Prepoznavanje govora

DFT

Discrete Fourier Transform

- Cilj nam je da od sirovog zvučnog signala dobijemo spektrogram.
- Osnovni periodični signali – sinusoidalni.
- Zbir više periodičnih signala je takođe periodičan signal.
- Moguće je razložiti zvučni signal na njegove sabirke – generalno se primenjuje Furijeova transformacija, iako postoje i drugi pristupi.

Ulaz sistemu

- Niz semplova. Postoje implementacije DFT koje kao ulaz prihvataju niz realnih brojeva, ali tipično DFT kao ulaz prihvata niz kompleksnih brojeva. Ako radimo sa realnim brojevima (kao što bi npr. bili semplovi glasovnog signala), dovoljno je popuniti imaginarni deo input-a nulama.

Izlaz sistema

- Niz brojeva iz istog domena kao ulaz. Ako su to realni brojevi, ne moramo dodatno da ih obrađujemo. Ako su kompleksni, ima više pristupa. Jedan tipičan je da se uzme apsolutna vrednost kompleksnog broja:

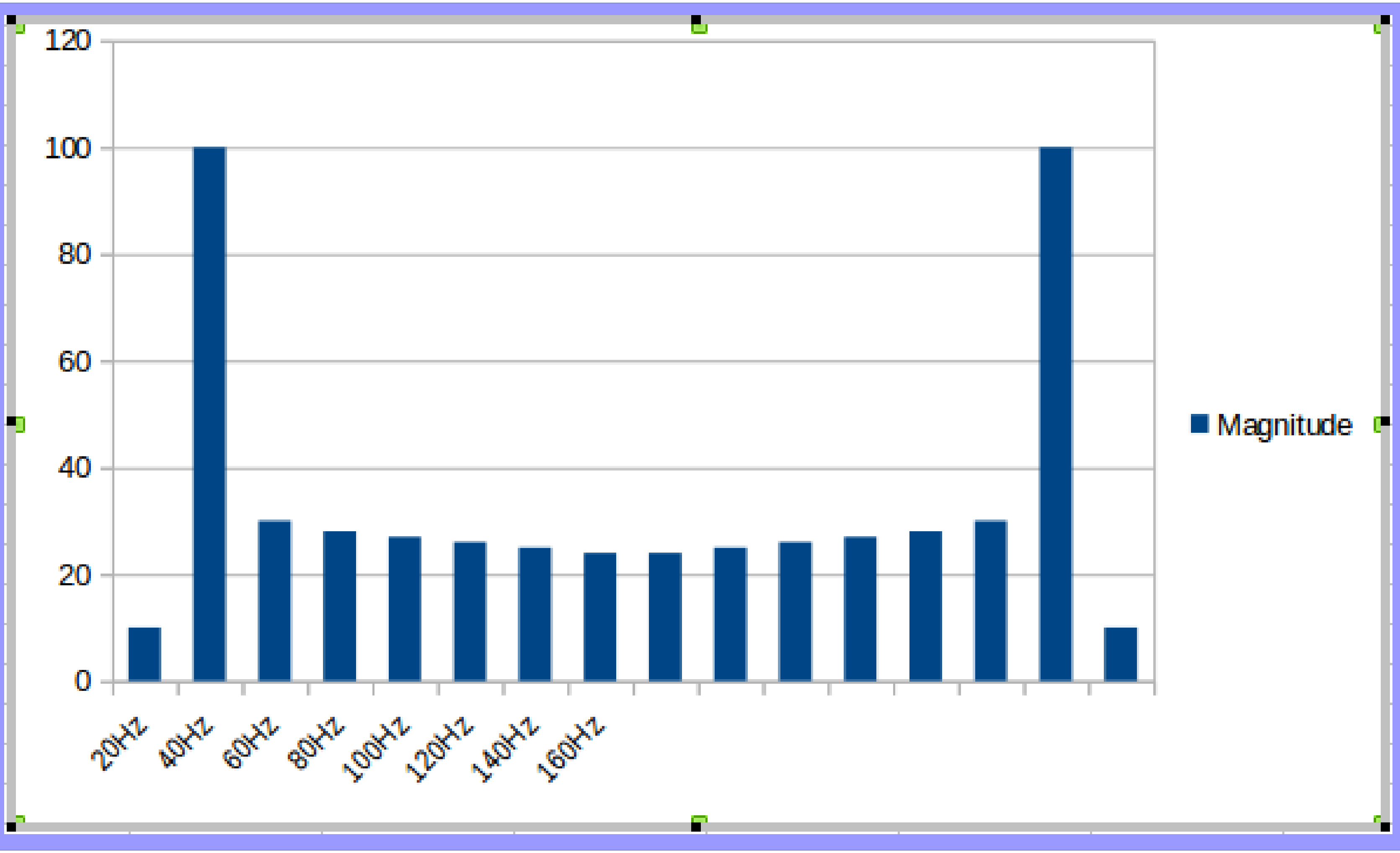
$$|z|=\sqrt{x^2+y^2}$$

- Ovu vrednost u analizi signala nazivamo magnituda, i ona predstavlja snagu signala za frekvenciju koja se određuje na osnovu indeksa. (vidi sledeći slajd)

Izlaz sistema

- Nulti element je izuzetak i predstavlja srednju vrednost ulaznog signala.
- Niz će biti iste dužine kao ulazni niz. Druga polovina izlaza može da se odbaci, pošto je to uvek identičan mirror prve polovine.
- Pretpostavimo da je dužina semplovanog signala T
- Prva vrednost u rezultujućem nizu predstavlja amplitudu za harmonik koji ima frekvenciju 1/T, druga vrednost za harmonik koji ima frekvenciju 2/T, potom 3/T, itd.

Izlaz sistema



Matematički zapis

- Direktni DFT:
- $X(n) = \frac{1}{N}\sum_{k=0}^{N-1} x(k)e^{-jk2\pi n/N}$
- Inverzni DFT:
- $X(n) = \frac{1}{N}\sum_{k=0}^{N-1} x(k)e^{jk2\pi n/N}$
- Pri implementaciji koristimo identitet:
- $e^{j\theta} = \cos(\theta) + jsin(\theta)$
- Gde je:
- $j = \sqrt{-1}; N - broj\ semplova;$

FFT

- Varijanta DFT koja pruža drastično poboljšanje brzine.
- DFT je $O(n^2)$, dok je FFT $O(n*\log_2 n)$
- Jedini zahtev je da ulazni niz mora da bude dužine koja je stepen dvojke. Pošto se popunjavanjem nulama ne gubi preciznost, ovaj problem generalno nije teško rešiti.

DFT primer

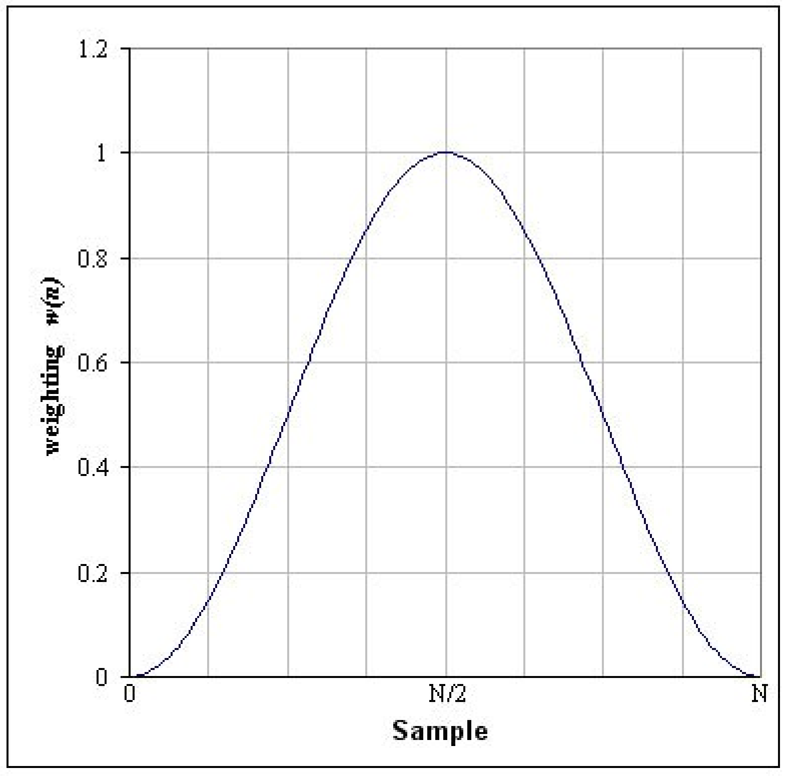
- Ulaz:
  - Sinusoida sa periodom T
  - Sinusoida sa periodom T/2
  - Zbir prethodna dva

Prozorske funkcije

- DFT pretpostavlja da je signal savršeno periodičan van okvira semplovanja.
- Pošto ne možemo biti sigurni koji deo signala smo uhvatili, korisno nam je da se „na silu“ pobrinemo da imamo glatki prelaz između različitih perioda. Na ovaj način može da se dobije vernija aproksimacija prave funkcije.
- Primenjuje se prozorska funkcija na input.

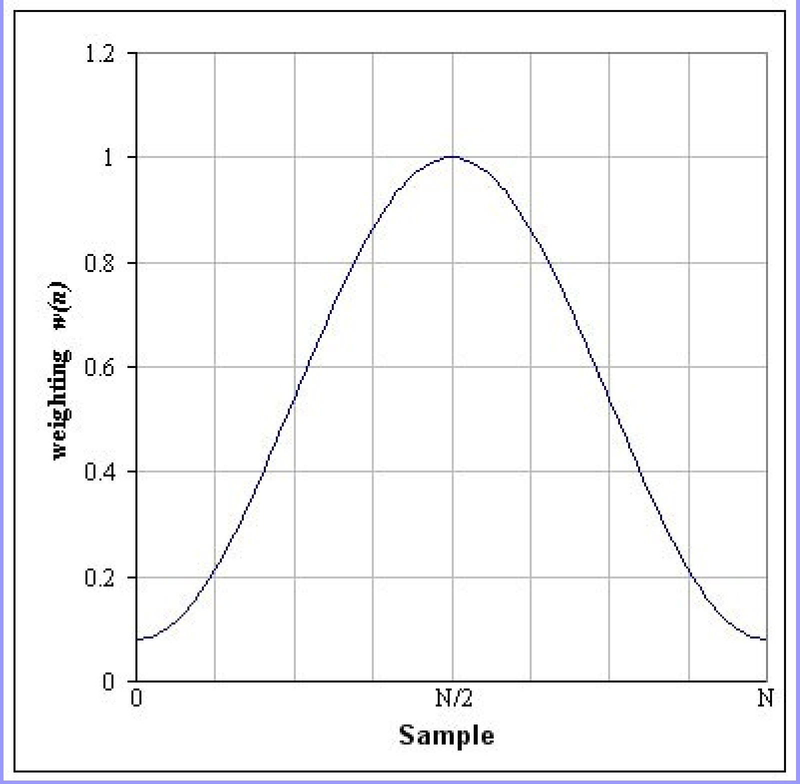
Hanning

$$w(n) = 0.5\left(1 - \cos\left(\frac{2\pi n}{N-1}\right)\right)$$



Hamming

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right)$$



Vežba

- Generisati faux semplovan signal dužine 1s koji se sastoji od bar 20 harmonika. Frekvencija semplovanja je 8KHz (tj. generisati 8K semplova).
- Uzimati prozore od po 15ms.
- Izvršiti DFT na input-u kod kog je primenjena prozorska funkcija, i na input-u kod kog nije primenjena prozorska funkcija i uporediti rezultate.