

Mgh-2 fémrész (fémrész)

F1 : nyelő függőleges rezgés
 F2 : -||- vízszintes - || -

képződött
 2000

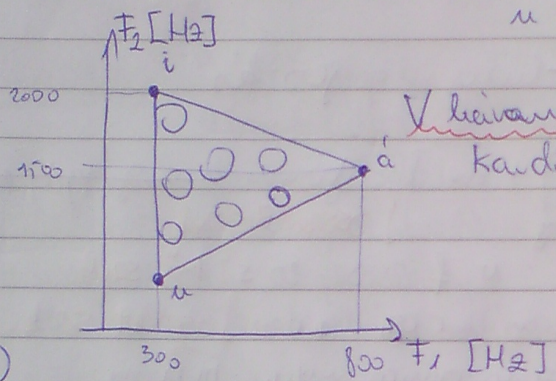
képződött
 2000
 SEMLEGES MGH
 500

- f1 hangot fémrész alacsonyabb a vízszintes.

MODEL \longleftrightarrow

$$F_n = (2n-1) \frac{c}{4L} \quad / c = 350 \frac{m}{s}$$

$L = 17 \text{ cm}$ $F_1 = 500 \text{ Hz}$
 $n = 15 \text{ cm}$ $F_2 = 1500 \text{ Hz}$
 $F_3 = 2500 \text{ Hz}$



Válasz
 Kardinális mgh-2

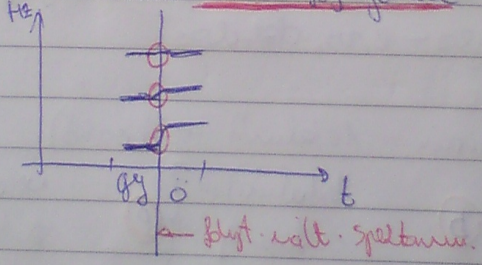
(P)

- $F_1 = 100 \text{ Hz} - 1100 \text{ Hz}$
 - $F_2 = 500 \text{ Hz} - 2400 \text{ Hz}$
 - $F_3 = 1500 \text{ Hz} - 3400 \text{ Hz}$
 - $F_4 = 2000 \text{ Hz} - 4500 \text{ Hz}$
- } ill. tudni

Abstrakció jelölés = az aktuális # hang-jánal megjelölés

1. az: tartalom. ■

$\vec{0} : f_1 = 500$
 $f_2 = 1500$
 $f_3 = 2500$



(fizikai-függ. összekap.)

hangok hálóját \Rightarrow intenzitás vektor \vec{E}

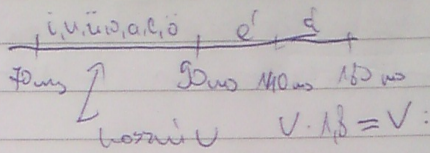
8. ca - 2009 - 10. 05

V - 3 param

form. számban

1xü szd.

form. elhely.? időszak? időszak?



A magyar nássalhangzók

•• 3 param - képzési mód, -hely, géjesség

• képzési mód (h?)

⊕ zárlathangok: z_{iv} (képzési) + z_{iv} (hely) + b, d, g, gy, ty, p, t

↳ $b, p | d, t | g, z | gy, ty$

összetett szó

⊕ nyílshangok → nyílelem: v, f, z, sz, sz, s kü szó

⊕ zánós hang → z_{iv} + nyílelem: cs, dz, dz összetett szó.

⊕ nasális hangok → z_{iv} + nyílelem: m, n, ny kü szó

⊕ egyéb → kísérítő hangok - j, l, E

↳ perzselé r.

• képzési hely: CD 251. oldalán (2. apr. for. // illék tudni)

• géjesség: zárlat - b, d, g, gy, m, n, ny, v (szó) l, j, r

tubuláris $p, t, z, ty, f, sz, s, cs$

nyílelem - z, sz, dz, dzs

•• jellemzés (3 param - hely, mód + dozbd)

(P) bilabialis

zárlat

zöngés

bilabialis

zárlat

tubuláris

velális

zárlat

zöngés

CD 138/1

• intenzitás kü
• nasális formáció

•• időtartamok: 30ms - 90ms

akumulációs sebesség = 10 hang/sec

z_{iv}

z_{iv}

zárlat

zárlat

zöngés zöngés
 b, d, g, gy, dz, dzs

nyílelem
 p, t, z, ty, v, cs

Hangszerszámok: CVCVVCVC

jellemzés: - frekvencia - különböző hangok, nem illeszkedő hangok

- intenzitás $\sqrt{2}$

- 1+2. időszoros

jobb
CVCVVCVC
illegális

j, gy, ty, sz, m, n - stabil felh. szó

magasabb intenzitás a V-t.

$\sim V$ (de cannot not 1 vs)

Voice Onset Time (VOT) : zöngellen kezdés + zöngés megkezdésig hang

↳ időmenny.: (ms)

zöngelpannas és a zöngés kezdés között mennyi idő telik el.

p: 10-12 ms b: 30-40 ms

t: 20-75 ms k: 40-50 ms

zöngés ~~teljes~~ hely \rightarrow ad utat

Prosódia : \hat{a} dallam (hanglejtés) $\rightarrow F_0$ (Hz)

\hat{a} hangszínt $\leftarrow \begin{matrix} F_0 \text{ (Hz)} \\ I \text{ (dB)} \end{matrix} \rightarrow$

- ritmus - intenz. seb. váltás, szünetek

- (hangszínzet) - spektrum

Szöveg-dallam kapcsolata:

• megjelölés kezdés: mennyit nem függ a dallam

• eldöntendő -|| : — | | — a dallam

• kijelölés módok: függőség \Rightarrow kell szövegalkalmazás

Hangszínt : (magyar kötött hangszínt: 1. mérték kivételével)

↑ nyomaték $\rightarrow F_0$ elem

$\rightarrow I_0$ intenzitás

időtartam (nyújtás)

(240. oldal)

Beszélés négy !

2. gyakorlat - 2009.10.12 : Beszélkés

beszél = jel + szünet + jel + ...

↑ információs

↑ pragmatikus
feldolgozás

kommunikáció \rightarrow DIALOGUS

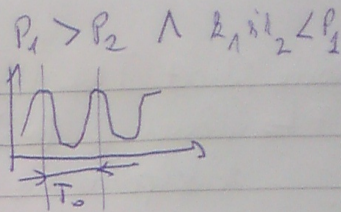
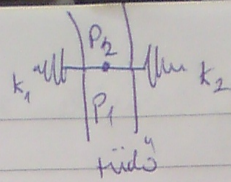
szp-vel fontos !

• beszélkéses köz. eszk.: agy, onírány, májírás, nyelv, yllcsap, hangszólag, légszö

• ha nincs jel: szünet : beszélés részekben nincs jel

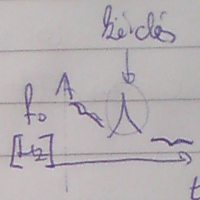
nincs felírás: hang A, részekben nincs jel

- Hangszalagok **NEM** rezegnek
 Isalása...: ill. tudni (CD: 15.2)



- sége szinten megjelenő paraméterek. $n = 200 \text{ Hz}$

a) $1/T_0 = f_0 (F_0) \rightarrow$ faji v. nö $f = 100 \text{ Hz}$
 $g_j = 3-500 \text{ Hz}$



b.) hol zöngés / zöngétlen hangok

c.) dallam ; .?

d.) Szünet (hangon belüli + mandatoru belüli)

- hol lehet még tubulus áramlás előáll: ahol szűkületet lehet léteztetni.

• h' allás: ut. utat. szűkület, gőzök sülken képződik

- ut. utat. szűk. tubulus áramlás \rightarrow szűkület - változó hely
 tubulus áramlás intenzitása: frekvencia gőzök (s, sz)

$f_{max} = 6-10000 \text{ Hz} // C, Sz // Z, Sz //$

300-3400 Hz \rightarrow telefonátvitel (adattovábbítás)

- NEMA FAZIS : zöngés } véna fázis + zöngés felismerés (45 ms / ms)
 zöngés } zöngés h. } szűk
 (P, F, Z, ty, C, Cs) "partikularitása"

- folytató zöngé : zöngés páros (ill. tudni)
 zöngés } - intenzitása kicsi, formájában minirek.
 zöngés }
 (mit a formájában? ...)

- Gőzszó isz és a működési isz. **FÜGGETLEN**
 Isalásait: gőz jel a működési jel szüneteltetés / formás szüneteltetés

- megismerés, v. közelebbi munka meg.

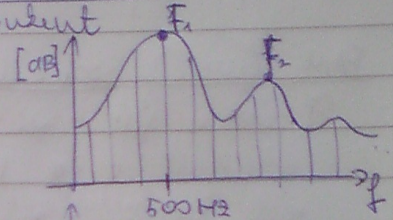
* jellemző: alapfok. pálya/sec

Beszédszintézis:

1) Gépi: formális szintézis - egyesítés + formáisszerűsítés = hang

Adatok: besz.

időrités: 5-10 ms-onként
 jelenergia
 formális - szintézis
 -||- - sínszerűség



Felhangszerkezet NINCS 85%

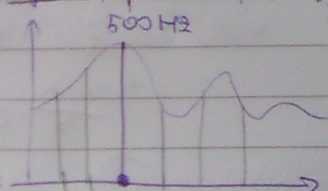
a formáisszerűsítéshez

felhang: $k \cdot f_0$, $k \in \mathbb{Z}^+$

VONALAS SPEKTRUM!

$f_0 = 100 \text{ Hz}$

$f_0 = 180 \text{ Hz}$



2) Euleri beszédből:

Adatok: hullámfunkció értékek

3) Kópusz alapszintézis:

adatok: szavak - hangadatok

CD: cheap compact disc
 -> formális szintézis
 -> szavak: hangadatok
 -> szavak: hangadatok
 CD-R: 2009

BESZÉDKÖDOLÁS

8. eq

1) Bevezetés:

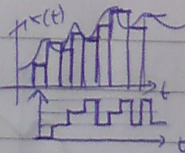
Beszédjel: $R \rightarrow R$ // folytonos ÉT ÉK

Digitális ábrák

kezelés

idődiszkrétiz.

Műveletsorozat: mintavételezés
 kvantálás
 kódolás

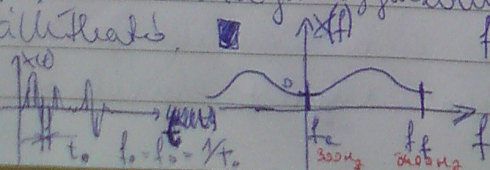


Szempontok: jel hűsége -> mérete
 sebesség
 komplexitás

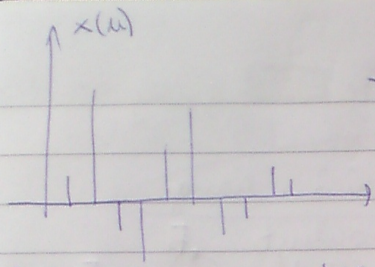
egyszerű
 optimalizálás

Shannon-tétel (Shannon-kódolás): B sávú korlátozott jel

$B = f_0 \geq 2B$ sűrűségű egyenlően mértékű egységesen elosztott -
 állítható.



$f_H - f_L$: sávszélesség = B_1
 $B_2 = f_H$: legnagyobb frekvenciájú komponens

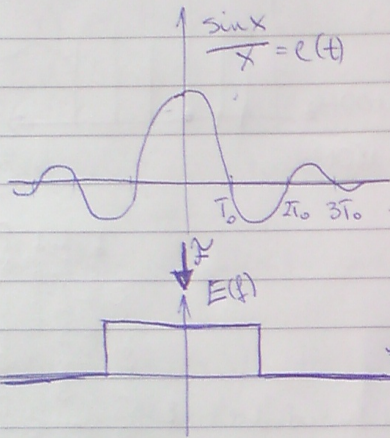


→ Hoogman
allitkelt
vissza?

allit

PAM tipusi visszaallitás (Pulse audio modulation)

B sávú jelből jel mintából a PAM sávú visszaallitással, hely-
reallitással a jel, ha $f_c > 2B$ $\Rightarrow H(f) = M(f)G(f) - B \dots + B$
tant. sávú \Rightarrow a f_c -t B -szel közelebbre \Rightarrow

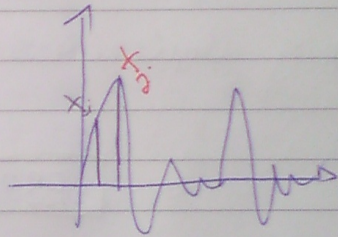


Egyenlítő minták

$\frac{\sin x}{x}$ interpolálás

$2(\frac{\sin x}{x})$ összerajz a helyreallitással

helyreallit.: $-\infty$ -ben is van sok
 \rightarrow $\frac{\sin x}{x}$ sávú a minták
sávú jelét is visszahatározza

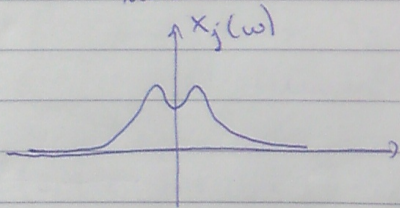


Nyquist-ekvivalens:

$$\sum_{k=-\infty}^{\infty} X_k(\omega - k\omega_0) = \sum_{l=-\infty}^{\infty} X_j(\omega - l\omega_0)$$

$\frac{2\pi}{T_0}$

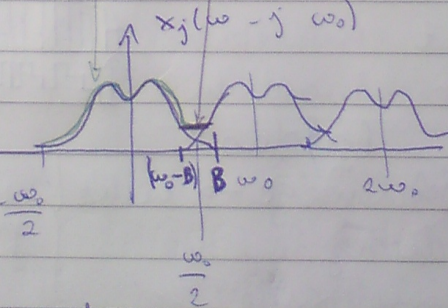
$$\sum_{n=-\infty}^{\infty} X^2(nT_0) < \infty$$



Nyquist-ekvivalens

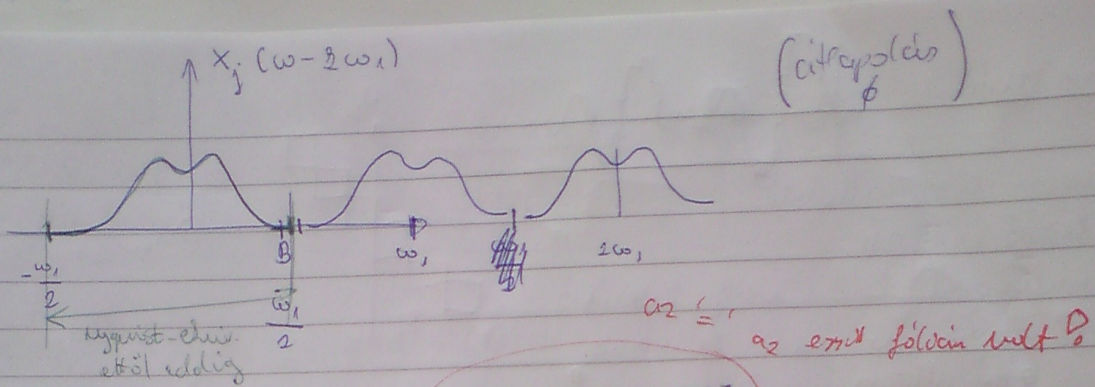
átlapátolás

hálósávú
jelnek



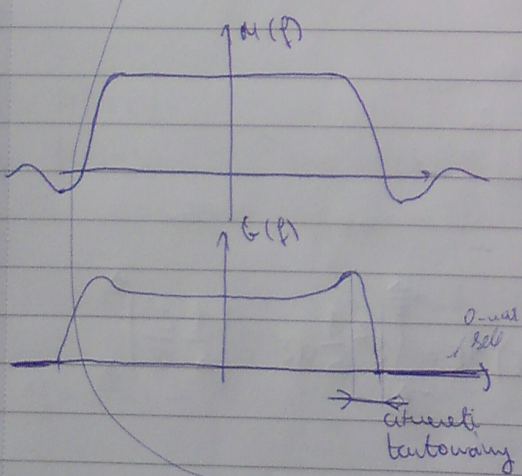
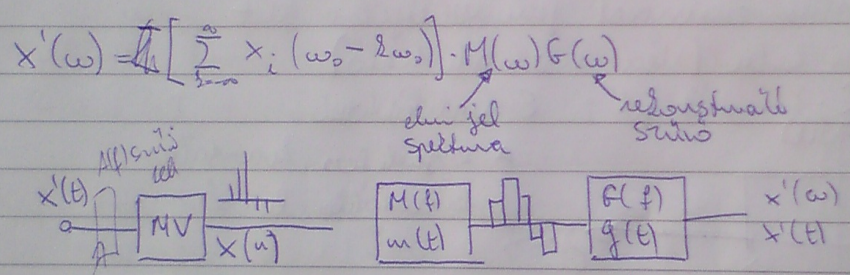
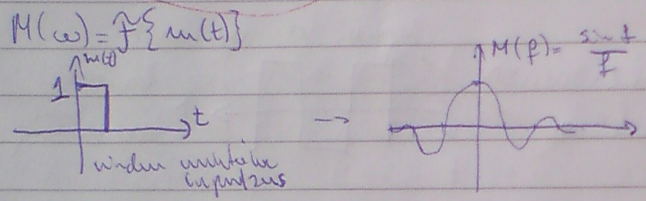
- hálósávú spektrum $-\frac{f_0}{2} \leq f \leq \frac{f_0}{2}$ tartományát nevezik
Nyquist-ekvivalensnek

- ha a jel B sávú hálósávú és a mintavételi felvétel
nagyobb $2B$ -nél, akkor spektrumának Nyquist-ekvivalense
összege. ■



Üssraal (toht) paan jelt: $X(\omega) = T_0 \left[\sum_{n=-\infty}^{\infty} x(nT_0) e^{j\omega n T_0} \right] M(\omega) G(\omega) =$

$= \left[\sum_{i=-\infty}^{\infty} X_i(\omega - 2\omega_0) \right] M(\omega) G(\omega)$

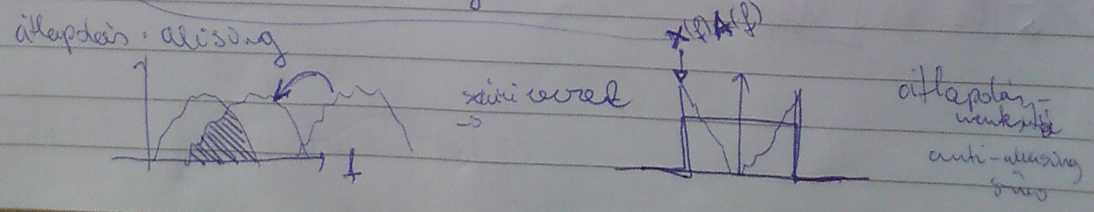


$M(f) \cdot G(f) = 1$ - see keel lüü, ahel ar eeldati jelt spektrum van

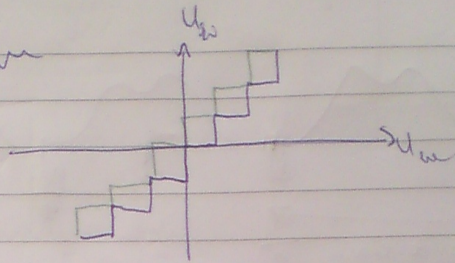
↓

de et lahvaret spektrumid ocl "leigum"

6800 → 8000 Hz a unntakele jelt.



Kvantálás.



nem visszacélítással
 lépés mérték
 nullszintet több nagy lépéssel
 lineáris \approx nem lineáris

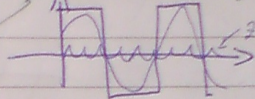
- ⊕ kis mérték
- ⊖ nagyobb mérték: kis mért. nagy korlátok

Sec: kvantálás

a) lineáris kv.

- Amplitúdó érték diszkrétizálása
 Reeves - 1938 : Bell Labs

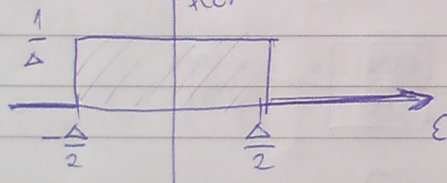
kvant. jel reprezentáció



zaj: ha a kv. szint közel van az 0-hoz

- Véges mértékű minibólummal jellemzhető a jel
 Pr.: kvantálási hiba (nem visszacélítással)

$\hat{x} = x + \varepsilon$ \hat{x} : kvantált jel; ε : valahány



Δ : egyenlő lépés kvantálás
 kvantálási lépés

ε_i és ε_j : időben más ut. hív h. való.

∫ f(\varepsilon) d\varepsilon
 0-valószínűségi sűrűség

~~Zaj k~~

Kvantálási zaj teljes.

∫ d\varepsilon
 ∫ \varepsilon^2

$$= \frac{1}{\Delta} \left[\frac{\Delta^3}{8 \cdot 3} + \frac{\Delta^3}{8 \cdot 3} \right] = \frac{\Delta^2}{12}$$

$$M(\varepsilon^2) = \int_{-\infty}^{\infty} \varepsilon^2 f_{\varepsilon}(\varepsilon) d\varepsilon = \int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} \varepsilon^2 \frac{1}{\Delta} d\varepsilon = \left[\frac{1}{\Delta} \frac{\varepsilon^3}{3} \right]_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} =$$

Amplitúdóján min.: $P_s = \frac{C^2}{2}$

Jel-zaj viszony: $NSR = \frac{P_s}{P_E} = 6 \frac{C^2}{\Delta^2} = \frac{3}{2} N^2 = \frac{3}{2} 2^{2n}$

N db kvantálási szint = $2C = N\Delta$

$N = 2^n$

$SNR_{[dB]} = 10 \lg SNR = 17.4 + 6.02 \cdot n [dB]$

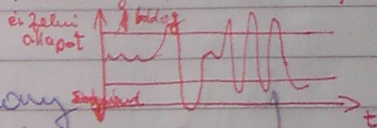
↳ kúles kvantálás + megfelelő bitszám esetén (sinusjelre)

PL $n=8 \Rightarrow SNR = 49.7 dB$

$n=16 \Rightarrow SNR = 91.7 dB$

Összabály: dB-ben vett jel-zaj viszony: $SNR \approx 6 \cdot n [dB]$

Gordos I. tv. (miniót jelölt ↔ mikroelő)



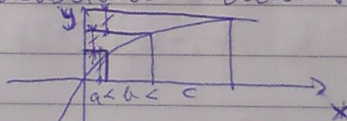
b) logaritmikus kvantálás:

- telefonos átvitelhez 60-70 dB jel-zaj viszony
- kúles kvantálásnál 12 bit kellene $12 \text{ bit} \cdot f [8 \text{ kHz}] \approx 96 \text{ kbit/s}$
- Fül nagy amplitúdóig érzékeny a hibákra
- $\frac{\Delta^x}{x} = c$: konstans
- nem kell lineárisan közeledni

Gordos II. tv. Élelet előn 2 digitális útvonal:
1) PCM
2) Számítás

Kapacitás növelés: pszichodisztribúció: STEVENS (M.I.T)

↳ jel érzékenysége: $\frac{\Delta y}{\Delta x} = C_1$



$\Delta y = C_2$

$f(x) = y$

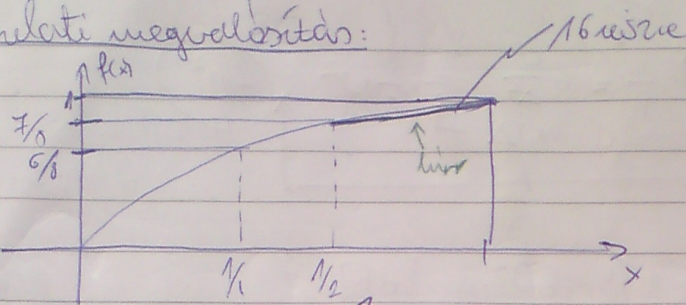
$\frac{\Delta y}{\Delta x} x = C_3 \quad \frac{\Delta y}{\Delta x} = \frac{C_3}{x}$

$f(x) = \ln x + C_4 \Rightarrow$ 1950 évi szabvány (CCITT)

$y = f(x) = \begin{cases} \frac{1 + \ln A \cdot |x|}{1 + \ln A} & \text{ha } |x| > \frac{1}{A} \\ \frac{A|x|}{1 + \ln A} \cdot \frac{x}{|x|} & \text{ha } |x| \leq \frac{1}{A} \end{cases}$ $A = 87.6$ EU A-law

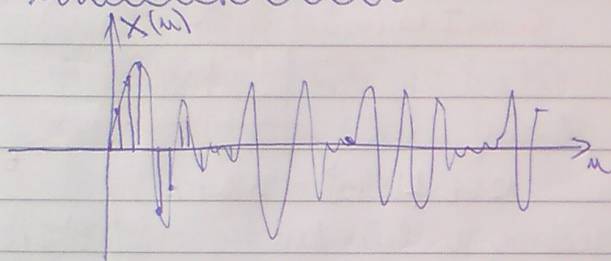
$y = f(x) = \begin{cases} \frac{\ln(1 + \mu|x|)}{\ln(1 + \mu)} \cdot \frac{x}{|x|} & \text{ha } |x| > \frac{1}{\mu} \\ \frac{x}{|x|} & \text{ha } |x| \leq \frac{1}{\mu} \end{cases}$ $\mu = 255$ USA, Japan μ -Law

Gyakorlati megvalósítás:



12 bit len kw → 96 kbit/s } MOS (Mean Opinion Score)
 8 bit log kw → 64 kbit/s } h12 (1-5) → meggyezik
 ...
 5-10 kbit/s (GSM)

C lineáris predikció:



$$\hat{x}(n) = \sum_{i=1}^P d_i x(n-i)$$

d_i : lin predikciós e.-h. ← linear prediction coefficient
 P: predikciós felszám

LPC: linear predictive Coding

$e(n) = x(n) - \hat{x}(n)$: predikciós hibája

Ha $x(n) = \hat{x}(n)$; $n_0 \leq n \leq n_1$, rabszon
 és $(n_1 - n_0) \gg 2P$, akkor LPC kinyelgkémelő.

$e(n)$ legyen kicsi! **Hogyan?**

• $E_1 = \sum_{n=n_0}^{n_1} e^2(n) = \sum_{n=n_0}^{n_1} \left[x(n) - \sum_{i=1}^P d_i x(n-i) \right]^2 \Rightarrow \min$
 determinisztikus
 minimális ~~hibája~~ ~~jelé~~

• $E_2 = \frac{E_1}{n_1 - n_0} \Rightarrow \min$ det.

• $E_3 = M \{ e^2(n) \} \Rightarrow \min$ ~~det~~

• lehetőleg más lehetőségek is, de ált. bonyi megoldások.

Adott $x(0), x(1), \dots, x(n-1)$ valamilyen jelvélet
 E_1 és E_2 hibákért tekintve min.

$$E = \sum_{n=n_0}^{n_1} \left[x(n) - \sum_{k=1}^m d_k x(n-k) \right]^2 = \sum_{n=n_0}^{n_1} \left[\sum_{i=0}^p d_i x(n-i) \right] \left[\sum_{j=0}^p d_j x(n-j) \right] =$$

$d_0 = 1$

$$\sum_{i=0}^p \sum_{j=0}^p d_i \left[\sum_{n=n_0}^{n_1} x(n-i) x(n-j) \right] d_j = \textcircled{*}$$

$$\Phi(i, j) := \sum_{n=n_0}^{n_1} x(n-i) x(n-j) = \Phi(j, i)$$

$$\textcircled{*} = \sum_{i=0}^p \sum_{j=0}^p d_i \phi(i, j) d_j = d_m^2 \phi(m, m) + \sum_{\substack{i=0 \\ i \neq m}}^p d_i \phi(i, m) d_m + \sum_{\substack{j=0 \\ j \neq m}}^p d_m \phi(m, j) d_j +$$

máshová tartozó

$$+ \sum_{\substack{j=0 \\ j \neq m}}^p \sum_{\substack{i=0 \\ i \neq m}}^p d_i \phi(i, j) d_j$$

$$\frac{\partial E}{\partial d_m} = 2 d_m \phi(m, m) + \sum_{\substack{i=0 \\ i \neq m}}^p d_i \phi(i, m) + \sum_{\substack{j=0 \\ j \neq m}}^p d_j \phi(m, j) + 0 =$$

$$= 2 \sum_{i=0}^p d_i \phi(i, m)$$

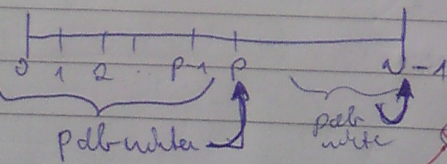
$$\sum_{i=1}^p d_i \phi(i, \xi) = \Psi(\xi) \quad \xi = 1, \dots, p$$

ahol $\Psi(\xi) = \phi(0, \xi)$
 vagy $\underline{\underline{\Phi d = \Psi}}$

$$\underline{\underline{\Phi d = \Psi}}$$

LP változatos:

• Kovariancia mátrix:



$$\phi(i, j) = \sum_{n=n_0}^{n_1} x(n-i) x(n-j)$$

gondolj kovariancia mátrixra
 diszkrét

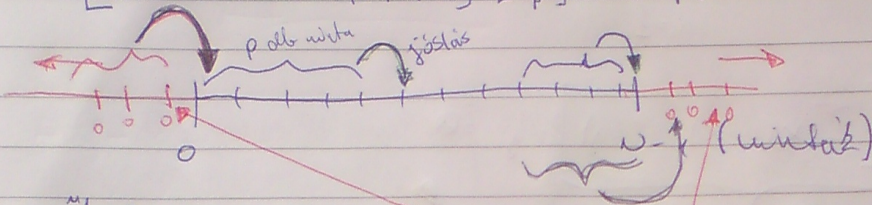
$n_0 = p \quad n_1 = N - 1$

gyakorlat kihasznál

10. ea - 2009. 09. 30

$$\underline{\Phi} \underline{\alpha} = \underline{\varphi}$$

$$\begin{bmatrix} \phi[1,1] & \phi[1,2] & & \\ & \phi[2,2] & & \\ & & \ddots & \\ & & & \phi[p,p] \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix}, \quad \varphi(i) = \underline{\Phi}(0, i)$$



$$\sum_{n=n_0}^{n_1} e^{2} = \dots$$

$$n_0 = p \quad n_1 = N-1$$

$$\underline{\Phi}(i, j) = \sum_{n=n_0}^{n_1} x(n-i) \cdot x(n-j) \rightarrow \text{egyszerűsített képlet}$$

M0 van, ha $n_0 = -\infty$ $n_1 = +\infty$ (a határolt/ kinyújtott hiba)

Megoldható, ha $N \gg p \rightarrow$ (a hiba ~~hagyható~~)

$$R(z) = \sum_{n=-\infty}^{\infty} x(n-i) x(n-j) = \sum_{n=-\infty}^{\infty} x(n) x(n-z) \quad z := |i-j|$$

↳ bevezetve $\underline{\Phi}(i, j)$ helyett $R(z)$ szerepel

$$\begin{bmatrix} R(0) & R(1) & \dots & R(p-1) \\ R(1) & R(0) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ R(p-1) & \dots & \dots & R(0) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix} \rightarrow \text{Lin. pred. alapozott autokor. módszer esetén}$$

$$\underline{R} \underline{\alpha} = \underline{v}$$

Toeplitz mátrix

Egyéb módok szerint: szám

• Kov. módszer: $\frac{p^2}{2} + \frac{p}{2} + p = \frac{p^2}{2} + \frac{3}{2}p$

• autokor. módszer: $p+1$

(p) $p=10: 65 > M$

Megoldások:

- zov. módszer ← Cholesky-f. dekompozíció
- autózov. -"- ← Levinson-Durbin rekurziós
- Parcor módszer

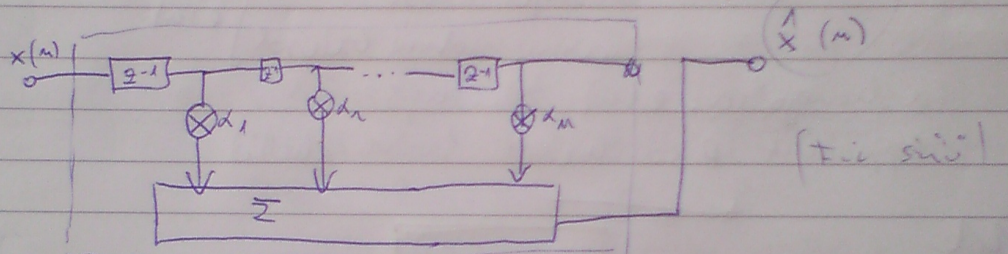
Legyenek modellek: Lu. ped. bz. mod.

Beadási modell: $x(n)$ $\hat{x}(n)$

$$\hat{x}(n) = \sum_{i=1}^t d_i x(n-i)$$

$$\hat{x}(n) = f(n) * x(n)$$

$$\hat{X}(z) = F(z) \cdot X(z) \quad F(z) = \sum_{i=1}^m d_i z^{-i}$$

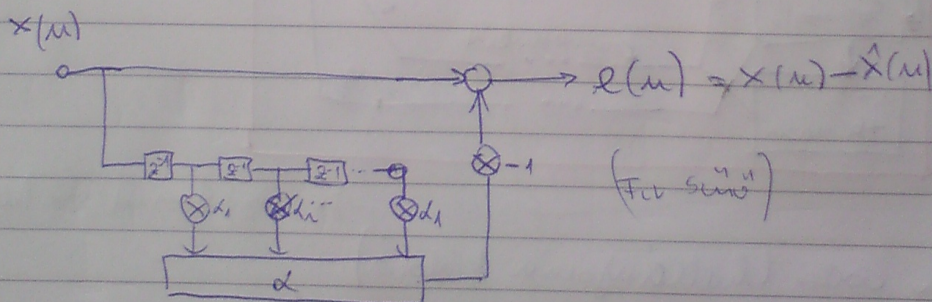


Analízis modell

$$x(n) \quad \hat{e}(n) = x(n) - \hat{x}(n)$$

$$E(z) = 1 - F(z) X(z) = A(z) X(z)$$

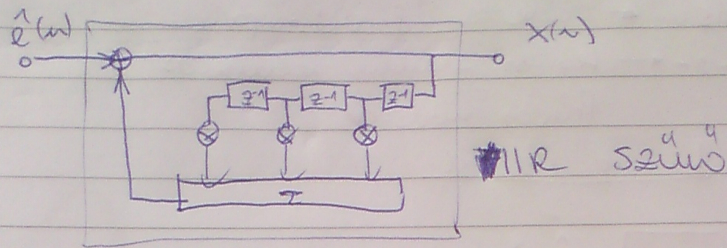
$$A(z) = 1 - F(z)$$



Szintézis modell: $x(n) = \hat{x}(n) + \hat{e}(n)$

$$e(n) \quad x(n) \quad \left. \begin{aligned} X(z) &= S(z) E(z) \\ E(z) &= A(z) X(z) \end{aligned} \right\} S(z) = \frac{1}{A(z)}$$

$$S(z) = \frac{1}{1 - \sum_{i=1}^p d_i z^{-i}}$$

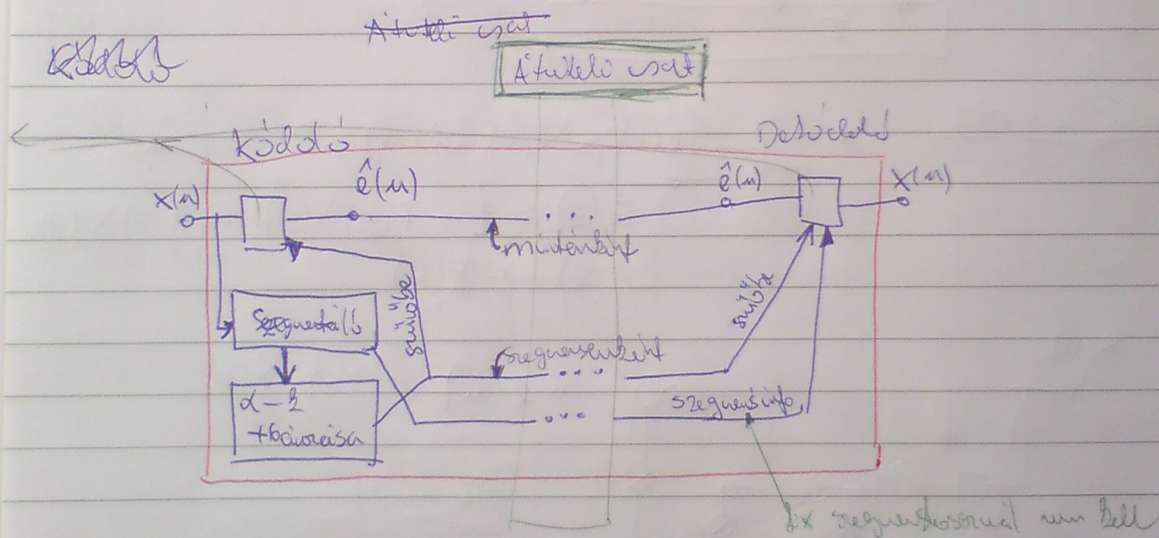


All-pole modell // $S(z)$ csak pólyákat tartalmaz

Beszédátvitel lineáris predikcióval:

- szegmmentális: 10-20 ms
- főszólam (v. hatékony szólam): 10-12 zöngés hangzóla hibajel kódi maximummal bírhat } $e(n)$

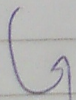
úshangzóla: $e(n)$ jelkészszerű egyébként p. üveléssel kódi lehet

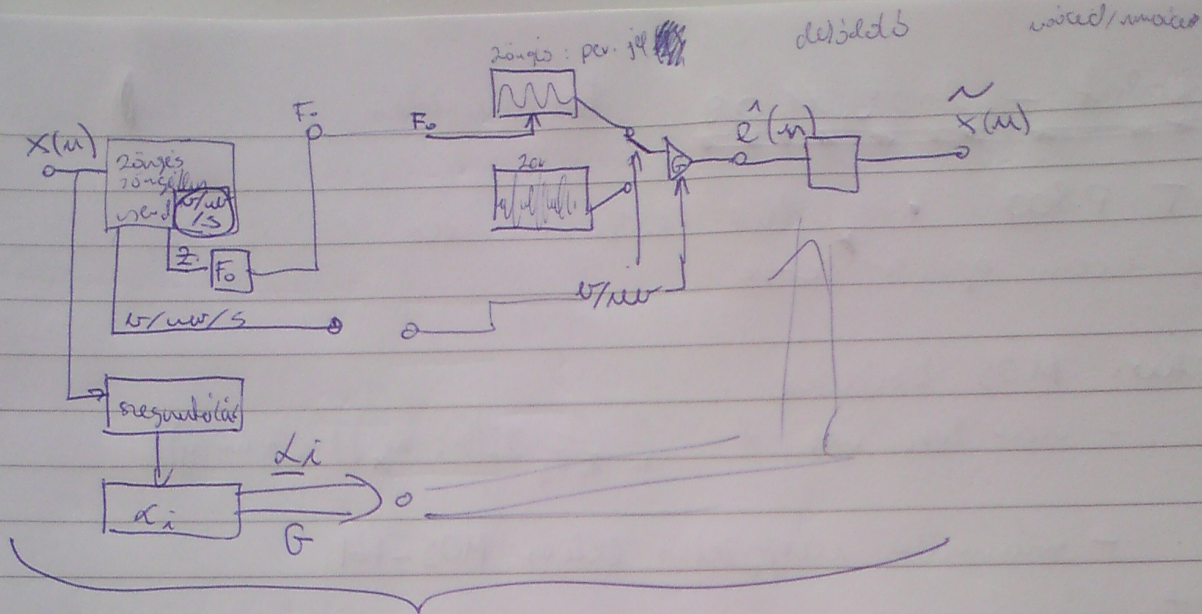


Waveform coder (hullámforma kódoló)

újság: $e(n)$ kódi \rightarrow kódi kódi ábrázolható

Hibajelen spóroláshoz: forrás kódolás (source coder)





LPC-10 : ~ 27 cível eredő: 2,4 bit/s (200 és 2000)

e(n) modellezése :
 • Egyszerű : LPC-10
 • Multiplex : MPE [-LP] alsó és felső
 • Vékony kvantálás : CELP
 ↳ változik a minta mélyebb a legközelebbi ponttól ismét át

- GSM-RPE-LTP
- 3G : AMR : adaptive multirate

ÉRZETI KÓDLÁS (perceptual coding)

Részrész kódolás (subband coding)

gyakorlat - 2009. 11. 09

Beszámoló

- a) ITU-T P.800 : • nyelvi beszámoló
- b) MOS : mean opinion score

Objektív MOS becslés

- max. szór. nos - d: $\frac{1}{n} \frac{1}{\sigma_{\text{obj}}} \sum (x_i - \mu_x) (y_i - \mu_y)$

- minimális negyzetes eltérés MOS-tól

PSQM+

~~Zelődés kiegészítés~~

Beszámoló: 2009. 11. 16

CD 299. 5.

III Beszédválasztó rendszer

- kötött séta

írásfelismerés

Modellzés: kezdő
lény

elér. + beosztás

ultrás (2-ös) ← lekijezs
klondó

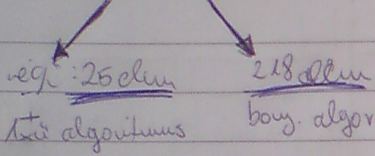
egy
egykor
← eljött

→ 218 elem

típusú

2+m írás. kell megval.

Rendszermodulok:



- milyen elemekből áll felismerés?

- működését...

10.652.822

FOLYTATÁSRA bizt.

kötött séta rendszer előnye/hátránya

⊕ 1-ös séta

⊕ 1-ös alg.

⊕ jó hangminőség is lehet

⊖: kötött témakör

⊖: nehéz hirtelen új írást

⊖: sok változás miatt ezeken nehéz

megvalósítani (sok munka/hosszútáv)

kezdő, ködös megval.

- Szövegfelismerés (gpi:) //Text-to-speech//

- köv.: - kötetlen szövegfelismerés: nyitott dolog (konv. szöveg)
 - nép hangon: kezdő, szöveg...
 - több hangon ←

• megoldási formák: - aditáláció → mechanikus: 1791^{KF}

↳ elektronikus: a model: 1939: VODER

↳ diff. mech. JAPAN

↳ software aditálációs szintetizátor JÓVÓ

2009

TTS

• Akusztikai produkció

↳ kvantifikálás: adatok; LPC, HMM

↳ hullámforma kódolás: elemösszevétel

↳ kisebb, hangjelöltek (2-3)

↳ nagyobb; 1520, 10000 Hz

szegmentális - szupraszegmentális

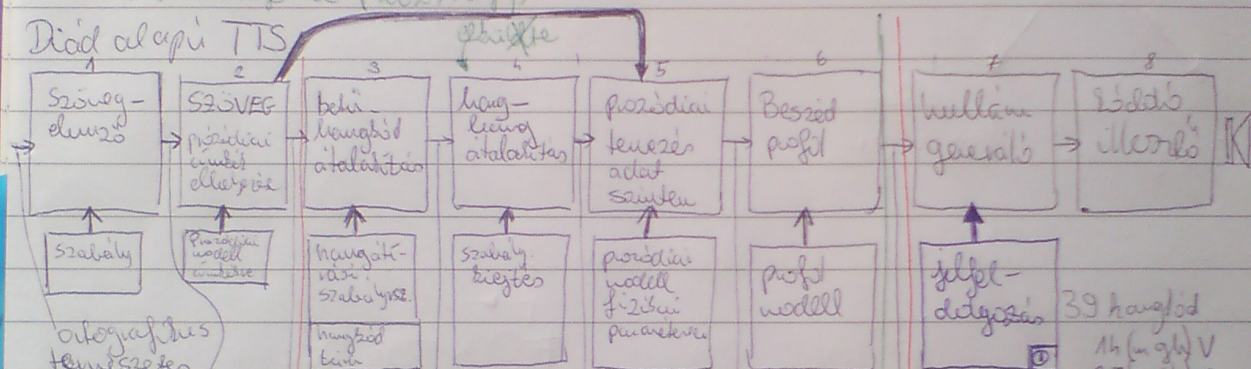
a.) együtt kezeljük - 1 lépés a szintézis

b.) külön - " - 2 - " - " - "

Szabály - gépi tanulás → HMM, ... / jellemzők

↳ hullámforma (kis/nagy)

Díád alapú TTS



ortografikus természetű szöveg.

fonetikus szöveg: v. helyesírás

fonetikus szöveg pozícionálási címekkel

hangjelöltek (A-E)

helyesírás hangjelöltes szöveg

hangjelöltek + fiz. paraméterek

alkotott beszéd paraméterekkel

hangjelöltes szegmentális szupraszegmentális

39 hangjelöltes
1h (magh) V
25 (msh) C

FIZIKAI + VAL

SZÖVEG ELŐKÉSZÍTÉS

HANGSZINTŰ ELŐKÉSZÍTÉS

ADATSZINTŰ ELŐKÉSZÍTÉS

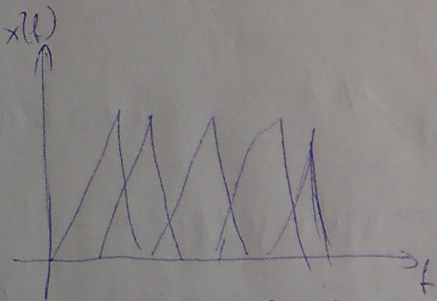
szintézisbe való adat bevezetés
← ADAT MÁTRIX

Hanguly: F_0 elem
 (szimmetria) | intenzitás
 időtartam (nyújtás)

Beszorítás vége

2009. 11. 02.

18025



Motion Picture Expert Group [ISO/IEC JTC1/SC23/W6M]

MPEG 1 1988 L-1, L-2, L-3
 Layer-3, MP3 → Fraunhofer

MPEG 2 1994

MPEG 4 1993-97

MPEG 7 1997 Jantalan keres www.windroom.hu

MPEG 11 2000-

Standard típus	ITU szab.	Adatsebesség (kb/s)	MOS
PCM	G711	64	4,3
ADPCM	G726	16, 24, 32, 40	4,1
Embedded —	G727	—	
Low-delay CELP	G728	16	4,1
ACELP	G729	8	4,0
ACELP+MP#-LPC	G729.1	53/63	
RPE-LTP	GSM	10	3,5
	GSM-EFR		3,8

LPC-10	GSM-E - half note	5	3,3
		2,4	2,6

Arcefoloklar:

- Mintavétel
- Időskis jelátvitel feltétele
- Előantálak Lin-log
- Előantálak raj
- Lineáris predikció
- ⋮

III. Beszésváltozó rendszerek

Feladat: Gépi beszédreállítás

- Brúni máci?
 - Magnetofon
 - Diktafon
- } Későbbes verziók

→ Többes verzió, többes nyelvű többes nyelvű beszédreállítás gépi megvalósítása.

- irakítás
- tematika

Spektrális

- hang: karakter
- nyelv
- tematika
- HW + SW közeg
- idő - válság

Eddig (történelmi látleírás)

- Szempelen János beszédgépe
- Bánó Miklós gépe (1916)
- Dudley → VODER (1939) → Bell-labs
- Hungarovox

- Értékelési rendszerek

Egy üreget (rúlestartó)

Walter üreget (diagramgyenerátor)

- Értékelési rendszerek

Szóegyföldolvasó (Text-To-Speech)

Szószintézis (illesztett írásmű)

- Nevezetes

Elektronikus beszédfeldolvasó

Hinckley x szószintézis / tematikus változtatás = K