УДК 621.39

СЕГМЕНТАЦІЯ МОВЛЕННЄВИХ СИГНАЛІВ ПРИ СТВОРЕННІ АКУСТИЧНИХ БАЗ ДАНИХ

Дзісь О. В.
Науковий керівник – доц., к.т.н. Ткаченко О. М.

Дослідження у таких галузях, як розпізнавання, ущільнення та синтез мовлення потребують накопичування великої кількості мовленнєвих фрагментів, що супроводжується описом відповідних деталей цих фрагментів (розміткою). Створення та розмітка достатньо повних акустичних баз даних є однією з головних передумов успішного розвитку сучасних мовленнєвих технологій. Відсутність розміченої акустичної бази даних великого обсягу для українського мовлення зумовлює труднощі у процесі навчання та оцінювання якості систем автоматичного розпізнавання та синтезу мовлення, фонетичних вокодерів тощо. Більшість експериментів у цих дослідженнях виконується на іншомовному матеріалі. Тому існує необхідність створення аналогічної бази даних для української мови. Передумовою її створення є розробка ефективних систем сегментації, орієнтованих на українське мовлення.

У роботі розглянуто шляхи підвищення точності встановлення границь між фонемами при сегментації мовлення.

Для сегментації при відомій фразі запропоновано застосовувати алгоритм Вітербі, що дозволяє значно скоротити кількість варіантів розміщення границь. Пропонується також застосовувати методи розпізнавання мовлення при визначенні оціночних функцій для алгоритму Вітербі.

Найточнішою вражається ручна розмітка мовлення, тому з нею доцільно порівнювати результати автоматичної розмітки. Для характеристики результатів введено два показники – показник точності сегментації (CS), що характеризує відношення тривалості вірно розмічених сегментів, до загальної тривалості фрази, та показник що характеризує кількість правильно встановлених границь (BC).

При застосуванні алгоритму Вітербі та методів розпізнавання отримано такі результати: за показником CS точність склала 87,5%, а за BC - 77,3%.

Для покращення результатів запропоновано використовувати статистичну інформацію про фонеми, а саме середню максимальну та середню мінімальну тривалість фонеми. За рахунок внесення таких уточнень в алгоритм Вітербі, показник CS зріс до 90%, а показник BC покращився на 6,5%. Таке покращення отримано за рахунок підвищення обчислювальної складності.

Отримані результати можуть застосовуватися для попередньої сегментації мовленнєвих даних в автоматизованій системі сегментації з метою створення україномовної бази мовленнєвих даних.

На основі запропонованих підходів розроблено програмне забезпечення для автоматизованої сегментації мовлення.