

Пошук ключових слів в потоці мовлення із записів судових засідань

¹В.В. Пилипенко, ²О.М.Радущький

¹Міжнародний науково-навчальний центр інформаційних технологій та систем
40 просп. Академіка Глушкова, Київ 03680

²ТОВ Спеціальні реєструючі системи
valeriy.pylypenko@gmail.com, alex@srs.kiev.ua

Анотація

Розглядається використання підходу фонетичного стенографа для пошуку ключових слів у потоці зв'язного мовлення на прикладі записів судових засідань. Фонем моделюються Прихованими Марківськими моделями. Ключове слово задається за допомогою фонетичної транскрипції. Розроблено демонстраційну програму пошуку ключових слів для української та російської мов. Наводяться результати експериментів з пошуку ключових слів у потоці мовлення.

1. Вступ

В даний час записується все більше мовленнєвої інформації в різних галузях господарства – телекомунікаційних мережах, телебаченні, радіо, телефонних центрах обслуговування, судах тощо. Вся ця інформація зберігається, але доступ до неї потребує значного часу. Наприклад, для пошуку того, що було сказано, потрібно прослухати практично всі записи. Автоматичне розпізнавання мовлення, зокрема пошук ключових слів, може значно полегшити цю задачу, надаючи текстове представлення мовленнєвої інформації.

Пошук ключових слів в потоці мовлення – це знаходження фрагментів звукового файлу, які містять ключове слово або фразу. Існує декілька підходів до пошуку ключових слів, зокрема, часто використовується автоматичне розпізнавання мовлення з великим словником. Якщо слово, яке шукається, належить до словника системи розпізнавання, то пошук даного слова зведеться до звичайного текстового пошуку з урахуванням часових міток, які надає автоматичне розпізнавання мовлення. Це потребує значного комп'ютерного ресурсу, оскільки словники сягають декількох сотень тисяч слів і більше, але це разова процедура, яку можна зробити заздалегідь.

У випадку, коли слово не належить до заздалегідь сформованого словника та заздалегідь розпізнаного мовленнєвого файлу, використовуються інші підходи, засновані на моделюванні оточуючого потоку слів. В даній роботі використовується підхід фонетичного стенографа [1], який висвітлюється у другому розділі. На базі цього підходу розроблено демонстраційну програму для української та російської мов, яка описана у третьому розділі. Наприкінці наведені результати

експериментальних тестових досліджень на матеріалі реальних звукових записів судових засідань.

2. Математична модель потоку слів

Здійснення пошуку одного або невеликої кількості ключових слів потребує моделювання решти слів мовлення, які не потрапили до списку ключових слів. Модель оточуючих слів, як довільної послідовності фонем, можна представити у вигляді графа, відображеного на рис. 1.

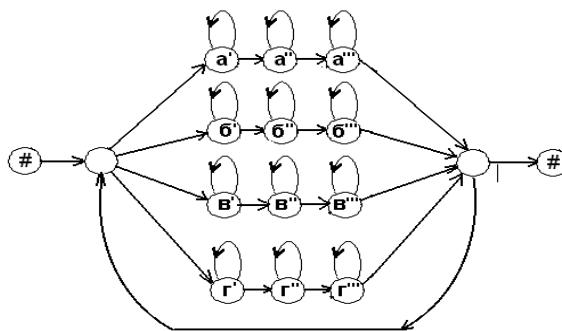


Рисунок 1. Граф довільної послідовності фонем з трьома станами

Кожна фонема моделюється трьома станами прихованої Марковської моделі зі 128 сумішами. Акустичні моделі фонем будувалися з використанням розміченого корпусу мовлення АКУЕМ [2] за допомогою інструментарію НТК [3].

Ключові слова для пошуку задаються послідовністю фонем, яка формується автоматично або експертом (у випадку, коли необхідно знайти слова з незвичайним вимовлянням).

3. Програма пошуку ключових слів для української та російської мов

Розроблено демонстраційну програму пошуку ключових слів. Інтерфейс користувача відображений на рис. 2. Верхнє вікно показує звуковий сигнал у вигляді осцилограми мовленнєвого сигналу, де вказано знайдені ключові слова. Усі знайдені слова також відображаються у списку з зазначенням релевантності пошуку.

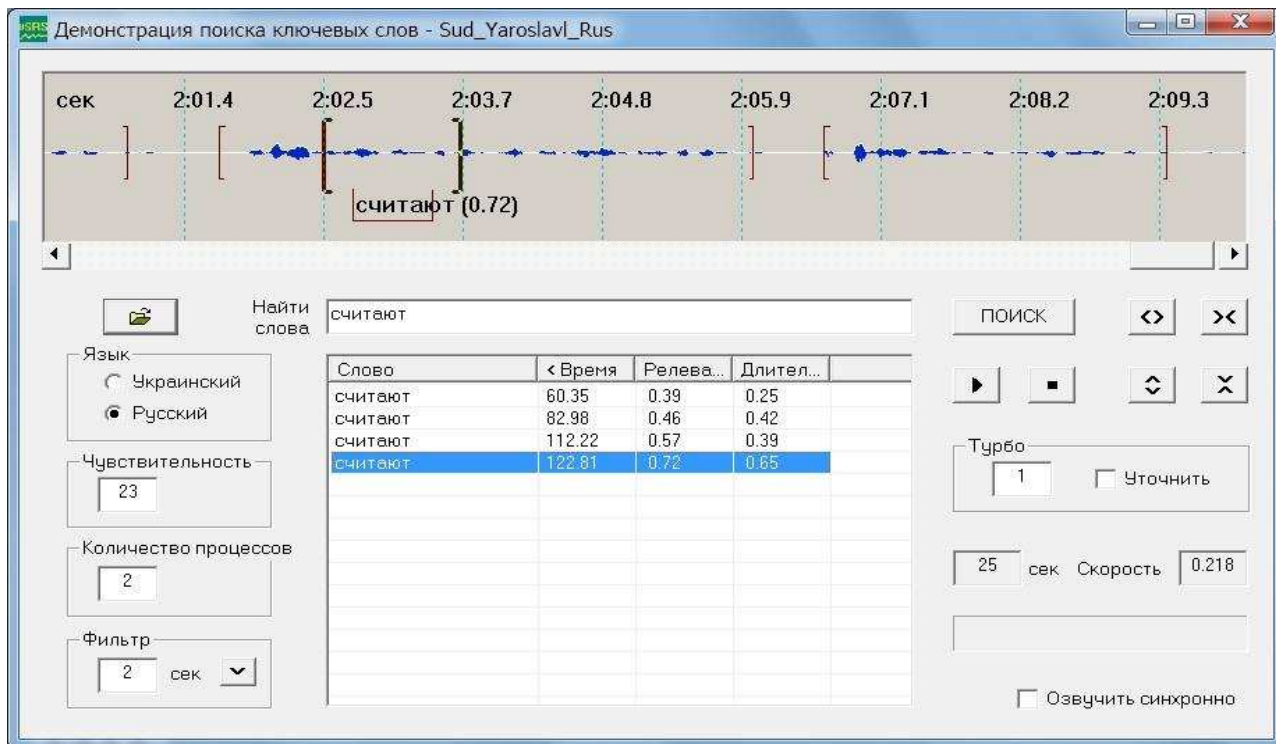


Рисунок 2. Демонстраційна програма пошуку ключових слів

Користувач задає чутливість пошуку, яка керує співвідношенням між точністю пошуку та кількістю помилкових тривог.

Сучасні комп'ютери часто мають декілька процесорів, і для прискорення пошуку користувач може використовувати паралельні обчислення, збільшуючи кількість задіяних процесорів. На сучасних процесорах типу Pentium i7 швидкість оброблення сягає більше ніж 5 секунд мовлення за 1 секунду обчислень для одного процесора. У випадку, коли необхідно ще швидше отримати відповідь, можна вибрати режим ТУРБО, який прискорює пошук приблизно в 10 разів, але при цьому знижується точність пошуку.

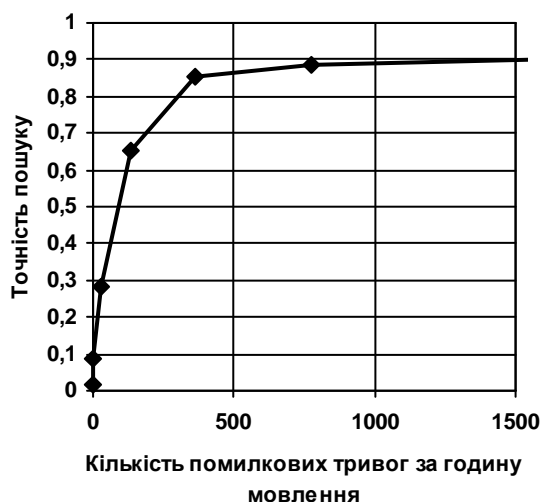


Рисунок 3. Характеристики пошуку ключових слів

4. Експериментальні дослідження

Для тестування програми використовувався запис виступів декількох учасників судового засідання колегії

Вищого арбітражного суду Російської федерації. Був проведений експеримент з пошуку всіх слів цього запису. Результати автоматичного пошуку порівнювалися з еталонним описом запису, зробленим експертом. Показники пошуку – точності пошуку та кількості помилкових тривог – відображені на рис. 3. Точність пошуку визначена як співвідношення правильно знайдених слів до кількості усіх реалізацій ключових слів, визначених експертом.

5. Висновки

Використання моделі оточуючих слів як довільної послідовності фонем дозволило розробити програму для пошуку ключових слів.

Експерименти з оптимізації набору фонем довели, що об'єднання наголошених та ненаголошених варіантів голосних фонем призводить до незначного зниження показників пошуку.

Експериментальні дослідження на реальних записах показали, що точність пошуку ключових слів на рівні 88% дає можливість використовувати розроблену програму для пошуку необхідної звукової інформації в архівах судових засідань.

7. Література

1. Пилипенко В.В. Распознавание ключевых слов в потоке речи при помощи фонетического стенографа. Искусственный интеллект. – Донецк, 2009. – № 4с. 220-224.
2. Н.Б.Васильева, В.В.Пилипенко, О.М.Радучький, В.В. Робейко, М.М.Сажок. Створення акустичного корпусу українського ефірного мовлення. – Оброблення сигналів і зображень та розпізнавання образів: Десята Всеукраїнська міжнародна конференція. – Київ, 2010, С.55-58.
3. Young S. et al. The NTK Book (for NTK Version 3.4) // Cambridge University Engineering Department: Cambridge, UK, 2009.