

Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Automatizálási és Alkalmazott Informatikai Tanszék

Vértessaljai Bálint

NOTE BLOCK HANGMINTÁK FELISMERÉSE ZENEFÁJLBAN

Bővített összefoglaló

KONZULENS

Grad-Gyenge László György

BUDAPEST, 2024

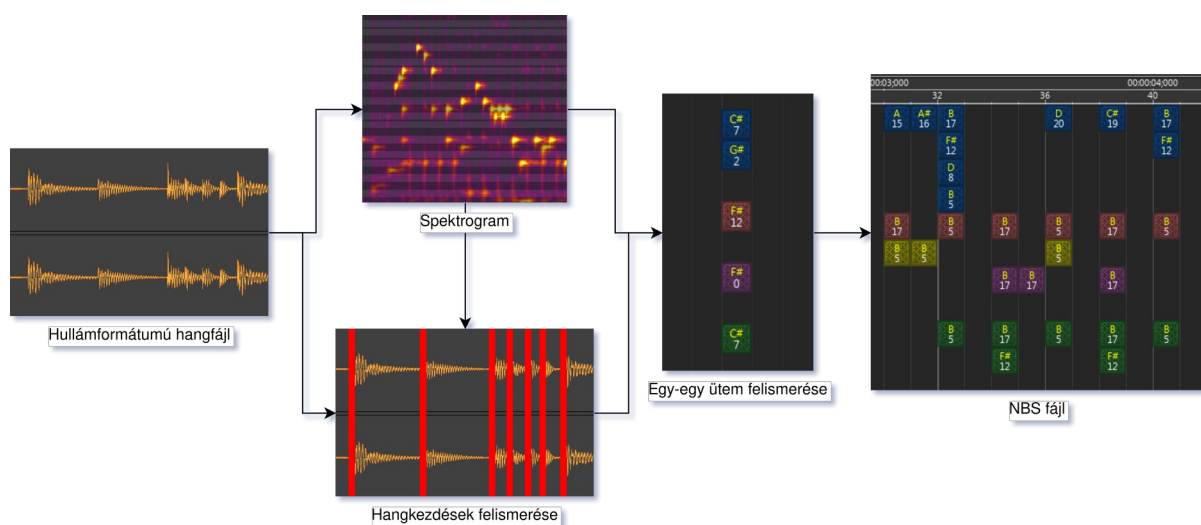


1 Bevezetés

A hullámformátumú hangfájlokból történő automatikus zenei hangvisszafejtés (Automatic Music Transcription, AMT) az egyik legösszetettebb kihívás a zenei informatika területén. A legtöbb elérhető szoftver a WAV/MP3 formátumú fájlokat MIDI fájlkká alakítja, amelyek tartalmazzák a hangok időpontját, hangmagasságát és dinamikáját. Azonban ezek az általános célú megoldások gyakran nem elég pontosak, és jelentős, elkerülhető zajt visznek az eredményekbe, különösen specifikus alkalmazások esetén. A dolgozat a problémát egy szűk területre korlátozza: a Minecraft nevű játék Note Block mintáinak automatikus felismerésére és NBS (Note Block Studio) formátumú fájlkká történő konvertálására.

A Minecraft részeként egy különleges zenei világ tárul elénk, a játékosok Note Block-ok segítségével hozhatnak létre zenéket. A Note Block Studio szoftver lehetőséget nyújt ezeknek a zenéknek a létrehozására, szerkesztésére és visszajátszására. Mivel a hangminták pontosan ismertek és rögzítettek, egy célzott algoritmus segítségével pontosabb és gyorsabb eredmény érhető el, mint egy általános AMT rendszerrel. A dolgozat célja tehát egy olyan parancssori alkalmazás fejlesztése, amely egy hullámformátumú fájlban a Note Block hangokat automatikusan felismeri, és NBS formátumba konvertálja.

2 Módszertan



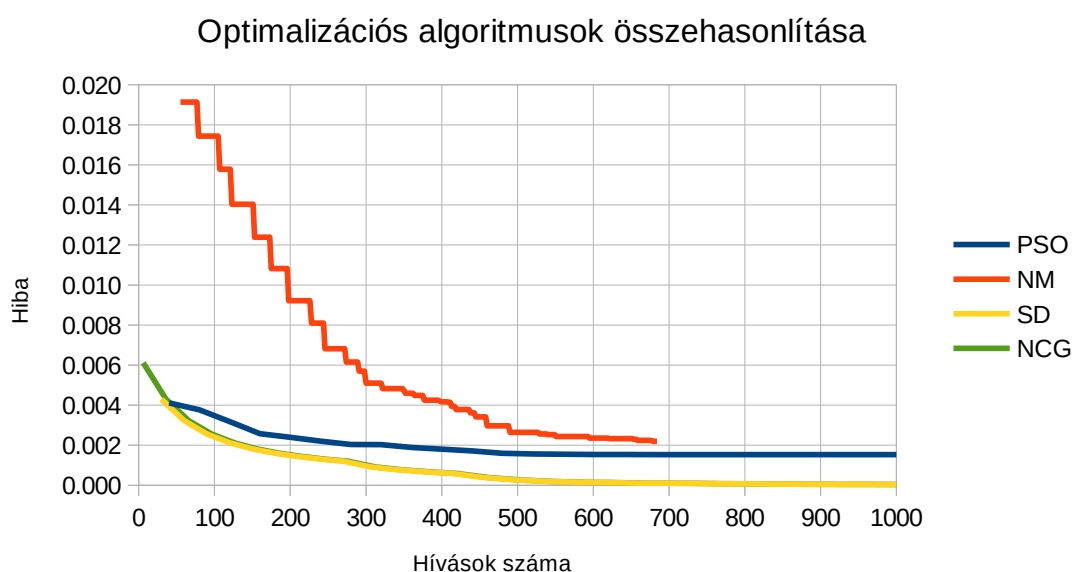
1. ábra: Áttekintés az architektúráról

A dolgozatban bemutatott megoldás (1. ábra) a spektrális analízis és az optimalizáció kombinációján alapul. A fejlesztett algoritmus időpillanatonként elemzi a bemeneti zenét. A zenefájl és a Note Block minták spektrogramjait összevetve egy speciálisan kialakított költségfüggvény segítségével keresi meg a minimumpontot. Ez határozza meg, hogy az adott időpillanatban melyik hangmintának kell megszólalnia, és milyen hangerővel, hogy az eredeti hanganyaghoz a lehető legközelebb álló eredményt kapjuk.

A megoldásban használt elemzést különböző Rust könyvtárak támogatják. A Babycat könyvtár biztosítja a hangfájlok kezelését, míg az Argmin optimalizációs könyvtár végzi a minimumkeresési folyamatot. A program fontos eleme a Note Block Studio formátum kezelésére írt könyvtár, amely lehetővé teszi az NBS fájlok hatékony generálását és szerkesztését.

3 Eredmények

A programot egy 9,5 órányi zenét tartalmazó, Note Block mintákat használó adathalmazon teszteltem. A rendszer négy különböző hangszer (hárfa, basszusgitár, pergődob és klikk) hangmintáit ismerte fel többszólamú formában. Az nbswave nevű, Note Block Studio fájlt hullámformává szintetizálni képes program kimenetén a program 1,6%-os négyzetes hangerőhibát (mean squared error) ért el a csendéhez viszonyítva, a teljes adathalmazon.



2. ábra: Optimalizációs algoritmusok összehasonlítása

Az értékelés során megvizsgáltam a különböző optimalizációs algoritmusokat, beleértve a Nelder-Mead és a Steepest Descent módszereket (2. ábra). Az összehasonlítás alapján a Steepest Descent algoritmus bizonyult jóval megbízhatóbbnak a hangminták felismerése szempontjából.

Ehhez szükséges volt a költségfüggvény gradiensének meghatározása, ezáltal váltak elérhetővé a gradiens alapú optimalizációs algoritmusok, amik közül a Steepest Descent-et, és a Nonlinear Conjugate Gradient-et elemeztem. Mindkét algoritmus hasonlóan jól teljesített, és jelentősen túlteljesítették a nem gradiens alapú algoritmusokat (Nelder-Mead, Particle Swarm Optimization).

Egy további fontos eredmény a formátumspecifikus tulajdonságok kihasználása. Az általános célú AMT rendszerek általában nem képesek figyelembe venni a Note Block minták egyedi jellemzőit, például a hangmagasságok korlátozott tartományát és a rögzített lecsengést. A dolgozatban bemutatott algoritmus viszont ezeket a tulajdonságokat maximálisan kihasználja, ami jelentősen javítja a felismerés pontosságát.

4 Következtetések

Az elért eredmények igazolják, hogy egy speciális célra tervezett algoritmus lényegesen jobb teljesítményt nyújthat, mint az általános megoldások. A dolgozat rámutat arra is, hogy a formátumspecifikus megközelítés nemcsak a Note Block zenék esetében, hanem más zenei területeken is ígéretes lehet.