

Frequency-Warped Linear Prediction and Speech Analysis

Diplomityöesitelmä

2.3.2004

Jouni Pohjalainen

Johdanto

- spektrin mallintaminen Bark-asteikolla hyödyllistä mm. puheanalyysisovelluksissa
- erityisesti puheentunnistuksen esikäsittely
- taajuusvarpattu lineaariprediktio (WLP)
 - yleisiä ominaisuuksia
 - perinteisten kehyspohjaisten ja laskennallisesti tehokkaampien adaptiivisten menetelmien vertailu

Työn tavoitteet

- analyysiparametrien valinta
 - asteluku
 - aikaikkuna
- menetelmien vertailu

Teoriaa:Lineaariprediktio (1/3)

- linear predictive coding (LPC), linear prediction (LP), suom. myös lineaarinen ennustaminen
- erityisesti koodausmenetelmä, käytössä myös muissa puheanalyysisovelluksissa
- yleistyi puheen käsittelyssä 1970-luvun alussa

Teoriaa:Lineaariprediktio (2/3)

- aika-alueessa:
$$s_n = \sum_{k=1}^p a_k s_{n-k} + G u_n$$

- taajuusalueessa:
$$H(z) = \frac{G}{1 - \sum_{k=1}^p a_k z^{-k}}$$

- signaalinäytteen **ennustaminen**:
$$\hat{s}_n = \sum_{k=1}^p a_k s_{n-k}$$

Teoriaa: Lineaariprediktio (3/3)

- ennustusvirhe eli ***residuaali***:

$$e_n = s_n - \sum_{k=1}^p a_k s_{n-k}$$

- normaaliyhtälöt (perusmuoto):

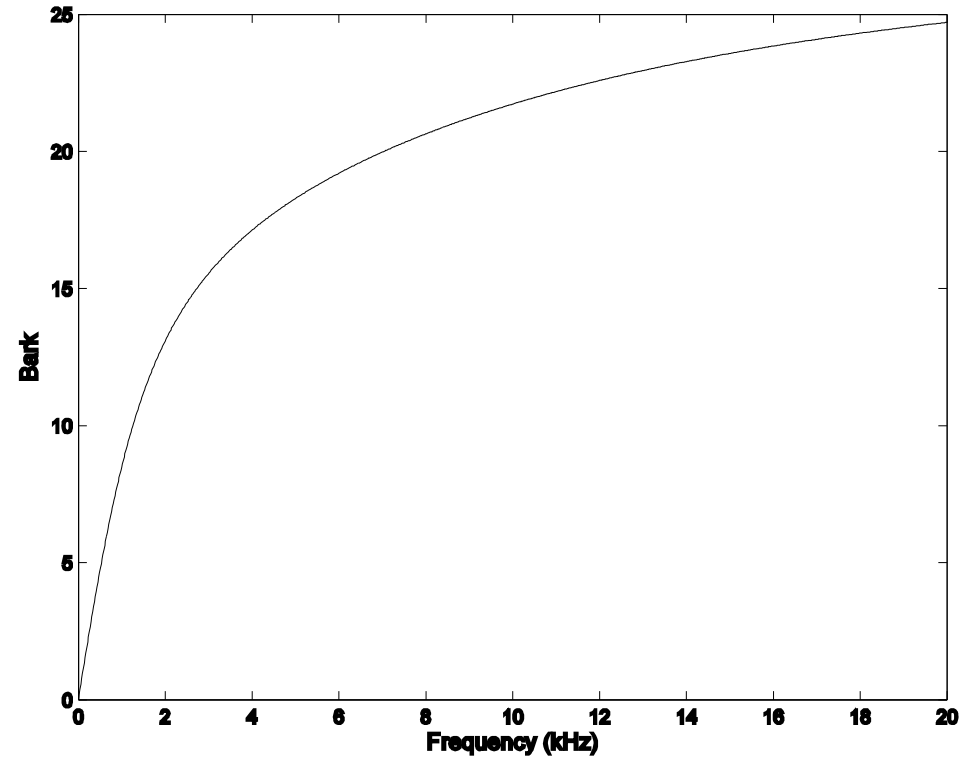
$$\sum_{k=1}^p a_k \sum_n s_{n-k} s_{n-i} = \sum_n s_n s_{n-i} \quad 1 \leq i \leq p$$

- kerrointen päivitys (adaptiivinen LMS):

$$\bar{a}_{n+1} = \bar{a}_n + \mu e_n \bar{x}_n$$

Teoriaa: Taajuusasteikot

- Bark-asteikon suhde lineaariseen taajuusasteikkoon



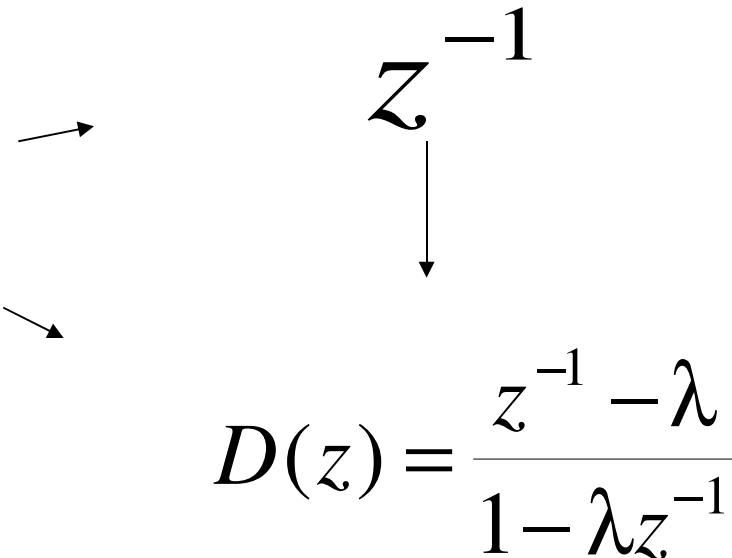
Teoriaa: Taajuusvarppaus (frequency warping)

- yksikköviive-elementit

korvataan allpass-suotimilla

- täten taajuusresoluutio voidaan saada vastaamaan esimerkiksi kuulonmukaista Bark-asteikkoa, kun λ valitaan sopivasti



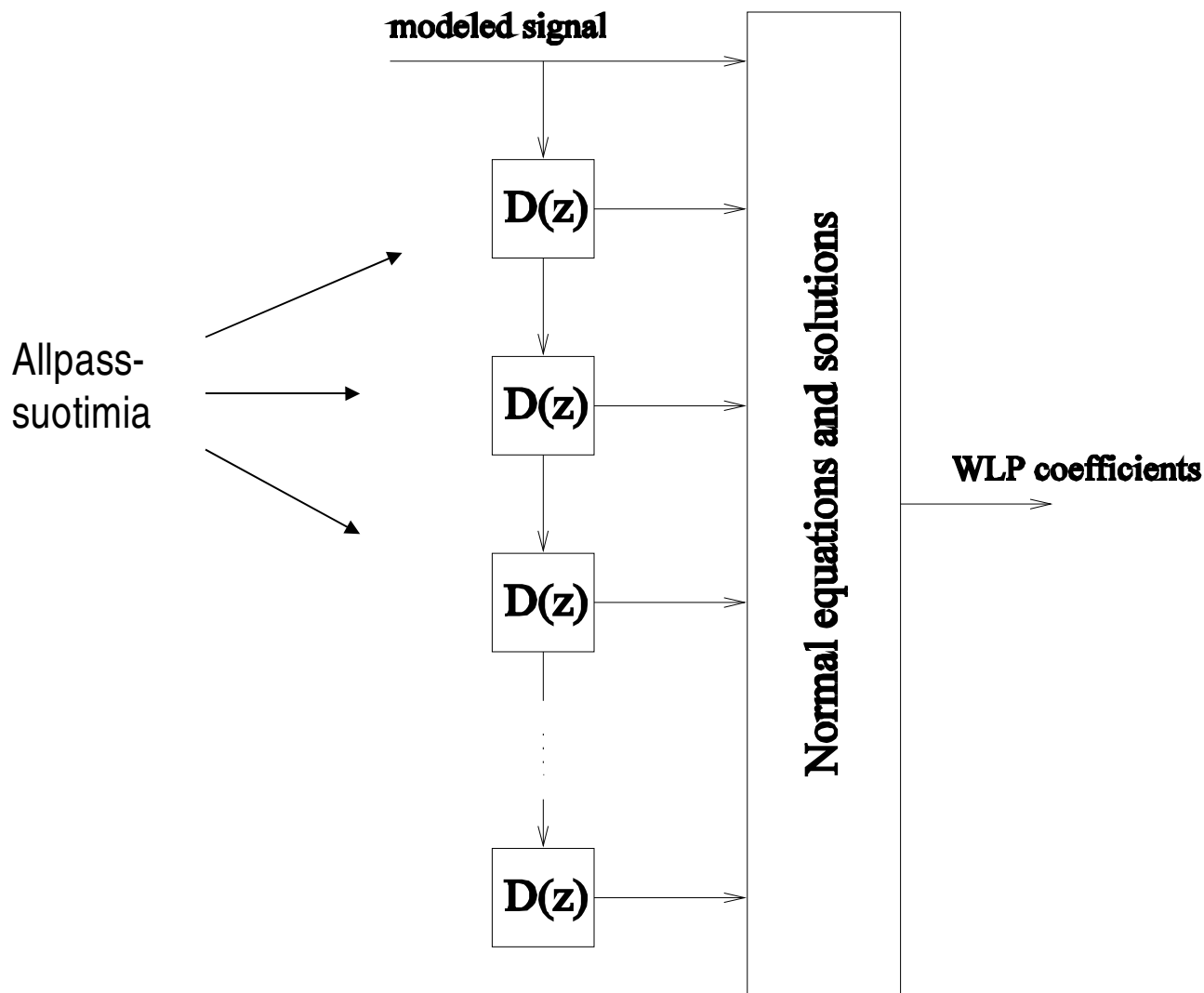


The diagram illustrates the frequency warping process. It shows the mapping of the unit delay element z^{-1} to the transfer function $D(z) = \frac{z^{-1} - \lambda}{1 - \lambda z^{-1}}$. A horizontal arrow points from the text 'yksikköviive-elementit' to the z^{-1} term. A vertical arrow points from z^{-1} down to the numerator of the transfer function. A diagonal arrow points from the text 'korvataan allpass-suotimilla' to the denominator of the transfer function.

$$D(z) = \frac{z^{-1} - \lambda}{1 - \lambda z^{-1}}$$

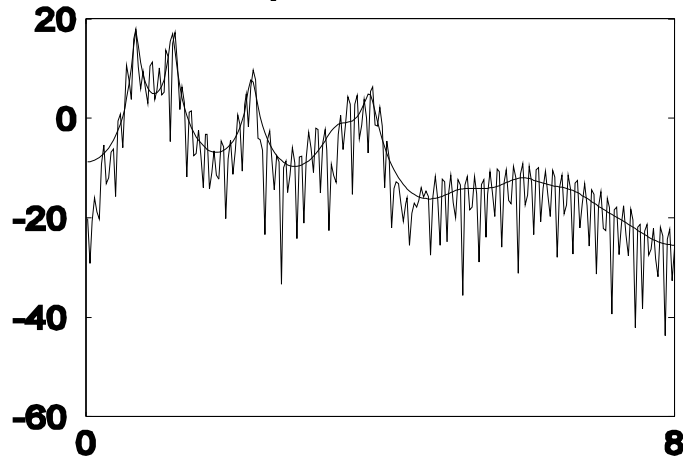
Taajuusvarpattu lineaariprediktio

Warped linear prediction (WLP)

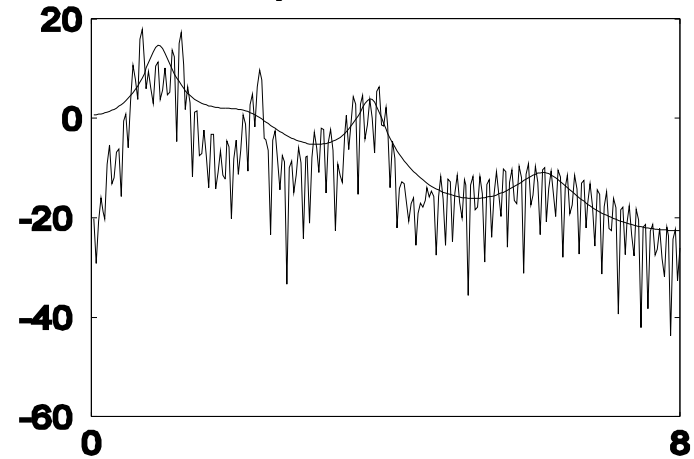


LP ja WLP: asteluku 20->10

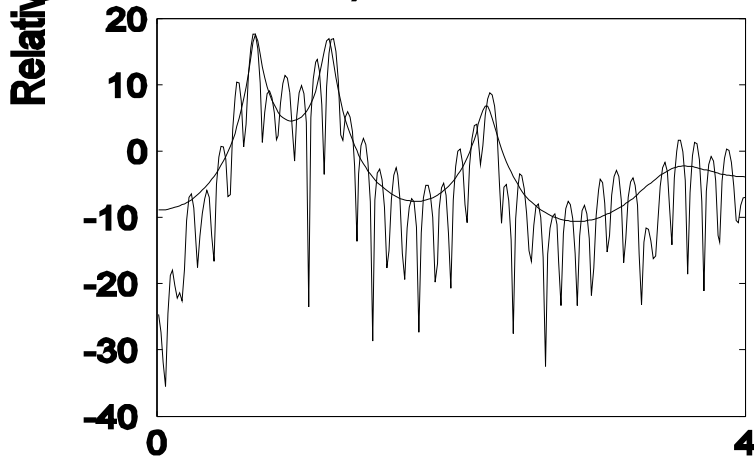
a) 16 kHz LP20



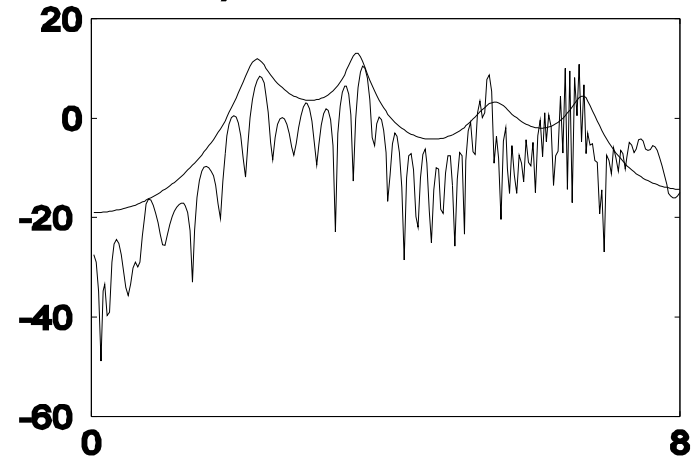
b) 16 kHz LP10



c) 8 kHz LP10



d) 16 kHz BarkWLP10



Frequency (kHz)

Tarkastellut WLP-menetelmät

- Kehyspohjaiset menetelmät:
 - modifioitu autokorrelaatiomenetelmä
 - modifioitu kovarianssimenetelmä
- Adaptiiviset menetelmät:
 - lattice-menetelmä (Gradient Adaptive Lattice, GAL)
 - suoramuotomenetelmä (Least-Mean-Squares, LMS)

Mitat spektrimallien vertailuun

1. normalisoitu ennustusvirhe-energia:

2. sama matalilla taajuuksilla (n.0-3 kHz):

3. ”varpattu” normalisoitu virhe:

4. residuaalin spektrin tasaisuus:

5. sama matalilla taajuuksilla (n. 0-3 kHz):

 V_L V_W F_L

$$V_N = \frac{\sum_n e_n^2}{\sum_n s_n^2}$$

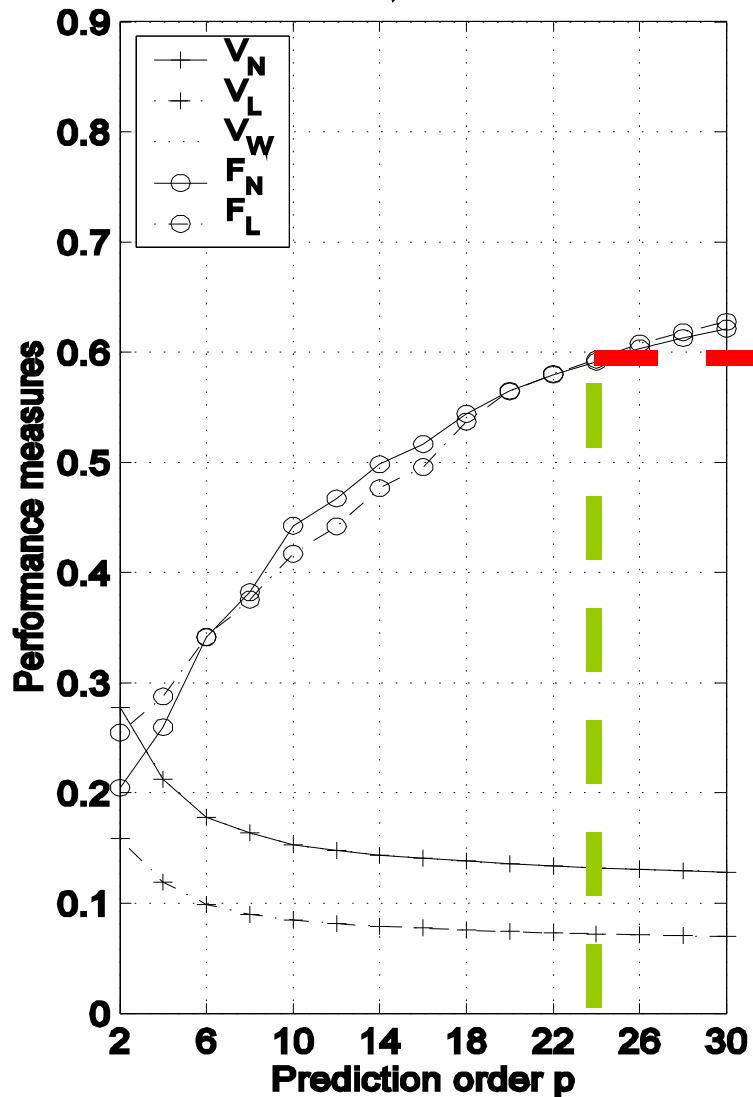
$$F_N = \frac{(\prod_{n=1}^N E_n)^{1/N}}{\frac{1}{N} \sum_{n=1}^N E_n}$$

Tulokset

1. LP/WLP vertailua ja asteluvun valinta
2. WLP-aikaresoluution säätö (erityisesti adaptiiviset menetelmät)
3. WLP-menetelmien keskinäinen vertailu

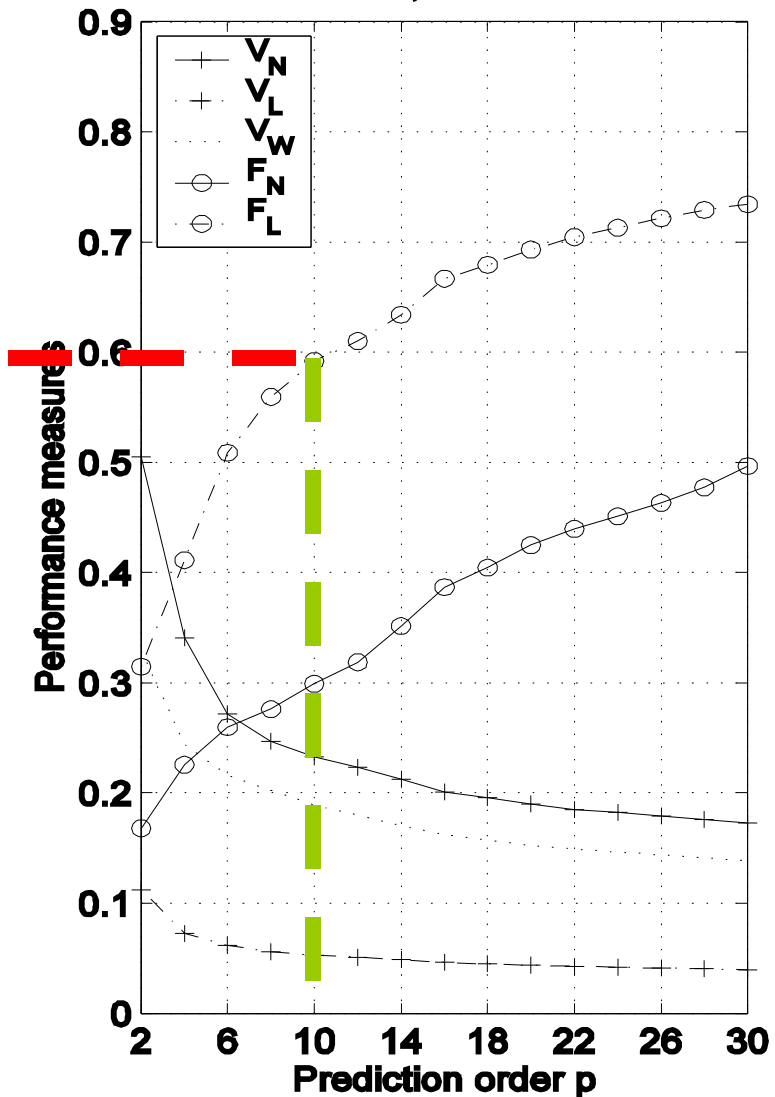
1. LP

Autocorrelation method
22 kHz, all frames



VS WLP

Bark-warped autocorrelation method
22 kHz, all frames

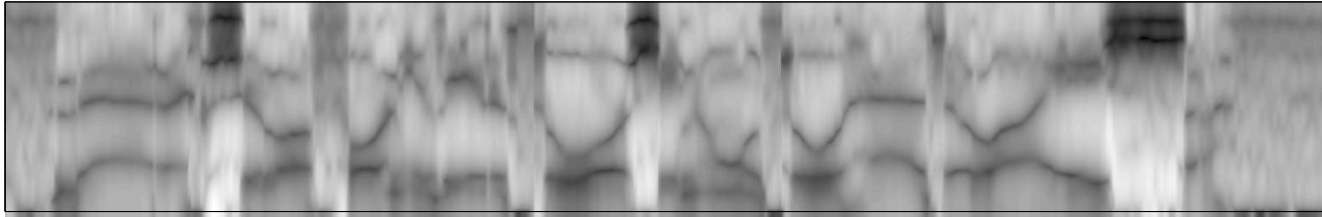


2. Aikaresoluutio

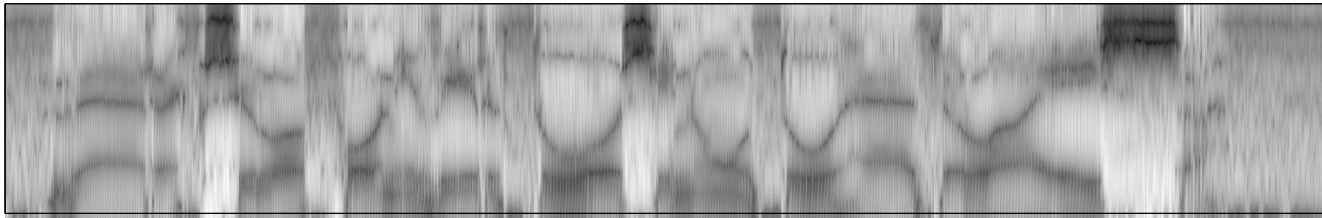
- Kehysmenetelmillä on tunnettu **kiinteä kehyspituus**, vaikuttaa aikaresoluutioon
- Adaptiivisilla menetelmillä vastaavasti **adaptaationopeusparametri** ("muisti")

Aikaresoluutio, esimerkki

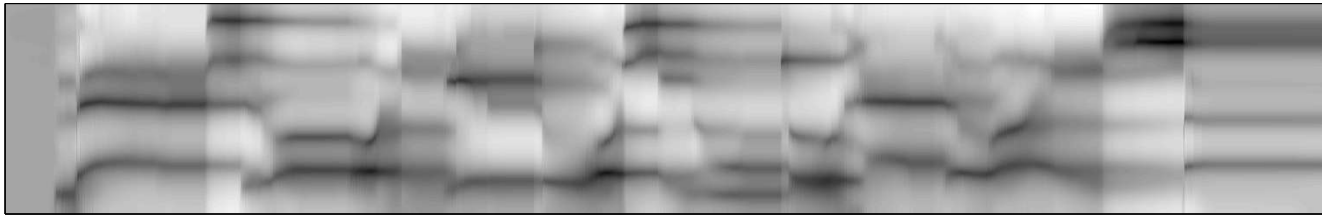
A) Autocorrelation method, estimation frame 50 ms



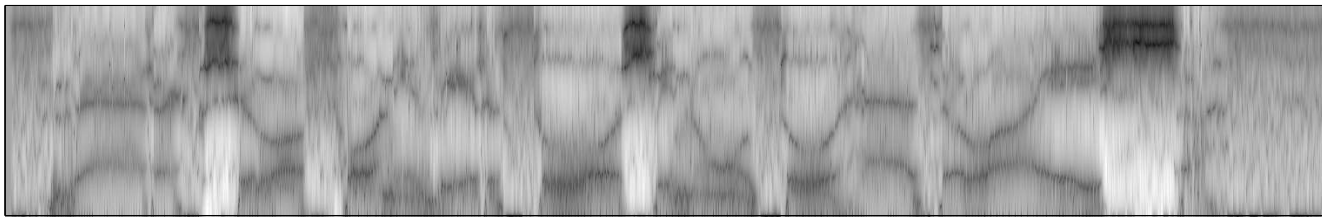
B) Autocorrelation method, estimation frame 10 ms



C) GAL, memory coefficient 0.9995



D) GAL, memory coefficient 0.9795



Time

Aikaikkunaan vaikuttavien analyysiparametrien valinta

25 millisekunnin tarkasteluikkunan kannalta parhaat
parametriarvot ($F_s=22$ kHz):

- Autokorrelaatiomenetelmä: 25 ms kehys
- Kovarianssimenetelmä: 20 ms kehys
- GAL: 0.993 muisti
- LMS: 0.06 päivitysaskel

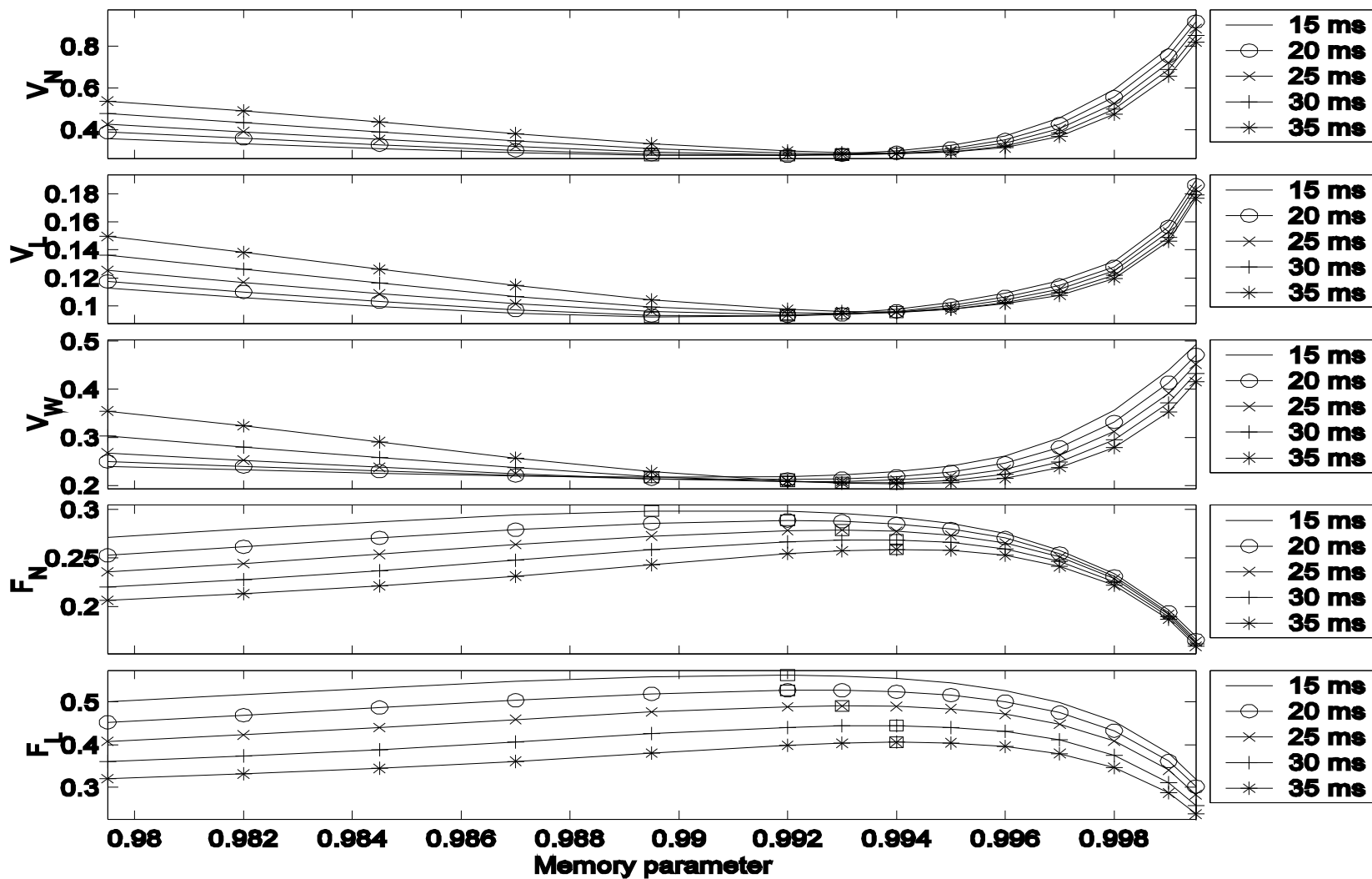
...analyysiparametrien valinta...

35 millisekunnin tarkasteluikkunan kannalta parhaat parametriarvot ($F_s=22$ kHz):

- Autokorrelaatiomenetelmä: 35 ms kehys
- Kovarianssimenetelmä: 30 ms kehys
- GAL: 0.994 muisti
- LMS: 0.05 päivitysaskel

Esim. aikaresoluutioparametrin valinta, gradient adaptive lattice

GAL, 22 kHz, all frames

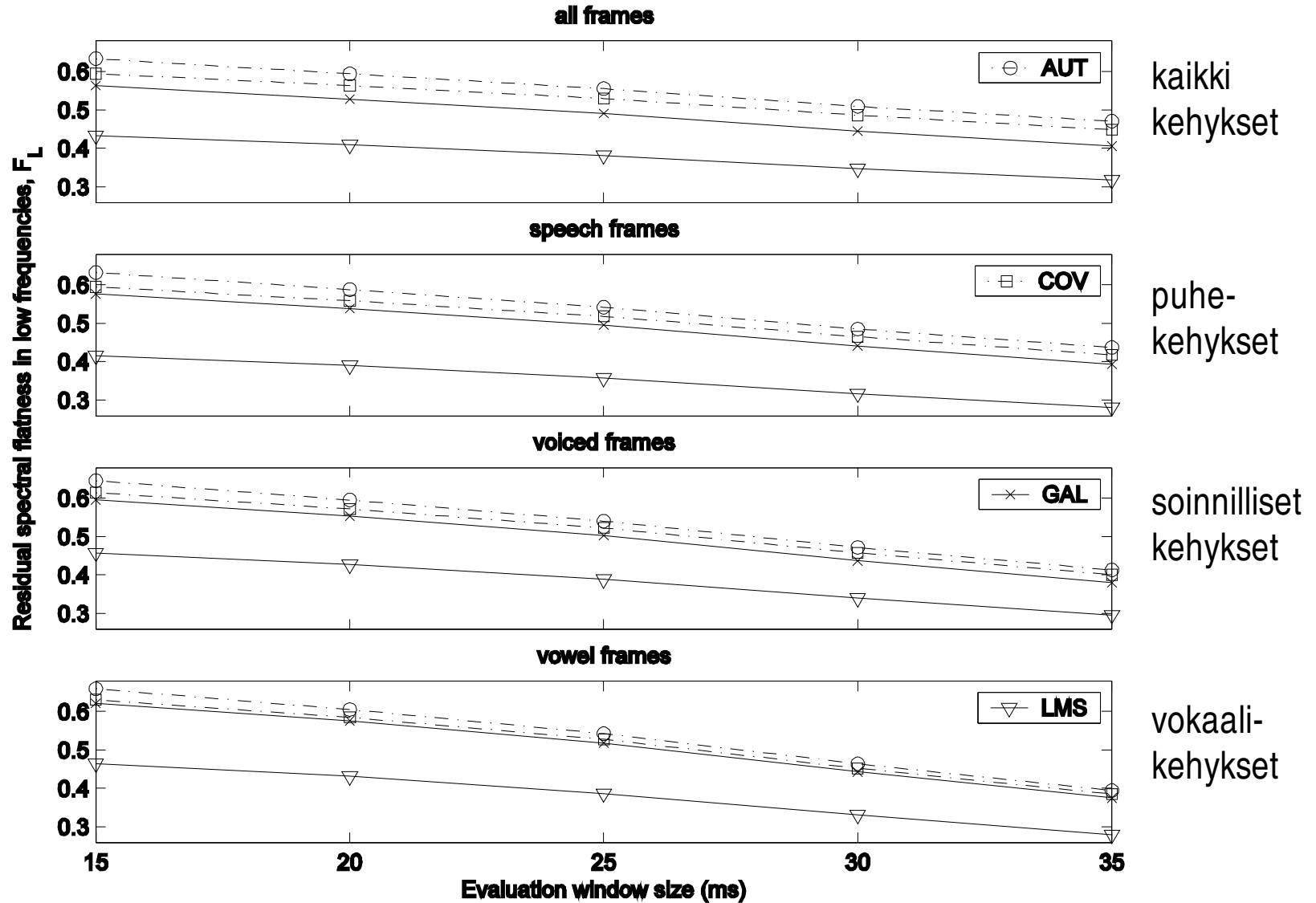


3. WLP-menetelmien vertailua

- kehysmenetelmät (autokorrelaatio ja kovarianssi) parhaita
- LMS huonoin
- GAL hyvä ainakin suhteellisen hitaasti muuttuvissa puhekohdissa

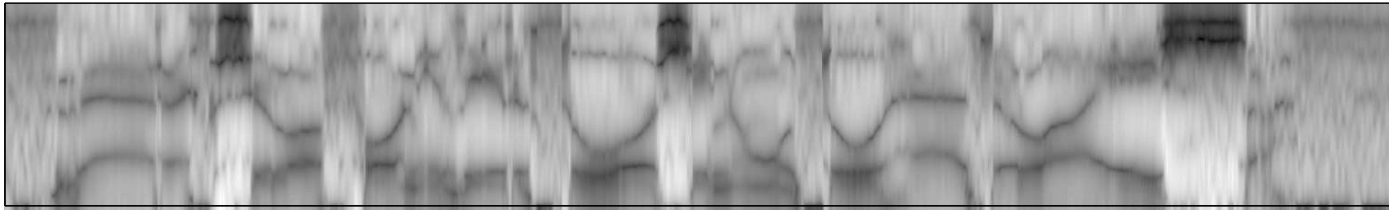
Menetelmien vertailu

F_L



Optimoidut WLP-menetelmät (25 ms)

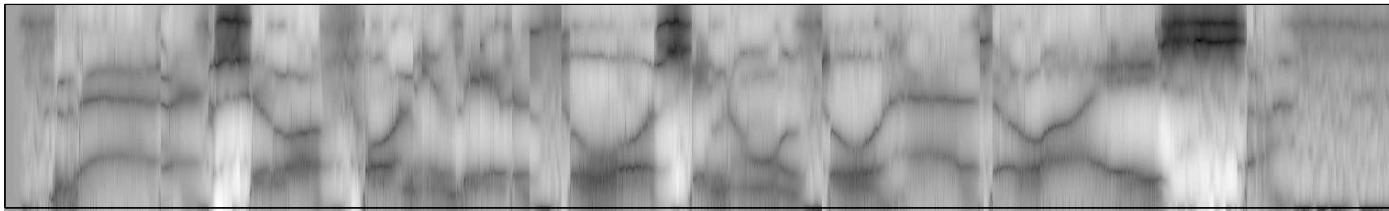
Autocorrelation method, estimation frame 25 ms



Covariance method, estimation frame 20 ms



GAL, memory coefficient 0.9930



LMS, update step size 0.06



Frequency

Time