

Razpoznavanje vzorcev

UN2-1-AV 2024/2025

Značilke vzorcev

Simon Dobrišek

POVZETEK

- Značilke vzorcev
- Približni zapis vzorcev z začetnimi členi ortogonalnih transformirank
- Hevristične značilke izsekov govornega signala
- Hevristične značilke področij slike

Značilke vzorcev (1/2)

- Razpoznavanje vzorcev je „bitka“ proti računski kompleksnosti (*J.C. Simon*).
- Računsko zahtevnost zmanjšujemo, če uporabljam le tiste **bistvene lastnosti** objektov, ki poudarjajo posebnosti posameznih razredov vzorcev .
- Take lastnosti objektov oziroma njihovih „slik“ - vzorcev imenujemo značilke.
- Izoblikovali so se trije pristopi zapisa preprostih vzorcev oziroma področij zapletenih vzorcev.
- Pri prvih dveh se pri „luščenju“ značilk ne oziramo na razvrstitev vzorcev v razrede, pri tretjem pa to upoštevamo.

Značilke vzorcev (2/2)

- **Prvi pristop** predpostavlja informacijsko bogat začetni opis objektov razpoznavanja, ki ga zgostimo s **približnim zapisom** (**aproksimacijo**) vzorca z določenim številom začetnih členov izbrane **ortogonalne razvrstitev**.
- **Drugi pristop** temelji na izkušnjah strokovnjakov za določeno področje uporabe. Kakovost tako določenih (**hevrističnih**) značilk ocenimo s poskusom razpoznavanja vzorcev ali z merami ločljivosti razredov vzorcev.
- **Tretji pristop** sloni na informacijsko zelo bogatem začetnem opisu objektov razpoznavanja, ki ga zgostimo tako, da **izločimo odvečne podatke** s postopki, ki temeljijo na optimizaciji ustrezzo izbrane kriterijske funkcije.

PRIBLIŽNI ZAPIS VZORCEV Z ZAČETNIMI ČLENI ORTOGONALNIH TRANSFORMIRANK

- Uvod
- Transformacija Karhunena in Loeveja
- Hitra Fourierjeva transformacija
- Hitra transformacija Walsha in Hadamarda,
- Hitra Haarova transformacija

Uvod

- Postopki luščenja značilk temeljijo na predpostavki, da lahko vsak vzorec opišemo z vrednostmi r spremenljivk - **meritev**, ki jih obravnavamo kot krajevni vektor $\mathbf{y} = (y_1, \dots, y_r)^T$.
- Pri postopkih luščenja značilk, ki temeljijo na uporabi **ortogonalnih linearnih transformacij**, dosežemo zgostitev informacije tako, da zapišemo vzorec \mathbf{y} z $n \leq r$ **začetnih členov** izbrane ortogonalne razvrstitve.
- Linearno transformacijo, ki preslikuje vektorje \mathbf{y} v vektorje \mathbf{c} , definiramo z matrično enačbo:

$$\mathbf{c} = \mathbf{A}\mathbf{y},$$

kjer **A** označuje **transformacijsko matriko**, ki jo po vrsticah sestavljajo linearno neodvisni r -razsežni vektorji $\mathbf{a}_1^T, \dots, \mathbf{a}_n^T$, torej

$$\mathbf{A}_{n \times r} = \begin{bmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nr} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}$$

Koeficienti ortogonalne razvrstitve

- Vektor \mathbf{c} je vektor značilk, ki so določeni kot **koeficienti ortogonalne razvrstitve** $\mathbf{c} = (c_1, \dots, c_n)^T$.
- Podana linearna transformacija je ortogonalna oziroma ortonormalna, če velja $\mathbf{A}\mathbf{A}^T = \mathbf{I}$.
- Podano linearno transformacijo lahko zapišemo tudi kot sistem linearnih enačb:

$$c_i = \mathbf{a}_i^T \mathbf{y} = \sum_{j=1}^r a_{ij} y_j ; \quad i = 1, \dots, n .$$

- V primeru, ko je $n < r$, pri linearni transformaciji izvirni vzorec \mathbf{y} izrazimo s približkom $\tilde{\mathbf{y}}$, ki je določen s koeficienti ortogonalne razvrstitve

$$\hat{\mathbf{y}} = \mathbf{A}^T \mathbf{c} .$$

Napaka izražave

- Napako, ki jo pri tej izražavi naredimo, pa določimo kot srednjo kvadratno napako

$$\|\Delta_y\|^2 = \|y - \hat{y}\|^2 = \|y - A^T c\|^2 = \|y\|^2 - \|c\|^2.$$

- V primeru, ko je $n = r$, pri izražavi ne naredimo napake in je $y = \hat{y}$.
- Na tem področju se uporablja postopke približnega zapisa vzorcev, ki temeljijo na ortogonalnih razvrsttvah **Karhunena** in **Loeveja**, **Fourierja**, **Walsha** in **Hadamarda** ter na valčni **Haarovi transformaciji** idr.

Transformacija Karhunena in Loeveja

- Transformacijo Karhunena in Loeveja lahko definiramo kot postopek **minimizacije srednje kvadratne napake** približnega zapisa vzorcev.
- Za vsako polno množico ortonormiranih vektorjev $\{\mathbf{e}_j\}$ lahko zapišemo vzorec $\mathbf{y} = (y_1, \dots, y_r)^T$ kot

$$\mathbf{y} = \sum_{j=1}^r c_j \mathbf{e}_j$$

- Vzorec \mathbf{y} približno zapišimo tako, da pri $n < r$ nadomestimo koeficiente razvrstitve $(c_n + 1, c_n + 2, \dots, c_r)$ s konstantami b_j , ki imajo **enako vrednost** pri vseh \mathbf{y} , to je:

$$\hat{\mathbf{y}} = \sum_{j=1}^n c_j \mathbf{e}_j + \sum_{j=n+1}^r b_j \mathbf{e}_j$$

Napaka približnega zapisa vzorca

- S tem smo zgrešili napako približnega zapisa vzorca:

$$\Delta_y = \mathbf{y} - \hat{\mathbf{y}} = \sum_{j=1}^r c_j \mathbf{e}_j - \left(\sum_{j=1}^n c_j \mathbf{e}_j + \sum_{j=n+1}^r b_j \mathbf{e}_j \right) = \sum_{j=n+1}^r (c_j - b_j) \mathbf{e}_j$$

- ozziroma srednjo kvadratno napako preko vseh vzorcev:

$$\bar{\Delta}_y^2 = E\{\|\Delta_y\|^2\} = E\{\Delta_y^T \Delta_y\} = E\left\{ \sum_{j=n+1}^r \sum_{k=n+1}^r (c_j - b_j)(c_k - b_k) \mathbf{e}_j^T \mathbf{e}_k \right\}$$

$$\bar{\Delta}_y^2 = \sum_{j=n+1}^r E\{(c_j - b_j)^2\}$$

$$\mathbf{e}_j^T \mathbf{e}_k = \begin{cases} 1, & \text{če } j = k \\ 0, & \text{če } j \neq k \end{cases}$$

- Iščemo torej takšne konstante b_j in takšne ortonormirane vektorje \mathbf{e}_j , da bo srednja kvadratna napaka $\bar{\Delta}_y^2$ najmanjša.

Minimizacija napake približnega zapisa vzorca (1/3)

- Srednja kvadratna napaka bo najmanjša, še izberemo konstante b_j tako, da odvajamo $E\{(c_j - b_j)^2\}$ po b_j in odvode izenačimo z 0:

$$\frac{\partial}{\partial b_j} E\{(c_j - b_j)^2\} = -2(E\{c_j\} - b_j) = 0.$$

- Sledi $b_j = E\{c_j\}$, oziroma, ker je $c_j = \mathbf{e}_j^T \mathbf{y}$, velja

$$b_j = \mathbf{e}_j^T E\{\mathbf{y}\}.$$

- Sedaj poiščemo še ortonormirane bazne vektorje \mathbf{e}_j .

Ker lahko pišemo $c_j = \mathbf{y}^T \mathbf{e}_j$, je srednja kvadratna napaka

$$\begin{aligned}\bar{\Delta}_y^2 &= \sum_{j=n+1}^r E\{(c_j - E\{c_j\})^2\} = \sum_{j=n+1}^r E\{(\mathbf{e}_j^T \mathbf{y} - E\{\mathbf{e}_j^T \mathbf{y}\})(\mathbf{y}^T \mathbf{e}_j - E\{\mathbf{y}^T \mathbf{e}_j\})\} \\ &= \sum_{j=n+1}^r \mathbf{e}_j^T E\{(\mathbf{y} - E\{\mathbf{y}\})(\mathbf{y} - E\{\mathbf{y}\})^T\} \mathbf{e}_j = \sum_{j=n+1}^r \mathbf{e}_j^T K \mathbf{e}_j,\end{aligned}$$

Kovariančna matrika

Minimizacija napake približnega zapisa vzorca (2/3)

- Najmanjšo vrednost srednje kvadratne napake dobimo z minimizacijo pomožne Lagrangejeve funkcije

$$\mathcal{L}_{\bar{\Delta}_y^2} = \sum_{j=n+1}^r \mathbf{e}_j^T \mathbf{K} \mathbf{e}_j + \sum_{j=n+1}^r \lambda_j (1 - \mathbf{e}_j^T \mathbf{e}_j),$$

Langrangejevi
multiplikatorji

ki jo odvajamo bo baznih vektorjih in enačimo z nič

$$\frac{\partial \mathcal{L}_{\bar{\Delta}_y^2}}{\partial \mathbf{e}_j} = \frac{\partial}{\partial \mathbf{e}_j} \left[\sum_{j=n+1}^r \mathbf{e}_j^T \mathbf{K} \mathbf{e}_j + \sum_{j=n+1}^r \lambda_j (1 - \mathbf{e}_j^T \mathbf{e}_j) \right] = 2 (\mathbf{K} \mathbf{e}_j - \lambda_j \mathbf{e}_j) = 0.$$

- Najmanjšo vrednost Lagrangejeve funkcije dobimo torej z vektorji \mathbf{e}_j , ki zadoščajo enačbi:

$$(\mathbf{K} - \lambda_j \mathbf{I}) \mathbf{e}_j = \mathbf{0}$$

ozziroma enačbi

$$\mathbf{K} \mathbf{e}_j = \lambda_j \mathbf{e}_j.$$

Minimizacija napake približnega zapisa vzorca (3/3)

- Če v izrazu za napako nadomestimo $\mathbf{K}\mathbf{e}_j$ z $\lambda_j \mathbf{e}_j$ in upoštevamo ortonormiranost vektorjev \mathbf{e}_j , dobimo:

$$\bar{\Delta}_y^2 = \sum_{j=n+1}^r \mathbf{e}_j^T \mathbf{K} \mathbf{e}_j = \sum_{j=n+1}^r \mathbf{e}_j^T \lambda_j \mathbf{e}_j = \sum_{j=n+1}^r \lambda_j.$$

- Najmanjšo srednjo kvadratno napako približnega zapisa vzorca torej dosežemo, če ga zapišemo s koeficienti tistih lastnih vektorjev \mathbf{e}_j , ki ustrezano **največjim lastnim vrednostim** matrike \mathbf{K} .
- Preostale koeficiente zamenjamo s konstantami $b_j, j = n + 1, \dots, r$, kjer je $b_j = \mathbf{e}_j^T E\{\mathbf{y}\}$, ki pa imajo vrednost 0, če velja $E\{\mathbf{y}\} = \mathbf{0}$.

Koeficienti razvrstitve Karhunena in Loeveja

- Koeficiente razvrstitve Karhunena in Loeveja $c_j, j = 1, 2, \dots, n$ pri ($n \leq r$) tako določimo s transformacijo:

$$\mathbf{c} = \mathbf{A}(\mathbf{y} - E\{\mathbf{y}\})$$

- transformacijsko matriko \mathbf{A} pa sestavimo iz začetnih n lastnih vektorjev koordinatnih osi $\mathbf{e}_j, j = 1, 2, \dots, n$ pri ($n \leq r$),

$$\mathbf{A} = \begin{bmatrix} \mathbf{e}_1^T \\ \vdots \\ \mathbf{e}_n^T \end{bmatrix}$$

- ki so urejeni po padajočem vrstnem redu lastnih vrednosti matrike \mathbf{K} :

$$\lambda_1 \geq \lambda_2 \geq \dots \lambda_n \geq \dots \lambda_r ,$$

Računski zgled (1/5)

- Izračunajte transformacijo Karhunena in Loeveja za vzorce iz učne množice $\mathcal{U}_2 = \{S_8, \Omega\} = \{U_1, U_2\}$, kjer so

$$S_8 = \{(1,3)^T, (2,2)^T, (3,3)^T, (2,4)^T, (4,1)^T, (5,2)^T, (5,4)^T, (6,1)^T\},$$

$$\Omega = \{\omega_1, \omega_2\},$$

$$U_1 = \{(1,3)^T, (2,2)^T, (3,3)^T, (2,4)^T\}$$

$$U_2 = \{(4,1)^T, (5,2)^T, (5,4)^T, (6,1)^T\}.$$

- Pri računanju te transformacije ne upoštevamo oznake vzorcev, zato kovariančno matriko $\widehat{\mathbf{K}} = E\{(\mathbf{y} - \widehat{\mathbf{m}})(\mathbf{y} - \widehat{\mathbf{m}})^T\}$ in vektor povprečnih vrednosti vzorcev $\widehat{\mathbf{m}} = E\{\mathbf{y}\}$ ocenimo iz celotne množice vzorcev S_N

$$\widehat{\mathbf{m}} = E\{\mathbf{y}\} = \frac{1}{N} \sum_{\mathbf{y} \in S_N} \mathbf{y} = \frac{1}{8} \sum_{j=1}^8 \mathbf{y}_j = \frac{1}{8} (28, 20)^T = (3.5, 2.5)^T$$

Računski zgled (2/5)

- Izračunamo še oceno kovariančne matrike

$$\begin{aligned}\widehat{\mathbf{K}} &= E\{(\mathbf{y} - \widehat{\mathbf{m}})(\mathbf{y} - \widehat{\mathbf{m}})^T\} = E\{\mathbf{y}\mathbf{y}^T\} - \widehat{\mathbf{m}}\widehat{\mathbf{m}}^T = \left(\frac{1}{N} \sum_{\mathbf{y} \in S_N} \mathbf{y}\mathbf{y}^T \right) - \widehat{\mathbf{m}}\widehat{\mathbf{m}}^T = \\ &= \frac{1}{8} \begin{bmatrix} 1 \cdot 1 + 2 \cdot 2 + \dots & 1 \cdot 3 + 2 \cdot 2 + \dots \\ 3 \cdot 1 + 2 \cdot 2 + \dots & 3 \cdot 3 + 2 \cdot 2 + \dots \end{bmatrix} - \begin{bmatrix} 12.25 & 8.75 \\ 8.75 & 6.25 \end{bmatrix} \\ &= \frac{1}{8} \begin{bmatrix} 120 & 64 \\ 64 & 60 \end{bmatrix} - \begin{bmatrix} 12.25 & 8.75 \\ 8.75 & 6.25 \end{bmatrix} \\ &= \begin{bmatrix} 15 & 8 \\ 8 & 7.5 \end{bmatrix} - \begin{bmatrix} 12.25 & 8.75 \\ 8.75 & 6.25 \end{bmatrix} = \begin{bmatrix} 2.75 & -0.75 \\ -0.75 & 1.25 \end{bmatrix}.\end{aligned}$$

- Nato izračunamo lastne vrednosti ocene kovariančne matrike

$$|\widehat{\mathbf{K}} - \lambda \mathbf{I}| = \det \begin{pmatrix} 2.75 - \lambda & -0.75 \\ -0.75 & 1.25 - \lambda \end{pmatrix} = 0,$$

ki nas pripelje do kvadratne enačbe

$$(2.75 - \lambda)(1.25 - \lambda) - 0.75^2 = \lambda^2 - 4\lambda + 2.875 = 0$$

Računski zgled (3/5)

- Rešitvi kvadratne enačbe in s tem lastni vrednosti sta

$$\lambda_{1,2} = \frac{4 \pm \sqrt{4^2 - 4 \cdot 2.875}}{2} \Rightarrow \lambda_1 \approx 3.0607, \lambda_2 \approx 0.9393$$

- Lastna vektorja \mathbf{e}_j pri $j = 1, 2$ izračunamo iz enačbe:

$$(\hat{\mathbf{K}} - \lambda_j \mathbf{I}) \mathbf{e}_j = \mathbf{0}.$$

- Najprej izračunamo prvi lastni vektor $\mathbf{e}_1 = (e_{11}, e_{12})^T$

$$\begin{bmatrix} 2.75 - \lambda_1 & -0.75 \\ -0.75 & 1.25 - \lambda_1 \end{bmatrix} \begin{bmatrix} e_{11} \\ e_{12} \end{bmatrix} \approx \begin{bmatrix} -0.3107 & -0.75 \\ -0.75 & -1.8107 \end{bmatrix} \begin{bmatrix} e_{11} \\ e_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow$$
$$-0.3107 e_{11} - 0.75 e_{12} = 0$$
$$-0.75 e_{11} - 1.8107 e_{12} = 0$$

- Iz obeh enačb lahko izpeljemo zvezo $e_{11} \approx -2.4142 e_{12}$, kar pomeni, da so vsi vektorji $\hat{\mathbf{e}}_1$, katerih komponenti sta različni od nič in v podani zvezi, lastni vektorji matrike \mathbf{K} , ki pripadajo prvi lastni vrednosti λ_1 .

Računski zgled (4/5)

- Če izberemo vrednost $e_{12} = 1$ dobimo vektor $\hat{\mathbf{e}}_1 = (-2.4142, 1)^T$, ki ga še normiramo

$$\mathbf{e}_1 = \frac{\hat{\mathbf{e}}_1}{\|\hat{\mathbf{e}}_1\|} \approx \frac{\hat{\mathbf{e}}_1}{2.6131} \approx (-0.9239, 0.3827)^T$$

- Na podoben način določimo še drugi lastni vektor $\mathbf{e}_2 = (e_{21}, e_{22})^T$

$$\begin{bmatrix} 2.75 - \lambda_2 & -0.75 \\ -0.75 & 1.25 - \lambda_2 \end{bmatrix} \begin{bmatrix} e_{21} \\ e_{22} \end{bmatrix} \approx \begin{bmatrix} 1.8107 & -0.75 \\ -0.75 & 0.3107 \end{bmatrix} \begin{bmatrix} e_{21} \\ e_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow$$

$$1.8107 e_{21} - 0.75 e_{22} = 0$$

$$-0.75 e_{21} + 0.3107 e_{22} = 0$$

- Iz enačb izpeljemo zvezo $e_{21} \approx 0.4142 e_{22}$ in dobimo lastni vektor vektor $\hat{\mathbf{e}}_2 = (0.4142, 1)^T$, ki ga še normiramo

$$\mathbf{e}_2 = \frac{\hat{\mathbf{e}}_2}{\|\hat{\mathbf{e}}_2\|} \approx \frac{\hat{\mathbf{e}}_2}{1.0824} \approx (0.3827, 0.9239)^T$$

Računski zgled (5/5)

- Iz obeh lastnih vektorjev dobimo celotno transformacijsko matriko

$$\mathbf{A} = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{bmatrix} = \begin{bmatrix} -0.9239 & 0.3827 \\ 0.3827 & 0.9239 \end{bmatrix},$$

s transformacijo

$$\mathbf{c} = \mathbf{A}(\mathbf{y} - E\{\mathbf{y}\}) = \mathbf{A}(\mathbf{y} - \hat{\mathbf{m}})$$

pa vektorje značilk

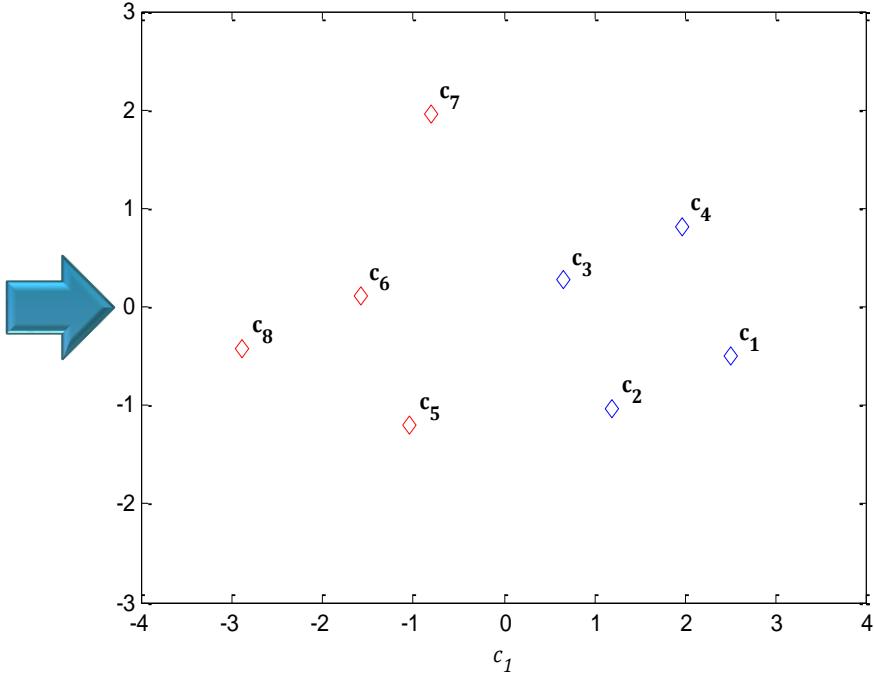
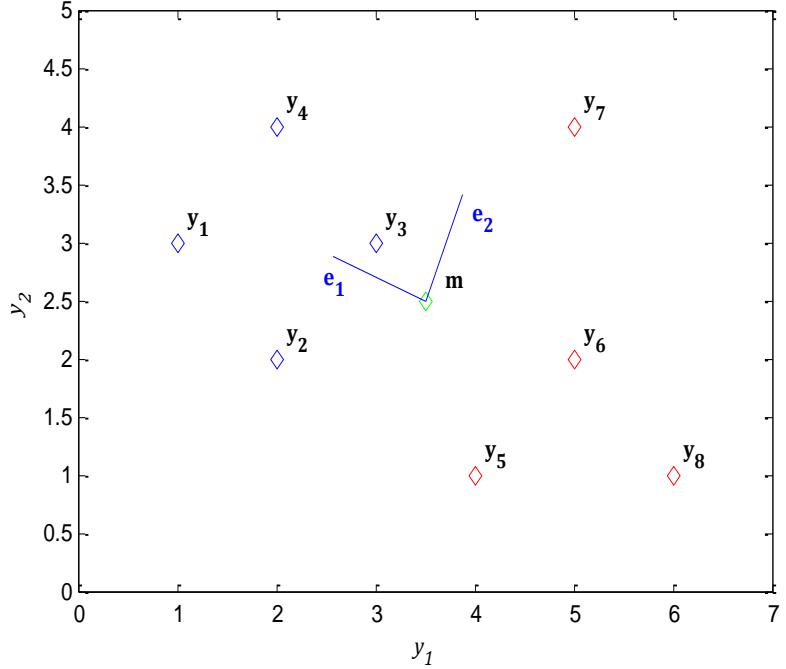
$$U'_1 \approx \{(2.50, -0.49)^T, (1.19, -1.04)^T, (0.65, 0.27)^T, (1.96, 0.81)^T\}$$

$$U'_2 \approx \{(-1.04, -1.19)^T, (-1.58, 0.11)^T, (-0.81, 1.96)^T, (-2.88, -0.43)^T\}.$$

- Povprečni vektor teh vektorjev značilk je **0**, kovariančna matrika pa je diagonalna matrika, ki po diagonali vsebuje lastne vrednosti, torej

$$\hat{\mathbf{K}} = \begin{bmatrix} 3.0607 & 0 \\ 0 & 0.9393 \end{bmatrix}.$$

Računski zgled – ponazoritev preslikave



$$\mathbf{c} = \mathbf{A}(\mathbf{y} - E\{\mathbf{y}\}) = \mathbf{A}(\mathbf{y} - \hat{\mathbf{m}})$$

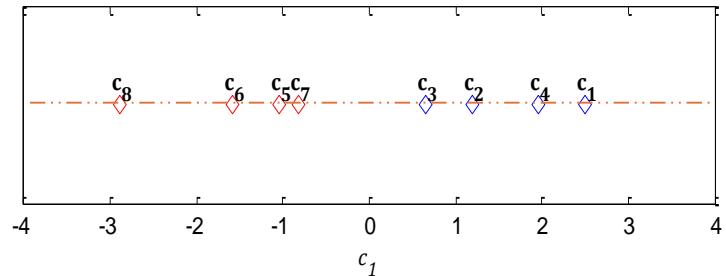
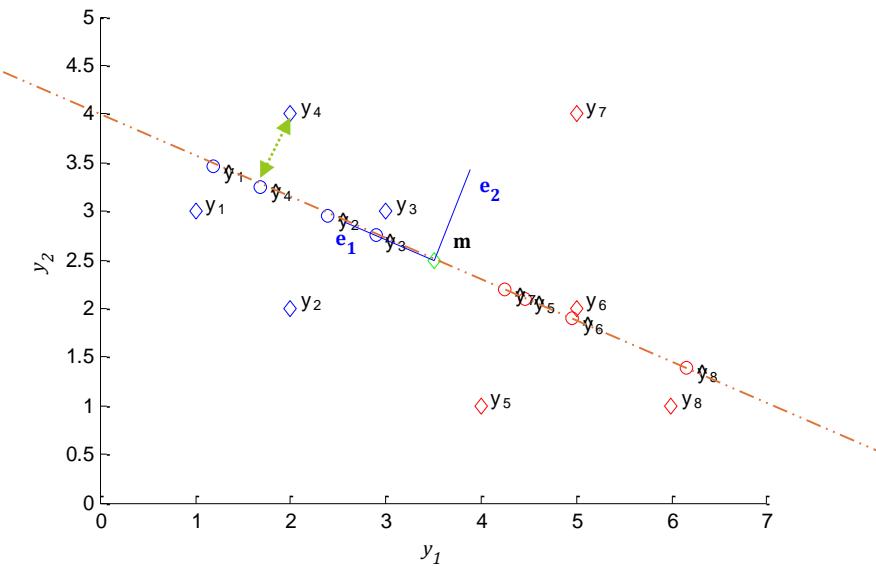
Računski zgled – luščenje ene značilke

- V primeru, ko želimo izluščiti eno samo značilko, izberemo za določitev linearne transformacije prvi lastni vektor \mathbf{e}_1 , ker mu pripada višja lastna vrednost λ_1 . Vrednosti značilk pridobimo s transformacijo

$$\mathbf{c} = \mathbf{e}_1^T (\mathbf{y} - E\{\mathbf{y}\}) = \mathbf{e}_1^T (\mathbf{y} - \hat{\mathbf{m}}),$$

torej

$$U'_1 \approx \{2.50, 1.19, 0.65, 1.96\}$$
$$U'_2 \approx \{-1.04, -1.58, -0.81, -2.88\}$$



Fourierjeva transformacija

- Fourierjeva transformacija vzorca $\mathbf{y} = (y_1, \dots, y_r)^T$ je definirana s transformacijsko matriko razsežnosti $r \times r$:

$$\mathbf{A} = \left[\exp \left\{ - \left(\frac{2\pi j}{r} \right) kl \right\} \right],$$

kjer so: $j = \sqrt{-1}$ ter $k = 0, 1, \dots, r-1$ in $l = 0, 1, \dots, r-1$.

- Zmnožek kl računamo po modulu r in če označimo $b = \exp\{-(2\pi j/r)\}$, dobimo za 8-razsežne vzorce Fourierjevo transformacijsko matriko

$$\mathbf{A} = \begin{bmatrix} b^0 & b^0 \\ b^0 & b^1 & b^2 & b^3 & b^4 & b^5 & b^6 & b^7 \\ b^0 & b^2 & b^4 & b^6 & b^0 & b^2 & b^4 & b^6 \\ b^0 & b^3 & b^6 & b^1 & b^4 & b^7 & b^2 & b^5 \\ b^0 & b^4 & b^0 & b^4 & b^0 & b^4 & b^0 & b^4 \\ b^0 & b^5 & b^2 & b^7 & b^4 & b^1 & b^6 & b^3 \\ b^0 & b^6 & b^4 & b^2 & b^0 & b^6 & b^4 & b^2 \\ b^0 & b^7 & b^6 & b^5 & b^4 & b^3 & b^2 & b^1 \end{bmatrix},$$

Transformacija Walsha in Hadamarda

- Transformacija Walsha in Hadamarda vzorca $\mathbf{y} = (y_1, \dots, y_r)^T$ je definirana s transformacijsko matriko razsežnosti $r \times r$:

$$\mathbf{A} = \left[(-1)^{\sum_{m=0}^{(\log_2 r)-1} k_m l_m} \right],$$

kjer sta k_m in l_m m -ta znaka binarnega zapisa indeksa vrstic in stolpcev matrike \mathbf{A} ter $r = 2n$ za $\forall n \in \mathbb{N}$.

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}$$

Zgled 8-razsežne
transformacijske matrike
Walsha in Hadamarda

- Transformiranko Walsha in Hadamarda lahko izračunamo z $r \log 2r$ seštevanji ali odštevanji.
- Tudi nekaj začetnih členov transformiranke Walsha in Hadamarda vsebuje največ informacije o vzorcu, vendar je zgostitev informacije na začetne člene manjša kot pri Fourierjevi transformiranki.

Haarova transformacija

- Haarova transformacija vzorca $\mathbf{y} = (y_1, \dots, y_r)^T$ je definirana s transformacijsko matriko razsežnosti $r \times r$, ki jo sestavimo iz vrednosti Haarovih funkcij χ_m reda $m = 0, 1, \dots, r - 1$ v točkah $x = m/r$
- Haarove funkcije reda m so določene kot

$$\begin{aligned}\chi(0, 0; x) &= \frac{1}{\sqrt{r}}; \quad x \in [0, 1] \\ \chi(p, q; x) &= \frac{1}{\sqrt{r}} \cdot \begin{cases} 2^{\frac{p}{2}}; & \frac{q-1}{2^p} \leq x < \frac{q-1/2}{2^p} \\ -2^{\frac{p}{2}}; & \frac{q-1/2}{2^p} \leq x < \frac{q}{2^p} \\ 0; & \text{x drugod v } [0, 1]. \end{cases}\end{aligned}$$

- Pri čemer red Haarovih funkcij m enolično razstavimo na celi števili p in q :

$$m = 2^p + q - 1, \quad m = 0, 1, \dots, r - 1 \quad \text{ter} \quad r = 2^n,$$

kjer sta

$$0 \leq p \leq n - 1, \quad 0 \leq q \leq 2^p \quad \text{za } p \neq 0 \quad \text{ter} \quad q = 0 \text{ ali } 1 \quad \text{za } p = 0.$$

Zgled Haarove transformacije

- Na primer, za 8-razsežne vzorce Haarovo transformacijsko matriko sestavimo tako, da najprej določimo parametra Haarovih funkcij:

m	0	1	2	3	4	5	6	7
p	0	0	1	1	2	2	2	2
q	0	1	1	2	1	2	3	4

izračunamo Haarove funkcije $\chi(p, q; x)$ v točkah $x = 0, 1/8, 2/8, \dots, 7/8$ ter izračunane vrednosti uredimo v matriko:

$$\mathbf{A} = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix}.$$

- Haarova transformiranka je zanimiva predvsem zaradi lokalne občutljivosti nekaterih njenih koeficientov. Lahko jo izračunamo z $2(r - 1)$ seštevanji ali odštevanji.

HEVRISTIČNE ZNAČILKE IZSEKOV GOVORNEGA SIGNALA

- Uvod
- Linearno napovedovanje izseka govornega signala
- Homomorfna analiza izseka govornega signala

Uvod

- Niz odtipkov t-tega izseka govornega signala,

$$\{f_t(nT)\} = f_t(0), f_t(T), \dots, f_t((N-1)T),$$

ki smo ga dobili z oknenjem, lahko strnjeno zapišemo s parametri sistema za tvorjenje govornega signala.

- Le-te najbolj pogosto poiščemo z:
 - linearnim napovedovanjem izseka govornega signala ali s
 - homomorfno analizo izseka govornega signala.
- Število parametrov prevajalne funkcije linearrega sistema, ki jih rodi obravnavani izsek govornega signala, je navadno med 10 in 15. So kvantitativnega značaja in ko jih uredimo v vektor značilk.
- Tem značilkam lahko dodamo še naslednje značilke:
 - frekvenca osnovnega tona zvenečih glasov,
 - moč P ali glasnost G izseka govornega signala in
 - resonančne frekvence (formanti) glasovnega trakta

$$P = \frac{1}{N} \sum_{n=0}^{N-1} f^2(nT)$$

$$G = 20 \log_{10} P$$

Linearno napovedovanje (LPC)

- Predpostavimo, da je zaporedje odtipkov izseka govornega signala $\{f(n)\}$ izhod iz diskretnega linearnega sistema, ki ga povzroči (neznano) zaporedje odtipkov signala generatorja $\{u(n)\}$:

$$f(n) = - \sum_{i=1}^p a(i)f(n-i) + \sum_{j=0}^q b(j)u(n-j).$$

- Ker zaporedja odtipkov signala generatorja $\{u(n)\}$ ne poznamo, lahko izračunamo odtipek $f(n)$ samo iz predhodnih p odtipkov $f(n-1), \dots, f(n-p)$:

$$\tilde{f}(n) = - \sum_{i=1}^p a(i)f(n-i),$$

- Če vzamemo, da je parameter p dan, lahko izračunamo napako med dejansko in približno vrednostjo odtipka kot:

$$e(n) = f(n) - \tilde{f}(n) = f(n) + \sum_{i=1}^p a(i)f(n-i).$$

Določanje koeficientov linearega napovedovanja

- Koeficiente $a(i)$; $i = 1, 2, \dots, p$ določimo tako, da minimiziramo kvadratno napako približka govornega signala

$$E = \sum_{n=0}^{N-1} e^2(n) = \sum_{n=0}^{N-1} \left(f(n) + \sum_{i=1}^p a(i)f(n-i) \right)^2$$

- Če E delno odvajamo na koeficiente $a(j)$; $j = 1, 2, \dots, p$ ter odvode izenačimo z nič, dobimo sistem p linearnih enačb:

$$\sum_{i=1}^p a(i) \sum_{n=0}^{N-1} f(n-i)f(n-j) + \sum_{n=0}^{N-1} f(n)f(n-j) = 0, \quad j = 1, 2, \dots, p,$$

ki jih v strnjeni obliki zapišemo kot:

$$\sum_{i=1}^p a(i)R(j-i) = -R(j), \quad j = 1, 2, \dots, p,$$

$$R(j) = \sum_{n=0}^{N-1} f(n)f(n-j)$$

kjer je $R(j)$ avtokorelacija govornega signala, pri kateri velja $R(j) = R(-j)$ in $R(j - i) = R(i - j)$.

Določanje koeficientov linearega napovedovanja

- Sistem enačb poznamo pod imenom Yule-Walkerjeve enačbe.
- Lahko jih rešimo za koeficiente $a(i); i = 1, 2, \dots, p$, če poznamo $(p + 1)$ avtokorelacijskih parametrov $R(i); i = 0, 1, \dots, p$.
- V matrični obliki zapišemo sistem enačb kot:

$$\begin{bmatrix} R(0) & R(1) & \dots & R(p-1) \\ R(1) & R(0) & \dots & R(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R(p-1) & \dots & R(1) & R(0) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ \vdots \\ a(p) \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix}$$

ozziroma kot $\mathbf{R} \cdot \mathbf{a} = \mathbf{r}$.

- Enačbo lahko rešimo za \mathbf{a} z invertiranjem matrike \mathbf{R} ali računsko bolj učinkovito z Durbinovim rekurzivnim postopkom, ki se opira na dejstvo, da je \mathbf{R} simetrična in Toeplitzeva.

$$\begin{bmatrix} a & b & c & d & \\ e & a & b & c & \\ f & e & a & b & \\ g & f & e & a & \end{bmatrix}$$

Durbinov rekurzivni postopek

1. korak: Inicializacija

$$\begin{aligned} E(0) &= R(0) \\ a_j(0) &= 0, \quad j = 0, 1, \dots, p. \end{aligned}$$

2. korak: Rekurzija

za $j := 1$ do p naredi

$$\begin{aligned} K(j) &= \frac{-R(j) - \sum_{i=1}^{j-1} a_{j-1}(i)R(j-i)}{E(j-1)} \\ a_j(j) &= K(j) \\ a_j(i) &= a_{j-1}(i) - K(j)a_{j-1}(j-i); \quad i = 1, 2, \dots, (j-1) \\ E(j) &= (1 - K^2(j)) E(j-1). \end{aligned}$$

konec-za

3. korak: Zaključek

Najboljši vektor koeficientov linearnega napovedovanja signala je:

$$\begin{aligned} \mathbf{a} &= (a_p(1), a_p(2), \dots, a_p(p)) \\ &= (a(1), a(2), \dots, a(p)). \end{aligned}$$

Določanje koeficientov linearne napovedovanja

- S tako določenimi koeficienti je najmanjša kvadratna napaka natanko:

$$E(p) = R(0) + \sum_{i=1}^p a(i)R(i).$$

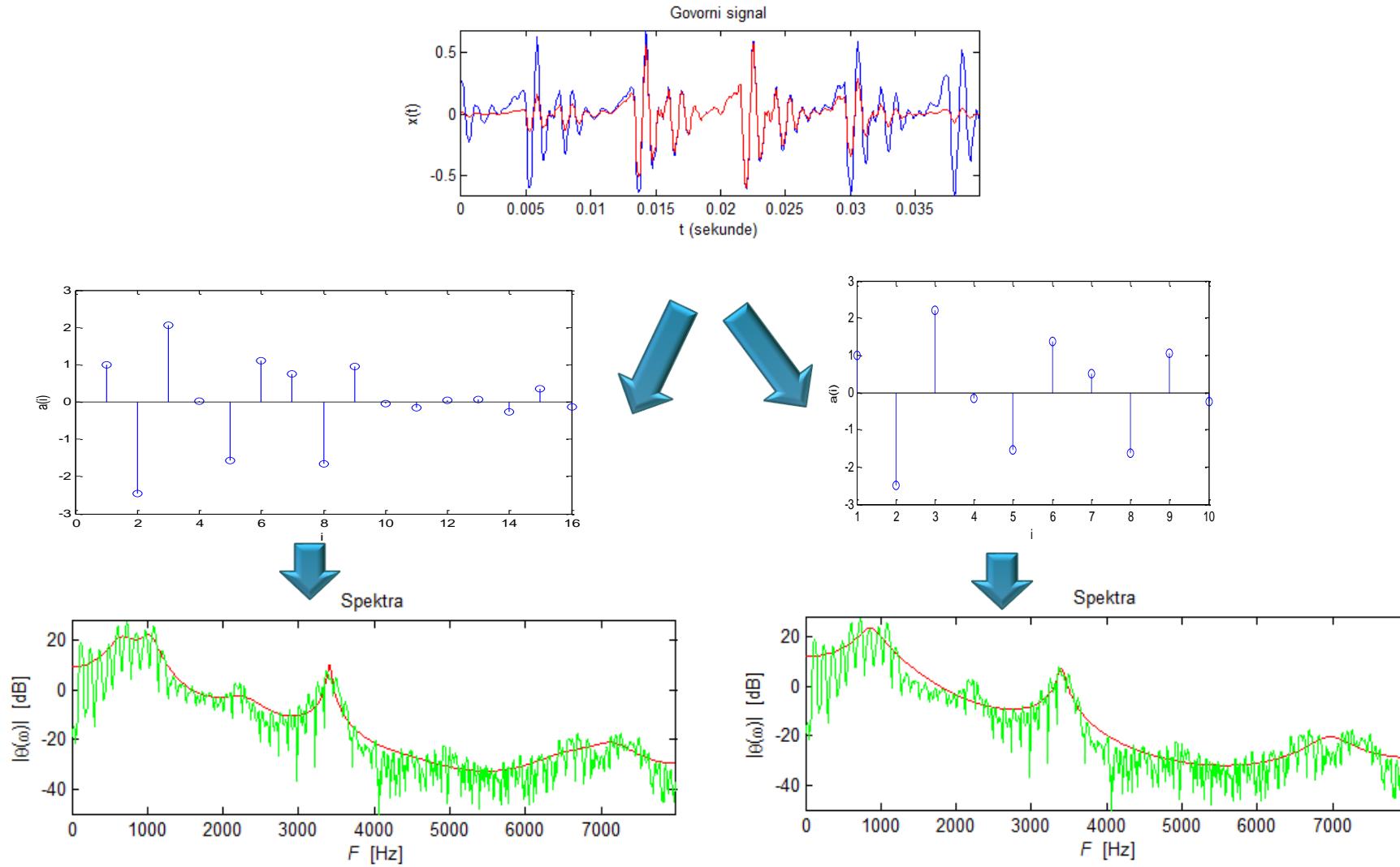
- Makhoul je pokazal, da glasnost izseka lahko določimo s pomočjo korelacijskih koeficientov in koeficientov modela glasovnega trakta z zvezo:

$$G^2 = R(0) + \sum_{i=1}^p a(i)R(i),$$

resonančne frekvence (formante) glasovnega trakta pa kot ničle polinoma:

$$A(e^{j\omega}) = 1 + \sum_{i=1}^p a(i)e^{-j\omega i}.$$

Zgled LPC analize izseka govornega signala



Homomorfna analiza

- Obravnavajmo izsek govornega signala $\{f(n)\}$ kot izhod iz diskretnega linearnega sistema, ki ga opišemo z odtipki odziva na enotin impulz $\{h(n)\}$.
- Odtipke signala vzbujanja sistema označimo z $\{u(n)\}$, ki ga pri zveničih glasovih sestavlja vlak enotskih impulzov $\delta_{T_0}(n)$.
- Govorni signal $f(n)$ lahko zapišemo kot **konvolucijo** vzbujanja sistema $u(n)$ in odziva sistema na enotin impulz $h(n)$ in s homomorfni analizo poskušamo ločeno pridobiti ta dva signala.
- Pri zveničih glasovih predpostavljamo, da je vzbujanje $u(n)$ kar periodična Kroneckerjeva funkcija $\delta_{T_0}(n)$ in relacijo zapišemo

$$f(n) = h(n) * \delta_{T_0}(n)$$

Homomorfna analiza

- Pri homomorfni analizi **operator konvolucije** z diskretno Fourierovo transformacijo najprej spremenimo v **operator množenja** in nato z logaritmiranjem še v **operator seštevanja**.

$$f(n) = h(n) * \delta_{T_0}(n)$$

$$F_D(k) = H_D(k) \cdot \Delta_{T_0}(k)$$

$$\ln F_D(k) = \ln H_D(k) + \ln \Delta_{T_0}(k)$$

- Če izvedemo še inverzno diskretno Fourierovo transformacijo logaritma spektra $\ln F_D(k)$ dobimo signal $\tilde{c}(n)$, ki ga lahko obravnavamo kot vsoto drugih dveh časovnih signalov

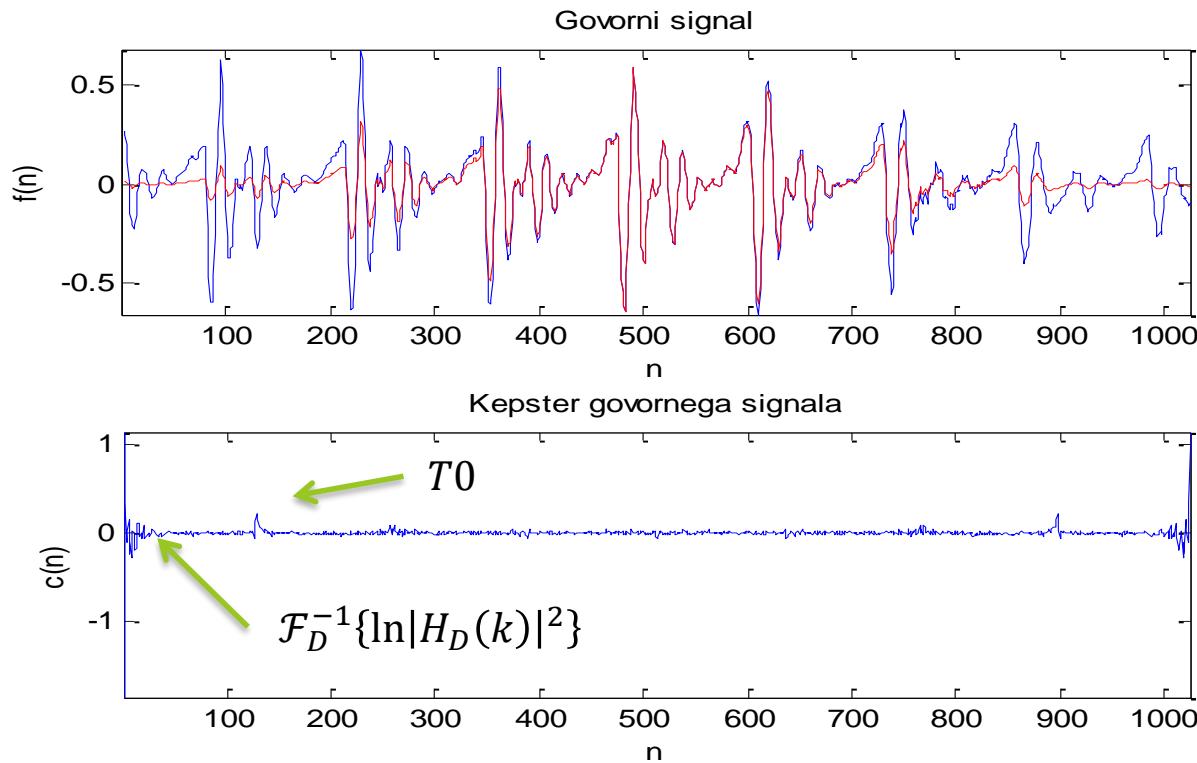
$$\{\tilde{c}(n)\} = \mathcal{F}_D^{-1}\{\ln F_D(k)\} = \mathcal{F}_D^{-1}\{\ln H_D(k)\} + \mathcal{F}_D^{-1}\{\ln \Delta_{T_0}(k)\}$$

- Signal $\tilde{c}(n)$ imenujemo **kepster** govornega signala.

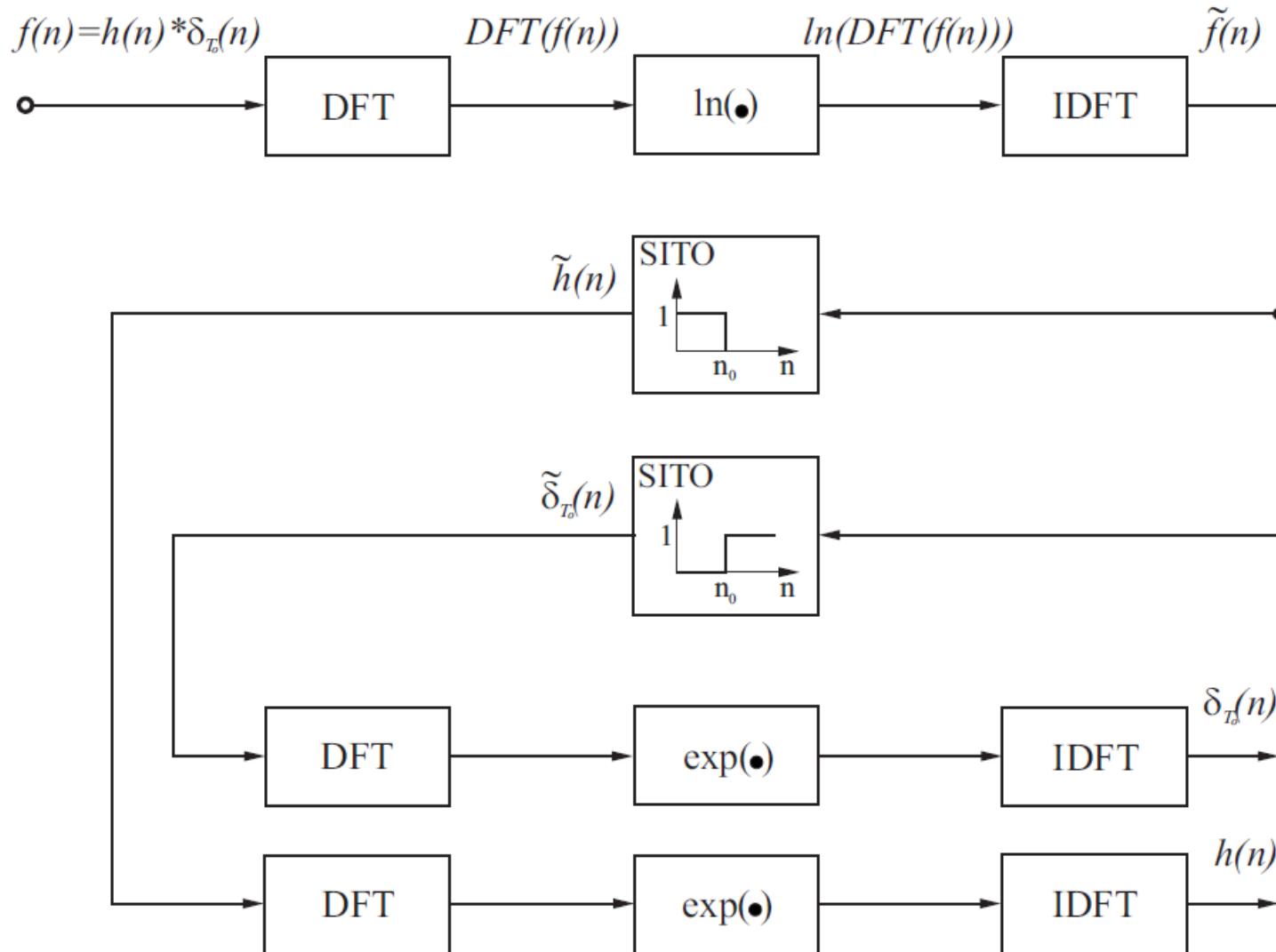
Homomorfna analiza

- Ker slušno zaznavanje ni občutljivo na fazne zamike med frekvenčnimi komponentami, navadno uporabljamo namesto kompleksnega le močnostni spekter govornega signala

$$\{\tilde{c}(n)\} = \mathcal{F}_D^{-1}\{\ln|F_D(k)|^2\}$$



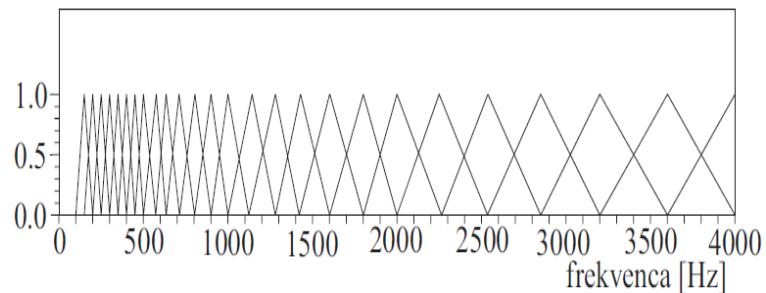
Homomorfna analiza govornega signala



Zgled melodične delitve močnostnega spektra

- Slušno zaznavanje pa je tudi neenakomerno (melodično) porazdeljeno po frekvenčni osi, zato se pri izračunu kepstra namesto logaritma močnostnega spektra uporablja logaritme povprečnih moči frekvenčnih območij, razporejenih po melodični delitvi.

m	$s(m)$ Hz	$d(m)$ Hz	m	$s(m)$ Hz	$d(m)$ Hz
0	150.0	50.0	13	1000.0	166.7
1	200.0	50.0	14	1166.7	166.7
2	250.0	50.0	15	1333.3	166.7
3	300.0	50.0	16	1500.0	166.7
4	350.0	50.0	17	1666.7	166.7
5	400.0	50.0	18	1833.3	166.7
6	450.0	50.0	19	2000.0	333.3
7	500.0	83.3	20	2333.3	333.3
8	583.3	83.3	21	2666.7	333.3
9	666.7	83.3	22	3000.0	333.3
10	750.0	83.3	23	3333.3	333.3
11	833.3	83.3	24	3666.7	333.3
12	916.7	83.3			



Izračun koeficientov melodičnega kepstra

$$\tilde{c}_{MEL}(n) = \sum_{m=0}^{M-1} \ln P_m \cos\left(\frac{n(2m+1)\pi}{2M}\right), \quad n = 1, 2, \dots, 10 \text{ do } 15,$$

kjer so:

$\tilde{c}_{MEL}(n)$ n -ti koeficient MEL kepstra t -tega izseka govornega signala,

$P_m = \sum_{k=0}^{N-1} \Delta_m(k) |F_t(k)|^2$ uteženo povprečje močnostnega spektra $|F_t(k)|^2$ t -tega izseka na m -tem frekvenčnem območju,

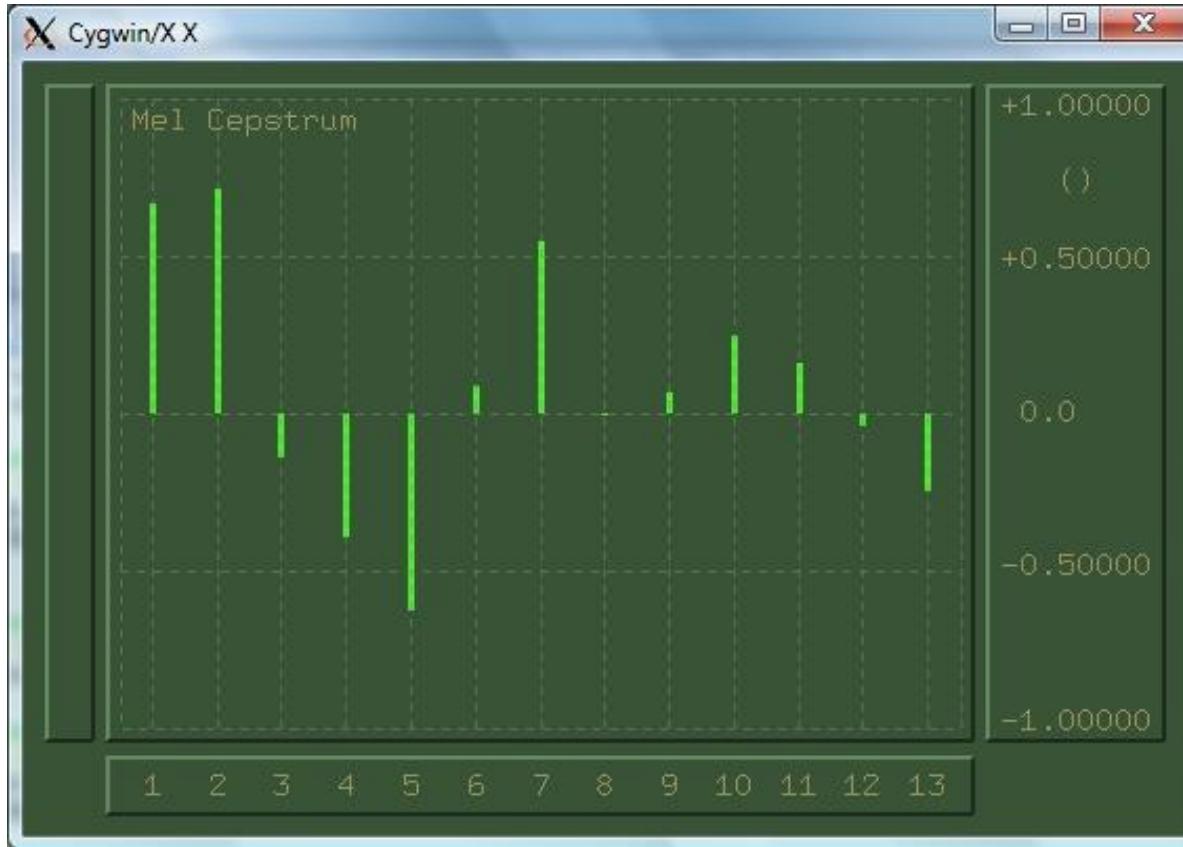
$F_t(k) = \frac{1}{N} \sum_{n=0}^{N-1} f_t(n) e^{-2\pi j nk/N}; k = 0, 1, \dots, N - 1$ diskretna Fourierjeva transformiranka t -tega izseka,

$\Delta_m(k)$ utežna funkcija m -tega frekvenčnega območja,

M število frekvenčnih območij (navadno med 20 in 30) na t -tem izseku in

N število koeficientov diskretne Fourierjeve transformiranke (ponavadi 128 ali 256) t -tega izseka govornega signala.

Demonstracija izračuna govornih značilk



Dinamične značilke izsekov govornega signala

- Izkazalo se je, da so pomembne tudi **dinamične značilke**, s katerimi opišemo spremembe, ki so se zgodile med izseki govornega signala.
- Izpeljemo jih iz “osnovnih” značilk in so lahko prvega ali drugega reda.
- Dinamične značilke **prvega** reda so definirane kot razlike:

$$\delta x_i^t = x_i^{t-\tau} - x_i^{t+\tau}, \quad i = 1, \dots, n$$

dinamične značilke **druge** reda pa kot razlike značilk prvega reda

$$\Delta x_i^t = \delta x_i^{t-\tau} - \delta x_i^{t+\tau}, \quad i = 1, \dots, n$$

kjer je

t indeks izseka govornega signala,

x_i^t “osnovne” značilke izseka govornega signala,

n število “osnovnih” in dinamičnih značilk ter

τ odmik od t -tega izseka, navadno med 1 in 3.

HEVRISTIČNE ZNAČILKE PODROČIJ SLIKE

- Značilke področij slike
- Opis obrisa
 - s funkcijami Fourierjevih koeficientov,
 - z morfološkimi značilkami, verižno kodo
 - točkami največje ukrivljenosti, mnogokotniki ter
 - funkcijo ukrivljenosti in podpisom
- Opis področij s funkcijami geometričnih momentov
- Opis področij z značilkami sestave površine področij

Značilke področij slike

- Rezultat razčlenjevanja slike na področja navadno zapišemo s seznamom naslovov (x_i, y_i) in sivim nivojem $f(x_i, y_i)$ vseh točk področij predmeta, ki ga želimo razpoznati, t.j.

$$\{(x_0, y_0, f(x_0, y_0)), \dots, (x_{U-1}, y_{U-1}, f(x_{U-1}, y_{U-1}))\}$$

ali s seznamom naslovov točk (x_i, y_i) krivulj, ki ograjujejo področja - obrisov področij predmeta, ki ga želimo razpoznati, t.j.

$$\{(x_0, y_0), \dots, (x_{S-1}, y_{S-1})\},$$

kjer sta

U število točk področja in

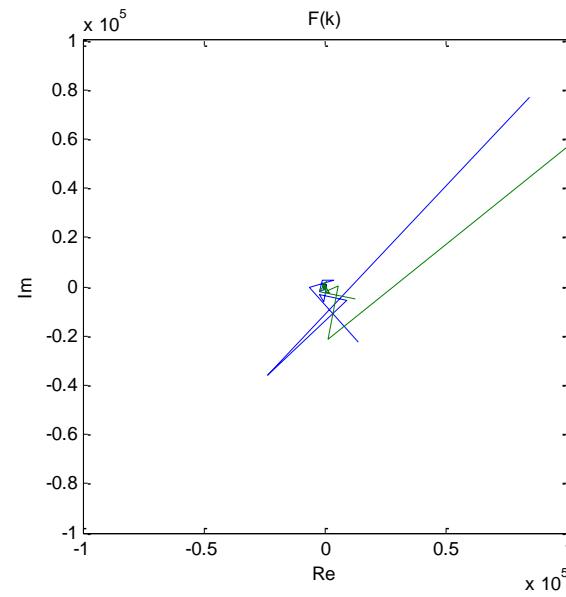
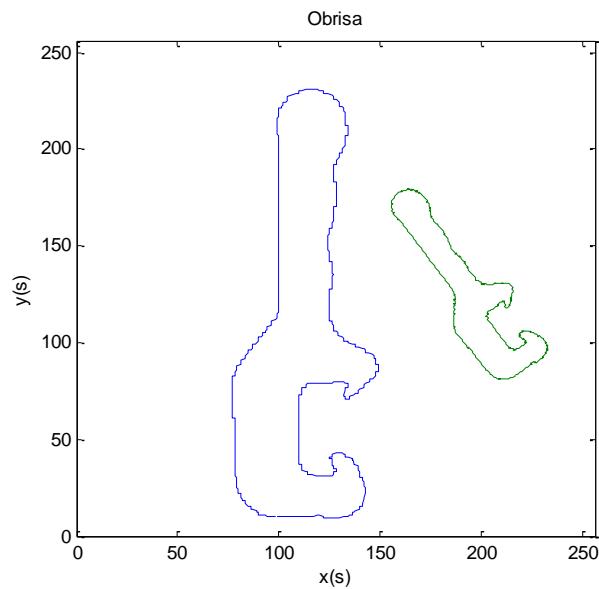
S število točk obrisa.

Značilke področij slike

- Predmete (objekte) na slikah, lahko razpoznavamo na podlagi **oblike obrisa** predmetov, ali na podlagi **sestave površine ploskev** predmetov.
- Glede na to delimo postopke opisa področij z značilkami na tiste, ki temeljijo na:
 - opisu obrisa področja, to je na **opisu krivulje**, ki ograjuje področje in na tiste, ki temeljijo na
 - opisu "notranjosti" obrisov, to je na **opisu enovitega dvorazsežnega področja v ravnini**.



OPIS OBRISA S FUNKCIJAMI FOURIERJEVIH KOEFICIENTOV



Zgoščen opis obrisov področij slike

- Obrise področij digitalnih slik lahko zgoščeno opišemo s pomočjo koeficientov enorazsežne diskretne Fourierjeve transformacije.
- Pri tem obris obravnavamo kot enorazsežno diskretno kompleksno periodično funkcijo

$$\{d(s) : s = 0, \dots, S - 1\}$$

kjer je

$$d(s) = x_s + jy_s \quad \text{kompleksne točke obrisa}$$

S število točk obrisa

- Koeficiente diskretne Fourierjeve transformacije obrisa izračunamo iz enačbe:

$$F(n) = \frac{1}{S} \sum_{s=0}^{S-1} d(s) e^{-j \frac{2\pi n s}{S}}$$

Spektralna frekvenčna analiza obrisa

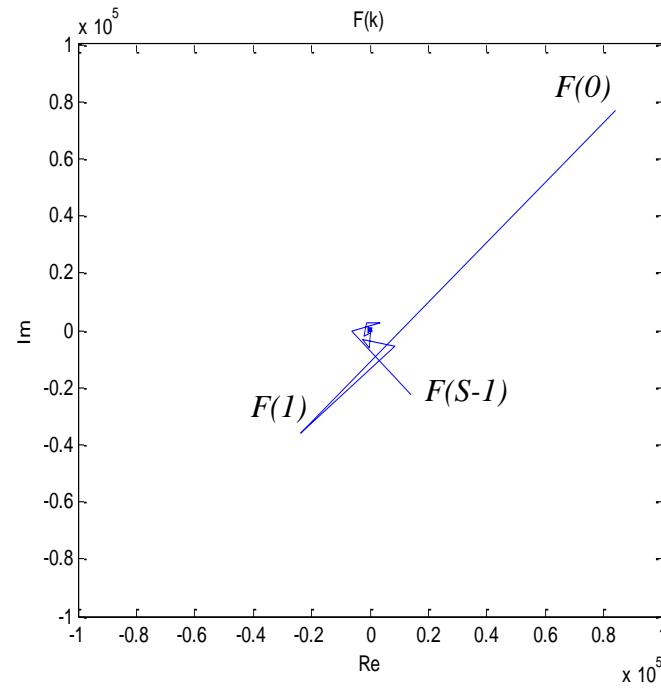
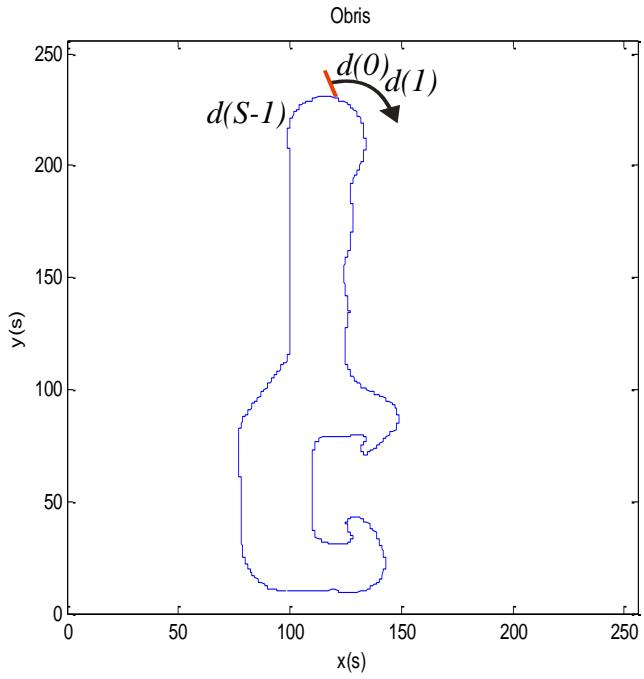
$$\{d(s) : s = 0, \dots, S - 1\}$$

$$d(s) = x_s + jy_s$$

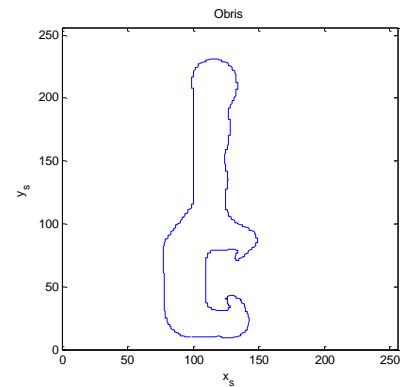


$$\{F(n) : n = 0, \dots, S - 1\}$$

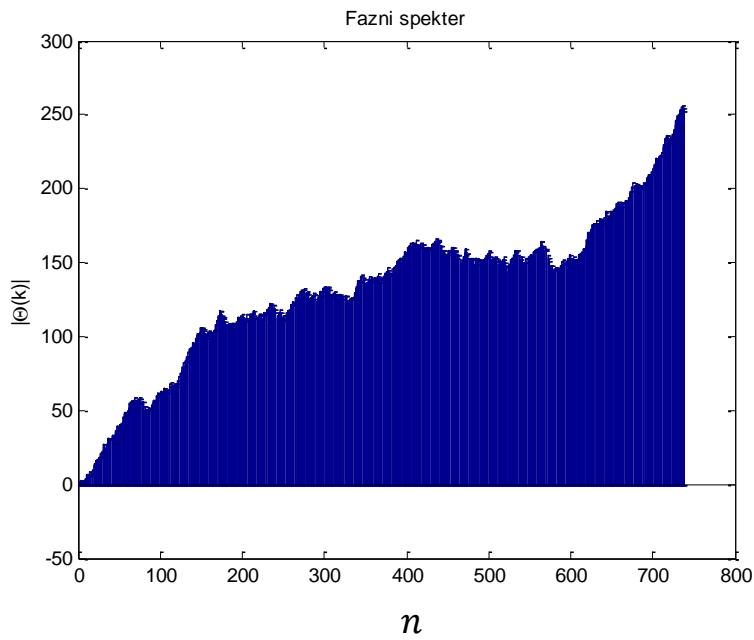
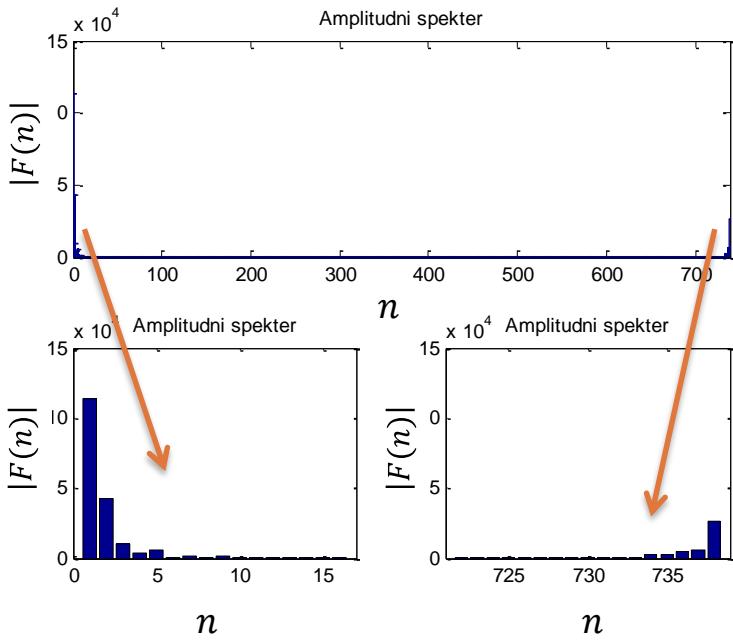
$$F(n) = \frac{1}{S} \sum_{s=0}^{S-1} d(s) e^{-j \frac{2\pi ns}{S}}$$



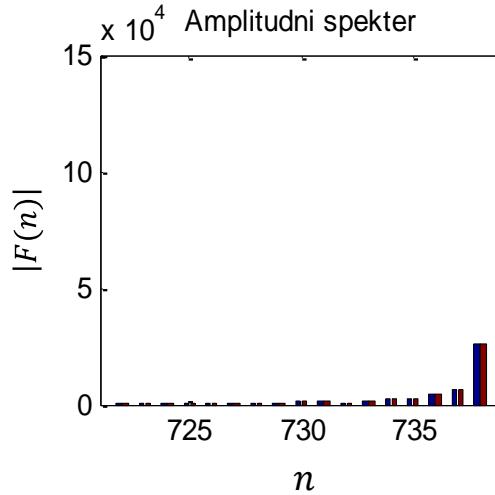
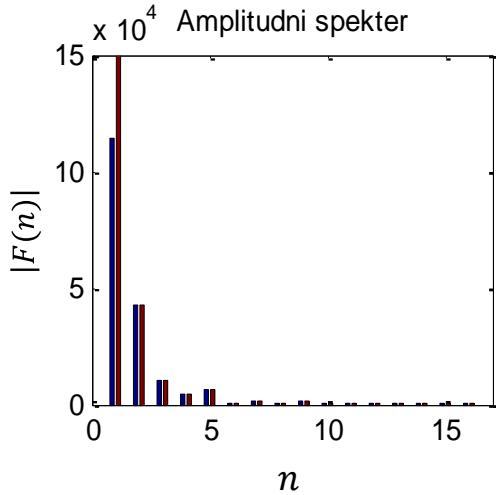
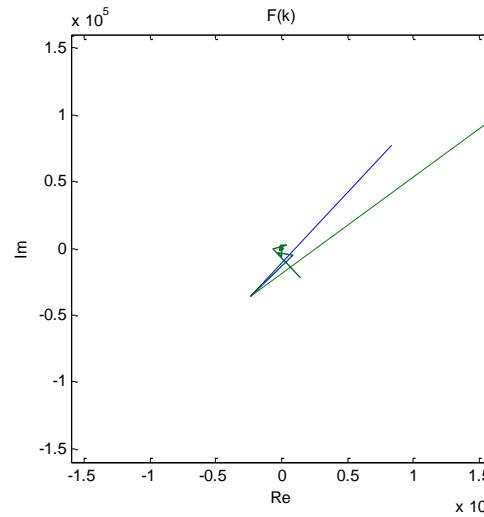
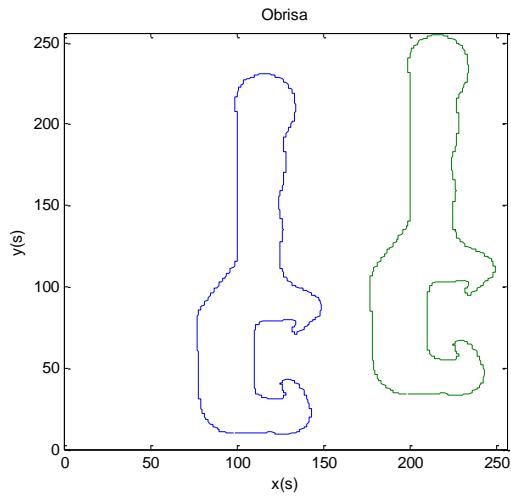
Spektralna predstavitev „signala“ obrisa



$$\{d(s): s = 0, \dots, S - 1\} \leftrightarrow \{F(n): n = 0, \dots, S - 1\}$$



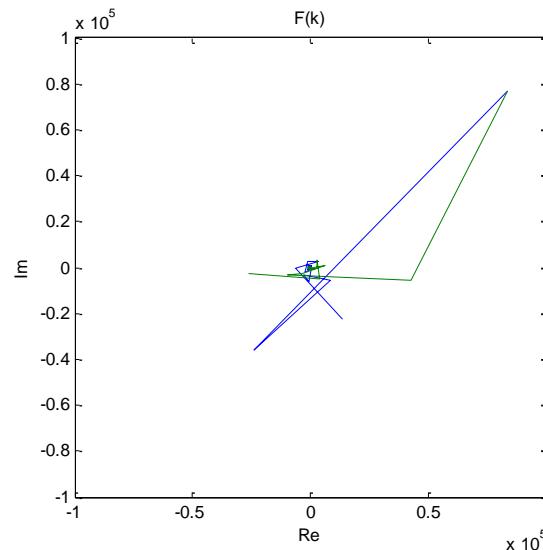
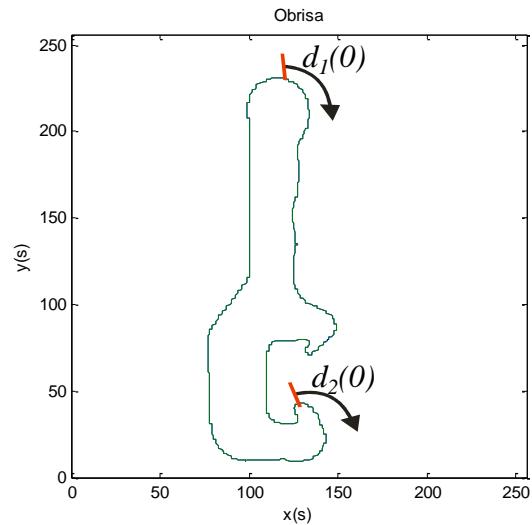
Vpliv premika obrisa na njegov frekvenčni spekter



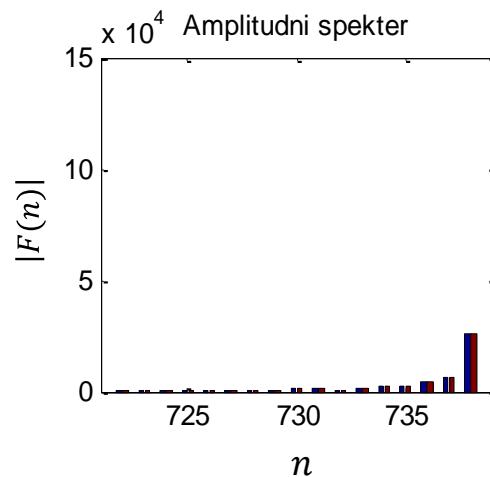
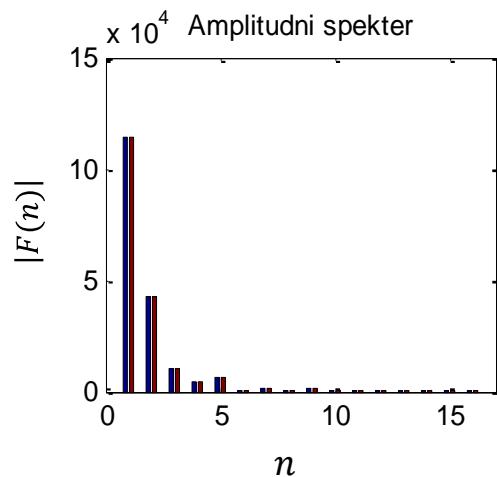
$$d_2(s) = d_1(s) + C$$

$$F_2(n) = \begin{cases} F_1(0) + C & n = 0 \\ F_1(n) & n > 0 \end{cases}$$

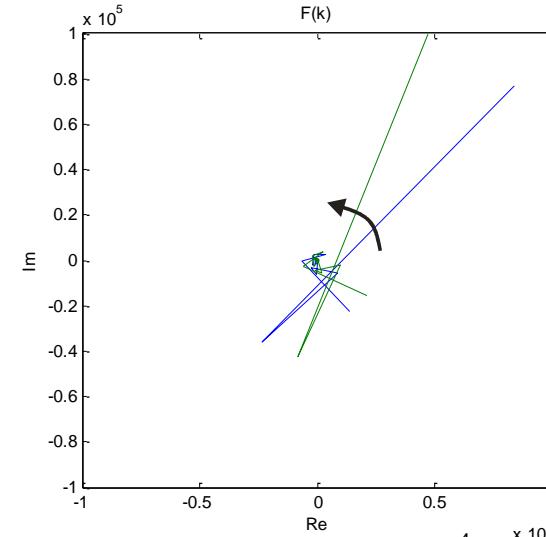
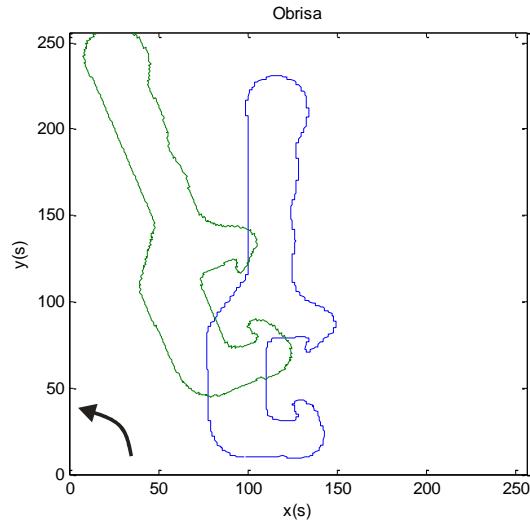
Vpliv premika začetne točke obrisa na njegov frekvenčni spekter



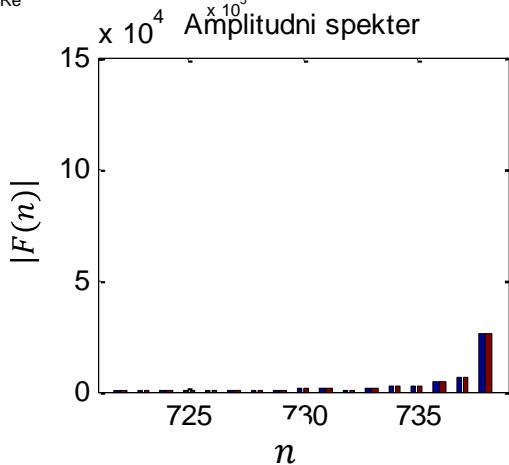
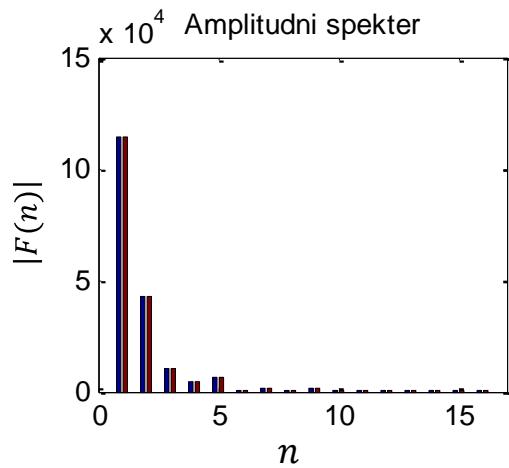
$$d_2(s) = d_1(s + \tau)$$
$$F_2(n) = F_1(n)e^{-j\frac{2\pi n \tau}{s}}$$



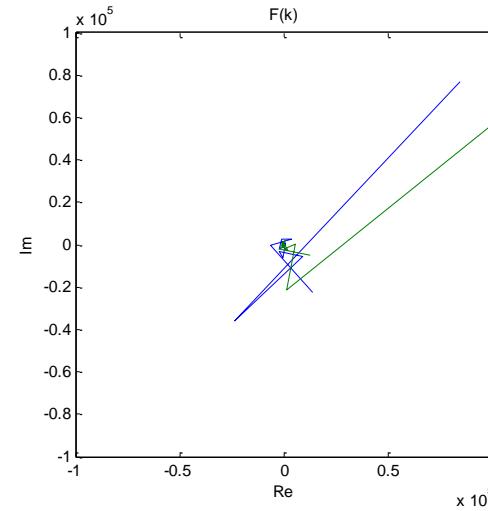
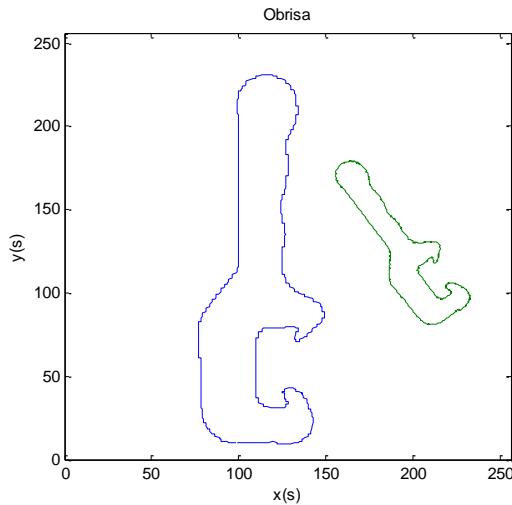
Vpliv zasuka obrisa na njegov frekvenčni spekter



$$d_2(s) = d_1(s)e^{j\varphi}$$
$$F_2(n) = F_1(n)e^{j\varphi}$$

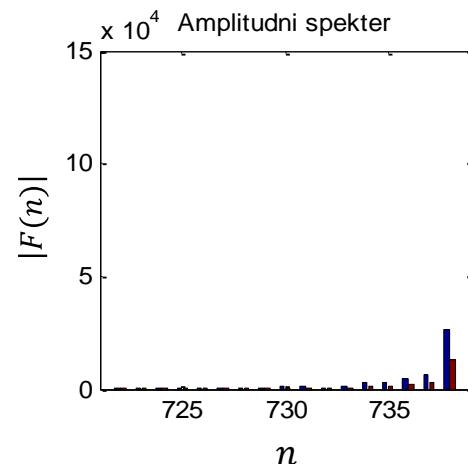
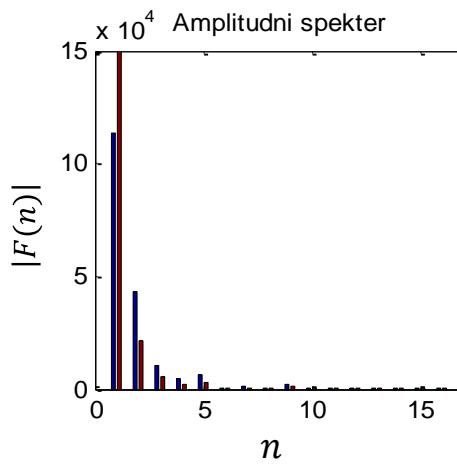


Vpliv sestavljenje spremembe obrisa na njegov frekvenčni spekter

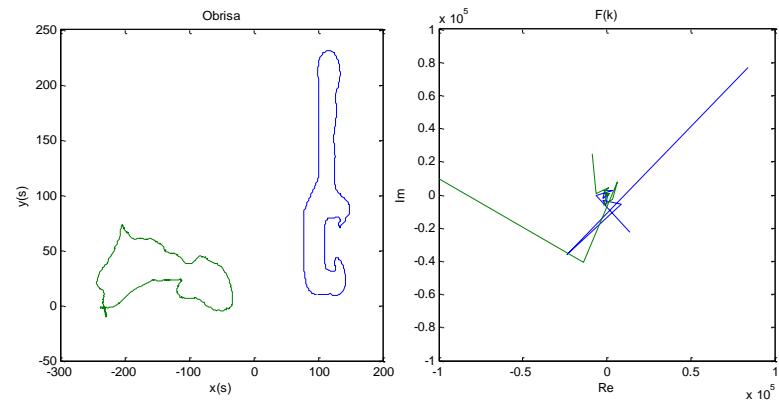
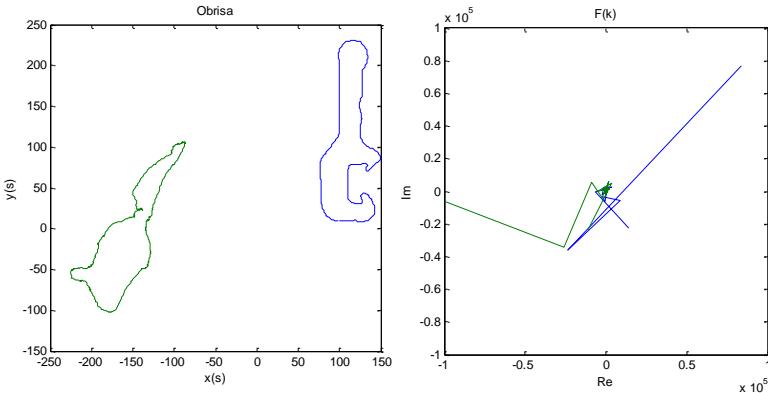


$$d_2(s) = R d_1(s + \tau) e^{j\varphi} + C$$

$$F_2(n) = \begin{cases} R F_1(0) e^{-j\left(\frac{2\pi n \tau}{s} + \varphi\right)} + C & n = 0 \\ R F_1(n) e^{-j\left(\frac{2\pi n \tau}{s} + \varphi\right)} & n > 0 \end{cases}$$

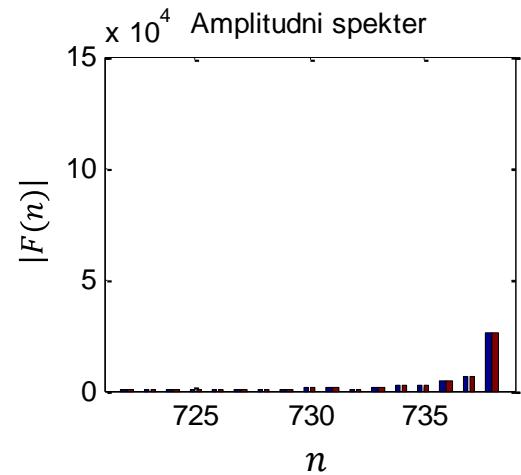
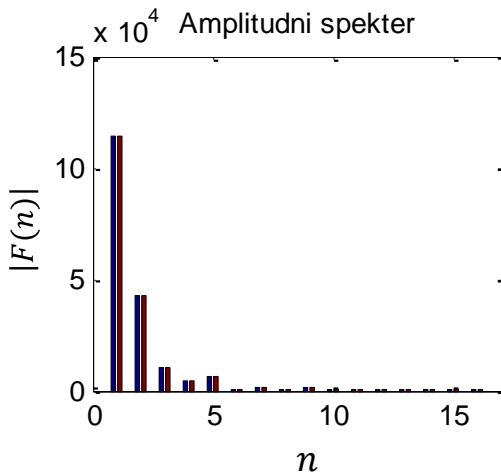


Pomen faznega spektra o obrisa



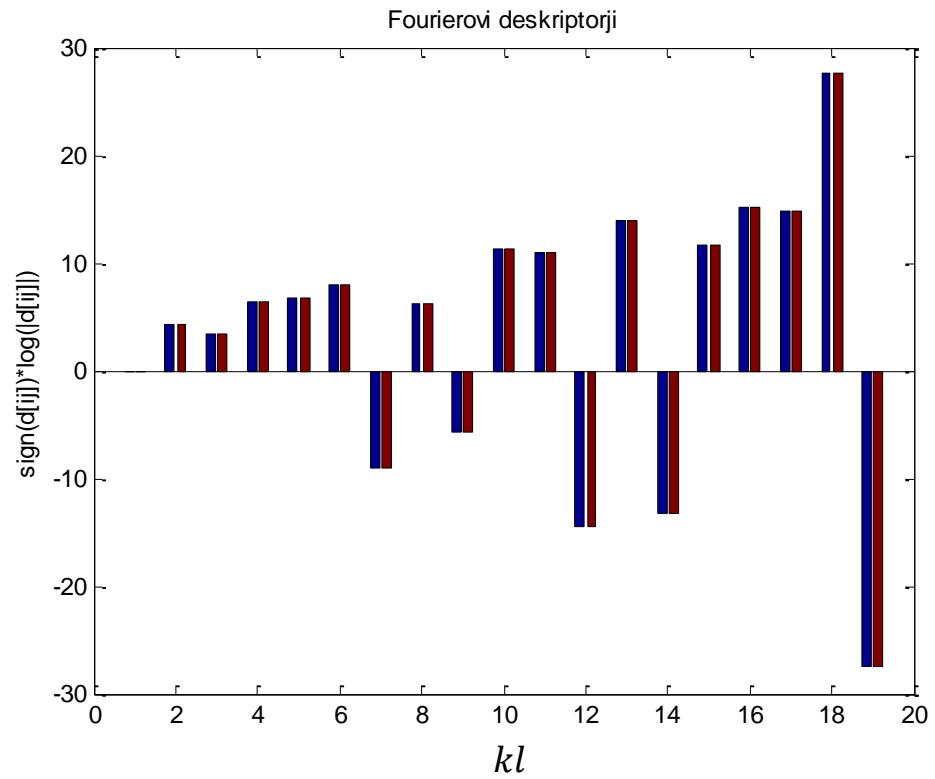
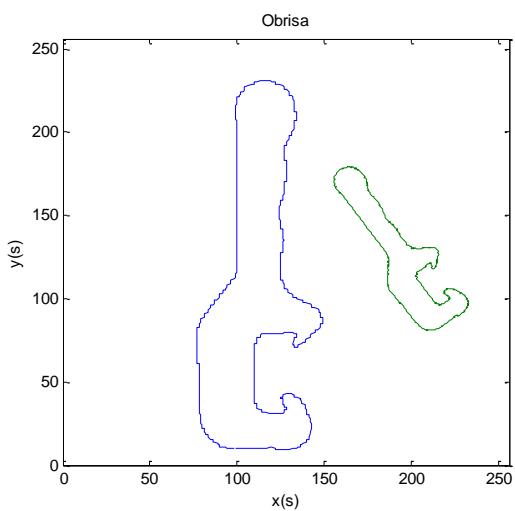
$$d_2(s) = R d_1(s + \tau) e^{j\varphi} + C$$

$$|F_2(n)| = \begin{cases} R|F_1(n)| + C & n = 0 \\ R|F_1(n)| & n > 0 \end{cases}$$



Funkcije Fourierovih koeficientov

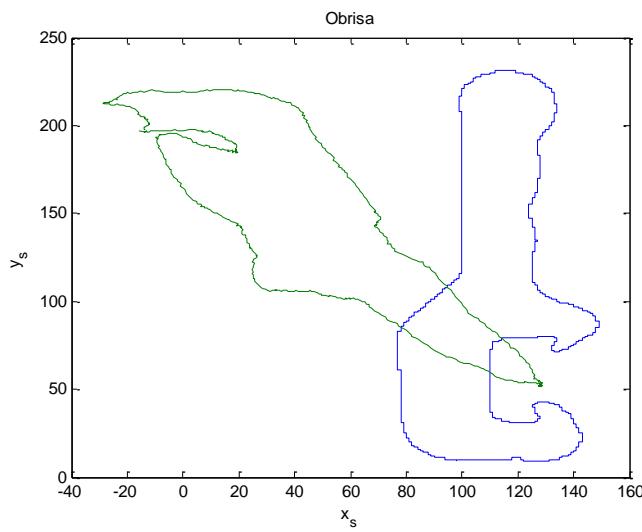
- » Granlund je definiral funkcije Fourierjevih koeficientov tako, da se izničijo vplivi premika začetne točke, premika, zasuka in razširitve osnovne krivulje obrisa.



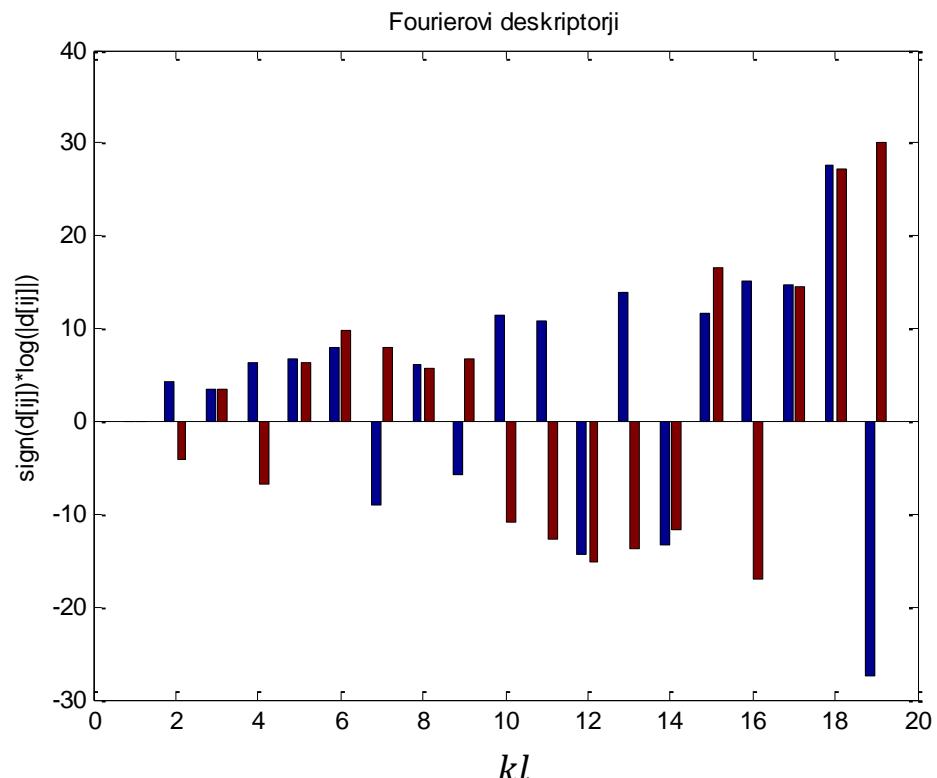
$$d_{kl} = \frac{F^l(k+1)F^k(S+1-l)}{F^{k+l}(1)}$$

Vpliv naključne spremembe faze na funkcije Fourierovih koeficientov

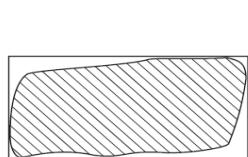
Naključna sprememba faze



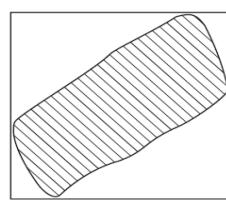
$$d_{kl} = \frac{F^l(k+1)F^k(S+1-l)}{F^{k+l}(1)}$$



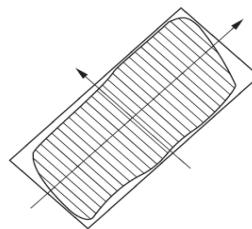
OPIS OBRISA Z NEKAJ DRUGIMI FUNKCIJAMI IN ZNAČILKAMI



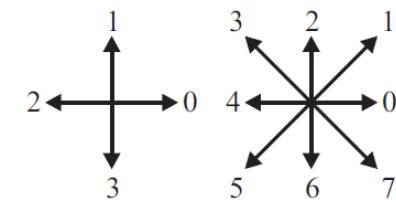
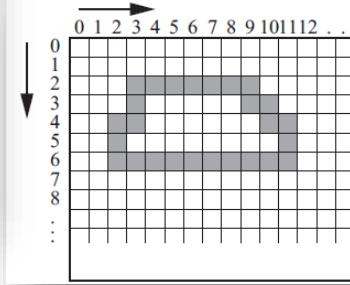
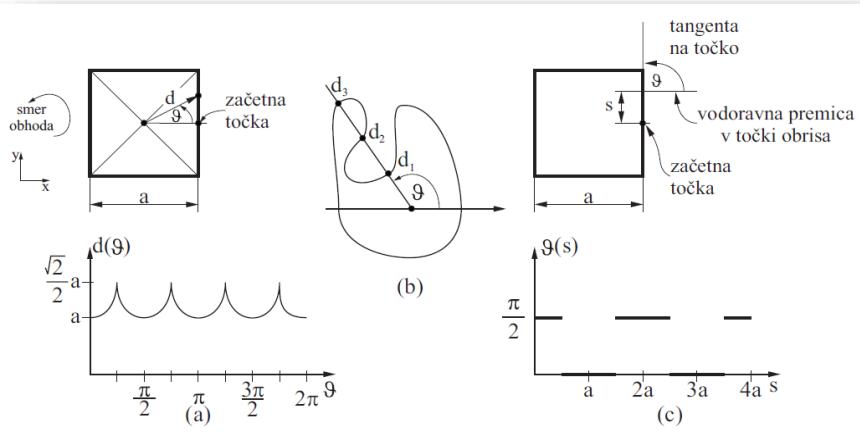
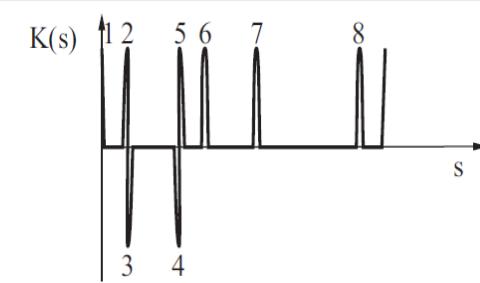
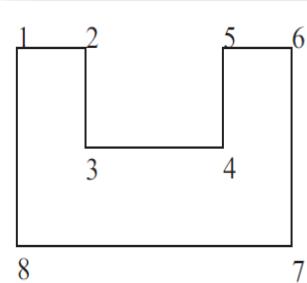
$e = 2,16$



$e = 1$



$e = 2,33$



Opis obrisa z morfološkimi značilkami

- Obrise področij v slikovni ravni lahko opišemo tudi z drugimi značilkami, ki so odvisne le od oblike obrisov, ne pa tudi od lege, rotacije ali povečave področja objekta na sliki.
- Primera takšnih značilk sta strnjenost in raztegnjenost področja.
- **Strnjenost** c (med 0 in 1) je razmerje med ploščino (P) področja (številom slikovnih elementov U , ki jih obris ograjuje) in obsegom (O) področja (številom slikovnih elementov obrisa S)

$$c = \frac{4\pi P}{O^2} = \frac{4\pi U}{S^2}$$

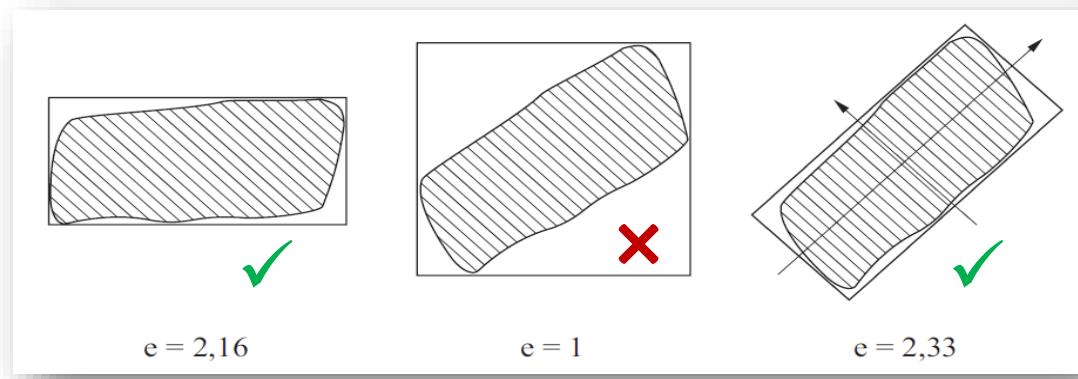
- Na primer, krog ima strnjenost 1, enakostranični trikotnik ima največjo strnjenost med trikotniki itd.

Raztegnjenost področja (1/3)

- Raztegnjenost e je razmerje med dolžino a in širino b pravokotnika, v katerega je področje vklenjeno (vrisano)

$$e = \frac{a}{b}$$

- Določitev raztegnjenosti obrisa zahteva določitev pravokotnika, ki oklepa njegovo področje.



- Stranice pravokotnika določimo tako, da so vzporedne z lastnimi vektorji matrike usredičenih momentov druge stopnje področja R

Raztegnjenost področja (2/3)

- Vrednost značilke e področja, ki je vklenjeno s pravokotnikom, ki ima stranice vzporedne z glavnimi osmi področja, je:

$$e = \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}}{U}$$

kjer so μ_{20} , μ_{02} in μ_{11} usrediščeni momenti drugega reda in U število slikovnih elementov področja.

- Za dano področje najprej izračunamo geometrične momente prve stopnje:

$$m_{pq} = \sum_{x \in R} \sum_{y \in R} x^p y^q, \quad p, q = 0, 1.$$

in težišči področja

$$(\bar{x}, \bar{y}) = \left(\frac{m_{10}}{U}, \frac{m_{01}}{U} \right)$$

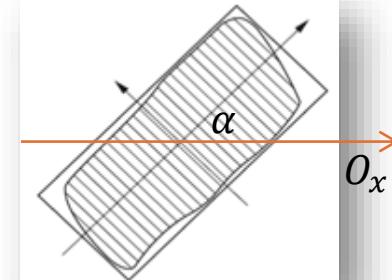
Raztegnjenost področja (3/3)

- Če premaknemo izhodišče koordinatnega sistema v težišče področja, dobimo usredишčene momente:

$$\mu_{pq} = \sum_{x \in R} \sum_{y \in R} (x - \bar{x})^p (y - \bar{y})^q.$$

- S pomočjo usredишčenih geometričnih momentov področja lahko izračunamo kot α med osjo inercije v smeri podaljšanja področja, in osjo O_x koordinatnega sistema:

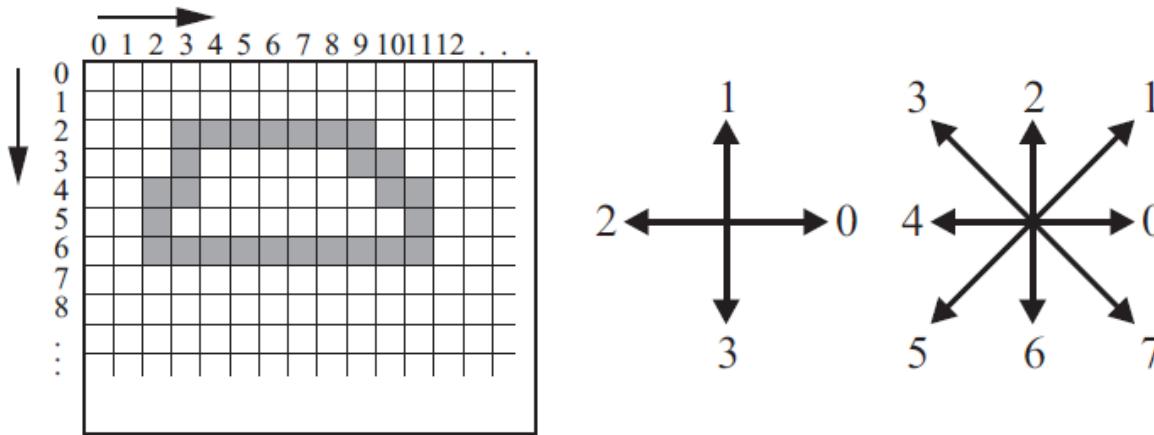
$$\alpha = \frac{1}{2} \arctan \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right).$$



- Kot α lahko uporabimo kot značilko, ki vsebuje informacijo o usmerjenosti (podolgovatega) področja.

Opis obrisa z verižno kodo

- Obrise področij na sliki navadno zapišemo s štiri- ali osem-smerno verižno kodo.



- Štiri-smerna verižna koda zgornjega primera obrisa je:
 $(2, 3) : 0000003030332222222211011$
- in osem-smerna
 $(2, 3) : 0000007766444444442212.$

Neodvisnost od začetne točke obrisa

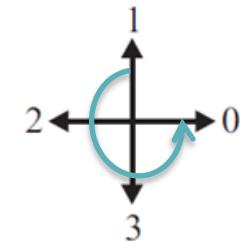
- Zapis obrisov z verižnimi kodi, lahko neposredno uporabimo za razpoznavanje področij slike, če ga **naredimo neodvisnega** od začetne točke in od rotacije področja na sliki
- Verižni kod obrisa je že neodvisen od lege področja na sliki.
- Od začetne točke ga lahko naredimo neodvisnega tako, da zapišemo niz številk koda v krožni register in premikamo številke toliko časa v desno, da zavzame kodni zapis najmanjšo možno desetiško vrednost.



Neodvisnost od rotacije obrisa

- Kod obrisa naredimo neodvisnega od zasuka obrisa, če izračunamo razliko med sosednjima elementoma koda.
- Le-to izračunamo s štetjem števila usmeritev, ki delijo dva zaporedna elementa kodnega zapisa v smeri, ki je obratna od vrtenja kazalca na uri.

3 1 3 1 3 0 0 1 3 1 3 3 1 3



- Prva številka tega niza predstavlja razliko med zadnjo in prvo številko (med **1** in **0**).
- Zgornji niz številk zapišemo še s krožnim premikanjem v desno kot najmanjšo možno desetiško **številko obrisa**.

3 1 3 1 3 0 0 1 3 1 3 3 1 3

→

0 0 1 3 1 3 3 1 3 3 1 3 1 3

Red številke obrisa

- Red številke obrisa n je enak številu številk v nizu.
- Red n je sodo število za vsako sklenjeno krivuljo.
- Njegova vrednost omejuje število možnih oblik obrisa.



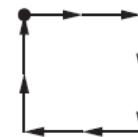
$n = 4$

verižna koda : 0 3 2 1
razlika : 3 3 3 3
št. obrisa : 3 3 3 3



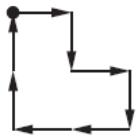
$n = 6$

verižna koda : 0 0 3 2 2 1
razlika : 3 0 3 3 0 3
št. obrisa : 0 3 3 0 3 3



$n = 8$

verižna koda : 0 0 3 3 2 2 1 1
razlika : 3 0 3 0 3 0 3 0
št. obrisa : 0 3 0 3 0 3 0 3



$n = 8$

verižna koda : 0 3 0 3 2 2 1 1
razlika : 3 3 1 3 3 0 3 0
št. obrisa : 0 3 0 3 3 1 3 3



$n = 8$

verižna koda : 0 0 0 3 2 2 2 1
razlika : 3 0 0 3 3 0 0 0
št. obrisa : 0 0 3 3 0 0 3 3

Opis obrisa s točkami največje ukrivljenosti

- Obrise področij lahko zgoščeno, vendar z dovolj informacije za razpoznavanje, opišemo s točkami največje ukrivljenosti obrisa.
- Točke največje ukrivljenosti obrisa navadno zapišemo z znakom a iz slovarja V , ki označuje tip in **stopnjo ukrivljenosti** (npr. rahlo vbočen), in koordinatami točke, torej s trojico (a, x, y) .
- Točke največje ukrivljenosti obrisa najlažje odkrijemo s postopkom, ki posnema zasnovno **razširjanja toplote** v telesih ali s približnim opisom obrisa z **mnogokotnikom**.
- Lahko jih odkrijemo tudi z opazovanjem **funkcij ukrivljenosti** in **podpisa obrisa**.

Razširjanje toplote (1/2)

- Postopek temelji na predpostavki, da se toplota razširja samo v notranjost telesa in ne v okolje, zato razširjanje toplote lahko zapišemo z rekurzivno enačbo

$$f_{k+1}(x_i, y_j) = f_k(x_i, y_j) + \\ + D \cdot (f_k(x_{i-1}, y_j) + f_k(x_{i+1}, y_j) + f_k(x_i, y_{j-1}) + f_k(x_i, y_{j+1}) - N(x_i, y_j)f_k(x_i, y_j))$$

kjer je

$f_k(x_i, y_j)$ svetilnost slikovnega elementa (x_i, y_j) v koraku k ,

D koeficient razširjanja toplote (navadno $1/4$) in

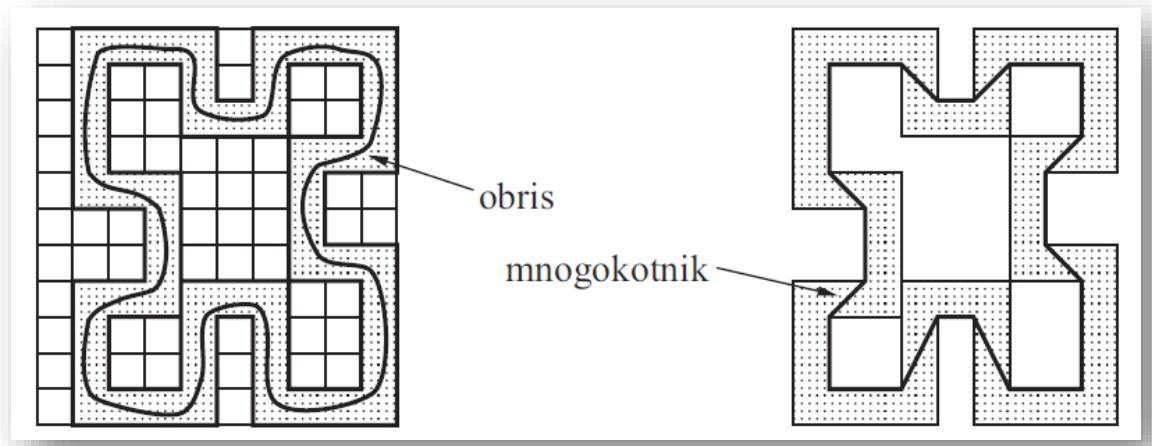
$N(x_i, y_j)$ število sosedov slikovnega elementa (x_i, y_j) , ki pripadajo področju, ki ga ograja obris

Razširjanje toplote (2/2)

- Na začetku postopka priredimo vsem točkam obrisa neko poljubno celo vrednost, večjo od nič, vsem ostalim slikovnim elementom področja pa vrednost nič.
- Po prvi iteraciji spremenijo svojo vrednost samo slikovni elementi obrisa in njihovi sosedji znotraj področja.
- Po drugi iteraciji spremenijo svojo vrednost še sosedji sosedov obrisa in tako naprej v notranjost področja.
- Po določenem številu iteracij predstavljajo slikovni elementi z največjimi vrednostmi f izbočene točke obrisa, slikovni elementi z najmanjšimi vrednostmi f pa vbočene točke.

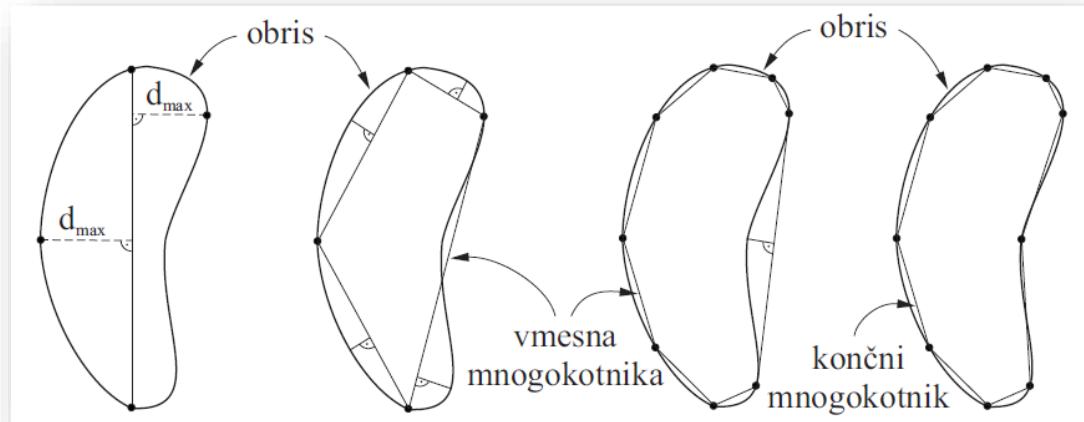
Opis obrisa z oglišči mnogokotnika

- Obris lahko približno opišemo z mnogokotnikom z najmanjšim obsegom.
- Obris področja prekrijemo z mrežo s pravokotnimi okni določene velikost in označimo zaporedje oken mreže, skozi katera prehaja obris.
- Predpostavimo, da je obris iz elastičnega materiala, ki mu dovolimo, da se skrči, kolikor mu to dopuščajo okna mreže, zato ta zavzame obliko mnogokotnika.



Določitev mnogokotnika z združevanjem in razdruževanjem razmakov

- Pri združevanju obris razdelimo na n majhnih razmakov, ki jih predstavimo s končnima točkama, ki določajo daljico oz. stranico mnogokotnika.
- Nato krožimo po obrisu in razmake združujemo po kriteriju napake predstavitve točk vmesnega dela obrisa z daljicami.
- Pri razdruževanju je postopek obraten, t.j. najprej določimo le dva razmaka, ki ju nato razdružujemo na več razmakov, dokler ne dosežemo dovolj majhno napako predstavitve delov obrisa z daljicami.



Funkcija ukrivljenosti obrisa

- Obris lahko predstavimo tudi s funkcija ukrivljenosti obrisa, ki je enorazsežna funkcija, definirana kot odvod tangente na obris:

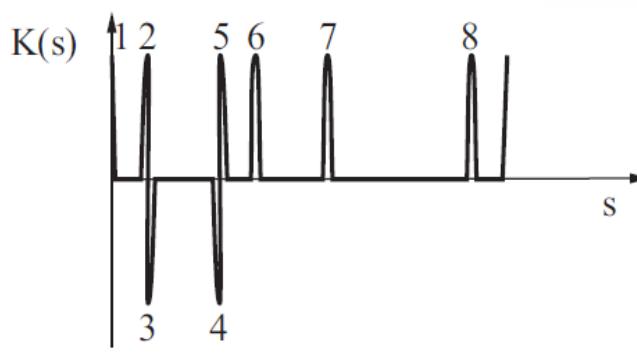
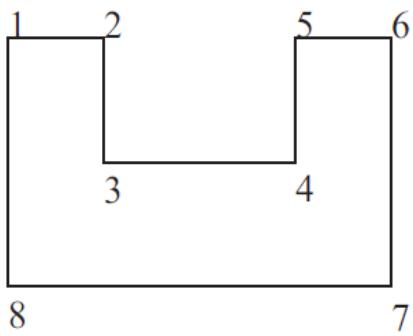
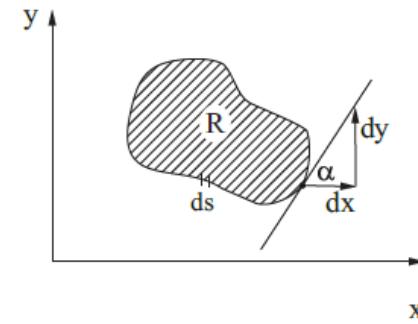
$$K(s) = \frac{d\alpha}{dt} = \frac{x'y'' - y'x''}{\sqrt{(x'^2 + y'^2)^3}},$$

kjer so:

x', x'' prvi in drugi odvod funkcije $x = x(s)$,

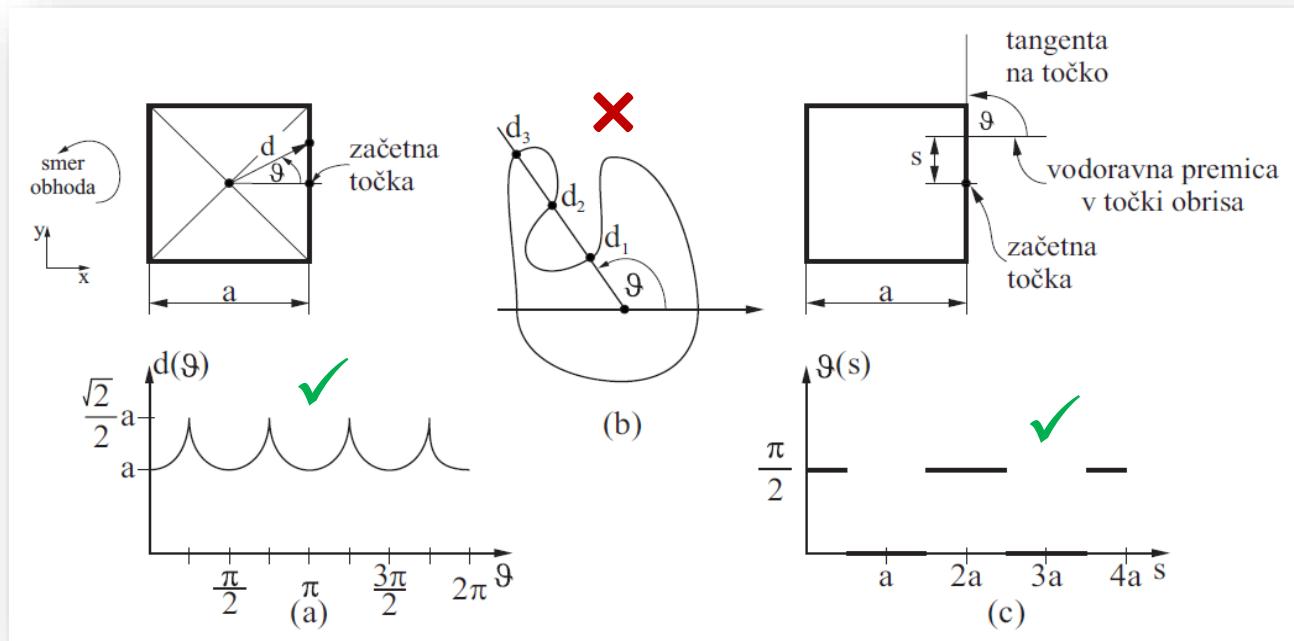
y', y'' prvi in drugi odvod funkcije $y = y(s)$,

t ločna dolžina



Opis obrisa z njegovim podpisom

- Podpis obrisa je z enorazsežno funkcionalna predstavitev obrisa.
- Obliko obrisa lahko opišemo kot funkcijo odvisnost razdalje d in kota ϑ med vodoravno premico, ki gre skozi težišče in daljico, ki povezuje težišče in točko obrisa.
- Druga možnost je, da podpis obrisa predstavimo kot kot med abcisno osjo in tangento v vsaki točki obrisa



OPIS PODROČIJ S FUNKCIJAMI GEOMETRIČNIH MOMENTOV

- Obliko področja slike lahko opišemo tudi s funkcijami geometričnih momentov svetilnosti vseh točk področja.
- Tudi funkcije geometričnih momentov so neobčutljive na spremembe lege področja v slikovni ravnini, na zasuk in na razširitev (skrčitev) področja.
- Geometrične momente stopnje $(p + q)$ definiramo z izrazom:

$$m_{pq} = \sum_{x \in R} \sum_{y \in R} x^p y^q f(x, y), \quad p, q = 0, 1, 2, \dots,$$

- Težišče področja je v točki:

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}.$$

Normiran usrediščeni momenti

- Če premaknemo izhodišče koordinatnega sistema v težišče področja, dobimo usrediščene momente stopnje ($p + q$):

$$\mu_{pq} = \sum_{x \in R} \sum_{y \in R} (x - \bar{x})^p (y - \bar{y})^q f(x, y).$$

- V novem koordinatnem sistemu so momenti področja neodvisni od lege področja na sliki.
- Neodvisnost od razširitve (skrčitve) pa dosežemo z normiranimi usrediščenimi momenti:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^r},$$

kjer je $r = (p + q)/2 + 1$, $(p + q) = 2, 3$.

Funkcije normiranih usrediščenih momentov

- Značilke področij slik končno določimo kot naslednje funkcije normiranih usrediščenih momentov:

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\begin{aligned}\phi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})((\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2) + \\ & +(3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})(3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2)\end{aligned}$$

$$\begin{aligned}\phi_6 = & (\eta_{20} - \eta_{02})((\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2) + \\ & + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})\end{aligned}$$

$$\begin{aligned}\phi_7 = & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})((\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2) - \\ & - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})(3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2).\end{aligned}$$

OPIS PODROČJA Z PARAMETRI SESTAVE POVRŠINE PODROČJA



Opis notranjosti področij slike

- Notranjost področij zgoščeno opišemo z značilkami, ki temeljijo na statistikah **prve in druge stopnje**.
- Statistika prve stopnje je statistična funkcija **posameznih** slikovnih elementov, statistika druge stopnje je statistična funkcija **parov sosednjih** slikovnih elementov.
- Značilke, ki temeljijo na statistikah prve stopnje, izračunamo iz porazdelitve relativnih frekvenc sivih nivojev danega področja:

$$P_i = \frac{n_i}{U} , \quad i = 0, 1, \dots, L - 1$$

kjer so:

n_i število slikovnih elementov področja R s sivim nivojem i ,

U število vseh slikovnih elementov na področju R ter

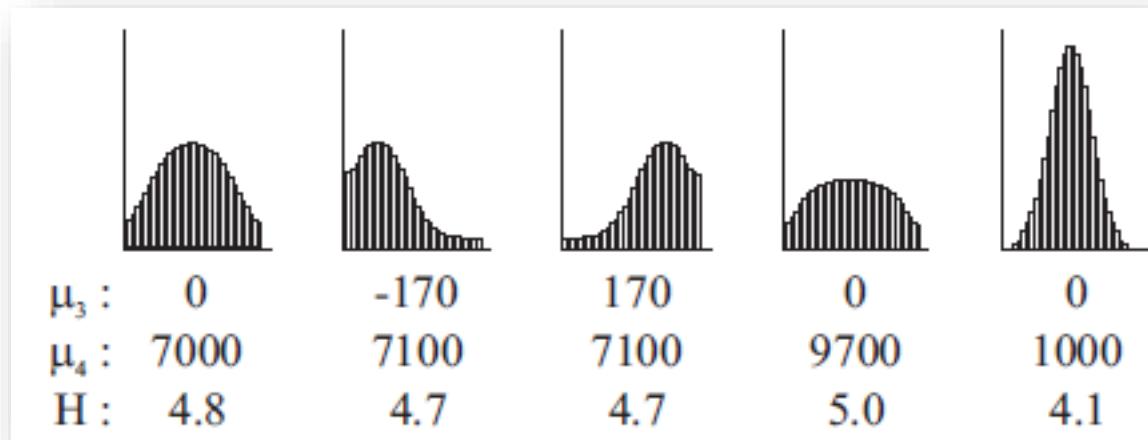
L število sivih nivojev.

Statistike prve stopnje (1/2)

- Najpomembnejše značilke, ki temeljijo na statistikah prve stopnje, so:
 - + prvi moment (povprečna svetilnost),
$$m_1 = \sum_{i=0}^{L-1} iP_i.$$
 - + 2,3 in 4 usredишčeni moment in
$$\mu_k = \sum_{i=0}^{L-1} (i - m_1)^k P_i, \quad k = 2, 3, 4.$$
 - + entropija (informacija).
$$H = - \sum_{i=0}^{L-1} P_i \log P_i.$$
- Značilka $\mu_2 = \sigma^2$ je mera širine, μ_3 je mera nesimetričnosti, μ_4 je mera sploščenosti in H mera enakomernosti porazdelitve relativnih frekvenc svetilnosti področja.

Statistike prve stopnje (2/2)

- Primeri histogramov sivih nivojev z značilkami μ_3, μ_4 in H :

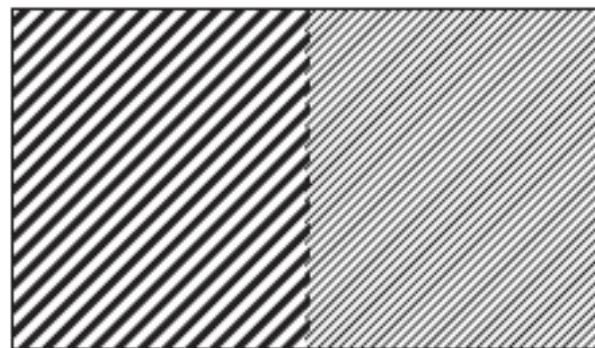
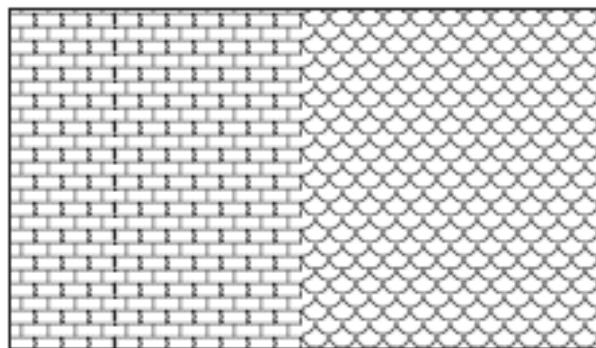


- Značilke μ_2, μ_3 in μ_4 navadno normiramo, in sicer takole:

$$\nu_2 = 1 - \frac{1}{1 + \mu_2}, \quad \nu_3 = \frac{\mu_3}{\mu_2^{3/2}} \quad \text{in} \quad \nu_4 = \frac{\mu_4}{\mu_2^2}.$$

Statistike druge stopnje (1/4)

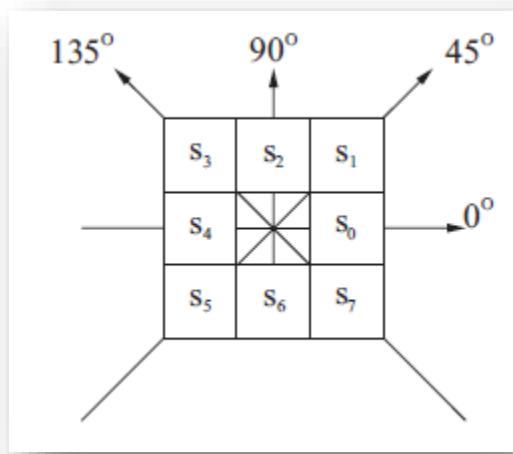
- Značilke, ki temeljijo na porazdelitvi relativnih frekvenc svetilnosti slikovnih elementov, ne vsebujejo informacije o medsebojnem položaju slikovnih elementov področja z enako ali skoraj enako svetilnostjo.



Nekaj primerov področij z enakimi povprečnimi vrednostmi svetilnosti.

Statistike druge stopnje

- Dodamo še značilke, ki temeljijo na matriki vezanih verjetnosti dveh svetilnosti sosednjih slikovnih elementov danega področja (**statistika druge stopnje**).
- Element takšne matrike predstavlja verjetnost, da sta dve vrednosti svetilnosti sosednji na **določeni razdalji** in **v določeni smeri**.
- Za različne smeri in za različne razdalje med sosedji tvorimo samostojne matrike.
- Navadno opazujemo sosedje v smereh 0° , 45° , 90° in 135° ter z razdaljo ena do osem slikovnih elementov.



Smeri opazovanja sosednjih elementov.

Matrika frekvenc parov

- Matriko vezanih verjetnosti določimo s pomočjo matrike F , ki vsebuje frekvence parov določene vrednosti svetilnosti.
- Oznaka $\#(i, j)$ na spodnji sliki pomeni število dogodkov, da sta slikovna elementa s svetilnostma i in j sosedna v določeni smeri na določeni razdalji.
- Dimenzija matrike frekvenc parov je definirana s številom diskretnih vrednosti svetilnosti slikovnih elementov L

svetilnost	0	1	2
0	$\#(0,0)$	$\#(0,1)$	$\#(0,2)$
1	$\#(1,0)$	$\#(1,1)$	$\#(1,2)$
2	$\#(2,0)$	$\#(2,1)$	$\#(2,2)$

Primer slike z tremi različnimi sivimi nivoji in ustrezna matrika frekvenc parov.

Primeri matrike frekvenc parov

- Matrike frekvenc parov, za podan primer področja za razdaljo en slikovni element ter za smeri 0° , 45° , 90° in 135° so:

0	0	1	1
0	0	1	1
0	2	2	2
2	2	1	1

$$\begin{aligned} \mathbf{F}_{1,0^\circ} &= \begin{bmatrix} 4 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 6 \end{bmatrix} & \mathbf{F}_{1,90^\circ} &= \begin{bmatrix} 6 & 0 & 2 \\ 0 & 4 & 4 \\ 2 & 4 & 2 \end{bmatrix} \\ \mathbf{F}_{1,45^\circ} &= \begin{bmatrix} 4 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 4 \end{bmatrix} & \mathbf{F}_{1,135^\circ} &= \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 0 \end{bmatrix} \end{aligned}$$

Iz posamezne matrike frekvenc parov \mathbf{F} dobimo matriko vezanih verjetnosti \mathbf{P} tako, da vsak element matrike frekvenc parov delimo z vsoto vseh elementov matrike frekvenc parov. Tako dobimo: $\mathbf{P}_{1,0^\circ} = 1/24 \mathbf{F}_{1,0^\circ}$, $\mathbf{P}_{1,90^\circ} = 1/24 \mathbf{F}_{1,90^\circ}$, $\mathbf{P}_{1,45^\circ} = 1/18 \mathbf{F}_{1,45^\circ}$ in $\mathbf{P}_{1,135^\circ} = 1/18 \mathbf{F}_{1,135^\circ}$.

Značilke matrik vezanih verjetnosti

- Matrike vezanih verjetnosti svetilnosti slikovnih elementov danega področja zapišemo v bolj strnjeni obliki s različnimi značilkami.
- Med temi so najbolj pomembne: energija, kontrast, homogenost, entropija in največji korelacijski koeficient, ki so definirani takole:

- Energija

$$z_1 = \sum_i \sum_j (p(i, j))^2$$

- Kontrast

$$z_2 = \sum_i \sum_j (i - j)^2 p(i, j)$$

- Homogenost

$$z_3 = \sum_i \sum_j \frac{p(i, j)}{1 + (i - j)^2}$$

- Entropija

$$z_4 = - \sum_i \sum_j p(i, j) \log(p(i, j))$$

- Korelacijski koeficient

$$z_5 = \sum_i \sum_j \frac{(i - \mu_x)(j - \mu_y)p(i, j)}{\sigma_x \sigma_y}$$

$\mu_x, \sigma_x, \mu_y, \sigma_y$ so povprečne vrednost i in standardni odklon i robnih porazdelitev

$$p_x(i) = \sum_{j=0}^{L-1} p(i, j)$$

$$p_y(j) = \sum_{i=0}^{L-1} p(i, j)$$

Zaključek

- Spoznali smo različne pristope k določanju hevrističnih značilke področij slike.
- Področja na sliki lahko opišemo z obrisi, ki ograjujejo področje in ki jih opišemo z različnimi funkcijami oziroma z verižno kodo.
- Področja slik lahko opišemo tudi s funkcijami svetilnosti točk notranjosti in z značilkami sestave njihovih površin.