

УДК 519.24

Савіцький О. – ст.гр. ПМ_п-62, Дедів Л. – ст. гр. ПМ-61

Тернопільський державний технічний університет імені Івана Пулюя

ДОСЛІДЖЕННЯ ТА ПОБУДОВА СПЕКТРОГРАМ ГОЛОСНИХ ЗВУКІВ МОВНОГО СИГНАЛУ

Науковий керівник: к.т.н., доц. Шадріна Г.М.

На сьогоднішній час існує гостра необхідність створення нових спеціальних програмно-апаратних технічних засобів і комплексів обробки мовної інформації на основі стандартних обчислювальних пристроїв, в яких може бути досягнена значна економія часових та матеріальних ресурсів. Створення ефективного алгоритму розпізнавання голосних звуків мовного сигналу, в першу чергу, вимагає дослідження та побудову спектрів цих сигналів.

Як і для більшості інших досліджень з даної тематики для полегшення розуміння подальших викладень можна ввести визначення фонооб'єкта, під яким розумітиметься реальний об'єкт, що генерує в звуковому діапазоні частот сигнали, які, будучи перетворені в цифрову форму, можуть записуватися і зберігаються в пам'яті комп'ютера у вигляді окремих файлів з метою подальшої обробки або передачі.

Сліди фонооб'єктів у вигляді параметрів амплітуд і фаз вузькосмугових сигналів виявляються на зображеннях динамічних спектрограм у вигляді сукупності контурів перепаду яскравості або треків локальних і глобальних екстремумів колірної насиченості в рівнях одного кольору. В даний час існує велика кількість програмних цифрових аналізаторів і редакторів аудіо сигналів, призначених для візуального аналізу звукових сигналів в тимчасовій (осцилограми, графіки рівня потужності сигналу і ін.) і, звичайно, частотній (сонограми, кепстри та ін.) областях. Серед програмних продуктів такого роду слід зазначити Cool Edit Pro 1.2, Dart Pro, Sound Forge, Wave Lab, Wave Studio та ін. У ряді звукових редакторів є можливість проводити деякі види обробки мовного сигналу.

Всю різноманітність спектрограм одного і того ж звуку ми намагаємося описати за допомогою статистичної моделі.

Практика досліджень показує, що розподіли спектрограм для одних і тих самих звуків у вимові різних дикторів різні. Отже, тут не може йтися про одномодальність розподілів. Крім того, спектральна картина звуку міняється залежно від оточуючих його звуків, які тим або іншим чином роблять вплив на його модифікацію. Враховуючи все це, ми всю різноманітність спектрограм кожного звуку розбиваємо по дикторах на групи, які, у свою чергу, розбиваються далі, скажімо, на звукотипи, що можуть виступати як еталони.

Відомо, що розподіл спектрограм одного звукотипу для одного диктора має один мод. Це характеризує чітку наявність деякої типової спектрограми, а розподіл спектрограм звукотипу для одного диктора може бути описаний багатовимірним нормальним розподілом (у загальному випадку із залежними компонентами). Отже, ми характеризуємо розподіл спектрограм звуку сумою нормальних розподілів, і таким чином апроксимуємо дійсний розподіл. Число доданків визначається кількістю дикторів і звукотипів. Правильність такої моделі, яка представляється достатньо правдоподібною, може бути перевірена за наслідками розпізнавання за допомогою алгоритму, витікаючого з цієї моделі.

Запропонована обробка спектрограм дає можливість виявити відмінність у голосних звуках мовного сигналу та подальшого їх розпізнавання.