

# *Szakmai gyakorlat a szakdolgozat előkészítésére*

Név: Nagy Csaba

Témavezető: Dr. Iclănzan David Andrei

Téma: Automatikus hangszerfelismerés zenefelvételek alapján

Számítástechnika III  
2024.08.29

## ***Tartalom***

|                                     |   |
|-------------------------------------|---|
| 1. Bevezető                         | 3 |
| 2. A szakmai gyakorlat menete       | 3 |
| 2.1 Fizikai alapok                  | 3 |
| 2.2 Hullámformák jellemzése         | 3 |
| 2.3 Akusztikai paraméterek          | 3 |
| 2.4 Audiojelek (analóg – digitális) | 3 |
| 2.5 Audio-jellemzők típusai         | 4 |
| 2.6 Audio-jellemzők kinyerése       | 4 |

# 1. Bevezető

Az Államvizsga-dolgozat előkészítését szolgáló szakmai gyakorlatom célja, hogy szilárd elméleti és gyakorlati alapot teremtsen a hullámok fizikai viselkedésének, a jelfeldolgozásnak és a korszerű hangfeldolgozási módszereknek a megértéséhez. Régóta foglalkoztat a zene világa, most pedig arra törekszem, hogy a hagyományos zenei képzésemet technikai és digitális nézőponttal egészítsem ki, előkészítve „Automatikus hangszerfelismerés” témájú államvizsga-dolgozatom kutatási munkáját.

## 2. A szakmai gyakorlat menete

A szakmai gyakorlatot tudatosan apró, egymásra épülő lépésekre bontottam. Először a hang fizikai alapjait tisztáztam – rezgés, frekvencia, amplitúdó –, majd ezekre építve haladtam az akusztikai mennyiségek, az analóg-digitális átalakítás, végül az audio-jellemzők elméletéig. Ezzel a „legelemibb kiindulóponttól” jutottam el odáig, hogy biztonsággal értelmezhessem a hangszerfelismeréshez szükséges fogalmakat és módszereket, megalapozva a későbbi gyakorlati kísérleteket.

### 2.1 Fizikai alapok

Kitekintés arra, hogyan keletkezik és terjed a hang, és miként alakul ki a hangérzet.

- A hang keletkezése, közegetípusok szerinti terjedés
- Rezgési alapfogalmak: frekvencia, amplitúdó, fázis
- Az emberi hallás határai (20 Hz – 20 kHz)

### 2.2 Hullámformák jellemzése

Áttekintem a hullámformák legfontosabb idő- és frekvenciasíkbeli tulajdonságait.

- Időbeli osztályozás: periodikus vs. aperiodikus
- Összetett hang és felharmonikusok
- Spektrogram-alapú vizualizációk

### 2.3 Akusztikai paraméterek

Bemutatom, hogyan vezethetők le a fizikai hangerősségi mennyiségek a halláslélektani érzetekhez.

- Hangteljesítmény, hangintenzitás, intenzitásszint
- Hallás- és fájdalomküszöb, decibel-skála
- Komplex hang, hangszín kialakulása

### 2.4 Audiojelek (analóg – digitális)

Részletezem az ADC/DAC (analóg-digitális és digitális-analóg átalakítás) lépéseit.

- PCM, mintavételezés, Nyquist-frekvencia
- Kvantálás, bitmélység, aliasing

- A hang rögzítésének és reprodukciójának folyamata

## 2.5 Audio-jellemzők típusai

Rendszerezem az audio-jellemzőket három szempont szerint.

- Absztrakció: *low*-, *mid*-, *high-level* jellemzők
- Időbeli lefedettség: pillanatnyi, szakaszos, globális
- Zenei szempont: hangszín, hangmagasság, időtartam, hangerő
- Jeltartomány: idő-, frekvencia-, idő-frekvencia (spektrogram)

## 2.6 Audio-jellemzők kinyerése

Összevetem a hagyományos gépi tanulás (Traditional Machine Learning) és a mélytanulás (Deep Learning) megközelítéseit.

- Keretezés (framing) → ablakozás (windowing) → átfedés (overlapping)
- Fourier-transzformáció, spektrális szivárgás csökkentése
- Feature-pipeline: számítás → aggregálás → értékesítés

## 3. Elméleti összegzés és tapasztalatok

A rendelkezésre álló 36 óra alatt kizárólag irodalom- és fogalomfeldolgozó munkára volt lehetőség. A fogalmi térkép elkészítése összekötötte a hangfizikai alapokat a jelfeldolgozás és a mélytanulás eszközeivel, egyértelmű kutatási irányt jelölve ki. Ezzel az elméleti áttekintő szakasszal sikerült egységes fogalmi keretet létrehozni, melyet a gyakorlati implementáció követ a nyári gyakorlat idején, a lefektetett elméleti alapokra építve.

A következő szakaszban az elméleti ismeretbővítést kódalapú, gyakorlati implementáció egészíti ki: az audio-jellemzők kinyerését és vizualizálását Python környezetben, a Librosa és TensorFlow csomagok segítségével végzem el.

**Felhasznált forrás:**

**Valerio Velardo - Audio Signal Processing for Machine Learning - The Sound of AI**  
<https://www.youtube.com/playlist?list=PL-wATfeyAMNqLee7cH3q1bh4QJFAaeNv0>

**Github:**

<https://github.com/nCsab/Instrument-Recognition>