

Semestrální zkouška ZRE, řádný termín, 7.5.2013, skupina A

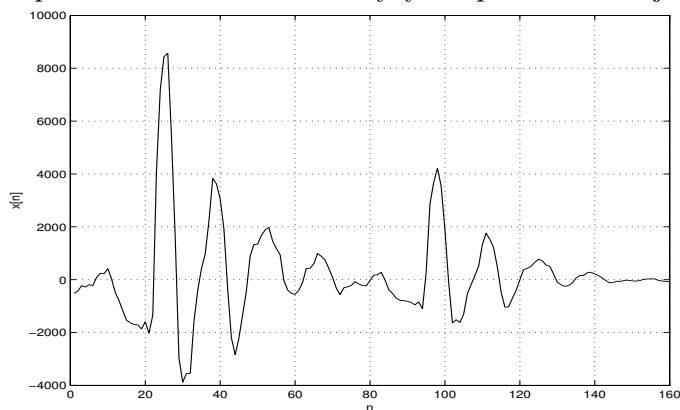
Login: Příjmení a jméno: Podpis:
(prosím čitelně!)

1. Signál je dán rovnicí $x[n] = \cos(\frac{2\pi}{8}n)$. Určete, kolik průchodů nulou bude na 160-vzorkovém rámci takového signálu.

-
2. Je dán signál, který má tři vzorky: $x[0] = -1$, $x[1] = 0$, $x[2] = 1$. Proveďte LPC analýzu řádu 2 — jakýmkoliv způsobem určete koeficienty a_1 a a_2 polynomu $A(z)$.

-
3. Ve spektru znělé řeči má první formant frekvenci $F_1 = 500$ Hz, druhý formant $F_2 = 2000$ Hz. Vzorkovací frekvence je $F_s = 8000$ Hz. Řeč modelujeme filtrem $H(z) = \frac{1}{A(z)}$ čtvrtého řádu. Nakreslete rovinu “z” a vyznačte do ní přibližně póly přenosové funkce $H(z)$.

-
4. Na obrázku vidíte rámeček znělého signálu $x[n]$ (hláska “a”). Určujete frekvenci základního tónu pomocí klipování a autokorelační analýzy. Klipovací úroveň je určena jako $c_L = 0.8 \max(|x[n]|)$. Jaký bude výsledek?



-
5. Proč není určování základního tónu 100% úspěšné? Stačí uvést několik příčin, nemusíte psát, jak určování zpřesnit.
-

6. Napište, jakých hodnot může nabývat logaritmická spektrální vzdálenost

$$d_2 = \sqrt{\int_{-1/2}^{+1/2} |V(f)|^2 df}, \text{ kde } V(f) = 10 \log G(f) - 10 \log \hat{G}(f),$$

$G(f)$ a $\hat{G}(f)$ značí spektrální hustoty výkonu původního a kódovaného rámce řeči.

7. Vysvětlete pojem “adaptivní kódová kniha” v CELP kódování řeči.

8. Napište kód v jazyce C nebo “pseudokód” pro implementaci základní varianty DTW. Předpokládejte, že v poli d jsou v prvcích $d[1][1] \dots d[R][T]$ uloženy lokální vzdálenosti vstupních vektorů každý s každým. R je počet referenčních vektorů, T je počet testovacích vektorů. Pole $g[0][0] \dots g[R][T]$ je již alokované a má v nultém řádku a nultém sloupci uloženy velmi vysoké hodnoty. Stačí vypočítat pouze DTW vzdálenost, back-tracing optimální cesty neřešte. Nezapomeňte na normalizaci.

9. HMM se 4mi stavy (1 je vstupní, 2 a 3 jsou vysílací, 4 je výstupní) má následující matici přechodových pravděpodobností:

$$\mathbf{A} = \begin{bmatrix} 0 & a_{12} & a_{13} & 0 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

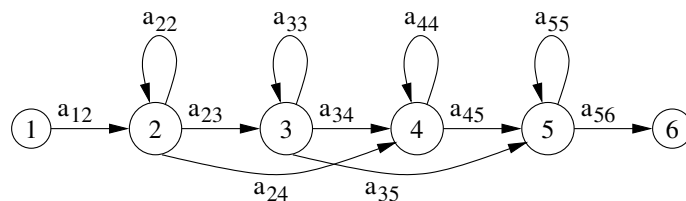
Nakreslete model a napište všechny možné stavové sekvence X pro 3 vstupní vektory. Pomůcka: stavová sekvence musí začínat stavem 1 a končit stavem 4.

10. HMM o $N = 5$ stavech je trénován na jedné nahrávce o délce $T = 50$ rámců. S předchozími parametry modelu je vypočítán průběh pravděpodobnosti 3. stavu (neboli state occupation probability): $\gamma_3(t)$. Uveďte, kolik je:

$$\sum_{t=1}^{50} \gamma_3(t) = \dots\dots\dots$$

Pokud nejde vyhodnotit, napište to.

11. S modelem



probíhá Viterbiho rozpoznávání algoritmem token passing. Hodnotu tokenu ve stavu j v čase t označíme $\Psi_j(t)$. Určete, jak se spočítá hodnota tokenu $\Psi_3(60)$. Předpokládejme, že nahrávka je delší než 60 rámců.

12. Nakreslete jednoduchý konečný stavový akceptor (FSA) a popište, co je co.

13. Základem mnoha rozpoznávačů s velkým slovníkem pro spojitou řeč je vztah: $HCLG = H \circ C \circ L \circ G$. Vysvětlete význam alespoň některých symbolů.

14. Uveďte několik typů nechtěné (škodlivé) variability v rozpoznávání mluvího.

15. V běžných systémech pro ověřování mluvího se většinou nepoužívá *trénování* modelu cílového mluvího, ale *adaptace* tzv. univerzálního modelu pozadí (universal background model, UBM) na cílového mluvího. Proč ?

16. Systémy pro ověřování mluvího (i mnoho jiných detekčních systémů) se vyhodnocují pomocí tzv. detection error tradeoff (DET) křivky. Popište, co je na osách, a proč je DET křivka a ne bod.

17. Co je podstatou fonotaktického přístupu k rozpoznávání jazyka ?

18. V čem se liší použití skrytých Markovových modelů v HMM-syntéze od rozpoznávání ?

19. Proveďte normalizaci textu a převod do fonetické podoby pro větu:

Doc. Jos. Tlustý skončil ve třídě 250 ccm na KTM na 23. místě z 50 závodníků.

Jako fonetickou abecedu použijte češtinu, dbejte, abyste nepoužili kontextové závislosti, tedy ne **tě**, ale **če**, atd.

20. Co se v syntéze řeči (a nejen v systémech rozumí prozódii) ?
