

Semestrální zkouška ZRE, řádný termín, 16.5.2019, skupina 333

Login: Příjmení a jméno: Podpis:
(prosím čitelně!)

1. Navrhněte, jaké příznaky (features) by se daly použít pro určení emočního stavu mluvčího.

2. V programu v jazyce C jsou pole X_r a X_i o velikosti $\frac{N}{2} + 1 = 129$ naplněny vypočítanou diskretní Fourierovou transformací od nuly do poloviny vzorkovací frekvence (reálná a imaginární složka). Napište kód pro převod na odhad spektrální hustoty výkonu (power spectral density, PSD).

3. Podle prací ve fonetice má foném “a” formantové frekvence $F_1 = 850$ Hz a $F_2 = 1610$ Hz. Nakreslete do komplexní roviny z póly IIR filtru $H(z) = \frac{1}{A(z)}$ čtvrtého řádu, který modeluje tvorbu tohoto fonému. Vzorkovací frekvence $F_s = 8000$ Hz. Pomůcka: pro filtr 4. řádu budou póly 4.

4. Je dáno 16 vzorků signálu $x[n]$: [1 0 -1 0 1 0 -1 0 1 0 -1 0 1 0 -1 0]
Určete koeficienty a_1 a a_2 prediktoru druhého řádu. Správnost výpočtu zkontrolujte pomocí intuice “jak mají vypadat koeficienty a_1 a a_2 filtru $1 - A(z) = -a_1z^{-1} - a_2z^{-2}$, aby tento filtr co nejlépe predikoval signál $x[n]$?

5. Nakreslete schéma kódování řeči DPCM (diferenční pulsní kódová modulace). Stačí pouze kodér.

6. Jaká je funkce dlouhodobého prediktoru (long-term predictor LTP) a jak se realizuje v CELP kodérech řeči ?

7. V tabulce jsou prvky dvou desetirozměrných vektorů \mathbf{x} a \mathbf{y} . Vypočítejte jejich Euklidovu vzdálenost.

i	1	2	3	4	5	6	7	8	9	10
x_i	4	3	2	1	6	-1	8	-2	5	3
y_i	9	3	-3	0	6	-1	15	-2	5	3

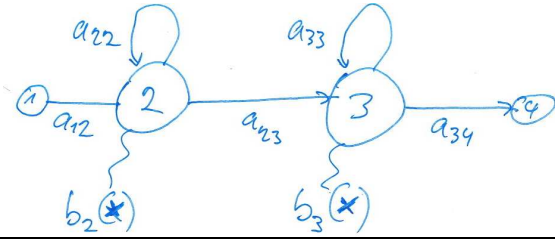
8. Mřížka lokálních vzdáleností u DTW je následující (reference svisle, test vodorovně). Dopočítejte mřížku částečných kumulovaných vzdáleností a určete DTW vzdálenost mezi testem a referencí. Nezapomeňte na normalizaci.

3	6	1
5	2	8
3	2	4
1	1	1

9. Při běhu DTW jsou do matice \mathbf{R} ukládána lokální rozhodnutí “odkud jsem se sem dostal”. Matice má N_{REF} řádků (délka referenční promluvy v rámcích) N_{TEST} sloupců (délka testovací promluvy v rámcích). Hodnota $R[i, j]$ je 0, pokud to bylo z $[i-1, j-1]$, hodnota 1 pro $[i, j-1]$ a hodnota -1 pro $[i-1, j]$. Napište C kód nebo pseudo-algoritmus pro určení optimální cesty pomocí back-tracingu: naplňte vektory \mathbf{r} pro indexování reference a \mathbf{t} pro indexování testu a najděte počet kroků cesty K . Hodnoty v \mathbf{r} a \mathbf{t} mohou být “odzadu”, nemusíte je reverzovat.

10. Nakreslete průběh 2D hustoty pravděpodobnosti daný modelem se směsí dvou Gaussovek (Gaussian Mixture model). Gaussovky mají střední hodnoty: $\boldsymbol{\mu}_1 = [1 \ 1]^T$, $\boldsymbol{\mu}_2 = [1 \ 2]^T$, kovarianční matice $\boldsymbol{\Sigma}$ je pro obě Gaussovky stejná — diagonální s hodnotami 1 na diagonále, váhy jsou $w_1 = 0.7$, $w_2 = 0.3$. Můžete použít 3D zobrazení, barvičky, jak chcete.

11. Levo-pravý skrytý Markovův model (HMM) má dva vysílací stavy s funkcemi hustoty rozdělení pravděpodobnosti $b_2(\mathbf{x})$ a $b_3(\mathbf{x})$. Vstupní sekvence má 3 vektory: \mathbf{x}_1 , \mathbf{x}_2 , \mathbf{x}_3 . Napište vzorec pro výpočet Baum-Welchovy likelihood vyslání této sekvence modelem.



12. Popište (slovně nebo pseudo-algortmem) princip token-passingu při Viterbiho dekódování pomocí HMM.

13. Je definován levo-pravý HMM se čtyřmi stavy, z toho 2 vysílací, přechodové log. pravděpodobnosti jsou:
 $\log a_{12} = 0$, $\log a_{22} = -0.51$, $\log a_{23} = -0.92$, $\log a_{33} = -0.36$, $\log a_{34} = -1.2$.

Tabulka logaritmů hodnot funkcí hustoty vysílacích likelihoodů je:

t	...	46	47	48	...
$\log b_2(\mathbf{x}_t)$...	-1	-2	-3	...
$\log b_3(\mathbf{x}_t)$...	-4	-5	-6	...

Provádíme Viterbiho algoritmus pomocí "token passing". Hodnota tokenu ve stavu 2 v čase 46 je $\Psi_2(46) = -21$. Určete hodnotu tokenu ve stavu 2 v čase 48.

$\Psi_2(48) = \dots\dots\dots$

14. K čemu je při rozpoznávání řeči s velkým slovníkem (large vocabulary continuous speech recognition, LVCSR) dobrý model jazyka (language model, LM) a jak se implementuje ?

15. Jak se v LVCSR řeší kontextová závislost fonémů (např. to, že 'm' uvnitř slova "máma" zní jinak než ve slově "mamka") ?

16. Co se v LVCSR označuje jako pruning a proč je potřeba ? Pomůcka: vzpomeňte si na projekt.

17. Uveďte příklad jakéhokoliv použití neuronových sítí (NN) v LVCSR.

18. Nakreslete blokové schéma jakékoliv metody pro ověřování mluvčího (speaker verification, SV).

19. Uveďte, jak lze měřit úspěšnost SV.

20. Vysvětlete, proč je potřeba v SV systémech normalizace skóre a jak se může realizovat.