



# Prediktivno kodiranje medijskih signala

Prof.dr.sc. Davor Petrinović,

29.veljače 2008.



# Linearna predikcija

prof.dr.sc. Davor Petrinović



# Predikcija procesa

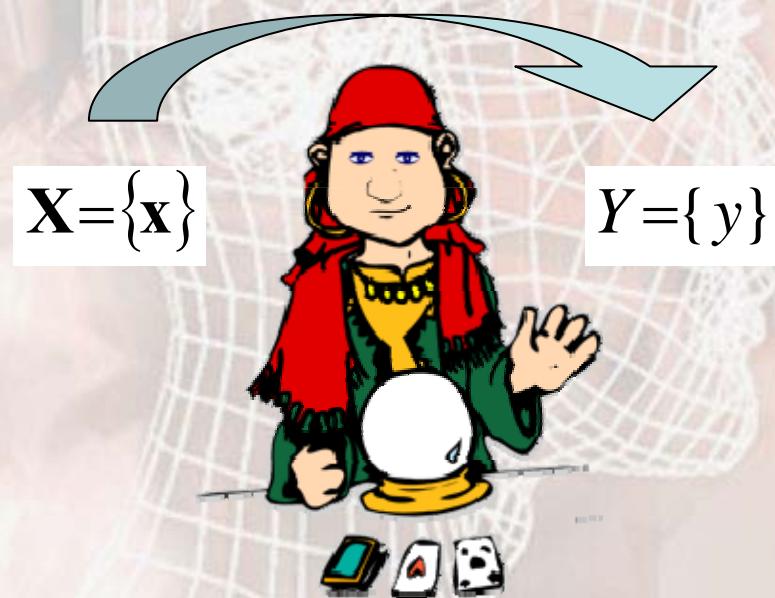
- Zadatak predikcije:
  - Na osnovu poznatog vektorskog procesa  $X$  potrebno je predvidjeti skalarni proces  $Y$ .
  - Temeljna pretpostavka je da postoje određene statističke zavisnosti između pojedinih komponenata procesa  $X$  i procesa  $Y$ .
  - Prediktor mora moći identificirati i modelirati takve statističke zavisnosti.
  - Uz primjenu prediktora, opis procesa  $Y$  biti će mnogo jednostavniji, jer ako poznamo proces  $X$ , veći dio informacije sadržane u procesu  $Y$  možemo “naslutiti” predikcijom iz procesa  $X$ , a samo eventualne razlike u odnosu na stvarnu vrijednost kodirati!



# Predikcija procesa

- Neka je proces  $\mathbf{X}$ , vektorski proces dimenzije  $p$ ,  $\mathbf{X}=\{\mathbf{x}\}$ , a  $Y$  ciljni skalarni proces  $Y=\{y\}$  kojeg želimo predvidjeti

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$



- Nažalost, u praksi je predikcija uvijek približna, pa je označavamo oznakom:  $\tilde{Y}=\{\tilde{y}\}$



# Predikcijska pogreška

- Cilj dobrog prediktora je da predikcija  $\tilde{Y}=\{\tilde{y}\}$  bude čim sličnija ciljnom procesu  $Y=\{y\}$  tj. da minimizira proces pogreške predikcije  $E=\{e\}$ 
$$e=y-\tilde{y}$$
- U tom smislu, potrebno je uvesti mjeru “sličnosti”, a u praksi se u tu svrhu najčešće koristi obična euklidska udaljenosti, tj. varijanca procesa  $E$ :
$$\sigma_e^2=E[e^2]=E[(y-\tilde{y})^2]$$
- ... uz pretpostavku da je očekivanje predikcijske pogreške jednako nuli,  $E[e]=0$  (nepristran pred.)



# Prediktor

- Prediktor se može realizirati na razne načine, ovisno o tipu statističkih zavisnosti koje pokušavamo iskoristiti.
- Najopćenitije, ... prediktor se može konstruirati iz zajedničke (ili uvjetne) funkcije gustoće vjerojatnosti izvornog vektorskog procesa  $X$  i ciljnog procesa  $Y$ , engl. ***joint*** (ili ***conditional***) ***probability density function***.
- U praksi, najčešće se koristi **linearni prediktor** (engl. ***linear predictor***) koji iskorištava linearne zavisnosti između izvornog i ciljnog procesa koje su opisane korelacijama.



# Linearni prediktor

- Linearni prediktor postupak predikcije ostvaruje običnim skalarnim množenjem izvornog procesa  $\mathbf{X}$  sa prediktorom  $\alpha$ :

$$\tilde{y} = [\alpha^\top] \cdot \begin{bmatrix} \mathbf{x} \end{bmatrix}$$

- Prediktor  $\alpha$  se odabire tako da minimizira varijancu predikcijske pogreške:

$$\bar{\alpha} = \underset{\alpha}{\operatorname{argmin}}(\sigma_e^2(\alpha))$$

$$\sigma_e^2(\alpha) = E[(y - \alpha^\top \mathbf{x})^2]$$



# Optimalni linearni prediktor

- Može se pokazati da prediktor koji minimizira  $\sigma_e^2(\alpha)$  mora zadovoljavati slijedeći uvjet:

$$E[\begin{bmatrix} \mathbf{x} \end{bmatrix} \cdot e] = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$
$$\begin{aligned} e &= y - \bar{\mathbf{a}}^\top \mathbf{x} \\ &= y - \mathbf{x}^\top \bar{\mathbf{a}} \end{aligned}$$

- Opisani uvjet traži da pogreška predikcije  $e$  bude linearno nezavisna sa svakom komponentom izvornog procesa  $X$ , pa se još naziva uvjetom ortogonalnosti.



# Optimalni linearni prediktor

- Uvrštavanjem izraza za pogrešku predikcije dobivamo:

$$E[\mathbf{x} \cdot (y - \mathbf{x}^\top \bar{\alpha})] = E[\mathbf{x} \cdot y] - E[\mathbf{x} \cdot \mathbf{x}^\top \bar{\alpha}] = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

– a obzirom da se prediktor može izvući ispred operatora očekivanja, dobivamo slijedeći izraz:

$$\Phi_{xx} \cdot \bar{\alpha} = \Psi_{xy}$$

$$\Phi_{xx} = E[\mathbf{x} \cdot \mathbf{x}^\top] = E\left[\begin{bmatrix} \mathbf{x} \end{bmatrix} \cdot \begin{bmatrix} & \mathbf{x}^\top & \end{bmatrix}\right]$$

$$\Psi_{xy} = E[\mathbf{x} \cdot y] = E\left[\begin{bmatrix} \mathbf{x} \end{bmatrix} \cdot y\right]$$



# Optimalni linearni prediktor

- Dakle, optimalni prediktor se nalazi rješavanjem linearog sustava jednadžbi.
    - Matrica sustava  $\Phi_{xx}$  se naziva kovarijacijska matrica jer opisuje korelacijska svojstva između svih parova komponenata izvornog procesa;
    - Ona je simetrična matrica dimenzije  $p \times p$ .
    - Stupac  $\Psi_{xy}$  opisuje korelacije između ciljnog procesa  $Y$  i pojedinih komponenti izvornog procesa  $X$  i ima dimenziju  $p \times 1$ .
    - Rješenje nalazimo kao:
- $$\bar{\alpha} = \Phi_{xx}^{-1} \Psi_{xy}$$



# Linearna predikcija

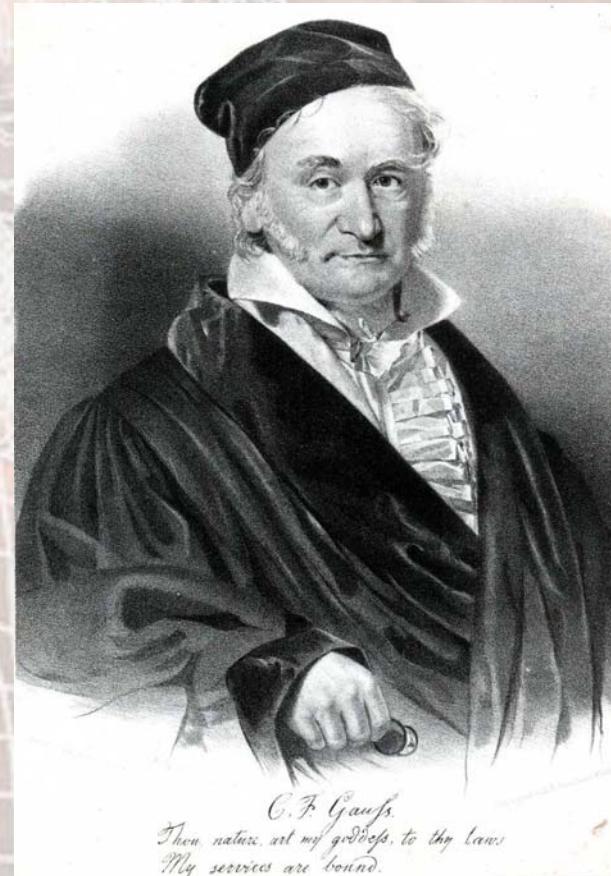
- Linearna predikcija je usko vezana sa **regresijskom analizom**, koja u svojoj najranijoj verziji proizlazi iz **metode najmanjih kvadrata**.
- Metodu najmanjih kvadrata razvijaju matematičari **Legendre** (1805) i **Gauss** (1809) za potrebe analize orbita nebeskih tijela; planeta i njihovih prirodnih satelita.



# Metoda najmanjih kvadrata



- Adrien-Marie Legendre



Carl Friedrich Gauss



# Metoda najmanjih kvadrata



- Gauss je primijenio postupak najmanjih kvadrata na estimaciju putanje patuljastog planeta Ceres-a
- *Theoria motus corporum coelestium in sectionibus conicis solem ambientum, 1801* (objavljeno 1809)



# Modeliranje vremenskih korelacija

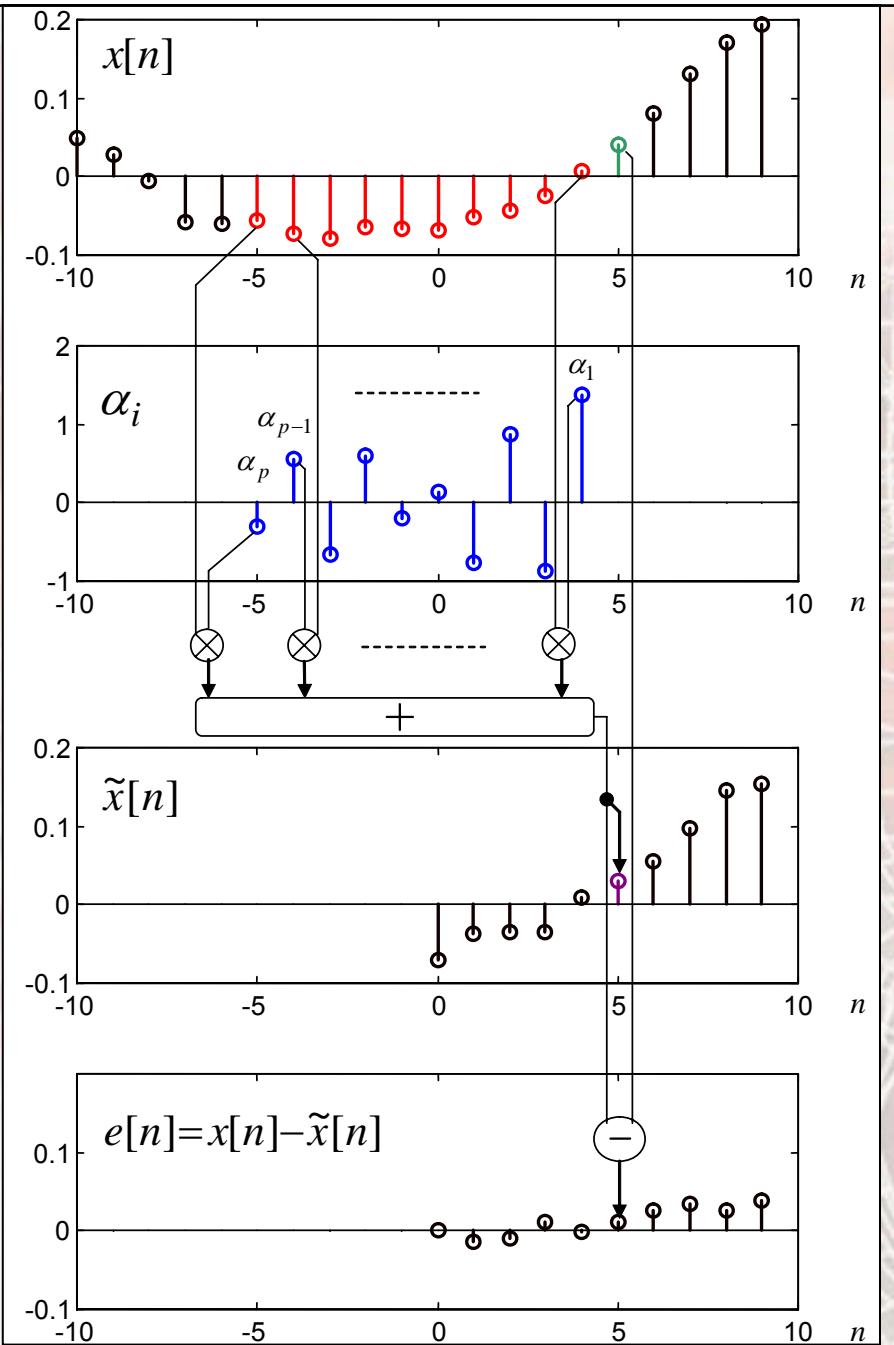
- Postupak linearne predikcije može se iskoristiti za **modeliranje vremenskih korelacija signala**.
- Pretpostavka ... trenutni uzorak signala  $x[n]$  je koreliran sa prethodnim uzorcima  $x[n-1], x[n-2], \dots, x[n-p]$ , te se može iz njih predvidjeti.
- Dakle izvorni proces  $X$  i ciljni proces  $Y$  formiraju se iz uzorka istog signala na slijedeći način:

$$\mathbf{x} = \begin{bmatrix} x[n-1] \\ x[n-2] \\ \vdots \\ x[n-p] \end{bmatrix}$$

$\mathbf{X} = \{\mathbf{x}\}$



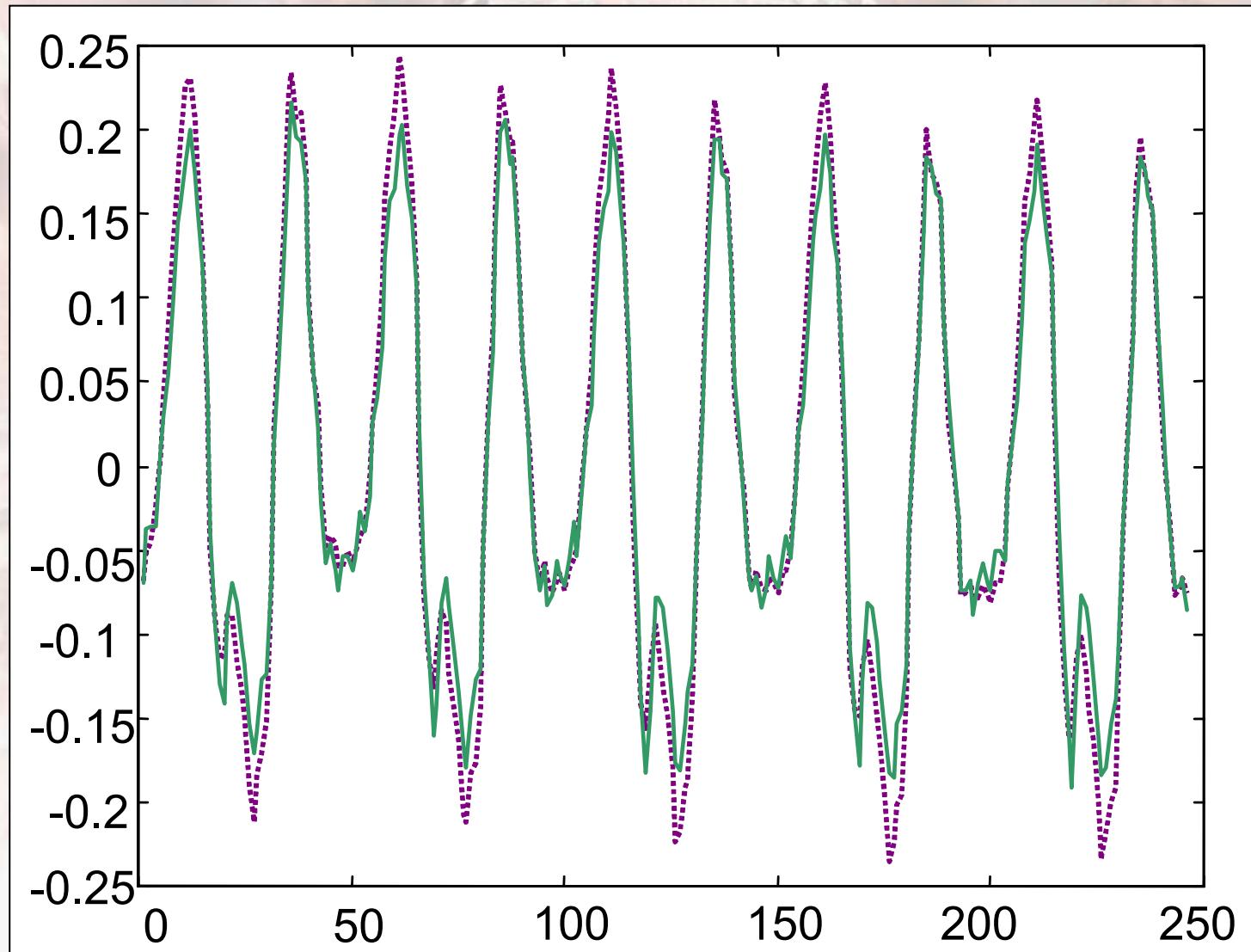
$Y = \{y\} = \{x[n]\}$



- signal
  - izvorni proces ..  $x$
  - ciljni proces .....  $y$
- koeficijenti prediktora  $\alpha$
- predikcija
- pogreška predikcije

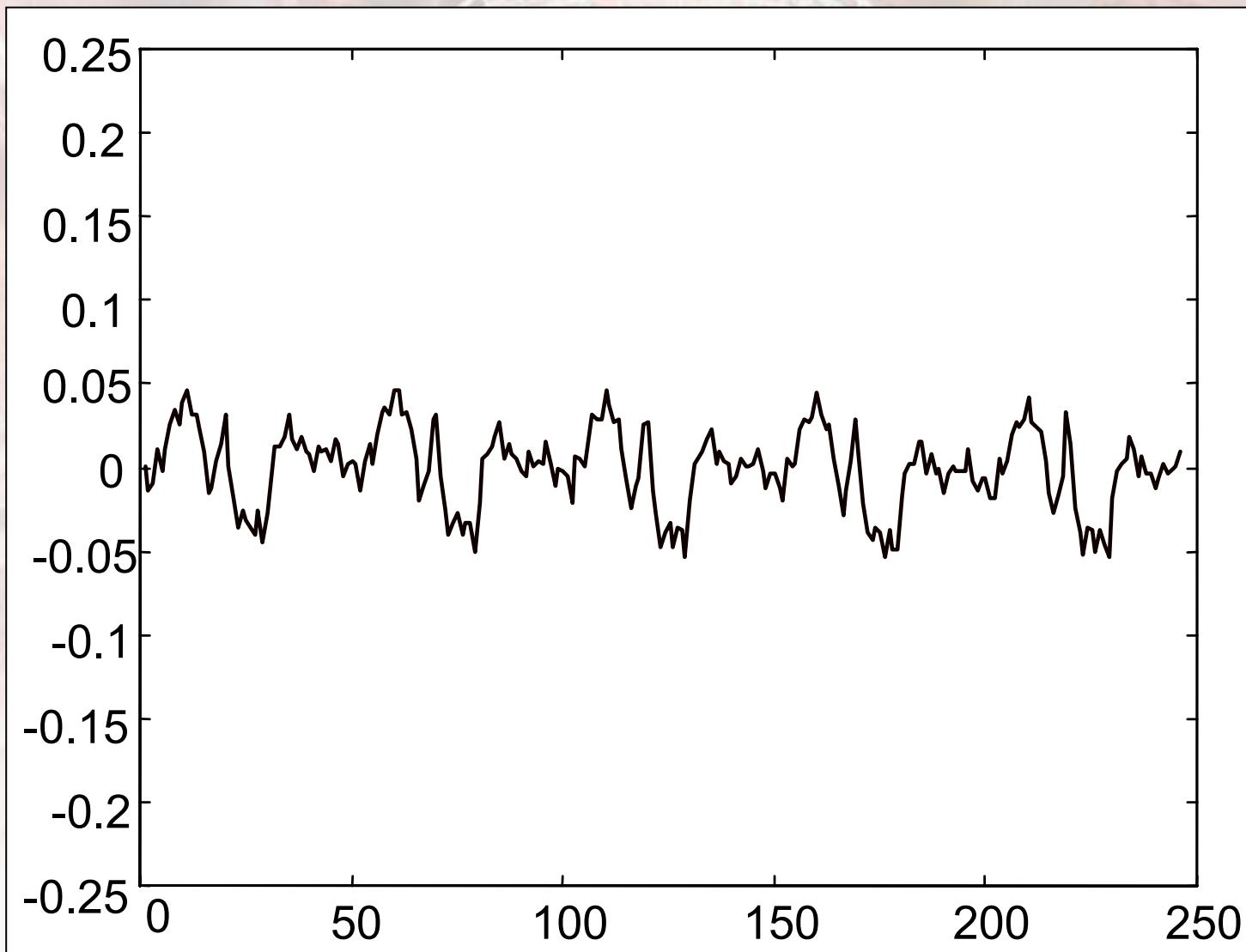


# Primjer signala i njegove predikcije





# Pogreška predikcije za isti primjer





# Linearna predikcija vremenski koreliranih signala

- Za opisani slučaj modeliranja vremenskih korelacija, matrica sustava jednadžbi je:

$$\Phi_{xx} = E[\mathbf{x} \cdot \mathbf{x}^\top] = E\left[ \begin{bmatrix} x[n-1] \\ x[n-2] \\ \vdots \\ x[n-p] \end{bmatrix} \cdot [x[n-1] \ x[n-2] \cdots \ x[n-p]]^\top \right]$$

- uz pretpostavku ergodičnosti procesa  $\mathbf{X}$  očekivanje ovog vanjskog produkta izračunava se usrednjavanjem preko svih vremenskih uzoraka  $n$  (dakle kao vremenska sr. vrijednost)



# Linearna predikcija vremenski koreliranih signala

- Analogno tome, stupac  $\Psi_{xy}$  sa desne strane matrične jednadžbe se nalazi kao:

$$\Psi_{xy} = E[\mathbf{x} \cdot y] = E\left[ \begin{bmatrix} x[n-1] \\ x[n-2] \\ \vdots \\ x[n-p] \end{bmatrix} \cdot x[n] \right]$$



# Linearna predikcija vremenski koreliranih signala

- U praksi će signal  $x[n]$  za koji se izračunava optimalni prediktor biti konačnog trajanja.
- Tada se uz pretpostavku stacionarnosti signala, operator očekivanja može zamijeniti običnim usrednjavanjem preko svih raspoloživih uzoraka konačne realizacije tog procesa:

$$\Phi_{xx}(i,k) = E[x[n-i] \cdot x[n-k]] \cong \phi(i,k) = \frac{1}{N} \sum_n x[n-i] \cdot x[n-k]$$

$$\Psi_{xy}(i) = E[x[n-i] \cdot x[n]] \cong \phi(i,0) = \frac{1}{N} \sum_n x[n-i] \cdot x[n]$$



# Linearna predikcija vremenski koreliranih signala

- To vodi na sustav jednadžbi oblika:

$$\begin{bmatrix} \phi(1,1) & \phi(1,2) & \dots & \phi(1, p) \\ \phi(2,1) & \phi(2,2) & \dots & \phi(2, p) \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \phi(p,1) & \phi(p,2) & \dots & \phi(p, p) \end{bmatrix} \begin{bmatrix} \bar{\alpha}_1 \\ \bar{\alpha}_2 \\ \dots \\ \dots \\ \bar{\alpha}_p \end{bmatrix} = \begin{bmatrix} \phi(1,0) \\ \phi(2,0) \\ \dots \\ \dots \\ \phi(p,0) \end{bmatrix}$$

$$\phi(i,k) = \frac{1}{N} \sum_n x[n-i] \cdot x[n-k]$$

– faktor  $1/N$  se može i izostaviti zbog kraćenja!

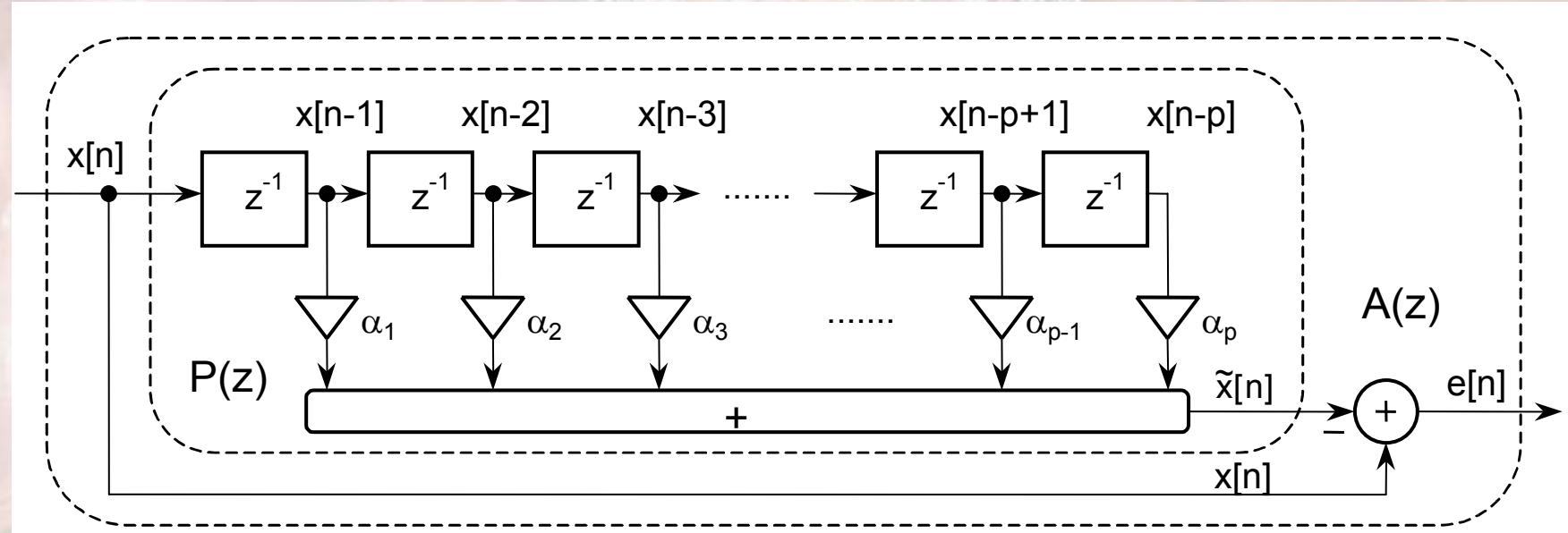


# Svojstva linearog prediktora

- Prediktor sa  $p$  koeficijenata može modelirati korelacije najdalje do  $p$  uzoraka unazad
  - Broj  $p$  se naziva redom prediktora (engl. ***prediction order***).
  - Povećanje reda monotono smanjuje pogrešku  $\sigma_e^2(\bar{a})$
- Linearni prediktor ima vrlo jednostavnu izvedbu kao digitalni filter s konačnim impulsnim odzivom reda  $p$ , s prijenosnom funkcijom označenom sa  $P(z)$ .
  - Izračunavanje predikcije traži  $p$  množenja i  $p - 1$  zbrajanja za jedan izlazni uzorak predikcije  $\tilde{x}[n]$



# Izvedba linearog prediktora



- prijenosna funkcija prediktora  $P(z)$

$$P(z) = \sum_{k=1}^p \alpha_k z^{-k}$$

$$e[n] = x[n] - \tilde{x}[n] = x[n] - \sum_{k=1}^p \alpha_k x[n-k]$$

$$A(z) = \frac{E(z)}{X(z)} = 1 - \sum_{k=1}^p \alpha_k z^{-k} = 1 - P(z)$$



# Rekonstrukcija polaznog signala iz predikcijse pogreške

- Pod određenim uvjetima, moguće je “invertirati” postupak predikcije na način da se originalni signal rekonstruira iz signala pogreške predikcije.
- To se provodi digitalnom filtracijom signala  $e[n]$  rekurzivnim filtrom  $H(z)$  koji ima beskonačni implusni odziv  $h[n]$  i prijenosnu funkciju:

$$H(z) = \frac{1}{A(z)} = \frac{X(z)}{E(z)} = \frac{1}{1 - P(z)} = \frac{1}{1 - \sum_{k=1}^p \alpha_k z^{-k}}$$

- ... pa se stoga  $A(z)$  naziva inverznim filtrom.



# Rekonstrukcija polaznog signala iz predikcije pogreške

- Filtar  $A(z)$  i  $H(z)$  su inverz jedan drugom, pa nul-točke filtra  $A(z)$  predstavljaju polove filtra  $H(z)$ .
- Filtar  $H(z)$  će biti stabilan i rekonstrukcija moguća samo ako su mu svi polovi unutar jedinične kružnice u  $z$ -ravnini, što traži da filter  $A(z)$  ima sve nul-točke po modulu manje od jedan.
- To je moguće postići primjenom odgovarajućih postupaka izračuna prediktora, a jedan od najpopularnijih je metoda autokorelacijske.



# Sustav jednadžbi za autokorelacijski postupak

- $R(j)$  je autokorelacija konačnog signala  $x[n]$  sa uzorcima na indeksima  $n=0$  do  $N-1$  za pomak  $j$

$$R(j) = \sum_{m=0}^{N-1-j} x[m] \cdot x[m+j]$$

$$\begin{bmatrix} R(0) & R(1) & R(2) & \dots & R(p-1) \\ R(1) & R(0) & R(1) & \dots & R(p-2) \\ R(2) & R(1) & R(0) & \dots & R(p-3) \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ R(p-1) & R(p-2) & R(p-3) & \dots & R(0) \end{bmatrix} \begin{bmatrix} \bar{\alpha}_1 \\ \bar{\alpha}_2 \\ \bar{\alpha}_3 \\ \dots \\ \dots \\ \bar{\alpha}_p \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ R(3) \\ \dots \\ \dots \\ R(p) \end{bmatrix}$$



# Kodiranje linearnom predikcijom

- Zbog svoje visoke učinkovitosti predviđanja na brojnim prirodnim procesima i signalima, postupak linearne predikcije nalazi čestu primjenu u postupcima kodiranja medijskih signala.
- Postupci kodiranja izvora koji koriste linearnu predikciju se zajednički nazivaju **Kodiranje linearnom predikcijom**, engl. ***Linear Prediction Coding, LPC***.
- Jedna od prvih primjena *LPCa* je bila u području kodiranja govora, ali je kasnije proširena i na druge vrste medijskih signala.



# Tipovi prediktivnog kodiranja

- Prediktivno kodiranje signala se može podijeliti u dvije osnovne grupe:
  - kodiranje u otvorenoj petlji (engl. ***open-loop prediction***)
  - kodiranje u zatvorenoj petlji (engl. ***closed-loop prediction***)
- Temeljna razlika je u izvornom procesu **X** od kojeg se računa predikcija:
  - Kod kodiranja u otvorenoj petlji, izvorni proces korišten od kodera i dekodera je sličan ali nije jednak, dok je
  - kod prediktivnog kodiranja u zatvorenoj petlji isti izvorni proces **X** raspoloživ i korišten na obje strane, koderu i dekoderu.



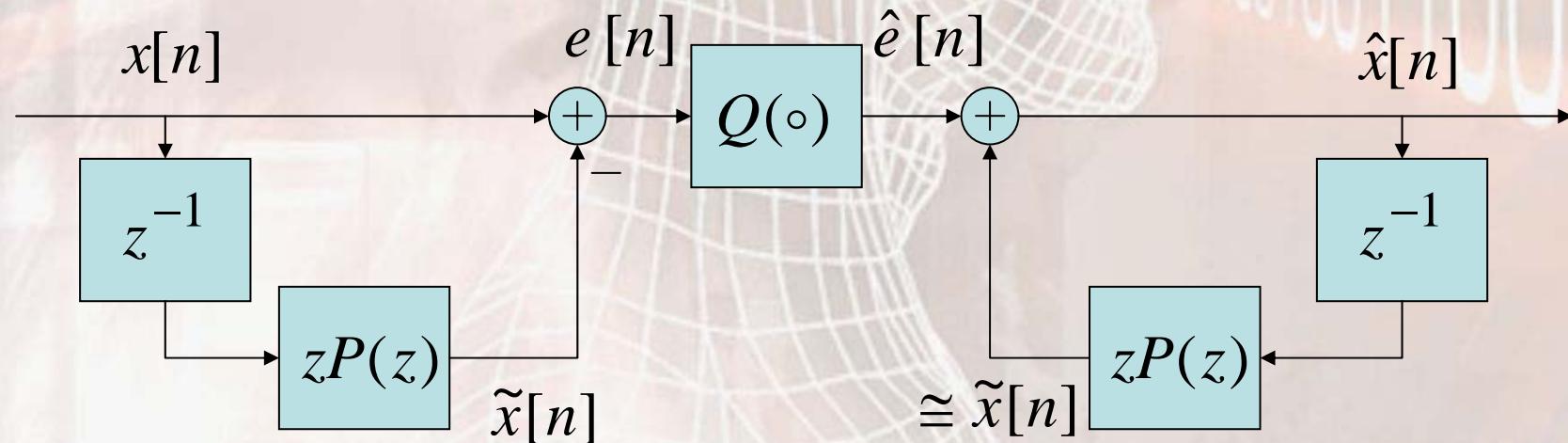
# Prediktivno kodiranje u otvorenoj petlji

- Koder provodi predikciju na osnovu zakašnjelih uzoraka originalnog signala i računa signal pogreške predikcije.
- Kvantizirani signal pogreške predikcije proslijeđuje dekoderu.
- Dekoder na kvantiziranim uzorcima predikcijske pogreške provodi inverzni postupak:
  - računa približnu predikciju na osnovu zakašnjelih rekonstruiranih uzoraka;
  - tu predikciju pribraja kvantiziranoj pogrešci predikcije da izračuna trenutni rekonstruirani izlazni uzorak.



# Prediktivno kodiranje u otvorenoj petlji

- Izvori pogrešaka rekonstrukcije:
  - pogreška kvantizacije:  $e[n] - \hat{e}[n]$ , ali i
  - razlika između predikcije na strani kodera i dekodera.
- Smanjenje pogreške kvantizacije automatski smanjuje i drugu pogrešku!





# Prediktivno kodiranje u **zatvorenoj** petlji

- Kod kodiranja **u zatvorenoj petlji** kao izvorni proces predikcije se umjesto zakašnjelih uzoraka ulaznog nekvantiziranog signala koristi  $P$  zadnjih rekonstruiranih uzoraka.
- U tu svrhu koder u svom sastavu ima “lokalnu repliku” dekodera, koji na izlazu generira identične rekonstruirane uzorke signala, kao i pravi dekoder na udaljenoj strani.
- Na ovaj način su prethodno rekonstruirani uzorci raspoloživi na obje strane i predstavljaju zajednički izvorni proces predikcije  $\mathbf{X}$  na osnovu kojeg se predviđa trenutni ulazni nekvantizirani uzorak  $x[n]$ .



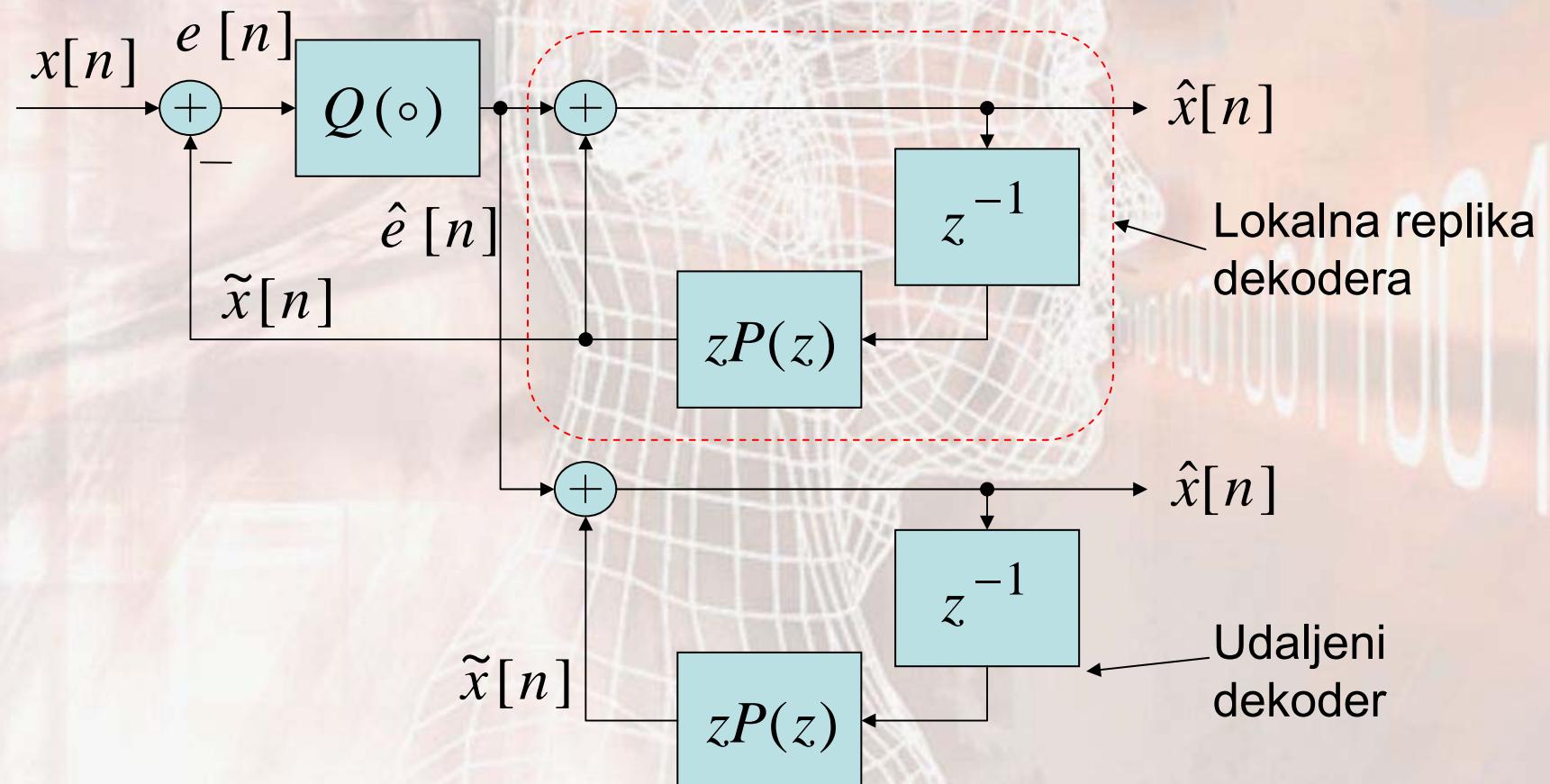
# Prediktivno kodiranje u **zatvorenoj** petlji, ... nastavak

- Koder računa razliku između trenutnog ulaznog uzorka i predikcije.
- Kvantiziranu pogrešku predikcije prosljeđuje dekoderu.
- Koder i dekoder ovu kvantiziranu vrijednost pribrajaju lokalno izračunatoj predikciji i tako dobivaju rekonstruiranu (kvantiziranu) vrijednost polaznog uzorka koju koriste kao izlaz dekodera, ali i kao ulaz prediktora za slijedeći uzorak.



# Prediktivno kodiranje u zatvorenoj petlji

- Koder i dekoder koriste istu predikciju!





# Adaptacija prediktora

- Medijski signal koji se predviđa prediktorom u pravilu ima **vremenski promjenjiva koreacijska svojstva**.
- Zbog toga je potrebno prediktor povremeno prilagoditi (adaptirati) svojstvima signala.
- Ta se informacija nažalost mora dojaviti i dekoderu da i on ima ažurnu repliku prediktora u svakom trenutku, radi mogućnosti jednoznačnog dekodiranja primljene poruke.
- Slanje ove dodatne informacije “troši” komunikacijski kanal i smanjuje efikasnost



# Adaptacija prediktora

- Opisani način adaptacije prediktora se naziva “adaptacija unaprijed”, engl. ***forward-adaptive prediction*** jer se dodatna informacija o prediktoru šalje direktno dekoderu (engl. ***side information***).
- Prediktor je kod takve adaptacije u pravilu izračunat tako da maksimizira učinkovitost prediktivnog kodiranja uzorka koji tek slijede.
- Alternativa jest “adaptacija unazad”, engl. ***backward-adaptive prediction*** kod koje se izbjegava potreba za slanjem prediktora dekoderu.



# Adaptacija unazad

- Kod adaptacije unazad koristi se sličan princip kao kod prediktivnog kodiranja u zatvorenoj petlji:
  - adaptacija se provodi iz uzoraka raspoloživih na obje strane, tj. iz prethodno rekonstruiranih izlaznih uzoraka.
- Problem, ...
  - Prediktor se optimira za uzorke koji su već poslati, a ne za one koje tek dolaze, pa
  - postupak ima smisla samo ako se korelacijska svojstva signala dovoljno sporo mijenjaju.
  - Prediktor se računa iz kvantiziranih izlaznih uzoraka pa pogreška kvantizacije utječe na učinkovitost određenog prediktora ... ne radi za male podatkovne tokove!



# Što smo naučili

- zadatak predikcije, izvorni i ciljni proces, pogreška predikcije
- prediktor, način određivanja
- linearni prediktor
- optimalno rješenje, uvjet ortogonalnosti
- metoda najmanjih kvadrata
- modeliranje vremenskih korelacija
- linearna predikcija vremenski koreliranih signala
- svojstva i izvedba prediktora
- rekonstrukcija signala iz predikcijske pogreške
- autokorelacijski postupak određivanja prediktora
- kodiranje linearnom predikcijom
- prediktivno kodiranje u otvorenoj i zatvorenoj petlji
- adaptacija prediktora



# Osnove kodiranja govora

prof.dr.sc. Davor Petrinović



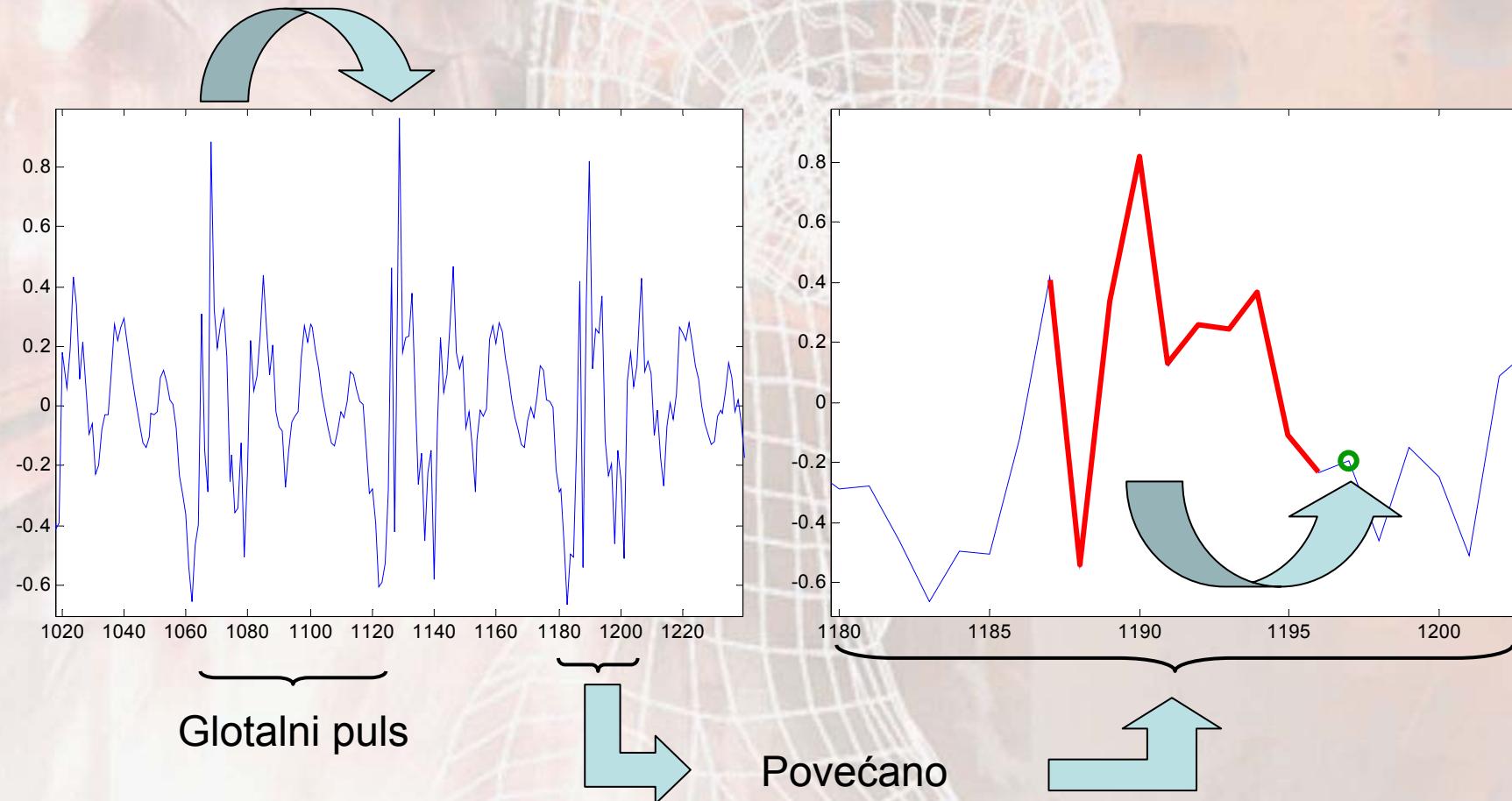
# Kodiranje govora

- Većina danas korištenih postupaka kodiranja govora temeljena je na linearnoj predikciji.
- LPC model izuzetno je blizak fizikalnom procesu nastajanja govora, pa stoga primjena LPC kodiranja na govor daje iznimno dobre rezultate!
- Kod govora postoje dvije vrste korelacija
  - vremenski **kratkotrajne** korelacijske vezane uz **djelovanje vokalnog trakta** i
  - vremenski **dugotrajne** korelacijske vezane uz proces **pobude vokalnog trakta** na glasnicama.
- LPC se uspješno koristi za obje.



# Korelaciјe u govoru

- dugotrajna korelacija:      kratkotrajna korelacija:



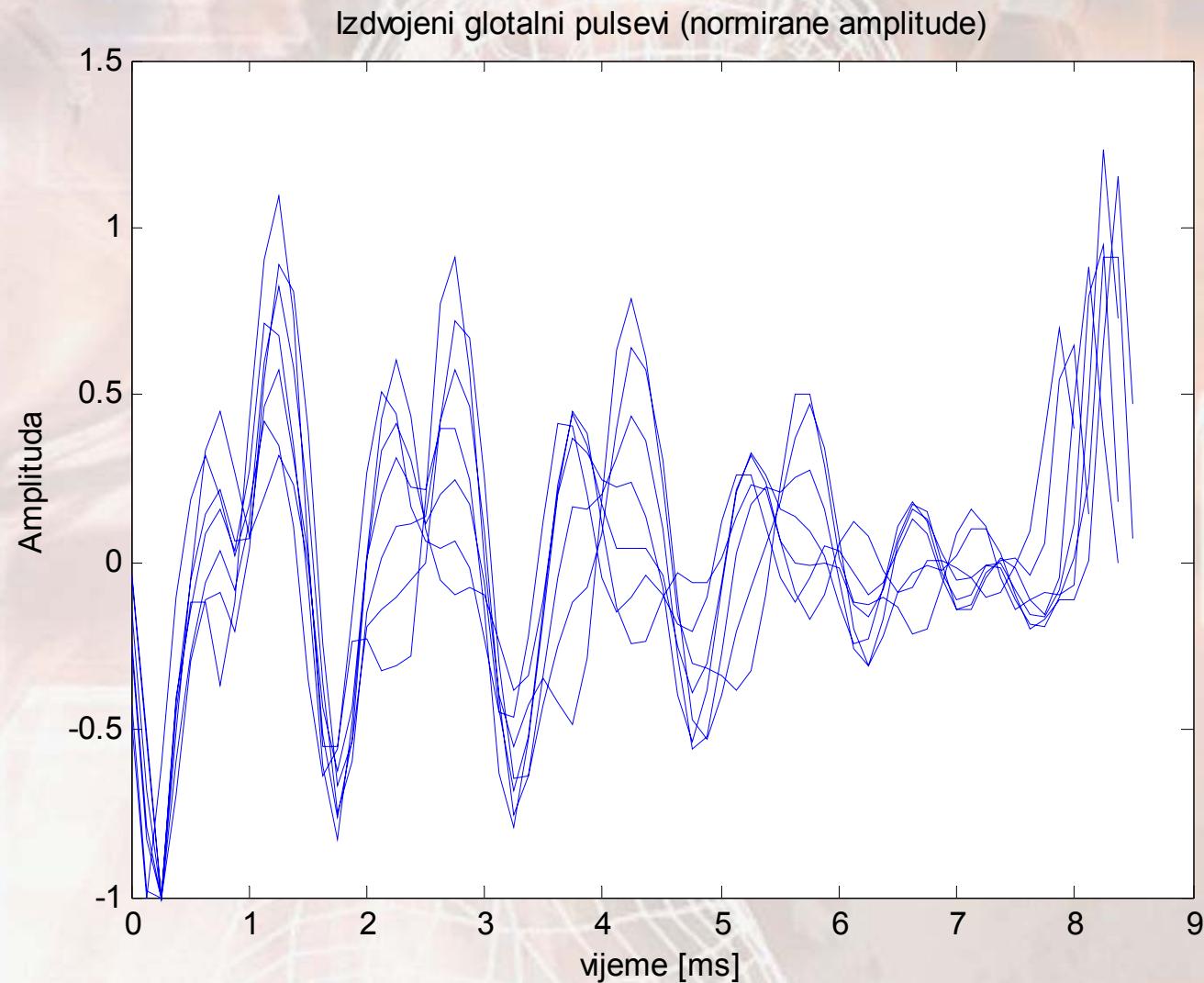


# Dugotrajne korelaciјe

- Dugotrajne korelaciјe (engl. ***long-term correlations***) očituju se kao sličnosti susjednih perioda (glotalnih pulseva) govornog signala.
- Svaka perioda odgovara jednom otvaranju glasnica, kada se na ulazu vokalnog trakta pojavljuje novi puls zračne struje iz pluća.
- Period ovisi o visini glasa i kreće se u rasponu od  $PP=3\text{ms}$  kod dječjih glasova, pa sve do  $PP=12-15\text{ms}$  za duboke muške glasove.
- Dugotrajna predikcija je to uspješnija, što su susjedne periode sličnije i što je njihov period vremenski stalniji.



# Sličnost susjednih perioda u govoru



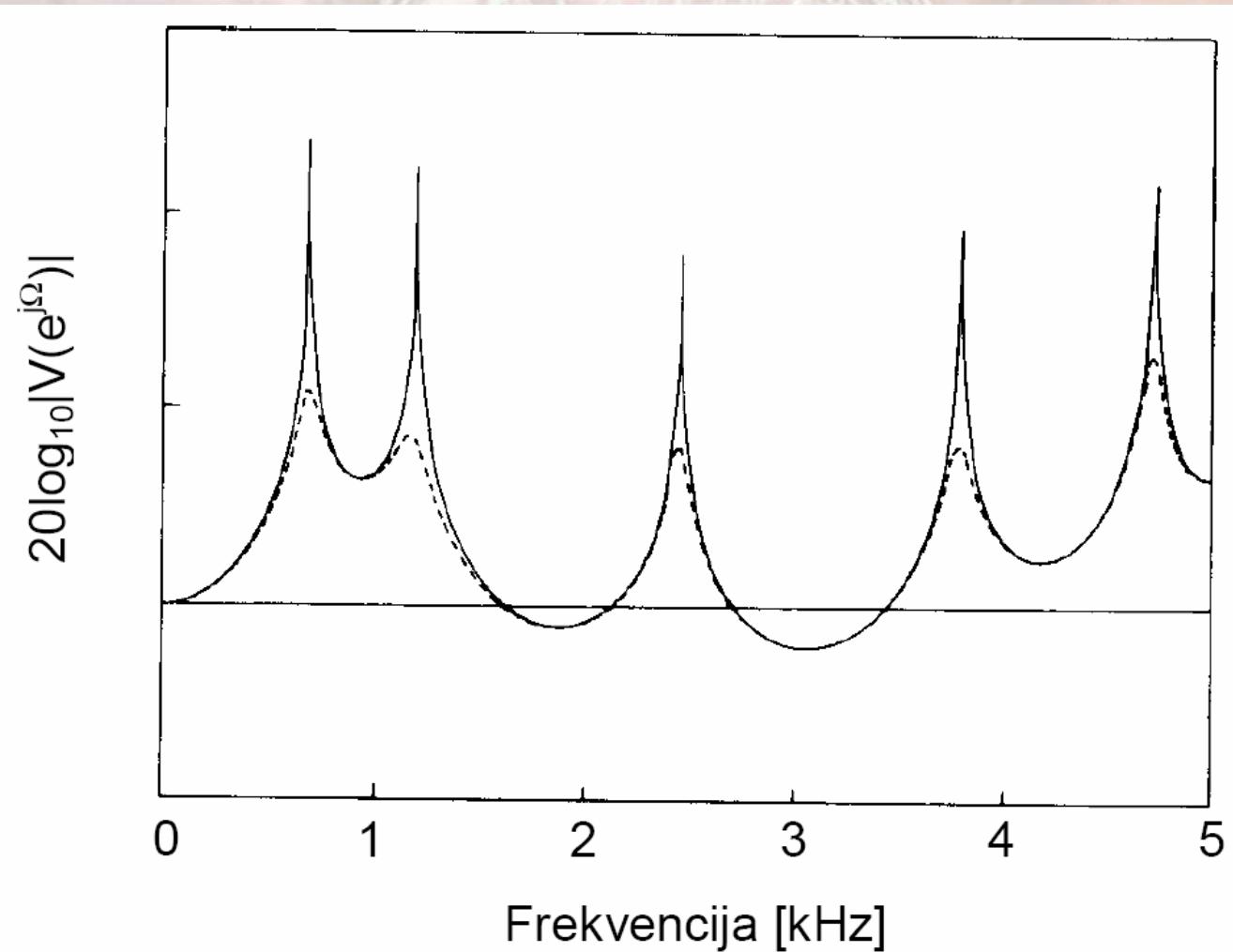


# Kratkotrajne korelaciјe

- Kratkotrajne korelaciјe (engl. ***short-term correlations***) vezane su uz promjene koje unosi vokalni trakt na signalu koji dolazi s glasnica i protežu se na vremenske periode unutar 1ms.
  - Vokalni trakt se ponaša kao niz selektivnih filtara koji pojačavaju pojedina spektralna područja ulaznog signala.
  - Pozicije i širine ovih filtara ovise o glasu koji se izgovara.
  - Npr. kod glasa “a” značajno se pojačavaju okoline sljedećih frekvencija: 800Hz, 1200Hz, 2500Hz, 3800Hz.
  - Takve rezonantne karakteristike u govoru se nazivaju **formanti**.



# Formanti glasa “a”





# Spektrogram govora

- Koristi se za vizualizaciju korelacija u govornom signalu.
- Prikazuje promjenu spektra, tj. modula Fourierove transformacije govora kroz vrijeme i sa frekvencijom.
- Amplituda (iznos spektra) prikazana je trećom dimenzijom pomoću boje (ili nijanse sive).
- Ovisno o frekvencijskoj selektivnosti, postoje dva tipa spektrograma:
  - uskopojasni spektrogram (engl. ***narrowband spectrogram***)
  - širokopojasni spektrogram (engl. ***wideband spectrogram***)

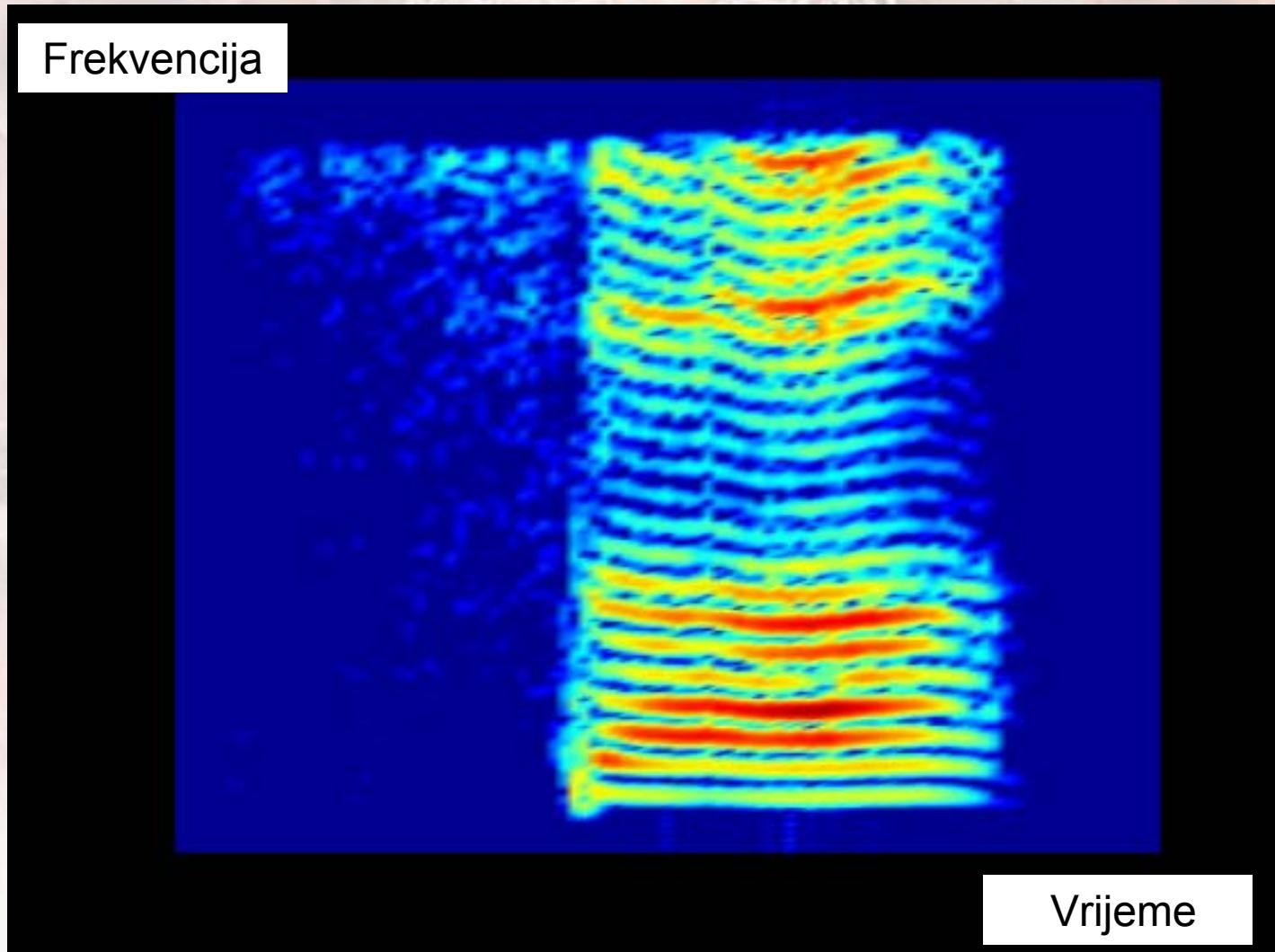


# Uskopojasni spektrogram govora

- Frekvencijska razlučljivost analize je prilagođena da se u spektrogramu očituje periodičnost signala vezana uz **dugotrajne korelacije**.
- Svaki pojedini vremenski otvor Fourierove analize mora zahvaćati više od dva glotalna pulsa.
- Dugotrajne korelacije se manifestiraju kao horizontalne pruge u uskopojasnom spektrogramu.
- Razmak pruga jednak je temeljnoj frekvenciji titranja glasnica (recipročna vrijednost perioda  $PP$ )
- Vertikalno kretanje ovih linija opisuje **promjenu temeljne frekvencije glasa**.



# Uskopojasni spektr. – glas “sa”



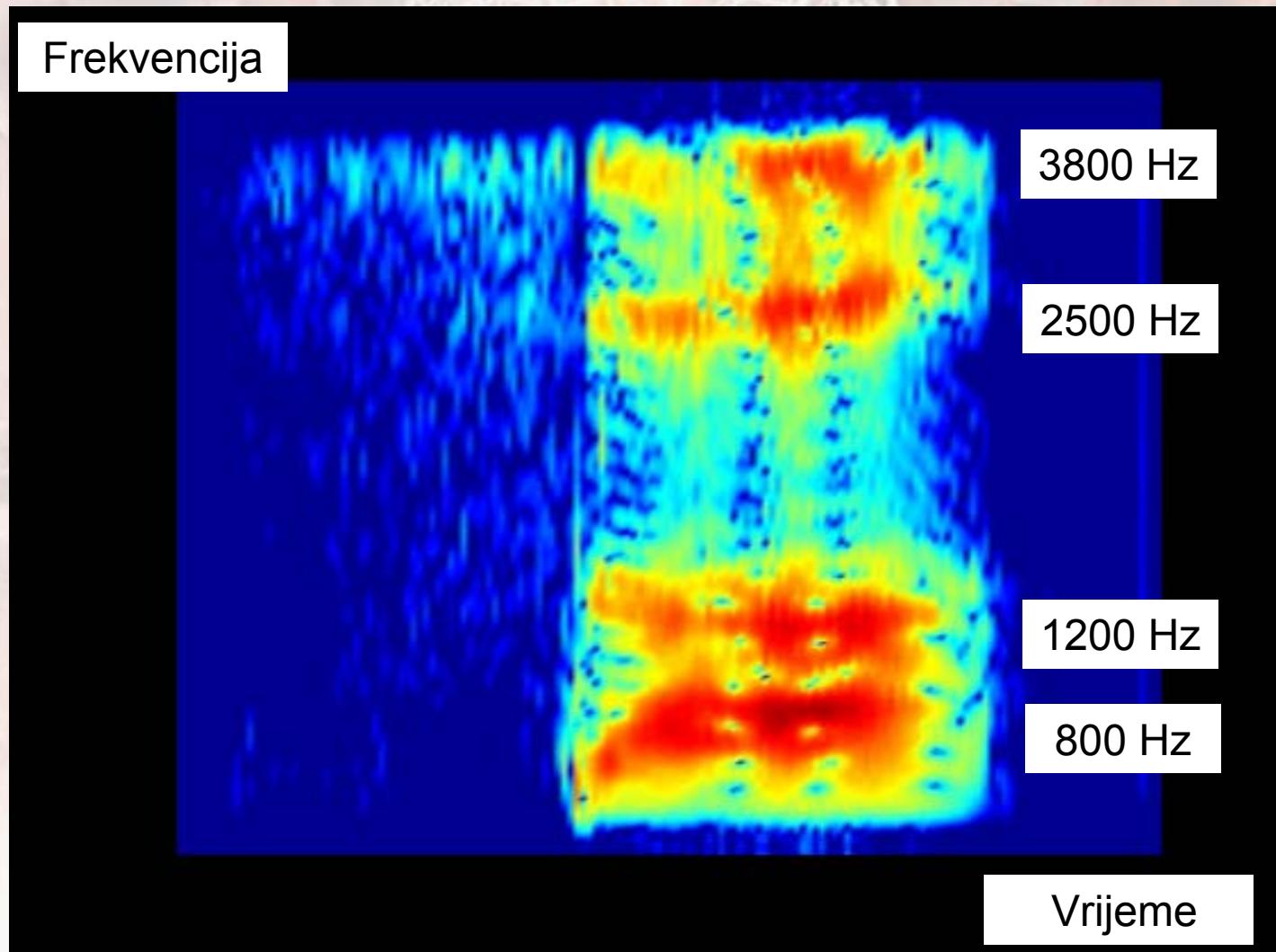


# Širokopojasni spektrogram govora

- Za vizualizaciju **kratkotrajnih korelacija** u govornom signalu koristi se širokopojasni spektrogram.
- Uklonjene su brze promjene duž frekvencijske osi koje su vezane uz periodični pobudni signal (titranje glasnica), jer je frekvencijska razlučljivost analize nedovoljna za razdvajanje pojedinih harmonika (spektralnih linija).
- U njemu su vidljiva spektralna nadvišenja koja odgovaraju **vremenskoj evoluciji formanata** tokom izgovora pojedinog glasa.



# Širokopojasni spektr. – glas “sa”





# Počeci modeliranja i kodiranja govora

- Istraživanja u području govora počela su vrlo rano.
- Teorija modeliranja i kodiranja govora rezultat je istraživanja fizičara i inženjera koji su već od 18 stoljeća analizirali i predlagali raznovrsne **fizikalne modele procesa nastajanja govora** u vokalnom traktu.
- Izgrađivali su i različite mehaničke i električke uređaje koji su pokušavali simulirati ove procese ... tj. sintetizirati govor.



# Mehanički sintetizator govora

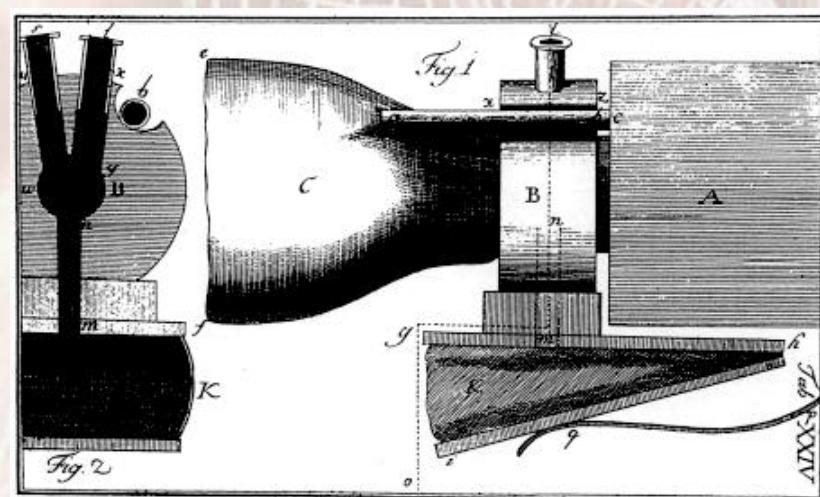
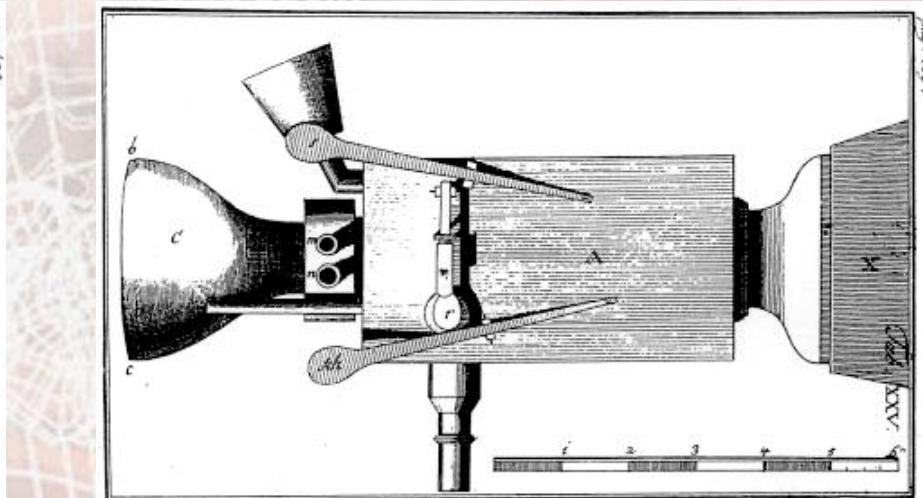
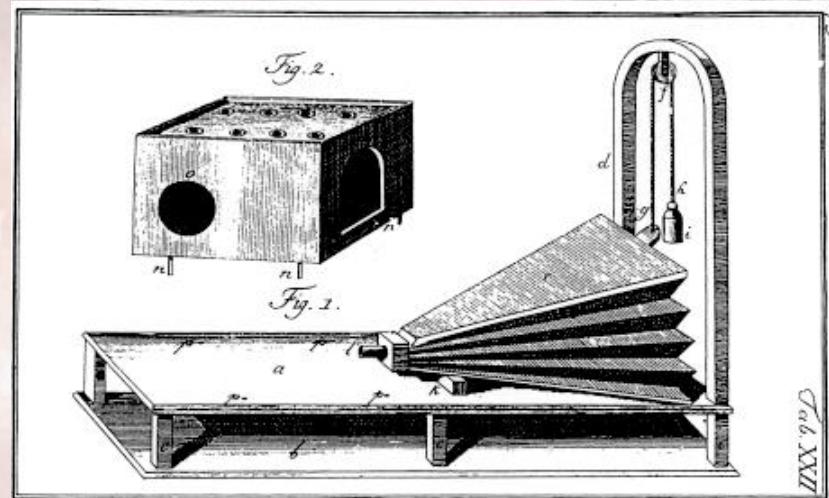


- 1791. gradi mehanički uređaj koji je mogao govoriti pod kontrolom operatera.
- Opisuje ga u knjizi: *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine*

- **Wolfgang von Kempelen**
- **1734-1804**

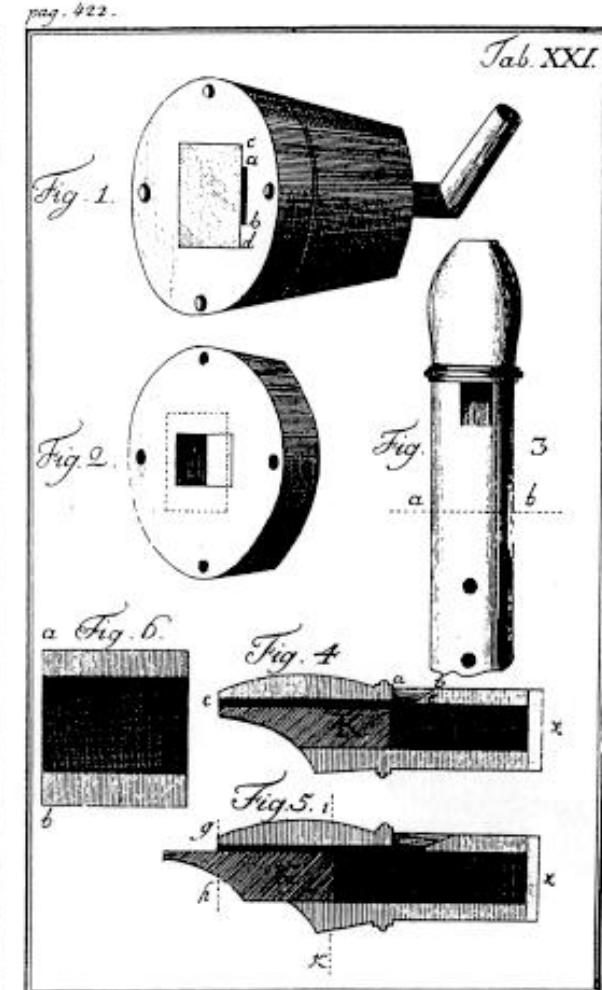
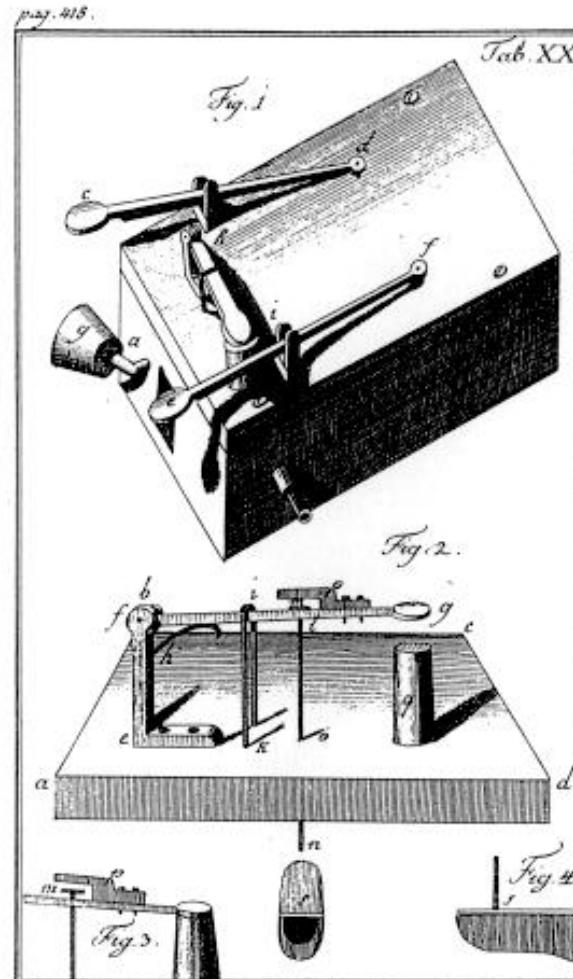
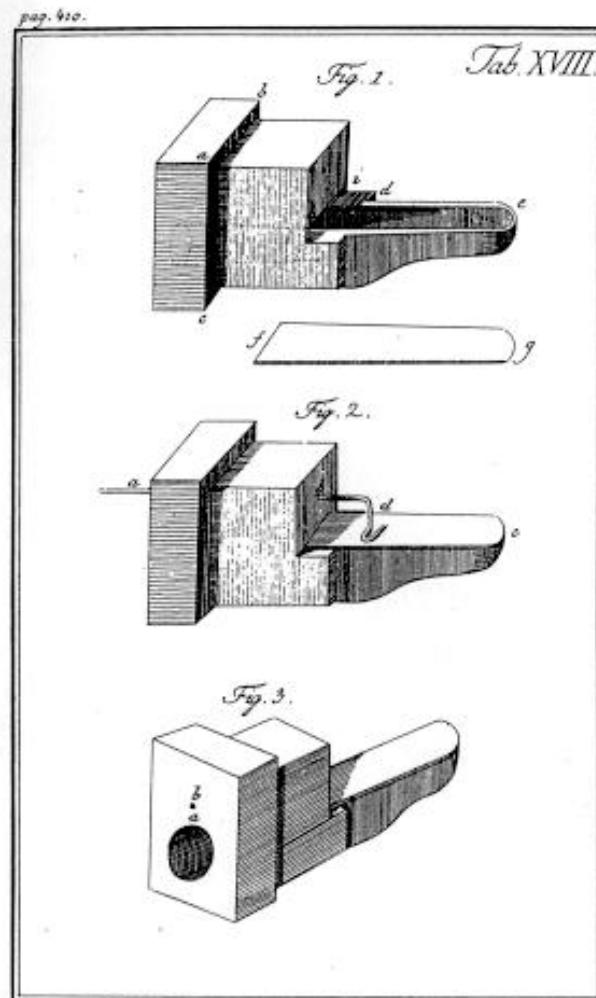


# Kempelenov sintetizator





# Kempelenov sintetizator





# Kempelenov sintetizator



- Originalni primjerak u:  
*Deutsches Museum  
von Meisterwerken der  
Naturwissenschaft und  
Technik in München*



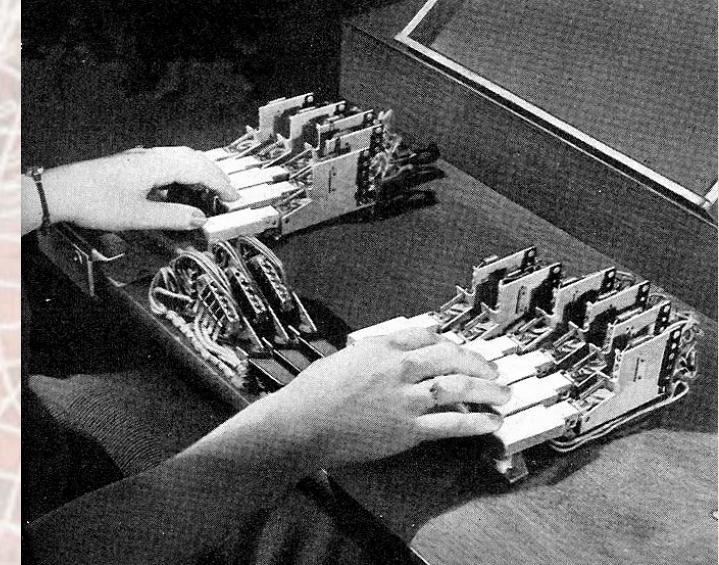
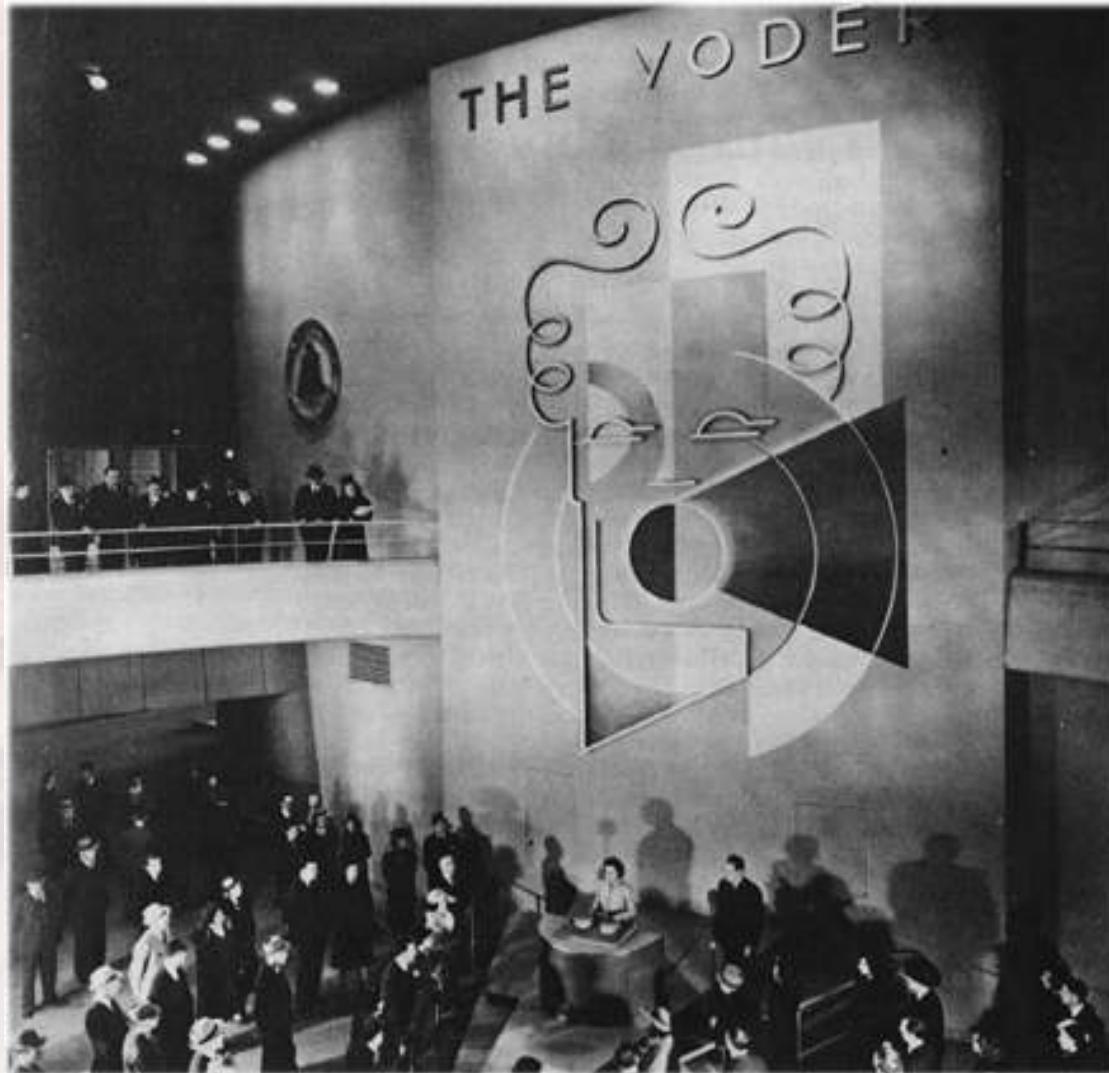
# Prvi električki sintetizator VODER



- **Homer Dudley**
- 1939. na Svjetskom sajmu (World Fair) u New Yorku prezentira potpuno električki sintetizator govora razvijen u "**Bell Telephone Laboratory**" pod nazivom **VODER** (**Voice Operated reCorDER**)
- Jedno od najvećih "čuda" tog doba.



# Prvi električki sintetizator VODER

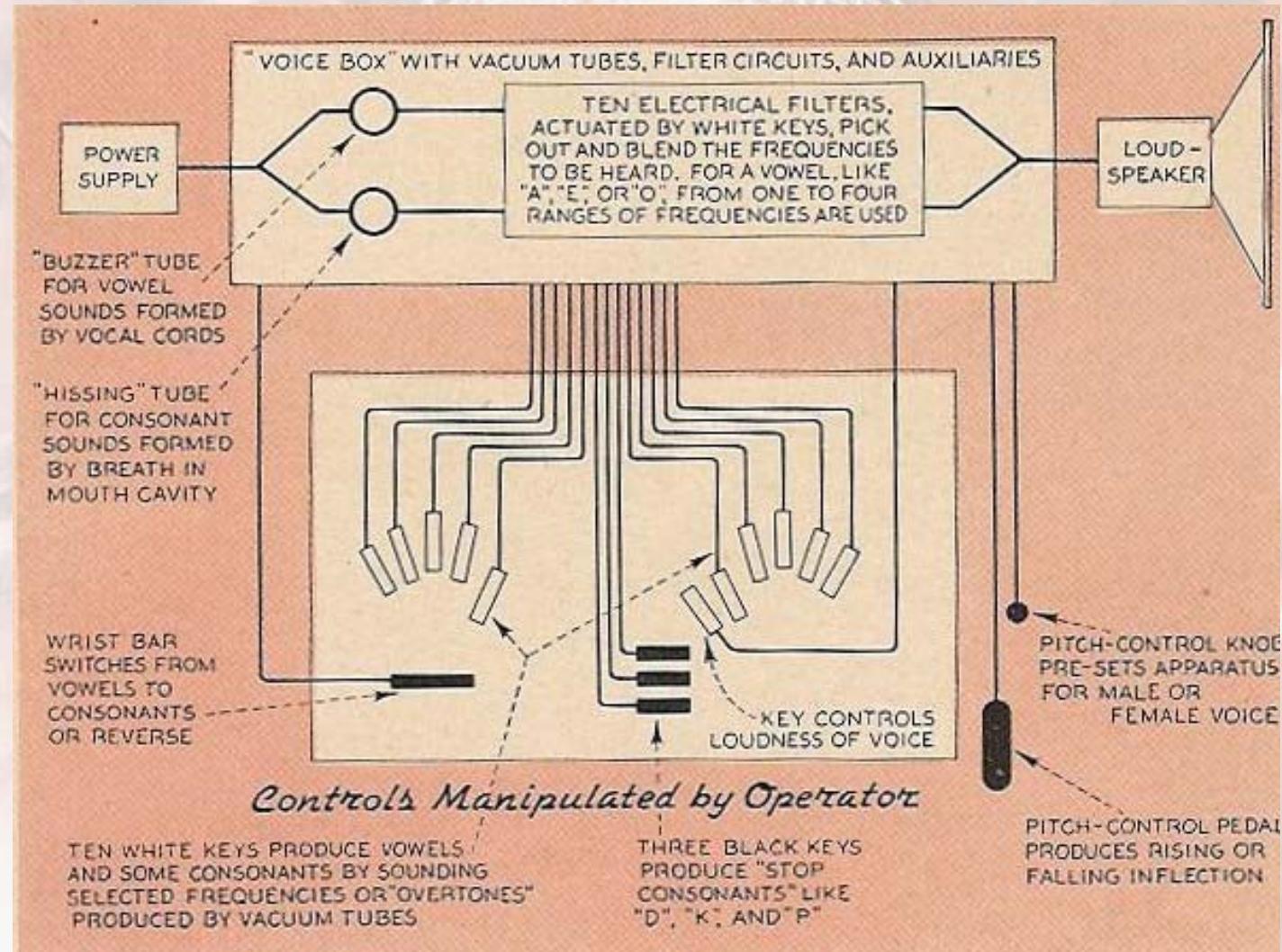


- Upravljanje tipkama
- Prezentacija na sajmu 1939



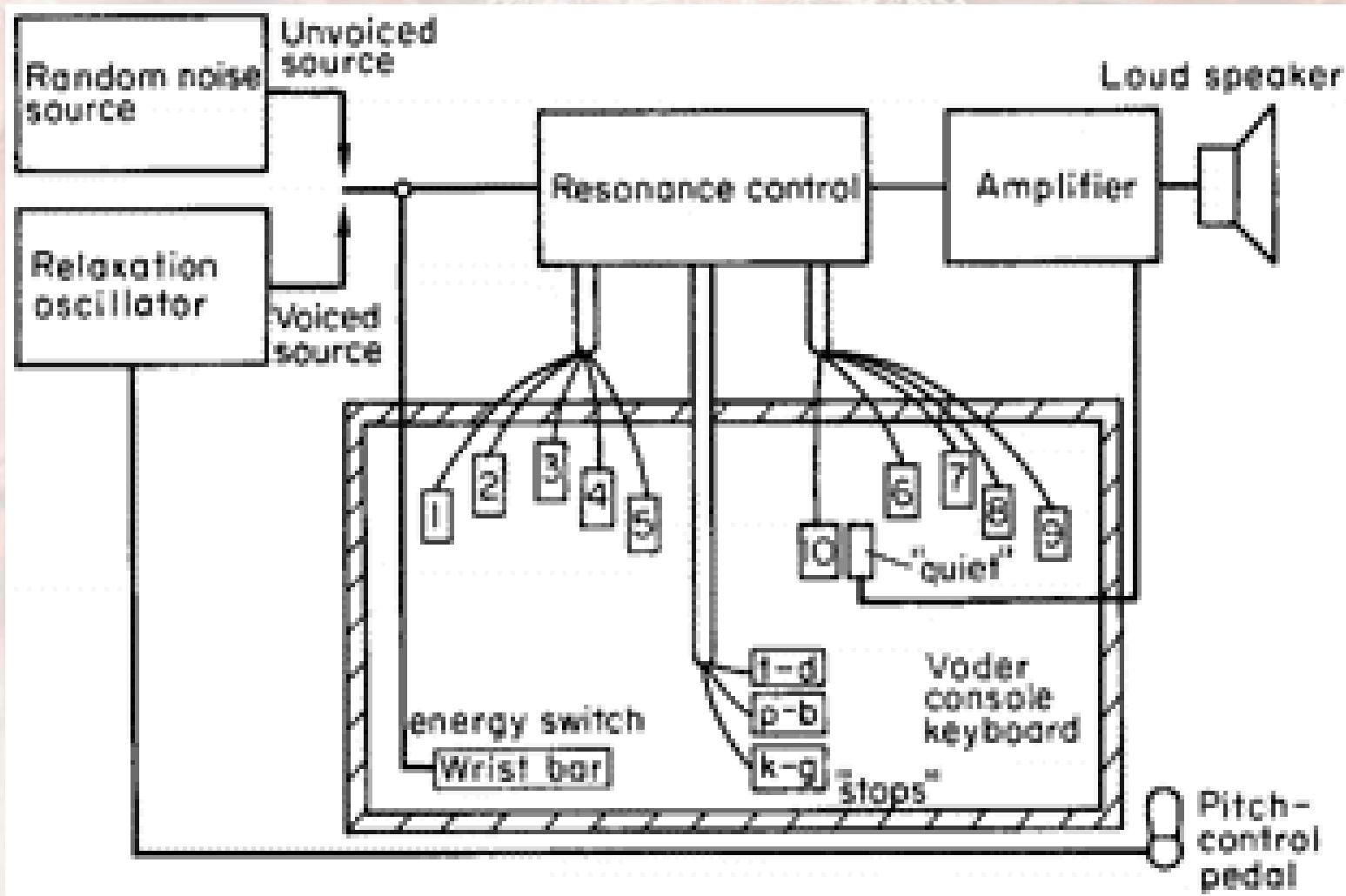


# Prvi električki sintetizator VODER





# Prvi električki sintetizator VODER





# Početci primjene linearne predikcije za govor

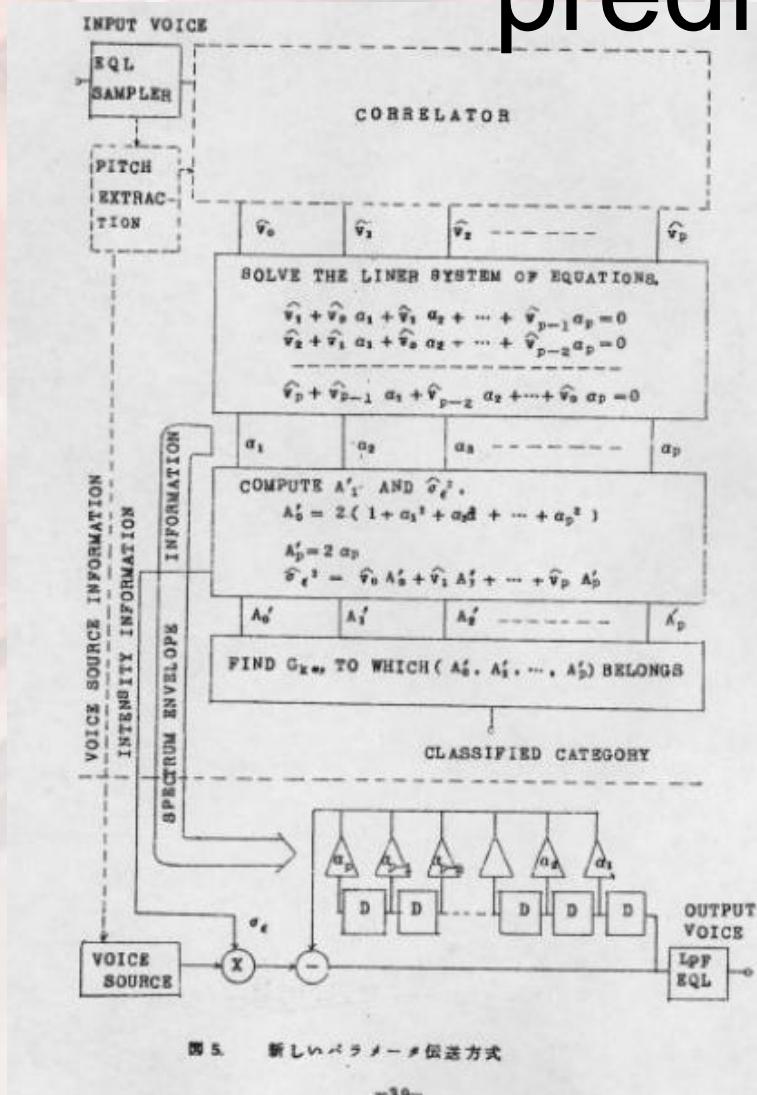


- **F.Itakura i S.Saito** radeći za NTT, Japan predlažu 1966 postupak kodiranja govora primjenom linearne predikcije
- Fumitada Itakura



# Početci primjene linearne predikcije za govor

- Slika iz rada:
  - S. Saito and F. Itakura,  
**“The theoretical consideration of statistically optimum methods for speech spectral density,”**  
Report No. 3107,  
Electrical Communication Laboratory, NTT, Tokyo,  
December 1966.



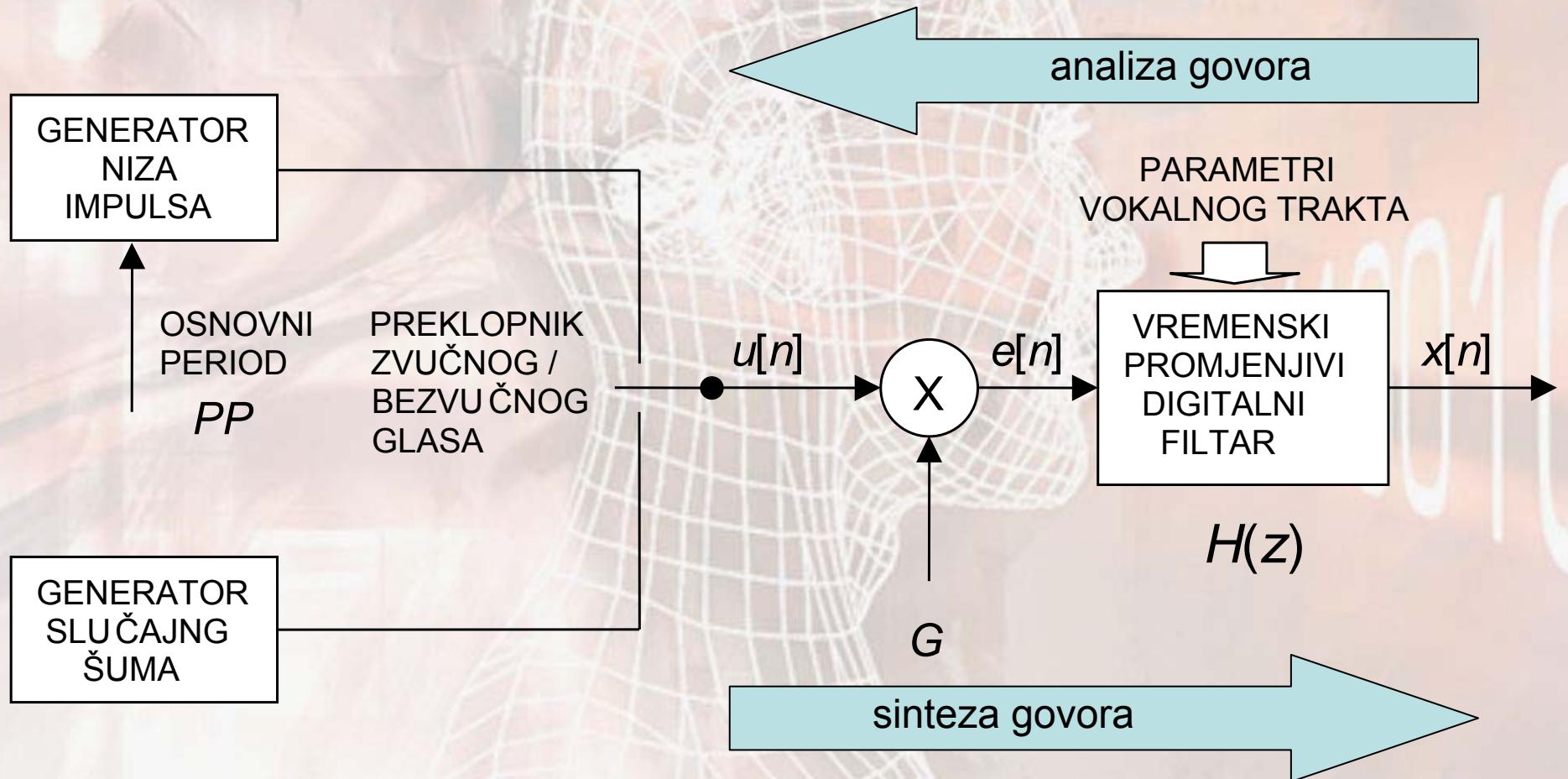


# LPC VOCODER

- Govor se može uspješno “oponašati” propuštanjem vremenski diskretnog pobudnog signala kroz digitalni filter s nekoliko rezonantnih karakteristika.
- Pobudni signal je:
  - periodični niz impulsa koji modelira signal na glasnicama,
  - ili šumovit signal koji modelira pobudu zvučnog trakta kod bezvučnih glasova.
- Parametri modela su:
  - svojstva filtra (pozicije i širine rezonantnih karakteristika),
  - amplituda i tip pobudnog signala (zvučni / bezvučni),
  - period impulsa pobude za zvučne glasove.



# LPC VOCODER





# LPC VOCODER

- Vocoder ne pokušava točno reprezentirati valni oblik govora!
- Njegova jedina težnja da govor sintetiziran opisanim modelom **zvuči što sličnije** ulaznom govornom signalu.
- Obzirom da se sami uzorci govora uopće ne kodiraju, već se za sintezu na strani dekodera koriste isključivo parametri ovog modela, Vocoder se ubraja u klasu parametarskih kodera (engl. **parametric source coders**)



# LPC VOCODER

- Sustav  $H(z)$  “oponaša” vokalni trakt, tj. vremenski kratkotrajne korelaciјe u govornom signalu.
- Govor se mijenja, pa se i korelaciјe mijenjaju s vremenom (od glasa do glasa).
- Filtar  $H(z)$  mora biti promjenjiv i kontinuirano modelirati promjene korelacija u signalu.
- I preostali parametri: (sklopka zvučno / bezvučno, pojačanje  $G$  i period  $PP$  zvučnih glasova) se moraju periodično ažurirati u skladu s promjenama ulaznog govora.



# Analiza i sinteza u parametarskom koderu govora

- Postupak određivanja “najboljih” vrijednosti parametara modela na osnovu kratkog segmenta ulaznog govornog signala zove se **analizom govora**, tj. **postupkom estimacije** parametara modela.
- **Sinteza** je obrnut postupak gdje se kratki segment govornog signala formira propuštanjem odabrane pobude kroz filter sa odabranim parametrima.
- Analiza se radi u sklopu kodera, koji kvantizirane parametre modela kodira i šalje dekoderu koji zatim provodi opisani postupak sinteze.



# Primjena linearne predikcije za estimaciju parametara Vocodera

- Kuharica ...:
  - Primjeni postupak linearne predikcije na kratkom segmentu govornog signala (20-30ms) i pronađi prediktor  $P(z)$  reda  $p$ , koji minimizira energiju predikcijske pogreške, opisan koeficijentima:  $\bar{\alpha}_1, \bar{\alpha}_2, \dots, \bar{\alpha}_p$ .
  - Pomoću tog prediktora izračunaj signal predikcijske pogreške  $e[n]$ , te odredi njegovu varijancu:  $\sigma_e^2(\bar{\alpha})$ .
  - Kao faktor pojačanja modela odaberi drugi korijen iz varijance: ...  $G = \sigma_e(\bar{\alpha})$



# Primjena linearne predikcije za estimaciju parametara Vocodera

- Kuharica ...nastavak:
  - Istraži da li u signalu predikcijske pogreške postoje periodičnosti, te ako postoje odredi njihov period  $PP$ .
  - Položaj sklopke zvučno/bezvučno, postavi ovisno o identificiranoj periodičnosti.
  - U tu svrhu se najčešće izračunava **autokorelacija** signala predikcijske pogreške te se traži pomak uz koji se postiže dobro poklapanje signala sa samim sobom.



# Primjena linearne predikcije za estimaciju parametara Vocodera

- Kuharica ...nastavak:
  - Estimirani parametri modela kvantiziraju i kodiraju se na pogodan način te ubacuju u izlaznu kodiranu poruku.
  - Cijeli postupak ponavlja se na novom segmentu govora, uz pomak “točke fokusa” analize od tipično 10 ili 20 ms.
  - Učestalost analize je recipročna vrijednost pomaka i uobičajeno se naziva engl. **frame rate** te iznosi 50 ili 100 okvira u sekundi.



# Postupak dekodiranja u Vocoderu

- Kuharica za dekoder:
  - Dekodiraj primljenu poruku i rekonstruiraj sve parametre modela.
  - Ovisno o informaciji o zvučnosti (periodičnosti) formiraj segment signala jedinične varijance koji se sastoji od:
    - pravilnog niza jediničnih impulsa na razmaku  $PP$ , ili
    - slučajnog bijelog šuma.
  - Pomnoži signal s faktorom pojačanja  $G$ , što daje “sintetički” signal  $e[n]$  na ulazu filtra  $H(z)$ .



# Postupak dekodiranja u Vocoderu

- Kuharica za dekoder ... nastavak:
  - Taj sintetički pobudni signal propusti kroz rekurzivni digitalni filter  $H(z)$  koji modelira vokalni trakt, sa kvantiziranim parametrima određenim u postupku analize:

$$\hat{\alpha}_k = Q(\bar{\alpha}_k)$$

$$H(z) = \frac{1}{1 - P(z)} = \frac{1}{1 - \sum_{k=1}^p \hat{\alpha}_k z^{-k}}$$

- Segment sintetiziranog signala pošalji na izlaz dekodera i ponovi postupak za novi okvir.

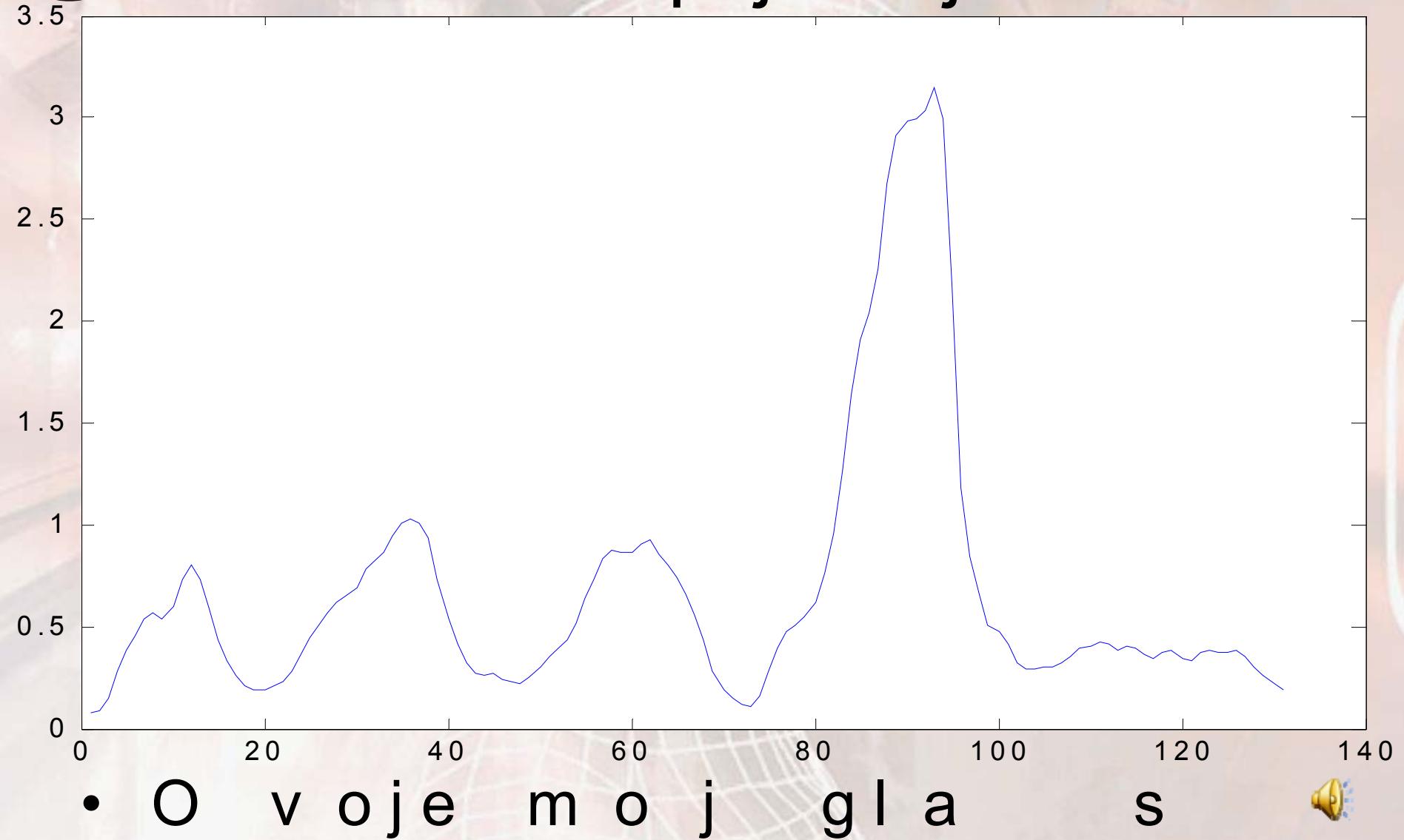


# Simulacija Vocoder-a u Matlabu

- Program korišten na Laboratorijskim vježbama iz Digitalne obrade govora:
  - frekvencija otipkavanja  $f_s=11050$  Hz,
  - širina otvora analize 46ms,
  - korak analize 10ms (100 okvira/s),
  - red prediktora  $p=12$ .
- Program radi analizu i sintezu, ali bez kvantizacije parametara modela – promatramo ograničenja samog modela!



# Faktor pojačanja $G$

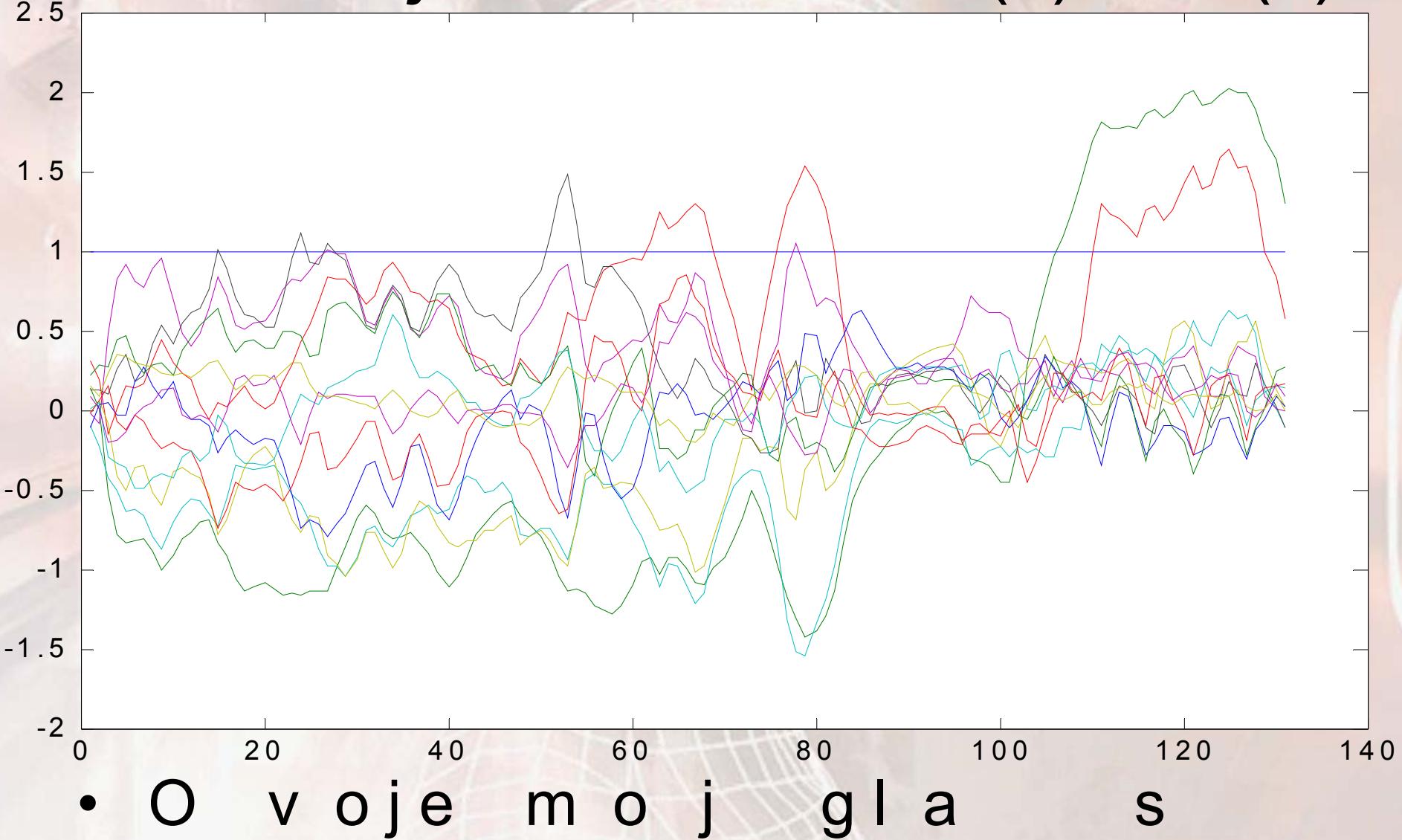


• O v o j e m o j g l a s



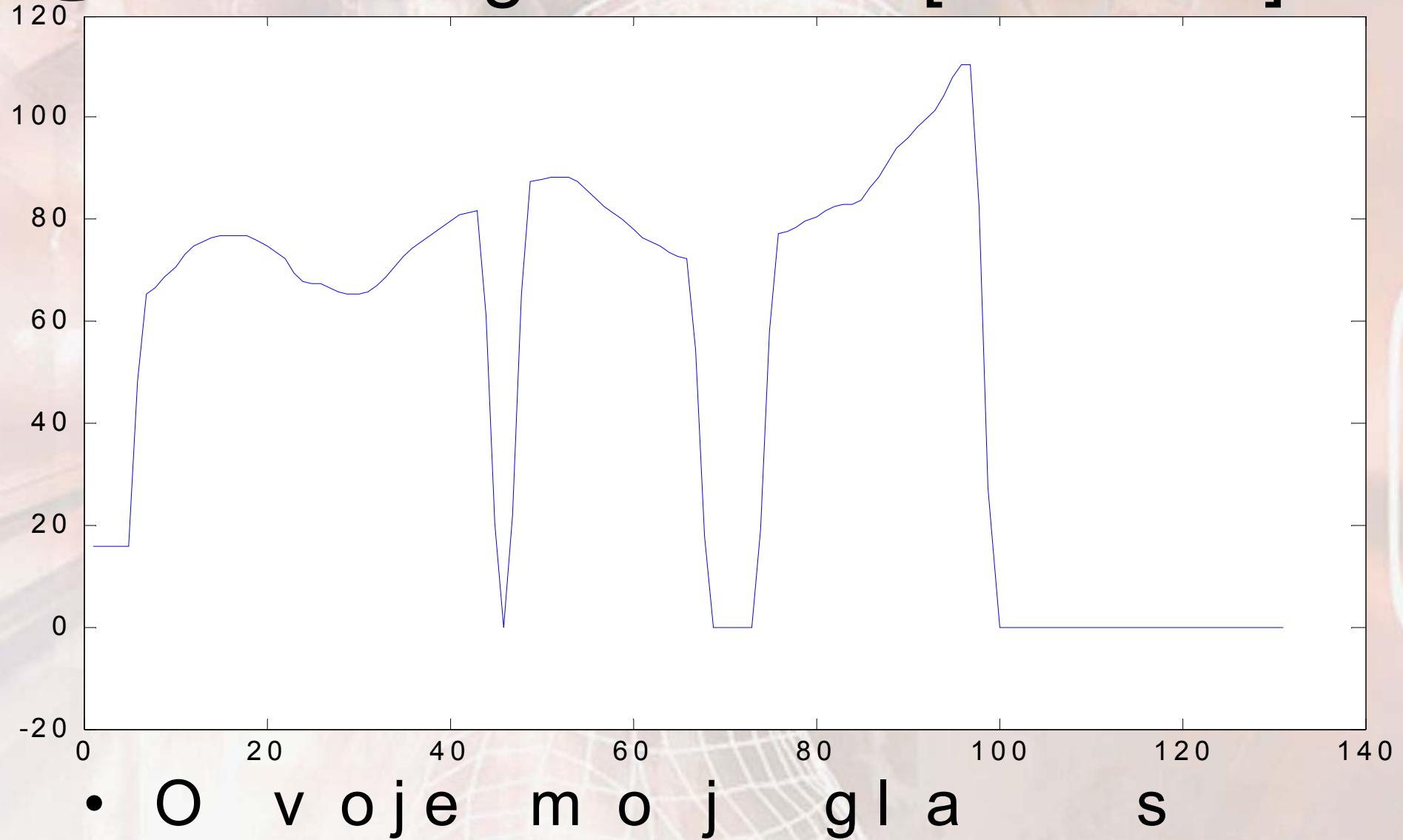


# Koeficijenti nazivnika $A(z)=1-P(z)$



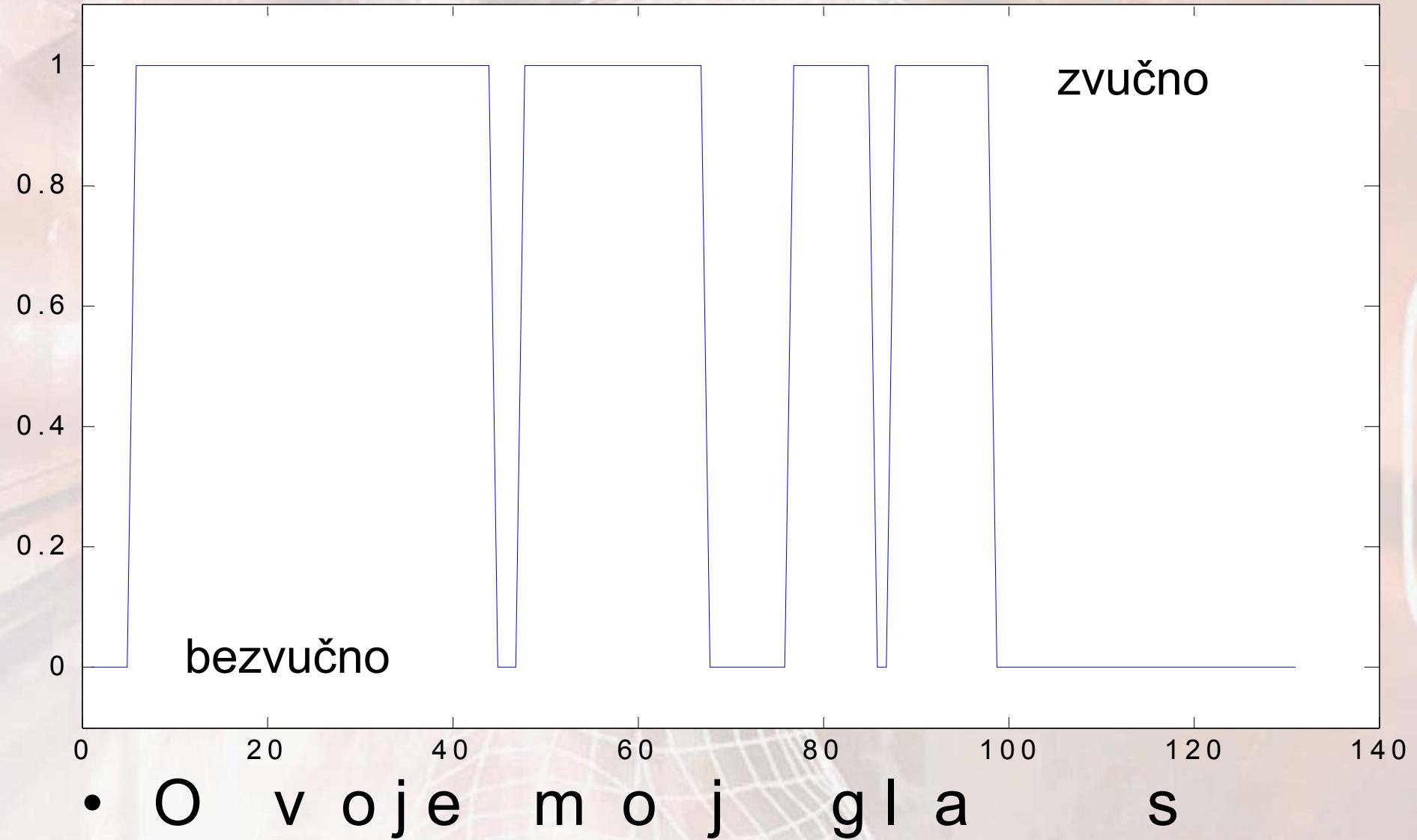


# Period glasnica *PP* [uzorcima]



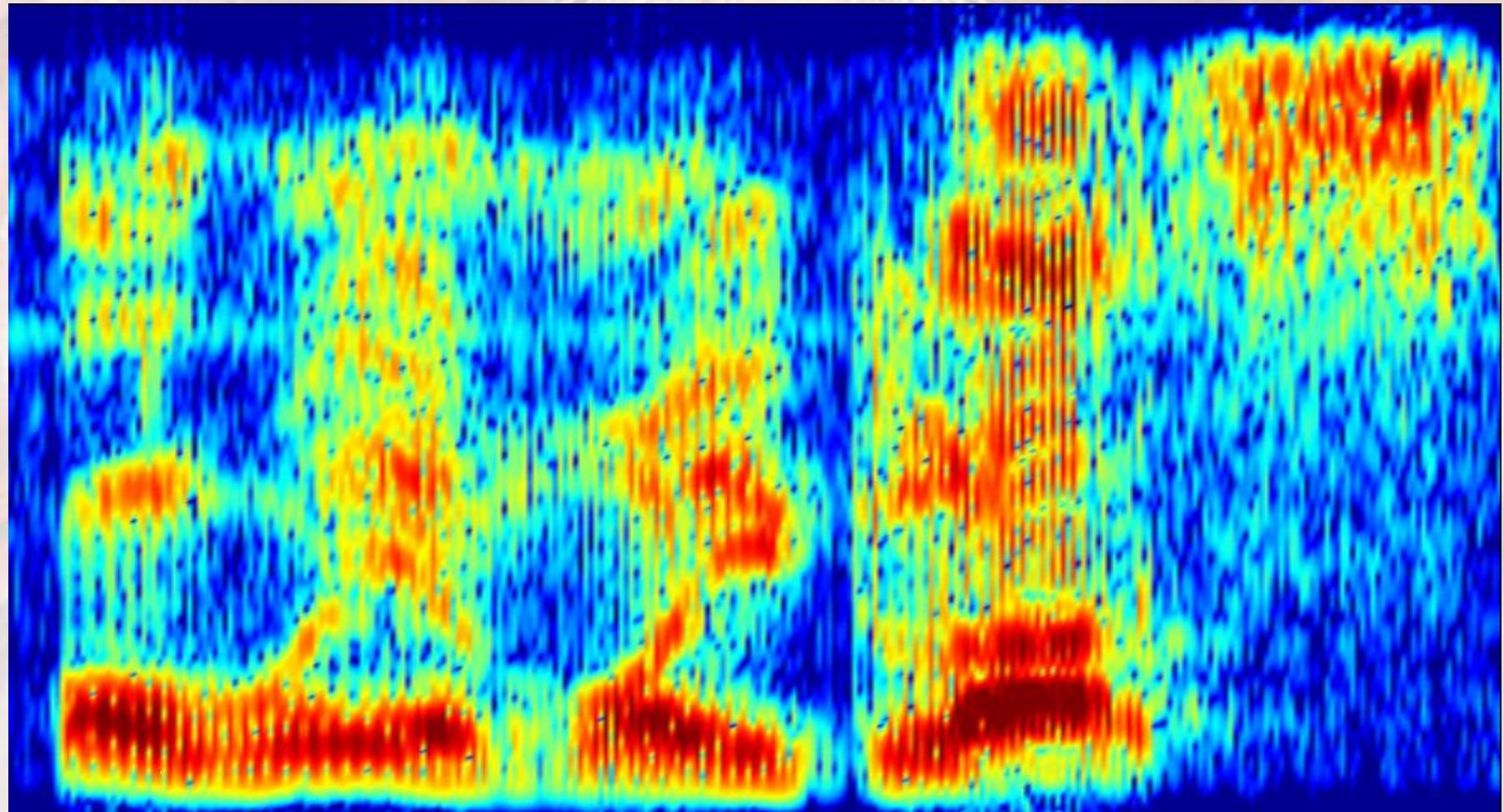


# Tip pobude zvučno/bezvučno





# Ulazni glas – širokopojasni sp.

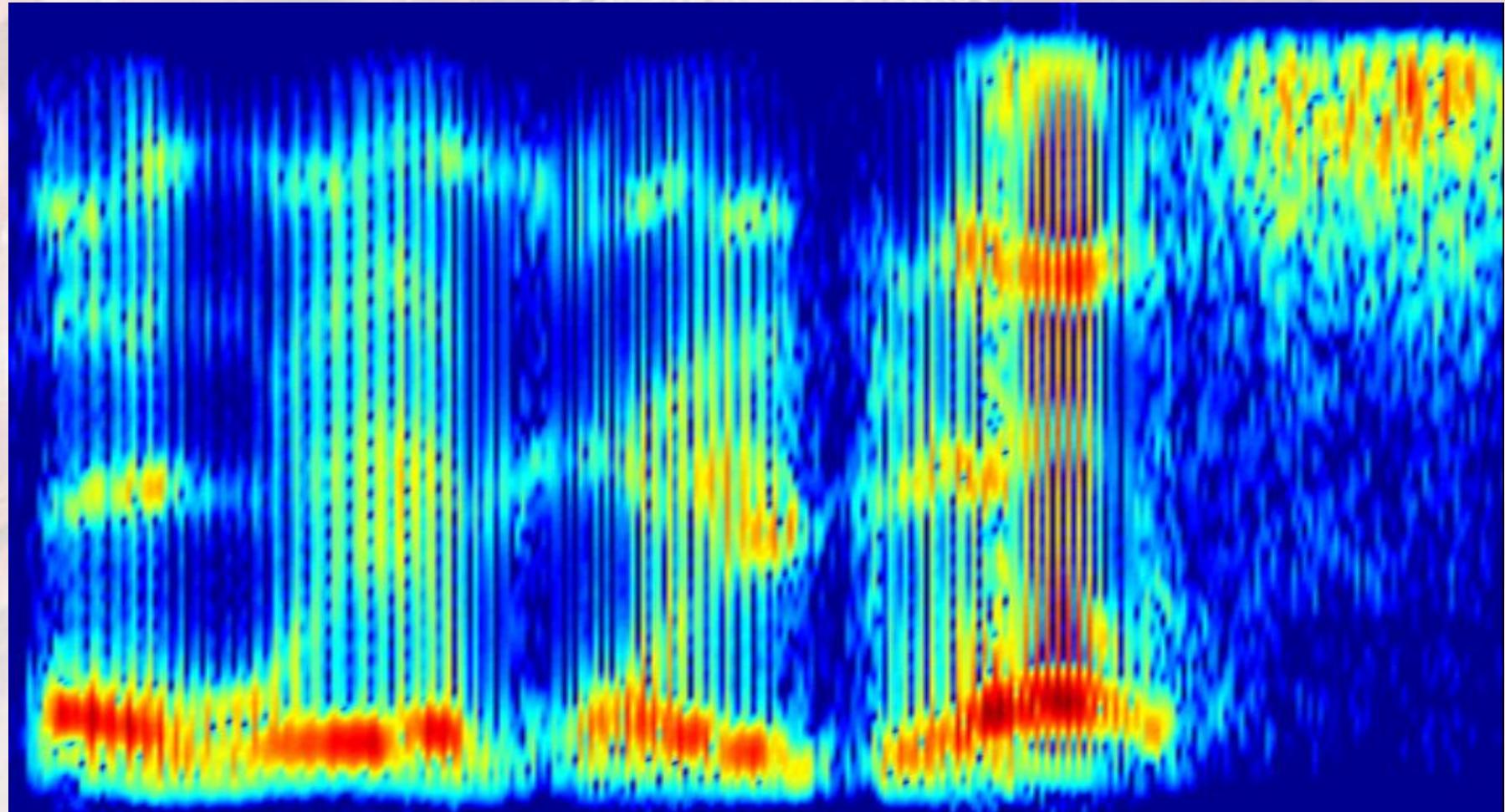


• O v o j e m o j g l a s





# LPC vocoder – širokopojasni sp.

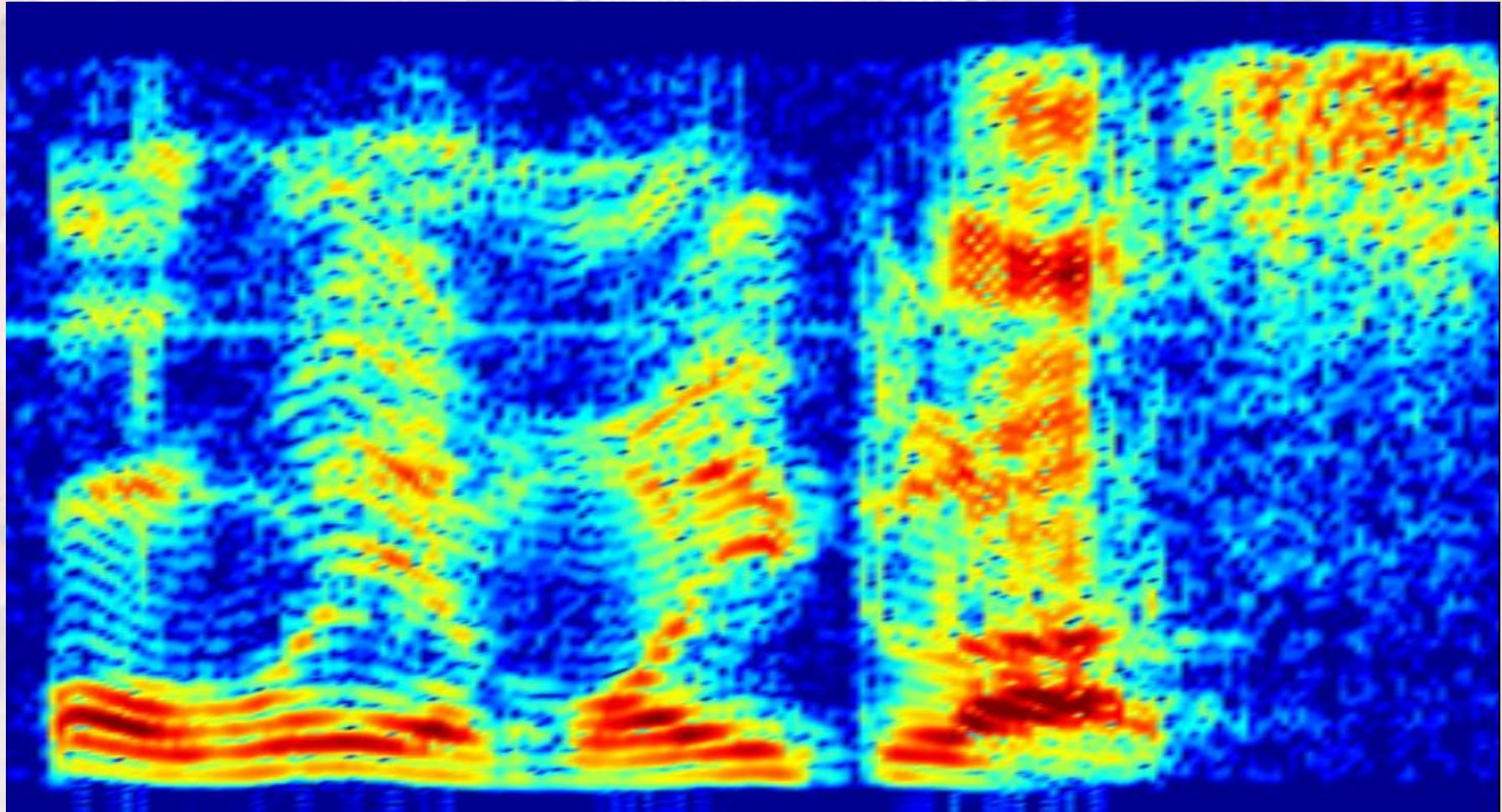


• O v o j e m o j g l a s





# Ulagni glas – uskopojasni sp.

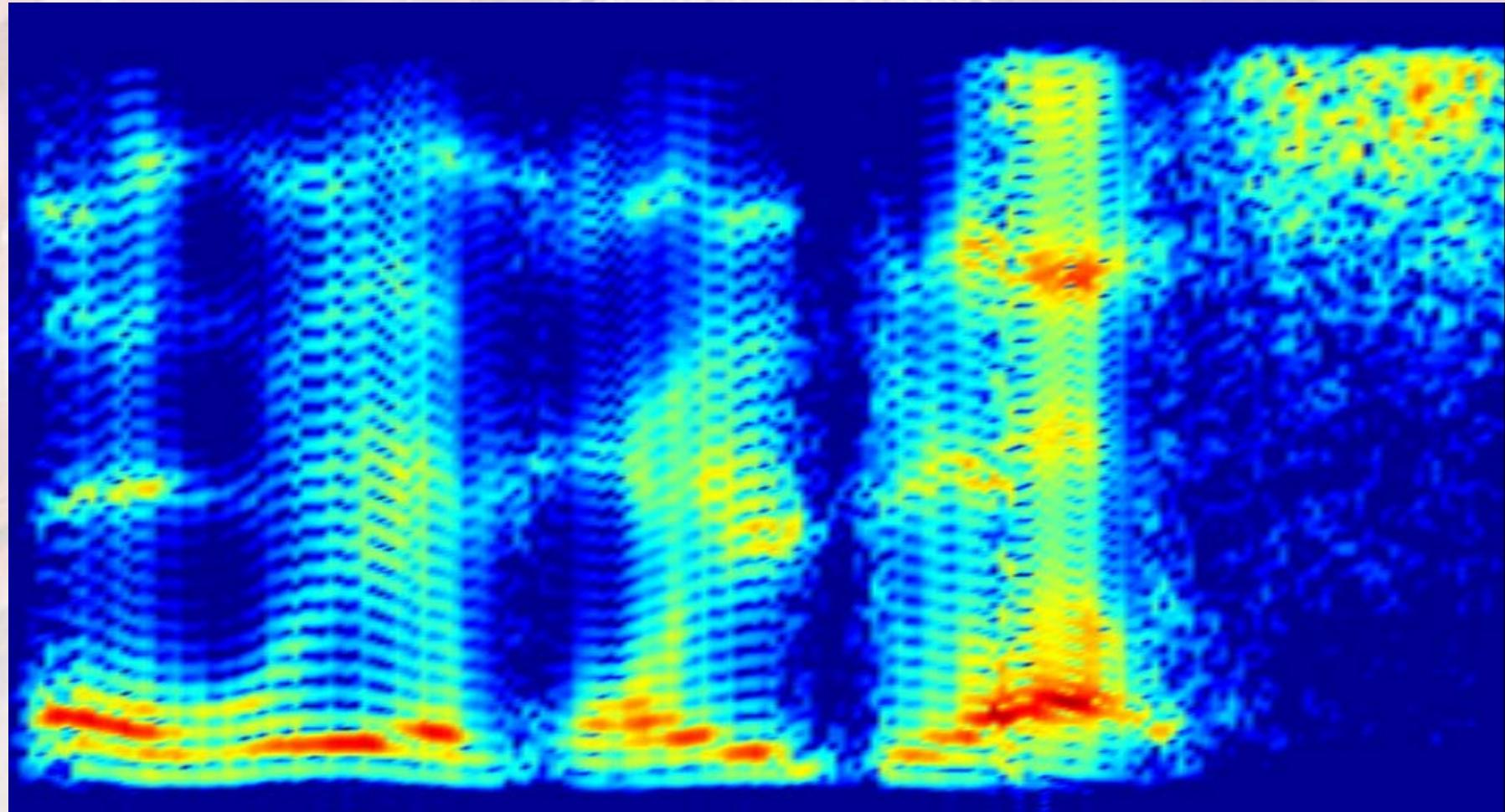


• O v o j e m o j g l a s





# LPC vocoder – uskopojasni sp.



• O v o j e m o j g l a s





# Simulacija Vocoder-a u Matlabu

- Govor povremeno zvuči sintetički, ali je razumljiv!
- Parametarski model omogućava raznovrsne manipulacije ovog signala:
  - ponavljamo svaki skup parametara dva puta i dobivamo dvostruko duži signal
  - izbacujemo svaki drugi skup parametara i dobivamo dvostruko kraći signal
  - lažno udvostručavamo period  $PP$
  - lažno prepoljavljamo period  $PP$



# LPC-10 (FS-1015)

- Jedan od prvih standarda za kodiranje govora temeljen je na Vocoder modelu.
- Standard razvijen za US DoD, kasnije preuzet od NATOa, za potrebe sigurnosne komunikacije.
- Koristi LPC prediktor 10 reda (... odatle ime).
- Ulazni govor otipkan sa  $f_s=8\text{kHz}$ .
- Podatkovni tok ulaznog signala nominalne točnosti je  $R=64\text{kbit/s}$ .
- Izlazni (kodirani) podatkovni tok  $R=2.4\text{kbit/s}$ ,
- ... ostvaruje sažimanje od 27x



# LPC-10 (FS-1015)

- Struktura okvira analize/sinteze u LPC-10
  - raspored bita, (engl. ***bit allocation***):

54 bits/frame

Pitch + U/V->7bits

G->5bits

K1 a K4->5bits

K5 a K8-> 4bits

K9->3bits

K10->2bits

Fs= 8000 samples/sec

54bits/frame

180 samples/frame

(22.5 ms/frame)

$$54 \times 8000 / 180 = 2400 \text{ bits/sec}$$

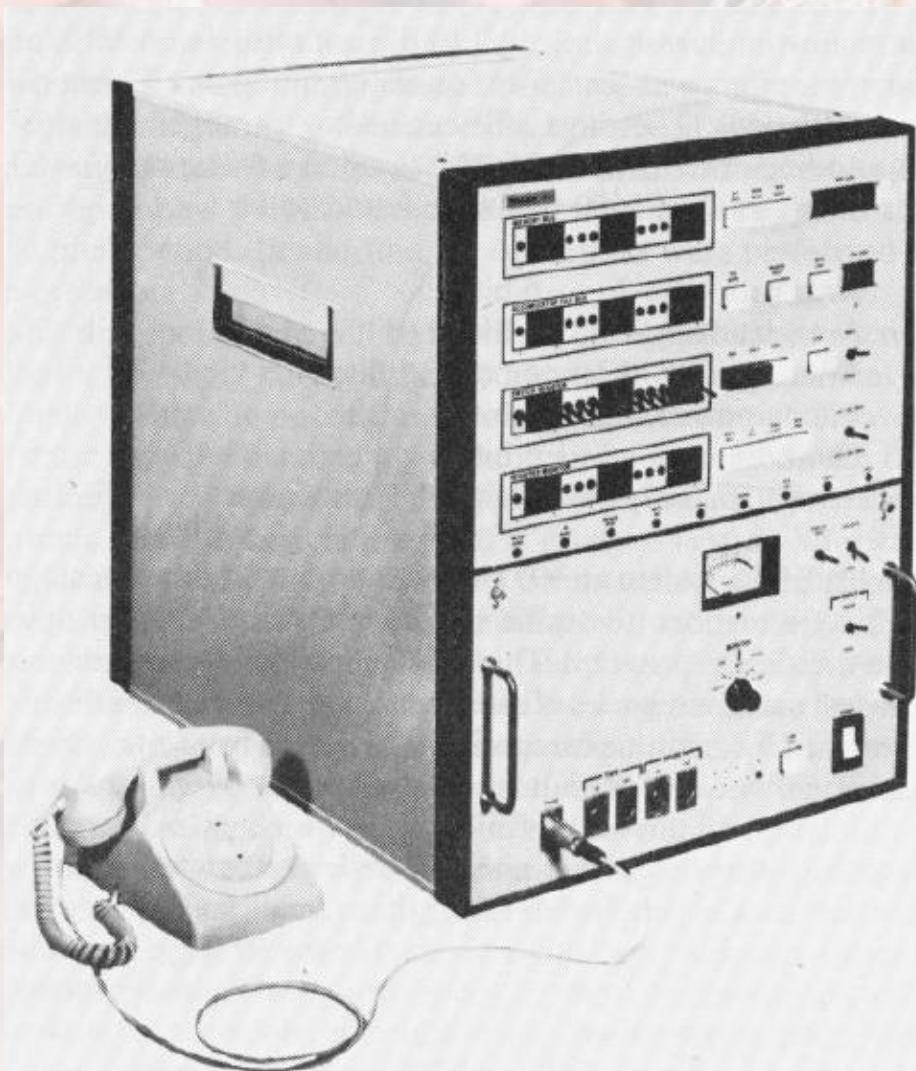


# LPC-10 (FS-1015)

- Izvorni Fortran i C kod za LPC-10 koder dostupan na više stranica.
- Npr.
  - [http://www.speech.cs.cmu.edu/comp.speech/  
Section3/Software/celp-3.2a.html](http://www.speech.cs.cmu.edu/comp.speech/Section3/Software/celp-3.2a.html)
- Kvaliteta govora je prilično loša, ali i dan-danas spada u kategoriju kodera sa vrlo niskim izlaznim podatkovnim tokom ...
  - novije inačice rade sa tokom od 1.2kbit/s



# LPC koder za rad u stvarnom vremenu



- Philco-Ford,  
**LONGBRAKE II**, 1974
- Prodana 4 primjerka  
(Navy, NSA)
- Teži 120kg
- Koristi procesor za  
obradu signala **PFSP**  
**(Philco-Ford Signal  
Processor)**



# Vocoder

- Vocoder korišten 1973 od grupe Kraftwerk na albumu "Ralf und Florian"





# Sigurnosni telefoni STU-III

- Temeljeni na koderu FS1016 novije generacije tipa CELP

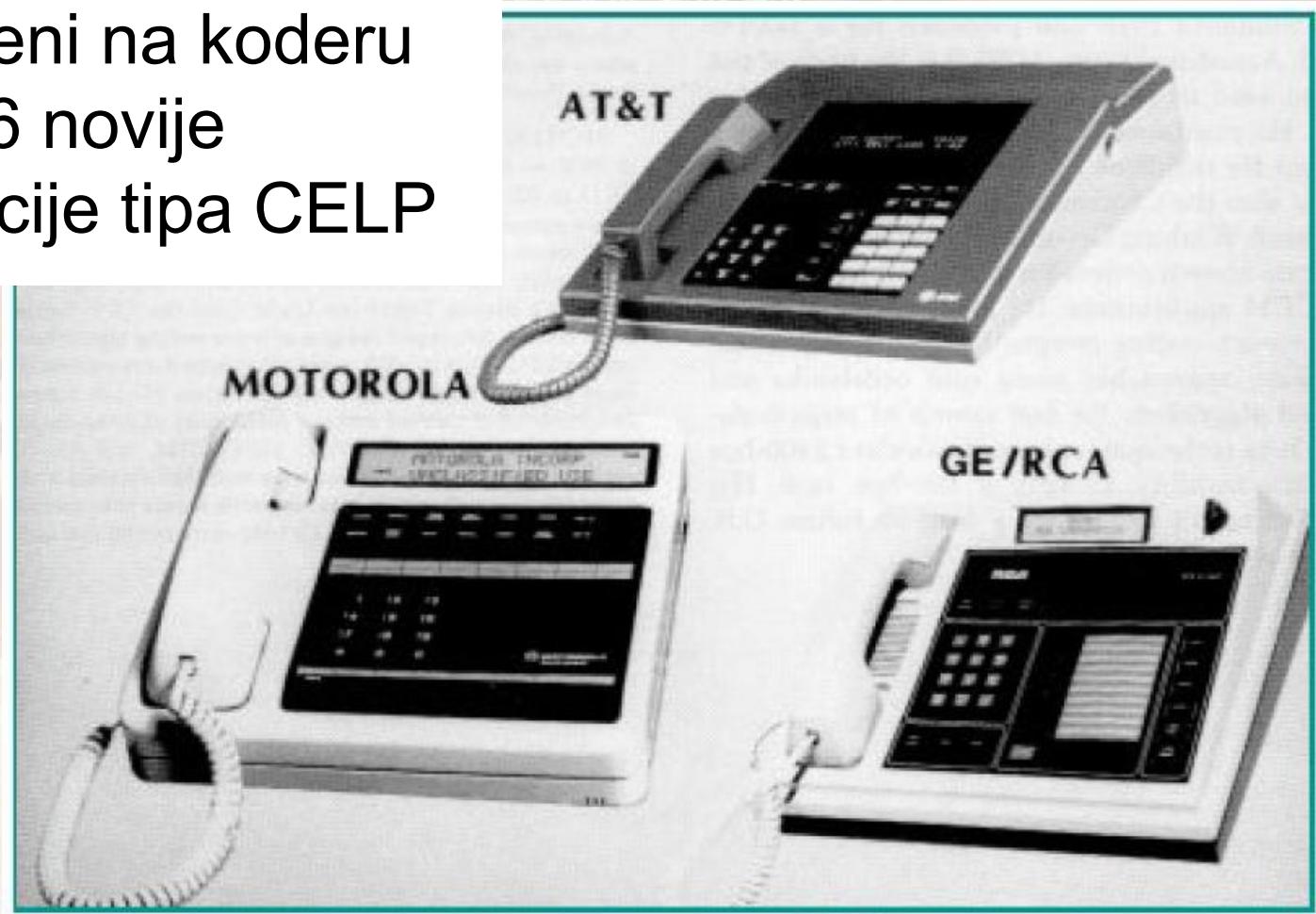


Fig. 4. The STU-III secure voice terminal family, circa 1986



# Što smo naučili

- korelacije u govoru i njihovi izvori
- formanti u govoru
- spektrogram; širokopojasni, uskopojasni
- mehanički sintetizator govora
- električki sintetizator, Voder
- parametarsko kodiranje, Vocoder
- analiza i sinteza kod parametarskih kodera
- primjena LPCa u Vocoderu
- primjer Vocoder-a
- LPC-10 standard kodiranja govora
- primjene i stvarni uređaji