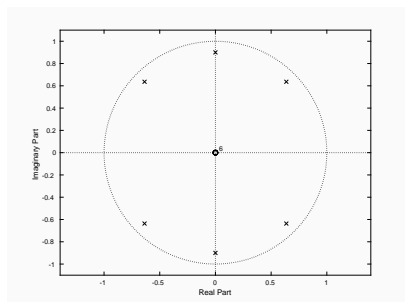


Semestrální zkouška ZRE, řádný termín, 16.5.2016, skupina 07r

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

1. Na obrázku jsou vyznačeny póly číslicového filtru $H(z) = \frac{1}{A(z)}$. Nakreslete přibližně modul jeho kmitočtové charakteristiky od 0 do poloviny vzorkovací frekvence.



-
2. Je dáno 10 vzorků signálu $x[n]$: [1 2 4 8 16 32 64 128 256 512 1024].
Určete koeficient prediktoru prvního řádu (nejlépe bez výpočtu, z hlavy!). Pomůcka: $1 - A(z) = -a_1 z^{-1}$.

-
3. Analýzou rámce signálu byly spočítány následující nenormované autokorelační koeficienty:
 $R[0] = 10$, $R[1] = 0$, $R[2] = 0$, $R[3] = 5$.
Koeficienty prediktoru 2. řádu spočítané algoritmem Levinsona-Durbina jsou $a_2^{(2)} = 0$, $a_1^{(2)} = 0$. Při výpočtu prediktoru 3. řádu vyšel koeficient $k_3 = a_3^{(3)} = -0.5$. Dopočítejte další potřebné koeficienty prediktoru 3. řádu. Nepovinná doplňující otázka: Vysvětlete hodnoty dopočítaných koeficientů.

-
4. Vzorkovací frekvence je $F_s = 8$ kHz. Znělý úsek mužského hlasu má frekvenci základního tónu $F_0 = 100$ Hz. Jsou z něj odhadnuty autokorelační koeficienty $R[k]$. Nakreslete, jak bude vypadat jejich průběh v závislosti na k a vyznačte, jak se z nich určí F_0 .

-
5. Vysvětlete, jak se pro určení frekvence základního tónu dá využít dlouhodobý prediktor. Není potřeba psát rovnice, stačí princip (můžete využít text, graf, schéma, atd).

6. Nakreslete schéma LPC dekodéru a napište, jakou informaci musí přijímat z přenosového kanálu.

7. Máme k dispozici populaci N trénovacích vektorů \mathbf{x}_n pro trénování kódové knihy vektorové kvantizace. Toto trénování začíná tvorbou kódové knihy o velikosti $L = 1$. Napište, jak získáme hodnotu jejího jediného kódového vektoru.

8. Kódujeme pomocí VQ vektory o rozměru $P = 10$. Kódová kniha má velikost $L = 1000$. Optimální kódový vektor vybíráme pomocí minimální Euklidovy vzdálenosti. Určete, kolik operací zabere zakódování jednoho vstupního vektoru. Předpokládejte, že operace sčítání, odečítání, druhé mocniny, odmocniny a porovnání mají všechny stejnou složitost: 1 OP.

9. Při kódování typu CELP hledáme v jedné nebo několika kódových knihách nejlepší buzení. Jak? Není potřeba psát rovnice, stačí princip (můžete využít text, graf, schéma, atd).

10. Vysvětlete, jak se zjistí impulsní odezva s nulovým vstupem (zero impulse response, ZIR), která se používá při kódování CELP. Nemusíte se rozepisovat, proč je to nutné.

11. Referenční i testovací sekvence nejsou sekvence vektorů, ale skalárů: $R = [1 \ 2 \ 3 \ 2]$, $O = [1 \ 1 \ 2 \ 3 \ 0]$. Proveďte plné srovnání pomocí DTW včetně výpočtu mřížky lokálních vzdáleností “každý s každým”, výpočtu mřížky částečných kumulovaných vzdáleností a normalizace výsledku. Vzdálenost dvou skalárů je jednoduše absolutní hodnota jejich rozdílů.

12. Mřížka lokálních vzdáleností u DTW je následující (reference svisle, test vodorovně). Dopačítejte mřížku částečných kumulovaných vzdáleností a určete průběhy indexovacích funkcí $r(k)$ a $t(k)$.

9	16
6	8
1	2

13. Stav skrytého Markovova modelu obsahuje 2-rozměrnou Gaussovku s těmito středními hodnotami a směrodatnými odchylkami: $\boldsymbol{\mu} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\boldsymbol{\sigma} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ (jedná se o Gaussovku s diagonální kovarianční maticí, takže to takto stačí). Vypočítejte skóre této Gaussovky pro vstupní vektor $\mathbf{o} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Pomůcka: $b(\mathbf{o}) = \frac{1}{\sqrt{2\pi} \sigma_1} e^{-\frac{(\sigma_1 - \mu_1)^2}{2\sigma_1^2}} \times \frac{1}{\sqrt{2\pi} \sigma_2} e^{-\frac{(\sigma_2 - \mu_2)^2}{2\sigma_2^2}}$.

14. Je definován levo-pravý HMM se čtyřmi stavy, z toho 2 vysílací, log. přechodové 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 pravděpodobností je:

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

Hodnota tokenu ve stavu 2 v čase 46 je $\Psi_2(46) = -20$. Určete hodnotu tokenu ve stavu 2 v čase 48: $\Psi_2(48) = \dots$

15. Nakreslete strukturu HMM, která bude reprezentovat promluvy “bla”, “blabla”, “blablabla”, atd. Předpokládejte, že jeden stav HMM reprezentuje jeden foném.

16. Dekódovací síť pro rozpoznávání řeči s velkým slovníkem LVCSR se staví jako kompozice váhovaných konečných stavových převodníků WFST (weighted finite state transducers): $HCLG = H \circ C \circ L \circ G$. Vysvětlete, zda je možné některý z převodníků vynechat pro totálně fonetický jazyk (přesně ty znaky, které čteme, také říkáme).
-
17. Uveďte, jak se dají v rozpoznávání řeči modelovat pravděpodobnosti fonémových stavů jinak než Gaussovskými.
-
18. Navrhněte systém pro detekci hanáčtiny. Hanáčtina je charakteristická prodlužováním samohlásek a záměnou “ou” za “ó” (“padnou → “padnó”, atd.)
-
19. Jaké jsou při rozpoznávání mluvího největší zdroje variability pocházející od samotného mluvího (tedy ne mikrofon, přenosový kanál, kodek, atd) ? Uveďte alespoň tři.
-
20. Máte k dispozici LVCSR systém natrénovaný na obecné češtině. Chcete jej prodat do call centra banky a klient má samozřejmě zájem o co největší přesnost. Navrhněte kroky, jak tuto přesnost zvýšit. U každého kroku odhadněte pracnost.
-