

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 637

**PROCEDURALNO GENERIRANJE ZVUČNIH
EFEKATA U RAČUNALNIM IGRAMA**

Ivor Rendulić

Zagreb, siječanj 2014.

Sadržaj

Uvod.....	5
1. Povijest zvuka u igrama	6
2. Proceduralna sinteza zvučnih efekata	8
2.1. Prednosti i mane korištenja snimljenih efekata	8
2.2. Ideja proceduralne sinteze i usporedba s korištenjem snimljenih zvukova.....	9
2.2.1. Zahtjevi na funkcionalnost	9
2.2.2. Prednosti i mane proceduralne sinteze	11
3. Koraci u stvaranju sustava za sintezu zvuka	12
3.1. Analiza zvuka i procesa	12
3.2. Stvaranje i pojednostavljivanje modela	12
3.3. Sinteza i stvaranje valnog oblika	13
4. Pregled nekih metoda sinteze zvuka	14
4.1. Metode u frekvencijskoj domeni	14
4.1.1. Modalna analiza i sinteza	14
4.1.2. Modeliranje spektra.....	23
4.2. Metode u vremenskoj domeni	26
4.3. Fizikalno modeliranje	26
5. Sinteza nekih zvučnih efekata	27
5.1. Zvonjenje staklene čaše	27
5.1.1. Analiza zvuka i izvlačenje parametara	29

5.1.2.	Sinteza	38
5.2.	Udarac mača o metalni štit ili oklop.....	40
5.2.1.	Analiza zvuka i izvlačenje parametara	40
5.2.2.	Konačan rezultat i subjektivni dojmovi	46
5.3.	Naleti vjetra	47
5.3.1.	Analiza zvuka.....	47
5.3.2.	Rezonantni filtri drugog reda za modeliranje spektra	48
6.	Zaključak.....	53
7.	Literatura	55
	Sažetak	56

Uvod

U samim začecima industrije računalnih igara, u doba Tetrisa, Ponga, Pacmana, Super Maria, zvukovi i glazba u računalnim igrama sintetizirani su na računalima. Radilo se o vrlo jednostavnim zvučnim efektima koji su zvučali izrazito umjetno i bili neusporedivi sa stvarnim zvukovima, ali niti grafički doživljaj nije bio realističan i nitko zapravo nije očekivao od igara da vjerno modeliraju stvarni svijet.

Porastom računalnih mogućnosti kućnih računala grafika u igrama iznimno je brzo napredovala i danas možemo reći da prilično realistično prikazuje svijet. Međutim, sinteza zvučnih efekata nije se razvijala jednakom brzinom. Razvijani su fizikalni modeli sinteze zvukova, ali su za realističan rezultat zahtijevali dugo izvođenje i nisu mogli biti sintetizirani u stvarnom vremenu. Ostale metode sinteze također nisu bile zadovoljavajuće te se industrija okrenula unaprijed snimljenim efektima koji bi zatim bili reproducirani tijekom igre.

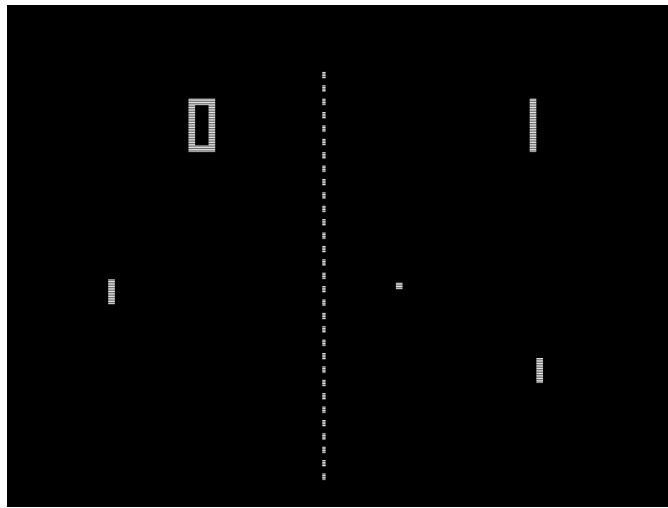
S obzirom na daljnji razvoj kućnih računala i konzola, koji su danas u stanju u stvarnom vremenu izvoditi iznimno velik broj operacija, sinteza zvučnih efekata je u zadnjih nekoliko godina ponovno postala zanimljivo područje.

U ovom radu će na primjeru tri konkretna zvučna efekta biti razvijene neke metode prikladne za korištenje i sintezu u računalnim igrama. Proučavani zvukovi bit će analizirani, i za svaki od njih bit će korištena i implementirana prikladna metoda, te će dobiven zvuk biti subjektivno uspoređen s originalnom. Također će biti diskutirana upotrebljivost svake od metoda u smislu računske složenosti, potrebne memorije za pohranu parametara i mogućnosti variranja sintetiziranih zvukova.

U poglavlju 1 bit će napravljen kratak pregled povijesti zvuka u računalnim igrama. Poglavlje 2 objasniti će ukratko ideju proceduralne, umjetne sinteze zvukova, te dati neke zahtjeve na postupke analize i sinteze. U poglavlju 3 će se malo dublje ući u korake izrade sustava za sintezu zvukova. Poglavlje 4 dat će pregled nekih metoda koje se mogu koristiti za sintezu. Konačno, u poglavlju 5 bit će analizirana tri različita zvuka, razvijene metode za sintezu i sintetizirane umjetne verzije promatranih zvukova.

1. Povijest zvuka u igrama

Teško je reći koju igru možemo smatrati prvom ikada izdanom[1], ali prva koja je uključivala zvučne efekte je *Pong*[2]. Radi se o arkadnoj igri nalik stolnom tenisu u kojoj igrač kontrolira platformu i mora vratiti lopticu protivniku. Udarac loptice u platformu bio je popraćen jednostavnim zvučnim efektom.



Slika 1 – originalna verzija igre Pong

Neke od značajnijih ranijih igara sa stajališta zvuka su *Pacman*, poznat po karakterističnom "waka-waka" zvuku, *Space Invaders*, gdje prvi puta dolazi do prilagodbe zvuka trenutnoj situaciji u kojoj se igrač nalazi, te *Super Mario Brothers*, u kojem je svaka razina igre imala karakterističnu zvučnu podlogu.

U tom ranom razdoblju računalnih igara, sedamdesetih i početkom osamdesetih godina 20. stoljeća, za sintezu zvukova bili su zaduženi zvučni čipovi koji su generirali valne oblike ili kratke odsječke šuma (eng. *Noise bursts*) koje bi zatim oblikovali ADSR ovojnica (eng. *Attack-Decay-Sustain-Release*) kako bi proizveli konačan zvučni efekt [3]. Program za te čipove bio je sličan MIDI formatu, u obliku liste s informacijama o vremenu i parametrima zvuka.

Tijekom osamdesetih godina generiranje zvučnih efekata seli se na glavnu procesorsku jedinicu, dedikirani zvučni čipovi nestaju, a pojavljuju se zvučne kartice. Umjesto generatora valnih oblika, generiraju se uzorci zvučnog signala koji se potom digitalno-analognom pretvorbom prevode u konačni valni oblik. Radili su sa vrlo niskim frekvencijama očitavanja i brojem bitova za kvantizaciju uzorkovanja, tipično 11 kHz i 8 bitova po uzorku. Iako su studijski alati dozvoljavali i više te davali mnogo veću kvalitetu zvuka, osobna računala i zvuk u igrama kaskali su nekoliko godina za njima, prvenstveno zbog ograničenja procesorske moći i ograničenom procesorskom vremenu dodijeljenom zvučnom dijelu igara (grafika je koristila najveći dio).

Krajem devedesetih godina dolazi do nove velike promjene. Umjesto generiranja zvukova, počinju se koristiti unaprijed snimljeni zvukovi koji se tijekom igre po potrebi reproduciraju, eventualno filtrirani kako bi vjerno prikazali udaljenost i položaj izvora zvuka. Na taj način dobiva se visoka kvaliteta zvuka uz malu potrošnju procesorskog vremena. Ovakav se oblik zvučnih efekata u igrama zadržao sve do današnjeg dana. Glavni proces u stvaranju takvih efekata jest snimanje što vjernijih i kvalitetnijih zvukova, te snimanje čim većeg broja različitih zvukova. Na taj se način može zvučnim efektom popratiti čim više mogućih radnji, interakcija ili zvukova iz okoliša u igri te, ako se radi o realističnim zvukovima, doprinijeti realizmu same igre.

U zadnjih se nekoliko godina ponovno razmatra računalna sinteza zvukova. U idućem će se poglavlju diskutirati o prednostima i manama unaprijed snimljenih efekata, uvesti pojam proceduralne sinteze zvukova te napraviti subjektivnu usporedbu između te dvije metode.

2. Proceduralna sinteza zvučnih efekata

U ovom će se poglavlju diskutirati o prednostima korištenja snimljenih efekata u igrama te o manama te metode. Uvest će se pojam proceduralne sinteze zvukova te će se napraviti usporedba idejne razlike, prednosti i mana u odnosu na korištenje snimki.

2.1. Prednosti i mane korištenja snimljenih efekata

U prošlom su poglavlju već spomenute neke od prednosti korištenja unaprijed snimljenih zvučnih efekata. Tako dobiveni zvukovi su, prije svega, vrlo realistični. Kvaliteta zvuka najviše ovisi o samom procesu snimanja, odnosno koliko blisko zvuk koji snimamo odgovara procesu koji se događa u igri. Primjerice, neka se radi o zvuku pucanja iz različitih pušaka, nečemu što je prisutno u gotovo svim igrama jednog od najpopularnijih žanrova igara. U slučaju da se radi o modernim puškama, moguće je otići u streljanu i snimiti stvarni zvuk točno određene puške koja se pojavljuje u igri. Ako se, pak, radi o igri u kojoj se radnja odvija u budućnosti sa izmišljenim oružjima, osobi odgovornoj za zvučne efekte daje se kreativna sloboda da raznim manipulacijama, efektima i filtriranjima iz nekog snimljenog efekta dođe do željenog rezultata.

Osim realističnosti, prednost su i mali zahtjevi na procesor koji je vrlo opterećen računalnom grafikom. Potrebno je samo učitati snimke koje je u tom trenu potrebno reproducirati, ovisno o formatu u kojem su snimke spremljene ih dekodirati i reproducirati. S obzirom na veliku količinu radne memorije u današnjim računalima, moguće je učitati prilično veliku količinu podataka istovremeno i tako reproducirati mnogo različitih zvukova istovremeno.

Iako današnja računala imaju velike količine radne memorije, kao i veliki kapacitet tvrdih diskova, ipak je potrebno racionalno ju koristiti. Ukoliko se želi imati vrlo velik broj zvučnih efekata visoke kvalitete, kako bi se moglo dočarati čim više različitih događaja u igri, potrebno je mnogo memorije za pohranu. Osim toga, potrebno je snimiti sve te zvukove, što samo po sebi predstavlja prilično velik posao. Zbog tih se razloga ograničava na zvuk onih događaja koji su najvažniji i koji se najčešće događaju. Tako će u modernim igrama gotovo uvijek biti prisutan,

primjerice, zvuk koračanja, udaraca mača o oklop, razbijanja stakla, ali će pomicanje virtualnog stolca ili nekog drugog nevažnog objekta rijetko biti popraćeno škripanjem i struganjem o pod.

Još jedna stvar koja se zamjera ovoj metodi jest jednoličnost, statičnost, ne-interaktivnost zvuka. Barata se snimljenim zvukom koji se potom koristi u mnogim situacijama u kojima bi se možda očekivali različiti zvukovi, te je filtriranje prije reprodukcije jedina opcija da bi se napravila neka manja promjena. Ova je mana zapravo usko vezana za prethodno naveden problem snimanja velike količine različitih zvukova. Kako bi se dočarali zvukovi svih mogućih interakcija u igri, teoretski je potrebna gotovo beskonačna količina prostora za pohranu jer je u današnjim igrama broj mogućih interakcija igrača i okoline često neizmjerljivo velik.

Konačno, zanimljivo je usporediti način na koji su se razvijali grafika i zvuk u igrama. Ideja je kod stvaranja grafike oduvijek u osnovi bila ista – stvarati na ekranu oblike koji će što vjernije predstavljati stvarne predmete. Da se u slučaju grafike krenulo drugim putem, kao što je slučaj sa reprodukcijom snimljenog zvuka, igre bi zapravo bile video snimke, gdje bi se igrač mogao kretati samo putovima kojima je prošao kameraman.

2.2. Ideja proceduralne sinteze i usporedba s korištenjem snimljenih zvukova

Imajući na umu sve navedeno u prethodnom potpoglavlju, mnogi ljudi danas zagovaraju povratak na sintetičko generiranje zvukova. U nastavku će biti razmatrane moguće prednosti toga, kako je sinteza napredovala u zadnjih nekoliko desetljeća i koliko je tako nešto ostvarivo, te će se ukratko diskutirati o manama i problemima.

Pod pojmom "proceduralna sinteza zvuka" podrazumijeva se metoda umjetnog generiranja zvuka pozivanjem neke funkcije ili procedure koja iz određenih parametara stvara valni oblik željenog zvučnog efekta. Da bi se bolje opisala ova ideja, bit će razmotrena sa stajališta korisnika (programera računalne igre) te sa stajališta developera, tj. osobe koja mora razviti te procedure.

2.2.1. Zahtjevi na funkcionalnost

U ovom će se potpoglavlju diskutirati koje bi sve funkcionalnosti trebale imati procedure ili funkcije za sintezu zvučnih efekata. To će biti učinjeno iz perspektive korisnika – programera

koji razvija računalnu igru. Kao primjer će biti korištena dva konkretna zvučna efekta koji se često koriste u igrama.

Prvi efekt koji će biti razmatran jest zvuk udarca mača ili sjekire o štit ili oklop. Radi se o vrlo čestom događaju u mnogim igrama smještenima u ranija povijesna razdoblja, ili u izmišljene fantastične svijetove (eng. *fantasy*).



Slika 2 – scena iz računalne igre Morrowind netom prije udarca sjekire o štit

U najosnovnijoj verziji sinteze zvuka, u trenutku udarca poziva se procedura koja generira pripadni zvučni efekt. Želimo li naglasiti jačinu udarca, moglo bi se dodati jedan parametar koji se šalje proceduri i utječe na intenzitet zvuka udarca. Također, potencijalno bi se mogla dodati i točna pozicija na kojoj je štit udaren, ili razni drugi parametri koji su odmah poznati u igri i mogu se iskoristiti. U poglavlju 5.2 bit će provedena detaljna analiza ovog zvuka i razvijen model sinteze u frekvencijskoj domeni.

Potpuno drugačija klasa zvukova koja se javlja u igrama su pozadinski zvukovi. Često su to kiša ili vjetar. Kod takvih je zvukova prirodno inicijalizirati proceduru s parametrima koji opisuju trenutno stanje (npr. lagana kiša), nakon čega će se kontinuirano sintetizirati zvuk lagane kiše, sve dok ne dođe do novog poziva s drukčijim parametrima.

2.2.2. Prednosti i mane proceduralne sinteze

Nakon kratkog pregleda ideje proceduralne sinteze, sada će ukratko biti opisano koje su, teoretski, prednosti i mane. U daljnjim poglavljima, u kojima će biti riječ o samoj sintezi zvuka, još će se malo proširivati ovaj popis iskustvom iz prakse.

Neke od prednosti sinteze zvuka.

- Uvelike smanjeni zahtjevi za memorijskim prostorom – proceduralnom sintezom zvukova nije potrebno imati pohranjen ogroman broj snimki zvukova, već se spremaju samo određeni parametri, formule ili eventualno mnogo kraći zvučni segmenti.
- Interaktivnost – varijacijom parametara u postupku sinteze moguće je mijenjati zvukove kako ne bi svaki puta zvučali isto. Tu se varijaciju može povezati s nekim parametrima u igri kako bi sintetizirani zvuk bio u skladu s događanjima u igri čime se dobiva realističniji doživljaj zvuka.
- Recikliranje modela – različiti procesi često se mogu opisati sličnim modelima, bilo da se radilo o fizikalnim ili spektralnim modelima. Iz tog se razloga razvijeni modeli mogu, uz manje modifikacije, ponovno iskoristiti čime se preskače dugotrajan proces razvoja novih modela.

Mane:

- Komplicirana analiza i razvoj modela – često je iznimno teško razviti modele koji će dati zadovoljavajuće rezultate, dobru kvalitetu zvuka i učinkovitu sintezu. Ova je faza mnogo dugotrajnija i teža od snimanja efekata i često obeshrabruje već na samom početku kod razmatranja sintetiziranog zvuka.
- Lošija kvaliteta zvuka – ovisi o korištenom modelu, ali za velik broj zvukova (posebice kod kompliciranijih procesa) još se uvijek primijeti razlika u kvaliteti i realističnosti naspram snimki.

3. Koraci u stvaranju sustava za sintezu zvuka

Postupak izgradnje sustava za sintezu zvučnih efekata sastoji se od dvije vrlo različite faze s drugačijim zahtjevima – analiza zvuka koji želimo sintetizirati, te sinteza na temelju rezultata analize.

3.1. Analiza zvuka i procesa

Prvi korak jest analiza zvuka, ili klase zvukove, koje se želi sintetizirati. Osim samog zvuka, analizira se i proces kojim nastaje zvuk, što ćemo posebno opisati u poglavlju o sintezi temeljenoj na fizikalnim modelima.

Proces analize vjerojatno je i najvažniji za konačnu kvalitetu sintetiziranog zvuka. Ovdje se traže oni parametri kojime je moguće čim preciznije opisati zvuk, analizira proces i opisuje jednadžbama širenja zvuka. Pritom treba imati na umu ideju proceduralne sinteze i nastojati pronaći parametre kojima će biti moguće manipulirati u svrhu modificiranja zvuka (a ne odabrati kao parametre, primjerice, sve uzorke valnog oblika). Također treba imati na umu i iduće korake, stvaranje modela i sintezu, kako već u startu sustav ne bi bio pretjerano komplicirano zamišljen i kao takav neostvariv.

Faza analiza ujedno je i najdugotrajnija. Ako se u njoj koriste neka programska rješenja, kod njih vrijeme izvođenja nije kritično, odnosno nije bitno da rade u stvarnom vremenu.

3.2. Stvaranje i pojednostavljivanje modela

Kao međukorak i svojevrsni most između analize i sinteze, potrebno je prilagoditi podatke dobivene analizom kako bi bili pogodni u postupku sinteze koja se mora izvršavati u stvarnom vremenu. Primjerice, ako smo u postupku analize razvili model s parcijalnim diferencijalnim jednadžbama, potrebno je diskretizirati model kako bi se mogao koristiti na računalu.

3.3. Sinteza i stvaranje valnog oblika

Konačno, potrebno je implementirati postupak same sinteze zvuka i stvaranja valnog oblika koji će biti reproduciran. Kod ovog je koraka učinkovitost izvođenja iznimno bitna. Potrebno je da sinteza radi u stvarnom vremenu te da čim manje opterećuje procesor. Uz manju računsku zahtjevnost sinteze svakog od zvukova moguće je sintetizirati veći broj zvukova i time dobiti realističniju zvučnu sliku događaja u igri.

4. Pregled nekih metoda sinteze zvuka

U ovom će poglavlju biti napravljen pregled najpopularnijih metoda koje se koriste u sintezi zvuka. Ugrubo će biti podjeljene u tri kategorije – one temeljene na postupcima u frekvencijskoj domeni, one vezane uz vremensku domenu, te metode temeljene na fizikalnim modelima. Najveći naglasak bit će na metodama u frekvencijskoj domeni koje će kasnije biti i korištene u postupcima sinteze nekih zvukova.

Treba naglasiti da je ovakva podjela za neke od metoda dosta neprecizna. Naime, često se kombiniraju postupci vezani za sve tri kategorije, posebice ako se jedni koriste u analizi, a drugi u sintezi. Većina je metoda stoga svrstana s obzirom na postupke analize (npr., fizikalno modeliranje služi kod analize da bi se dobile jednadžbe širenja zvuka, ali se sinteza provodi u vremenskoj domeni).

4.1. Metode u frekvencijskoj domeni

Glavni fokus ovog rada bit će korištenje metoda u frekvencijskoj domeni kako se modelirao spektar zvuka. U ovom će poglavlju biti objašnjene neke od tih metoda, koje će u kasnijem poglavljima i biti implementirane i detaljnije proučene na realnim primjerima.

4.1.1. Modalna analiza i sinteza

Prva metoda koju ćemo opisati, a kasnije i koristiti, jest temeljena na modovima. Pojam modova i modalne analize i sinteze vezan je uz frekvencijsku domenu, ali isto tako i uz fizikalnu analizu i modeliranje, dok se sinteza uobičajeno odvija u vremenskoj domeni. U ovom će se radu modalna analiza i sinteza koristiti u pojednostavljenom obliku i algoritam koji će biti implementiran temeljen je na Fourierovoj transformaciji na vremenskom otvoru, te je stoga modalna analiza obrađena u ovom potpoglavlju. Ipak, bit će opisana i veza modova s fizikalnim modelom, kao i proces sinteze korištenjem filtara u vremenskoj domeni.

Modalna analiza bavi se proučavanjem dinamičkih svojstava objekata kod pobuda koje uzrokuju vibraciju. Pojam "mod" opisuje svaku od karakterističnih frekvencija prisutnih u titranju

određenog krutog objekta, s time da svaka čestica titra tom frekvencijom i da postoji stalna razlika između faza različitih čestica. Primjerice, zvono udareno čekićem odzvanja točno određenim frekvencijama koje se prigušuju s vremenom. Svaka od njih opisana je jednim modom koji modelira eksponencijalno prigušenu sinusoidu određene amplitude. To su tri parametra koji karakteriziraju modove – frekvencija moda f_i , amplituda A_i te faktor prigušenja (eng. *damping factor*) α_i . Valni oblik zvuka dobiva se kao suma svih prisutnih modova:

$$s(t) = \sum_i m_i(t) = \sum_i A_i e^{-\alpha_i t} \sin(2\pi f_i t + \theta_i).$$

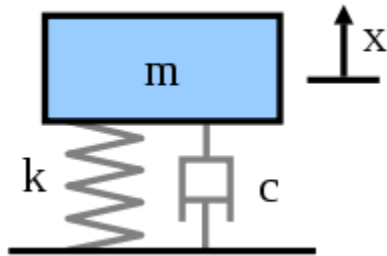
Zašto se koriste upravo eksponencijalno prigušene sinusoide postat će jasnije u idućem potpoglavlju o fizikalnoj pozadini modova.

Karakteristične frekvencije modova te faktori prigušenja ovisni su o karakteristikama materijala predmeta (npr. elastičnost, gustoća). Kad bi se zvuk mogao potpuno opisati s ova dva parametra, odnosno kad bi samo ovisio o obliku i vrsti materijala, proces određivanja parametara bio bi mnogo jednostavniji. Međutim, zvuk zvona zvuči drugačije ovisno o poziciji udarca čekića. Amplitude modova, koje zapravo označavaju s kolikim će faktorom određeni mod biti prisutan u signalu, ovise i o mjestu pobude te osiguravaju te razlike u zvuku.

4.1.1.1 Fizikalna pozadina vibracija i modova

U ovom će poglavlju biti objašnjena fizikalna pozadina vibracija, modalne sinteze i zašto se koriste upravo eksponencijalno prigušene sinusoide.

Udarcem u neki kruti objekt njegove čestice počinju titrati, međusobne utječući jedna na drugu. Ukupno gibanje čestica objekta dovodi do nastanka zvuka. Kao osnovni fizikalni model bit će analiziran sustav s masom, oprugom i prigušnicom (eng. *Mass-Spring-Damper*) prikazan na Slika 3. Radi se o masi m koja je povezana s čvrstom plohom oprugom koeficijentom elastičnosti k i prigušnicom konstante gušenja ζ .



Slika 3 – sustav s masom, oprugom i prigušnicom

Elastična sila opruge proporcionalna je pomaku (sabijanju ili rastezanju opruge) s koeficijentom elastičnosti:

$$F_{el} = -kx.$$

Sila prigušenja proporcionalna je brzini:

$$F_d = -\zeta \frac{dx}{dt}.$$

Negativni predznaci u oba slučaja ukazuju da sile djeluju u smjeru suprotnom od vektora pomaka, odnosno vektora brzine.

Izjednačavanjem sila po drugom Newtonovom zakonu gibanja dobiva se:

$$-\left(mg + \zeta \frac{dx}{dt} + kx\right) = m \frac{d^2x}{dt^2}.$$

Uz zanemarivanje težine tijela u usporedbi s ostalim članovima:

$$m \frac{d^2x}{dt^2} + \zeta \frac{dx}{dt} + kx = 0$$

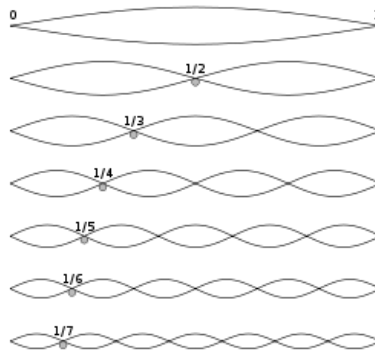
rješenje se nalazi kao:

$$x(t) = Ae^{-\frac{\zeta t}{2m}} \cos\left(t \sqrt{\frac{k}{m} - \left(\frac{\zeta}{2m}\right)^2}\right).$$

Vidljivo je da se radi upravo o eksponencijalno prigušenoj sinusoidi koja je ranije spomenuta kao osnovni oblik modalne sinteze. Gušenje (eksponent) ovisi o masi i koeficijentu prigušenja, a frekvencija titranja o masi, koeficijentu elastičnosti i koeficijentu prigušenja.

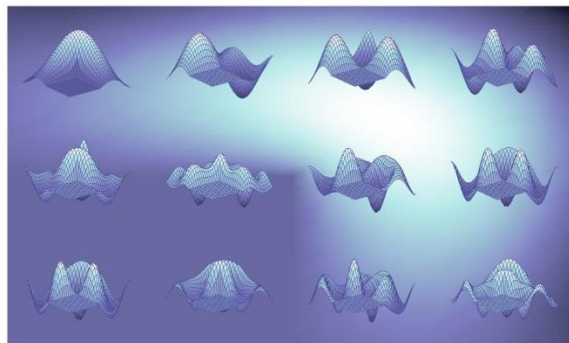
4.1.1.2 Modovi žice i plohe

Slika 4 prikazuje prvih sedam harmonika koji se javljaju kod titranja žice. Harmonici su upravo posljedica modova prisutnih kod titranja. Primjerice, pobudom (trzajem) žice točno na sredini, žica će titrati neparnim modovima i javit će se neparni harmonici.



Slika 4 – prvih sedam harmonika titranja žice

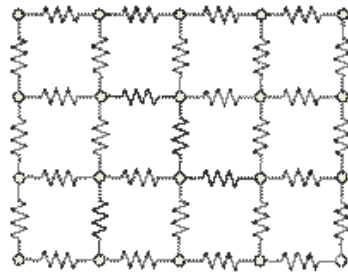
Slika 5 prikazuje modove vibracije plohe. U slučaju žice titranje je nešto lakše za predočiti, ali vibracija plohe je veoma slična. Daljnjim poopćavanjem dolazi se i do vibracija trodimenzionalnog tijela.



Slika 5 – modovi vibracije plohe

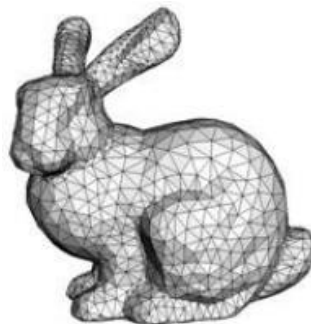
4.1.1.3 Metoda konačnih elemenata

Naravno, svi sustavi koji su od interesa za sintezu zvuka su značajno kompliciraniji od ovakvog jednostavnog modela mase, opruge i prigušnice. Korištenjem metode konačnih elemenata (postupka iz numerike za rješavanje problema opisanih diferencijalnim jednačinama) promatrani predmeti mogu se raščlaniti na konačan broj diskretnih čestica koje su u međusobnoj interakciji. Slika 6 prikazuje razmještaj čestica u dvodimenzionalnoj plohi povezanih oprugama.



Slika 6 – skup čestica u plohi povezanih oprugama

Ovakav se postupak poopćava na i na trodimenzionalne predmete koji se mogu podijeliti u konačan broj tetraedara ili kocaka različitih dimenzija. Slika 7 prikazuje model zeca podijeljenog u konačno mnogo tetraedara. Takvom analizom predmeta mogu se dobiti podatci o ukupnim vibracijama i modovima u ovisnosti o mjestu udarca.



Slika 7 – model zeca podijeljen u konačno mnogo tetraedara

Metoda konačnih elemenata mogla bi biti iznimno primjenjiva u području sinteze zvukova u igrama budući da su svi predmeti u igrama već renderirani i mogu se koristiti za modalnu analizu. Metoda konačnih elemenata je zbog složenosti izvan dosega ovog rada, ali po mišljenju autora predstavlja jedan od najboljih smjerova u sintezi zvukova. Primjeri korištenja metode konačnih elemenata i modalne sinteze mogu se naći u [8] [9].

4.1.1.4 Postupci sinteze

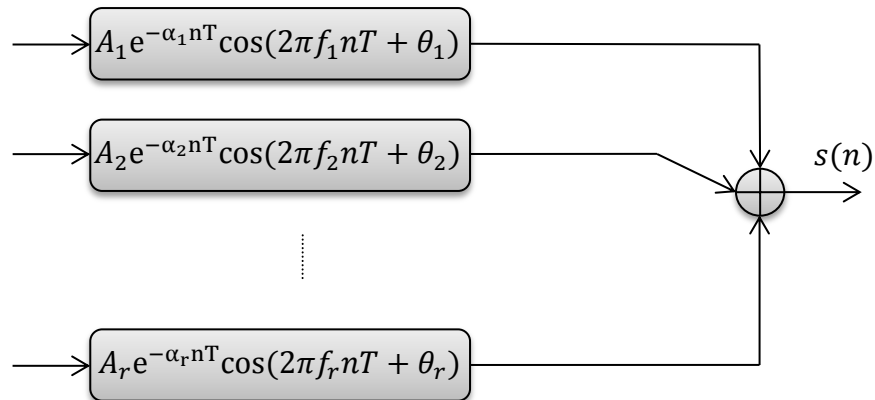
U nastavku ćemo razmotriti samu sintezu, odnosno generiranje valnog oblika, iz poznatih modova.

4.1.1.4.1 Generiranje valnog oblika u vremenskoj domeni

Najjednostavniji način sinteze, uz poznate modove i njihove frekvencije, faktore prigušenja i amplitude, jest jednostavno generirati doprinos svakog od njih te rezultat zbrojiti. S obzirom da vrijednosti asimptotski teže nuli, ali ju, u teoriji, nikad ne dostižu, potrebno je samo odabrati neku minimalnu vrijednost ispod koje će se doprinos svakog od modova zanemariti.

Neka je r broj modova, te n (prosječan) broj uzoraka u kojem računamo vrijednosti za svaki mod. Ovakva izravna sinteza ima složenost $O(nr)$ te zahtjeva toliki broj računanja vrijednosti sinusa (koje se mogu čitati iz pohranjene tablice) i eksponencijalne funkcije.

Slika 8 prikazuje broj dijagram ovakve sinteze. S obzirom na velik broj računanja eksponencijalnih i trigonometrijskih funkcija, kao i ostalih operacija, vrijedi istražiti može li se postupak sinteze nekako ubrzati kako bi bio učinkovitiji i manje opterećivao procesor.



Slika 8 – blok dijagram modalne sinteze

4.1.1.4.2 Paralelni rezonantni filtri

Sumu r ekponencijalno prigušenih sinusoida moguće je dobiti pobuđivanjem r paralelnih filtara drugog reda koji imaju samo polove (eng. *all-pole*). Naime, odgovarajućim smještajem polova u kompleksnoj ravnini moguće je ostvariti rezonanciju na željenoj frekvenciji i kontrolirati prigušenje.

Opći oblik prijenosne funkcije $H(z)$ *all-pole* filtra drugog reda jest:

$$H(z) = \frac{b_0}{1 + a_1 z^{-1} + a_2 z^{-2}}.$$

Jednadžba diferencija koja opisuje odnos izlaza i ulaza glasi:

$$y[n] = b_0 x[n] - a_1 y[n - 1] - a_2 y[n - 2].$$

Ukoliko se želi ostvariti rezonancija na frekvenciji $\omega_c = 2\pi f_c$ potrebno je postaviti pol pod kutem $\theta_c = \omega_c T_s = \frac{\omega_c}{f_s}$. Kako bi se osiguralo da je filter realan, odnosno da za realan ulaz daje realan izlaz, potrebno je postaviti i drugi pol kompleksno konjugiran prvome.

Polovi se stoga biraju kao:

$$p_1 = R e^{j\theta_c}, p_2 = R e^{-j\theta_c}.$$

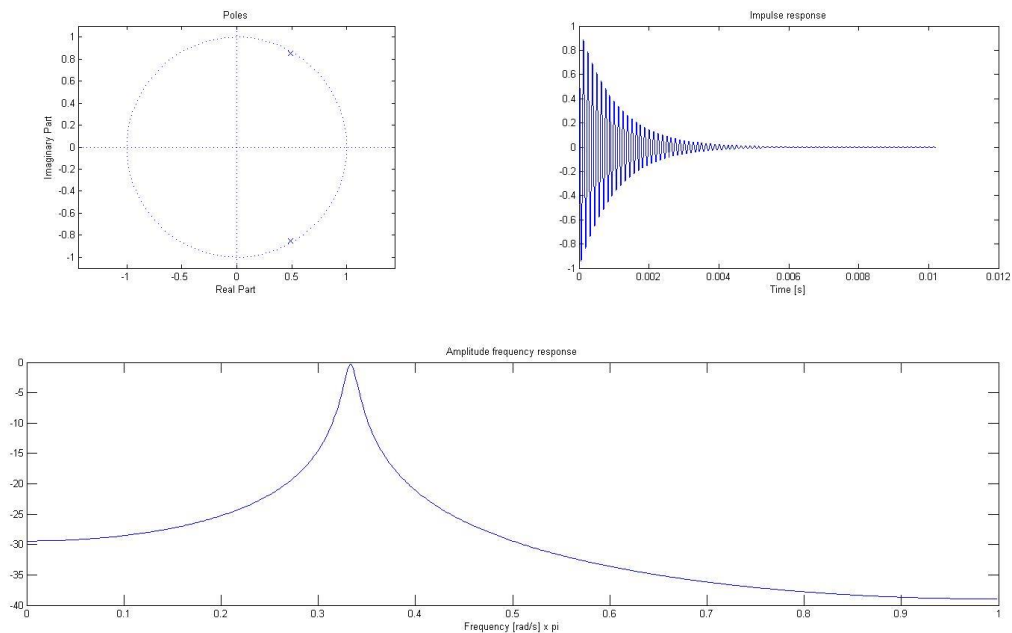
Kako bi se osigurala stabilnost filtra nužno je da vrijedi $|R| < 1$. Uz kaskadu dvije sekcije prvog reda s polovima p_1 i $p_2 = \overline{p_1}$ dobiva se:

$$\begin{aligned} H(z) &= \frac{b_0}{(1 - p_1 z^{-1})(1 - p_2 z^{-1})} \\ &= \frac{b_0}{(1 - R e^{-j\theta_c} z^{-1})(1 - R e^{j\theta_c} z^{-1})} \\ &= \frac{b_0}{1 - 2R \cos(\theta_c) z^{-1} + R^2 z^{-2}} \end{aligned}$$

Usporedbom s prethodnim, općenitijim izrazom vidi se da vrijedi $a_1 = -2R \cos(\theta_c)$, $a_2 = R^2$ te jednačba diferencija glasi:

$$y[n] = b_0 x[n] + 2R \cos(\theta_c) y[n - 1] - R^2 y[n - 2].$$

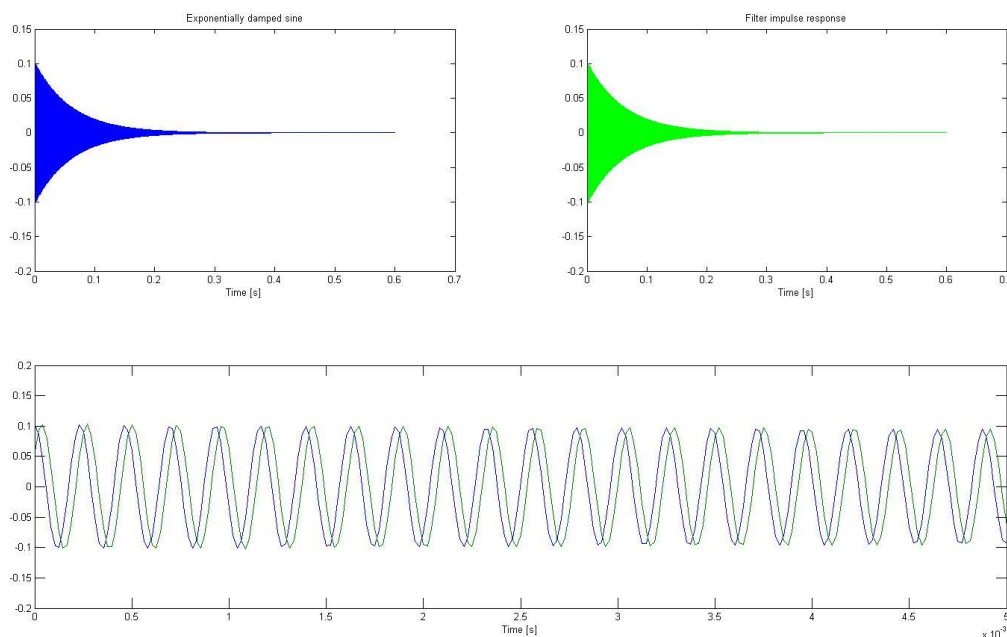
Slika 9 prikazuje položaj polova $p_{1,2} = 0.98e^{\pm \frac{j\pi}{3}}$ u kompleksnoj ravnini, te impulsni odziv i frekvencijsku karakteristiku rezonantnog filtra drugog reda.



Slika 9 – polovi, impulsni odziv i amplitudna frekvencijska karakteristika *all-pole* filtra drugog reda s polovima $p_{1,2} = 0.98e^{\pm \frac{j\pi}{3}}$

Odabirom $R = e^{-\frac{\alpha}{f_s}}$, te odabirom b_0 kako bi se skalirala amplituda, dobiva se impulsni odziv koji odgovara eksponencijalno prigušenoj sinusoidi frekvencije f_c , s faktorom gušenja α i željenom amplitudom A .

Slika 10 prikazuje eksponencijalno prigušenu sinusoidu oblika $s[n] = Ae^{-\frac{16.4967}{f_s}n} \cos\left(\frac{2\pi 4305}{f_s}n\right)$, te istu eksponencijalu dobivenu pobudom rezonantnog filtra s parametrima $b_0 = 0.0553$, $a_1 = -1.6901$ i $a_2 = 0.9993$. Ispod grafova valnih oblika prikazana je usporedba prvih 0.005 sekundi gdje se vidi kašnjenje odziva iz filtra.



Slika 10 – eksponencijalno prigušena sinusoida, impulsni odziv rezonantnog filtra s dva pola i usporedba u prvih 0.005s

Ovim postupkom pokazan je učinkovit način modalne sinteze korištenjem rezonantnih filtara s karakteristikama modova.

4.1.2. Modeliranje spektra

Metoda koja će u ovom radu biti najviše korištena (u kombinaciji s modalnom sintezom) bazira se na modeliranju spektra signala, odnosno tri specifične klase u spektru – harmonijski dio, neharmonijski ili zašumljeni dio spektra te tranzijenti ili prijelazne pojave u signalu.

Ove su metode korištene kod jačih kompresija audio ili govornog signala, kao i kod raznih efekata i manipulacija signalom kao što su promjena pitcha (eng. *pitch shifting*). Mi ćemo ih koristiti kako bi iz određenog signala (ili klase signala) izvukli neke parametre iz kojih je moguće sintetizirati zvuk, te koje je moguće varirati kako bi se dobile varijacije u zvuku kod svakog postupka sinteze.

4.1.2.1 Sinusoidalno modeliranje

Ideju modeliranja harmonijskog dijela spektra korištenjem diskretne Fourierove transformacije na vremenskom otvoru razvili su MacAulay i Quartieri sredinom osamdesetih godina prošlog stoljeća.

Neka je $x(n)$ promatrani audio ili govorni signal. Njegova diskretna Fourierova transformacija na vremenskom otvoru definirana je kao

$$X(m, k) = \sum_{n=-\infty}^{\infty} x(n)w(n - m)e^{\frac{2\pi jkn}{N}},$$

gdje je $w(n)$ vremenski otvor, N broj točaka u kojima računamo Fourierovu transformaciju te m diskretni pomak vremenskog otvora. Tipičan odabir za vremenski otvor su pravokutni, Hannov, Hammingov i Gaussov.

Idući korak jest traženje "šiljaka" (eng. *peak*) u amplitudnom dijelu izračunatog spektra. Naime, šiljci u spektru na određenim pozicijama možda su znak da je u signalu prisutna sinusna komponenta pripadajuće frekvencije. Detekcija šiljaka potrebna je u mnogim postupcima kod obrade signala te je razvijeno mnogo algoritama (primjerice, i MPEG standard kodiranja audio signala koristi detekciju šiljaka u psihoakustičkom modelu), a njihovu uspješnost najlakše je ocijeniti subjektivnim promatranjem spektra i detektiranih šiljaka.

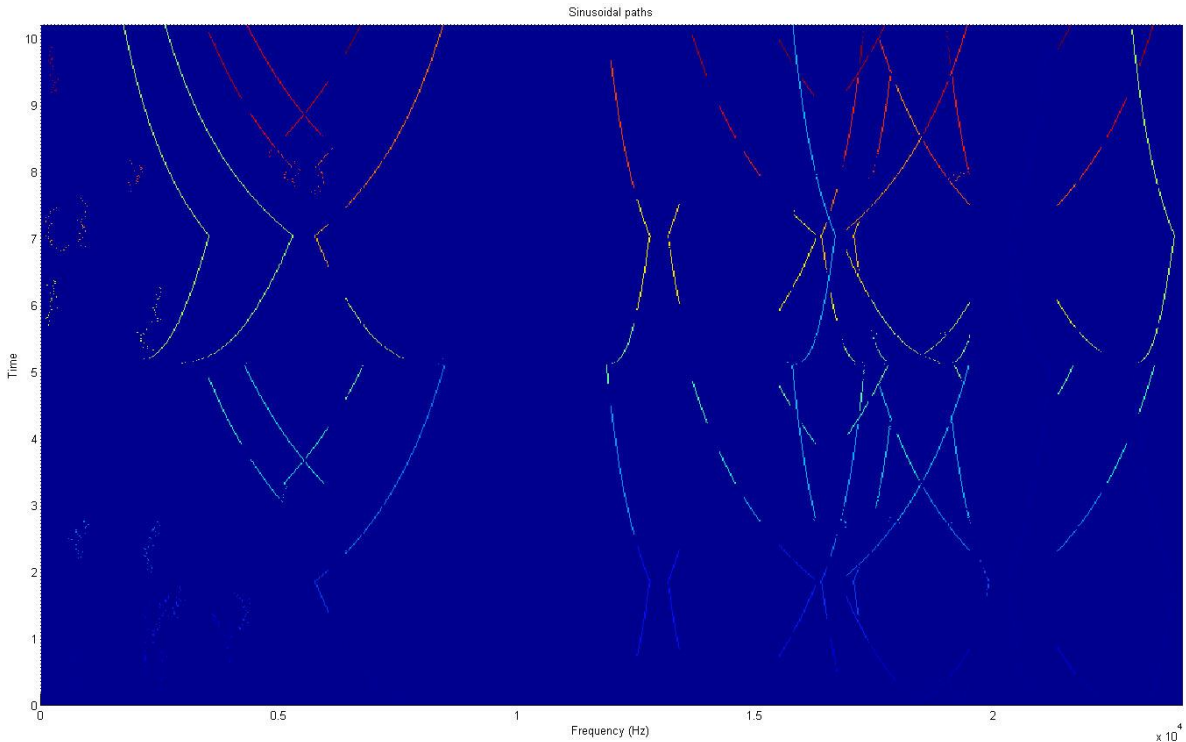
Nakon što su šiljci detektirani u amplitudnim spektrima za sve vrijednosti koraka m , metoda sinusoidalnog modeliranja koju predlažu MacAulay i Quartieri temelji se na pronalaženju šiljaka u spektrima susjednih vrijednosti koraka m koji odgovaraju istoj sinusoidi. Proces počinje označavanjem svakog šiljka u spektru u prvom koraku analize kao zasebnu sinusnu komponentu.

Neka je sa $p_i^{(k)}$ označen i -ti šiljak u k -tom koraku analize, a sa $\omega_j^{(k)}$ j -ta sinusna komponenta koja postoji u k -tom okviru. U prvom koraku svaki šiljak postaje nova sinusna komponenta:

$$p_i^{(1)} \rightarrow \omega_i^{(1)}, \forall i.$$

Zatim se promatraju šiljci u spektru koji odgovara idućoj vrijednosti koraka analize m . Ovdje se mogu dogoditi tri različite situacije. Ako je neki od šiljaka, koji je najbliže nekoj od zabilježenih sinusnih komponenti, dovoljno blizu, smatramo da je on posljedica istog sinusa čija se frekvencija eventualno malo promijenila u odnosu na prošli okvir analize. Ako, pak, nekom od šiljaka iz prethodnog okvira ne pronađemo kandidata koji bi mogao biti njegov nasljednik, onda ta sinusna komponenta "umire". I zadnje, ako neki šiljak u trenutnom okviru analize nema prethodnu sinusnu komponentu, to označavamo kao "rođenje" nove sinusne komponente. Ovaj se postupak zatim ponavlja za svaki slijedeći okvir.

Slika 11 prikazuje rezultat ovakvog algoritma u slučaju zvuka policijske sirene. Različitim bojama označene su različite sinusne komponente.



Slika 11 – detektirane sinusne komponente u zvuku policijske sirene

4.1.2.2 Modeliranje šuma

U slučaju spektralno bogatih zvukova, sinusoidalno modeliranje nije prikladno jer bi bio potreban iznimno velik broj sinusoida. Za modeliranje takvog, šumovitog dijela spektra mogu se koristiti razne tehnike. Jedna od njih, koja će i biti korištena u konkretnim slučajevima analize i sinteze u idućem poglavlju, jest modeliranje šumom iste spektralne razdiobe. Računa se energija šuma u signalu u diskretnom broju frekvencijskih pojaseva te se zatim sintetizira šum koji ima jednaku razdiobu energije.

4.1.2.3 Modeliranje tranzijenata

Fourierova transformacija na vremenskom otvoru često nije prikladna za detekciju i sintezu tranzijenata, kratkotrajnih i intenzivnih promjena u signalu. Primjerice, kod bilo kakve vrste udarca sam trenutak udarca traje iznimno kratko, ali izaziva naglu i intenzivnu promjenu u signalu.

Modeliranje tranzijenata korištenjem Fourierove transformacije na vremenskom otvoru zahtijevalo bi vrlo kratak vremenski otvor, praktički neupotrebljiv za analizu ostatka signala. Tranzijent se stoga modelira posebnim metodama – ekspanzijom valića, određivanjem amplitudne omotnice signala, korištenjem snimljenog valnog oblika tranzijenta ili promjenom duljine vremenskog otvora. U svakom slučaju potrebno je detektirati trenutke tranzijenata, što se u najjednostavnijoj varijanti može učiniti detekcijom velike promjene energije u signalu.

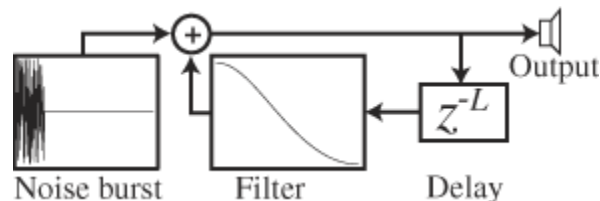
4.2. Metode u vremenskoj domeni

Metode sinteze zvuka koje su većinski (ili isključivo) vezane uz vremensku domenu obično se svode na obradu snimljenog zvuka. Jedna od zanimljivijih i popularnijih među njima je metoda granularne sinteze čije je temelje postavio Dennis Gabor. Ona se temelji na reprodukciji određenog broja vrlo kratkih snimki zvuka (granula, duljine 1-50 ms). Međutim, granule se reproduciraju s varijabilnim preklapanjima, fazama, promjenama frekvencije i amplitudne ovojnice.

4.3. Fizikalno modeliranje

Fizikalno modeliranje već je spomenuto u kombinaciji s modalnom sintezom. U „najekstremnijim“ primjenama razrađuje se fizikalni model promatranog procesa kojim nastaje zvuk, rezultat čega su obično parcijalne diferencijalne jednadžbe tlaka zraka. Primjer takve analize i sinteze u svrhu sintetiziranja zvuka kiše moguće je naći u [10].

Druga često korištena metoda koja se može svrstati u ovo poglavlje su tzv. *Digital Waveguide* modeli, modeli s linijama za kašnjenje kojima se opisuju putujući valovi kao D’Alembertova rješenja valne jednadžbe. Vjerojatno najpoznatiji takav model jest Karplus-Strong algoritam za sintezu zvuka titranja gitare nastao analizom titranja žice.



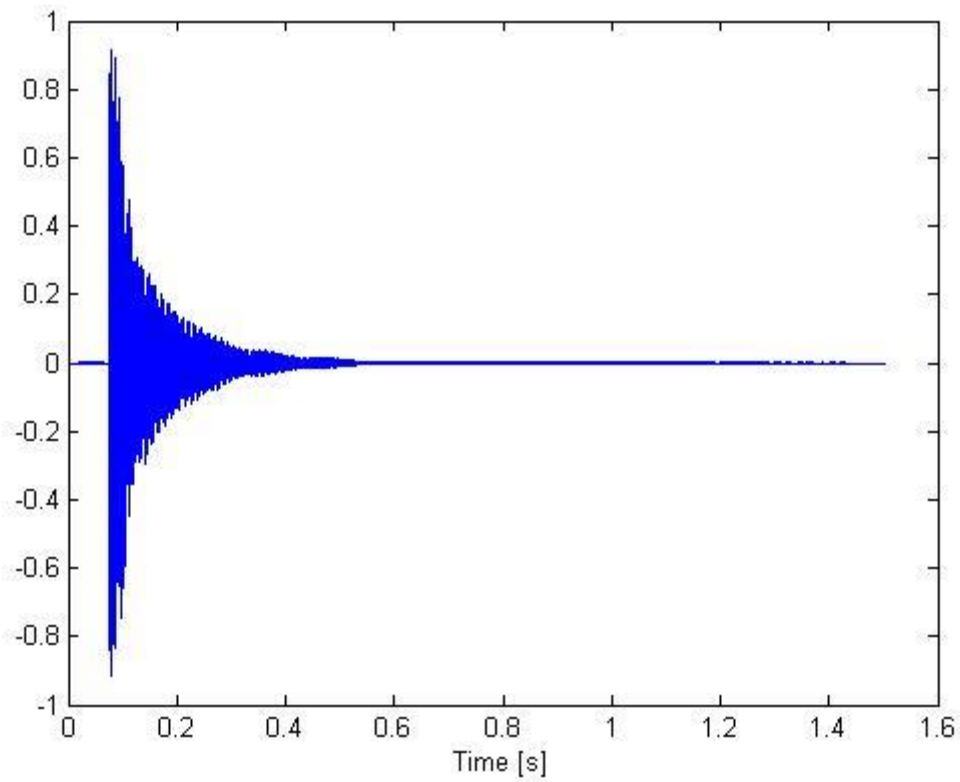
Slika 12 – *digital waveguide* model Karplus-Strong algoritma

5. Sinteza nekih zvučnih efekata

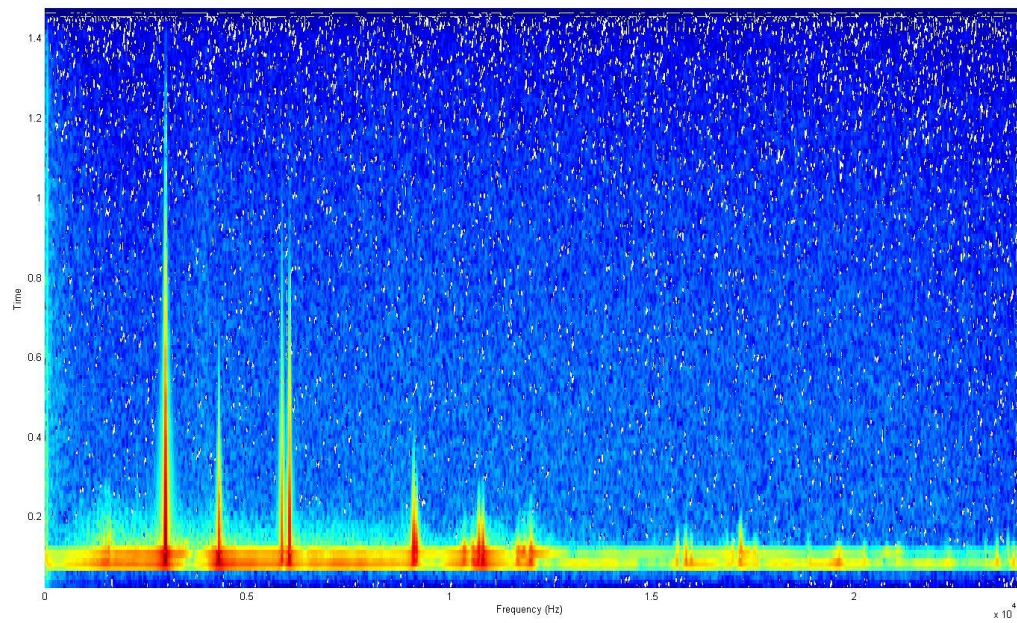
U ovom ćemo poglavlju analizirati nekoliko različitih zvučnih efekata, te na temelju spektralnih značajki i subjektivnog dojma odabrati metodu sinteze za svaki od njih. Fizikalni modeli procesa nastajanja tih zvučnih efekata i njihov razvoj neće biti ovdje razmatrani iz nekoliko razloga. Prvo, kompliciran postupak analize razvojem fizikalnih modela stvara modele specifične za relativno usku klasu zvukova. S obzirom da se često radi o iznimno kompliciranim analizama, a da developeri igara žele imati na raspolaganju iznimno velik broj različitih zvukova, postupak fizikalnog modeliranja zahtjeva iznimno puno posla. Umjesto toga, orijentirat ćemo se na implementaciju nekih postupaka spektralne analize, a subjektivnom analizom zvuka odabrat ćemo koju ćemo od implementiranih metoda koristiti.

5.1. Zvonjenje staklene čaše

Prvi primjer koji ćemo sintetizirati jest vrlo jednostavan zvuk dobiven udarcem noža o staklenu čašu (na način kao što se, primjerice, na proslavama daje do znanja da netko želi održati govor). Slika 13 prikazuje valni oblik snimke toga zvuka. Primjećujemo da su sinusne komponente iznimno izražene, da ne mijenjaju frekvenciju s vremenom već se samo prigušuju, te da je prisutna vrlo mala količina ne-harmonijskog dijela signala. Slika 14 prikazuje spektrogram promatrane snimke, i potvrđuje subjektivne zaključke donešene na temelju slušanja. Očito je da postoji nekoliko više ili manje izraženih harmonijskih komponenti koje opadaju s vremenom. Također se vidi da je u trenutku udara prisutno mnogo više frekvencija koje ubrzo iščeznu. Za generiranje spektrograma korišten je Hannov vremenski otvor duljine 2048 uzoraka, pomak otvora od 512 uzoraka, odnosno faktor preklapanja (eng. *overlap factor*) od 1536 uzoraka.



Slika 13 – valni oblik zvuka staklene čaše



Slika 14 – spektrogram zvuka staklene čaše

S obzirom na navedena razmatranja, razumno rješenje jest upotrijebiti modalnu analizu i sintezu kako bi se opisali prisutni harmonici. Iz kretanja amplitude svakog od harmonika bit će vidljivo je li opravdano koristiti eksponencijalno prigušenje.

Nakon modalne sinteze, slušanjem sintetiziranog zvuka čut ćemo je li potrebno dodatno modelirati tranzijent, odnosno trenutak udarca.

5.1.1. Analiza zvuka i izvlačenje parametara

U nastavku ćemo opisati implementaciju modalne analize, te u sklopu nje i detekcije šiljaka koju ćemo koristiti i u kasnijim postupcima.

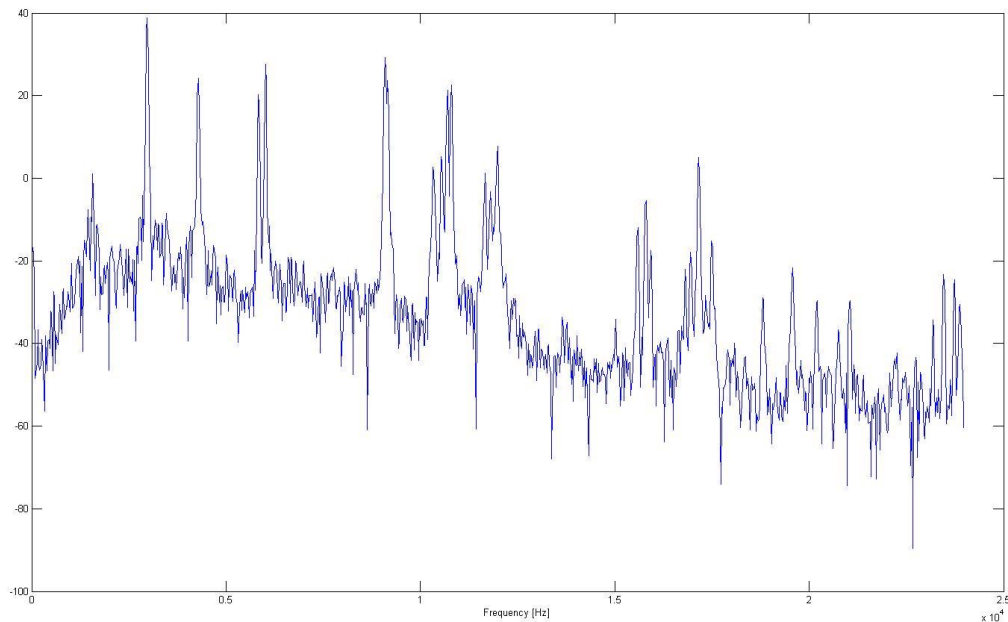
5.1.1.1 Traženje modova

Cilj je razviti algoritam za automatsku detekciju modova, što uključuje i računanje svih njihovih parametara (frekvencija, amplituda i faktor prigušenja).

5.1.1.1.1 Detekcija šiljaka

Postupak traženja modova započinje traženjem šiljaka u amplitudnom spektru signala za svaki vrijednost pomaka vremenskog otvora. Postoji mnogo različitih algoritama za to, a ovdje ćemo razviti i prezentirati jedan relativno jednostavan koji daje zadovoljavajuće rezultate.

Slika 15 prikazuje amplitudni dio spektra jednog okvira promatranog zvuka. Na ovom okviru bit će demonstriran proces detekcije šiljaka.



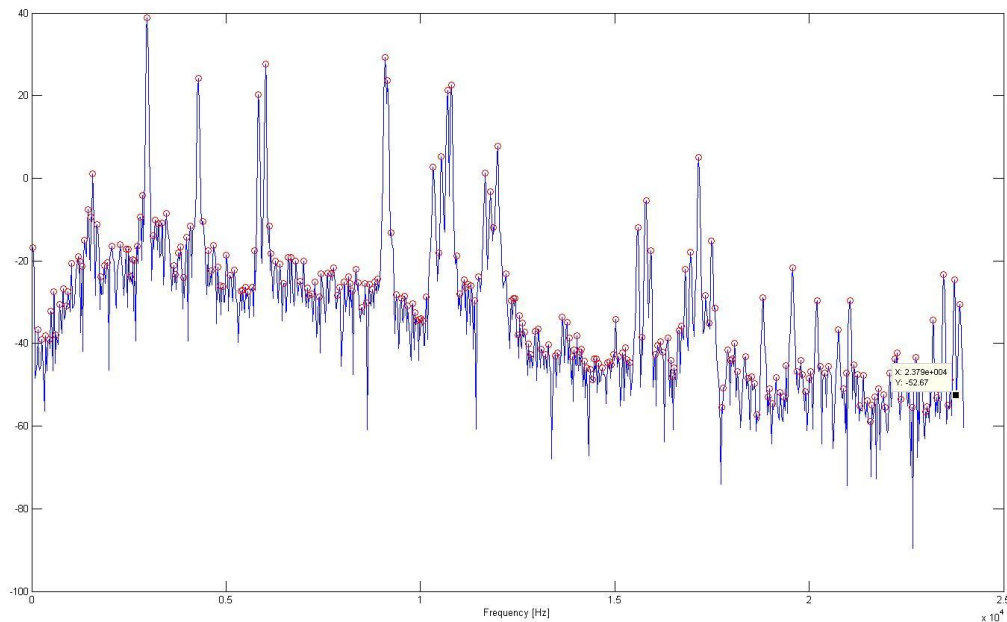
Slika 15 – amplitudni spektar jednog okvira zvuka zvonjenja čaše

U prvom koraku, označit ćemo kao potencijalne kandidate za šiljak (odnosno sinusnu komponentu signala) svaki indeks u spektru koji odgovara lokalnom maksimumu, odnosno za koji je zadovoljeno:

$$|X(k)| > |X(k + 1)|,$$

$$|X(k)| > |X(k - 1)|.$$

Slika 16 prikazuje tako dobivene indekse, označene crvenim kružićem na grafu. Vidljivo je da je takvih indeksa iznimno puno i da većina od njih nikako nije posljedica sinusne komponente.



Slika 16 – amplitudni spektar okvira s označenim kandidatima

Iz ovako dobivenih indeksa, kojih može biti čak i do $\frac{NFFT}{2}$, potrebno je na neki način izdvojiti one za koje je vjerojatnije da su rezultat sinusne komponente u signalu, odnosno one koji su "jači".

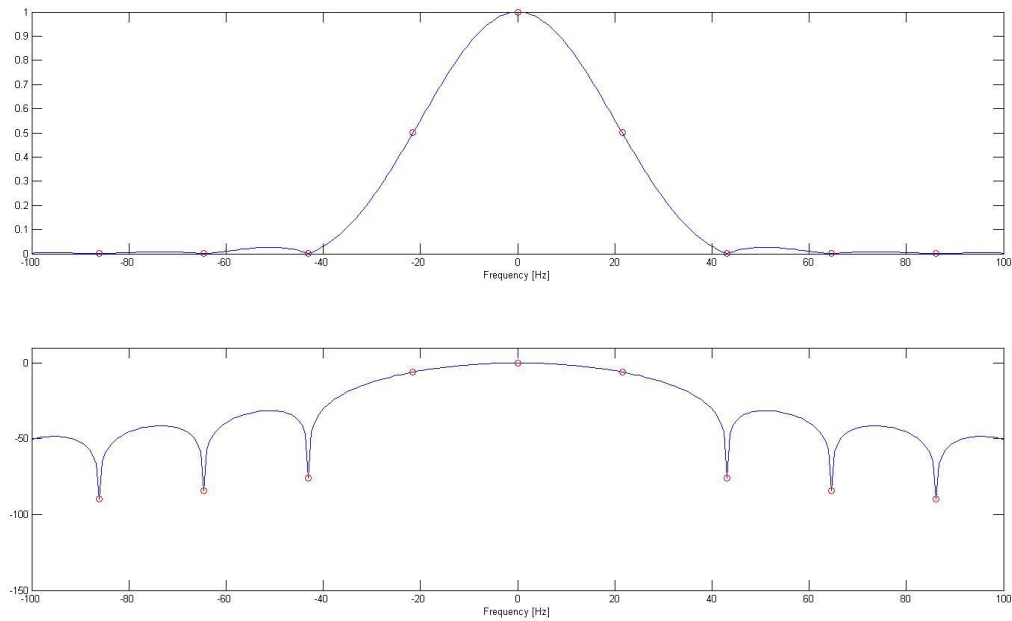
Pri traženju sinusnih komponenti treba imati na umu kakav je spektar korištenog vremenskog otvora. Naime, znamo da množenju u vremenskoj domeni odgovara konvolucija diskretnih Fourierovih transformacija:

$$x[n]w[n] \Leftrightarrow X[e^{j\omega}] * W[e^{j\omega}]$$

Neka su u signalu $x[n]$ prisutne sinusne komponente frekvencija redom $\omega_1, \dots, \omega_r$. Množenjem $x[n]$ i $w[n]$ dolazi do konvolucije njihovih spektara. S obzirom da je Fourierova transformacija sinusoide frekvencije ω_i jednaka Diracovoj funkciji na $-\omega_i$ i ω_i , Fourierova transformacija te sinusoide pomnožene s vremenskim otvorom odgovara preslikavanju spektra vremenskog otvora na frekvencije $-\omega_i$ i ω_i .

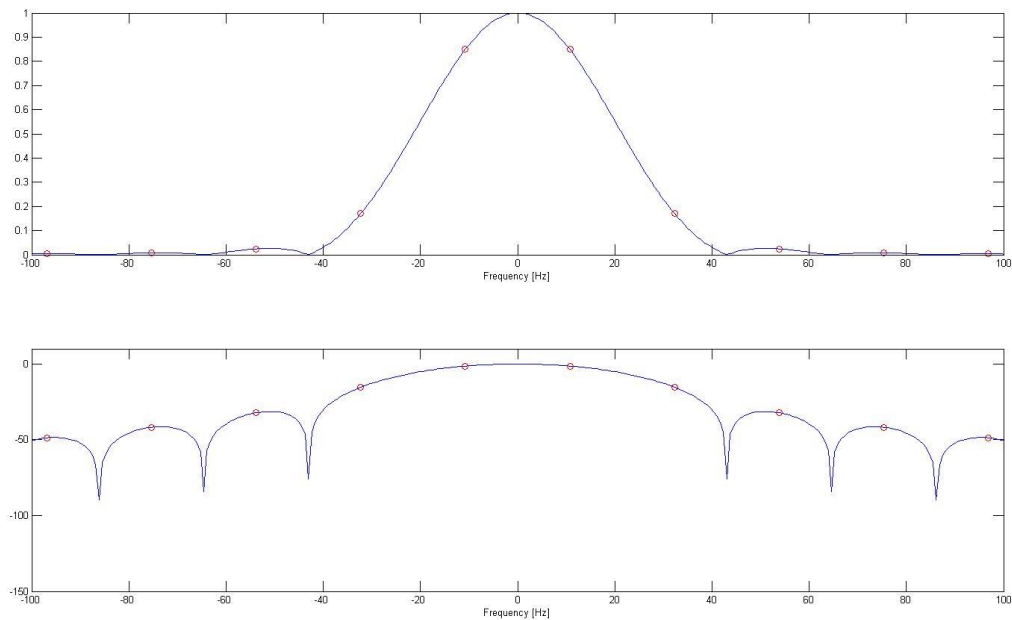
$$\mathcal{F}(w[n] \sin(\omega_i n)) = W(e^{j(\omega-\omega_i)}) + W(e^{j(\omega+\omega_i)}).$$

Dakle, u spektru signala možemo očekivati spektar vremenskog otvora na frekvencijama sinusoida prisutnih u signalu. Slika 17 prikazuje amplitudni spektar Hannovog prozora te njegov logaritam. Plavom linijom označen je interpolirani spektar koji odgovara prozoru izračunatom uz dodavanje velikog broja nul-uzoraka, dok je crvenim kružićima označen DFT spektar samog prozora od 2048 uzoraka.



Slika 17 – amplitudni spektar Hannovog prozora u linearnom i logaritamskom mjerilu i uzorci u slučaju množenja sa sinusom frekvencije koja odgovara frekvenciji indeksa u DFT-u

Iz ovoga je vidljivo da su od najvećeg interesa kod traženja sinusnih komponenti u spektru svega tri dominantna uzorka otvora. Središnji uzorak ima amplitudu od oko 6 dB iznad susjedna dva, dok svi ostali padaju točno u nul-vrijednosti spektra. Međutim, takav odnos bi se pojavio samo u idealnom slučaju u kojem frekvencija sinusoide točno odgovara frekvenciji jednog indeksa u kojem se računa diskretna Fourierova transformacija. U najgorem slučaju frekvencija sinusoide bit će točno između frekvencija dvaju susjednih indeksa spektra. Slika 18 prikazuje takav slučaj. Dva su indeksa jednakih amplituda, gušenje susjedna dva je 14 dB ispod, a sada je prisutno konačno gušenje i na svim ostalim indeksima.



Slika 18 – amplitudni spektar Hannovog prozora u linearnom i logaritamskom mjerilu i uzorci u slučaju množenja sa sinusom frekvencije koja se nalazi točno između frekvencija dvaju indeksa u DFT-u

Odabir šiljaka koji vjerojatno pripadaju značajnoj sinusnoj komponenti bit će proveden u tri faze. Prvo će biti odabrani šiljci koji imaju amplitudu dovoljno veliku u odnosu na prve i druge susjede, odnosno koji zadovoljavaju:

$$2 \cdot 20 \log_{10}|X[k]| - 20 \log_{10}|X[k-1]| - 20 \log_{10}|X[k+1]| \geq 8 \text{ dB},$$

$$2 \cdot 20 \log_{10}|X[k]| - 20 \log_{10}|X[k-2]| - 20 \log_{10}|X[k+2]| \geq 15 \text{ dB}.$$

Ovim postupkom često neće biti obuhvaćeni šiljci koji imaju sporiji pad kod najbližih članova. Dodatan uvjet će stoga biti da šiljci koji su barem 20 dB iznad lijevog i desnog lokalnog minimuma budu odabrani, odnosno oni za koje vrijedi:

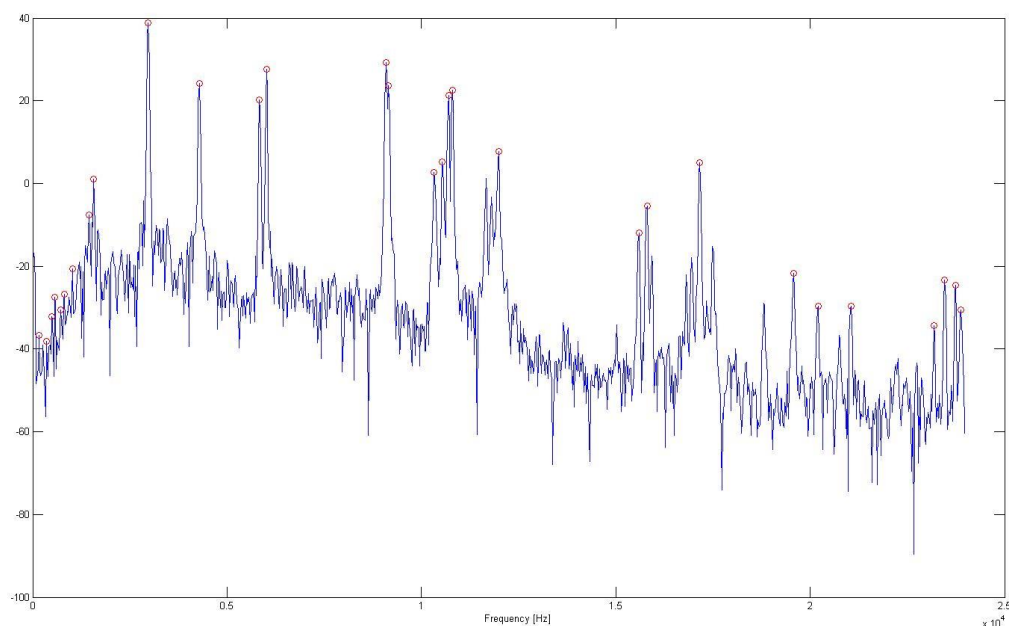
$$20 \log_{10}|X[k]| - 20 \log_{10}|X[k-l]| \geq 20 \text{ dB},$$

$$20 \log_{10}|X[k]| - 20 \log_{10}|X[k-r]| \geq 20 \text{ dB},$$

gdje su l i r indeksi lijevog i desnog lokalnog minimuma.

Konačno, na ovaj će način biti odabran i velik broj šiljaka u području šuma, vrlo niske amplitude. Zbog toga će se odbaciti svi šiljci koji su preko 20 dB ispod najjačeg šiljka u svakom od kritičnih pojaseva.

Slika 19 prikazuje šiljke odabrane nakon prethodno navedenih koraka detekcije. Vidljivo je da je mnogo šiljaka odabrano pri niskim frekvencijama na kojima su širine kritičnih pojaseva vrlo malene, ali oni će biti ignorirani kasnije u procesu određivanja modova.



Slika 19 – odabrani šiljci nakon svih koraka

5.1.1.1.2 Određivanje parametara modova

Iz određenih šiljaka za sve korake analize m (odnosno, za sve korake Fourierove transformacije na vremenskom otvoru) bit će određeni svi značajni modovi – njihove frekvencije, amplitude i faktori gušenja.

Kod određivanja modova iz šiljaka valja uzeti u obzir dvije stvari:

- zbog malih promjena u frekvenciji titranja, te zbog nepreciznosti pogađanja točne frekvencije od strane diskretne Fourierove transformacije, često se može dogoditi da se

istitavanje jednog moda (jedne frekvencije) očituje kao šiljci na nekoliko susjednih indeksa.

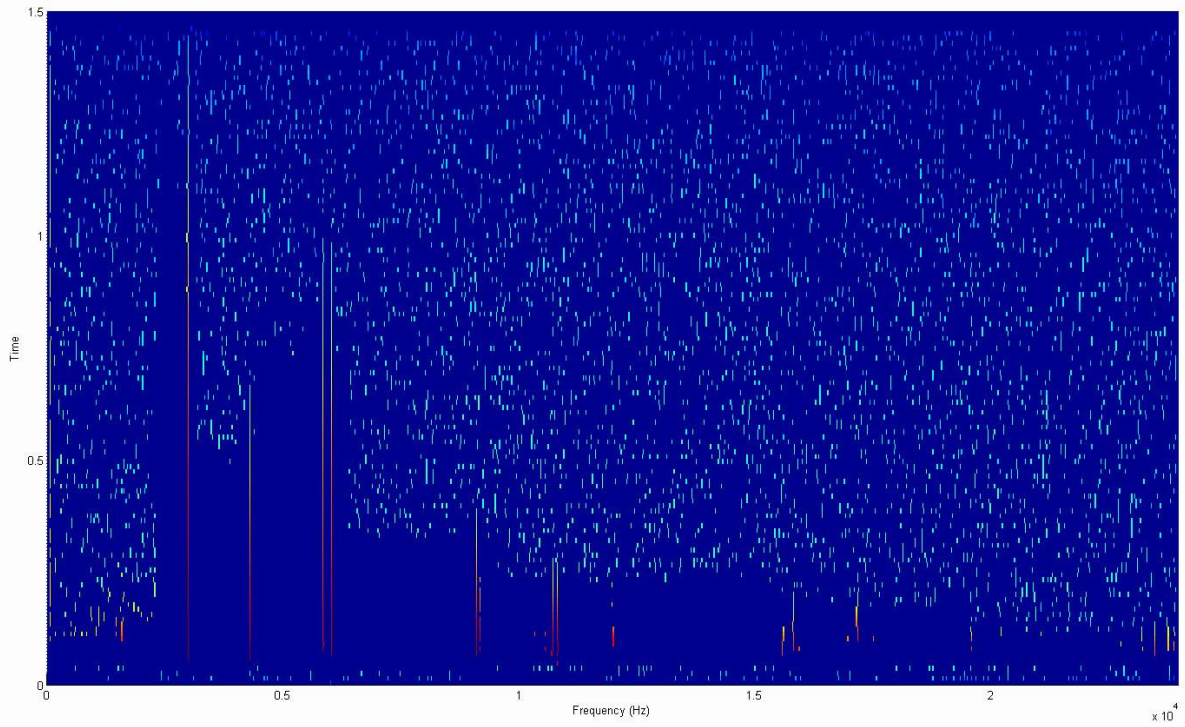
- očekuje se i da u određenim okvirima neka frekvencija ne bude otkrivena, bilo zbog nesavršenosti algoritma za detekciju šiljaka ili zbog stvarnog izostanka te frekvencije u određenom okviru analize.

Razvijeni algoritam koji uzima u obzir ta dva posebna slučaja jest sljedeći:

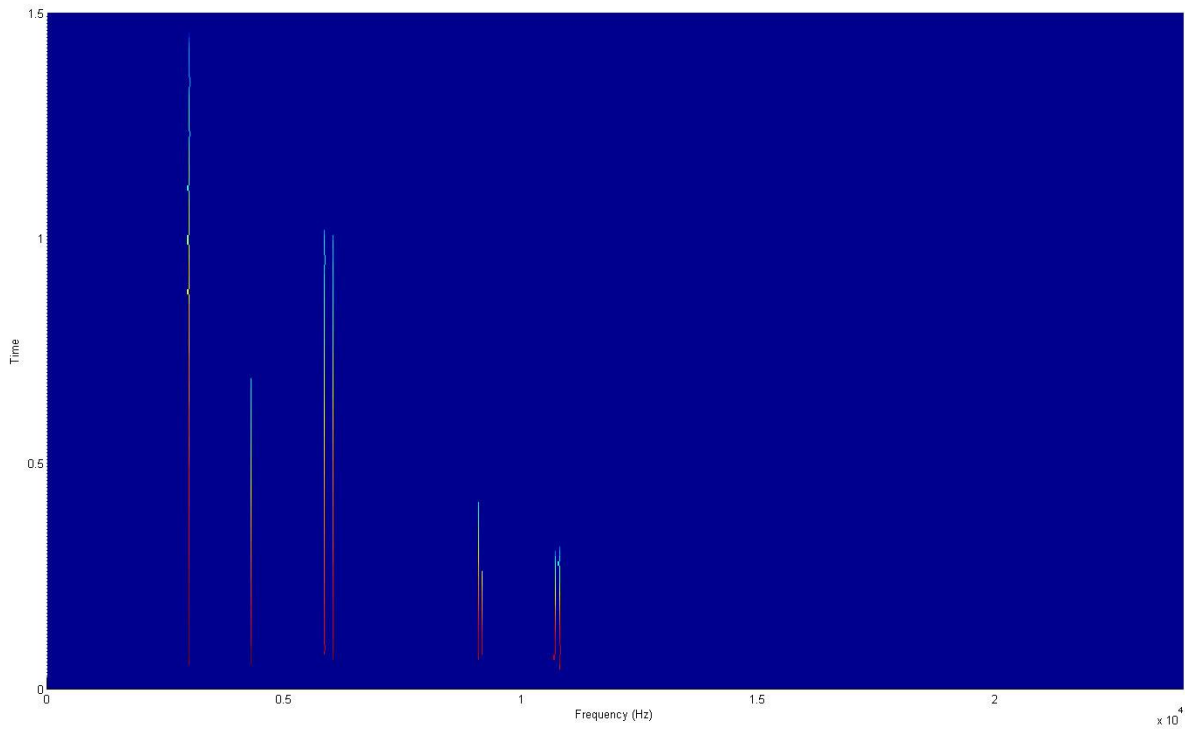
- Za svaki indeks u spektru počni od prvog okvira.
- Pronađi prvi sljedeći šiljak na tom indeksu. Ako ga nema, pređi na sljedeći indeks.
- Pregledavaj promatrani indeks, te njegove neposredne susjede u idućem okviru analize. Ako je na nekom od njih pronađen šiljak, mod se nastavlja i kroz idući okvir. Pomakni okvir analize i ponovi postupak.
- Dozvoli odsustvo šiljka u jednom okviru analize (prethodni i sljedeći okvir moraju imati šiljak).
- Ako se pojave dva uzastopna okvira bez šiljaka, usporedi duljinu pronađenog moda s najduljom detektiranom za ovaj indeks. Ako je dulji, zapamti njega.
- Ako promatrani okvir nije posljednji, vrati se na drugu točku. Ako jest, kreni sa sljedećim indeksom od prve točke.

Nakon što su izvršeni navedeni postupci, potrebno je ukloniti modove kraće ili tiše od zadane vrijednosti. Također, s obzirom da promatramo uvijek i dva susjedna indeksa, za očekivati je da će često biti detektirani modovi na dva ili tri susjedna indeksa. Zadržava se onaj srednji (u slučaju tri indeksa) ili onaj veće duljine (u slučaju dva susjedna indeksa).

Slika 20 prikazuje intenzitet svih šiljaka izračunatih za promatrani signal, a Slika 21 modove izračunate iz tih šiljaka predstavljenim algoritmom.



Slika 20 – prikaz svih šiljaka i njihovih intenziteta



Slika 21 – prikaz modova izračunatih iz šiljaka

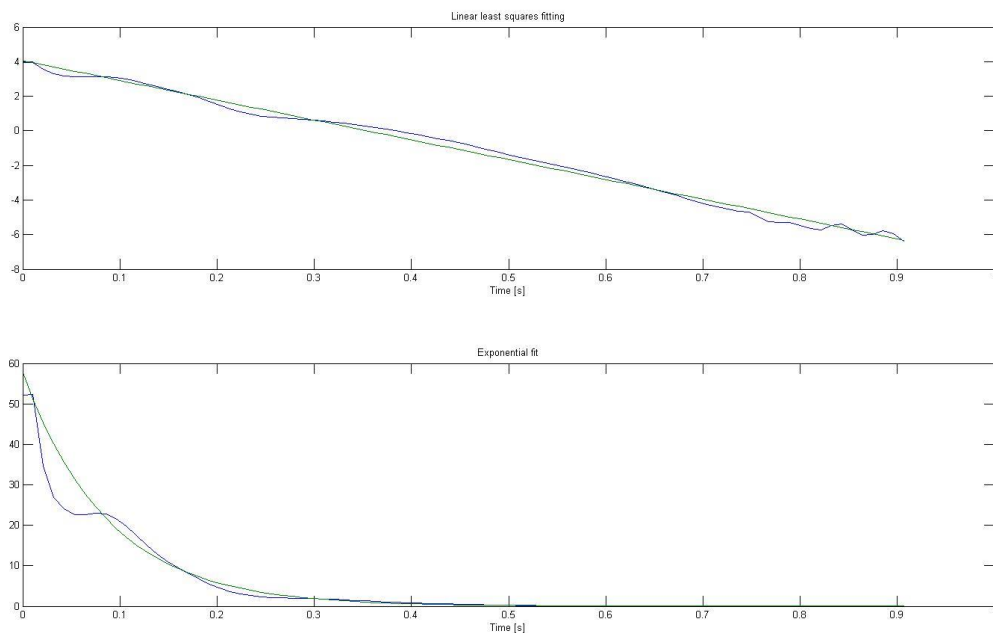
Sad kad su poznate frekvencije modova, potrebno je još izračunati amplitude i faktore gušenja, odnosno potrebno je amplitude modelirati eksponencijalno opadajućom eksponencijalnom funkcijom.

U tu svrhu prvo će biti izračunat prirodni logaritam amplituda modova:

$$Y[mod_i] = \ln|Amp[mod_i]|,$$

gdje je $Amp[mod_i]$ vektor amplituda i -tog moda. Nakon toga metodom najmanjih kvadrata određuje se pravac oblika $f = a + bt$ koji najbolje opisuje vektor Y . Eksponenciranjem točaka izračunatog pravca dobiva se eksponencijalna funkcija $Ae^{-\alpha t}$ koja opisuje kretanje amplitude moda. Iz poznatih točaka funkcije računaju se parametri A (amplituda moda) i α (faktor gušenja).

Slika 22 prikazuje postupak modeliranja amplitude. Na prvom grafu prikazan je logaritam amplitude moda i njegov model dobiven metodom najmanjih kvadrata, a na drugom prave amplitude moda i eksponencijalna funkcija dobivena eksponenciranjem pravca s prvog grafa.



Slika 22 – modeliranje amplitude metodom najmanjih kvadrata

Sad su poznate sve vrijednosti modova potrebne za sintezu – frekvencija, amplituda, faktor gušenja i vrijeme početka moda (koje je u promatranom slučaju gotovo jednako za sve modove budući da svi počinju titrati prilikom udarca).

5.1.2. Sinteza

Sad kad su poznati svi modovi moguće je sintetizirati zvučni signal jednim od dva ranije obrađena postupka – generiranjem valnog oblika u vremenskoj domeni, ili pobudom paralelnih rezonantnih filtara koji odgovaraju modovima. U ovom će poglavlju biti implementirana oba primjera, uz mogućnost dodavanja slučajnih promjena variranjem parametara pri sintezi.

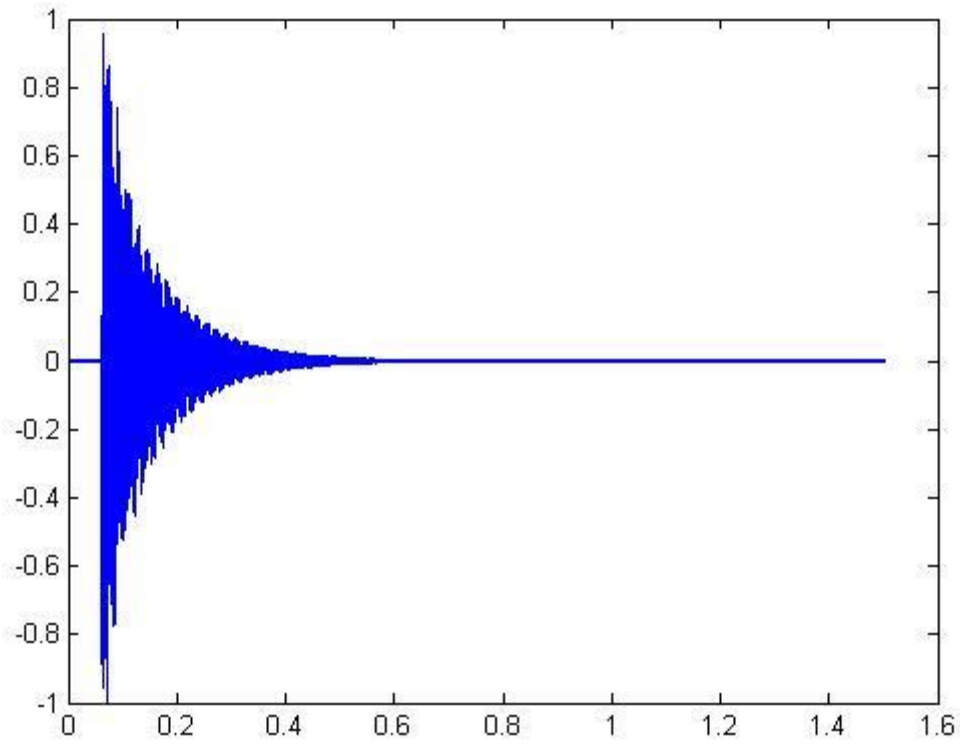
5.1.2.1 Generiranje valnog oblika sumom prigušenih sinusoida

Najjednostavnija metoda sinteze jest "po definiciji", generiranje valnog oblika sumiranjem eksponencijalno prigušenih sinusoida:

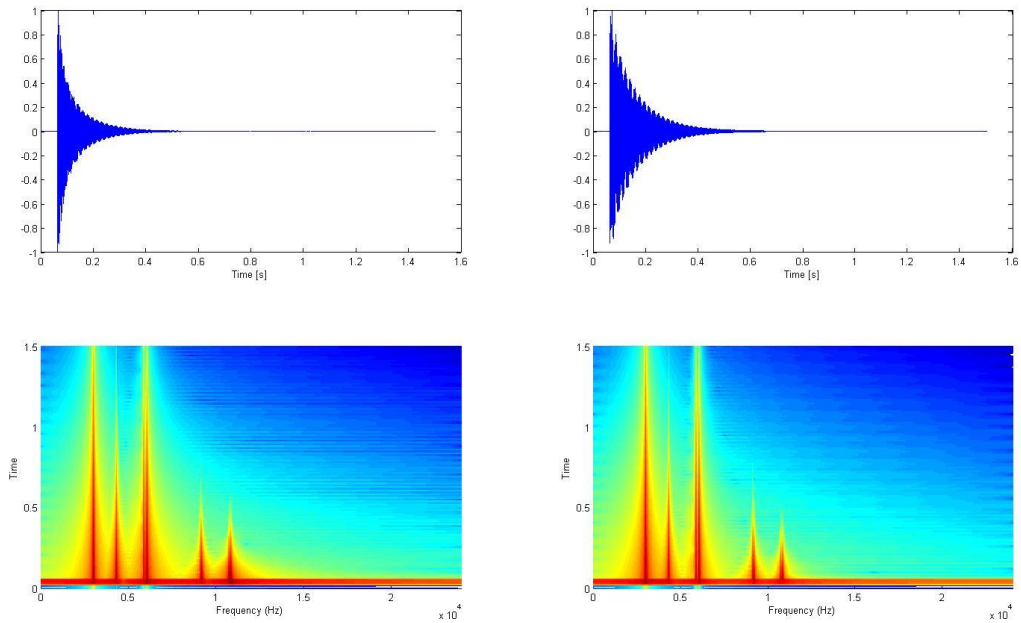
$$s(n) = \sum_i m_i(n) = \sum_i A_i e^{-\alpha_i n T_s} \sin\left(\frac{2\pi f_i}{f_s} n + \theta_i\right).$$

Slika 23 prikazuje valni oblik nastao ovakvim postupkom sinteze iz poznatih modova. Ovakav postupak ima broj računskih operacija (računanje eksponencijala, sinusa, množenja i sumiranja) proporcionalan ukupnoj duljini svih modova. U slučaju nešto kompleksnijih signala s više značajnih modova, te kod potrebe za istovremenom sintezom više različitih zvukova, ovakav način sinteze može dovesti to relativno visokih zahtjeva za procesorskim vremenom.

Kao što je u poglavlju 4.1.1 spomenuto, pobudom (udaranjem) predmeta na različitim mjestima dolazi do različitih amplituda titranja pojedinih modova. Variranjem parametra A_i moguće je sintetizirati različite zvukove koji i dalje zvuče kao udarac u staklenu čašu, ali na različitim mjestima. Slika 24 prikazuje valni oblik i spektrogram zvukova sintetiziranih slučajnim variranjem amplituda modova u rasponu $\pm 50\%$.



Slika 23 – valni oblik sintetiziranog zvuka



Slika 24 – valni oblik i spektrogram dvaju valnih oblika sintetiziranih variranjem amplituda modova

5.1.2.2 Sinteza zvuka pobuđivanjem rezonantnih filtara

U poglavlju 4.1.1.4.2 opisan je postupak modalne sinteze korištenjem paralelnih rezonantnih filtara. Ovdje će taj postupak biti iskorišten za sintezu zvuka zvonjenja čaše iz poznatih modova i njihovih parametara.

Za svaki od modova m_i iz njegovih parametara A_i , f_i i α_i potrebno je izračunati odgovarajuće koeficijente filtara b_0 , a_1 i a_2 , prema formulama:

$$a_{i1} = -2e^{-\frac{\alpha}{f_s}} \cos\left(\frac{2\pi f_i}{f_s}\right)$$
$$a_{i2} = e^{-\frac{2\alpha}{f_s}}$$
$$b_{i0} = \frac{A}{e^{-\frac{2\alpha}{f_s}} \left(4 \cos^2\left(\frac{2\pi f_i}{f_s}\right) - 1\right)}.$$

Promjena amplitude pojedinog moda, te slučajno variranje iste radi dobivanja dojma različitih udaraca ostvaruje se skaliranjem parametra b_0 .

Zvuk se dobiva pobuđivanjem svakog od filtara Kroeneckerovim delta impulsom i sumiranjem rezultata (odnosno, sumiranjem impulsnih odziva projektiranih filtara).

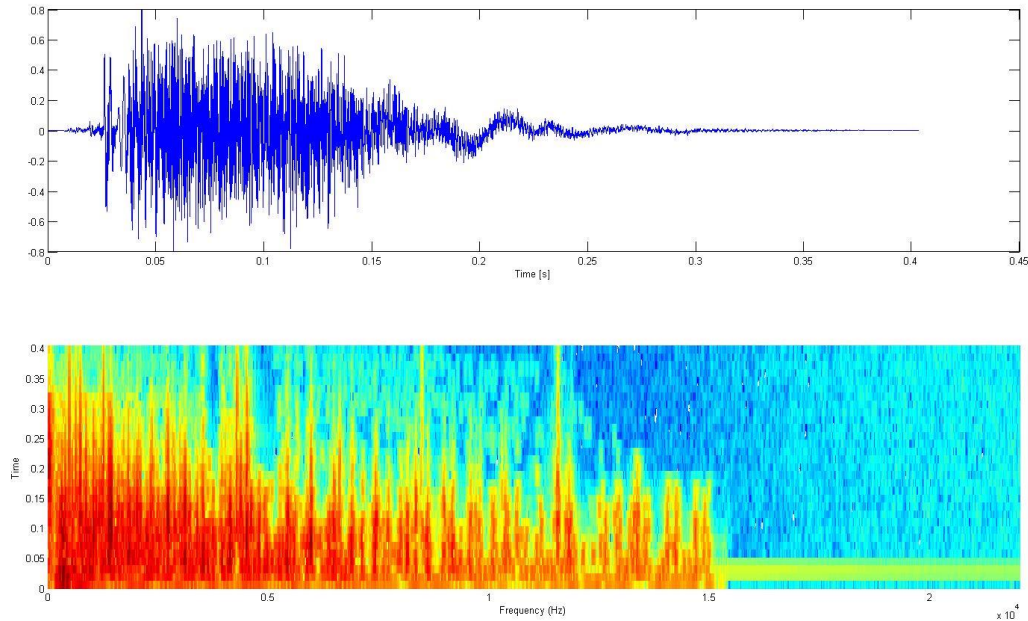
5.2. Udarac mača o metalni štit ili oklop

Idući cilj jest sintetizirati zvuk udarca mača o metalni štit ili oklop. Ovaj se zvuk često susreće u računalnim igrama. Po načinu nastajanja vrlo je sličan prethodnom (impulsna pobuda nakon koje slijedi istitravanje vlastitim frekvencijama), ali je kompliciraniji jer je prisutno mnogo više sinusnih komponenti, ali i šumovitog, ne-harmonijskog dijela koji je također potrebno sintetizirati.

5.2.1. Analiza zvuka i izvlačenje parametara

Slika 25 prikazuje valni oblik i spektrogram zvuka udarca mača o metalni štit. Spektrogram je napravljen s istim parametrima kao i ranije - Hannovim vremenskim otvorom duljine 2048 uzoraka, pomakom od 512 uzoraka (faktorom preklapanja od 1536 uzoraka). Iz spektrograma se

odmah uočava postojanje velikog broja izraženih frekvencijskih komponenti. To, a i preslušavanje snimke, jasno ukazuje da će biti potrebno i modelirati ne-harmonijski dio spektra.

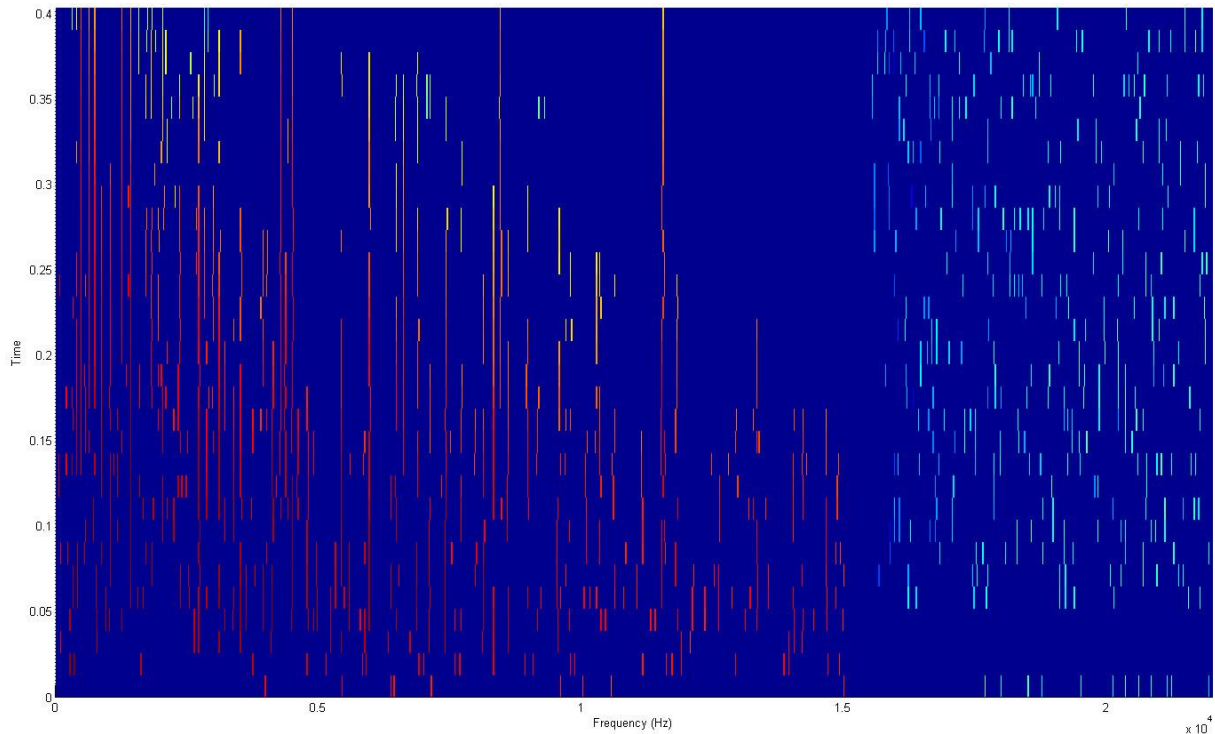


Slika 25 – valni oblik i spektrogram zvuka udarca mača o štit

5.2.1.1 Sinusoidalno modeliranje harmonijskog dijela spektra

Iako su harmonici prisutni u ovome signalu nastali udarcem, "ravni", odnosno ne mijenjaju značajno frekvenciju, te su samim time pogodni za modalnu sintezu, harmonijski dio spektra bit će modeliran općenitim sinusoidama promjenjive frekvencije i amplitude (postupak opisan u poglavlju 4.1.2.1). Jedan razlog za to jest mogućnost detaljnijeg uvida u algoritam sinusoidalnog modeliranja, a drugi spektralno bogatstvo signala zbog kojeg se mogu očekivati bolji rezultati i vjerniji zvuk uz korištenje općenitih sinusnih modela proizvoljnih amplituda i promjenjivih frekvencija, nego eksponencijalno prigušenih sinusoida fiksne frekvencije.

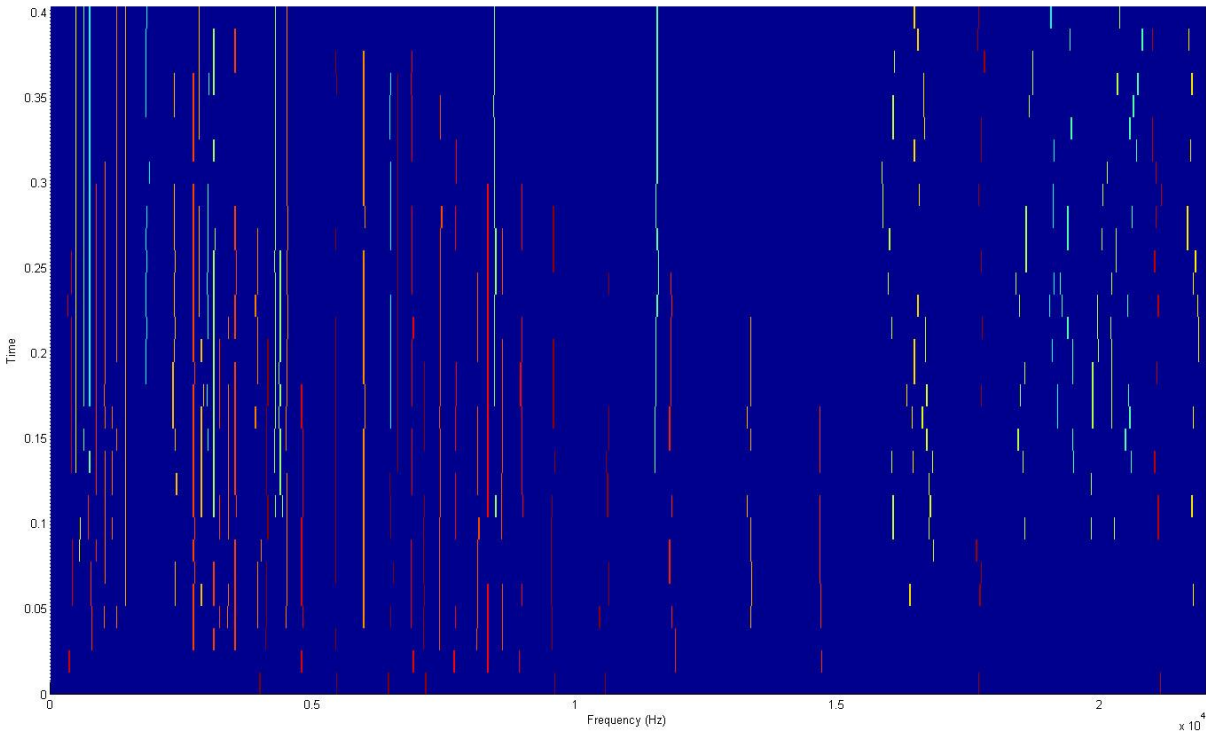
Prvi korak u sinusoidalnom modeliranju jest identičan onome kod modalne analize i sinteze – detekcija šiljaka. Korišten je algoritam detaljno opisan u poglavlju 5.1.1.1.1. Slika 26 prikazuje detektirane šiljke u amplitudnom spektru te njihov intenzitet. Vidljivo je nekoliko vrlo dugačkih i izraženih linija na nižim i srednjim frekvencijama, ali i mnogo kraćih linija u čitavom spektru.



Slika 26 – detektirani šiljci

Nad detektiranim šiljcima provodi se algoritam koji traži sinusoidalne putove spajajući šiljke u susjednim koracima okvira koji su međusobno najbliže (i dovoljno blizu). U implementiranom algoritmu za svaki se šiljak u koraku okvira j traži najbliži sinusoidalni put (jedan šiljak ili više njih već spojenih u put) iz okvira $j - 1$. Ukoliko je neki put najbliži za dva šiljka u trenutnom koraku, za nastavak toga puta bira se onaj šiljak koji mu je bliži.

Slika 27 prikazuje detektirane sinusoidalne putanje koje traju dulje od 10 okvira analize. Svaka od putanja iscrtana je svojom bojom i za svaku su pohranjene trojke *okvir, frekvencija, amplituda*. Frekvencija i amplituda izračunate su parabolnom interpolacijom šiljka, a okvir je broj okvira za koji vrijedi ta amplituda i frekvencija. Naime, algoritam radi na način da jedna sinusoidalna putanja može i preskočiti određen broj okvira ako tako dolazi do najbližeg šiljka.



Slika 27 – detektirani sinusoidalni putevi

5.2.1.2 Sinteza sinusoida u vremenskoj domeni

Sinteza detektiranih sinusoida moguće je provesti u frekventijskoj i vremenskoj domeni. Frekventijska domena, korištenjem inverzne Fourierove transformacije na vremenskom otvoru, omogućava brzu sintezu svih komponenti istovremeno. Ovaj se postupak ujedno i najčešće koristi. Međutim, problem je što se frekvencije sinusa mogu mijenjati samo u diskretnim trenucima pomaka vremenskog otvora što postaje primjetno kod signala s bržim promjenama u frekvenciji.

Ovdje će biti predstavljen postupak sinteze u vremenskoj domeni korištenjem sinusoida linearno vremenski promjenjive frekvencije.

Za svaku od detektiranih sinusoidalnih komponenti m_i poznat je niz trojki $okvir_j, freq_j, amp_j$, a duljina tog niza razlikuje se među sinusoidalnim komponentama. Algoritam koji će biti korišten za sintezu generirat će sinusnu funkciju koja će odgovarati vremenu od $okvir_j$ do $okvir_{j+1}$, linearno promjenjive frekvencije od f_j do f_{j+1} , linearno promjenjive amplitude od amp_j do

amp_{j+1} . Pamtit će se konačna faza svakog od odsječaka kako bi se osigurala neprekinutost sintetizirane sinusoide.

Neka su n_i diskretne točke u vremenskoj domeni koje se nalaze između početaka j -tog i $j+1$ -og okvira analize, te neka postoji p uzoraka u promatranom odsječku između ta dva okvira. Uzorci odsječka sinusa u tim točkama računaju se kao:

$$s_{jj+1}[i] = A_i \sin \left(\phi_j + 2\pi \left(\frac{f_j n_i}{f_s} + \frac{q \left(\frac{n_i}{f_s} \right)^2}{2} \right) \right),$$

gdje je q koeficijent linearne promjene frekvencije od f_j do f_{j+1} dobiven kao:

$$q = \frac{f_{j+1} - f_j}{n_p - n_0} f_s.$$

Amplituda trenutnog uzorka odsječka računa se kao:

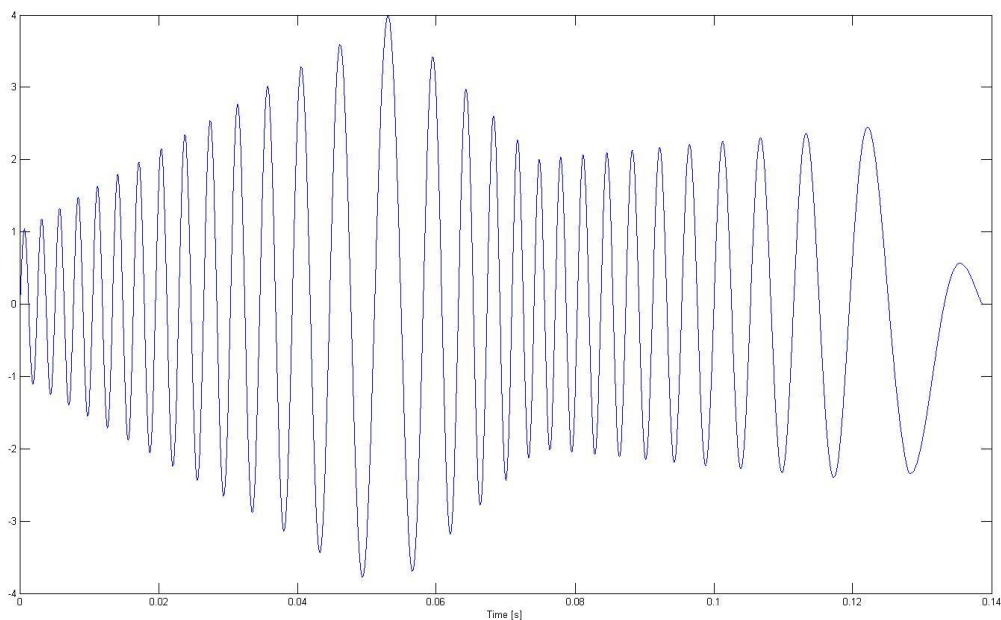
$$A_i = \left(1 - \frac{i}{p} \right) amp_j + \frac{i}{p} amp_{j+1}.$$

Konačno, računa se početna faza za idući okvir kako bi se osigurala neprekinutost:

$$\phi_{j+1} = \phi_j + 2\pi \left(\frac{f_j n_p}{f_s} + \frac{q \left(\frac{n_p}{f_s} \right)^2}{2} \right).$$

Slika 29 prikazuje primjer sinusoide s linearno promjenjivom frekvencijom i amplitudom koja je nastala ovim postupkom. Početna amplituda sinusoide je 1, zatim linearno raste do 4, pa pada na 2 i ponovno raste na 2.5. Početna frekvencija iznosi 412 Hz, pada na 125 Hz, raste na 334 Hz i pada na 60 Hz.

Treba napomenuti da ovakav postupak sinteze korišten u kombinaciji s ranije opisanom detekcijom sinusoida nije posve točan u energijskom smislu. Naime, kod detekcije sinusoida amplituda se određuje kao prosječna unutar jednog okvira analize, dok se u ovom algoritmu koristi kao početna amplituda za taj okvir. Zbog toga ukupna energija sintetiziranog signala nije točna.

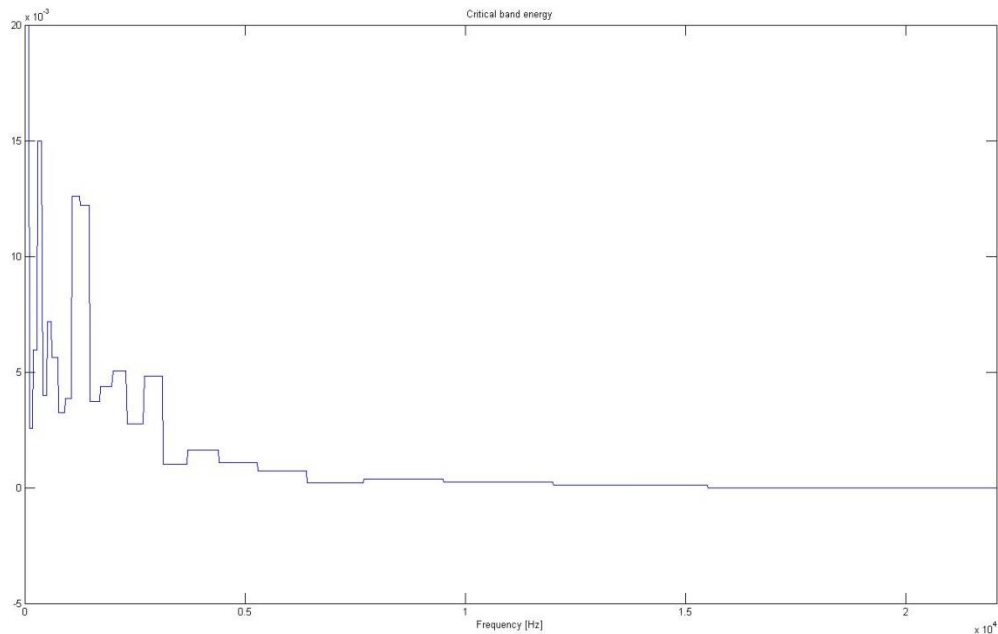


Slika 28 – sinteza sinusoide s linearno promjenjivom frekvencijom i amplitudom

5.2.1.3 Modeliranje ne-harmonijskog dijela spektra

Ne-harmonijski dio spektra bit će modeliran razinama energije u kritičnim pojasevima. Za svaki korak okvira bit izračunata ukupna energija u svakom od kritičnih pojaseva, ali ne uključujući indekse koji su bili obuhvaćeni sinusoidalnim modeliranjem. Indeks šiljka i njegova dva susjeda bit će zanemareni, a amplitude na ostalim indeksima u pojedinom kritičnom pojasu posumirane kako bi se dobila energija. Slika 29 prikazuje raspodjelu energije po kritičnim pojasevima u jednom okviru analize. Za svaki od okvira analize pohranjuje se vrijednost energije u svakom od kritičnih pojaseva, što se kasnije koristi prilikom sinteze.

$$E(cb_i) = \sum_{\substack{j \in cb_i \\ j \neq \text{sinepeak}}} Amp(j).$$



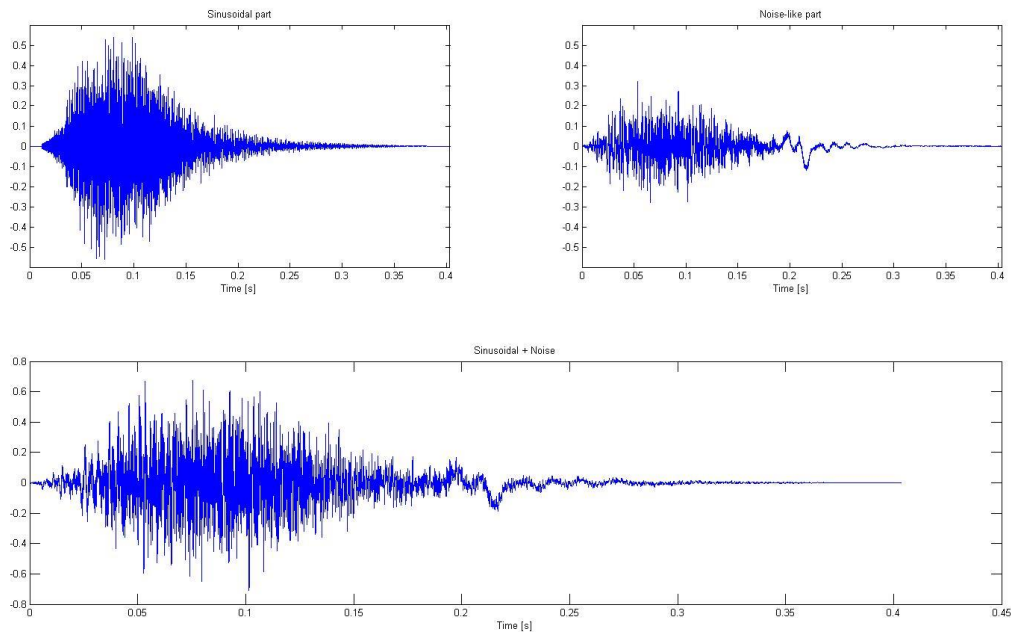
Slika 29 – raspodjela energije ne-harmonijskog dijela signala po kritičnim pojasevima

5.2.1.4 Sinteza ne-harmonijskog dijela signala

Iz poznatih raspodjela energije po kritičnim pojasevima za svaki korak okvira analize valni oblik koji odgovara ne-harmonijskom dijelu signala može se sintetizirati filtriranjem bijelog šuma (koji ima približno ravan amplitudni spektar) kroz pojasno propusne paralelne filtre, svaki od kojih predstavlja jedan kritični pojas. Druga mogućnost jest za svaki okvir analize napraviti inverznu Fourierovu transformaciju nad spektrom koji je opisan s energijama po kritičnim pojasevima, ali uz slučajnu fazu na svim indeksima spektra.

5.2.2. Konačan rezultat i subjektivni dojmovi

Slika 30 prikazuje valne oblike sinusoidalnog dijela, ne-harmonijskog dijela i njihove sume. Dobiveni zvuk je zadovoljavajuće kvalitete, ali osjetan je nedostatak tranzijenta, odnosno trenutak samog udarca nije dovoljno intenzivan. Variranjem amplituda svake od sinusoidalnih komponenti ponovno se dobiva jednostavna metoda za variranje zvuka kod svake sinteze.



Slika 30 – sinusoidalni i ne-harmonijski dio sintetiziranog zvuka, te konačni valni oblik

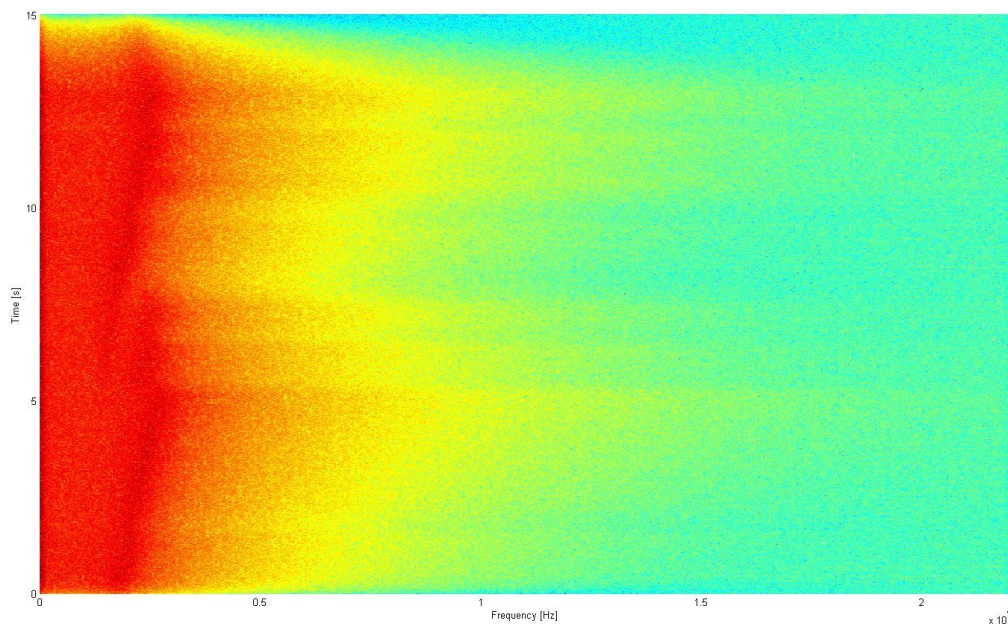
5.3. Naleti vjetra

Zadnji zvučni efekt koji će biti obrađen jest zvuk snažnog vjetra. Radi se o još jednoj klasi zvukova prisutnih u ogromnom broju računalnih igara. Kod takvih pozadinskih, ambijentalnih zvukova posebno može doći do izražaja nepromjenjivost uzrokovana ponavljanjem jedne snimke.

5.3.1. Analiza zvuka

Očito je da je zvuk vjetra u bilo kojem obliku iznimno šumovit i da metode koje se baziraju na modeliranju sinusnih komponenti nisu upotrebljive. Slušajući nekoliko snimki zvuka vjetra i analizom spektrograma (Slika 31) kao razuman odabir nameće se oblikovanje spektra bijelog šuma projektiranjem odgovarajućih filtara. Konkretna primjer koji će biti promatran i čiji je spektrogram prikazan jest zvuk vrlo snažnog vjetra koji "fijuče", odnosno kod kojeg je dio spektra na određenoj, promjenljivoj frekvenciji čujniji od ostatka. Na spektrogramu se jasno vidi

ta linija koja ide od oko 2000 Hz do 2500 Hz u prvoj polovici signala, te ponovno u sličnom frekvencijskom rasponu i u drugoj polovici signala.

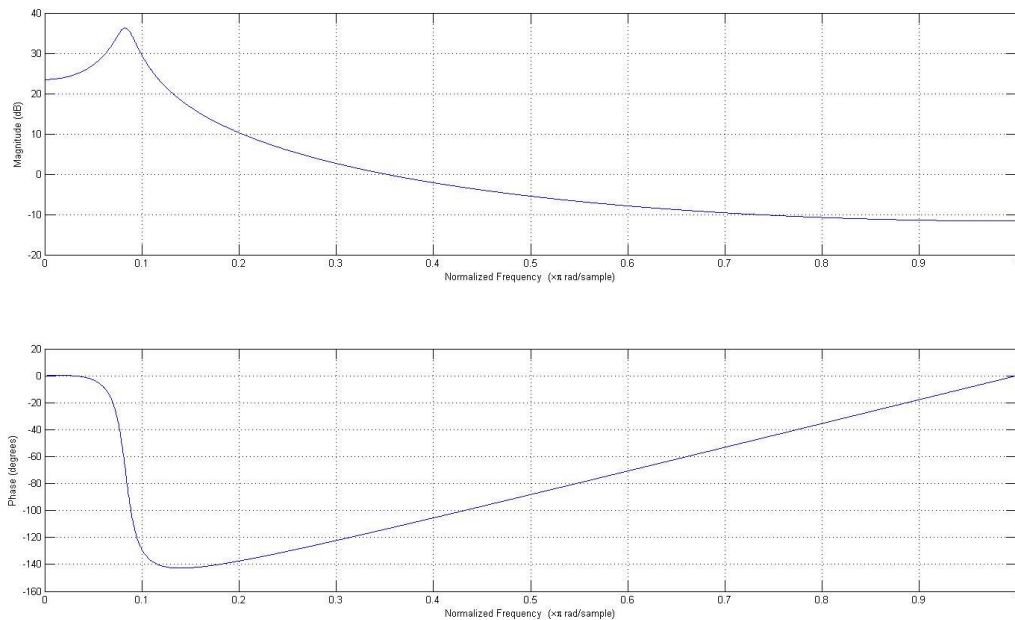


Slika 31 – spektrogram zvuka snažnih udara vjetra

5.3.2. Rezonantni filtri drugog reda za modeliranje spektra

Za oblikovanje spektra bijelog šuma i dobivanje zvuka snažnog vjetra koji "fijuče" koristit će se rezonantni filter drugog reda opisan u poglavlju o modalnoj sintezi (4.1.1.4.2). Ovaj puta neće se koristiti njihov impulsni odziv već će se filter konstantno pobuđivati novim uzorcima bijelog šuma.

Iz frekvencijske karakteristike rezonantnog filtra (Slika 32, uz rezonantnu frekvenciju 2000 Hz i radijus polova $R = 0.97$) vidi se da ona približno odgovara spektru analiziranog zvuka vjetra.



Slika 32 – frekvencijska karakteristika rezonantnog filtra s dva pola

5.3.2.1 Modeliranje promjene frekvencije i amplitude slučajnom šetnjom i Brownovim gibanjem

Kako bi se postigla veća realističnost sintetiziranog zvuka filter opisan u prošlom potpoglavlju mora imati vremenski promjenjive parametre. U suprotnome dobiveni zvuk ima konstantan "fijuk" na fiksnoj frekvenciji i ne zvuči uvjerljivo.

U prvom promatranom slučaju promjena rezonantne frekvencije filtera bit će modelirana slučajnom šetnjom.

Neka je f_i rezonantna frekvencija filtera u i -tom diskretnom vremenskom trenutku. Rezonantna frekvencija u idućem trenutku, f_{i+1} , računat će se kao:

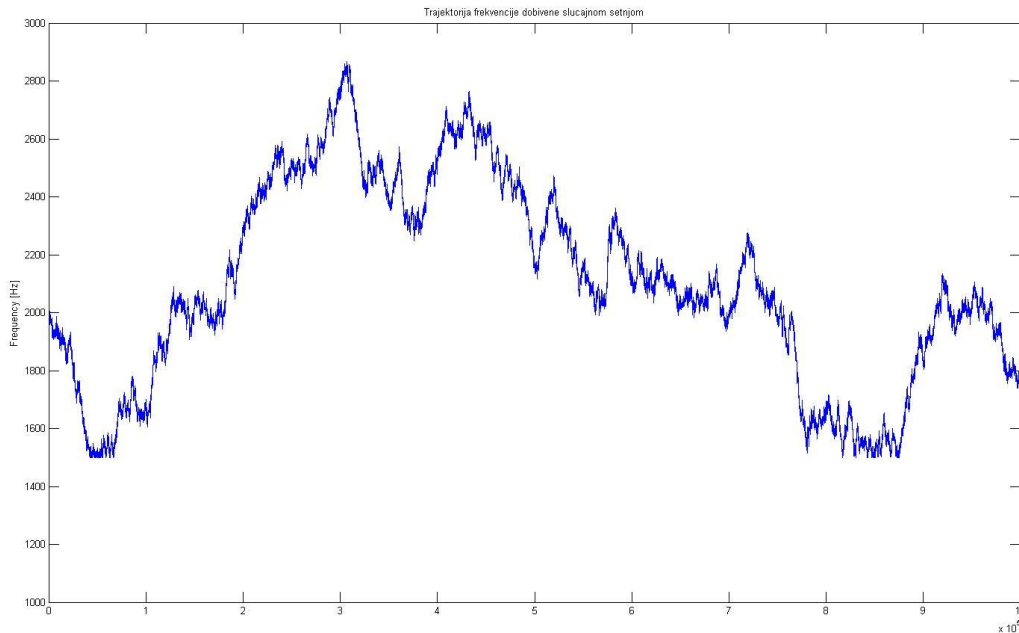
$$f_{i+1} = f_i + Y_i \Delta f,$$

gdje je Δf fiksni korak u Hz, a Y_i diskretna slučajna varijabla s razdiobom

$$Y_i \sim \begin{pmatrix} -1 & 1 \\ p_i & q_i \end{pmatrix}.$$

Drugim riječima, Y_i poprima ili vrijednost -1 , s vjerojatnošću p_i , ili vrijednost 1 , s vjerojatnošću $q_i = 1 - p_i$.

Slika 33 prikazuje trajektoriju frekvencije dobivene slučajnom šetnjom uz $f_0 = 2000 \text{ Hz}$, $p_i = q_i = 0.5$ i diskretni korak frekvencije $\Delta f = 1 \text{ Hz}$.



Slika 33 – trajektorija rezonantne frekvencije filtra dobivena slučajnom šetnjom

Drugi mogući način modeliranja promjenjive frekvencije jest koristiti ideju Brownovog gibanja. Brownovo gibanje (ili Wienerov proces) jest slučajni proces $W(t)$ za koji vrijedi [7]:

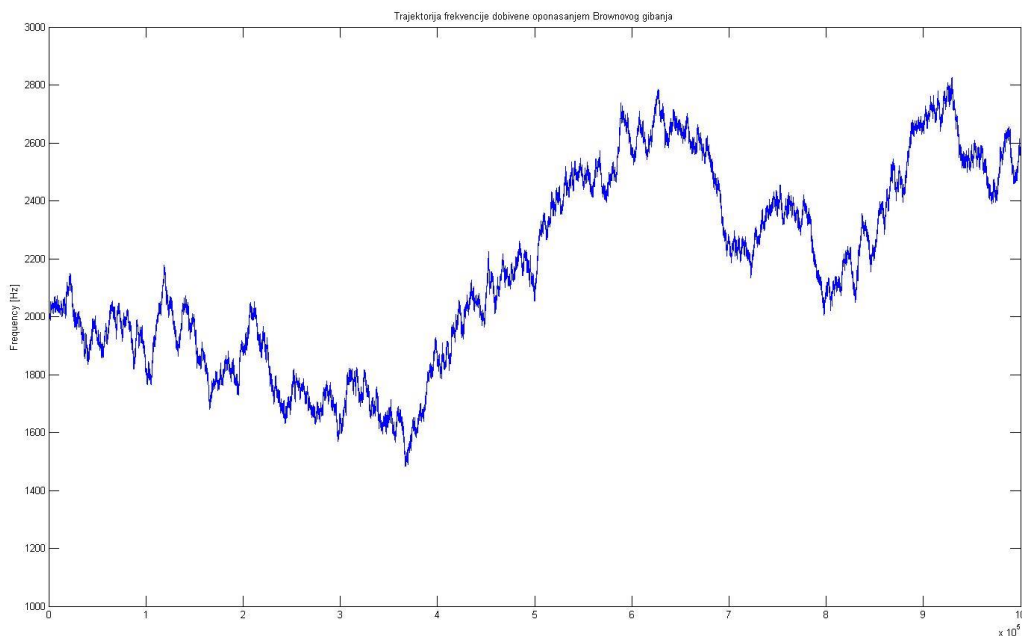
1. $W(0) = 0$.
2. W ima nezavisne priraste.
3. Prirast $W(t) - W(s)$ ima normalnu razdiobu $N(\mu(t - s), \sigma^2(t - s))$.

Model koji će opisivati promjenu frekvencije u vremenu neće biti u strogo smislu Brownovo gibanje jer će početna vrijednost frekvencije biti pozitivna, te će se prirasti dešavati samo u diskretnim, ekvidistantnim vremenskim trenutcima.

Promjena frekvencije inspirirana Brownovim gibanjem bit će opisana kao:

$$f_{i+1} = f_i + X_i \Delta f,$$

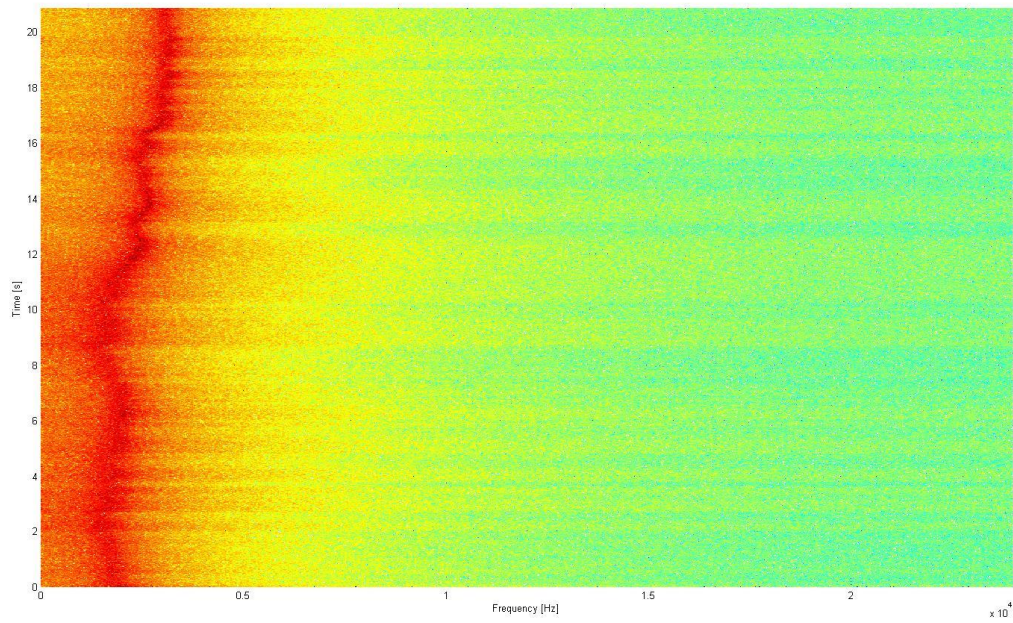
dakle jednako kao i kod slučajne šetnje. Međutim, u ovom slučaju X_i je kontinuirana slučajna varijabla s normalnom razdiobom. Parametre normalne razdiobe (srednja vrijednost i standardna devijacija) te konstantnu Δf potrebno je odabrati na način da se osigura da frekvencija uvijek ostane u rasponu očekivanom za zvuk vjetra.



Slika 34 – trajektorija frekvencije dobivena oponašanjem Brownovog gibanja

Slušanjem zvuka koji nastaje kod snažnog vjetra primjećuje se porast frekvencije kod snažnijih udara. Modeliranje te pojave može se postići povezivanjem frekvencije i amplitude. Uz linearnu vezu frekvencije i amplitude rezonantnog filtra postižu se vrlo dobri rezultati.

Slika 35 prikazuje spektrogram sintetiziranog zvuka snažnog vjetra uz promjenu rezonantne frekvencije i amplitude filtra po modelu s Brownovim gibanjem. Iz spektrograma je uočljivo da gušenje rezonantnog filtra s dva pola na višim frekvencijama nije toliko snažno te su visoke frekvencije izraženiji u usporedbi s analiziranom snimkom, te da su niske frekvencije nešto prigušenije, no subjektivan dojam kvalitete zvuka je svejedno odličan.



Slika 35 – spektrogram sintetiziranog zvuka snažnog vjetra

6. Zaključak

U ovome radu razvijene su i analizirane tri metode sinteze zvukova temeljene na analizi spektra. Provedena je analiza na tri konkretna zvučna efekta koji su zatim i sintetizirani iz detektiranih parametara.

Detekcija modova i njihovih parametara iz spektra zvuka daje mogućnost iznimno brze i procesorski nezahtjevne sinteze korištenjem paralelnih rezonantnih filtara. Broj parametara koje je potrebno pamtit za sintezu je iznimno malen, a moguće je i vrlo jednostavno varirati zvuk u samom procesu sinteze kako bi se dobili malo drugačiji zvukovi pri svakoj sintezi. Mana ove metode jest što je upotrebljiva samo za klasu zvukova koji su nastali udarcem i vibracijama nekog objekta, te što je za velik broj zvukova potrebno dodatno modelirati rezidual (šum) i tranzijent signala.

Druga obrađena metoda, sinusoidalno modeliranje, jest nešto općenitija i prigodna za širu klasu zvukova. Omogućava modeliranje i vremenski promjenjivih frekvencija proizvoljnih amplituda. Međutim, proces sinteze je složeniji nego u slučaju modalne sinteze, a često je potrebno i pohraniti mnogo veći broj parametara. Uz kombinaciju modeliranja šuma i tranzijenata predstavlja jednu iznimno općenitu metodu prikladnu za gotovo sve klase zvukova.

Treća korištena metoda, oblikovanje spektra bijelog šuma filtrima, korištena za sintezu zvuka snažnog vjetra također pokazuje da se individualnim pristupom za svaki od zvukova koje se želi sintetizirati ipak mogu dobiti bolji rezultati nego što bi to bilo u slučaju korištenja neke općenite metode za sve zvukove.

Na tri primjera zvukova pokazalo se da je i ovakvim jednostavnim i općenitim postupcima moguće dobiti zadovoljavajuće rezultate i da je korištenje umjetno sintetiziranih zvukova u računalnim igrama vrlo realna alternativa reprodukciji snimljenih zvukova. Kompliciranijim metodama, a posebice kombinacijom modalne sinteze i metode konačnih elemenata, mogao bi se razviti sustav za analizu i sintezu odličnih performansi i lako primjenjiv u igrama. Nažalost, dobivanje sintetiziranih zvukova koji su kvalitetom usporedivi sa snimljenim zvučnim efektima i

dalje je vrlo velik posao i vjerojatno najveća prepreka ka značajnijoj tranziciji prema umjetnoj sintezi.

7. Literatura

- [1] First Video Game, Wikipedia, http://en.wikipedia.org/wiki/First_video_game, 18.11.2013.
- [2] CHAN, NORMAN: "A Critical Analysis of Modern Day Video Game Audio", <http://www.gamessound.com/texts/chan.pdf>, 18.11.2013.
- [3] FARNELL, ANDY: "Procedural Audio", <http://obiwannabe.co.uk/html/papers/proc-audio/node1.html>, 09.10.2013.
- [4] FARNELL, ANDY: "Designing Sound ", The MIT Press, 2010.
- [5] COOK, PERRY R.: "Real Sound Synthesis for Interactive Applications", CRC Press, 2002.
- [6] MACAULAY, ROBERT J., QUARTIERI, THOMAS F.: "Magnitude-Only Reconstruction Using a Sinusoidal Speech Model", Proc. IEEE ICASSP-84, vol. 2, pp. 27.6.1-27.6.4, 1984.
- [7] ELEZOVIĆ, NEVEN: "Gaussovi procesi", skripta iz predmeta Stohastički procesi, http://www.fer.unizg.hr/_download/repository/procesi05a-101112.pdf, 21.12.2013.
- [8] REN, ZHIMIN, LIN, C. MING: "Example-Guided Physically Based Modal Synthesis", http://gamma.cs.unc.edu/AUDIO_MATERIAL/examplebasedsoundsynthesis.pdf, 20.11.2013.
- [9] VAN DEN DOEL, KEES, KRY, PAUL G., PAI, DINESH K.: "FoleyAutomatic: Physically-based Sound Effects for Interactive Simulation and Animation", <http://www.cs.ubc.ca/~kvdoel/publications/foleyautomatic.pdf>, 19.11.2013.
- [10] ZITA, ANDREAS: „Computational Real-Time Sound Synthesis of Rain“, <http://www.diva-portal.org/smash/get/diva2:19156/FULLTEXT01.pdf>, 29.11.2013.
- [11] SMITH, JULIUS O. III.: „Spectral Audio Signal Processing“, <https://ccrma.stanford.edu/~jos/sasp/>, 18.10.2013.
- [12] SMITH, JULIUS O. III.: „Physical Audio Signal Processing“, <https://ccrma.stanford.edu/~jos/pasp/>, 18.10.2013.

Sažetak

Proceduralno generiranje zvučnih efekata u računalnim igrama

U začetima industrije računalnih igara zvuk je nastajao sintetički, generiranjem odabranog valnog oblika. Razvojem računala grafika i sam doživljaj u igrama postajali su sve realističniji. Umjesto razvoja u sintezi zvuka počeli su se koristiti nasnimljeni zvučni efekti.

Sinteza zvučnih efekata danas je ponovno aktualna tema. Iako su prepoznate mnoge prednosti (ušteda memorije, mogućnost ozvučenja više događaja, lakše varijacije u zvuku i sl.), sinteza zvuka prilično sporo pronalazi povratak u računalne igre. Mogući razlog za to jest što je potrebno mnogo istraživanja i eksperimentiranja kako bi se postigla zadovoljavajuća kvaliteta zvuka, posebice u slučaju razvijanja fizikalnih modela.

Neke metode analize i sinteze u frekvencijskoj domeni omogućuju jednostavnu sintezu široke klase zvukova. Modalna sinteza uz vrlo mali broj parametara može opisati vibracije predmeta, a raznim poboljšanjima (npr. modeliranje šuma te metoda konačnih elemenata za točniju detekciju modova) nudi mnogo mogućnosti i dobre rezultate.

Sinusoidalne modeliranje je nešto općenitija metoda prikladnija za veći broj zvukova. U kombinaciji s modalnom sintezom i modeliranjem šuma omogućava sintetiziranje još nekih klasa zvukova. Korištenjem algoritama temeljenih na modalnoj sintezi i sinusoidalnom modeliranju dobiveni su zadovoljavajući rezultati kod sinteze nekih zvukova udaraca.

Filtriranje spektra bijelog šuma je metoda sinteze prikladna za šumovite zvukove poput vjetra i kiše. Uz variranje parametara filtara uvođenjem slučajnih procesa moguće je vjernije oponašati zvukove iz prirode.

Ključne riječi: zvuk, sinteza, modalno, vibracije, filtri, modeliranje.

Procedural sound synthesis in video games

In the very beginning of computer games sound was generated synthetically by generating a pre-selected waveform. With the development of computers and processing power computer graphics

evolved, making computer games much more realistic. However, sound synthesis didn't follow that path and recorded sounds got used instead.

Today sound synthesis in computer games is again a topic of interest. Although a lot of advantages have been recognized (such as lower memory consumption, allowing more different sounds to be synthesised, easy sound variation etc.), sound synthesis is very slow in making a return to computer games. Possible reason for that is that it requires a lot of research and experimenting in order to create sound of desired quality, especially in case of developing physical models.

Some frequency-domain-based methods allow fairly simple synthesis of a wide variety of sounds. Modal synthesis can be used to describe vibrations of struck objects with just a few parameters. With some improvements (noise modelling, finite elements method) modal synthesis can be used for a lot of sounds.

Sinusoidal modelling is a more general method appropriate for more sounds, but having higher memory consumption and more complex synthesis. Combined with modal synthesis and noise modelling some satisfactory sounds of struck objects were synthesised.

Finally, white noise filtering is appropriate for synthesising more noise-like sounds such as wind or rain. Variation of filter parameters and characteristics by adding a random process gives a more natural sound.

Keywords: sound, synthesis, modal, vibration, filters, modeling.