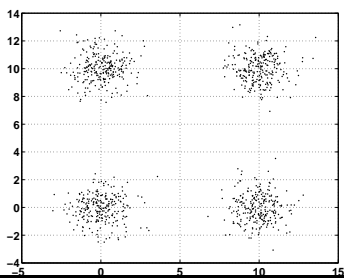


Semestrální zkouška ZRE, řádný termín, 12.5.2015, skupina W

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

1. Nakreslete blokové schéma dekodéru diferenční pulsní kódové modulace (DPCM).

-
2. Pro trénování vektorové kvantizace je k dispozici 1000 dvourozměrných trénovacích vektorů, viz obrázek. Kódová kniha má mít velikost $L = 8$ kódových vektorů. Nakreslete jejich polohy po natrénování.



-
3. Součástí moderních kodérů řeči je dlouhodobý prediktor (long-term predictor, LTP). Uveďte, co se rozumí touto “dlouhou dobou” ?

-
4. O co se v kodérech řeči typu Codebook-Excited Linear Prediction (CELP) stará perceptuální filtr $W(z)$?

-
5. Jak se v kodérech řeči typu Algebraic Codebook-Excited Linear Prediction (A-CELP) konstruuje stochastická kódová kniha, aby se v ní dalo rychle vyhledávat ?
-

6. Je dána matice (mřížka) lokálních vzdáleností DTW (reference svisle, test vodorovně, postup zleva nahoru). Vypočtěte DTW vzdálenost (nezapomeňte na normalisaci!) a vyznačte optimální DTW cestu.

9	5
6	4
1	1

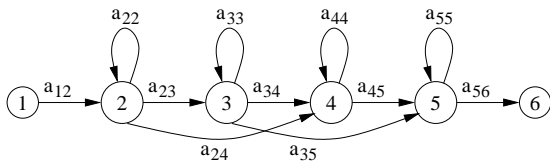
-
7. Máte za úkol zarovnat dva odstavce textu - jeden čistý a druhý s překlipy. Uveďte, zda a jak by se úloha dala řešit pomocí DTW.

-
8. Jedna z technik pro vyhledávání informací v řeči je tzv. "Query by Example" - mluvený dotaz (např. jedno klíčové slovo nebo krátká fráze) se vyhledává v mluvené řeči. Jako jedna z technik pro QbE se dá použít DTW. Napište jak.

-
9. V rozpoznávání řeči pomocí statistických modelů (a prakticky v jakékoliv detekci či rozpoznávání) je základním teoretickým východiskem Bayesův vzorec: $P(W|X) = \frac{p(X|W)P(W)}{p(X)}$. Popište, co jsou $P(W|X)$, $p(X|W)$ a $P(W)$. Věnujte pozornost tomu, co jsou pravděpodobnosti a co ne.

-
10. Máte k dispozici 4 trénovací vektory, které tvoří matici: $\mathbf{X} = \left[\begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix} \begin{bmatrix} -2 \\ -2 \end{bmatrix} \right]$. Tyto vektory modelujete jednou dvourozměrnou Gaussovskou. Vypočtěte její kovarianční matici. Pomůcka: $\Sigma = \frac{1}{T} \sum_{t=1}^T (\mathbf{x}(t) - \boldsymbol{\mu})(\mathbf{x}(t) - \boldsymbol{\mu})^T$.
-

11. Určete všechny možné sekvence stavů pro 4 vstupní vektory pro skrytý Markovův model HMM na obrázku. Stavy 1 a 6 nejsou vysílací a nepatří k nim žádný vektor, ale 1 je povinný na začátku a 6 na konci stavové sekvence.



-
12. Popište princip trénování HMM v nejjednodušší verzi, tedy jedna promluva, jeden model. Není nutné vypisovat detailně rovnice, stačí princip.

-
13. Nakreslete rozpoznávací síť pro rozpoznávání sekvencí slov “VPRAVO” a “VLEVO”, slova se mohou opakovat. Jazykový model neuvažujte. Slova jsou složena z modelů fonémů.

-
14. Co je to kontextová závislost fonémů a jak s ní v rozpoznávání řeči pracujeme ?

-
15. 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 význam, vstup a výstup alespoň dvou z těchto převodníků.
-

16. Co se v rozpoznávání řeči s velkým slovníkem LVCSR označuje pojmem “lattice” a co taková lattice obsahuje ? Pomůcka: jedná se o jednu z možných forem výstupu rozpoznávače.

17. Srovnajte rozpoznávání osob podle hlasu s rozpoznáváním pomocí otisků prstů z hlediska získání vstupních dat.

18. Uveďte, co jsou v ověřování mluvčího podle hlasu (a obecně v jakékoliv detekční úloze) chyby typu “Miss” a “False Alarm” a na čem závisí počty těchto chyb.

19. Uveďte princip ověřování mluvčího pomocí projekce do podprostoru totální variability, tzv. iVektorů.

20. Jak vypadá a z čeho se dá odhadnout fonotaktický model daného jazyka v rozpoznávání jazyka (language identification, LID) ?
