DT1130 Spektrala Transformer
Tentamen 200114

Tentamen består av fem uppgifter där varje uppgift maximalt ger 4 p. Normalt gäller följande betygsgränser: **E: 9 p, D: 11.5 p, C: 14 p, B: 16 p, A: 18 p**
Tillåtna hjälpmedel: räknare, formelblad (bifogat)

Lycka till!

1

Som en del i antagninsprocessen vid ett företag inom musikteknologi får de sökande uppgiften att skriva kod i python för att sampla ned en ljudfil från 44100 Hz till 22050 Hz. Nedan listas kod från två av de sökande.

Sökande A

```
import numpy as np
import scipy.io.wavfile as wav

fs,X = wav.read('44100.wav')
Y = np.zeros(int(len(X)/2))

for n in range(0,len(Y)):
  Y[n] = X[n*2]

wav.write('22050.wav',
          int(fs/2),Y.astype('int16'))
```

Sökande B

```
import scipy.io.wavfile as wav
import numpy as np

fs,X = wav.read('44100.wav')
Y = np.zeros(int(len(X)/2))

for n in range(0,len(Y)):
  Y[n] = (X[n*2] + X[n*2+1])/2

wav.write('22050.wav',
          int(fs/2),Y.astype('int16'))
```

(a) Förklara vad A:s kod gör, i egna ord (1p)

(b) Vad verkar B ha tänkt på som A har missat? (1p)

(c) Som en del i utvärderingen provkör man koden med en 10 sekunder lång testfil som innehåller en sinuston vars frekvens ökar konstant från 0 till 22 kHz. När man lyssnar på utsignalen från A:s kod, hörs en uppåtgående ton, men efter ca 5 sekunder vänder tonen och går *nedåt*, hela tiden med konstant ljudstyrka. För B händer i princip samma sak men tonens styrka avtar under den nedåtgående fasen. Förklara vad det kan bero på. (1p)

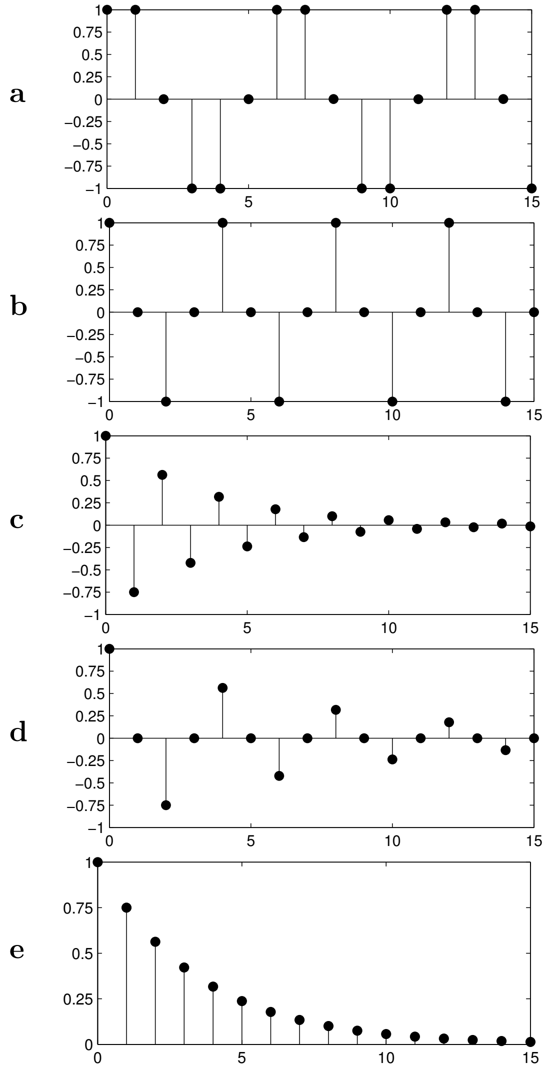
(d) Hur skulle det låta om nedsamplingen var korrekt gjord, och hur ge ett förslag på hur programmet skulle kunna ändras för att åstadkomma detta (i ord, ej kod). (1p)

1 $H(z) = 1 + \frac{3}{4}z^{-1}$

2 $H(z) = 1 - z^{-1} + z^{-2}$

3 $H(z) = 1 + \frac{3}{4}z^{-2}$

4 $H(z) = 1 - \frac{3}{4}z^{-1}$



Figur 1. Vilken av signalerna till höger ska skickas till vilket filter till vänster för att utsignalen ska bli en impuls?

2

I figur 1 ser du överföringsfunktioner för ett antal filter samt ett antal signaler. För varje filter till vänster finns en signal till höger som, när den matas in i filtret som $x(n)$, resulterar i en utsignal som är en *impuls*, dvs $y(n) = \delta(n)$. Dessa ska paras ihop, med tydlig motivering!

(1p/ korrekt motiverat par)

3

Eva testar en ny metod för att göra syntetiskt tal, med en så kallad *neural vocoder* där ett djupt neuralt nätverk, som tränas på hundratals timmar tal, genererar talsignalen[1]. Hon är i stort sett nöjd med vad som kommer ut, bortsett från ett ettrigt, högfrekvent missljud som skär igenom alla ljudfiler som genereras. Eva kontaktar upphovsmannen till medtoden som meddelar att han känner till störningen men säger att den vanligtvis försvinner efter några hundratusen ytterligare

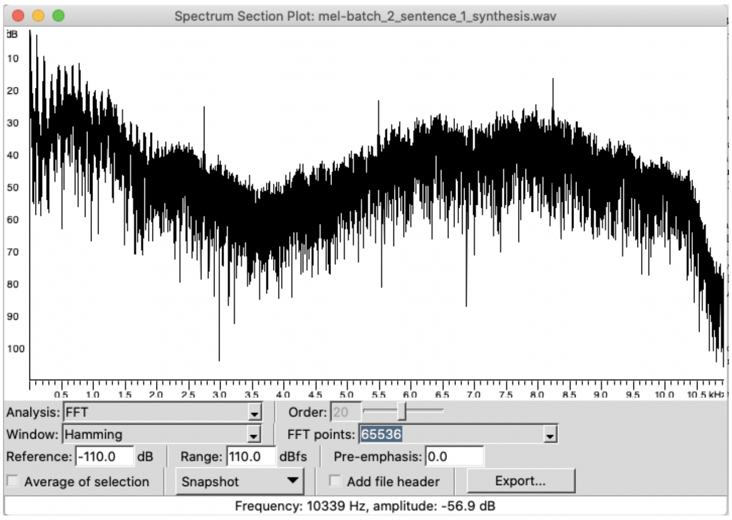
[1] https://github.com/NVIDIA/waveglow

träningscykler. Eva är inte nöjd med det beskedet eftersom det skulle ta ca två månader på hennes dator... Hon bestämmer sig därför för att undersöka om det finns något annat att göra. Hon plottar långtidsspektrum över signalen (se figur 2) och ser genast vad problemet är: tre smala toppar i spektrum som verkar ligga exakt på frekvenserna $\frac{f_s}{8}$, $\frac{f_s}{4}$ och $\frac{3f_s}{8}$.

Hjälp Eva att göra ett filter som tar bort dessa tre frekvenser, genom att placera nollställen på rätt ställen i z-planet!

(a) Ange filterekvationen $y(n) = ...$ för det resulterande filtret (3p)

(b) Vad är filtrets förstärkning vid $\omega = 0$? (1p)



Figur 2. Spektrumplot av ljudsignalen från Evas neurala vocoder.

4

En tvådimensionell filterkärna beskrivs av ekvationen $h(x, y) = e^{-(x^2+y^2)/r^2}$.

(a) Vad kallas ovanstående typ av filterkärna, som ofta används vid lågpassfiltrering av bilder? (1p)

(b) Filterkärnan ovan är linjärt separerbar, vilket innebär att funktionen $h(x, y)$ kan uttryckas som en produkt av två andra funktioner, av x respektive y. Visa detta. (1p)

(c) Filterkärnan är även cirkulärt symmetrisk, viket innebär att den kan uttryckas som en funktion av avståndet från mitten av kärnan. Visa det. (1p)

(d) Beskriv hur du skulle gå tillväga för att göra ett högpassfilter med utgångspunkt i denna filterkärna. (1p)

5

Tre linjära filter $H_1(z)$, $H_2(z)$ och $H_3(z)$ kaskadkopplas enligt nedan. Insignalen till det första filtret, $x_1(n)$ är en impuls, och utsignalen från det andra filtret, $y_2(n)$ ges av figur 3 I övrigt gäller att

$$H_1(z) = \frac{1 - z^{-3}}{1 - z^{-1}}$$

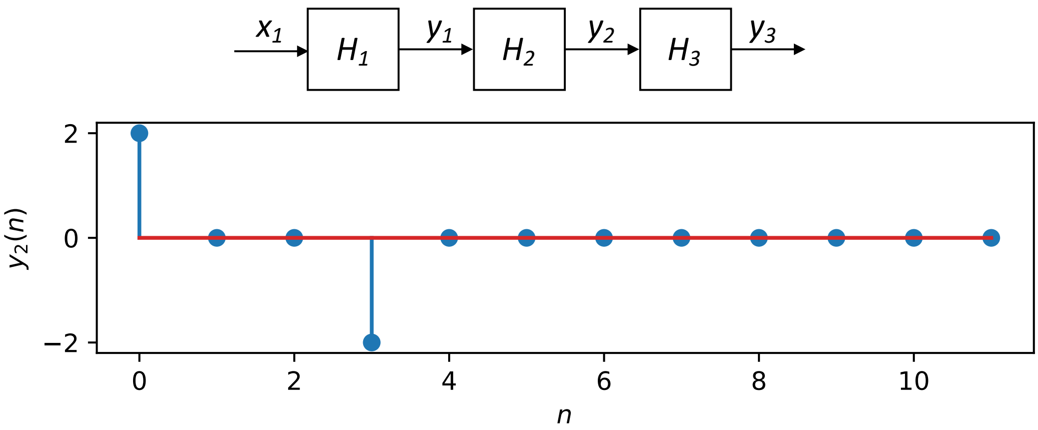
samt

$$H_3(z) = \frac{1}{1 - z^{-3}}$$

(a) Vad är utsignalen från H_1, $y_1(n)$? (1p)

(b) Vad är $H_2(z)$? (2p)

(c) Vad är utsignalen från H_3, $y_3(n)$? (1p)



Figur 3. Tre kaskadkopplade filter, samt utsignalen från det andra filtret i kedjan.

Lösningar

1

(a) en ljudfil läses in i vektorn X och en tom vektor Y med halva längden skapas. Vartannat sampel från X läggs in i Y, varpå Y sparas i en ny fil.

(b) B har försökt ta hänsyn till vikning/aliasing genom att sätta varje sampel Y till medelvärdet av två närliggande sampel i X.

(c) Efter ca 5s går frekvensen över 11025, dvs halva den nya samplingsfrekvensen, vilket leder till vikning, därav den nedåtgående tonhöjden. I B:s implementation tas ett medelvärde av två sampel i X vilket motsvarar filtrering med $H(z) = \frac{1+z^{-1}}{2}$ av X. Detta är ett lågpassfilter (rullande meldelvärde av längden 2) som kommer dämpa de höga frekvenserna innan de viks ned.

(d) Korrekt nedsampling innebär att inga frekvenser över $f_s/2(11025Hz)$ får förekomma. Detta kan åstadkommas med ett bättre lågpassfilter som appliceras på X *innan* nedsamplingsloopen där vartannat sampel plockas ut. Då skulle man endast höra en uppåtgående ton fram till 5s, sedan tystnad.

2

Om vi antar att varje signal till höger är ett impulssvar från ett visst filter, så ska vi hitta ett filter till vänster som utgör *inversen* till detta. Då kommer utsignalen bli en impuls.

Vi noterar att alla filter till höger är icke-återkopplde, så strategin är att hitta *nollställena* för dessa filter, och sedan leta efter impulssvar från filter med *poler* på dessa platser (inversen till ett filter fås genom att byta poler mot nollställen och vice versa).

1.

$$H(z) = 1 + \frac{3}{4}z^{-1} = \frac{z + \frac{3}{4}}{z}$$

Nollställe: $z_1 = -\frac{3}{4}$. Ett filter med en pol i z_1 ger polvinkel π och radie 0.75 $\rightarrow$ dämpad svängning med $\omega = \pi$ och periodtid $\frac{2\pi}{\omega} = 2$ sampel, stämmer på c.

2.

$$H(z) = 1 - z^{-1} + z^{-2} = \frac{z^2 - z + 1}{z^2}$$

Nollställen (använd pq-formel): $z_{1,2} = \frac{\sqrt{3} \pm j}{2}$. Ett polpar med polvinkel $\omega = \frac{\pi}{3}$ och radie 1 $\rightarrow$ svängning med periodtiden $\frac{2\pi}{\omega} = 6$ sampel och ingen dämpning. Stämmer på a.

3.

$$H(z) = 1 + \frac{3}{4}z^{-2} = \frac{z^2 + \frac{3}{4}}{z^2}$$

Nollställen: $z_{1,2} = \pm\frac{j\sqrt{3}}{2}$ Ett polpar med polvinkel $\omega = \frac{\pi}{2}$ och radie mindre än ett $\rightarrow$ dämpad svängning med periodtiden $\frac{2\pi}{\omega} = 4$ sampel. Stämmer på d.

4.

$$H(z) = 1 - \frac{3}{4}z^{-1} = \frac{z - \frac{3}{4}}{z}$$

Nollställe: $z_1 = \frac{3}{4}$. Ett filter med en pol i z_1 ger polvinkel $\omega = 0$ och radie $0.75 \rightarrow$ avklingning utan svängning, stämmer på e.

3

För att eliminera de tre störfrekvenserna vid $\frac{f_s}{8}$, $\frac{f_s}{4}$ och $\frac{3f_s}{8}$ behöver filtret ha nollställen på enhetscirkeln vid $\omega_1 = \pi/4$, $\omega_2 = \pi/2$ och $\omega_3 = 3\pi/4$. Vi åstadkommer detta med tre filter i serie:

$$H(z) = H_1(z)H_2(z)H_3(z)$$

där varje filter kan skrivas som

$$H_k(z) = \frac{(z - e^{j\omega_k})(z - e^{-j\omega_k})}{z^2}$$

Filter 1:

$$\omega_1 = \pi/4 \rightarrow H_1(z) = \frac{(z - e^{j\pi/4})(z - e^{-j\pi/4})}{z^2} = \frac{z^2 - \sqrt{2}z + 1}{z^2}$$

Filter 2:

$$\omega_2 = \pi/2 \rightarrow H_2(z) = \frac{(z - e^{j\pi/2})(z - e^{-j\pi/2})}{z^2} = \frac{z^2 + 1}{z^2}$$

Filter 3:

$$\omega_3 = 3\pi/4 \rightarrow H_3(z) = \frac{(z - e^{j3\pi/4})(z - e^{-j3\pi/4})}{z^2} = \frac{z^2 + \sqrt{2}z + 1}{z^2}$$

Multiplicera ihop polynomen i täljare respektive nämnare för att få totala överföringsfunktionen $H(z)$:

$$H(z) = H_1(z)H_2(z)H_3(z) = \left(\frac{z^2 - \sqrt{2}z + 1}{z^2}\right)\left(\frac{z^2 + 1}{z^2}\right)\left(\frac{z^2 + \sqrt{2}z + 1}{z^2}\right)$$

$$= \frac{z^6 + z^4 + z^2 + 1}{z^6} = 1 + z^{-2} + z^{-4} + z^{-6}$$

(a) detta ger oss filterekvationen

$$y(n) = x(n) + x(n-2) + x(n-4) + x(n-6)$$

(b) förstärkningen vid $\omega = 0$ får vi genom att sätta in $z = e^{j0} = 1$ i uttrycket för $H(z)$:

$$H(1) = 1^6 + 1^4 + 1^2 + 1 = 4$$

4

(a) Gaussfilter eller gausskärna

(b)

$$h(x,y) = e^{-\frac{x^2+y^2}{r^2}} = e^{-\frac{x^2}{r^2}} e^{-\frac{y^2}{r^2}}$$

där första faktorn i uttrycket beror av x och den andra av y *Q.E.D.*

(c) Pythagoras stats ger avståndet från origo: $d = \sqrt{x^2 + y^2} \to d^2 = x^2 + y^2$ Om vi byter ut $x^2 + y^2$ mot d^2 i uttrycket för $h(x,y)$ får vi en funktion av avståndet från mitten dvs cirkulär symmetri *Q.E.D.*

(d) Tar vi skillnaden mellan en originalbilden och en lågpassfiltrerad bild återstår endast de höga frekvenserna (förutsatt att lågpassfilterkärnan är normaliserad, dvs kärnan summerar till ett) Vill vi göra om en filterkärna $h_{LP}(x,y)$från lågpass till högpass gör vi motsvarande operation: *originalbild* motsvaras då av en 2D-impuls (värdet ett vid $x = 0, y = 0$ och noll annars - dvs en filterkärna som inte förändrar bilden alls), så vi tar skillnaden mellan impulsen och lågpassfilterkärnan:

$$h_{HP}(x,y) = \delta(x,y) - \frac{h_{LP}(x,y)}{\sum_x \sum_y h_{LP}(x,y)}$$

5

(a) För att få impulssvaret från H_1 tar vi fram filterekvationen

$$H_1(z) = \frac{1 - z^{-3}}{1 - z^{-1}} = \frac{Y(z)}{X(z)}$$

$$\to Y(z)(1 - z^{-1} = X(z)(1 - z^{-3})$$

$$\to y(n) = x(n) - x(n-3) + y(n-1)$$

och sätter in $x(n) = \delta(n)$:

$$y(0) = x(0) - x(-3) + y(-1) = 1 - 0 + 0 = 1$$

$$y(1) = x(1) - x(-2) + y(0) = 0 - 0 + 1 = 1$$

$$y(2) = x(2) - x(-1) + y(1) = 0 - 0 + 1 = 1$$

$$y(3) = x(3) - x(0) + y(2) = 0 - 1 + 1 = 0$$

$$y(4) = x(4) - x(1) + y(3) = 0 - 0 + 0 = 0$$

osv...

alltså:

$$y_1(n) = [1, 1, 1, 0, 0, 0, 0, 0.....]$$

(b) från figur:

$$y_2(n) = [2, 0, 0, -2, 0, 0, 0, 0...]$$

Vi ska bestämma filtret H_2 givet insignal $(y_1(n))$ och utsignal $(y_2(n))$. y_1. y_1 har längden 3 och y_2 har längden 4. alltså vet vi att H_2 måste innehålla en fördröjning (men inte fler, för då skulle y_2 vara ännu längre) dvs vi kan skriva filtret H_2 som ett icke återkopplat filter av längden ett:

$$y(n) = b_0 x(n) + b_1 x(n-1)$$

där $x(n) = y_1(n)$ och $y(n) = y_2(n)$, dvs

$$y_2(n) = b_0 y_1(n) + b_1 y_1(n-1)$$

Vi sätter in kända värden vid $n = 0$:

$$y_1(0) = 1, y_2(0) = 2 \rightarrow 2 = b_0 \cdot 1 + b_1 \cdot 0 \rightarrow b_0 = 2$$

samt vid $n = 3$:

$$y_1(3) = 0, y_1(2) = 1, y_2(3) = -2 \rightarrow -2 = 2 \cdot 0 + b_1 \cdot 1 \rightarrow b_1 = -2$$

vilket ger

$$H_2(z) = 2 - 2z^{-1}$$

(c) Vi ska bestämma hela filterkaskadens utsignal när insignalen är $\delta(n)$. Vi vet nu H_1,H_2 och H_3 så vi kan få fram totala överföringsfunktionen:

$$H(z) = \left(\frac{1 - z^{-3}}{1 - z{-1}}\right)(2 - 2z^{-1})\left(\frac{1}{1 - z^{-3}}\right) = 2$$

Detta ger

$$y_3(n) = 2\delta(n)$$