

МЕТОД СЕГМЕНТАЦІЇ СИГНАЛІВ У МОВЛЕННЄВОМУ КОРПУСІ НА ОСНОВІ АЛГОРИТМУ ВІТЕРБІ

Дзісь О. В.

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

Дослідження у галузях пов'язаних з обробкою та аналізом мовленнєвих сигналів потребують накопичування великої кількості мовленнєвих фрагментів, з додатковою інформацією (анотацією). Створення та розмітка достатньо повних мовленнєвих корпусів (МК) є однією з головних передумов успішного розвитку сучасних мовленнєвих технологій. Відсутність розміченого МК 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. Передумовою створення аналогічного МК для української мови 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.

Оскільки найточнішою вважається ручна розмітка мовлення, то результати автоматичної розмітки порівнюються саме з нею за показником точності сегментації (CS), що характеризує відношення тривалості вірно розмічених сегментів, до загальної тривалості фрази, та показником 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 тривалість фонем. 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.

Якщо розглядати границю як миттєву зміну однієї фонем на іншу, то частина розмітки у МК буде неадекватно описувати сегменти, оскільки насправді між фонемами є досить тривалі ділянки, які не можна однозначно класифікувати. Запропоновано шляхом додаткового аналізу оцінок у дереві Вітербі визначати ділянки переходів між фонемами, які можна використовувати у якості границь. За рахунок такого уточнення 91% відсоток інформації про границі фонем у МК буде достовірним.

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