy fázis-alakító formára, illetve a hullámformálás analóg módon nem-lineáris amplitúdójú torzítási módszerére. Az eredeti kivitelezésében, a PD szintézist szakaszonként használtjuk, lineáris függvényeknek kettő vagy több szegmensét, hogy végrehajtsuk a nem-lineáris leképezést.

Legfontosabb előnye, hogy *all-pass* filtert alkalmaz, ami jól ismert módszer, hogy alacsony késleltetés és fázis korrekció jusson a jelekbe.

Hátránya a kidolgozatlanságának köszönhető, ugyanis bonyolult matematikai összefüggéseket és számításokat alkalmaz.

2.2. Waveshaping (Hullámalakítás)

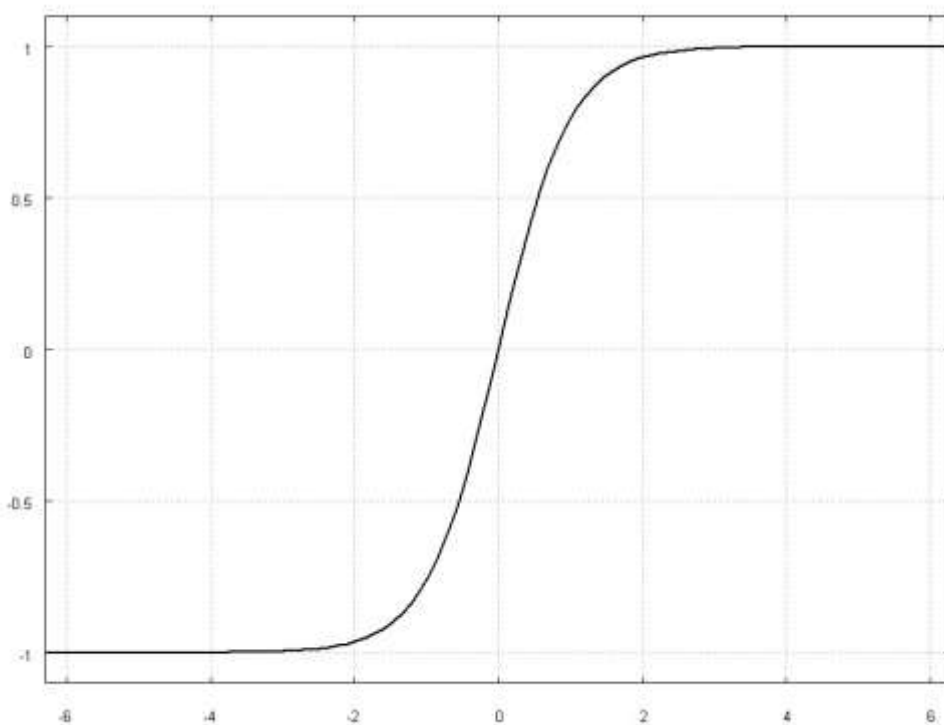
A *hullámalakításos* technika egy jel amplitúdójának nem lineáris torzításán alapul. A hullám bemeneti formája leírható, általában egy egyszerű szinuszos jellel, és egy olyan függvénnyel, ami átalakítja a kívánt kimeneti hullámformára. A legelterjedtebb módszer az úgynevezett átviteli függvény, ami a spektrális polinom egyezés révén állapítható meg.

A fő előnye ennek a megközelítésnek, hogy a polinom függvény pontosan elkészíti a megfelelő sávkorlátozott spektrumot, egy adott szinusz jel adott amplitúdójához.

Hátránya azonban, hogy a polinomok hajlamosak, hogy amplitúdóban változó természetellenes hangzást hozzanak létre, ha csupán időben változó spektrum felel meg az elvárásainknak. Újabb kutatásokkal igazolták, hogy bizonyos hullámalakokhoz hasonló függvények előnyös tulajdonságait is ki tudjuk használni, akár a trigonometrikus, hiperbolikus függvények repertoárját, melyek közül néhány egyenletes spektrális változásokat szolgáltat. Ezeknek az ötleteknek a hasznos alkalmazása az analóg szintetizátor oszcillátorok modelljeiben van jelen (*Virtual Analogue Models* néven is ismertek ezen a kutatási területen).

Hyperbolic Tangent Waveshaping (Tangens hiperbolikus hullámalakítás)

Bemutatásának egyszerű módja, hogy generálunk egy nem sávkorlátozott négyszög hullámot a *Signum* függvény használatával, a változó bipoláris bemeneteket leképezve. Ez a szakaszos függvény 1-et ír minden nem nulla, pozitív bemeneti értékekre, 0-át egy null bemenetre és -1-et minden negatív argumentumba. Összegezve tehát, levágja a jelet, de ennek során sok összetevőt generál a *Nyquist*-en felül, amelyek megfelelően rejtettek maradnak. Ennek a legfőbb oka, hogy 0-nál nem folytonos, ahol a kimenet teljesen negatívból teljesen pozitívba mozdul.



14. ábra Tangens hiperbolikus függvény

A tangens hiperbolikus egy olyan függvény, amelyet a *Signum* függvény helyett alkalmazhatunk, ugyanis egyenletes jel átvitel van egy adott pontnál, de a levágási tulajdonságokat is pontosan őrzi.

Assymetrical FM Synthesis (Aszimmetrikus FM szintézis)

Az *FM* technika és bizonyos mértékig a fázis moduláció (*PM*) jól ismertek, igazán pontos tulajdonságait kimerítették a tudósok. Számos, régebben ismeretlen módszerrel nem törődtek, mint például a majdnem feledésbe merült Palamin által kifejlesztett Aszimmetrikus *FM* szintézis módszerrel. Palamin megfogalmazásában az eredeti *FM* modell (vagy *PM*) egy exponenciális jel által alkotott *ring-modulated* lesz. Ennek az a hatása, hogy egy új paramétert (r) bevezetve szabályozza a spektrális szimmetriát, amely lehetővé teszi a csúcsoknak, hogy a vivőfrekvencia fölé vagy alá mozduljanak el. A kifejezés, a normalizáló tényezővel ($\exp(0,5k(r-1/r))$) a következő:

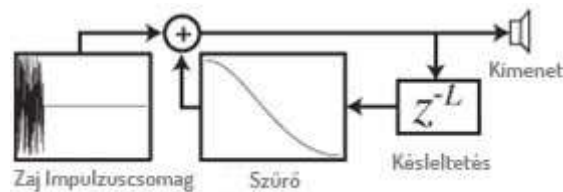
$$s(t) = \exp\left(0.5k\left(r - \frac{1}{r}\right)\cos(\omega_m t)\right) \sin(\omega_c t + 0.5k\left(r + \frac{1}{r}\right)\sin(\omega_m t)) = \sum_{n=-\infty}^{\infty} r^n J(k) \sin(\omega_c t + n\omega_m t)$$

Az új paraméter az r , ami a szimmetria szabályzó ahol, ha $r < 1$, akkor a ω_c vivőfrekvencia alá húzza a spektrális csúcsot, ha pedig $r > 1$, akkor a vivőfrekvencia fölé nyomja.

A szimmetria szabályzás nagyon hasznos és jó tulajdonság, ahol a kevés összetett és néhány extra kikeresési tábla rovására hozzáadódik a koszinusz- és exponenciális értékekhez. A fentiek értelmében állíthatjuk, hogy egy hullámalakító kimenetének *ring* modulációja az exponenciális átviteli függvényt használja, és az *FM* jel értékeiből tevődik össze. Hatalmas előnye még, hogy bármely két torzítási technika összekötése lehetségessé válik.

2.3. Karplus-Strong Algorithm Basics (Karplus-Strong algoritmus alapjai)

Egyszerű algoritmus, mely zenei hangokat készít zajból. Egy zaj rövid töredékéből csinál egy hurkot, egy szűrt késleltetőn keresztül, hogy egy leütött hangot, pengetett húrt vagy bizonyos ütős hangszereket szimuláljon, hasonlóan egy *comb* szűrőhöz (z-transzformációs analízis). A hullámtábla módosítású algoritmus legegyszerűbb osztályának is tekinthető (ahol a hanghullám periódikusan olvasott a számítógép memóriájából), mely *Digital Waveguide* szintézis néven ismert, amiben a késleltető hat a jel egy periódusára. Lényegében egy *szubtraktív* szintézis technika, amely visszacsatolt hurkon alapul.



15. ábra Karplus-Strong algoritmus blokkvázlata

3. MINTAVÉTELEZÉS ÉS FELDOLGOZOTT FELVÉTELEK

3.1. Sampling Synthesis (Minta-alapú szintézis)

Egy olyan *audio* szintézis technika, amely bemintázott hangokat vagy hangszereket használ új hangok előállítására. A *minta alapú* szintézist nem szabad összetéveszteni a *hullámtábla* szintézissel. A korai mintavételezők igencsak memória korlátozottak voltak, ezért a lehető

legrövidebb mintákat használták. Ciklus, illetve hangmagasság változtatás elérése által lehetőség volt a minták hosszának a növelésére, a mintákat megnyújtva számos hangnemen keresztül. A későbbi mintavételezők sokkal több memóriát és tároló képességet biztosítottak, megengedve a *sound design*eknek, hogy *multisample*-öket alkalmazzanak, hogy sokkal valószerűbb változásokat nyújtsanak dinamikában és hangszínen. Sok gyártó sokféle név alatt kínál minta alapú szintetizátorokat, például Korg's HI (Hyper-Integrated), Yamaha's AWM (Advanced Wave Memory. A szoftver eszközök, mint például a TASCAM GigaStudio és Native Instruments Kontakt szintén sokféle minta-szerkesztési funkciót nyújtanak.

A minta-alapú szintézis fő előnye az egyéb digitális szintetizáló módszereken túl, mint a *fizikai modellezéses* szintézis vagy az *additív* szintézis, hogy sokkal alacsonyabb feldolgozási teljesítményt igényel. Ez azért van, mert a legtöbb hang modell árnyalatait tartalmazzák az előre felvett minták, nem úgy mintha valós időben számolnánk őket.

Az analóg szintetizátorokkal szemben az áramköröket nem kell sokszorozítani, hogy egyszerre több hangot játszunk le. Ezért a minta-alapú gépek polifóniája általában sokkal magasabb.

Hátránya azonban amiatt, hogy „gazdagabb” legyen, egyszerre több minta szükséges a lejátszáshoz (egy trombita tartalmazhat lélegzet zajt, morgást és hurkolt hanghullámot a folyamatos lejátszásnál). Ez újra csökkenti a polifóniát, ahogy a minta-alapú szintetizátorok megbecsülik azt, ami a minták számán alapul, hogy egyszerre lejátszhatóak legyenek.

3.2. Wavetable (Hullámtábla), Vector (Vektor) Synthesis

A *hullámtábla* szintézis megértése nagyon egyszerű. A korai szintetizátorok oszcillátorai csak korlátozott számú hullámformát voltak képesek szolgáltatni, amelyek nem meglepően a szinusz, háromszög, fűrész és a négyszögjel vagy impulzus. Ezekkel a hullámformákkal meglepően az alapvető hangoknak széles választékát játszhatjuk le, különösen ha különféle hullámformákat különböző módokon kombinálunk. Az 1970-es években már léteztek olyan hullámtábla szintetizátorok, amikben már a három vagy négy hullámforma helyett 64 hullámforma állt rendelkezésre, melyek digitálisan tárolva voltak egy úgynevezett hullámtáblába és ezeket a kezelő panelen megkönnyítették a kezelést. Így már rendkívül széles palettája nyílt meg a hanghullámoknak, melyekkel a saját hangjainkat megalkotjuk.

Előnye ennek (azon kívül, hogy 60 plusz hullámformánk van), hogy használhatunk LFO-t, burkoló görbét vagy MIDI vezérlőt ezeken a hullámformákon. Ha ezeket a hullámformákat okosan rendezzük képesek vagyunk a harmónikusok mozgását megalkotni a hangban. Például ha az első hullámunk szinuszos hullám és a 64. erős négyszög hullám és a másodiktól a 63.-ig fokozatosan egyre erősödik, mint extra harmónikusok hozzáadása a hullámtábla minden lépésében. Ahogyan mozgunk a hullámtáblán, úgy közeledünk egy hagyományos *sweep* szűrőhöz. Azonban ennek van hátránya is, mert a *sweep* nem lesz sima. A mai szintetizátorokban akár 32 hullámtábla is lehet, mindegyik 64 hullámformával.

A *hullámtábla* szintézis megmutatja az egy függvényben számolt egymástól független változókat, amelyek egymást követő memóriahelyeken vannak tárolva (számítógépen). A hallható hanghullám mintáit ez a memória darab (a hullám tábla) tárolja, majd beolvassa, vagy a periodikus *ütemezett* függvény segítségével hozza létre. A *hullámtábla* szintézist a Waldorf és a PPG hangszerek használták először.

Előre kiszámolja a hullám spektrum inverz diszkrét Fourier transzformáltját, mielőtt azt visszajátszaná, ahelyett, hogy valós időben számolná az inverz DFT-t, mint az additív szintézis. A Roland LA (lineáris aritmetikus) szintetizátorok (D-50) ilyen elven működnek. Ezekben a szintetizátorokban azonban az összetett minták *attack* fázisai kombinálva vannak az egyszerű *sustain*-el vagy *decay* fázisaival, hogy létrehozzák a hangot. Lényegében ez egy egyszerű hullámtábla, ami két mintából áll. Amiben az LA és a hullámtábla szintetizátorok különböznek azaz, hogy az utóbbit új, eredeti, digitális hangok készítéséhez tervezték.

Hasonlóság van az egyszerű digitális szinuszos hullám létrehozásánál, vagy az *additív* szintézisnél, de minimum két szempont figyelembevétele szükséges:

A hullámforma kikeresési tábla: nem csak egy egyszerű szinuszos függvény periódus mintáit tartalmaz, hanem sokkal egyszerűbb és általánosabb hullámalakokat is.

A mechanizmus: a zenei hangok kialakulásával, a dinamikusan változó hullámalak miatt egy kvázi-periodikus időfüggvényt hoz létre.

A *hullámtábla* szintetizátorok egy hangszer hangjának két részét tárolják:

Egy minta *attack* szekcióját (pl: zongorabillentyű leütésének a hangja egy kalapács ütéssel).

Egy szeletet a hangszer hangjának a kitartási részéből.

változtatni. Az összefűzendő hullámok számát tekintve bármennyit képesek vagyunk ezzel a módszerrel megtenni, ami természetesen csak a számítógépünk memóriájától függ.

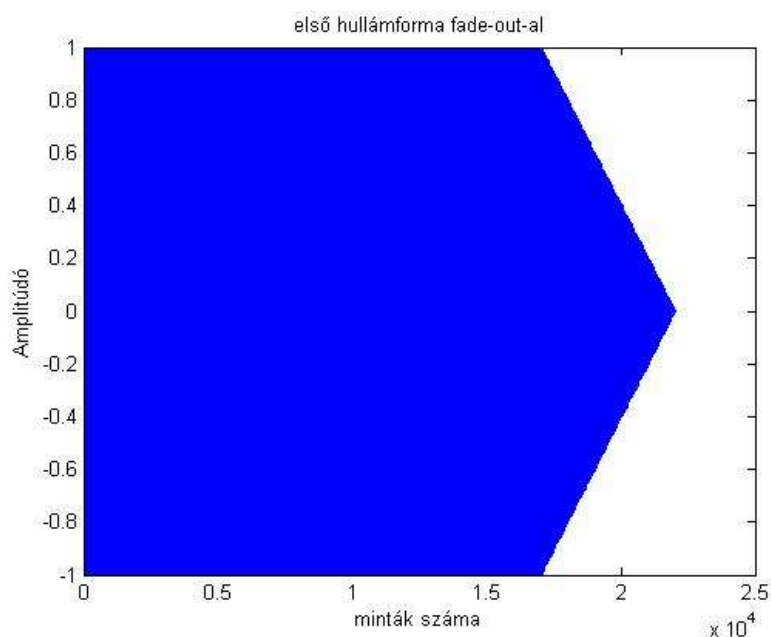
A következő példában a szemléletesség kedvéért három különböző hullámformát fűztem egymáshoz.

A három hullámforma megvalósítása MATLAB-ban:

```
%%
xfadewidth = 5000; % cross fade szélessége mintákban
N = length(sounds); % hangok száma
Fs = 22050; % mintavételi frekvencia
all_sample = 0; % kimeneti vektor kinullázása
for i=1:N
    all_sample = all_sample + (N+1) * xfadewidth;
    y = zeros(1,all_sample); % az összes hang ide jön
    ramp_in = (1:xfadewidth)/xfadewidth; % fade-in
    ramp_out = 1- ramp_in; % fade-out
    prev_end = 0;
    for i=1:N
        dt = 1/Fs; % másodperc per minta
        StopTime = sounds(i,2); % időtartam másodpercben
        T = (0:dt:StopTime-dt)'; % másodpercben
        switch sounds(i,3)
            case 1
                temp = sin(2*pi*sounds(i,1)*T)';
            case 3
                temp = square(2*pi*sounds(i,1)*T)';
            case 2
                temp = sawtooth(2*pi*sounds(i,1)*T)';
            otherwise
                temp = sin(2*pi*sounds(i,1)*T)';
        end
    end
```

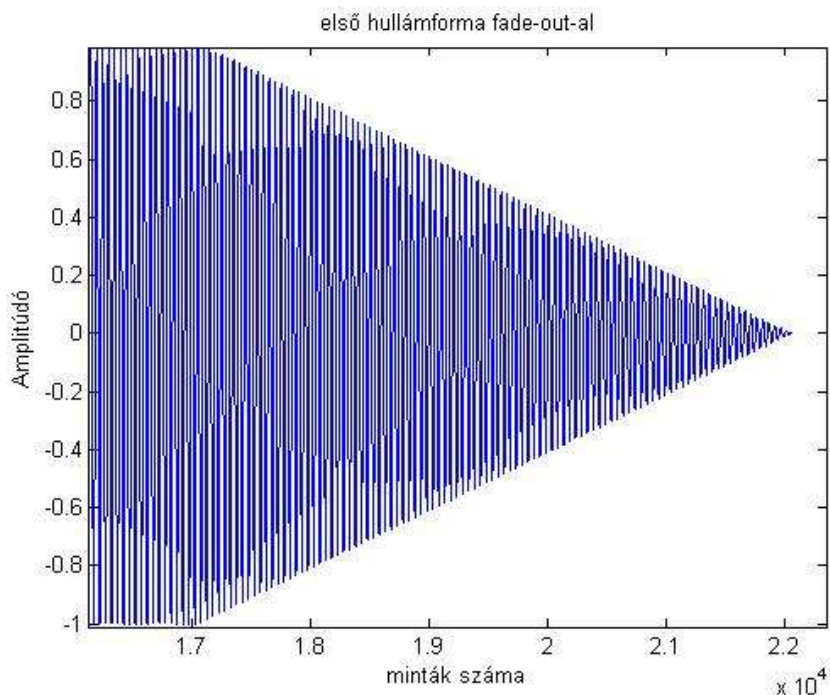
16. ábra Hullámformák megvalósítása

Az első hullámforma (szinusz) a következőképpen alakul:



17. ábra 440Hz-es szinusz hullám fade-out-tal

Az első hullámforma fade-out része kinagyítva:



18. ábra 440Hz-es szinusz hullám fade-out kinagyítva

A crossfade megvalósítása MATLAB-ban:

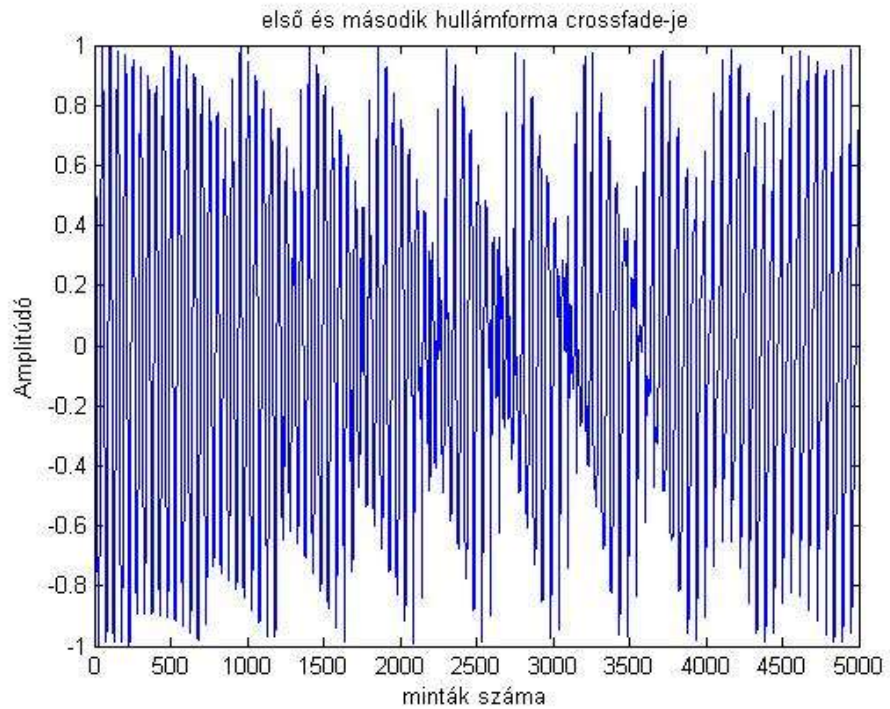
```
%első hang
% \
% \
% \
% Bejövő össze keverése
y(prev_end:prev_end+xfadewidth-1) = y(prev_end:prev_end+xfadewidth-1)+temp(1:xfadewidth);
plot(y(prev_end:prev_end+xfadewidth-1));
title('első és második hullámforma crossfade-je')
xlabel('minták száma')
ylabel('Amplitúdó')
pause;

% kimenő össze keverése
y(prev_end+xfadewidth : prev_end+xfadewidth + length(temp) - xfadewidth-1) = temp(xfadewidth+1 : length(temp));
prev_end = prev_end + length(mid) + xfadewidth;
plot(y);
title('első és második hullám összekeverve')
xlabel('minták száma')
ylabel('Amplitúdó')
pause;

end
```

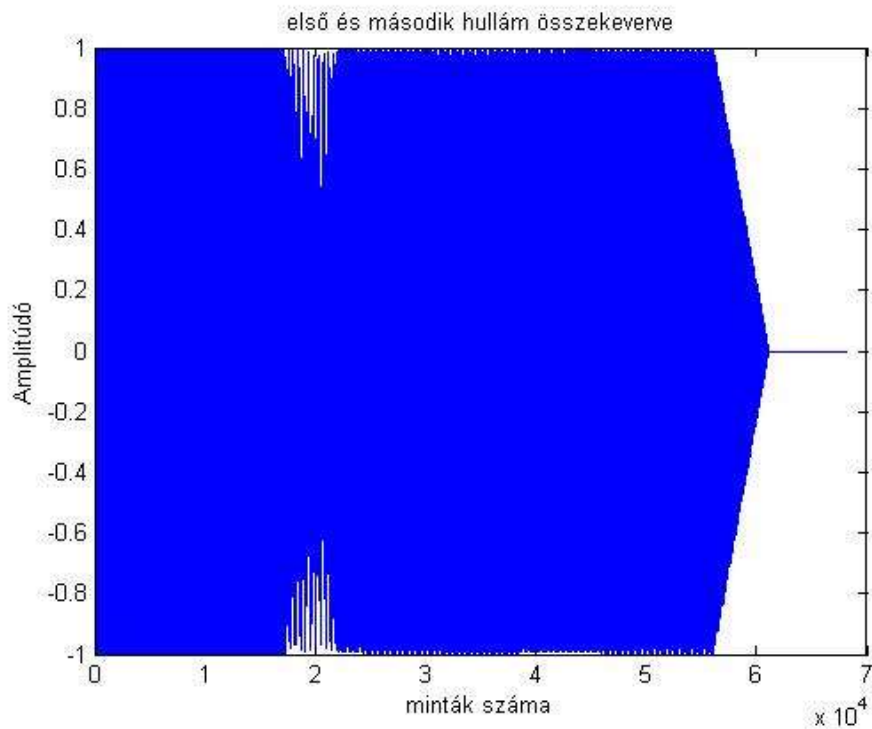
19. ábra Crossfade megvalósítása

Az első és második hullámforma crossfade szakasza:



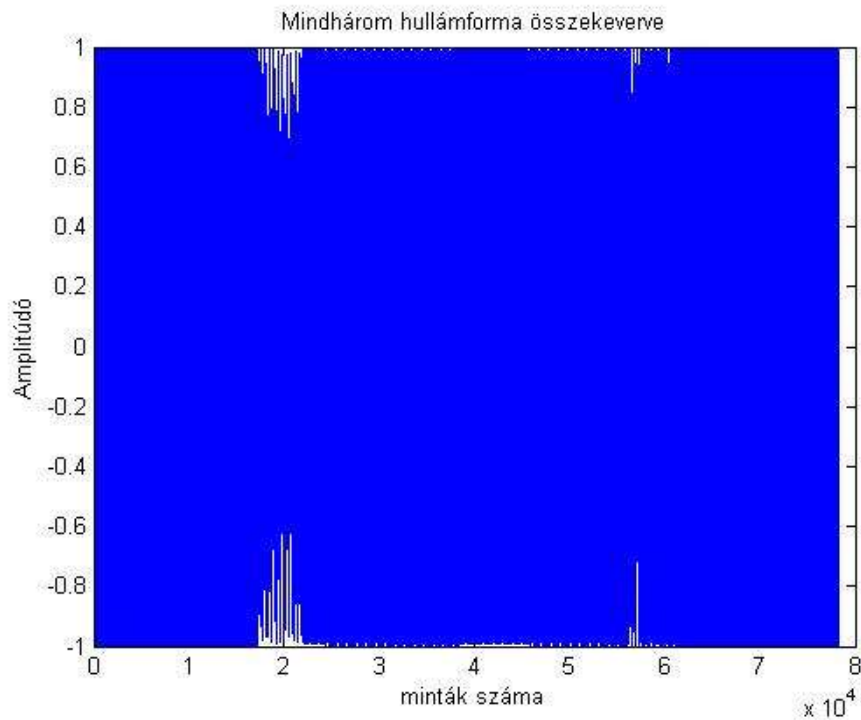
20. ábra 440Hz-es szinusz hullám és 392Hz-es fűrészhullám crossfade szakasza

Az első és második hullámforma crossfade-el összefűzve a következő képpen alakul:



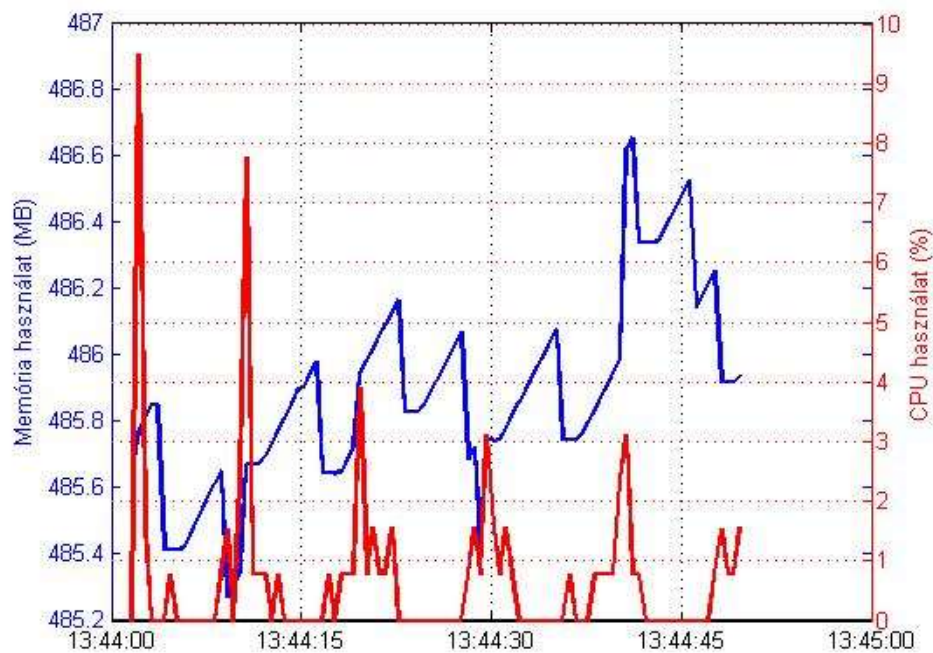
21. ábra 440Hz-es szinusz hullám és 392Hz-es fűrészhullám összefűzve

Mindhárom hullámforma crossfade-el összekeverve:



22. ábra Szinuszos, fűrész, négyszög jelek crossfade-el összefűzve

A módszer számítási és memória felhasználását a következő grafikon mutatja:



23. ábra Hullámtábla szintézis CPU és RAM felhasználása

Az adott hullámtábla szintézisről több mérést is szimuláltam, különböző számú és időtartamú jelekkel, melyek alapján készítettem el a szintézis RAM, illetve CPU értékeléseket. A fenti diagram a következő hullámtábla szintézis CPU és RAM felhasználtságát mutatja: az első hullámforma szinusz, mely 1 másodpercig tart, a második hullámforma fűrészjel, mely 2 másodpercig tart és a harmadik hullámforma négyszögjel, mely 1 másodpercig tart.

A CPU kezdeti megugrása csak a hangok előre definiálása miatt történt. A RAM felhasználást elsősorban a jelek száma, azok hossza, illetve a mintavételezési frekvencia határozza meg.

Maximális CPU felhasználtság: 9.5 %

Maximális RAM felhasználtság: 486,6 Mb

Átlagos CPU felhasználtság: 4,8 %

Átlagos RAM felhasználtság: 486,1 Mb

3.3. Granular Synthesis (szemcsés szintézis)

A szemcsés szintézis egy mindent átfogó kifejezés számos különböző audio rendszerekhez, amely a hangok egy kis töredékét felhasználva dolgozik, amelyeket egyénileg tudjuk manipulálni és rekombinálni, ezáltal létrehozva a végső kimenetet. A legtöbb szemcsés rendszer audio fájlokat, hangmintákat használ alapanyagként. A mintákat sok apró töredékre szeleteli fel, amelyek általában nem több, mint 10-50 milliszekundum hosszúak. Minden egyes szelet úgy ismert, mint 'szemcse' és ezek szemcsék sorozatát, pedig 'szemcsetáblának' hívjuk. Ha a szoftver szemcsetáblát készít, amely lejátsza az összes szemcse kivonatát az eredeti sorrendben és az eredeti sebességgel egy adott mintából, akkor az eredeti minta reprodukcióját hallhatjuk. Ha a szoftver lassabban játssza vissza a szekvenciát, akkor rések tűnhetnek fel a szeletek között, ezért az aktuális szelet a szemcsetáblázatban általában hurkolt. Gyorsabban visszajátszva minden szemcse átfedi a következőt vagy néhány szelet kimarad attól függően milyen a szoftver működése. Annak érdekében, hogy kattanások és hibák ne legyenek minden minta fade-in és fade-out-tal van ellátva egy burkoló görbe segítségével, ezt más néven simításnak is nevezzük. Az egyes szeletek átszerveződése szoftverenként különbözhet.

Sok szempontból *hullámtábla* szintézisnek is tekinthető, de sokkal finomabb skálán működik. Ahogy az várható, ez a módszer ideális folyamatosan változó hangok és igazán egyedi hangok előállításához.

Fizikai háttere:

A kvantum fizika megmutatta, hogy a hang automatikusan csökkenthető fizikai részecskékre.

A hang fizikai formáját először Isaac Beeckman álmodta meg 1618-ban.

1947-ben Gábor Dénes javasolta a szemcsés ötletet, mint a hang kvantumját.

Xenakis (1971) volt az első zenész, aki granulár szintézist használt.

Curtis Roads (1988), digitális granulár szintézise.

Barry Truax (1990), valós idejű granulár szintézise.

A szemcsék összetevői:

Burkoló: Használatával nem lesz torzított és ropogó zaj a minták elején és végén. A burkoló alakjának jelentős hatása van a szemcsés hangon.

- A bemintázott hanghoz egy rövid lineáris *attack* és *decay* akadályozza meg a hanghoz hozzáadott "kattanást"
- Változtatva a szemcse meredekségének a burkolóját megváltozik a kapott szemcse spektruma. (Például élesebb *attack* esetén szélesebb sávzélességet állít elő, mint egy nagyon rövid szemcse időtartammal)

Audio tartalom: bármilyen forrásból származhat: alapvető hullámformák vagy minták.

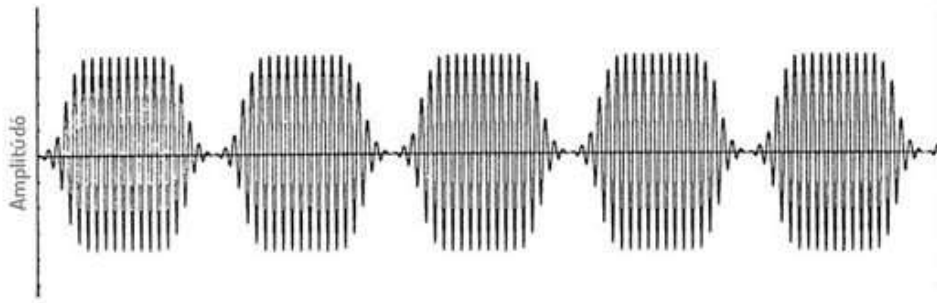
Hangalkotás:

Több ezer rövid szemcsés hang keletkezéséből állnak elő:

- Lineárisan kombinált nagyméretű hangeseményeket alakít,
- Három lehetséges kombináció:
 - I. kvázi-szinkron granulár szintézis
 - II. aszinkron granulár szintézis
 - III. *pitch/tempo* granulár szintézis
- A szemcsék jellemzői is meghatározhatóak és együttesen befolyásolják a hangzást

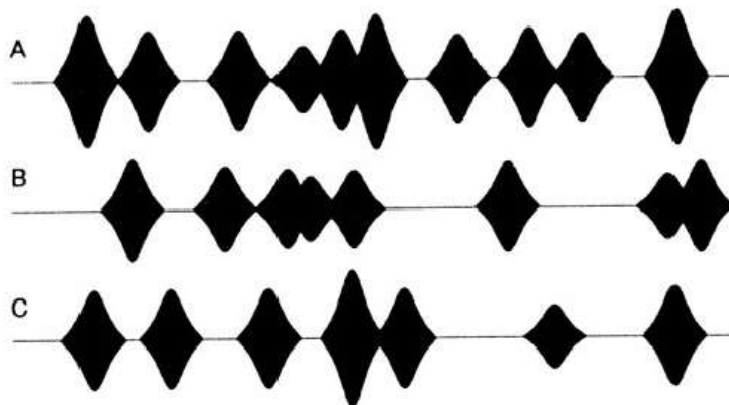
Kvázi-szinkron granulár szintézis:

- Az azonos méretű szemcsék folyama amplitúdó modulációt okoz kevesebb, mint 50ms-os szemcse időtartammal.



24. ábra Azonos méretű granulátumok

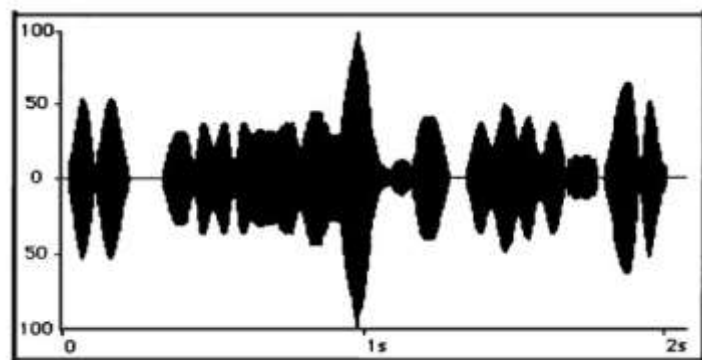
- Szemcse folyamok változó késleltetési idővel a szemcsék között: Az összegük az aszinkron granulár szintézisre hasonlít.



25. ábra Különböző késleltetésű granulátum folyamok

Aszinkron granulár szintézis:

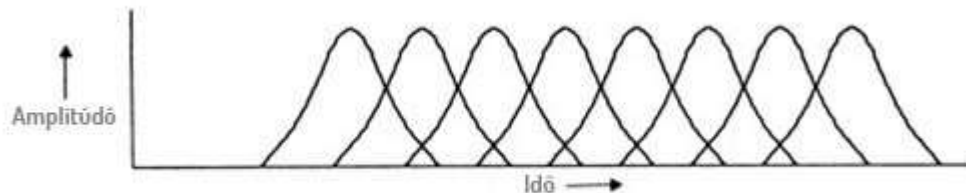
A szemcsék sztochasztikusan oszlanak el kvázi rendszeresség nélkül.



26. ábra Aszinkron granulár szintézis

Pitch/Tempo-szinkron granulár szintézis:

- Megőrzi a *pitch/tempo*-t miközben megváltoztatja a minta lejátszás sebességét. (Például: Kontakt, Intakt)
- Az átfedő szemcse burkolókat arra tervezték, hogy szinkronban legyenek a szemcse hullámforma frekvenciájával, ezáltal kevesebb melléktermék áll elő.



27. ábra Granulátumok átfedése

A granulár szintézis előnye, hogy a mai számítógépek már elegendő feldolgozási teljesítménnyel bírnak, így ez a szintézis módszer praktikus, és számos kereskedelmi termék számára már elérhető. Különleges hangokat lehet létrehozni a segítségével, még hozzá úgy, hogy a bemeneti mintákon számos paraméter megváltoztatható.

Hátránya, hogy a szemcsés szintézis rendkívül processzor igényes és nem lehetett valós időben használni a közelmúltig, ezért nagyjából mindenki figyelmen kívül hagyta. Néhányan a tudományos intézetek kutatói közül azonban a szemcsés szintézist tanulmányozták és fejlesztették tovább.

3.3.1. Granulár szintézis megvalósítása MATLAB programmal:

Egy egyszerű hang szintetizálásához készítettem egy függvényt, amelyben négy bemeneti paramétert vagyunk képesek változtatni:

- bemeneti jelet,
- első mintát,
- granulátum hosszát a mintában,
- fade-in és fade-out hosszát a mintában.

Ez a függvény a következőképpen néz ki MATLAB-ban:

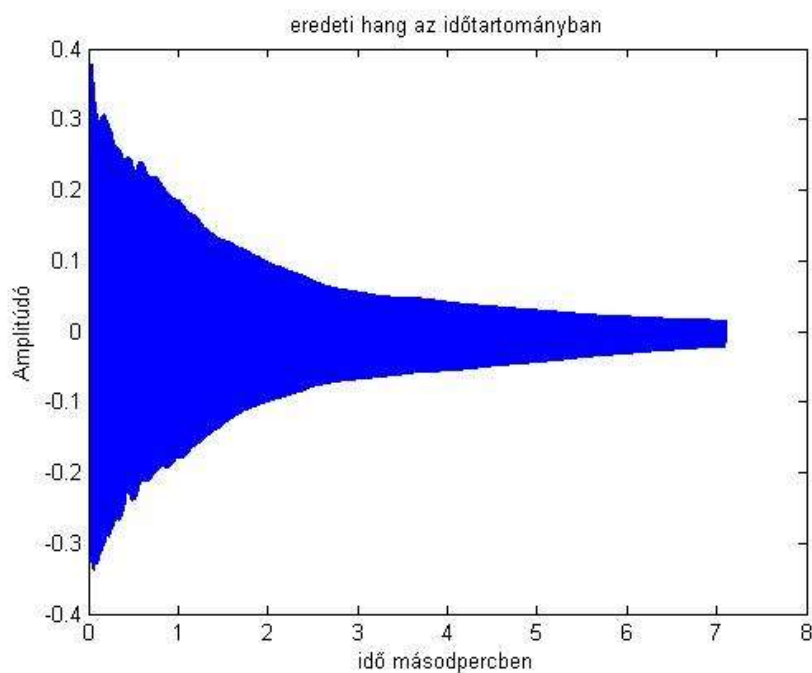
```
function y = granulumLn(x,init,L,Lw)
%
% Egy hosszú granulátum kivonata
%
% x - bemeneti jel
% init - első minta
% L - granulátum hossza a mintában
% Lw - fade-in és fade-out hossza a mintában

if length(x) <= init+L , error('x hossza túl rövid'), end

y = x(init:init+L-1); % bemeneti jel szemesse kivonata
w = hanning(2*Lw+1); % ablak
y(1:Lw) = y(1:Lw).*w(1:Lw); % fade-in
y(L-Lw+1:L) = y(L-Lw+1:L).*w(Lw+2:2*Lw+1); % fade-out
```

28. ábra Szemcséző függvény

Egy akusztikus 'D' gitárhang a következőképpen néz ki az időtartományban:



29. ábra 'D' Gitárhang

A szintézis megvalósítása MATLAB-ban a következőképpen néz ki:

```
%%
Ly=length(x);
y=zeros(Ly,1); % kimeneti jel kinullálása (memória lefoglalás)

timex = Ly/fs;

% Konstansok

nEv=10000; % granulátumok száma
maxL=fs*0.02; % granulátum maximális hossza
minL=fs*0.01; % granulátum minimális hossza
Lw=fs*0.01; % ablakméret

% Inicializálás

L = round((maxL-minL)*rand(nEv,1))+minL; % granulátum hossza a mintában
initIn = ceil((Ly-maxL)*rand(nEv,1)); % granulátum eleje (random)
initOut= ceil((Ly-maxL)*rand(nEv,1)); % granulátum vége (random)

endOut=initOut+L-1; % granulátumok egymáshoz csatolása

% Szintézis
for k=1:nEv,
    grain=granulatumLn(x,initIn(k),L(k),Lw); % adott szemcses
    %figure(2)
    %plot(grain);
    y(initOut(k):endOut(k))=y(initOut(k):endOut(k))+ grain;
end

doit = input('\nLejártessa és kirajszolja a granulált szintetizált hangot? Y/[N]:\n','s');

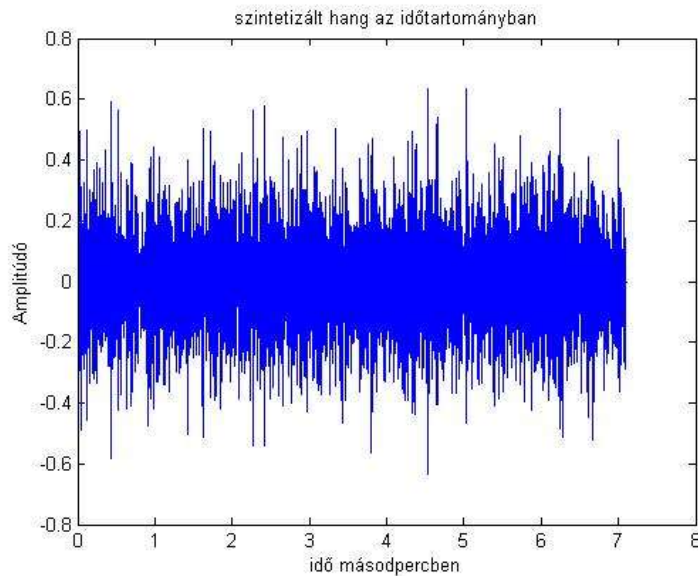
if doit == 'y',

    figure(2)
    plot(Time,y);
    title('szintetizált hang az időtartományban')
    xlabel('idő másodpercben')
    ylabel('Amplitúdó')
    sound(y,fs);
end

stopRecordAndDisplay;
```

30. ábra Granulár szintézis megvalósítása

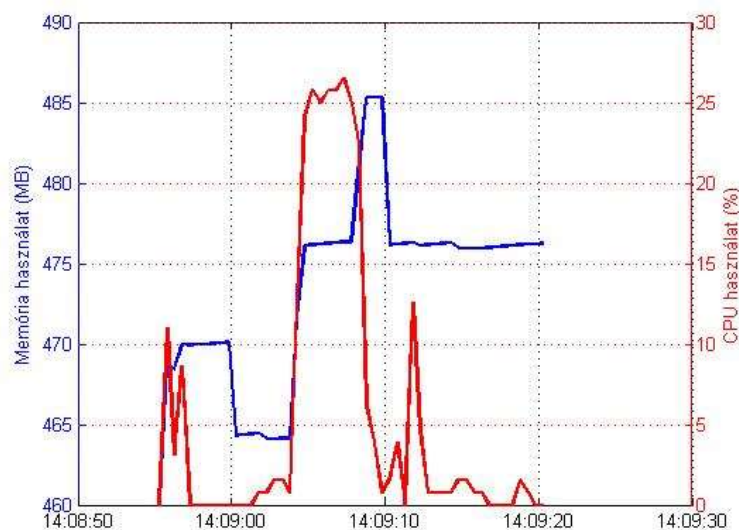
Véletlenszerűen granulált 'D' gitárhang pedig a következőképpen alakul:



31. ábra 'D' Gitárhang modulált szemcséiből szintetizált hang

E szintézis módszert úgy valósítottam meg, hogy az eredeti mintából készítettem nEv számú szemcsét (10000), ezeket a szemcséket véletlenszerűen moduláltam, majd ismét véletlenszerűen egymásután raktam őket. Az ábrán jól látszik, hogy a szemcsék elrendezése viszonylag egyenletesen alakul.

Az általam készített mérőprogram segítségével szemléltetem a szintézis CPU, illetve RAM felhasználását:



32. ábra Granulár szintézis CPU és RAM felhasználása

Az adott granulár szintézis méréséről több szimulációt is végrehajtottam, melyben a legfőbb paraméterek változtatásával vizsgáltam a CPU és RAM felhasználtságot. A fenti grafikon 10000 szemcse számú és a maximális, illetve minimális szemcseméret a mintavételezési frekvencia 2, illetve 1 százaléka.

A granulátumok száma, illetve a minimális és maximális szemcseméret befolyásolja a számítási igényt. A granulátumok optimális értékei nagyságrendileg 1000 és 100000 között találhatók. A minimális és maximális szemcseméret optimális értékei 1 és 10 százalék közöttiek.

Maximális CPU felhasználtság: 26,4 %

Maximális RAM felhasználtság: 485,5 Mb

Átlagos CPU felhasználtság: 14,6 %

Átlagos RAM felhasználtság: 476,9 Mb

Scanned Synthesis (Ütemezett szintézis)

Az *ütemezett* szintézis egy olyan lassú dinamikus rendszer, amelynek rezgő frekvenciái 15 Hz vagy ezen érték alattiak. A rendszert közvetlenül a cselekvő személy mozgása befolyásolja. Az *ütemezett* szintézis a hallásunk pszicho akusztikáján alapul, függvény segítségével megbecsüli a hangszíneket és a mozgató izmunk vezérlésének a képességét (hektikus), így befolyásolja a mérhető teljesítményű hangszíneket. Az *ütemezési* függvény periódus ideje a szintetizált hang periódusa. A előállított hangszínek nem hasonlíthatóak a hagyományos hangszerek hangszíneihez, teljes mértékben eltérőek.

Az *ütemezett* szintézis lényege az, hogy lassú és alacsony rezonáns frekvenciájú tárgyat használ a közvetett személy, aki közvetlenül befolyásolja az objektum rezgéseit (a testének mozgásával). A továbbiakban egy periódusos vizsgáló függvényt használ. Megvizsgálja az objektum alakját egy periodikus pálya mentén. A pálya periódusa megegyezik a hang alapfrekvenciájával, amit létre szeretnénk hozni. A vizsgáló függvény átalakítja az objektum lassan változó térbeli hullám alakját, egy hallható frekvenciás hanghullámmá.

Feltételek a módszer használatához:

Lassú rezgésű dinamikus rendszer,

Közvetett személy, aki hektikus mértékben befolyásolja a rendszert,

Ütemezési folyamat, rendszeresen hallható mértékű hangszín.

A rendszer rezgései, a kezdeti feltételek függvényei, az energia rendszerek dinamikájával írható le a szereplő személy és a rendszer között. A hallható frekvencia elkészítéséhez a felhasznált dinamikus rendszer alakjának egy zárt pálya mentén periódikusan ütemezettnek kell lennie. Az *ütemezési* függvény sebessége határozza meg a hangmagasságot. A hangmagasság vezérlése teljesen különáll a dinamikus rendszerünktől, tehát a hangszín és a hangmagasság függetlenek.

A fentiek értelmében a rendszerre tekinthetünk úgy, mint a vezérlők engedélyezésével, egy közvetett személy által ellenőrzött dinamikus hullámtáblázatra. A dinamikus hullámtáblázatban lévő számoknak egyezniük kell a hasznos térbeli frekvenciákkal.

Az *ütemezett* szintézis hatalmas előnye, hogy az *ütemezett* folyamat számításigénye egyszerű és hatékony, csupán egyszer kell elvégezni, így igen alapos is. A szintézis egyedi jellemzője, hogy fokozott hangsúlyt fektet a cselekvő személy által generált hangszín vezérlésére.

Hátránya, hogy nem elegendő, hogy egy szám adott helyen, és a tömbben változzon a zavart idejű frekvencián. Egy egyedi számnak kell kapcsolódnia a szomszédjához az *ütemezett* pálya mentén, a kívánt térbeli frekvenciát használva. Ez a térbeli frekvencia, idő-frekvenciává van átalakítva az *ütemezett* függvény segítségével. Ezt a tulajdonságot elérhetjük a matematikai függvények választásával, amit a zavar-generátor számol.

4. SPEKTRÁLIS MODELLEK:

A spektrális hang szintézis módszerek a hanghullámok tulajdonságainak a modellezésén alapulnak, mivel ezek a hallgató által érzékeltek. Ezek közül számos figyelembe veszi a pszicho akusztikus ismereteket. A spektrális hangszintézis módszerek általánosak abban, hogy a legkülönbözőbb hangok modellezésében alkalmazhatók.

4.1. Spectral Modelling Synthesis (Spektrális modellező színtézis)

A Spektrális modellező színtézis (SMS) technikát az 1980-as években fejlesztették ki a Stanford Egyetemen. Serra (1989) kikísérletezett egy módszert, mellyel a hanghullámokat lebontotta determinisztikus és sztochasztikus összetevőkre. A determinisztikus összetevőket a MQ algoritmussal (McAulay and Quatieri, 1986), illetve amplitúdó analízissel határozta meg. A determinisztikus rész kivonásra kerül az eredeti jel idő, illetve frekvencia tartományából. Ezáltal a fennmaradó jel megfelel a sztochasztikus összetevőknek.

Ez a módszer számos zenei analízis alkalmazásban használatos, beleértve a pengetős hangok felvett analízisét, hogy ebből levezesse a megfelelő gerjesztési jeleket a pengetős hangok természetes modellezéséhez.

Az SMS technika alapja, hogy feltételezi, hogy a bemenő hangok leírhatóak két jel összetevőinek az összegeként, melyek a determinisztikus és sztochasztikus összetevők. Meghatározás szerint a determinisztikus jel olyan jel, amely teljes mértékben számolható. Az SMS modell szinuszos összetevőkre korlátozza a determinisztikus részt szakaszosan lineáris amplitúdó és frekvencia változataival. Ez befolyásolja a modell általános jellegét és emiatt néhány hangot nem lehet pontosan modellezni ezzel a technikával. A sztochasztikus összetevők ebben a módszerben a spektrális energia sűrűséggel írhatóak le. Ezért nem szükséges a sztochasztikus összetevők fázis információit megőrizni. A sztochasztikus összetevőket hatékonyan lehet reprezentálni az egyes DFT keretek fennmaradó spektrum burkoló nagyságával.

4.2. Transient Modeling Sythesis (Tranziens modellező színtézis)

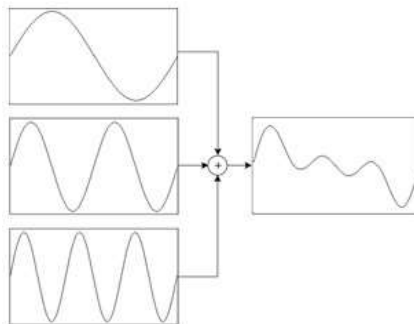
Ebben a megközelítésben, a maradék jelet az eredeti jelből kivont szinuszos modellel kapjuk meg, ami két részből áll: tranziensek és állandó zajos összetevők. A Tranziens Modellező színtézis (TMS) a tranziens összetevők parametrikus ábrázolását teszi lehetővé.

A TMS az idő és frekvencia tartomány közötti kettőségen alapul. A tranziens jelek intuitívak az időtartományban, ezáltal nem könnyen paraméterezhetők. Azonban megfelelő módosítással átalakíthatóak frekvencia tartománybeli jelekké, amelyek szinuszos jellegűek. Ez azt jelenti, hogy a frekvenciatartományban szinuszos modellezés alkalmazható, hogy az impulzív jeleknek a parametrikus ábrázolását megkapjuk.

4.3. Additive synthesis (Összegezési szintézis)

Az additív szintézis alapelve, hogy hangtónusokat ötvöz (tipikusan változó amplitúdójú harmónikusokat.)

- Az additív szintézis lényege a frekvenciák „összekeverése”.
- Egy hang minden egyes frekvencia összetevőjének megvan a saját amplitúdó burkolója.
- Ez lehetővé teszi a független viselkedését ezen összetevőknek.
- A források különfélek lehetnek a szintézisek vagy minták formái tekintetében.



33. ábra Additív szintézis

Az additív szintézis az első széles körben elterjedt technika, amit hangszintézisre és analízisra használtak a digitális zenében. Az ötletet a *Fourier-elmélet* ihlette (Daniel Bernoulli meglátásai alapján), amely kimondja, hogy matematikailag minden hangot ki lehet fejezni szinuszos hangok összegeként, ez az *összeadási* szintézis. A hang kialakulása több komponensből, viszonylag lassan változó amplitúdójú és frekvenciájú modulációból következik. A szinusz hullámok összetevőinek elemzésével rekonstruálhatjuk a komplex hangszíneket. Az elmélet minden eleme valós mennyiség, ezért minden szinusz-görbének

egymástól független amplitúdója (A) és/vagy fázis változása (w_0) lehet. A görbéket rövid idejű Fourier analízis határozza meg.

$$y(t) = \sum_{i=1}^N A_i(t) \sin[\theta_i(t)]$$

Kiemelkedő tulajdonsága, hogy minden egyes hang frekvencia komponensének megvan a saját amplitúdó burkolója, amelyek függetlenek egymástól. Lényegében az *összeadási* szintézis a frekvenciák összekeverése.

Előnye, hogy rengeteg mikro-eltérés található az egyedi hangok frekvenciájában és amplitúdójában.

Hátrányát viszont a hatékonyság hiánya képezi: nagy mennyiségű adatot kell megadni, hogy meghatározhassuk egy hang bonyolultabb részleteit.

4.3.1. Additív szintézis megvalósítása MATLAB programmal:

Egy egyszerű hang szintetizálásához készítettem egy függvényt, amelyben öt bemeneti paramétert vagyunk képesek változtatni:

- frekvenciát Hertz-ben,
- időtartamot másodpercben,
- amplitúdót (0 és 1 tartományban),
- mintavételezési frekvenciát Hertz-ben,
- típus: szintézis típusának a kiválasztása (aktuális opciók: 'fm', 'szinusz' és 'fűrész')

Ennek megvalósítása MATLAB-ban a következőképpen alakul:

```
function y=ssintezis(freq,dur,amp,Fs,type)
% y=synth(freq,dur,amp,Fs,type)
%
% Egyes hangok szintetizálása
%
% Bemeneti adatok:
% freq - Frekvencia Hertz-ben
% dur - Időtartam másodpercben
% amp - Amplitúdó [0,1]
% Fs - Mintavételezési frekvencia Hertz-ben
% type - Hullám típusok, jelenlegi lehetőségek:
%       'fm', 'szinusz', vagy 'fűrész'
%
%-----
%-----

if nargin<5
    error('Öt bemeneti érték szükséges a szintetizáláshoz');
end

N = floor(dur*Fs);
n=0:N-1;
if (strcmp(type,'szinusz'))
    y = amp.*sin(2*pi*n*freq/Fs);

elseif (strcmp(type,'fűrész'))

    T = (1/freq)*Fs; % Minták periódusa
    ramp = (0:(N-1))/T;
    y = ramp-fix(ramp);
    y = amp.*y;
    y = y - mean(y);

elseif (strcmp(type,'fm'))

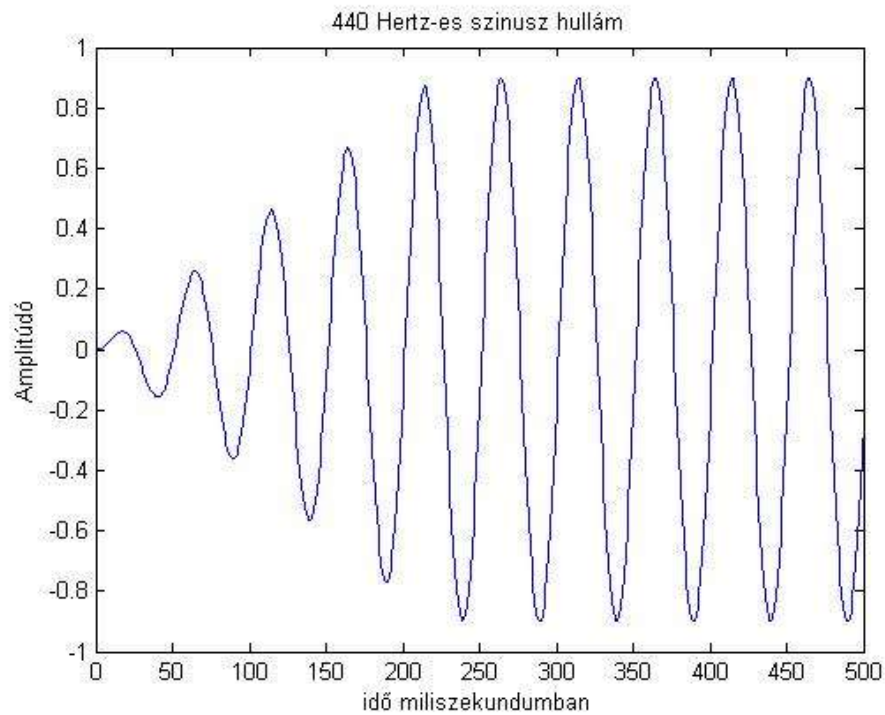
    t = 0:(1/Fs):dur;
    envel = interp1([0 dur/6 dur/3 dur/5 dur], [0 1 .75 .6 0], 0:(1/Fs):dur);
    I_env = 5.*envel;
    y = envel.*sin(2.*pi.*freq.*t + I_env.*sin(2.*pi.*freq.*t));

else
    error('Ismeretlen szintetizálási típus');
end

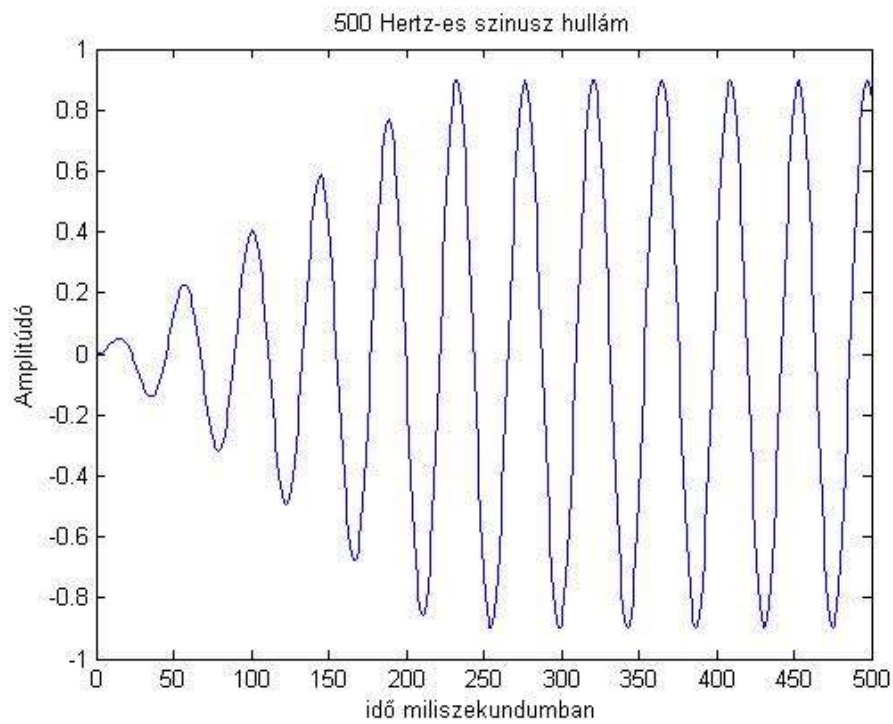
% széllek szimitása 10ms-os meredekséggel
if (dur > .02)
    L = 2*fix(.01*Fs)+1; % L páratlan
    ramp = bartlett(L); % páratlan hossza
    L = ceil(L/2);
    y(1:L) = y(1:L) .* ramp(1:L);
    y(end-L+1:end) = y(end-L+1:end) .* ramp(end-L+1:end);
end
```

34. ábra Szintézis függvény

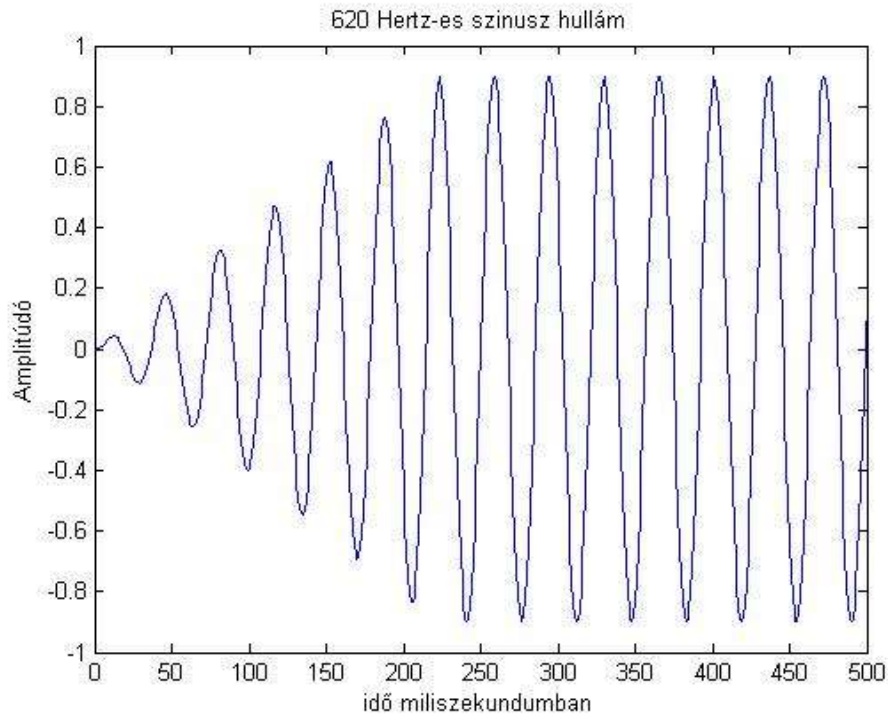
Egy egyszerű szinusz hullám generátor kimenetei a következőképpen néznek ki:



35. ábra 440Hz-es szinusz hullám



36. ábra 500Hz-es szinusz hullám



37. ábra 620Hz-es szinusz hullám

Az összeadást a következő MATLAB részlet szemlélteti:

```
% Hullámok összekeverése

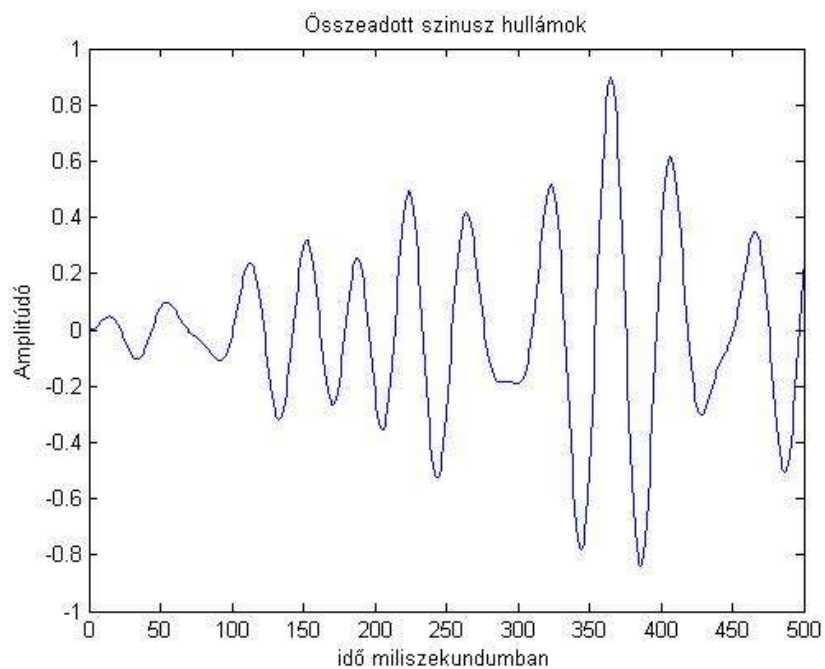
yadd = (y1 + y2 + y3)/3;

doit = input('\nLejátassa és kirajszolja az összekevert hullámokat? Y/[N]:\n', 's');

if doit == 'y',
figure(4)
plot(yadd(1:500));
title('Összeadott szinusz hullámok')
xlabel('idő milliszekundumban')
ylabel('Amplitúdó')
%sound(yadd,Fs);
end
```

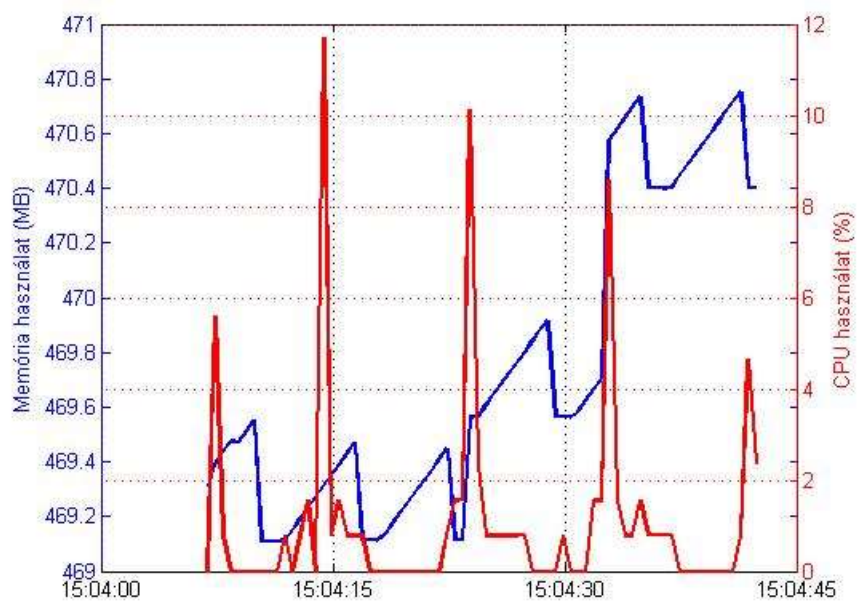
38. ábra Hullámok összeadása

A különböző frekvenciájú szinusz hullámok összekeverése látható az alábbi ábrán:



39. ábra Összeadott szinusz hullámok

Az általam készített mérőprogram segítségével szemléltetem a szintézis CPU, illetve RAM felhasználását:



40. ábra Additív szintézis CPU és RAM felhasználása

Az adott additív szintézis CPU, illetve RAM méréseinek a szimulációját többször végeztem el különböző számú bemeneti jelekkel. A fenti grafikonon három szinusz hullám (jelenként 2 másodpercesek, 0.9 amplitúdójúak és a mintavételezési frekvencia 22050) összeadásának a CPU és RAM felhasználtságát mértem.

A bemeneti paraméterek változtatásával (frekvencia, időtartam, amplitúdó, mintavételezési frekvencia és a bemeneti jel) nem változik jelentősen a számítási igény, ezért ezeket az ábrákat nem szemléltetem. A bemeneti jelek számát tekintve nagyjából lineárisan nő a számítási igény, hiszen minél több a bemeneti jelek száma, annál több összegzés szükséges. Tehát a számítási igényt tekintve a bemeneti jelek száma a meghatározó paraméter.

Maximális CPU felhasználtság: 11,9 %

Maximális RAM felhasználtság: 470,78 Mb

Átlagos CPU felhasználtság: 8,3 %

Átlagos RAM felhasználtság: 469,87 Mb

4.4. Subtractive Synthesis (Kivonó szintézis) – Analóg szintézis

Működésének alapelve, hogy összetett hangok szűrésével harmonikus spektrumot alakít ki, amelyek tipikusan geometriai hullámokkal kezdődnek.

Az analóg szintézis elavult, az 1960-'70-es évek szintetizátoraira fejlesztették ki. Természetesen magában foglalja az oszcillátorokat, amelyek időben változó nyomású hanghullámokat generálnak, amelyeket ezután feldolgoz és módosít: időben változtatja a hangmagasságot, frekvencia összetevőket és az amplitúdót. Az eredeti hullámforma elemei kivonásra kerülnek.

A szubtraktív szintézis még mindig használatos a modern digitális szintézisben is (Pl.: Moog Minimoog (1969), Sequential Circuits Prophet 5 (1978), Roland Juno 60 (1981)).

A szubtraktív szintetizátorok 3 alap eleme:

Hangforrás: Az oszcillátorok olyan elektronikus hangforrások, ahol nyers elektromos nyomású hanghullám jön létre. A legtöbb oszcillátor felkínálja a választható hullámformát a hang alkotásához. Ezek lehetnek: négyszög, pulzus, fűrész, illetve háromszög hullámok. A modern szintetizátorok (digitális vagy szoftveres)

tartalmazhatnak összetettebb hullámformákat vagy lehetővé teszik tetszőleges hullámformák feltöltését.

Szűrés: A vágási frekvencia és a szűrő rezonanciája vezérelhető, hogy szimulálni tudjuk egy adott hangszer természetes hangszínét.

Amplitúdó burkoló: A jel amplitúdójának további burkoló vezérlése (szigorúan véve ez nem szubtraktív szintézis, de gyakran használják). Ezt szintén lehet használni más szintézis technikáknál is.

Előnye, hogy akár minden változó értéket az idő függvényében módosít. Több évtizede ismert és használt, tehát a módszer alkalmazásában nem szükséges bonyolult számítások alkalmazása.

Hátránya a modern szintézis módszerekkel szemben, hogy a mai digitalizált világban már elavultnak tekinthető az analóg szintézis.

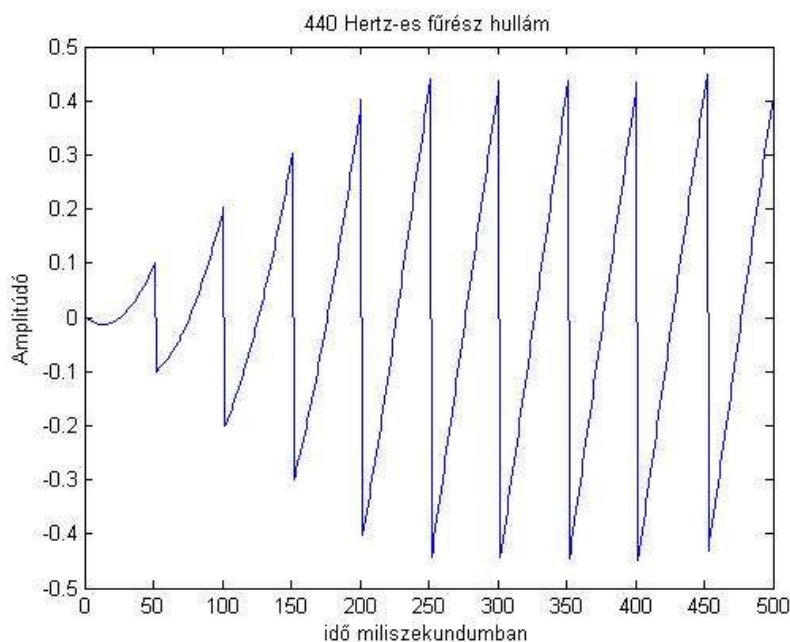
4.4.1. Szubtraktív szintézis megvalósítása MATLAB programmal:

Egy egyszerű hang szintetizálásához készítettem egy függvényt, amelyben öt bemeneti paramétert vagyunk képesek változtatni:

- frekvenciát Hertz-ben,
- időtartamot másodpercben,
- amplitúdót (0 és 1 tartományban),
- mintavételezési frekvenciát Hertz-ben,
- típus: szintézis típusának a kiválasztása a húrhoz (aktuális opciók: 'fm', 'szinusz' és 'fűrész')

Az ebben a szintézis módszerben előállított jeleket ugyanúgy állítottam elő, ahogyan az additív szintézisnél már bemutattam. Ezért ennek a forráskódja itt nem kerül bemutatásra.

Egy egyszerű fűrészel generátor kimenete a következőképpen néz ki:



41. ábra 440Hz-es fűrészel hullám

Az alul-áteresztő szűrő megvalósítása a következőképpen néz ki MATLAB-ban:

```
% Aluláteresztő szűrővel megvalósított szubsztitív szintézis

Fs = 22050; % Mintavételezési frekvencia

%y = szintesis(440,2,0.9,22050,'fm');
%y = szintesis(440,2,0.9,22050,'sinus');
y = szintesis(440,2,0.9,22050,'fűrészel');
[y, Fs] = wavread('gitar_D_rythm.wav');

% Lejátsza és kirajszolja a fűrészel hullámformáját

doit = input('\nLejátsza és kirajszolja a generált fűrészeljelet? Y/[N]:\n','s');

if doit == 'y',
    figure(1)
    plot(y(1:500));
    title('440 Hertz-es fm hullám')
    xlabel('idő milliszekundumban')
    ylabel('Amplitúdó')
    sound(y,Fs);
end

% Aluláteresztő szűrő készítése és y szűrése

[B, A] = butter(1,0.04, 'low');

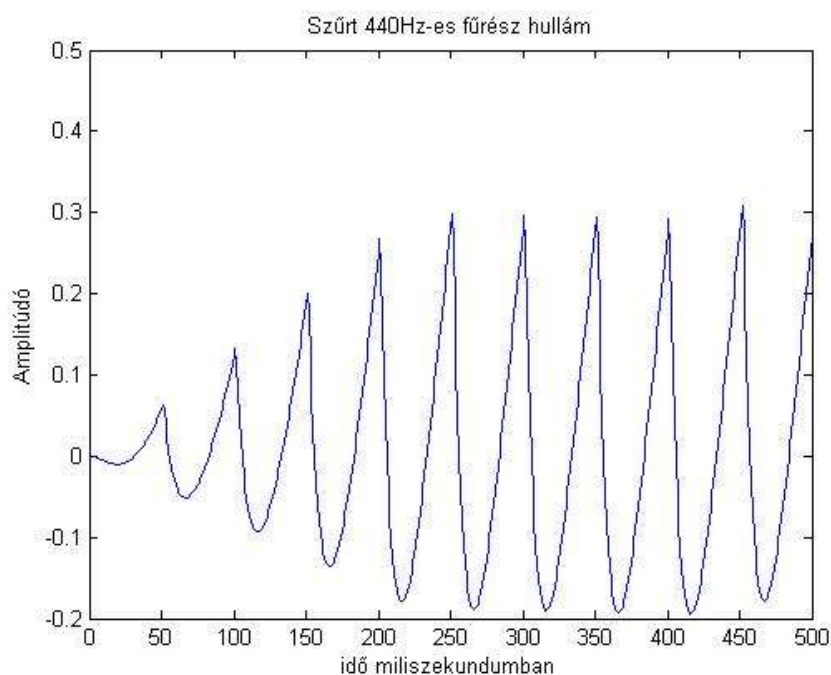
yf = filter(B,A,y);

[B, A] = butter(4,0.04, 'low');

yf2 = filter(B,A,yf);
```

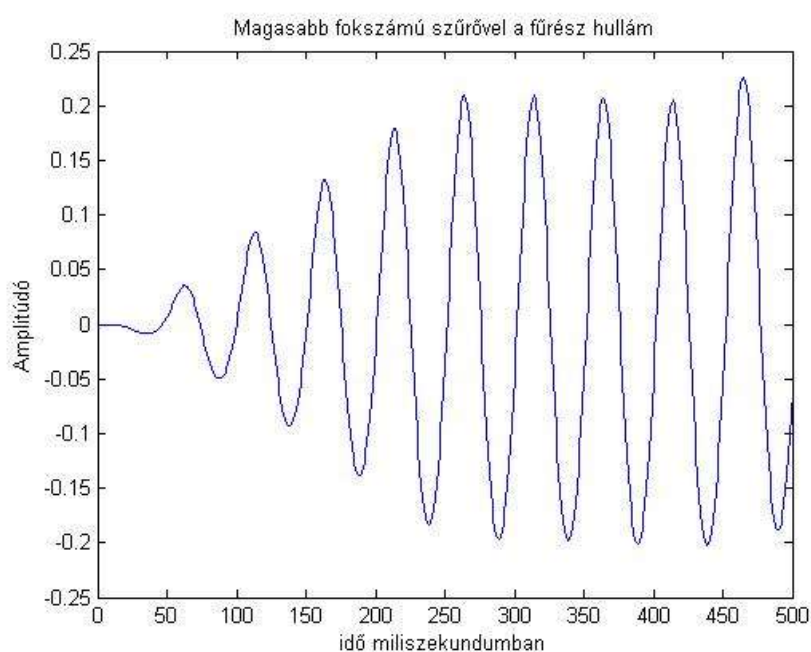
42. ábra alul-áteresztő szűrő megvalósítása

Egy alul-áteresztő - végtelen impulzus válaszu Butterworth - szűrő alkalmazásával kivonom az eredeti jelből a magas frekvenciás összetevőket:



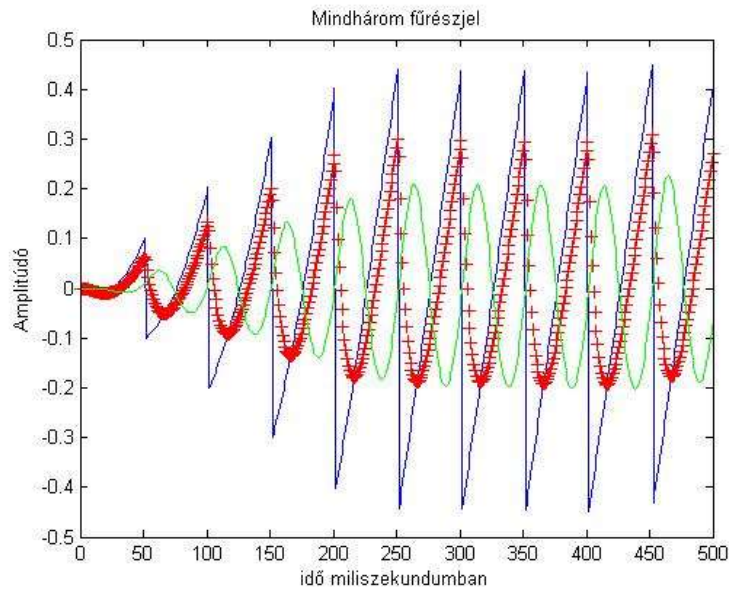
43. ábra Szűrt 440Hz-es fűrész hullám

Az eredmény még kifinomultabban alakul, ha magasabb fokszámú szűrőt használunk:



44. ábra Negyedfokú szűrővel a 440Hz-es fűrész hullám

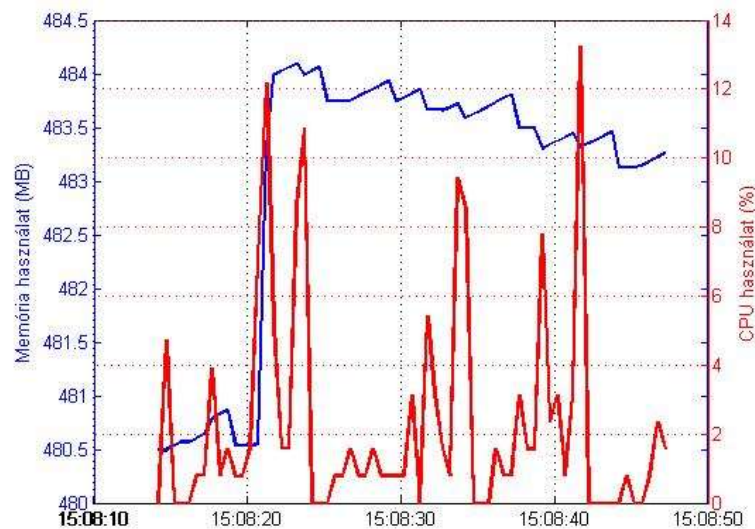
A szemléletesség kedvéért mindhárom hullámforma egy ábrán:



45. ábra Mindhárom fűrészhullám

Az ábra azt szemlélteti miként alakul az eredeti fűrészjel alul-áteresztő szűrő használatával. A piros hullámforma az elsőfokú alul-áteresztő szűrővel kivont jelet, míg a zöld egy magasabb fokszámú alul-áteresztő szűrővel kivont jelet ábrázol.

Az általam készített mérőprogram segítségével szemléltetem a szintézis CPU, illetve RAM felhasználását:



46. ábra Szubtraktív szintézis CPU és RAM felhasználása

Az adott szubtraktív szintézis CPU, illetve RAM mérésének a szimulációját többször elvégeztem különböző jelekkel és különböző szűrési beállításokkal. A fenti grafikon 0.04 rad/s körfrekvenciájú először egy elsőfokú, azután pedig egy negyedfokú szűrő alkalmazásával szűrt fűrészjelnek a CPU és RAM felhasználtságát mutatja. (A fűrészjel 440Hz-es, 2 másodpercig tart, 0.9 amplitúdóval és 22050 mintavételezési frekvenciával)

A számítási igényt tekintve a meghatározó paraméterek a szűrő paraméterei. Egy egyszerű 'butter' szűrőt használtam, ami beépített függvényként szerepel a MATLAB programban, amelynek három paramétere változtatható (a szűrő fokszáma, a szűrő vágási frekvenciája és a típusa). Minél magasabb fokszámú szűrőt használunk, annál simább lesz a kimeneti jelünk, ezért nem használhatunk végtelen nagy fokszámú szűrőt, mert a végére nem maradna semmi a jelből. Ebben a példaprogramban a maximális fokszám, amit még elbír a rendszer kb. 14.

Maximális CPU felhasználtság: 13,2 %

Maximális RAM felhasználtság: 484,15 Mb

Átlagos CPU felhasználtság: 9,76 %

Átlagos RAM felhasználtság: 483,07 Mb

4.4.2. Szubtraktív szintézis egy emberi példán keresztül:

Erre úgyis tekinthetünk, amiben az emberek zajokat generálnak, mint a szubtraktív szintézis:

- oszcillátor: a hangszalagokra tekinthetünk úgy, mint hangforrás és
- szűrő: a torok, illetve száj módosítja a hangot.
 - „Ó” vagy „Á” hangot kimondva vagy énekelve (ugyanazon a hangmagasságon)
 - A hangszalagok nagyjából ugyanezt a fűrészjelet generálják, amelyek gazdagok harmonikus hangzásban. A különbség a kettő között a szűrésből adódik, amelyet a száj és a torok alkalmaz.
 - A száj alakjának a változtatásával módosul a szűrő levágási frekvenciája, így eltávolítva (kivonva) néhányat a harmónikusok közül.
 - Az „Á” hangban van jelen mindig a legtöbb eredeti harmonikus.
 - Az „Ó” hangban pedig a legtöbb eltávolításra kerül. (Pontosabban, csökkentett amplitúdóval)

Summation Formulae (összegzéses formula)

A *closed-form* összegzéses formula, számos lehetőséget nyújt a komplex spektrum létrehozásához. Kihasználva a harmonikus sorozatokat és a jól ismert matematikai kifejezéseket, amelyek kompakt módon adják vissza a számtani sorozatokat.

Band-limited Pulse (Sávkorlátozott impulzus)

A legegyszerűbb esetben egy összetevőt a harmonikus komponensekből, hasonló súllyal növeljük (*scaled/normalised*):

$$\frac{1}{N} \sum_{k=1}^N \cos(k \omega_0)$$

Továbbá a következő változókat vezetjük még be: a t időváltozót, w (néhol q) radiánban mért frekvenciát, egy egyszerű változóban $w=2\pi ft$, ami időben változó fázis. A növelést követően, egy sávkorlátozott impulzust kapunk. Ez az összegzés könnyen előállítható, figyelembe véve az egyik legjobban ismert aritmetikus sorozatot, amit egy *closed-form*-en értelmezünk. Megkapjuk a következő feltételezett egyenlőséget, ezt *Dirichlet-kernel* kifejezésnek nevezzük, amit először Windham és Steiglitz állapított meg:

$$\frac{1}{N} \sum_{k=1}^N \cos(k \omega_0) = \frac{1}{2N} \left[\sum_{k=-N}^N e^{-ik \omega_0} \right] - 1 = \frac{1}{2N} \left[\frac{\sin((2N+1)\omega_0/2)}{\sin(\omega_0/2)} - 1 \right]$$

Az egyetlen probléma, ami itt felléphet, az a nevezőben esetlegesen megjelenő nulla. Ebben az esetben egyetlen 1-essel helyettesíthetjük az egész kifejezést, ugyanis rendkívül bonyolult számításokat igényelne a megvalósítás.

Moorer's generalised Discrete Summation Formulae Synthesis (Moorer féle generált Diszkrét Összegzéses Formula szintézis)

J.A.Moorer meghatározott számunkra egy másik összegzéses képletet, amely rugalmasabb szintézis vezérlést tesz lehetővé. Jelentősége, hogy lehetőséget kínál a spektrum *roll-off* vezérlésére, illetve a diszharmonikus részek kialakulására.

Így a szintézis kifejezése, azaz a sávkorlátozott jel:

$$\sum_{k=0}^N a^k \sin(\omega + k\theta) = \frac{\sin(\omega) - \sin(\omega - \theta) - a^{N+1}(\sin(\omega + (N+1)\theta) - a \sin(\omega + N\theta))}{1 - a \cos(\theta) + a^2}$$

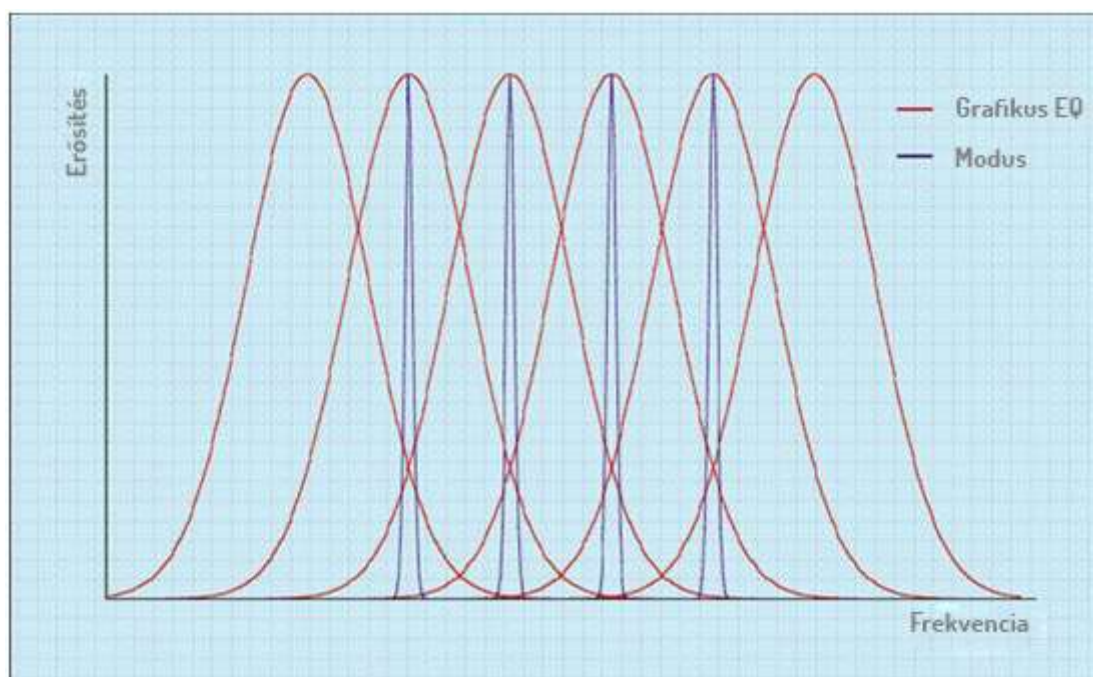
Ahol az a és N módosításával, megtudjuk változtatni a spektrális *roll-off*-ot és a sáv szélességet. A t időváltozó lehetővé teszi, hogy egy alul-áteresztő szűrő viselkedését utánozzuk. Különböző w és q értékeket választva különböző típusú harmonikus és diszharmonikus spektrumokat képezhetünk.

4.5. Formant Synthesis

A formáns a hangszerek és az emberi beszédhang jellegzetes színét adó rezonanciás úton felerősített felhangtartomány.

Azok a módszerek, amelyek a beszéd kialakulásának mechanizmusát szimulálják a *source-filter* elméleten alapulnak. Ebben a megközelítésben, egy lineáris szűrő szimulálja a vokális traktust. A *Formant* szintézisnél a vokális traktus átviteli függvénye kielégítően modellezhető, a formáns frekvenciák és amplitúdók szimulálása által.

Felmerül a kérdés, hogy miért nem használunk grafikus *equaliser*-t (vagy multi-band szűrőt), hogy létrehozzuk ezeket a megkülönböztethető csúcsokat a hang spektrumában. Ennek az oka egyszerű: Az *equaliserek*nek a csúcsai, amelyek magába foglalják a hagyományos *filter-bank*okat, túl durvák (nem elég pontosak), ezáltal minden szűrő széles frekvencia tartományban átereszt. Ezért szükséges egy speciálisabb szűrési módszer, ami lyukszerűen szűr. Ehhez több sáváteresztő lánc szükséges.



47. ábra Formánsok

A való életben, minden ember vokális traktusa különbözik egymástól, így a pontos pozíciója a formánsoknak emberről emberre változik. Továbbá a formánsok amplitúdója is különböző és a formánsok szélessége is más és más az embereknél (Q). Q -val fejezzük ki a sávszűrőnek az élességét. Úgy tudjuk kiszámolni a Q értékét, hogy az EQ görbének a fél sávszélessége által szétosztjuk a görbének a közép frekvenciáját.

5. FIZIKAI MODELLEZÉS:

A hangszerek fizikai modellezése az egyik legaktívabban fejlődő területe a hangszíntézisnek. A fizikai modellező alkalmazások egyre népszerűbbek, ezáltal a felhasználóknak egyre több lehetőségük van, hogy mind hagyományos, mind új szintetizált hangokat hozzanak létre.

A modell célja, hogy az aktuális hangszer alapvető természetes viselkedését szimulálja. Ezt úgy érhetjük el, hogy a fizikai törvényeket alkalmazzuk, amiket mozgások és kölcsönhatások szabályoznak a vizsgált rendszeren belül, majd kifejezzük őket matematikai formulákkal és képletekkel. Ezek a matematikai modellek nyújtanak eszközt a fizikai modellezéshez.

Alapvetően két fő oka van a fizikai alapú modellek fejlesztésének: Az első maga a tudomány, azaz, hogy a modellek fizikai jelenségeit megértsük. A másik a szintetizált hang előállítása.

A fizikai modellező módszereket öt kategóriába csoportosíthatjuk (Valimaki and Takala).

Parciális differenciálegyenletek numerikus megoldása.

Source-filter modellezés. (Forrás-szűrő)

Rezgő tömeg-rugó hálózatok

Modal szintézis (Alaki)

Waveguide szintézis

A *Waveguide* szintézis az egyik legszélesebb körben elterjedt fizika-alapú hang szintézis módszer napjainkban. Rendkívül hatékonyan szimulálja a hullám terjedést az egydimenziós homogén rezgő rendszerekben.

A módszer jellemzően digitális jelfeldolgozás irányultságú és számos valós idejű megvalósítása létezik a *Waveguide* szintézisnek.

A *Source-filter* modelleket különösen az emberi hangképzés mechanizmus modellezésére használják. A hangszálak és vokális traktus kölcsönhatása egy előrecsatolt rendszerrel modellezhető. A hatékony digitális szűrési technikák a *Source-filter* modellezésre lettek kifejlesztve, legfőképpen az emberi beszéd továbbítása céljából.

A modellezési módszerek a rendszert szimulálják, mind idő, mind pedig frekvencia tartományban. A frekvencia tartományú módszerek nagyon hatékonyak a lineáris rendszerű

modellekhez. A hangszerek sajnos nem közelítik meg elég pontosan ezt a linearitást. A nemlineáris rendszerek a frekvenciatartománybeli megközelítést viszont

megvalósíthatatlanná teszik. Az itt bemutatott módszerek az időtartománybeli vizsgálatát modellezzik a rendszernek.

Három fizikai modellező módszerrel kezdődik ez a fejezet, ami számszerűsített akusztikát használ. Az első véges különbségi módszert használ, a húros, valamint az ütős hangszerek bemutatására szánt alkalmazásokkal. Második a *Modal* szintézis. Harmadik a CORDIS, ami a *tömeg-rugó hálózatok* révén a rezgő tárgyak modellezési rendszere.

5.1. Finite Difference Methods (Parciális Differenciálegyenletek Numerikus Megoldása)

Ezek a modellező módszerek a véges differenciaegyenleteken alapulnak. Leginkább pengetős hangszereken használatos ez a módszer.

A módszer általában bármely rezgő tárgyra alkalmazható, például húrra, membránra, stb. (Hiller and Ruiz, 1971). Az alapelve, hogy egy matematikai egyenletet kapjunk, ami leírja az adott tárgy rezgőmozgását. Ezeket a hullám egyenleteket megoldva kapjuk meg a differenciaegyenletet. A differenciaegyenletek használata egy megismétlődő egyenlethez vezet, amit úgy értelmezhetünk, mint a hullámterjedés szimulálását egy rezgő objektumban. A véges differencia módszer a leghatékonyabb az egydimenziós vibrátorokkal és természetesen, ahogy növeljük a dimenziók bevezetését, úgy többszöröződik a számítási igény is. A térbeli pontok száma arányosan növeli a dimenziók számának a módját. Természetesen minden egyes ponton a számítási műveletek számának a megnövelésével az effektív mintavételezési frekvencia is megnő.

Hiller és Ruiz (1971) voltak az elsők, akik a rezgő húr differenciálegyenleteinek megoldásával foglalkozott hangszintézis céljából. Leginkább a pengetős, ütős, és hajlított húros modelleket fejlesztették. A húr merevségét modellezték, valamint a frekvenciafüggő veszteségeket. Hiller és Ruiz egy számítógépes program segítségével képesek voltak szintetizált hangok előállítására és a kapott hullámformák megjelenítésére (Chaigne and Askenfelt, 1994).

5.2. Modal Synthesis (Alaki szintézis)

A modális szintézis módszert elsősorban a párizsi IRCAM-ban fejlesztették. (Adrien, 1989), (Adrien, 1991). Az általuk gyártott Modalys (Eckel et al., 1995) kereskedelmi szoftver korábban Mosaic (Morrison and Adrien, 1993) néven volt ismert. Ezt az alkalmazást a rezgő rendszerek szimulálására használták a felhasználók. A felhasználó leírja a struktúrát a programhoz, majd a program kiszámítja a modális adatokat és a kimenő jelet egy megfigyelési ponton, amit a felhasználó határoz meg.

A Modális szintézis azon a feltevésen alapul, hogy bármely hangkeltő tárgy leírható egy sor rezgő al-struktúrával, amelyeket a modális adatok határoznak meg. (Adrien, 1991). Az al-struktúrák kapcsolatban vannak és reagálnak a külső gerjesztésekre. Ezek a csatlakozó kapcsolatok energiaáramlást tesznek lehetővé az al-struktúrák között. A jellemző al-struktúrák a következők:

- testek és áthidalások
- vonók
- akusztikus csövek
- membránok és lemezek
- csengők

A szimulációs algoritmus minden alstruktúra és azok kölcsönhatásainak az információit használja fel.

A módszer általánosnak tekinthető, mivel tetszőleges bonyolultságú struktúrákat lehet vele alkalmazni. A számításigénye gyorsan növekszik a bonyolultsággal, így a módszernek gyakorlati korlátai vannak.

5.3. Mass-Spring Networks – the CORDIS System (Tömeg-rugó hálózatok)

Cadoz et al. (1983) megpróbálta megvizsgálni az akusztikai rendszereket, egyszerű ideális elemeket használva, mint például a tömeg, a csillapító és a rugó. Céljuk az volt, hogy kidolgozzanak egy paradigmát, amely tetszőleges akusztikai rendszereken alkalmazható. A CORDIS rendszer volt az első, amely képes volt valós időben fizikai modellen alapuló hang előállítására (Florens és Cadoz, 1991).

A legkezdetlegesebb és alapvetőbb elemei a rendszernek a következők:

- pontszerű tömegek

- ideális rugók

- ideális csillapítók

Amikor ezek az összetevők megfelelő számban kombinálódnak, a térbeli folytonosság és az akusztikai jel reprodukálása lehetővé teszi a kívánt mintavételt (Florens és Cadoz, 1991). A vizsgált objektumot ezekkel az elemekkel különállóan elosztva közelíthetjük a felületen. Jelentős egyszerűsítést úgy kapunk, hogy az egyes elemek egy dimenzióban vannak, azaz minden elem csak egy dimenzióban mozoghat, illetve működhet. Az időben változó kölcsönhatások modellezésére (például a vonóvezetés, kalapáccsal ütés) egy feltételes kapcsolatot vezettünk be. Ez egy rugóból és egy csillapítóból áll, ami állítható paraméterekkel párhuzamosan van kapcsolva.

5.4. Numerikus módszerek összehasonlítása

Egy húr rezgésének a szimulálását már bemutattuk, korábban felhasználva a véges differenciaegyenleteket. A véges differencia módszer igen pontos az eredeti hullámalak reprodukálására, ha a modell paraméterei helyesek. Tudományos értelemben a vibrációs mozgás megfigyelhető bármely különálló ponton (például egy húron). Továbbá a modell paraméterei az aktuális valós hangszer paraméterei, mint például a merevség és a húr

veszteséges paraméterei, illetve a bemeneti admittancia az áthidalásnál. Ezeket a paramétereket a mérésekkel, analízisekkel kapjuk meg a hangszeren és az általa létrehozott jeleken.

A valós idejű hangszíntézis céljából a véges differencia modell nem a legkecsegtetőbb. A modellt csak egyszerű struktúrákra lehet alkalmazni, mint például egy rezgő húr, valós időben; ideértve egy gitár testét a modellben azt jelentené, hogy szükség van hibrid rendszerekre is. A húron minden egyes N ponthoz, 5 szorzásra és 8 összeadásra van szükség és egy további szorzásra és összeadásra, ha gerjesztést is alkalmazunk. A jó térbeli felbontáshoz N -nek nagynak kell lennie. Így több száz művelet szükséges minden kimeneti mintához. Kurz és Feiten (1996) írt példa programot egy húr rezgésének, valamint az egyes paraméterek hatásának a szimulálásához. A gitár *Modal* színtézissel három alstruktúrára osztható (minden funkcionális része a hangszernek), ezek a gerjesztés, a rezgő hurok és a hangteret sugárzó test. A gerjesztő alstruktúra csak akkor lép kölcsönhatásba más részekkel, amikor a húr meg van pengetve. A gerjesztés bármely ponton alkalmazható a másik két alstruktúrán is. A rezgő húr N párhuzamos független másodrendű rezonátorral van szimulálva, mindegyik egyetlen harmonikust állít elő. A rezonátorokat csak számításigényesen lehet megvalósítani, de a nagy részük jó minőségű színtézist tesz lehetővé. A hangszer testének a modelljét a struktúra modális analízisével kapjuk meg. Ez nagyon időigényes folyamat, különösen egy bonyolult szerkezetnél, mint a hegedű (Adrien, 1991).

Ha egy valós hangszer testét jelátalakítóként használjuk, a sugárzó hangteret a rezgő húr és a test összekapcsolása által lehet előállítani, szimulálni. Természetesen ez a módszer is használható más módszerekkel, amik képesek előállítani jelet az áthidalásnál.

A CORDIS rendszer minden rezgő szerkezetet idealizált elemekre osztja, azaz a pont tömegek rezgése egy irányban kapcsolódik ideális csillapítókkal és rugókkal. A rezgő húr, tehát N pont tömegek együttes kapcsolata $N-1$ összeköttetéssel, ami egy csillapítóból és egy húr párhuzamos kapcsolatából áll. Ez a szerkezet képes N harmonikus előállítására. A számítási műveletek száma minden egyes cellához viszonylag alacsony. Egy kimeneti mintához kb. $3N$ szorzás és $6N$ összeadás szükséges. Sajnos a becslése, hogy hány pont szükséges egy gitár testének a szimulációjához, nem állt rendelkezésre.

Az összefoglaló készítéséhez több megfigyelés készült. A véges differencia módszert a rezgések szimulálásához használják és így az egydimenziós tárgyak rendkívül pontosan leírhatóak. A többi módszer általánosabb a vibrációs jelenségek részletes matematikai

bemutatása és pontossága árán. A véges differencia módszer és a *Modal* szintézis módszer eszközöket nyújt a valódi hangszerek tanulmányozásához. Azonban egyik módszer sem alkalmazható kiválóan a valós idejű hangszintézis célokra. Az első probléma a számítási költség, ha jó minőségű szintézis kívánatos. Másodszor, a modell paraméterei zenei értelemben nem intuitívak és nehéz ugyanúgy irányítani, mint az aktuális kontrollált hangszert, különösen az élő helyzetekben. Végül a hangszintézis módszerek hatékonyabb számítással és ellenőrzéssel léteznek, különösen a vonósok és fafúvósok kúpos furatokkal.

5.5. Digital Waveguides (Digitális hullámvezetés)

Digital waveguides és a *single delay loop* (SDL) modellek a leghatékonyabb módszerek a fizikai alapú valós idejű hangszerek modellezéséhez. Jó minőségű modellek léteznek számos zenei hangszerhez és ezen a területen a kutatás még mindig aktív.

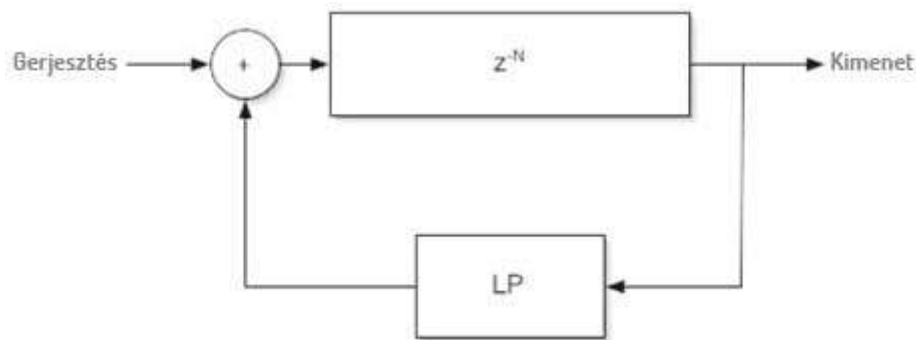
Ebben a fejezetben először a digitális hullámvezetőket tárgyaljuk. Másodszor a *digital waveguide meshes*-eket, amelyek 2D és 3D modellek, majd a *banded digital waveguides* is említésre kerül. A kétirányú digitális hullám modellt és az SDL modellt, amit Karjalainen et al., 1998-ban részletezett, röviden bemutatom.

A digitális hullámvezető koncepcióját Smith fejlesztette ki (1987, 1991, 1997). A digitális hullámvezetés és más módszerek alapján a véges differenciák szorosan összefüggnek, abban, hogy mindkettő a hullámegyenletek megoldásának az előfeltételével kezdődik. Minden időpillanatban egy fizikai jellemző, mint például az elmozdulás, kiszámítható minden egyes pontban. Ez azt jelenti, hogy az egész diszkrétizált rezgő objektum vibrációs mozgása könnyen megfigyelhető. Bár ez lehet, hogy vonzó a rezgő objektum és a vibrációs mozgás tanulmányozásához, hatékonyabb módszerekre van szükség hangszintézis célokra.

Az egy dimenziós *waveguide*-t széles körben használják a számítógépes zenében, hogy különféle közegekben terjedő akusztikus hullámokat modellezzenek. A *waveguides* megközelítés lehetővé teszi a különböző zenei hangszerek szimulációját, beleértve a pengetős, ütős, vonós és fúvós hangszereket. A *digital waveguides* hatékonyak (olcsó és valós idejű végrehajtás) és sokatmondó fizikai értelmezést biztosítanak.

A *digitális waveguides* előnye, hogy hatékony szintetizálási eszközt nyújtanak a kvázi-harmonikus rezonátoroknak, amelyek magukban foglalják a rezgő húrokat is (ha elhanyagolható vagy gyenge a diszperzió). Azokban az esetekben, ahol a merevség érezhető, de nem túl nagy, mint például a zongorahúroknál, *all-pass* szűrőket alkalmaznak. Az *all-pass* szűrők szerepe, hogy frekvenciafüggő terjedési sebességet hozzon létre.

A nagyon merev rendszerek azonban, a *hullámvezetők* egy kombinációja és az *all-pass* szűrők kevésbé hatékony módszert biztosítanak a hang szintézishez. Néhány diszharmonikus struktúrát, mint például a harangokat és a magasabb dimenziójú struktúrákat ennek ellenére *all-pass* szűrőkkel modellezzik. Ezekben az esetekben, más szintézis technikát, mint például a *spektrális szintézist* vagy *modell* alapú szintézist használnak.



48. ábra Digitális hullámvezető egyszerű blokkvázlata

Ez egy egyszerű *hullámvezető* modellje egy rezgő húrnak. Az N mintából álló késleltető művonal jelöli a húrt és az LP, pedig az alul áteresztő szűrőt, ami magyarázza a veszteséget.

5.5.1. Waveguide:

A digitális hullámvezetés a hullám egyenlet általános megoldásán alapul egydimenziós homogén közegben. Egy rezgő húr esetében a veszteségmentes hullámegyenlet a következőképpen fejezhető ki (Morse és Ingard, 1968):

$$K \frac{\partial^2 y}{\partial x^2} = \varepsilon \frac{\partial^2 y}{\partial t^2}$$

,ahol a K a húrnak a feszültsége, ε a lineáris tömegeloszlás, és y a húr elmozdulása. Ez az egyenlet bármely veszteségmentes vibrációs mozgásnál alkalmazható, mint például a levegőoszlop egy hengeres fafúvós hangszerben. Természetesen a paramétereket és a hullám változókat megfelelően kell értelmezni.

A megoldás, hogy a hullámvezető által az egydimenziós hullámegyenlet igen pontos a diszkrét pontokon veszteségmentes esetben mindaddig, amíg a hullámfrontok eredendően sávkorlátozva vannak a mintavételezési sebesség felére. Sávkorlátozott interpolációt alkalmazhatunk, hogy megbecsüljük a haladó hullámok értékeit a késleltető nem elválaszthatatlan pontjain. A *Fractional Delay Filter*-ek kényelmes megoldást nyújtanak a sávkorlátozott interpolációhoz. (Laakso et al., 1996) és (Valimaki, 1995).

5.5.2. Waveguide alkalmazásai:

A digitális hullámvezetőket számos hangszintézis problémánál alkalmazhatjuk (Smith, 1996). Az első fizikán alapuló megközelítés, miszerint modellezni lehet egy zenei hangszert digitális szűrő használatával, az a hegedű volt Smith által (1983). Jaffe és Smith (1983) számos kiterjesztést mutatott be a Karplus-Strong algoritmushoz, amelyek lehetővé tették a megpengetett húr jó minőségű szintetizálását, beleértve az *all-pass* szűrőket a késleltetett hurokban, közelítve a késleltető nem beépített részét.

Mióta ezek az úttörő munkák elindultak, számos fejlesztést és további kiterjesztéseket nyújtottak a megpengetett húr szintézisekhez. Ezek közé tartozik a Lagrange interpoláció, hogy finom hangolni lehessen a hangmagasságot (Karjalainen és Laine, 1991), és az *all-pass* szűrési technikák, hogy szimulálni lehessen a diszperziót, amit a húr merevsége okoz. (Smith, 1983), (Paladin és Rocchesso, 1992), (Van Duyne és Smith, 1994). A felcserélt hullámvezető szintézis technika hatékony módja a hullámvezető szintézisnek, mely magába foglalja a hangszer testének a kiváló minőségű modelljét. Valimaki et al. (1995) bemutatott egy módszert, mellyel sima *glissandi*-t állított elő részleges *all-pass* késleltető szűrővel. Az STFT alapján egy paraméter kalibrációs módszert fejlesztett ki Karjalainen et al. (1993), illetve további részleteit Valimaki et al. (1996). Ezeket a munkákat kiegészítették és implementálták egy automatikus kalibrációs rendszerbe (Tolonen és Valimaki által, 1997). A húr modell többsebességű megvalósítását és a különálló alacsony sebességű rezonátor testet Smith (1993), Valimaki et al. (1996), és Valimaki és Tolonen (1997) mutatták be.

A pengetős húr algoritmust szintén használták az elektronikus hangszerek hangjának a szimulálásához. Sullivan (1990) kiterjesztette a Karplus-Strong algoritmust, hogy elektronikus gitár hangokat szintetizáljon torzítással, illetve visszacsatolással.

Smith és Van Duyne (1995) mutatták be a *waveguide* szintézist a zongorákhoz, a zongorabillentyű modelljét (1994), 2D digitális *waveguide mesh* (Van Duyne és Smith, 1993) és az *allpass* szűrő technikát, hogy együtt kombinálva szimulálják a húr merevségét és a *soundboard*-ot. A zongorabillentyű modelljének egy másik fejlesztését Borin és Giovanni mutatták be (1996).

A *waveguide* szintézist számos fúvós hangszer esetében is alkalmazták. Az első alkalmazása a klarinét volt, melyben számos személy működött közre Smith, (1986), Hirschman, (1991), Valimaki et al., (1992), Roccesso és Turra, (1993). A fuvola *waveguide* modelljének a javaslatát Karjalainen és Laine, (1991) és Valimaki et al. (1996-97) mutatták be. Valimaki et al. (1993) javasolt egy modellt a fafúvós furatokban lévő ujj lyukakra. Cook (1991), Dietz és Amir (1995), Msallam et al. (1997), Vergez és Rodet (1997) a rézfúvós hangokat szimulálta. Cook már 1992-ben létrehozott egy eszközt, amely a fafúvós modelleket képes irányítani.

Az SPASM egy DSP program (Cook, 1993), amely az emberi hangszabályozási mechanizmusát modellezi valós időben. Grafikus felhasználói felületet biztosít a vokális traktus alakjának a megjelenítésével.

5.6. Waveguide Meshes („Waveguide hálók”)

Az előző részben bemutatott *digital waveguide* rendkívül hatékonyan modellezi az egydimenziós rezgéseket. Ha 2D-ben vagy 3D-ben szeretnénk rezgőmozgást modellezni, akkor a *digital waveguide* egy fejlettebb változatát kell használnunk, a *waveguide mesh*-t („waveguide háló-t”). A *waveguide háló* alkalmazásait megtaláljuk, például a membránok, *soundboard*-ok, cintányérok, gongok és terem akusztikában.

Van Duyne és Smith (1993a, 1993b), Fontana és Rocchesso (1995), Savioja és Valimaki (1996, 1997) adták a 2D *waveguide* háló különféle kivitelezéseit. Viszonylag egyszerű a háromdimenziós hálóra bővítés, de a 3D hálókat részletesen itt nem tárgyaljuk. A kétdimenziós hullámegeyenletnek a mozgó síkhullám megoldását Morse és Ingrad határozták meg (1968) :

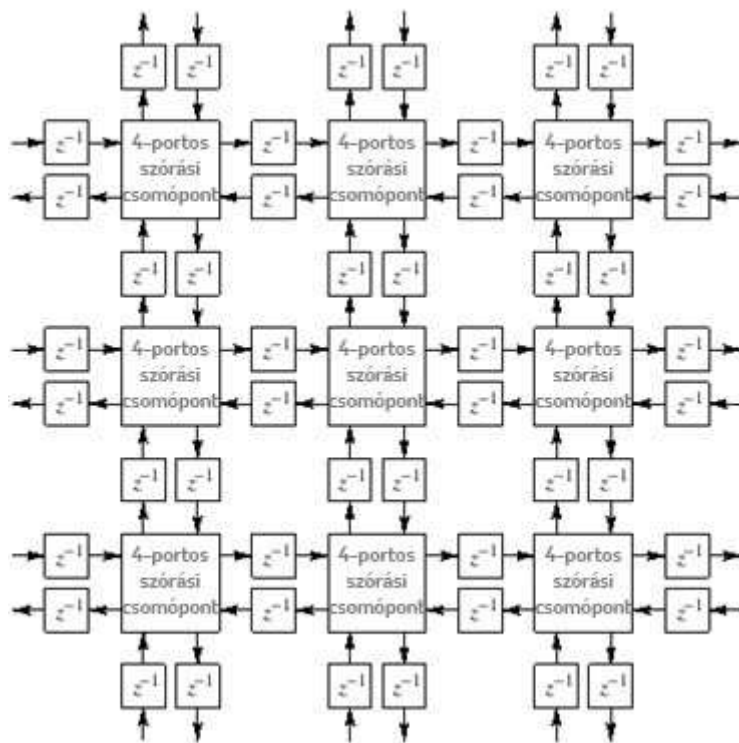
$$\frac{\partial^2 u(t, x, y)}{\partial t^2} = c^2 \left[\frac{\partial^2 u(t, x, y)}{\partial x^2} + \frac{\partial^2 u(t, x, y)}{\partial y^2} \right]$$

$$u(t, x, y) = \int f_{\alpha}(x \cos(\alpha) + y \sin(\alpha) - ct) d\alpha$$

Ahol α jelöli a síkhullám irányát. Az integrál magában foglalja a végtelen számú mozgó hullámokat, amelyek összetevőkre van osztva az x és y irányokban.

5.6.1. 2D Waveguide mesh:

A kétdimenziós hullámegyenlet egyenes vonalú *waveguide* háló formulája késleltető elemekből és 4-portos szórásos csomópontokból áll. Az egymást követő szórásos csomópontok közötti két késleltető elem egy kétirányú késleltető egység.

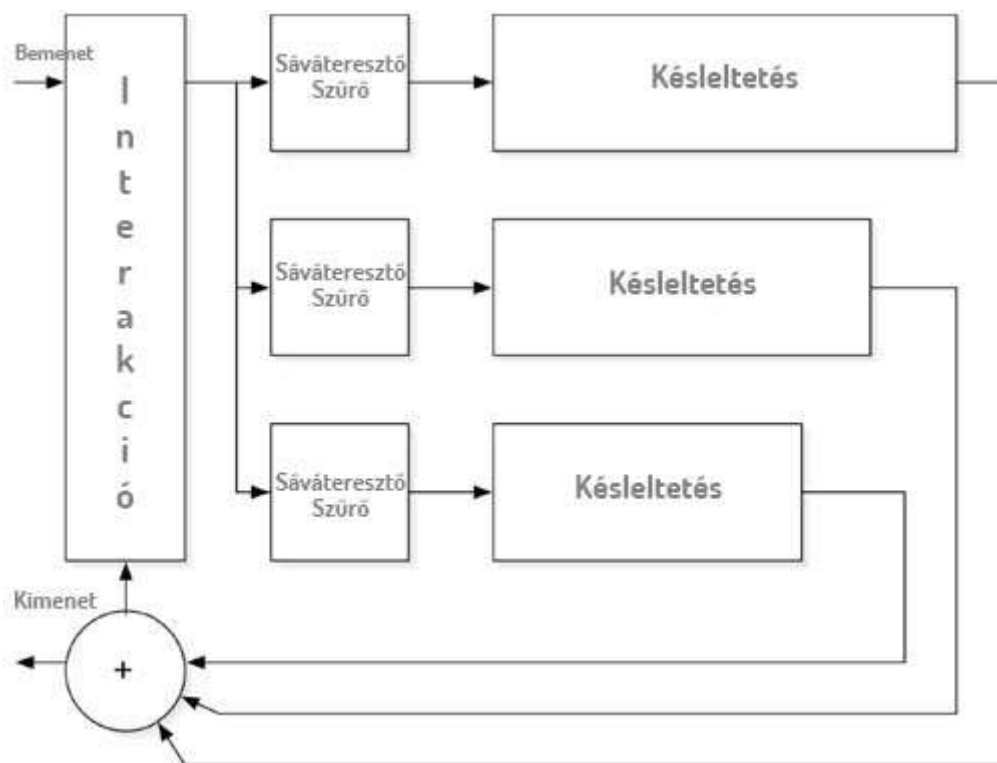


49. ábra 2D hullámvezető háló blokkvázlata

5.6.2. Banded waveguide:

A *Banded Waveguide Synthesis*, egy olyan fizikai modellező módszer, mely a szórt hangzású tárgyak, vagy erősen diszharmonikus rezonáns frekvenciájú tárgyak hatékony szimulálására való. 1999-ben Georg Essel és Perry Cook fejlesztette ki az eredeti módszert, hogy szintetizálja a hangvilla két ágán keletkező hajlított hangokat.

Ahogy a neve is sugallja, a *banded waveguide* szintézisben egy passzív vagy hirtelen gerjesztett vibrációs rendszernek a spektruma van frekvencia sávokra osztva, ahol minden sáv tartalmaz egy rezonáns üzemmódot. Minden egyes sávban a *digitális hullámvezetőt* arra használják, hogy modellezzék a hullám dinamikáját és az üzemmód rezonáns frekvenciáját. A hullám dinamika megtartása által, a szintézis algoritmus egy nemlineáris gerjesztési modellel használható (a hegedű vonójának a súrlódási modellje). A módszer lehetővé teszi még az *idiophone* modellezését is, mint például a *marimba*.



50. ábra Sávhatárolt hullámvezető blokkvázlata

A képen egy alap sávós *hullámvezető* struktúra látható. Ez hasonlít az alap *digitális hullámvezető* struktúrájához, azonban itt 3 késleltető van egy helyett, és egy interakciós modell adódott még hozzá, az alul-áteresztő szűrő pedig ki lett cserélve egy sáváteresztő szűrőre. Az ábra négy különálló részt mutat, amiket meg kell határozni. Mutatja a sávkorlátozó (sáváteresztő) műveletet (BP), a késleltető műveletet (Delay) és az interakciós műveletet (interaction). Továbbá a sáváteresztő/késleltető műveletet elvégző párokat.

5.7. Single Delay Loop Models (Egyszerű késleltető húrok modellek)

Az előző fejezetben bemutatott digitális *waveguide* alapján bizonyos helyzetben kialakítható egy SDL modell. Karjalainen et al. (1998) a gitárból származtatatta az SDL modellt. Csak egy erős jelű húr esetében lesz figyelembe véve a kimenet az átkötéseknél. Ez megfelel a klasszikus akusztikus gitár szerkezetének. Karjalainen et al. (1998) mutatta be a hangszedő kimenetének az esetét, amely megfelel az elektromos gitárnak. Egy folytonos idejű *waveguide* modellel kezdődik a Laplace tartományban, és egy diszkrét idejű modell kialakításával, amelyet egy SDL modellel lehet azonosítani.

A modell összetevőinek az átviteli függvénye a Laplace transzformációs tartományban van meghatározva. A Laplace transzformáció hatékony eszköze a lineáris folytonos idejű rendszerek elméletének. Különösen az időtartománybeli integráció és a származékos műveletek alakulnak a Laplace s változó által osztással és szorzással. A komplex s Laplace változó megváltoztatható $j\omega$ -val (ahol j az imaginárius része a $\sqrt{-1}$ -nek, ω a radian frekvencia $\omega=2\pi f$, és f a frekvencia Hertzben), annak érdekében, hogy levezethessük a megfelelő ábrázolást a Fourier-transzformációs tartományban, azaz a frekvencia tartományban. A diszkrét idejű megvalósításhoz, a folytonos idejű rendszer végül jól megközelíthető egy diszkrét idejű rendszerrel a z transzformációs tartományban.

6. ÉRTÉKELÉSI RENDSZER:

Ebben a dolgozatban bemutatott hangszíntézis módszerek a különféle típusú szintézis problémák miatt kerültek bemutatásra. Így ezeket a módszereket nem célszerű összehasonlítani az értékelési szempontok miatt. Az értékelés célja, hogy iránymutatást adjon, mely módszerek a legalkalmasabbak az adott hangszíntézis problémához.

Az itt bemutatott módszereket négy csoportba osztva tárgyaltam Smith (1991) taxonómiáján alapulva, hogy jobban össze lehessen hasonlítani azokat a technikákat, amelyek szorosan kapcsolódnak egymáshoz. A csoportok a következők: absztrakt algoritmusok, mintavételezett és feldolgozott felvételek, spektrális modell szintézisek és fizikai modellezések. A csoportosítás az egyes hangszíntézis módszerek előfeltételein alapulnak. Az absztrakt algoritmusok érdekes hangokat hoznak létre azokkal a módszerekkel, amelyeknek igazából kevés köze van a hangkeltés mechanizmusának a valós fizikai világban. A mintavételezett és a feldolgozott felvételek szintézise figyelembe veszi a meglévő hangeseményeket és akár közvetlenül is képes reprodukálni őket vagy új hangokat állíthat elő a további feldolgozás céljából. A spektrális modellező szintézis a hang tulajdonságainak az információit használja - a hallgató által érzékelt információkat. A fizikai modellező szintézis egy valódi hangszer hangkeltő mechanizmusát kíséri meg szimulálni. Ez a taxonómia úgyis értelmezhető, mint a szintézis rendszer felhasználója által generált feladatok figyelembe vétele.

Az értékelés szempontjából célszerű azonosítani a hangszíntézis problémákat. A teendők, hogy mely módszer a legalkalmasabb, a következők:

Absztrakt algoritmusok

- tetszőleges új hangok létrehozása,
- számítástechnikailag hatékony a meglévő zenei hangszerek mérsékelt minőségű szintéziseihez.

Mintavételezett és feldolgozott felvételek szintézise

- felvett hangok reprodukálása,
- felvett hangok összevonása és változtatása,

- rövid hang impulzuscsomagokat és felvételeket használva előállítani új hangeseményeket,
- kiváló hangminőséget igénylő alkalmazásokhoz.

Spektrális modellezés

a meglévő hangok szimulálásához és analíziséhez

- másoló szintézis (audio kódolás),
- hang jelenségének a tanulmányozása,
- hangmagasság-változtatás, időskála módosítás.

Fizikai modellek

- természetes hangszerek analíziséhez és szimulálásához
másoló szintézis,
hangszerek fizikai tanulmányozása,
fizikailag megvalósíthatatlan hangszercsaládok létrehozása.
- magas minőség irányítást igénylő alkalmazásokhoz

A hangszintézis módszereket tipikusan analízis és szintézis eljárásokra lehet osztani. Ezek a technikák külön-külön vannak értékelve, mivel általában különbözőek a követelményeik. Sok esetben az analízis elvégezhető *offline* és így a számítási idő költségeinek köszönhetően pontosabbak is. A szintézis rész jellemzően valós időben kell, hogy történjen, illetve rugalmasan, hogy irányítani is lehessen a szintézis folyamatot.

A hangszintézis módszer értékelésének megvitatásával Jaffe (1995) foglalkozott. Jaffe tíz szempontot javasolt, amit a következő fejezetben tárgyalunk néhány további kiegészítéssel. Ezeket a szempontokat hasznosítva hozzuk létre az értékelési rendszert. A következő fejezet az értékelési rendszert alkalmazva mutatja be a szintézis módszereket. Az eredmények egy táblázatba vannak összegyűjtve, hogy megkönnyítsék a módszerek összehasonlítását.

A tíz szempont foglalkozik a paraméterek használhatóságával; az előállított hangok minőségével, sokféleségével és természetességével, valamint a megvalósítással kapcsolatos kérdésekkel. Még egy kritériumot azonban szükséges megemlítenem az összehasonlítási táblázatban. A szintézis módszer párhuzamos végrehajtásának az alkalmasságát. Ezeket a szempontokat háromféleképpen osztályozhatjuk: gyenge, elfogadható és jó.

6.1. A paraméterek használhatósága:

Négy szempontból tárgyaljuk a paramétereket: intuitivitás (ösztönösség), természetesség, a paraméterek viselkedése, valamint a paraméter változásának az észlelhetősége.

Az ösztönösség azt jelenti, hogy egy vezérlő paraméter intuitív módon feltérképez egy zenei tulajdonságot vagy a hangszín minőségét. Az intuitív paraméterekkel a felhasználó könnyen képes megtanulni, hogyan lehet ellenőrizni a szintetikus hangszereket. Egy jelentős paraméterváltozásnak érzékelhetőnek kell lennie, hogy értelme is legyen. Ezeket a paramétereket *erősnek* tekintjük, ellentétben a *gyenge* paraméterekkel, amelyek alig hallható változásokat okoznak (Jaffe, 1995). A tendencia az, hogy minél több paramétere van egy szintézis rendszernek, az annál gyengébb. Azonban a túl nehéz paramétereket nehéz irányítani, ellenőrizni, mivel már egy kis változás a paraméter értékén drasztikus változást eredményez az előállított hangon, így nem számít, hogy mennyire intuitív a paraméter. A természetes vagy más néven fizikai paraméterek látják el a szintetikus hangszeren játszó zenészt egy valós hangszer viselkedését befolyásolva. Ezek olyan mennyiségeknek felelnek meg, amikben egy igazi hangszer zenésze járta, mint például, a húr hossza, a vonó vagy az ütő sebessége, a száj nyomása egy fúvós hangszernél, stb. A paraméter viselkedése szorosan összefügg a paraméter „linearitásával”. Egy paraméterben bekövetkező változásnak arányos változást kellene előállítani a létrehozott hangban.

Az ebben a leírásban bemutatott kritériumokat egy táblázatban foglaljuk össze az értékeléseknek megfelelően.

	gyenge	elfogadható	jó
Ösztönösség	+	++	+++
Észlelhetőség	+	++	+++
Természetesség	+	++	+++
Viselkedés	+	++	+++

6.2. A létrehozott hangok minősége és sokfélesége

Ebben a részben a létrehozott hangok tulajdonságainak a szempontjait tárgyaljuk. Beleértve a hang identitásának a robosztusságát, a módszer általánosságát és a rendelkezésre álló elemzési módszereket.

A hang robosztusságát az határozza meg, hogy mennyire jól tartják meg a hang eredeti hullámalakját (azonosságát), amikor módosításra kerülnek a paraméterek. Ez azt jelenti, hogy például egy klarinét modellnek úgy kell megszólalnia, mint egy klarinétnak, amikor különböző dinamikával és játéktípussal játszunk le, vagy például akkor is, ha a zenész úgy dönt, hogy a paraméter értékeinek változtatásával kísérletezik. Egy általános hangszíntézis módszer képes előállítani kiváló minőségű tetszőleges hangeseményeket. Minden létező hangszíntézis módszernek megvannak általában a hiányosságai. Kedvező, ha ismerjük ezeket a hiányosságokat.

Számos hangszíntézis módszerhez létezik analízis módszer is, hogy származtathassuk a színtézis paramétereket a felvett hangok jeleiből. Ez megkönnyíti a színtézis módszerek használatát, amely biztosítja az alapértelmezett paramétereket, amelyek ez után módosíthatóak, hogy lejátszható legyen a szintetikus hangszer. Az elemzési rész a módszerek számos szempontjából elengedhetetlen. Elméletileg másoló színtézist vagy egyéb optimális paramétereket lehet levezetni a legtöbb színtézis módszerekkel különböző optimalizálási módszereket használva. Jellemzően az elemzési szakasz gyakran használja a színtézis rendszerek ismereteit, hogy megbízható paramétereket kaphassunk.

Sok esetben az elemzés elvégezhető *offline* és tipikusan csak egyszer kell elvégezni minden egyes hangszernél. Így a precizitás kedvéért sokkal több hangsúlyt fordítunk a számítási idő költségeire, és ezzel összefüggésben a számítási hatékonyságot elvethetjük, mint szempontot az analízis módszereknél. A jelen dolgozatban a színtézis rendszerek analízis eljárásai alapján kell megítélni a pontosságot, az általánosságot és az igényeket a speciális eszközökhöz vagy hangszerekhez.

	gyenge	elfogadható	jó
A hang identitásának a robosztussága	+	++	+++
Általánosság	+	++	+++
Analizálási módszerek	+	++	+++

6.3. Végrehajtással kapcsolatos kérdések

Egy szintézis módszernek a végrehajtására vonatkozóan, számos fontos szempontnak kell megfelelnie. A technikák hatékonyságának a megítélése, a késleltetés és a vezérlés mértéke kerülnek becslések alá. A párhuzamos végrehajtás alkalmazhatósága is említésre kerül.

A hatékonyság további három részre osztható: számítási kapacitás, memória használat és a módszer vezérlése következtében a terhelés. Sok esetben a memória követelményeket kompenzálhatjuk a számítási költségek növelésével (Jaffe, 1995). A számítási költség *jó*, ha a módszer egy vagy több esetben könnyen fut valós időben, egy nem drága processzort meghajtva, *elfogadható*, ha a módszer csak egy esetben fut valós időben egy asztali gépen vagy PC-n és *gyenge*, ha a valós idejű végrehajtás nem lehetséges ajánlott hardver vagy szuperszámítógép nélkül.

A módszer vezérlő-adatfolyama hatással van mind az expresszivitásra és a számítási kapacitásra is. Általában több vezérlés lehetséges sűrű vezérlő-adatfolyammal, mint egy ritkával (Jaffe, 1995). A vezérlő-adatfolyam feldolgozása sokkal inkább költségesebb, hiszen magában foglalja az I/O-t külső eszközökkel vagy fájlokkal. Ebben az összefüggésben a vezérlő-adatfolyam kerül megítélésre, hiszen a vezérlés mennyisége lehetővé teszi az adatfolyam sűrűségének a vizsgálatát.

A valós idejű szintézis rendszerben mindig is jelen lesz késleltetés, mivel a rendszer alkalmas kell, hogy legyen a megvalósításra. Így a késleltetés a fő probléma a módszereknél, amelyek

blokk-számításokat alkalmaznak, csakúgy, mint a DFT. Más költséges számításigényes szintézis módszerrel is lehet előnyös, ha szűk hurkokat futtatunk, vagy akár száz kimeneti mintával felgyorsítsuk a kalkulációt. Alapvetően ez lehet az oka a késleltetési problémáknak. A minősítéseknél a *gyenge* jelenti, hogy a rendszer késleltetési ideje 10-100 milliszekundum vagy több, az *elfogadható* azt jelenti, hogy ha gyakorlatilag nincs semmi extra *overhead*, például az operációs rendszer késleltetése nem érzékelhető és a *jó* azt jelzi, hogy a módszer toleráns néhány *overhead* változatra.

	gyenge	elfogadható	jó
Számítási költség	+	++	+++
Memória használat	+	++	+++
Vezérlő-adatfolyam	+	++	+++
Késleltetés	+	++	+++
Párhuzamos feldolgozás	+	++	+++

A párhuzamos végrehajtáshoz való alkalmasság bizonyos helyzetekben fontos tényező. E tekintetben azt feltételezzük, hogy gyors kommunikáció érhető el a párhuzamos folyamatok között. A rendszer minősítése *jó* a párhuzamos végrehajtás alkalmasságát tekintve, ha könnyen osztható több folyamatra, hogy a kommunikáció az folyamatok között hozzávetőlegesen történjen a rendszer mintavételezési sebességének a szintjén. *Elfogadható* értékelést kap, ha a rendszert két kommunikáló folyamatra lehet osztani, vagy ha előnyös a számítást szétosztani magasabb kommunikációs szinteken. A módszer *gyenge* megítélésű, ha kevés vagy semmilyen előnye nincs a párhuzamos feldolgozásnak.

7. NÉHÁNY HANGSZÍNTÉZIS MÓDSZER ÉRTÉKELÉSE

Ebben a részben a hangszíntézis módszerek az előző fejezet szempontjai alapján kerülnek kiértékelésre. Az értékelések csoportosítva vannak módszerek szerint és táblázatba vannak gyűjtve, hogy lehetővé váljon a módszerek összehasonlítása. Meg kell jegyezni, hogy a cél nem az, hogy eldöntsük, melyik színtézis módszer a legjobb általánosságban, mert ez lehetetlen lenne. Az értékelésnek inkább egy iránymutatást kell adnia, hogy az adott hangszíntézis problémára a megfelelő módszert válasszuk.

Néhány módszer esetében olyanok a szempontok, hogy úgy érezzük, nem tudjuk elvégezni az értékelést.

7.1. Absztrakt algoritmusok vizsgálata

7.1.1. FM színtézis

Az FM színtézis paraméterei határozottan nem megfelelőek néhány szempontból: intuíció, fizikai állapot és viselkedés, mivel a modulációs paraméterek nem felelnek meg a zenei paraméterekhez vagy a zenei hangszerek paramétereihez, ugyanis a módszer igencsak nemlineáris. Így *gyenge* az értékelése az összes kategóriában. Figyeljük meg azonban, hogy az *I* modulációs indexű paraméter közvetlenül kapcsolódik az előállított jel sávszélességéhez. A módszer erős paraméterekkel rendelkezik, például a paraméterváltozások jól hallhatóak. Az észlelhetőség minősítése ezért *jó*.

Az FM színtézis azonban nem viselkedik jól, ha valódi hangszerek utánzására használjuk változó dinamikával és játék stílussal. A módszer paramétereit nagyon óvatosan kell

megváltoztatni, hogy ne veszítse el a hangszer identitását. A módszer minősítése *gyenge* az identitás robosztusságát tekintve. Az FM színtézis általánossága *jó*. Javasoltak analízis módszert az FM-hez, de a módszerek általános esetekben nem alkalmazhatóak megfelelően, így az analízis módszer tekintetében a minősítése *gyenge*.

Az FM hatékony megvalósítására készítettek egy népszerű módszert. Olcsó a megvalósítása, kevés memóriát használ és a vezérlő-adatfolyama is ritka. A minimális késleltetés teszi a módszert vonzóvá valós idejű színtézis célokra. A módszer ezekben a kritériumokban *jó* értékelést kapott. Az FM színtézis számítástechnikailag annyira olcsó, hogy egy FM hangszer forgalmazása nem kivitelezhető.

7.1.2. Waveshaping színtézis

A *waveshaping* paraméterei intuitívabbak (*elfogadhatóan*), mint az FM paraméterei, különösen, ha *Chebyshev* polinomokat használnak alakító függvényként. Egy egyszerű *Chebyshev* polinom súlyozásának a méretezése csak egy hasznos harmonikus változást idéz elő. Így a paraméterek nem nagyon érzékelhetőek és nem is annyira természetesek (*gyenge*). A paraméterezéstől függően a paraméterek viselkedése *elfogadhatóan* jónak tekinthető.

A *waveshaping* *elfogadhatóan* általános, abban, hogy könnyen előállítható tetszőleges harmonikus spektrum. *Waveshaping* színtézis után amplitúdó moduláció hozzáadásával diszharmonikus spektrum állítható elő. Zajos jeleket nem lehet könnyen előállítani. Spektrális analízist könnyen alkalmazhatunk, hogy megkapjuk minden harmonikus amplitúdóját. Ezeket az adatokat azért használják, hogy kinyerjék a *Chebyshev* polinomokat. Az analízis módszer minősítése így *jó*.

Ahogy az FM színtézis, a *waveshaping* is nagyon hatékony, viszont a forgalmazása nem kivitelezhető. A módszer minősítése számítási, memória, vezérlő-adatfolyam, valamint a késleltetés szempontjából *jó*.

7.1.3. Karplus-Strong színtézis

Bár kevés paramétere van a Karplus-Strong színtézisnek, azonban nagyon intuitívak, a változások jól hallhatóak és a viselkedésük is jó. Ezért a minősítése ezek szempontok alapján *jó*. Alapvetően a módszernek csak egy paramétere van a hangmagasságra és egy a hangszín meghatározására (például húr vagy ütő). A természetesség minősítése ezért *elfogadható*.

A KS színtézis robusztus abban, hogy úgy szólal meg, mint egy pengetős húr vagy egy dob, még ha a paraméterek meg is változnak, így az identitás robusztussága *jó*. Általánosságban a módszer *gyenge*. A KS színtézishez analízis technikák nem érhetőek el, de kapcsolódó hangszíntézis módszerek léteznek.

Más absztrakt algoritmusokhoz hasonlóan, a KS nagyon érdekes/vonzó a valós idejű alkalmazásokban. A végrehajtással kapcsolatos minősítése ugyanaz, mint az FM színtézis és *waveshaping* esetében.

7.2. Mintavételezés és feldolgozott felvételek vizsgálata

7.2.1. Mintavételezés

A mintavételezési színtézisben egy felvett hangjel lehetőleg hurokkal van visszajátszva az állandósult állapotú részben. A mintavételezés csak a *gain* paraméterekkel és a hangjegyekkel vezérelhető. Úgy döntöttünk, hogy nem adunk értékelést ezekhez a triviális paraméterekhez, hogy ne zavarjuk meg az egyéb színtézis módszerek értékelését.

A mintavételezés nagyon általános (*jó*) abban, hogy minden hangot rögzíteni lehet és mintát lehet venni belőlük. A hang „identitása” megmarad a különböző játéktílussal és feltételekkel. A hangalak (identitás) robusztusságának az értékelése *elfogadható*. Elérhetőek analízis módszerek, melyek meghatározzák a hurok határértékeit és a harmonikus hangokkal általában jó eredményeket adnak lejátszáskor.

A mintavételezés számításigényét tekintve nagyon hatékony (*jó*), viszont nagyon sok memóriát igényel (*gyenge*). A vezérlő adatfolyama ritka (*jó*) és a késleltetése is alacsony (*jó*).

Egy mintavételező hangszer forgalmazása nem teljesíthető maradéktalanul, mert egy szerverre lenne hozzá szükség a nagy memória kapacitás miatt.

7.2.2. Többszörös hullámtábla szintézis

A többszörös hullámtábla szintézis módszer sokféle képen paraméterezhető és a szintézis végeredménye nagyban függ a hullámtáblában tárolt jelektől. Ezért úgy döntöttünk, hogy nem adunk értékelést a módszer paramétereire vagy a hang identitásának a robosztusságára.

A módszer általános (*jó*) és az analízálási módszer számos végrehajtáshoz elérhető (*elfogadható*).

A módszer számítási szempontból viszonylag könnyen megvalósítható, de sok memóriát igényel (*gyenge*). A vezérlő-adatfolyam számítási szempontból nem túl költséges (*jó*) és a késleltetési idő is alacsony szinten tartható (*jó*). Csakúgy, mint a mintavételezés, különálló hullámtábla szerver képes csökkenteni a többszörös hullámtábla szintézis memória követelményeit, feltéve, hogy gyors kapcsolatok állnak rendelkezésre. A megosztott párhuzamos feldolgozáshoz való alkalmassága *elfogadható*.

7.2.3. Granulár szintézis

A granulár szintézis olyan technikák készlete, ami igen sokban tér el egymástól a paraméterezésben és implementációban. A koncepció általános értékelését megkíséreljük. A legprimitívebb alakban a granulár szintézis paraméterei közvetlenül vezérlik a szemcséket. A szemcsék száma jellemzően nagyon nagy és bonyolultabb eszközök vezérlik őket, amit ki kell használni. A paraméterek értékelése tehát *gyenge* az intuitivitásban, az észlelhetőségben és a természetességben is. A rendszer lineáris és a paraméterek viselkedése *jó*.

Léteznek analízálási módszerek a szinkron hangmagasság granulár szintézishez (PSGS), melyek igen hatékonyak (*jó*). Mivel az aszinkron módszer (AGS) nem kísérli meg modellezni vagy reprodukálni a felvett hangjeleket, ezért nincs szükség analízálási eszközökre. A

granulár szintézis módszerek általánosak (*jó*), és a PSGS-el az identitás robosztusságát jól megőrzi (*jó*).

A módszer végrehajtása elége hatékony, ahogyan a memória követelmények is (*elfogadható*), hiszen a szemcsék rövidek és így jellemzően azt feltételezzük, hogy a jelek néhány alap szemcséből állnak. Az alacsony szintű vezérlő-adatfolyam nagyon sűrűvé válhat, különösen az AGS-el (*gyenge*).

7.3. A Spektrális modellek vizsgálata

7.3.1. Additív szintézis

Az egyszerű additív szintézis paraméterei közvetlenül vezérlik a szinuszos oszcillátorokat. A vezérlőadatok csökkentésére számos lehetőség van és néhány más spektrális modellező módszerekről is szó fog esni. Ebben a dolgozatban csak az egyszerű additív szintézist tárgyaltam. A paraméterek *elfogadhatóan* intuitívak, mivel a frekvenciák és az amplitúdók könnyen megadhatóak. A paraméterek viselkedése *jó*, mivel a módszer lineáris. A paraméterek észlelhetősége és természetessége viszont *gyenge*.

Az additív szintézis elméletben tetszőleges hangokat képes szintetizálni, ha korlátlan számú oszcillátor áll a rendelkezésre. Ez persze hamar lehetetlenné válik, mivel a zajos jeleket nem lehet hatékonyan modellezni és így az általánosság értékelése *elfogadható*. Az analízis módszere (*jó*), például az STFT, könnyen elérhető additív szintézis, amire tekinthetünk úgy, mint néhány más spektrális modellező módszer szintézis részeként. Az identitás robosztussága nincs értékelve, mivel egy szintetikus hangszer vezérléséhez bonyolultabb vezérlési technikákra lenne szükség.

Egy egyszerű szinuszos oszcillátor hatékonyan megvalósítható, de az additív szintézisben jellemzően számos oszcillátorra van szükség. A számítási igény értékelése *elfogadható*. A vezérlő adat nagy memóriát igényel (*gyenge*), és a vezérlő-adatfolyam is nagyon sűrű (*gyenge*). A késleltetési idő alacsony, mivel az oszcillátorok párhuzamosan működnek (*jó*). A párhuzamos végrehajtás lehetetlenné válhat, ha túl sok oszcillátort szeretnénk használni. Az

osztott végrehajtásban az oszcillátorokat és a megfelelő vezérléseket csoportosítani kell. Az osztott párhuzamos feldolgozás alkalmazhatóságának az értékelése *elfogadható*.

7.3.2. Szubtraktív szintézis

A legegyszerűbb szubtraktív szintézis nagyon hasonlít az előbbi szintézishez azzal a különbséggel, hogy a kivonásra egy szűrőt használ. A szubtraktív szintézis paraméterei tartalmazzák a gerjesztési jel választhatóságát és az időben változó szűrő együtthatóit. A paraméterek ezért itt is elfogadhatóan intuitívak. A paraméterek viselkedése *jó*, de a szűrő tulajdonságait megfelelően kell paraméterezni. A tulajdonságok észlelhetősége a szűrő parametrizálásától függ, ezért az értékelése *elfogadható*, de a természetességük *gyenge*.

Az általánosság tekintetében e szintézis módszer is csak *elfogadható*, habár ezzel a módszerrel is bármilyen hang szintetizálható, ha korlátlan számú oszcillátor áll a rendelkezésre. Az analízisi módszer tekintetében az értékelés *jó*, hasonlóképpen az additív szintézishez. Az identitás robusztussága mérsékelten *elfogadható*, mivel függ a paraméterek helyes beállításától.

A gerjesztés és a szűrő *elfogadható* mértékben hatékony a végrehajtáshoz. A szubtraktív szintézisben is jellemzően számos oszcillátorra lehet szükség, ezért a számítási igény értékelése *elfogadható*. A vezérlő adatfolyam függ a modellezett jeltől, ezért az értékelése *elfogadható*. A szubtraktív szintézis késleltetési ideje *jó*, mivel alacsony. A párhuzamos végrehajtás itt is lehetetlenné válhat, ha túl sok oszcillátort szeretnénk használni, ezért az értékelése *elfogadható*.

7.3.3. Spektrális modellező szintézis

A spektrális modellező szintézis additív szintézist használ, hogy determinisztikus (harmonikus) komponenseket állítson elő és *source-filter* szintézist, hogy a szintetikus jel sztochasztikus (zajos) összetevőit állítsa elő. A paraméterek a determinisztikus összetevők amplitúdó és frekvencia pályákból és a sztochasztikus rész spektrális burkológörbéből állnak. Ezen paraméterek értékelése *elfogadható* az intuitivitást tekintve. A vizsgált jel

módosításához, hogy értelme is legyen, magasabb szintű vezérléseket kell használni, hogy csökkentsék a vezérlő adatokat. Az észlelhetősége és a természetessége így *gyengének* értékelhető. A paraméterek viselkedés *jó*.

Az identitás robosztussága *jónak* értékelhető, mivel a jel összetétele biztosítja azt, hogy a determinisztikus és sztochasztikus rész külön-külön legyen összeállítva. Ezáltal az *attack* és az egyensúlyi állapot részei jobban vezérelhetőek. A spektrális modellező szintézis megítélése általánosságban *jó*, és az analízisi módszert tekintve szintén.

A számítási költség *elfogadható*. A módszer több memóriát igényel, mint például a MQ algoritmus, de így is még mindig *elfogadható* a memória használatát tekintve. A vezérlő adatfolyam *elfogadhatóan* ritka, mivel az additív *source-filter* paraméterei interpolálva vannak az STFT frémek között. A késleltetési ideje szintén *elfogadható*. Az additív és a *source-filter* szintézis felosztható különböző párhuzamos feldolgozásokra és így az alkalmassága az osztott párhuzamos feldolgozásokhoz *elfogadható*.

7.3.4. Tranziens modellező szintézis

A tranziens modellező szintézis a spektrális modellező szintézis egy kiterjesztése, a tekintetben, hogy különálló zajként és tranziens jelként veszi számításba a maradék jel további feldolgozását. Így azt mondhatjuk, hogy a TMS általánosabb az SMS-nél és több számítással is jár. A két módszer elég közel áll egymáshoz az azonos minősítés tekintetében.

7.3.5. Formant szintézis

A *formant* szintézis paraméterei szabályozzák az alap frekvenciát és a szintetizált hang jelek formánsainak a struktúráját. A paraméterek megítélése *elfogadható* intuitivitás és természetesség tekintetében különösen, amikor a módszert az emberi hangkeltő rendszer szimulálására használjuk. Az észlelhetősége *jó*, hiszen általában csak néhány formáns van jelen a beszédben vagy az énekben. A paraméterek viselkedése *jó*.

A módszer *elfogadhatóan* általános, mivel képes előállítani jó minőségű harmonikus énekhang jeleket, illetve zenei hangszereket. A lineáris előrejelzés gondoskodik az analízisi módszerről (*jó*) és az előállított hang jól megőrzi az identitását (*jó*).

A módszer hatékonyan végrehajtható, ha különböző *formant* szintézisek vannak eltárolva a hullámtáblákban (*jó*). Ez természetesen megnöveli a szükséges memóriát (*elfogadható*). A vezérlési arány meglehetősen kevés és a késleltetési idő is alacsony (*jó*). Egyetlen *formant* hangszer párhuzamos feldolgozása sem tűnik megvalósíthatatlannak, mivel a gerjesztő jel megosztott az összes *formant* generátorok által (*gyenge*).

7.4. Fizikai modellek vizsgálata

7.4.1. Véges differencia módszerek

A véges differencia módszerek paraméterei közvetlenül megfelelnek a modellezett hang fizikai paramétereinek. Így a paraméterei nagyon természetesek, intuitívak és az észlelhetőség, illetve a viselkedés értékelése is szintén *jó*, mivel a vibrációs mozgás feltételezhetően lineáris.

Bár a módszer elméletben alkalmazható tetszőleges hang előállítási rendszerekhez, jellemzően egy új végrehajtás szükséges. Így a véges differencia módszerek általánosságát tekintve *elfogadhatóak*. Egy hangszer hangolt modellje sokkal jobban viselkedik, mint az eredeti és az identitását is jól megtartja (*jó*). Léteznek analízisi módszerek, és bár lehet, hogy az eredmény nagyon jó és gyakran tartalmaz speciális mérési eszközöket, sok időt és erőfeszítést igényel (*elfogadható*).

A véges differencia módszerek számítástechnikailag nem hatékonyak (*gyenge*), és *elfogadható* mennyiségű memória szükséges hozzájuk. A vezérlő adatfolyam a gerjesztéstől függ, de nagyon ritka a megpengetett vagy ütött húrok esetén (*jó*). A módszer nem okoz problémát a késleltetési idő szempontjából, ha elegendő számítási kapacitás áll a rendelkezésre (*jó*). A módszer kiválóan alkalmas az osztott párhuzamos feldolgozáshoz (*jó*),

mivel jelentős javulásokat valósíthat meg azáltal, hogy több alstruktúrába osztja a rendszert, különböző eljárásokat futtatva.

7.4.2. Modal szintézis

A modális szintézis paraméterei a modellezett struktúra modális adatából és a gerjesztésből vagy egyéb adatból állnak. A paraméterek nem eléggé intuitívak (*elfogadható*) és az egy hangnemben bekövetkező változás nehezen érzékelhető, ha a hangnemek száma nagy (*gyenge*). Ezek közvetlenül megfelelnek a fizikai struktúráknak, a természetességnek és a viselkedésnek (*jó*).

Léteznek analízis módszerek a rendszerhez, amelyek megbízható és pontos eredményeket nyújtanak. Ezek azonban nagyon bonyolultak és drágák (*elfogadható*). A rendszer általános, mivel akármilyen rezgő objektumot ki lehet fejezni modális adatként. A módszer mechanizmusa miatt nem állíthatóak elő tetszőleges hangok a nem természetes objektumok vonatkozásában. Az általánosság értékelése *elfogadható*. A modellezett struktúra nagyon jól megtartja az identitását (*jó*), mivel jellemzően a gerjesztési jel vezérli.

A módszer megvalósítható párhuzamos másodrendű rezonátorok készletével, amelyek számítástechnikailag hatékonyak. A rezonátorok száma nagyon nagyra nőhet és ezért a számítási hatékonyságának az értékelése csak *elfogadható*. A modális adat számos memóriát igényel (*gyenge*). A gerjesztő jel határozza meg a vezérlő adatfolyamot (a statikus struktúra érdekében) és ezért ennek az értékelése *gyenge*. Az alstruktúrák hatékonyan megoszthatóak és párhuzamosan feldolgozottak (*jó*).

7.4.3. CORDIS

A CORDIS rendszer paramétereire nem volt kapcsolódó irodalom, ezért ezek nincsenek értékelve.

Mivel a módszer egy fizika-alapú rendszer leírását használja (ami rezeg), jól megtartja a hang identitását (*jó*) különböző jelentős gerjesztésű jelekkel. Nincs irodalomban leírt analízis

módszer hozzá (*gyenge*). A rendszer általánosságát véve *elfogadható*, mivel elméletben képes modellezni tetszőleges rezgő tárgyakat. A módszer nem nyújt könnyű megoldást arra nézve, hogy tetszőleges hangokat állítsunk elő.

A módszer számítástechnikai megítélése szerint *elfogadható*, habár az alapvető elemeket hatékonyan számolja. Ez magyarázza az *elfogadható* minősítést a memória igényt tekintve. A vezérlő sebesség a paraméterezéstől függ, ami itt nincs értékelve. A CORDIS késleltetési ideje kicsi (*jó*). A CORDIS rendszer jól alkalmazható a párhuzamos feldolgozásokhoz (*jó*) (Rocchetto, 1998).

7.4.4. Digital Waveguides szintézis

A digitális *waveguide* szintézis paraméterei jól megfelelnek a hangszer természetes struktúrájának, ezért intuitivitás és a természetesség szempontjából az értékelésük *jó*. A paraméterváltozások általában hallhatóak (*jó*) és jól is viselkednek a lineáris modellekkel. Mivel néhány *waveguide* modell nemlineáris csatolású, a viselkedés értékelése *elfogadható*.

A hangszer identitása nagyon jól megmarad (*jó*). A módszer hangszerek szimulálására használható egydimenziós rezgő objektumnál, mint például a húros vagy fúvós hangszerek és így az általánosságát tekintve az értékelése *elfogadható*. A lineáris modellekhez tartozó analízis módszerek hatékonyak, de ezek sajnos nem érhetőek el a nemlineáris modellek esetén (*elfogadható*).

A digitális *waveguide* rendkívül hatékonyan megvalósítható, de a modell egyéb struktúrákat is egyesít, amelyek növelik a számítási igényt. Ezért a számítási hatékonyságának az értékelése *elfogadható*. A digitális *waveguide* modellek kevés memóriát igényelnek a gerjesztő jelen kívül. Egy néhány másodperces jó minőségű pengetős húr hangja igen kevés memória felhasználásával is előállítható (*jó*). A vezérlő adatfolyam a modellezendő hangszertől függ és ezért itt az értékelése *elfogadható*. A módszernek nem jelent problémát a késleltetés, mivel a rezgő struktúrájú modellek hatékonyan szétoszthatóak alstruktúrákra, amelyek párhuzamosan számolnak (mindkettő értékelése *jó*).

7.4.5. Waveguide Meshes

A *waveguide* hálók paraméterei *elfogadhatóan* intuitívak, mivel ezek megfelelnek a 2D- vagy 3D rezgő rendszer tulajdonságainak és a gerjesztésnek. A paraméterek fizikailag is megfeleltethetőek, észlelhetőek és a viselkedésük is jó, mivel a háló lineáris, így az értékelése az összes kategóriában *jó*.

Analizálási módszer nem volt található az irodalomban, ezért az értékelése *gyenge*. A módszer általánosságát tekintve *elfogadható*, mivel a 2D vagy 3D tárgyak szimulálására alkalmazható. Az identitás robusztussága nincs értékelve.

A módszer számítástechnikailag drága és sok memóriát igényel (mindkettő értékelése *gyenge*). A vezérlő adatfolyam *elfogadhatóan* ritka, mivel csak a gerjesztésből áll. A módszernek nincsenek késleltetési problémái (*jó*), habár az összetettebb struktúrák valós idejű implementációja elérhetetlen egy drága szuperszámítógép nélkül. A modell egyik legfontosabb előnye az, hogy szétosztható tetszőleges alstruktúrákra, amelyek számítása párhuzamosan történik (*jó*).

7.5. Értékelések összefoglalása

Ahogy már korábban említettem a különféle szintézis módszerek mind elveikben, mind szintetizálási folyamataikban rendkívül eltérőek, ezért ezek objektív összehasonlítása nem célszerű, de a meghatározott szempontok alapján értékelésre kerültek.

Az előző fejezetben tárgyalt értékelések eredményeit táblázatba foglaltam. Tapasztalataim azt mutatják, hogy az absztrakt algoritmusok és a mintavételező technikák a legerősebbek, legalkalmasabbak a végrehajtási kategóriában. Leginkább a digitális szintetizátorokban és számítógépes zenealkotásban használatosak a mai napig és nem is véletlen, hogy ezek a hang szintetizálási módszerek a legjobbak a számítógépes szoftvereket használó zenei producerek számára.

A spektrális modellek nagyrészt általánosak (mivel a hangok széles skáláját képesek modellezni), robusztusak (mivel a hangok rendkívül összetett hullámalakokkal írhatóak le) és analízis módszerek is elérhetőek számukra, ezért ezek a legerősebbek a hang kategóriában. A fizikai modellek rendkívül hatékony parametrizálást tesznek lehetővé, ezért ez a legerősebb a paraméter kategóriában. Ezek a modellek a legalkalmasabbak, legjobbak a valódi

hangszerek hangjainak a megvalósítására, hiszen a neve is mutatja, hogy a valódi hangszerek fizikai háttérét használja a hangok szintetizálására.

	Paraméterek				Hang				Megvalósítás															
	Intuitív Észlelés				Természeti Visekeds				Rövid és Hosszú Állapotok				Számítás				Mentő Vektorok				Készlet			
Abstract																								
FM	+	++	+	+	+	++	+	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++		
Wave shaping	++	+	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++		
K5	++	++	++	++	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++		
Mintavétel																								
Mintavétel					++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		
Többszörös Mintavétel									++	++	++	++	++	++	++	++	++	++	++	++	++	++		
Granular	+	+	+	++	++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		
Szerkezet																								
Additiv	++	+	+	++	++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		
Subtraktív	++	++	+	++	++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		
SW5	++	+	+	++	++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		
TMS	++	+	+	++	++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		
Formant	++	++	++	++	++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		
Ezért																								
Model	++	+	++	++	++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		
CORIS					++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		
FD módszer	++	++	++	++	++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		
Waveguide	++	++	++	++	++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		
MG Meszes	++	++	++	++	++	++			++	++	++	++	++	++	++	++	++	++	++	++	++	++		

51. ábra Szintézisek értékelése

7.6. Megvalósított színtézisek összefoglalása CPU és RAM felhasználtságuk alapján

	Max. CPU [%]	Max. RAM [Mb]	Átl. CPU [%]	Átl. RAM [Mb]
Hullámtábla színtézis	9,5	486,6	4,8	486,1
Granulár színtézis	26,4	485,5	14,6	476,9
Additív színtézis	11,9	470,8	8,3	469,9
Szubtraktív színtézis	13,2	484,15	9,76	483,1

Az általam megvalósított színtézisek CPU, illetve RAM felhasználtságát foglalja össze a fenti táblázat. Összehasonlítani nem érdemes ezeket a színtézis módszereket, mivel mindegyik különböző alapelvet követ, ez alatt azt értem, hogy valamelyik csak 1 mintát használ, míg a másik több ezret, az egyikhez kevés művelet szükséges, míg a másikhoz több száz vagy ezer.

8. Összefoglalás

A diplomamunkámban számos hang szintézis módszert mutattam be és hasonlítottam össze. A módszereket Smith taxonómiája alapján négy csoportra osztottam. Minden kategóriába reprezentatív példákat választottam és ezeknek a módszereknek a működését részletesen leírtam. Néhány kiválasztott hangszintézis módszert MATLAB program segítségével meg is valósítottam.

Az absztrakt algoritmusokon alapuló módszerek közül hármat választottam: az FM szintézist, a hullámalakításos szintézist és a Karplus-Strong algoritmust. Szintén három módszert tárgyaltam az előre rögzített hangok felhasználásával: a mintavételezéses szintézist, a többszörös hullámtábla szintézist és a granulár szintézist.

A spektrális modellező kategóriában két hagyományos hang szintézis módszer került bemutatásra, melyek a következők: additív szintézis, szubtraktív szintézis. Ebben a fejezetben tárgyaltam a spektrális modellező, illetve a tranziens modellező szintéziseket is. Végül az emberi beszéddel kapcsolatos szintézis módszer is említésre került, ami a formáns szintézis.

Három fizikai modellező szintézis vizsgálata is bemutatásra került, amelyek numerikus akusztikát használnak. Először azok a modellek kerültek bemutatásra, amelyek véges differencia módszert használnak. Itt a húros, illetve ütős hangszerek megvalósításához alkalmazható módszereket mutattam be. Ezek után az alaki szintézis módszert tárgyaltam. Végül a CORDIS rendszer zárta ezt a kategóriát, amely a rezgő objektumokat modellez tömeg-rugó hálózatok segítségével.

A fizikai modellező kategóriát folytatva bemutattam a digitális hullámvezetőket, illetve alkalmazásait. A hullámvezető hálók, amelyek tipikusan 2D vagy 3D modellek is tárgyalásra kerültek. Végezetül a *single delay loop* (SDL) módszer, más néven egyszerű késleltető hurok leírása zárta a szintézis módszerek tárgyalását.

A négy kategóriába csoportosított hangszintézis módszereket követően, *Jaffe* értékelési szempontjait felhasználva kerültek elbírálásra a módszerek. A módszerek alkalmassága a párhuzamos feldolgozás szempontjával került kiegészítésre az értékelési rendszer. Az egyes

módszerek értékelésre kerültek a vizsgálati szempontok megbeszélésével kapcsolatosan. A szempontok értékelését minőségi mérésekkel végeztem minden egyes módszerhez. Végül az értékelést egy táblázatban foglaltam össze, megkönnyítve az összehasonlítást az egyes módszerek között. Megfigyeléseim szerint az absztrakt algoritmusok és mintavételezési technikák a legerősebbek az implementációs kategóriában. A spektrális modellek általánosak, robusztusak és analízis módszerek is elérhetőek hozzá. Ezek a legerősebbek a hang kategóriában. A fizikai modellező algoritmusok rendkívül 'öszönös' parametrizálást alkalmaznak és ezért ezek a legerősebbek a paraméter kategóriában.

9. Summary

In my thesis, a lot of contemporary sound synthesis method have been talked over and compared. The methods are based on Smith's taxonomy and tabulated into four category. All categories chosen and described in detail the operation of these methods of illustrative examples. Some of the selected sound synthesis methods are implemented in MATLAB.

The methods based on algorithms from abstract chosen three: FM synthesis, waveshaping synthesis and Karplus-Strong algorithms. Also discussed three ways of using recording of sounds: sampling synthesis, multiple wavetable synthesis and granular synthesis.

In the spectral model category two traditional linear sound synthesis were presented which are follows: additive synthesis, subtractive synthesis. In this chapter I introduced the spectral modeling and the transient modeling synthesis as well. Finally, methods for synthesizing the human speech has been mentioned which is the formant synthesis.

Study of three physical modeling synthesis have been shown that use numerical acoustics. Firstly, the models were presented which uses finite difference methods. I have presented the string and percussion instruments used for the implementation of the methods. After this, the modal synthesis method was discussed. Finally, the CORDIS system closed this category, which is modeled by a mass-spring vibrating objects through the network. Continuing the physical modeling category introduced the digital waveguides and its applications. The waveguide meshes that have been discussed typically 2D or 3D models. Finally, the single delay loop (SDL) method closed the paragraph.

Following the four categories grouped by sound synthesis methods were evaluated by Jaffe's assessment criteria. The various methods have been evaluated regarding to the test criteria. The evaluation carried out aspects of quality measurements for each method. Finally, the evaluation are summarized in a table to facilitate comparisons between different methods. According to observations the abstract algorithms and sampling techniques were the strongest category for implementation. The spectral models are the strongest in the sound category. The physical modeling algorithms are very intuitive parameterization is used and therefore they are the strongest of the parameter category.

10. Irodalomjegyzék

- [1] Adrien, J. M. 1989. Dynamic modeling of vibrating structures for sound synthesis, modal synthesis, *Proceedings of the AES 7th International Conference*, Audio Engineering Society, Toronto, Canada, 291-300.
- [2] Adrien, J. M. 1991. The missing link_ modal synthesis, in: G. D. Poli, A. Piccialli and C. Roads (eds), *Representations of Musical Signals*, The MIT Press, Cambridge, Massachusetts, USA, 269-297
- [3] Arfib, D. 1978, "Digital Synthesis of Complex Spectra by Means of Multiplication of Non-Linear Distorted Sine waves, *Journal of the Audio Engineering Society* 27(10): 757-768.
- [4] Bate, J. A. 1990. The effect of modulator phase on timbres in FM synthesis, *Computer Music Journal* 14(3): 38-45.
- [5] Bennett, G. and Rodet, X. 1989. Synthesis of the singing voice, in: M. V. Mathews and J. R. Pierce (eds), *Current Directions in Computer Music Research*, The MIT Press, Cambridge, Massachusetts, chapter 4, 19-44.
- [6] Borin, G. and Giovanni, D. P. 1996. A hysteretic hammer-string interaction model for physical model synthesis, *Proceedings of the Nordic Acoustical Meeting*, Helsinki, Finland, 399-406.
- [7] Borin, G., de Poli, G. and Sarti, A. 1997b. Musical signal synthesis, in: C. Roads, S. T. Pope, A. Piccialli and G. De Poli (eds), *Musical Signal Processing*, Swets & Zeitlinger, Lisse, the Netherlands, chapter 1, 5-30.
- [8] Cadoz, C., Luciani, A. and Florens, J. 1983. Responsive input devices and sound synthesis by simulation of instrumental mechanisms: the CORDIS system, *Computer Music Journal* 8(3): 60-73.
- [9] Cavaliere, S. and Piccialli, A. 1997. Granular synthesis of musical signals, in: C. Roads, S. T. Pope, A. Piccialli and G. De Poli (eds), *Musical Signal Processing*, Swets & Zeitlinger, Lisse, the Netherlands, chapter 5, 155-186.
- [10] Chaigne, A. 1992. On the use of finite differences for musical synthesis. Application to plucked stringed instruments, *Journal d'Acoustique* 5(2): 181-211.

- [11] Chaigne, A. and Askenfelt, A. 1994a. Numerical simulations of piano strings. I. A physical model for a struck string using finite difference methods, *Journal of the Acoustical Society of America* 95(2): 1112-1118.
- [12] Chaigne, A. and Askenfelt, A. 1994b. Numerical simulations of piano strings. II. Comparisons with measurements and systematic exploration of some hammer-string parameters, *Journal of the Acoustical Society of America* 95(3): 1631-1640.
- [13] Chowning, L. M. 1973. The synthesis of complex audio spectra by means of frequency modulation, *Journal of the Audio Engineering Society* 32(7): 526-534. Reprinted in C. Roads and J. Strawn, eds. 1985. *Foundations of Computer Music*. Cambridge, Massachusetts: The MIT Press. 6-29.
- [14] Cook, P. R. 1991. TBone: an interactive waveguide brass instrument synthesis workbench for the NeXT machine, *Proceedings of the International Computer Music Conference*, Montreal, Canada, 297-299.
- [15] Cook, P. R. 1992. A meta-wind-instrument physical model, and a meta-controller for real time performance control, *Proceedings of the International Computer Music Conference*, International Computer Music Association, San Francisco, CA., 273-276.
- [16] De Poli, G. 1983. A tutorial on digital sound synthesis techniques, *Computer Music Journal* 7(2): 76-87.
- [17] De Poli, G. and Piccialli, A. 1991. Pitch-synchronous granular synthesis, in: G. D. Poli, A. Piccialli and C. Roads (eds), *Representations of Musical Signals*, The MIT Press, Cambridge, Massachusetts, USA, 391-412.
- [18] Dolson, M. 1986. The phase vocoder: a tutorial, *Computer Music Journal* 10(4): 14-27.
- [19] Eckel, G., Iovino, F. and Caussé, R. 1995. Sound synthesis by physical modelling with Modalys, *Proceedings of the international Symposium on Musical Acoustics*, Dourdan, France, 479-482.
- [20] Evangelista, G. 1993. Pitch-synchronous wavelet representation of speech and music signals, *IEEE Transactions on Signal Processing* 41(12): 3312-3330.
- [21] Fletcher, N. H. and Rossing, T. D. 1991. *The Physics of Musical Instruments*, Springer-Verlag, New York, USA, 620.
- [22] Florens, J.-L. and Cadoz, C. 1991. The physical model: modeling and simulating the instrumental universe, in: G. D. Poli, A. Piccialli and C. Roads (eds), *Representations of Musical Signals*, The MIT Press, Cambridge, Massachusetts, USA, 227-268.

- [23] Fontana, F. and Rocchesso, D. 1995. A new formulation of the 2D-waveguide mesh for percussion instruments, *Proceedings of the XI Colloquium on Musical Informatics*, Bologna, Italy, 27-30.
- [24] Goodwin, M. and Rodet, X. 1994. Efficient Fourier synthesis of nonstationary sinusoids, *Proceedings of the International Computer Music Conference*, Aarhus, Denmark, 333-334.
- [25] Gordon, J. W. and Strawn, J. 1985. An introduction to the phase vocoder, in: J. Strawn (ed.), *Digital Signal Processing: An Anthology*, William Kauffmann, Inc., chapter 5, 221-270.
- [26] Harris, F. J. 1978. On the use of windows for harmonic analysis with the discrete Fourier transform, *Proceedings of the IEEE* 66(1): 51-83.
- [27] Hiller, L. and Ruiz, P. 1971a. Synthesizing musical sounds by solving the wave equation for vibrating objects: part 1, *Journal of the Audio Engineering Society* 19(6): 462-470.
- [28] Hiller, L. and Ruiz, P. 1971b. Synthesizing musical sounds by solving the wave equation for vibrating objects: part 1, *Journal of the Audio Engineering Society* 19(7): 542-550.
- [29] Hirschman, S. E. 1991. Digital Waveguide Modeling and Simulation of Reed Woodwind Instruments, *Technical Report STAN-M-72*, Stanford University, Dept. of Music, Stanford, California.
- [30] Horner, A., Beauchamp, J. and Haken, L. 1993. Methods for multiple wavetable synthesis of musical instruments tones, *Journal of the Audio Engineering Society* 41(5): 336-356.
- [31] Jaffe, D. A. and Smith, J. O. 1995. Ten criteria for evaluating synthesis techniques, *Computer Music Journal* 19(1): 76-87.
- [32] Jaffe, D. A. and Smith, J. O. 1983. Extensions of the Karplus-Strong plucked-string algorithm, *Computer Music Journal* 7(2): 56-69.
- [33] Lazzarini, V., J. Timoney and T. Lysaght 2007, "Adaptive FM synthesis". Proceedings of the 10th Intl. Conference on Digital Audio Effects (DAFx07). Bordeaux: University of Bordeaux: 21-26.
- [34] Lazzarini, V., J. Timoney and T. Lysaght 2008, "The Generation of Natural-Synthetic Spectra by Means of Adaptive Frequency Modulation". *Computer Music Journal*, 32 (2): 12-22.

- [35] Le Brun, M. 1977, "A Derivation of the Spectrum of FM with a Complex Modulating Wave", *Computer Music Journal*, 1(4): 51-52.
- [36] Le Brun, M. 1979, "Digital Waveshaping Synthesis". *Journal of the Audio Engineering Society*, 27(4): 250-266.
- [37] Moorer, J. A. 1976, "The Synthesis of Complex Audio Spectra by Means of Discrete Summation Formulas". *Journal of the Audio Engineering Society*, 24 (9): 717-727.
- [38] Moorer, J. A. 1977, "Signal Processing Aspects of Computer Music: A Survey", *Proceedings of the IEEE*, 65 (8): 1108 – 1141.
- [39] Palamin, J.P., P. Palamin and A. Ronveaux 1988, "A Method of Generating and Controlling Musical Asymmetric Spectra". *Journal of the Audio Engineering Society* 36 (9): 671-685.
- [40] Puckette, M. 1995, "Formant-Based Audio Synthesis Using Nonlinear Distortion." *Journal of the Audio Engineering Society* 43(1): 40-47.
- [41] Rodet, X. 1984. "Time Domain Formant-Wave-Function Synthesis". *Computer Music Journal*, 8 (3): 9-14.
- [42] Schottstaedt, W. 1977, "The Simulation of Natural Instrument Tones Using a Complex Modulating Wave". *Computer Music Journal* 1(4): 46-50.
- [43] Windham, G. and K. Steiglitz 1970, "Input Generators for Digital Sound Synthesis". *Journal of the Acoustic Society of America*, 47(2): 665-677.
- [44] http://www.acoustics.salford.ac.uk/acoustics_info/sound_synthesis/ 2013.06.10
- [45] <http://www.electricdruid.net/index.php?page=info.pdsynthesis> 2013.06.10
- [46] https://ccrma.stanford.edu/~jos/sasp/Additive_Synthesis_Early_Sinusoidal.html 2013.06.10
- [47] <http://www.planetoftunes.com/synth/subtractive.htm> 2013.06.10
- [48] <http://www.soundonsound.com/sos/apr00/articles/synthsecrets.htm> 2013.06.10
- [49] <http://www.clavia.se/nordmodular/Modularzone/FMsynthesis.html> 2013.06.10
- [50] http://eprints.nuim.ie/2335/1/VL_Adaptive_Phase_paper_12.pdf 2013.06.10
- [51] <http://www.csounds.com/scanned/> 2013.06.10

- [52] <http://cSounds.com/scanned/toot/> 2013.07.06
- [53] <http://www.sweetwater.com/expert-center/glossary/t--Sample-basedSynthesis>
2013.07.06
- [54] <http://documentation.apple.com/en/logicstudio/instruments/index.html#chapter=A%26section=4%26tasks=true> 2013.07.06
- [55] http://web.eecs.umich.edu/~gessl/georg_papers/CMJ04-theory.pdf 2013.07.06
- [56] http://music.columbia.edu/cmc/MusicAndComputers/chapter4/04_04.php 2013.07.06
- [57] <http://www.acoustics.hut.fi/research/asp/> 2013.07.06
- [58] <http://www.dei.unipd.it/english/csc/papers/dproc/dproc.html> 2013.07.06
- [59] <http://www.soundonsound.com/sos/dec05/articles/granularworkshop.htm> 2013.07.06

11. Ábrajegyzék

1. ábra Tiszta hang (szinusz hullám)
2. ábra Összetett hullámforma (tisza hangok összege)
3. ábra Harmónikusok a frekvencia tartományban
4. ábra Gerjesztés, hullámvezető, rezonátor
5. ábra Folyamatábra a hullámvezetőtől a hangszer kimenetéig
6. ábra Mintavételezés
7. ábra Kvantálás
8. ábra Lépcsők
9. ábra Folyamatábra a kimeneti hullámformáig
10. ábra CPU és RAM mérőprogram
11. ábra Fázis torzulási függvény és szinusz hullám működése
12. ábra Fázistorzítási függvény
13. ábra FM Fázis torzítás
14. ábra Tangens hiperbolikus függvény
15. ábra Karplus-Strong algoritmus blokkvázlata
16. ábra Hullámformák megvalósítása
17. ábra 440Hz-es szinusz hullám fade-out-tal
18. ábra 440Hz-es szinusz hullám fade-out-tal kinagyítva
19. ábra Crossfade megvalósítása
20. ábra 440Hz-es szinusz hullám és 392Hz-es fűréshullám crossfade szakasza
21. ábra 440Hz-es szinusz hullám és 392Hz-es fűréshullám összefűzve
22. ábra Szinusz, fűrés, négyszög jelek crossfade-el összefűzve
23. ábra Hullámtábla szintézis CPU és RAM felhasználása
24. ábra Azonos méretű granulátumok
25. ábra Különböző késleltetésű granulátum folyamatok
26. ábra Asszinkron granulár szintézis
27. ábra Granulátumok átfedése
28. ábra Szemcséző függvény
29. ábra 'D' Gitárhang
30. ábra Granulár szintézis megvalósítása
31. ábra 'D' Gitárhang modulált szemcséiből szintetizált hang

- 32. ábra Granulár színtézis CPU és RAM felhasználása
- 33. ábra Additív színtézis
- 34. ábra Színtézis függvény
- 35. ábra 440Hz-es szinusz hullám
- 36. ábra 500Hz-es szinusz hullám
- 37. ábra 620Hz-es szinusz hullám
- 38. ábra Hullámok összeadása
- 39. ábra Összeadott szinusz hullámok
- 40. ábra Additív színtézis CPU és RAM felhasználása
- 41. ábra 440Hz-es fűrész hullám
- 42. ábra Alul-áteresztő szűrő megvalósítása
- 43. ábra Szűrt 440Hz-es fűrész hullám
- 44. ábra Negyedfokú szűrővel a 440Hz-es fűrész hullám
- 45. ábra Mindhárom fűrész hullám
- 46. ábra Szubtraktív színtézis CPU és RAM felhasználása
- 47. ábra Formánsok
- 48. ábra Digitális hullámvezető egyszerű blokkvázlata
- 49. ábra 2D hullámvezető háló blokkvázlata
- 50. ábra Sávhatárolt hullámvezető blokkvázlata
- 51. ábra Színtézisek értékelése

12. Mellékletek

1. A számítógép paraméterei amin a szimulációkat, illetve méréseket végeztem:

Alaplap: ASUS H87M-PRO

Processzor: INTEL Core i5-4570 3.20GHz 1150 BOX

Memória: KINGSTON 8GB HyperX XMP DDR3 1600MHz CL9 KIT
KHX1600C9D3K2/8GX

SSD: KINGSTON 120GB SATA3 2,5" SSDNow V300 SV300S37A/120G