

# Diplomamunka

2004. június 18.

Dallamkövetés és osztályozás népdalokban

Feldhoffer Gergely, 2004

## Dallamkövetés és osztályozás népdalokban

### Kivonat

A népdalok kutatása Magyarországon nagy múltra tekint vissza. Európa többi országához viszonyítva népdalismeretünk kiemelkedő. Ez nem csak annyit jelent, hogy sokkal több népdalt tartunk számon, hanem azt is, hogy sokat tudunk a népdalokról. Mára a népdalkutatás eljutott odáig, hogy a gyűjtő munka után a vizsgálat került előtérbe. A népdalok osztályozása az egyik feladat, amit a népdaltípusok magas száma megnehezít. Innen származik az ötlet, hogy a népdalfelvételek kottázása közben a népdaltípus megállapításához nyújtson segítséget egy rendszer.

Ez a munka egy kísérlet, választ keresünk arra a kérdésre, hogy alkalmas-e az itt leírt néhány módszer arra, hogy népdaltípus-megállapításra lehessen használni.

## Tartalomjegyzék

<b>I Akusztika</b>	<b>4</b>
<b>1. A jelfeldolgozás első lépései</b>	<b>4</b>
1.1. Hullámforma (időjel) . . . . .	4
1.2. A fül működése . . . . .	5
1.3. A Fourier transzformáció . . . . .	6
1.4. Ablakok . . . . .	6
1.5. Ablakfüggvények . . . . .	7
1.6. Szűrők . . . . .	8
<b>2. A pitch fogalma</b>	<b>8</b>
2.1. Felharmonikusok . . . . .	8
2.2. Zenei hangok . . . . .	9
2.3. Énekhang . . . . .	9
<b>3. A dallamkövetés feladata</b>	<b>10</b>
3.1. Követelmények . . . . .	10
3.2. Hibalehetőségek . . . . .	11
3.2.1. Oktávhiba . . . . .	11
3.2.2. Gyenge $f_0$ . . . . .	11
3.2.3. Különböző zajforrások . . . . .	11
<b>II Dallamkövetési eljárások</b>	<b>11</b>
<b>4. Harmonic Power Spectrum</b>	<b>12</b>
<b>5. Maximum Likelihood</b>	<b>13</b>
<b>III Mérés és értékelés</b>	<b>15</b>

<b>6. Dinamikus idővetemítés</b>	<b>15</b>
6.1. Távolságmátrix . . . . .	16
6.2. Az idővetemítés . . . . .	16
<b>7. Transzponáció</b>	<b>18</b>
<b>8. Kísérletek, eredmények</b>	<b>18</b>
8.1. Teszt adatbázis . . . . .	18
8.2. HPS kísérletek . . . . .	19
8.3. ML kísérletek . . . . .	20
8.4. Kombinált dallamvonal . . . . .	21
8.5. Zajtesztek . . . . .	22
8.5.1. HPS . . . . .	22
8.5.2. ML . . . . .	23
8.6. Összehasonlítás . . . . .	25
<b>9. Matlab kód</b>	<b>26</b>

## rész I

# Akusztika

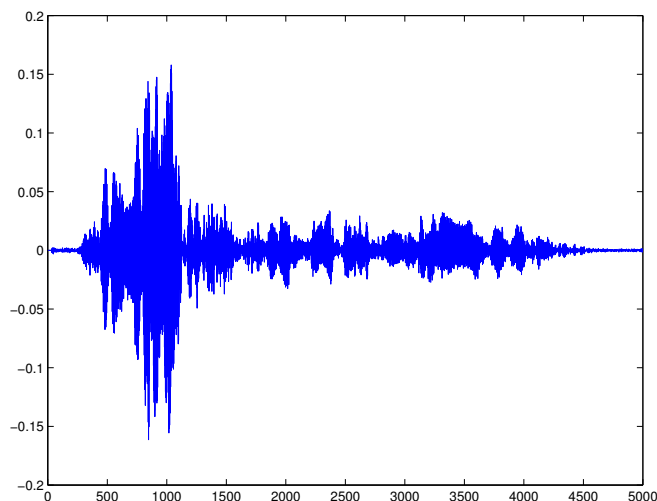
## 1. A jelfeldolgozás első lépései

### 1.1. Hullámforma (időjel)

A hangokat a fizikai hullámhoz hasonlóan, hullámalakban tároljuk. Az időjel egy pontjának jelentése az adott időben tapasztalható nyomás a mikrofon membránján. Ez tehát egy olyan jel, ami a 0, mint nyugalmi helyzet körül oszcillál. A digitalizálás azt jelenti, hogy a hullámalak minden pontját egy egész számmal közelítjük. Ezen egész szám egy adott tartományon belül marad. Ez a tartomány a digitalizálás felbontása, általában a leíráshoz használt bitekkel adjuk meg. A 16 bites hang tehát azt jelenti, hogy minden időpontban a jelet 65536 féle helyzetben tudja rögzíteni, ami a gyakorlatban elegendő.

Ezek a jelek időben ekvidisztáns felosztással keletkeznek, vagyis két jel rögzítése között az eltelt idő konstans. Az egy másodpercre eső jelmennyiséget mintavételezési gyakoriságnak, frekvenciának nevezik. Ma a CD-k a zenei felvételekhez 44100 Hz mintavételezési frekvenciát használnak. A mobiltelefonok nagyjából 8 kHz frekvenciával dolgoznak, mert a beszéd a magasabb frekvenciákban nem tartalmaz hasznos információt, nem úgy mint a zene.

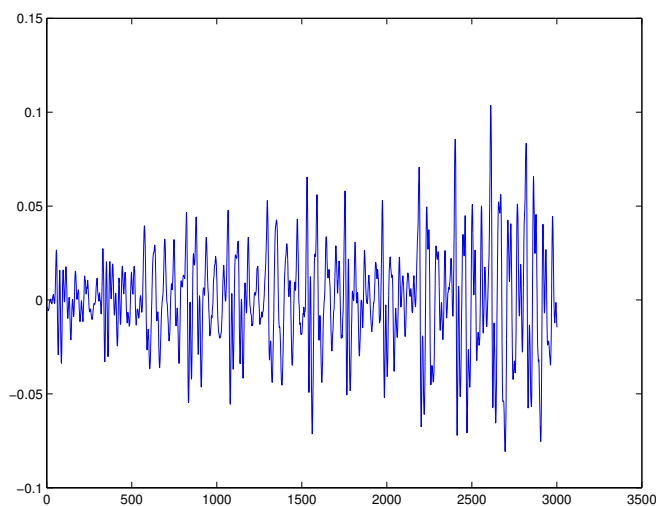
A továbbiakban a jelet mint sorozatot vagy vektort írjuk le, vagyis  $a_i$  jelentése az  $a$  jel  $i$ -edik időpillanatban felvett értéke.



1. ábra. Hullámforma

Minden ember által érdekesnek tartott hang, legyen az beszéd vagy zene, olyan jelet jelent, amiben ismétlődések fordulnak elő. Az ilyen ismétlődések nem a matematikai periodicitást jelentik, hanem ennek egy gyakorlati közelítését, miszerint a jel egy szakasza hasonlít egy korábbi szakaszára. Ezt a jelenséget kváziperiodicitásnak nevezzük. Erre egy példa lehet az énekhangban egy

tiszta szelet, ahol a jel szemre periodicitást mutat, közelről megnézve azonban eltéréseket tapasztalunk. Ezek az eltérések nem feltétlenül azt jelentik, hogy a jel az adott távolsággal öt megelőző jeltől mindig csak kicsit különbözik, ennél sokkal nagyobb is lehet az eltérés, például a digitalizálás miatt az adott görbét közelítve a magasabb frekvenciák miatt egészen eltérő fázisban vannak rögzítve. A kváziperiodicitást mégis jól jellemzi az autokorreláció, vagyis az a mérőszám, ami a jel egy adott hullámhosszhoz tartozó hibája az tökéletesen periodikus függvényétől.



2. ábra. Kváziperiodicitás

Ez egy  $\sum_{i=1}^n |a_i - a_{i+k}|$  összefüggést is jelenthet, de a hibát tetszőleges metrikával fel lehet írni.

A kváziperiodikus jelhez a fül hangmagasságot társít. Ez a hangmagasság kicsit különbözik a matematikai frekvenciától, mert a frekvencia fogalma tökéletesen periodikus függvényekről szól, azon egy bizonyos periódushossz a frekvencia. Hangoknál frekvenciatartalomról beszélünk.

## 1.2. A fül működése

A fül úgy működik, hogy egy közvetítő réteg a külső nyomásváltozást (akár a mikrofon membránjának mozgását) átadja egy csiga alakú elkeskenyedő, folyadékkal teli üregnek. Ebben az üregben a folyadék ettől mozogni kezd. Az elkeskenyedő, csigavonal alakú üreg faláról szőrszálak lógnak a folyadékba, amiknek az alján mozgásérzékelő idegsejtek találhatóak. Az idegsejtek érzékelik, hogy a szőrszál milyen erővel rezeg a folyadékban. Mivel az üreg elkeskenyedő, ezért a más-más pontokon elhelyezkedő szőrszálak más-más periodicitásra érzékenyek, mindegyik arra, ami az adott üregvastagságban állóhullámot tud kialakítani. Ez nem csak azt jelenti, hogy párhuzamosan sok adatot érzékel a fül, de azt is, hogy a leképezésnek a Fourier sorhoz igen sok köze van. Valójában az történik, hogy a jelet a szőrök mozgásával frekvenciák összességéként érzékelünk.