



Опыт автоматического стенографирования украинской парламентской речи

Пилипенко В.В.

Робейко В.В.

В статье рассматривается программа-стенограф для получения текста стенограммы из звукового файла на основе системы распознавания речи. Записанная фонограмма обрабатывается системой распознавания слитной речи многих дикторов с большими словарями (больше 10 тыс. слов). Оператор исправляет допущенные ошибки, а также вводит новые слова, незнакомые системе распознавания. На основе анализа ошибок и новых слов производится дообучение системы распознавания, что позволяет улучшать показатели надёжности распознавания речи в процессе эксплуатации системы. Применение индивидуализированных транскрипций для некоторых дикторов позволило повысить надёжность распознавания. Приведены результаты распознавания украинской парламентской речи.

In this work we present a research system of computerized stenographer. It produces texts out of sound records by means of human-aided speech recognition system. Continuous speech recognition system for a number of speakers with a large vocabulary (more than 10.000 words) is used to process recorded files. Human introduces out-of-vocabulary words and repairs errors to produce the final text. Retraining process is running to take into account repairs, thus improving system performance. Personal phonetic rules are listed and used to individualize transcriptions for different speakers. Experimental results for Ukrainian parliament speech recognition are presented.

Список ключевых слов

Распознавание слитной речи, автоматизированное стенографирование, многодикторное распознавание речи, большие словари, индивидуализированные транскрипции.

Continuous speech recognition, computer-assisted stenographer, multispeaker speech recognition, large vocabulary, transcription individualization.

Введение

Стенографирование широко используется для обработки и документирования материалов заседаний и совещаний различного уровня, для работы секретарей, журналистов и так далее. Во многих странах необходимо стенографировать заседания в парламентах и судах. Компьютеры значительно расширили возможности и увеличили гибкость применения систем стенографирования. На данный момент становится актуальной задача уменьшения доли ручного труда в таких системах. Для этого предлагается использовать автоматическое распознавание речи при преобразовании звука в текст.

Речь каждого человека сугубо индивидуальна. Поэтому перевести звук в текст по нажатию одной кнопки — задача довольно сложная для системы стенографирования. Такая система должна максимально упростить работу оператора и ускорить перевод звукового файла в текстовый, а также учесть все особенности речи диктора. Существует много программно-аппаратных комплексов автоматизированного стенографирования с различными возможностями, но даже самый простой позволяет увеличить скорость перевода звука в текст в несколько раз.

Автоматическое распознавание слитной речи многих дикторов с большими словарями значительно упрощает работу оператора, сводя её к исправлению ошибок, допущенных системой стенографирования. Дообучение системы позволяет сокращать количество ошибок в процессе эксплуатации.

1. Автоматизированная vs автоматическая

Системы стенографирования можно условно разделить на три категории в зависимости от соотношения участия человека и компьютера в процессе создания стенограмм:

- **автоматические** (без участия человека в процессе распознавания речи);
- **автоматизированные** (компьютер распознаёт поток речи, а человек участвует в этом процессе в той или иной степени);
- **стенографирование** при помощи компьютера (человек набирает текст, а компьютер используется как магнитофон и печатная машинка).

Разница между автоматической и автоматизированной системами заключается в надёжности автоматического распознавания речи.

Опыт эксплуатации показывает, что первичная стенограмма, созданная человеком, содержит ошибки, которые исправляются в процессе редактирования набранного текста. В среднем количество ошибок достигает 5 на одну страницу текста, что составляет надёжность 98%, поскольку одна страница содержит приблизительно 2000 знаков, или 250 слов. Таким образом, система стенографирования должна стать автоматической при надёжности распознавания речи выше 98%.

Такая надёжность уже сегодня достижима для автоматического распознавания речи при некоторых ограничениях [1]. При этом распознаётся речь только одного диктора. Для

изолированно произносимых слов словарь достигает 15 тыс. слов, а для слитной речи такая надёжность достигается при словаре в 1 тыс. слов.

Поэтому на настоящий момент актуальным является создание программ распознавания речи, свободных от таких ограничений. Для стенографирования необходимо достигнуть объёмов словаря от 10 тыс. слов до нескольких миллионов. Количество задействованных дикторов составляет от 100 до нескольких тысяч. При этом должна распознаваться слитная речь в реальном времени.

Автоматизированную систему стенографирования имеет смысл применять при надёжности 80% и выше. При этом оператору необходимо будет исправлять не более чем каждое пятое слово в тексте, что можно делать при прослушивании звуковой дорожки в процессе её воспроизведения.

2. Система распознавания слитной речи

В данной работе в качестве базовой системы используется инструментарий НТК [2] на основе скрытых Марковских моделей (СММ). Инструментарий НТК использовался для построения акустических и лингвистических моделей. Для распознавания речи был разработан программный комплекс, совместимый с акустическими и лингвистическими моделями НТК.

2.1. Пользовательский вид программы

Пользовательский вид программы стенографирования приведён на рисунке 1. В верхнем окне схематически изображена осциллограмма звуковой

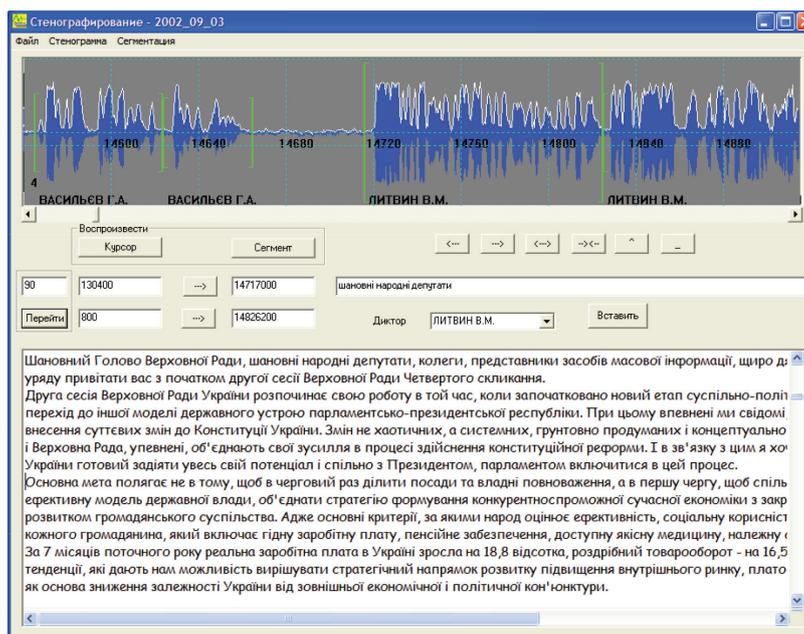


Рис. 1. Общий вид программы стенографирования

дорожки с автоматически выделенными сегментами речи (фразами или синтагмами). Оператор выделяет нужный ему сегмент и прослушивает его. При этом он может просмотреть ответ распознавания, который можно исправить в случае ошибки. После редактирования ответ добавляется в стенограмму и автоматически происходит переход к следующему сегменту.

Пользователь имеет возможность перейти к нужному диктору или прослушать необходимый сегмент стенограммы.

Распознавание производится автоматически в фоновом режиме работы программы. Все ошибки распознавания фиксируются и, после того как закончилось формирование стенограммы, используются для дообучения системы стенографирования. При этом в обучающую выборку добавляются новые слова и информация о новых дикторах. Таким образом, надёжность распознавания повышается в процессе эксплуатации системы стенографирования.

2.2. Предварительная обработка речевого сигнала

Речевой сигнал, оцифрованный на частоте 22050 Гц с точностью 16 бит, преобразуется в последовательность векторов признаков с интервалом анализа 25 мсек. и шагом анализа 10 мсек. Вначале речевой сигнал фильтруется фильтром высоких частот с характеристикой $P(z) = 1 - 0.97z^{-1}$. Затем применяется окно Хэмминга и вычисляется быстрое преобразование Фурье. Спектральные коэффициенты усредняются с использованием 26 треугольных окон, расположенных в мел-шкале, и вычисляются 12 кепстральных коэффициентов.

Логарифм энергии добавляется в качестве 13-го коэффициента. Эти 13 коэффициентов расширяются до 39-мерного вектора параметров путём дописывания первой и второй разностей от коэффициентов, соседних по времени. Для учёта влияния канала применяется вычитание среднего кепстра.

2.3. Акустическая модель

В качестве акустических моделей используются скрытые Марковские модели. 56 украинских фонем (включая фонему-паузу) моделируются тремя состояниями Марковской цепи без пропусков. Используется диагональный вид Гауссовских функций плотности вероятности.

Редко встречающиеся фонемы моделируются 64 смесями Гауссовских функций плотности вероятности, более часто встречающиеся фонемы моделируются большим числом смесей, наиболее часто встречающиеся фонемы используют 1024 смеси.

Словарь транскрипций создаётся автоматически из орфографического словаря с использованием контекстно-зависимых правил.

2.4. Многодикторная система

Распознавание речи независимо от диктора является задачей, не решённой до конца в распознавании речи. В [3] использовалась модель кооперативного распознавания, в



которой при обучении смешивалась речь разных дикторов. При этом речь разных дикторов рассматривалась как разные реализации речи одного диктора. Было показано, что надёжность распознавания улучшалась не только для дикторов, входящих в кооператив, но также и для дикторов, незнакомых системе. Скорее всего, это связано с тем, что речь знакомых системе дикторов похожа на речь других дикторов.

Опыт применения такого подхода показал, что при использовании больше 100 дикторов в кооперативе надёжность распознавания речи становится очень близкой к системе, независимой от диктора.

Методы работы с многими дикторами, заложенные в инструментарий НТК, такие, как нормализация длины речевого тракта (Vocal Tract Length Normalisation) и адаптация модели при помощи линейного преобразования максимального правдоподобия (Maximum Likelihood Linear Regression), позволяют улучшить надёжность распознавания речи для отдельных дикторов при условии, что каким-то независимым способом заранее определяется диктор. Предполагается использовать методы идентификации дикторов для автоматического определения говорящего.

3. Акустическое и текстовое наполнение

3.1. Обучающая выборка

Обучение производилось на выступлениях депутатов Верховной Рады Украины, записанных через телевизионную сеть. Парламентская речь характеризуется некоторыми особенностями.

- Это спонтанная речь. Встречаются отдельные доклады, зачитываемые по подготовленному заранее тексту, однако мало дикторов в точности придерживается этого текста.
- Из-за ограничения во времени выступления многих дикторов произносятся в слишком быстром темпе.
- Часто речь эмоционально окрашена.
- В основном, записи состоят из непрерывных выступлений дикторов, но в них встречаются реплики ведущего заседания или других депутатов.
- Качество записи достаточно высокое, поскольку каждое депутатское место оснащено микрофоном.

Для обучения использовались записи длиной в 197 тыс. секунд, в которых встретилось 427 576 слов. Всего было записано 287 дикторов. Дикторов с длиной больше 300 сек. оказалось 139, что позволяет сформировать качественные акустические модели для данных дикторов. Время записи сильно отличается для разных дикторов.

Обучение производилось на предварительно размеченной выборке. Для этого запись выступления автоматически разбивалась на фразы из нескольких слов, ограниченные паузами больше 400 мсек. Каждой фразе оператор ставил в соответствие метку в виде текста из стенограммы. Затем автоматически производилось преобразование текста в последовательность фонем в соответствии с контекстно-зависимыми правилами. Выборка, размеченная таким образом, использовалась для построения акустической модели при помощи инструментария НТК.

3.2. Контрольная выборка

Распознавание производилось на выступлениях депутатов, записанных в дни, отличающиеся от обучающей выборки. Для распознавания использовались записи длиной в 42 тыс. секунд, в которых встретилось 94 817 слов. Всего использовались записи 152 дикторов. Дикторов с длиной больше 300 сек. оказалось 45. Записи 41 диктора не встретились в обучающей выборке. Таким образом, эти дикторы оказались незнакомыми для системы распознавания.

3.3. Текстовый материал

Словарь был составлен из текстов стенограмм заседаний Верховной Рады Украины. С официального сайта Верховной Рады были загружены все стенограммы заседаний, начиная с 1991 года, что составило больше 100 МБ текста. Текст был модифицирован для того, чтобы убрать служебную информацию из стенограмм (например, обозначения аплодисментов и т.п.), записать числа в текстовом виде, а также отделить русский текст от украинского.

Результирующий текст разделён на две части: первая (обучающая выборка) содержит все тексты, кроме 2002—2003 годов, вторая содержит стенограммы 2002—2003 годов (контрольная выборка). Первая часть состоит из 14 629 111 словоупотреблений, во второй содержится 409 244.

Для первой части был составлен словарь из 156 108 слов и подсчитаны частоты встречаемости слов в этом словаре. Таблица 1 показывает долю текстов, которая покрывается словарём с определённой частотой встречаемости слов. Видно, что весь словарь (156 108 слов, частота встречаемости слов — один раз и больше) покрывает 99,6% нового текста. Доля текста из контрольной выборки, покрываемого словами с частотой встречаемости в обучающей выборке 50 раз и больше, составляет 94,9%. Для этого достаточно иметь словарь объёмом в 15 тыс. слов.

Исследовалась надёжность распознавания в зависимости от объёма частотного словаря с использованием биграммной модели языка. Результаты представлены в таблице 2, из которой видно, что надёжность незначительно увеличивается при увеличении размера словаря. «Закрытого» словаря объёмом в 15 тыс. слов достаточно для распознавания речи с небольшим (2%) уменьшением надёжности по сравнению с максимально возможным словарём. «Закрытый» словарь не предполагает возможность отказа от распознавания.

Таблица 1

Доля текстов, входящих в частотные словари

Частота встречаемости слов	Количество слов в частотном словаре	Встречаемость в текстах обучающей выборки	%	Встречаемость в текстах контрольной выборки	%
1	156 108	14 629 111	100,0	407 608	99,6
2	98 601	14 571 604	99,6	406 563	99,3
3	78 022	14 530 446	99,3	405 753	99,1
5	58 936	14 465 646	98,8	404 293	98,7
10	40 364	14 343 499	98,0	401 544	98,1
50	15 609	13 805 357	94,3	388 470	94,9
100	10 032	13 415 092	91,7	378 873	92,5
200	6 219	12 878 264	88,0	365 510	89,3
300	4 622	12 488 459	85,3	355 606	86,8

Таблица 2

Надёжность распознавания для разных объёмов частотного словаря

Объём словаря	64 000	50 000	30 000	20 000	15 000	10 000	5 000
Надёжность распознавания (%)	68,59	68,54	68,38	67,79	67,15	65,49	62,18

4. Биграммная модель языка

При распознавании речи использовалась биграммная модель речи, которая задавалась вероятностями появления пар слов. Поскольку в текстах, на которых вычислялись статистики, встретились далеко не все пары слов, возможные для данного словаря, то для аппроксимации *ненаблюдаемых* пар слов использовались так называемые *обратные* (back off) коэффициенты [2].

Биграммная модель языка позволила исправить достаточно много ошибок распознавания, которые не укладывались в модель языка. В таблице 3 показаны примеры исправления таких ошибок.

5. Индивидуализация транскрипций

5.1. Моделирование особенностей произношения дикторов

Для превращения орфографического текста в фонемный был разработан и сформирован набор контекстно-зависимых правил, по которым орфогра-

Таблица 3

**Примеры исправления ошибок распознавания
при помощи биграммной модели языка**

Было сказано	Результат распознавания с использованием свободного порядка слов	Результат распознавания с использованием биграммной грамматики
Доброго ранку	До в в о ранку	Доброго ранку
Шановні народні депутати запрошені та гості верховної ради	Шановні народі депутати запрошені та гості верховної ради	Шановні народні депутати запрошені та гості верховної ради
Прошу вас шановні колеги займати вас свої робочі місця	Прошу в о з шановні колег і й з е мати з в й й робоче місця	Прошу вас шановні колеги займати вас свої робочі місця
Прошу займати робочі місця	Прошу з е мав те й робоче й місця	Прошу займати робочі місця
Прошу підготуватися до реєстрації	Прошу б й готуватися до реєстрації б	Прошу підготуватися до реєстрації

фическое слово превращается в последовательность фонетических символов (путём преобразования одной последовательности символов в другую). Причём генерируется сразу несколько вариантов транскрипции для случаев неоднозначностей, заданных в правилах.

Для всех дикторов был создан общий вариант транскрибирования на основе правил литературного произношения для украинского языка. Кроме этого, все дикторы были разделены на группы, для которых разработаны свои правила индивидуализированного транскрибирования, которые заменяют или дополняют основной вариант.

Результаты изучения речи многих дикторов свидетельствуют, что ни один из них не придерживается орфоэпических правил произношения в полном объёме. В первую очередь это касается запрещённых литературной нормой регрессивной ассимиляции по глухости в паре фонем «звонкая + глухая» и оглушение согласных перед паузой. Дикторы с такими особенностями произношения были выделены в отдельную группу. Обработанный материал свидетельствует, что звонкие согласные в речи таких дикторов в позиции перед глухими оглушаются: (*тобто* → *т о п т о*; *підтримати* → *п' і т т р И м а т и*). Случаи оглушения звонких согласных перед паузой встречаются у большинства дикторов: (*робив* → *р о б И ф*).

Были выделены и многие другие характерные черты произношения разных дикторов: редуцирование окончаний некоторых слов (прилагательных, глаголов) в слитной речи (*шановний* → *ш а н О в н и*; *доброго* → *д О б р о*), «акание» (*робити* → *р а б И т и*), твёрдое произношение мягких согласных (*синього* → *с И н о го*) и др.

Такие тенденции моделируются путём изменения правил перехода от одних последовательностей символов к другим и расширением существующих правил.

В таблице 4 приведены примеры индивидуализированных транскрипций для нескольких слов. В основном, здесь задействованы правила оглушения и редуцирования окон-

чаний в словах. Для некоторых слов (служебных в том числе) задаётся несколько вариантов транскрипций — с ударением на разных слогах (если в языке возможны разные варианты прочтения таких слов) или вообще без ударения (*коли* → *к о л И*; *к О л и*; *к о л и*).

Таблица 4

Примеры модификации транскрипций слов

Слово	Литературная транскрипция	Модифицированная транскрипция
відповідно	в' і д п о в' і д н о	в' і т п о в' і д н о
головуючий	г о л о в У й у ч и й	г о л о в У й у ч и
доброго	д О б р о г о	д О б р о
коли	к о л И	к о л И к О л и к о л и
народного	н а р О д н о г о	н а р О д н о
підтримати	п' і д т р И м а т и	п' і т т р И м а т и п' і т р И м а т и
при	п р И	п р И п р и
робив	р о б И в	р о б И ф
робити	р о б И т и	р а б И т и
синього	с И н' о г о	с И н о г о
тільки	т' і л' к и	т' і л' к и т' і л' к и
тобто	т О б т о	т о п т о
шановний	ш а н О в н и й	ш а н О в н и

5.2. Правила индивидуализированной модификации транскрипций

Все правила индивидуализированной модификации транскрипций можно разделить на несколько групп.

К позиционным изменениям звуков в речевом потоке (изменения, которые зависят от таких условий, как позиция звука в слоге/слове, ударность/безударность, и др. [4]) относим:

- 1) кроме редукции безударных *e*, *i*, *o* (произносятся как *e^u*, *i^e*, *o^y*), также слабое произношение *o* как *a* в безударной позиции, реже встречается редукция безударных гласных до полного их исчезновения (*тепер* → *т и п Е р*, *зозуля* → *з у з У л' а*, *боротьба* → *б а р а д' б А* или *б р а д' б А*);
- 2) оглушение звонких согласных перед паузой (*бріd* → *б р' І т*, *зараз* → *з А р а с*);
- 3) редукцию в терминальных частях слов в процессе речи — исчезновение согласного звука в окончаниях - *ого*, - *их*, - *ий*, - *іх*, - *ій*, - *ії*, - *ої*, - *еї*,

-ою, -сю, -ити и подобных (коротший → к о р О ч ш и, синіх → с И н' і, безпекою → б е с п Е к о у); исчезновение конечного гласного звука в окончаниях -ою, -ею, -сю и подобных (доброю → д О б р о й, землю → з е м л Е й) и др.

К комбинаторным изменениям (качественные и количественные изменения соседних звуков [4]) относим:

- 1) полную регрессивную ассимиляцию по глухости в паре фонем «звонкая + глухая» на стыке любых морфем и на стыке слов (без причини → б е с п р и ч И н и, розсунути → р о с с У н у т и, книжка → к н И ш к а, сядьте → с' А т' т е);
- 2) ассимиляцию по мягкости свистящих и шипящих согласных, губных и заднеязычных согласных (злі → з' л' І, шлях → ш' л' А х, квітка → к' в' І т к а);
- 3) произношение удлинённых и удвоенных согласных звуков как одного короткого звука, произношение двух гласных звуков как одного звука (віддати → в' і д А т и, знання → з н а н' А, зоопарк → з о п А р к, аеропорт → а р о п О р т);
- 4) неполное упрощение сочетаний согласных (чесний → ч Е с т н и й) и др.

5.3. Индивидуализированные транскрипции словарей для групп дикторов

Отбор групп дикторов из обучающей выборки в связи с особенностями произношения проходил в несколько этапов. Сначала были внесены в отдельную группу дикторы с литературным произношением, после этого производилась классификация индивидуальных особенностей речи оставшихся дикторов. В результате мы получили 15 групп дикторов с разными особенностями произнесения, для каждой группы были сгенерированы свои правила транскрибирования словарей (15 типов затранскрибированных словарей).

Для проверки соответствия индивидуализированных транскрипций словарей и живой речи была проведена процедура распознавания речи каждого диктора (вне зависимости от группы, в которую он попал) с использованием всех 15 типов словарей.

После анализа результатов эксперимента каждому диктору были приписаны те правила транскрибирования, которые повысили надёжность распознавания его речи [5]. В результате для всех групп дикторов сформировался свой словарь с индивидуализированными правилами транскрибирования, который повысил надёжность распознавания речи (таблица 5).

Таблица 5

Надёжность распознавания с применением индивидуализированных словарей

Диктор	Длительность записи (сек.)	Надёжность распознавания с использованием лит. словаря (%)	Надёжность распознавания с использованием инд. словаря (%)	Изменение надёжности распознавания (%)
STO	2184	58,61	59,34	0,73
TER	1124	75,18	75,53	0,35
CER	1756	67,14	68,05	0,91
VAS	1629	67,40	69,00	1,6
BAB	1194	58,12	59,72	1,6

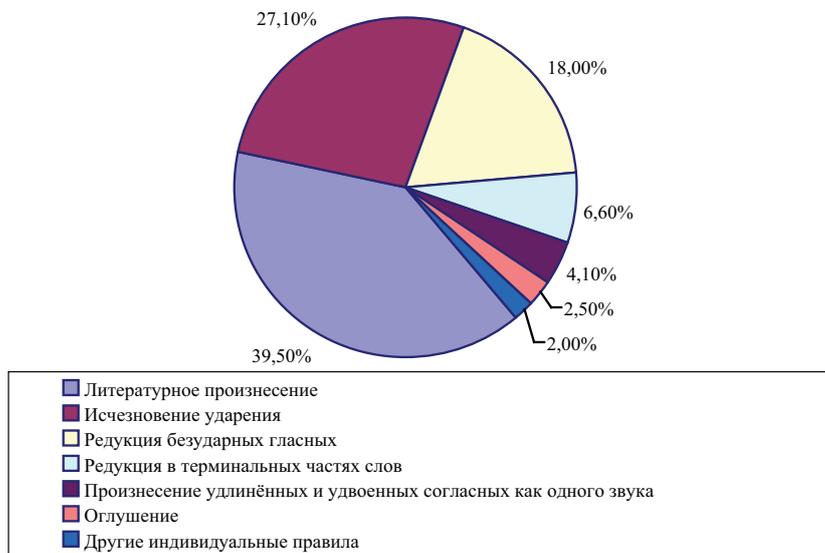


Рис. 2. Соотношение литературной и индивидуализированных транскрипций речи дикторов из обучающей выборки

Эксперимент показал, что существуют нелитературные правила, которые могут быть приписаны большинству дикторов, например: редукция окончаний слов, обусловленная слишком быстрым темпом речи или эмоциональностью высказывания; ассимиляция по глухости, обусловленная влиянием русского языка, и др. Для многих дикторов характерно произношение удлинённых и удвоенных звуков как одного короткого звука. Часто встречается произнесение ударных гласных как безударных (в речевом потоке ударение исчезает не только в односложных словах, но и в многосложных). Остальные правила характерны для одного или нескольких дикторов (рисунок 2).

5.4. Генерирование словаря транскрипций для обучающей выборки

Словарь транскрипций для обучающей выборки был сформирован на основе 15 типов индивидуализированных словарей. Каждое слово в словаре имеет то количество транскрипций, сколько существует гипотетических вариантов произнесения этого слова каждым диктором из обучающей выборки. До использования правил индивидуализации словарь транскрипций состоял приблизительно из 36 000 литературных транскрипций слов; после использования этих правил, а также комбинаций правил словарь увеличился до 285 000 транскрипций слов.

6. Результаты экспериментов по распознаванию слитной речи

Эксперименты проводились на описанной контрольной выборке в виде записей заседаний в течение одного дня. Надёжность распознавания сильно отличается в зависимости от того, какие дикторы попали в выборку.

Таблица 6 представляет результаты распознавания для некоторых дикторов, где также приведены длина обучающей выборки и темп произнесения для каждого диктора. Анализ результатов показывает, что в среднем указанные факторы (длина обучающей выборки и темп речи) влияют на надёжность распознавания.

Надёжность распознавания для отдельных дикторов сильно отличается — от 50% до 90%. В среднем по всей контрольной выборке она составляет 72.3%. Использование индивидуализированных транскрипций улучшило надёжность распознавания в среднем на 1,5%.

Таблица 6

Надёжность распознавания для некоторых дикторов

Диктор	Длина обучающей выборки (сек.)	Длина контрольной выборки (сек.)	Количество слов в контрольной выборке	Темп (слов/сек.)	Надёжность распознавания (%)
LIT	15 805	2336	5 721	2,45	79,85
POR	3 715	411	853	2,08	80,30
MOR	1 728	362	950	2,62	70,74
SIM	1 484	125	255	2,04	80,00
MAT	1 305	174	292	1,68	80,14
KLU	998	107	209	1,95	86,60
KIN	585	223	417	1,87	64,27
ONI	483	100	209	2,09	79,90
MIS	195	148	312	2,11	69,87
ZVA	25	101	205	2,03	80,00
GOL	0	379	790	2,08	78,48
KAP	0	375	927	2,47	80,91

Время распознавания для компьютера Pentium 2GHz составляет около 10 секунд для одной секунды речи. Применение алгоритмов ускорения принятия решений [6] позволит достичь распознавания речи в реальном времени.

Заключение

Статья описывает экспериментальную систему автоматизированного стенографирования.

Показана возможность построения таких систем при условии повышения надёжности распознавания речи до необходимых для практических применений показателей.

Предложено использовать индивидуальную информацию о дикторах для улучшения надёжности распознавания.



Литература

1. V. Pylypenko. Extra Large Vocabulary Continuous Speech Recognition Algorithm based on Information Retrieval // Proc. of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007). Antwerp, Belgium, 2007.
2. S. Young, G. Evermann, D. Kershaw, G. Moore, J. Odell, D. Ollason, V. Valtchev, P. Woodland. The HTK Book. Cambridge University Engineering Department, 2002.
3. Винцюк Т.К. Анализ, распознавание и смысловая интерпретация речевых сигналов. Киев: Наукова думка, 1987. 264 с.
4. Сучасна українська літературна мова. Фонетика: Навч. посібник для студентів-філологів. К.: Видавничо-поліграфічний центр «Київський університет», 2002. С. 60.
5. V. Pylypenko, V. Robeiko. Experimental System of Computerized Stenographer for Ukrainian Speech // Proc. of the 13th International Conference «Speech and Computer: SPECOM'2009». St.-Petersburg, Russia, 2009. P. 67—70.
6. Пилипенко В.В. Распознавание дискретной и слитной речи из сверхбольших словарей на основе выборки информации из баз данных // Искусственный интеллект. № 3. 2006. С. 548—557.

Пилипенко Валерий Васильевич

Старший научный сотрудник Международного научно-учебного центра информационных технологий и систем (г. Киев, Украина). Окончил Московский физико-технологический институт в 1984 году. Специалист в области распознавания речи.

Робейко Валентина Васильевна

Окончила Институт филологии Киевского национального университета им. Т. Шевченко. Специалист в области современного украинского языка. Работает младшим научным сотрудником в Международном научно-учебном центре информационных технологий и систем (г. Киев, Украина), аспирантка. Область интересов: синтез и распознавание речи, фонетика.