

## NUTQ SIGNALINI PARAMETRLASHNING MFCC ALGORITMI

**Xujayarov Ilyos Shiraliyevich**

Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti  
Samarqand filiali  
[i.khujayorov@samuit.uz](mailto:i.khujayorov@samuit.uz)

**Xafizova Shahnoza G'ulomovna**

Sharof Rashidov nomidagi Samarqand davlat universiteti  
[xafizovashahnoza@gmail.uz](mailto:xafizovashahnoza@gmail.uz)

### ANNOTATSIYA

*Nutq signalini parametrik ifodalashda MFCC algoritmining qo'llaniliishi*

**Kalit so'zlar:** MFCC, xemming oynasi, FFT, mel-kepstral koeffitsient, mel-chastota.

### KIRISH

Nutqni tanib olish tizimlarini ishlab chiqishda nutq signallarini parametrik ifodalashda optimal usullardan foydalanish maqsadga muvofiq sanaladi.

Nutq signalini parametrik ifodalashda signal fragmentarini spektral tahlil qilish hamda ularning kepstral xususiyatlarini hisoblashda eng ma'quli.

**Mel-chastota kepstral koeffitsientlar(mel-frequency cepstral coefficients, MFCC) algoritmi.** MFCC belgilari yozib olingan nutq signallaridan ajratib olinadi. MFCC algoritmi fonogramma va spektral almashtirish algoritmlari natijasidan foydalanadi.

Quyida algoritmnining bajarilish bosqichlari keltirib o'tilgan.

1. *Dastlabki filtrlash.* Nutq signali diskretlash jarayonidan o'tgandan so'ng, nutq signalining silliqlangan spektral shaklini olish uchun filtrnadi. Dastlabki filtr quyidagi tenglamada keltirilgan:

$$y(n) = x(n) - ax(n-1), \quad (1)$$

bunda  $a$  dastlabki filtrning konstanta qiymati  $0.9 < a < 1.0$ .

2. *Freynga ajratish bloki.* Dastlabki filtrlash amalga oshirilgandan so'ng nutq signali davomiyligi 16 msek bo'lgan freymlarga ajratiladi. Har bir freym(birinchi freymdan tashqari) o'zidan oldingi freymning oxirgi 10 msek qismini o'ziga qo'shib oladi. Ushbu jarayon signal oxirigacha amalga oshiriladi. Ushbu ishda nutq signalining diskretlash chastotasi 16 kGs bo'lganligi sababli bizda freym uzunligi  $N=256$  ta siljitish uzunligi  $M=160$  ta qiymatlardan iborat. Siljitish uzunligi(overlap) freym

uzunligining 62.5% ni tashkil etadi. Odatda freym uzunligining 50 - 75% gacha qamrab olishi tavsiya etiladi.

3. *Oynadan o'tkazish.* Ajratilgan freymlar bo'yicha buzilishlarni kamaytirish va ularni silliqlash uchun vazn oynasidan foydalaniladi:

$$w_n = 0.54 - 0.46 * \cos\left(2\pi \frac{n}{N-1}\right), n = 0, \dots, N - 1, \quad (2)$$

bu yerda  $N$  - oyna uzunligi.

4. *Spektral almashtirish.* Algoritmning uchinchi bosqichida vazn oynasidan o'tkazilgan freymlarda FFT spektral almashtirish prosedurasini qo'llaniladi.

$$X_k = \sum_{n=0}^{N-1} x_n w_n e^{-\frac{2\pi i}{N}kn}, k = 0, \dots, N - 1. \quad (3)$$

$k$  indeks qiymatlari chastotalarga mos keladi.

$$f_k = \frac{F_s}{N}k, k = 0, \dots, N/2, \quad (4)$$

bu yerda  $F_s$  — signalning diskretlash chastotasi.

5. *Mel almashtirish.* To'rtinchi bosqichda esa chastota sohasiga o'tkazilgan signal uchburchakli filtrlar yordamida diapazonlarga ajratiladi. Filtr chegaralari mel shkalasi bo'yicha hisoblanadi.

Mel-chastota sohasiga o'tish quyidagi formula asosida amalga oshiriladi.

$$B(f) = 1127 * \ln\left(1 + \frac{f}{700}\right) \quad (5)$$

Teskari almashtirish esa quyidagicha amalga oshiriladi.

$$B^{-1}(b) = 700 \left(e^{b/1127} - 1\right). \quad (6)$$

Agarda biz  $N$ -ni filtrlar soni (odatda 26 ta filtrdan foydalaniladi),  $(f_{low}, f_{high})$ -ni tadqiq qilinayotgan chastota diapazoni deb qabul qilsak, unda bu diapazon Mel-shkalasiga o'tkaziladi hamda  $N$  ta teng taqsimlangan bir-biri bilan kesishadigan diapazonlarga bo'linadi va chiziqli chastota sohasidagi tegishli chegaralari topiladi. Filtrlash asosida olingan vazn koeffitsiyentlarini  $H_{m,k}$  -deb belgilash kiritamiz. Filtrlar Furre almashtirishdan olingan koeffitsiyentlarining kvadrat moduliga qo'llaniladi. Olingan qiymatlar quyidagi ifoda orqali logarifmlanadi.

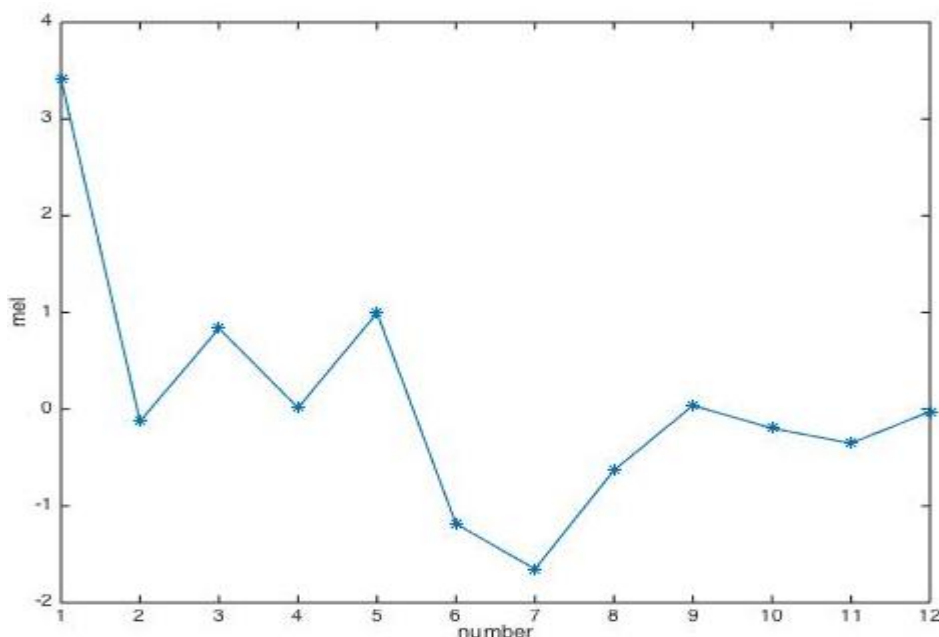
$$e_m = \ln\left(\sum_{k=0}^N |X_k|^2 H_{m,k}\right), m = 0, \dots, N - 1. \quad (7)$$

MFCC koeffitsiyentlarini hisoblashning oxirgi bosqichida esa DCT algoritmi amalga oshiriladi.

$$c_i = \sum_{m=0}^{N-1} c_m \cos\left(\frac{\pi i(m + 0.5)}{N}\right), i = 1, \dots, N_{MFCC}. \quad (8)$$

$c_0$  koeffitsiyenti ishlatilmaydi, chunki  $c_0$  signal energiyasini anglatadi. Odatda amaliyotda MFCC koeffitsiyentlar soni  $N_{MFCC}$  13 ta qiymat olinadi.

Mel-kepstral koeffisientlarning 13 filtr uchun olingan natijalari 1-rasmda keltirilgan.



1-rasm. Mel-kepstral koeffisientlarning olinishi

Barcha filtrlarni hisoblash uchun mel diapazondagi boshqa nuqtalarni belgilash quyidagi tartibda amalga oshiriladi. Misol uchun 10 ta filtr uchun 12 nuqta belgilash kerak. Bu shuni bildiradiki, 401,25 dan 2839,99 gacha diapazonda 10 ta nuqtani bir xil oraliqlarga joylashtirib chiqish kerak. Natijada quyidagicha ketma-ketlikdagi nuqtalar paydo bo'ladi:

$m(i) = [401,25; 622,5; 843,75; 1065; 1286,25; 1507,5; 1728,74; 1949,99; 2171,24; 2392,49; 2613,74; 2834,99]$ .

Nuqtalarni mel shkaladan asl chastotaga o'tkazish:

$h(i) = [300; 517,33; 781,9; 1103,97; 1496,04; 1973,32; 2554,33; 3261,62; 4122,63; 5170,76; 6446,7; 8000]$ .

Nutq signalidagi spektrdagi har bir nuqta ma'lum bir chastotaga to'g'ri keladi va u quyidagi formula yordamida hisoblanadi.

$$f(i) = \frac{2 \cdot (i+1)}{\text{sampleRate} \cdot n} \quad (9)$$

bu yerda sampleRate – signalning diskretlash chastotasi, n – signal uzunligi.

MFCC algoritmi yordamida mel-shkalaning o'lchamiga asosan 12,13 ta yoki 24 ta qiymatlardan iborat parametrlarni olish mumkin. Bu algoritm spektral chastotalarning eng muhim bo'lgan parametrlarini o'zida o'zlashtiradi. boshlang'ich va parametrik ishlov berish algoritmlar ishlov berish natijasida inson eshitish tizimiga moslashtirildi. Algoritmik nuqtai nazardan bunday uslub asosida yondashuv nutq tovushlarini, bo'g'inlarini va so'zlarini tanib olishda foydalaniladi.

## ADABIYOTLAR RO‘YXATI

1. Тампель И.Б. Автоматическое распознавание речи – основные этапы за 50 лет // Научно-технический вестник информационных технологий, механики и оптики. 2015. Том 15. № 6. С. 957–968.
2. Gulmezoglu M.B. A novel approach to isolated word recognition. IEEE Transactions on Speech and Audio Processing, 1999. 7(6): p. 620-628.
3. Xujayarov I.Sh, Kamolov R.K, Xafizova Sh.G‘ Call-markazi masalalari uchun mashinali o‘qitish usullarini qo‘llanilishi tahlili. “Zamonaviy axborot, kommunikatsiya texnologiyalari va at-ta’lim tadbiri muammolari” mavzusidagi respublika ilmiy-amaliy anjumani. Samarqand 7-8-aprel 2023-yil B 31-33.
4. Altınok, H., Kurdal, S., & Yucal, M.M. (2021). Turkish Speech Recognition in Call Centers. 2021 6th International Conference on Computer Science and Engineering (UBMK), 791-793.