

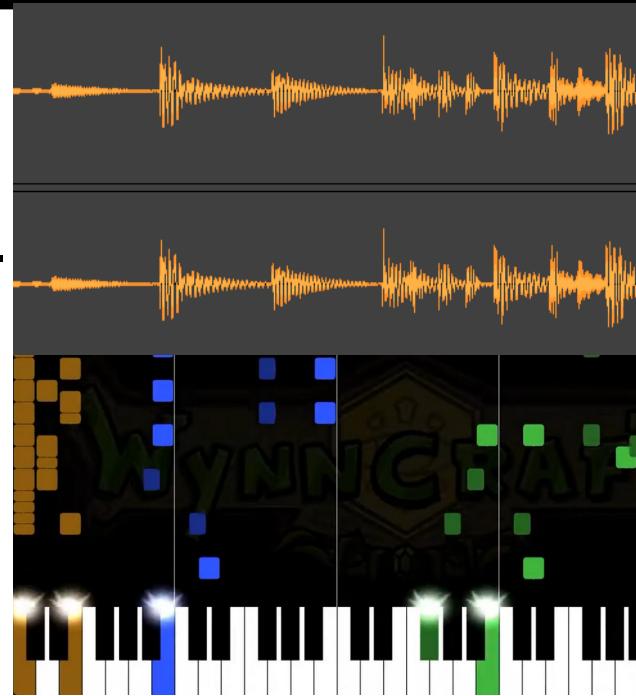
Note Block hangminták felismerése zenefájlban

Vértesaljai Bálint
Konzulens: Grad-Gyenge László György
Mérnökinformatikus BSc, Szakdolgozat



Hangfájlok formátuma

- Wave: wav/mp3/ogg/...
- Számértékek másodpercenként 44100-szor pl.
- Midi: ~kotta
- Zenehang kezdet- és végidőpontok
- Konverzió nem egyértelmű, nem könnyű (főleg wave → midi)



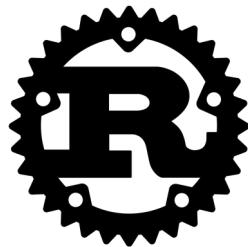
Note Block Studio

- Minecraft Note Block-ok mint minták
- Note Block Studio: szerkesztőprogram
- NBS fájlformátum
- ~Midi, de speciálisabb: egy időpillanat, konkrét minták



Projekt célja

- Bemenet: Wave fájl
- Hangminták felismerése adott időben, hangmagasságban, hangerővel
- Kimenet: Nbs fájl
- Hasonlít a zene lekottázásra, annak speciális esete
- Megvalósítás Rust-ban



Időbeosztás

- Előző félév (Önálló laboratórium):
 - Működő konverzió, két zenén tesztelve
 - Bonyolult akkordok esetén elég rossz felismerés
- Ez a félév (Szakdolgozat-készítés):
 - Gradiens a költségfüggvényhez
 - Sokkal pontosabb felismerés
 - Nagy adathalmazon tesztelés, elemzés



Videó

- Note Block hangminták
- Hullámos fájl (bemenet)
- Spektrogram (konverzióhoz fontos)
- Midi vizualizáció (motiváció része)
- Lejátszás a játékban (motiváció része)
- Eredeti Nbs fájl (elvárt kimenet)
- Felismert Nbs fájl (tényleges kimenet)



00:00:02,556 / 00:00:06,400 5.00 t/s 1.4.1

00:00:02,000 00:00:03,000 00:00:04,000 00:00:05,000 00:00:06,000 00:00:07,000

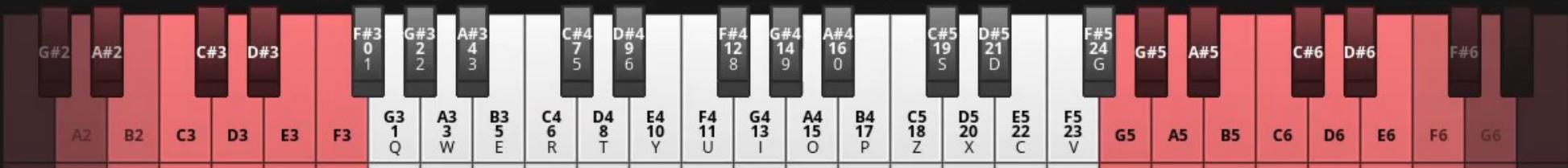
8 12 16 20 24 28 32 36

Harp Bass Snare Click

C 6 C 6 C 6 C 6

Layer 1 Layer 2 Layer 3 Layer 4 Layer 5 Layer 6 Layer 7 Layer 8 Layer 9 Layer 10 Layer 11 Layer 12 Layer 13 Layer 14

Note Block hangminták
Program: Open Note Block Studio



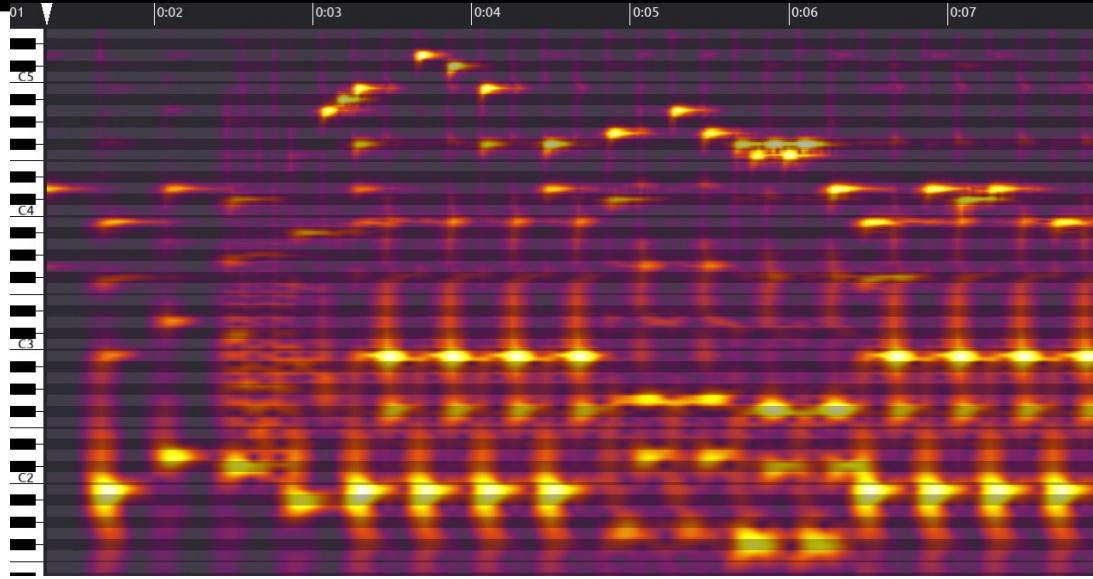
Adat tulajdonságai

- Különböző hangszerek (5..16)
- 2 oktávnyi hangmagasság (25)
- Hangerő 0.0 és 1.0 között, de lehet többszörös is
- Időzítés tps-ben (tick: legkisebb időegység), ONBS-ben oszlop

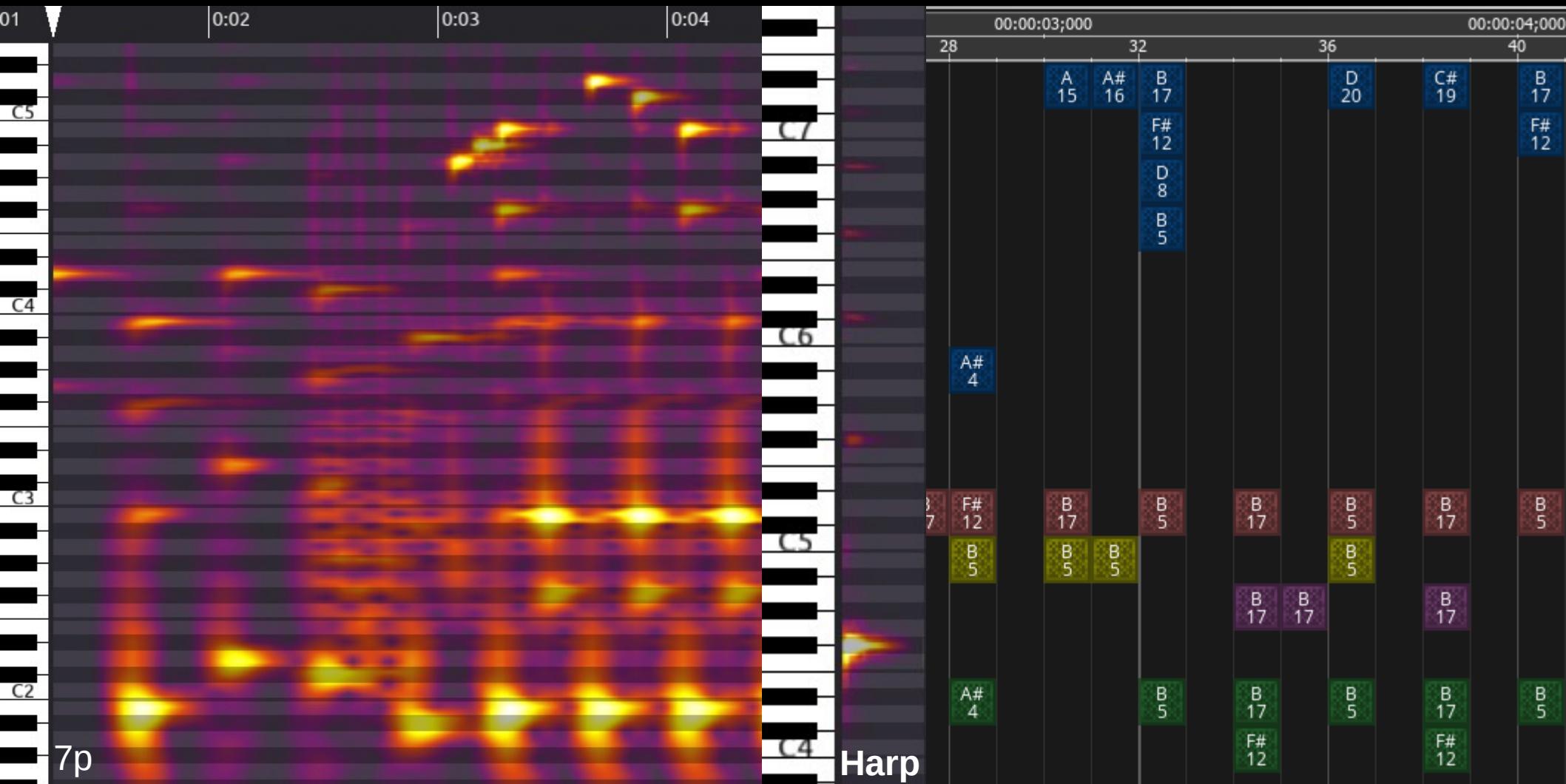


Konverzió, keresés módszere

- Spektrogram:
Fourier-transzformációk
fix időközönként
- Vizsgálandó időpillanatok
- Előfelismerés az állapottér
csökkentése érdekében
- Optimalizálás Steepest Descent módszerrel



Spektrogram (wave-ből generálható), és az eredeti NBS

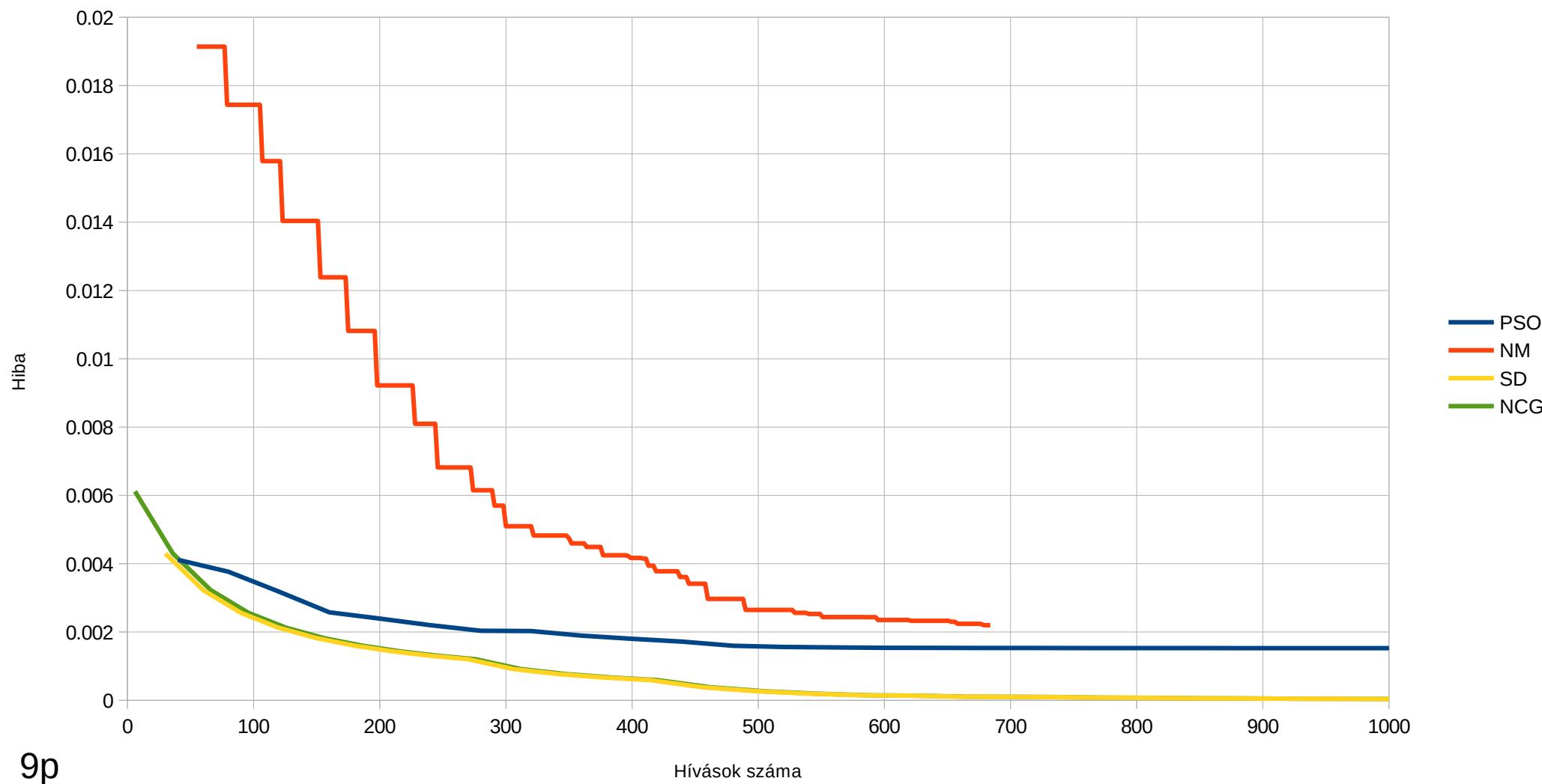


Költségfüggvény az optimalizációhoz

- Költségfüggvény:
bemenet: n db float a tippelt hangerőkkel
kimenet: egy float, a hiba / költség (kisebb = jobb)
- Számítása:
minták spektrogramjainak összeadása a tippelt hangerőn
előző hangok lecsengésének figyelembe vétele
különbség normanégyzete pixelenként
- Gradiens: nem könnyű meghatározni, de sokat pontosít



Optimalizációs algoritmusok összehasonlítása



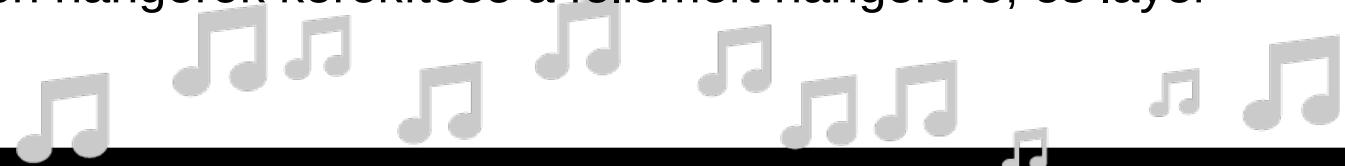
Értékelés

- nbswave kimenetéből 4 hangszert pontosan felismert
- 223 zenéből (631469 hangból) álló adathalmazon teszteltem
- többféle pontosság-mérték:
 - hangok hangerőkülönbségeinek négyzetösszege a csendéhez viszonyítva: 98.4%
 - Jól felismert hang / összes eredeti hang: 96.8%
 - Jól felismert hang / összes tippelt hang: 83.61%
 - Az utóbbi kettő a hangerőt nem veszi figyelembe



Fejlesztési lehetőségek, köszönöm a figyelmet!

- több hangszer felismerése
- tesztelés más adathalmazon
- tps felismerésének javítása, túlvezérlés-kompenzáció felismerése
- robusztusabb felismerés másfajta nbs to wave exporterek / módszerek által generált hangfájlokra
- tesztelés más resource pack-kel / soundfonttal
- futásidő optimalizálása: gyorsabb minimumkereső algoritmus, iterációszám okos megválasztása
- megpróbálni a layereket felismerni, amiken általában ugyanolyan hangerővel szerepelnek a hangok, ezen hangerők kerekítése a felismert hangerőre, és layer volume-ok használata



Kérdések

- Bíráló kérdései



Extra dia: motiváció

- Sok szerző készít ilyen zenét
 - Vizualizáció
 - Remix / átdolgozás
 - Kottakészítés
 - Lejátszani hangszeren
-
- Wave → Midi → Nbs nagyon pontatlan, közbeavatkozás kell

