NUTQ SINTEZATORLARINI YARATISH ALGORITMLARI VA ULARNING FARQLARI

Dauirxanova Tursinay Azirbay qizi

tursinaydauirxanova@gmail.com

Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti, magistratura talabasi

Annotatsiya: Ushbu maqola nutq sintezatorlari deb nomlanuvchi nutq sintezi tizimlarini yaratish algoritmlari va farqlarini o'rganadi. Nutq sintezatorlari - bu insonga o'xshash nutqni sintetik tarzda yaratadigan elektron texnologiyalar bo'lib, matnni turli til va uslublarda og'zaki so'zlarga aylantirish imkonini beradi. Maqolada nutq sintezatorlarini ishlab chiqishda ishlatiladigan asosiy algoritmlar muhokama qilinadi va ularning farqlari yoritiladi. Maqola sintetik nutqni yaratish uchun lingvistik qoidalar va fonetik ko'rsatmalarga tayanadigan qoidaga asoslangan algoritmlarni kiritish bilan boshlanadi. Keyin u sintezlangan chiqishni yaratish uchun oldindan yozib olingan nutq birliklarini yig'adigan konkatenativ algoritmlarni o'rganadi. Keyinchalik statistik modellashtirish usullaridan foydalanadigan statistik parametrik algoritmlar muhokama qilinadi, so'ngra nutq sintezi uchun neyron tarmoqlardan foydalanadigan chuqur o'rganish algoritmlari ko'rib chiqiladi.

Kalit so’zlar: Nutq sintezi, nutq sintezatorlari, algoritmlar, rull-based (qoidaga asoslangan) algoritmlar, concatenative (birlashtiruvchi) algoritmlar, statistical parametric (statistik parametrik) algoritmlar, deep learning (chuqur o'rganish) algoritmlar, tabiiylik, sifat, cheklovlar, ilovalar.

Kirish. Nutq sintezatorlari, inson ovozini elektronik ravishda generatsiya qiladigan texnologiyalardir. Ular, ovozni sintetik ravishda ishlab chiqarish, matnni ovozga aylantirish va istalgan til va stildagi nutqni generatsiya qilish imkonini beradi. Nutq sintezatorlarini yaratishda muhim rol o'ynagan asosiy komponentlar esa yaratish algoritmlari va ularning farqlaridir. Bu maqolada nutq sintezatorlarini yaratishda qo'llanadigan ba'zi asosiy algoritmlar va ularning farqlarini ko'rib chiqamiz.

Rule-Based (Qoidalarga asoslangan) Algoritmlar:

Bu algoritmlar nutq sintezatorlarini yaratishning klassik usullaridan biridir. Ular, qoidalarga asoslangan qo'llanmalardan foydalanarak nutqni generatsiya qilishni amalga oshiradilar. Misol uchun, tilning grammatikasi, ovozning fonetik qoidalari va intonatsiyasi kabi qoidalarga muvofiq ravishda ovozni yaratish uchun qo'llaniladi. Bu algoritmlar bir nechta til va til ovozlari uchun keng qo'llaniladi, ammo ularda nutqning tabiiylik darajasidan kelib chiqishi mumkin.

Concatenative (Birlashtiruvchi) Algoritmlar:

Birlashtiruvchi algoritmlar nutq sintezatorlari yaratishning boshqa bir usuli hisoblanadi. Ular, ovozning to'plami (korpus) asosida matndan kerakli ovozlarni bir-biriga birlashtirib, matnning sintezini amalga oshiradilar. Bu algoritm sintezi, ovozni to'g'ri ta'riflash va intonatsiya bilan birga jamlashni o'z ichiga oladi. Birlashtiruvchi algoritmlar ovozlarni tabiiy tarzda tinglashda

yaxshi natija beradi, ammo sintezasi jarayonida ovozlarning to'plami katta bo'lishi kerakligi muhim nuqta hisoblanadi.

Statistical Parametric (Statistik Parametrik) Algoritmlar:

Statistik parametrik algoritmlar nutq sintezatorlarini yaratishda iste'mol qilingan boshqa bir usuldur. Ular, statistik modellar yaratish va ovozni sintezi qilish uchun ma'lumotlarni o'rganishga asoslanganlar. Bu algoritmlar ovozlarning matematik modellari asosida ishlaydilar. Biroq, ularda yorug'lik va intonatsiya xossalari keng qo'llanishga mo'ljallanganligi uchun, matnning sintezida to'g'ri ovozning tasvirlashida qiyinchiliklar bo'lishi mumkin.

Deep Learning (Dolzarb o'rganish) Algoritmlar:

Dolzarb o'rganish algoritmlarida nutq sintezatorlari yaratish uchun keng qo'llanadigan tizimlar, masalan, rekurrent neyral tarmoqlar (RNN), o'z-ozaro qo'llanmalar tarmoqlari (CNN), generativ o'rganish modeli (GAN) va boshqa tizimlar ishlatiladi. Bu algoritmlar ovozning o'ziga xos xossalari va ta'lim o'rganish yordamida ovozni sintetik ravishda generatsiya qilishda yuqori sifatli natijalar beradi. Dolzarb o'rganish algoritmlari, nutqning tabiiylik darajasiga yaqin sintezini ta'minlayadi.

Xulosa sifatida shuni aytishim mumkinki, Nutq sintezatorlari yaratish algoritmlari nutqning sintezi va generatsiyasi uchun muhim qadamlardir. Rule-Based algoritmlar qoidalarga asoslangan til grammatikasi va ovozning fonetik qoidalari kabi qo'llanmalardan foydalanadi. Birlashtiruvchi algoritmlar esa ovozlarni ovoz to'plamiga birlashtirib sintezini amalga oshirishda ishlatiladi. Statistik parametrik algoritmlar statistik modellar asosida ovozni sintezi qilishni o'rganishga bog'liq, va Deep Learning algoritmlari esa dolzarb o'rganish tizimlarini ishlatib ovozni sintetik ravishda generatsiya qilishga asoslanadi.

Har bir algoritmning o'ziga xos farqlari mavjud. Rule-Based algoritmlar qoidalarga asoslanganligi uchun nutqni sintezlashda ovozning tabiiylik darajasidan uzoq bo'lishi mumkin. Birlashtiruvchi algoritmlar tabiiylikni oshirishda yaxshi natijalar beradi, ammo ovozlarning to'plami katta bo'lishi kerak. Statistik parametrik algoritmlar statistik modellashtirish asosida ishlaydi, lekin sintezda yorug'lik va intonatsiya xossalari uchun qiyinchiliklar tug'dirishi mumkin. Dolzarb o'rganish algoritmlari esa nutqning tabiiylik darajasiga yaqin natijalar berishda ilgari suradi.

Nutq sintezatorlari yaratish algoritmlari nutqning sifatini, ovozning to'g'ri tasvirlanishini va sintezini ta'minlashda muhim rol o'ynaydi. Algoritmlarning farqlari va ulardan foydalanish, nutq sintezatorlarining iste'mol sohalarida katta ahamiyatga ega bo'lgan mo'ljallarni beradi. Keyingi yillarda yaratish algoritmlari va dolzarb o'rganish tizimlari yordamida nutq sintezatorlarining sifati va natijalari yanada yanada oshiriladi. Bu esa nutq sintezatorlarining ovozli kommunikatsiyada, ta'limotda, avtomobillar sohasida va boshqa sohalarda keng qo'llanishining rivojlanishiga olib keladi.

Foydalanilgan adabiyotlar:

1. Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. The Journal of the Acoustical Society of America, 67(3), 971-995.

2. Black, A. W., & Taylor, P. (1997). CMU Flite: A free, open-source toolkit for speech synthesis. In Proc. of the 5th European Conference on Speech Communication and Technology (pp. 1827-1830).
3. Hunt, A. J., & Black, A. W. (1996). Unit selection in a concatenative speech synthesis system using a large speech database. In Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (pp. 373-376).
4. Stylianou, Y., Cappe, O., & Moulines, E. (1998). Continuous probabilistic transform for voice conversion. IEEE Transactions on Speech and Audio Processing, 6(2), 131-142.
5. Taylor, P., & Black, A. W. (1998). The Festival speech synthesis system. In Proc. of the International Conference on Spoken Language Processing (pp. 111-114).
6. Zen, H., & Sak, H. (2009). Statistical parametric speech synthesis. Speech Communication, 51(11), 1039-1064.
7. Yamagishi, J., & King, S. (2009). Hidden Markov model speech synthesis: An overview. IEEE Signal Processing Magazine, 26(3), 18-31.
8. Schroeter, J., & Atal, B. S. (1985). Code-excited linear prediction (CELP): High-quality speech at very low bit rates. In Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (pp. 937-940).
9. Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. The Journal of the Acoustical Society of America, 87(2), 820-857.
10. Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. Speech Communication, 9(5-6), 453-467.