

SGN-4010 Puheen käsittelyn menetelmät

Tentti 1.2.2008

Vastaa kaikkiin viiteen tehtävään. Kirjoita kaikki vastaukset vastauspaperille. Huomaa, että tentti jatkuu paperin kääntöpuolella.

Tehtävä 1

Mikä vaihtoehdoista pitää paikkansa? (Oikea vastaus: +1p/kohta, väärä vastaus: -0.5p/kohta, kuitenkin niin, että tehtävän kokonaispistemäärä on vähintään 0.)

- Mitä taajuusresoluutiolle tapahtuu, kun analyysi-ikkunan pituutta kasvatetaan?
 - Taajuusresoluutio paranee
 - Taajuusresoluutio huononee
 - Ikkunan pituuden kasvattaminen ei vaikuta taajuusresoluutioon
- Signaalin s_1 DFT (diskreetti Fourier-muunnos) on S_1 . Signaali s_2 muodostetaan kopiaimalla signaalia s_1 ja liittämällä kopiot alkuperäisen signaalin s_1 perään. Millainen on signaalin s_2 DFT S_2 ?
 - Kuin S_1 , mutta tiheämmin näytteistetty
 - Kuin S_1 , mutta jokaisen tapin väliin on lisätty nollia
 - S_1 :n perusteella ei voida päätellä mitään S_2 :sta
- Mikä siirtofunktioista kuvaa all-pole -suodinta?
 - $H(z) = C(z)$
 - $H(z) = 1/D(z)$
 - $H(z) = C(z)/D(z)$
- Puhujan perustaaajuus on 150 Hz. Mikä on ensimmäisen formantin taajuus?
 - 300 Hz
 - 450 Hz
 - Taajuutta ei voi päätellä annettujen tietojen perusteella
- Signaalin näytteenottotaajuus on 10 kHz. Järkevä valinta lineaarisen ennustuksen (linear prediction) asteeksi olisi
 - 5
 - 12
 - 100
- Tontun ääntöväylän pituus on 10 cm. Ääntöväylää mallinnetaan toisesta päästä avoimen ja toisesta päästä suljetun tasapaksun putken avulla. Millä taajuuksilla tontun puheen kolme ensimmäistä formanttia mallin mukaan ovat? Äänen nopeus ilmassa on 340 m/s.
 - 850 Hz, 1700 Hz, 2550 Hz
 - 425 Hz, 1275 Hz, 2125 Hz
 - 850 Hz, 2550 Hz, 4250 Hz

Tehtävä 2

Selitä lyhyesti seuraavat termit (1p/kohta) a) allofoni b) TTS c) formanttisynteesi d) Levinson-Durbin -rekursio e) ARMA-malli f) formantin kaistanleveys.

Tehtävä 3

- Miten äänihuulet käyttäytyvät, kun ihminen sanoo vokaalin [a]? Entä kun ihminen sanoo konsonantin [s]? (2p)
- Minkä kolmen ominaisuuden perusteella vokaalit voidaan luokitella? (2p)
- Minkä kolmen ominaisuuden perusteella konsonantit voidaan luokitella? (2p)

Tehtävä 4

Lyhyehkön puhekehyyksen $s(n)$ näytteet ovat:

n	0	1	2	3	4	5	6
$s(n)$	1	2	0	-2	4	2	3

- Laske kehyyksen autokorrelaatiokertoimet $r(0)$ ja $r(3)$. (1p)
- Laske kehyykselle $s(n)$ optimaaliset 1. asteen lineaarisen ennustuksen kertoimet $a_1(0)$ ja $a_1(1)$. (2p)
- Optimaaliset 3. asteen lineaarisen ennustuksen kertoimet kehyykselle $s(n)$ ovat $a_3(0) = 1$, $a_3(1) = -0.1994$, $a_3(2) = -0.0708$ ja $a_3(3) = 0.0359$.
 - Mikä on 3. asteen lineaarisen ennustuksen ennustusvirheen energia E_3 ? (1p)
 - Määritä optimaaliset 4. asteen ennustuskertoimet $a_4(0)$, $a_4(1)$, $a_4(2)$, $a_4(3)$ ja $a_4(4)$. (2p)

Tehtävä 5

- Miten puhesignaalin perustaaajuus saadaan määritettyä ja mihin käytetään SIFT-menetelmää (simple inverse filter tracking) (3p)
- Miten difonisynteesi ja unit selection -synteesi eroavat toisistaan? Mitä yhteistä menetelmissä on? (3p)