

Beszédinformációs rendszerek 2011. április 22. ZH

A feleletválasztós teszteknel a kihagyott válasz 0 pont, rossz válasz -1 pont. Előfordul, hogy több jó válasz is van.

1) Feleletválasztós kérdések, a példamondat: "10 cm hó esett.", a megvalósítás diádokat és cvc triádokat használó szintetizátorral történik (az első 4 kérdés ehhez kapcsolódik, a többi általános) (40p).

Hány helyre kell szünet a mondat elejét és a végét leszámítva?

3 --- 2 --- 1 --- 11 --- 0 --- 10

0 : nem tartasz szünetet, miközben kimondod a mondatot

Milyen hangsorozattal kezdődik a szintetizálás?

tízcm --- tízce --- tíce --- tízse --- tízze --- tízsce

tízsce: a z zöngétlenedését a szintetizálás során figyelembe kell venni

Melyik építőelem nem szerepel a szintetizált mondatban?

nt --- erh --- t_ --- mét --- set --- tim

A szintetizált mondatban melyik hangnál legnagyobb az alapfrekvencia?

s --- e --- h --- é --- í --- t

í: Kijelentő mondat elején magasabb az alapfrekvencia és mivel a *t* zöngétlen, rajta nem tudjuk értelmezni az alapfrekvenciát (ezért az *í* és nem a *t*).

Melyik érték lehet egy női beszélő magánhangzójának első formánsfrekvenciája?

200 Hz --- 50 Hz --- 3400 Hz --- 4500 Hz --- 800 Hz --- 100 Hz

800 Hz : A CD-ről: " A formánsokat sorszámmal jelölik, a legalacsonyabb az első formáns (F1). Ez a magyar beszédhangoknál a 200-800 Hz-es sávba esik." Tehát a 200 alattiak a mélyek, és mivel nyilván a férfiaké a mélyebb, ezért a 200 is, a 800 fölöttiek szintén kiesnek, így csak a 800 a jó megoldás.

Melyik szinuszjelnek a legkisebb az intenzitása?

5 kHz / 90 phon --- 2 kHz / 71 phon --- 22 kHz / 60 dB --- 1 kHz / 50 dB --- 500 Hz / 100 dB --- 20 Hz / 49 phon

1kHz / 50dB asszem.

Melyik esetben legnagyobb a kvantálási zaj a maximális kivezérlésű szinuszjelhez képest?

13 bit lineáris --- 32 bit lineáris --- 8 bit lineáris --- 16 bit lineáris --- A-törvényű log --- u-törvényű log

8 bit lineáris

Melyik hangnak nincs párja zöngésség szempontjából?

t --- g --- s --- k --- r --- p

t(d), g(k), s(zs), k(g), p(b), az *r* -nek nincs párja, tehát *r*.

7 kHz-es szinuszjelet mintavételezünk 10 kHz-en. Hol jelenik meg a bemeneti szinuszjel?

3 kHz-en --- Sehol --- 3,5 kHz-en --- 2 kHz-en --- 1 kHz-en --- 5 kHz-en

3 kHz

Melyiknek nincs se fojtott zöngéje, se néma fázisa?

p --- dz --- t --- sz --- k --- d

sz: réshangoknak nincs fojtott zöngéje/néma fázisa.

2) Egy jel 6,0 kHz és 10,8 kHz közötti komponensekből áll, melyet közvetlenül digitálissá kell tenni. Adja meg az összes célszerűen választható mintavételezési frekvenciát

a) végtelen (nem realizálható) szűrőmeredekség mellett (8p)

b) megvalósítható szűrőmeredekség mellett (8p)

3) Van egy hullámforma állomány A-törvényű kvantálással, 8 kHz-en mintavételezve: "Elindult a vonat." (16p)

a) Írja le pontokba szedve, hogy hogyan lehet ebből kérdő mondatot csinálni PSOLA-algoritmust használva.

b) Adja meg az alapfrekvencia változás törtvonalas közelítését (x: idő, y: alapfrekvencia) hangszinten az eredeti és az átalakított mondatban.

c) Adja meg a legkisebb és legnagyobb F_0 értéket, és ezek pontos helyét hangszinten az átalakított mondatban.

d) Mit kell változtatni a megoldáson, ha a felvétel 48 kHz-es stúdióban készült, de az átalakított változatot 16 kHz-en, 12 bites kvantálású kódolással akarjuk lejátszani?

4) Rajzolja fel egy emberi beszédelemekből építkező hullámforma összefűzéses szövegfelolvasó blokk-sémáját. Írja le az egyes blokkok működését. Hogyan lesz a szövegből beszéd? (10p)

5) Beszédadatbázis tervezésében működik közre. Az eredményt beszédfelismerő akusztikai modell betanításához szeretnék felhasználni, és Önt kérik fel az annotálásra.

a) Mit és hogyan kell annotálni (paraméterek, események)? (10p)

b) Automatizálható-e a folyamat? Ha igen, miért, ha nem miért? (8p)

Megoldásokat szívesen várunk, ahogy elnéztem elég kevesen fogunk átmenni... -- [Maestro](#) - 2011.04.23.