



Разработка методов биометрической идентификации личности по голосу



ЦЕЛЬ, ЗАДАЧИ

Цель – разработка эффективных (по точности и скорости вычислений) алгоритмов автоматического анализа, распознавания и обучения речи на основе кластерных моделей элементарных речевых единиц.

Задачи:

- Разработка алгоритма распознавания речи на основе кластерной модели элементарных речевых единиц и исследование его эффективности в задаче распознавания изолированных слов.
- Экспериментальная оценка выигрыша в скорости вычислений в задаче распознавания речи с использованием кластерной модели элементарных речевых единиц.
- Разработка адаптивной кластерной модели элементарных речевых единиц и ее экспериментальное исследование в задаче фонетического анализа речи.
- Разработка новой технологии информационного обеспечения с использованием кластерной модели речевых единиц для решения задачи обучения речи и постановки произношения.



Задача фонетического анализа речи

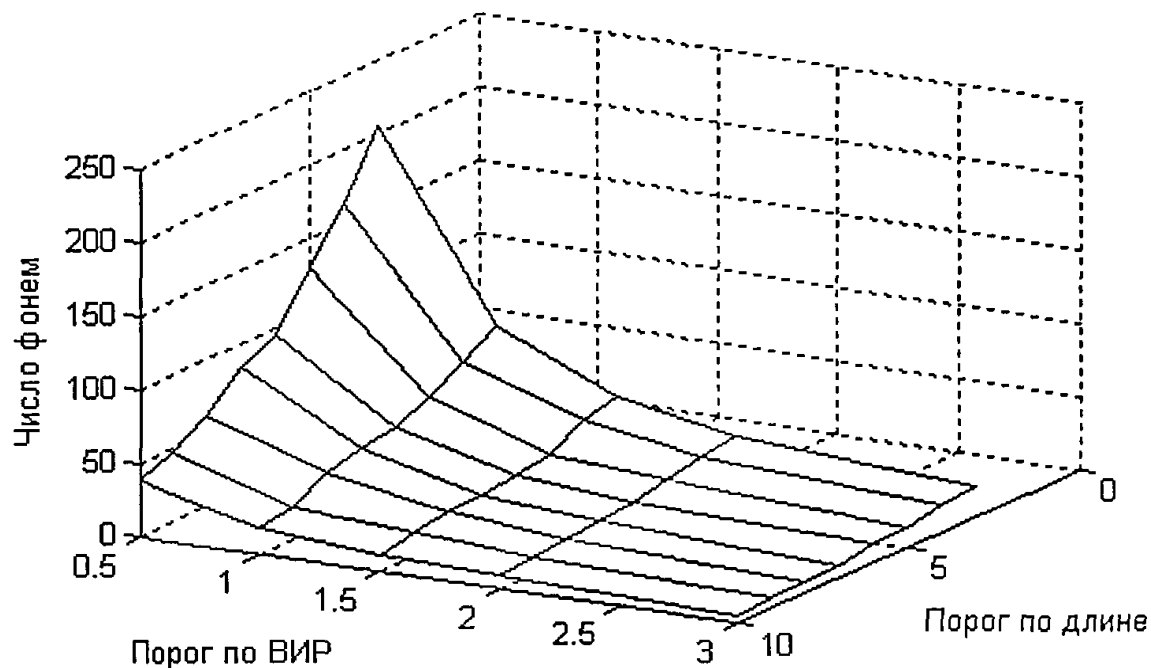
$$p(X_2, X_1) = p_{x,r} \Big|_{X=X_2}$$

$$p(X_1, X_2) \geq p_0$$

В выборе их порогов p_0 и V_0 , кроются и основные возможности, и основные проблемы практического осуществления метода переменного дерева. Рассмотрим некоторые из них по результатам проведенных ниже экспериментальных исследований.

Результаты математического моделирования алгоритма фонетического анализа речи на основе адаптивной кластерной модели

График числа выделенных фонем из речи первого диктора при различных порогах V_0 и p_0



Участки с выделенными фонемами фразы «Другого ментора я и не желал»

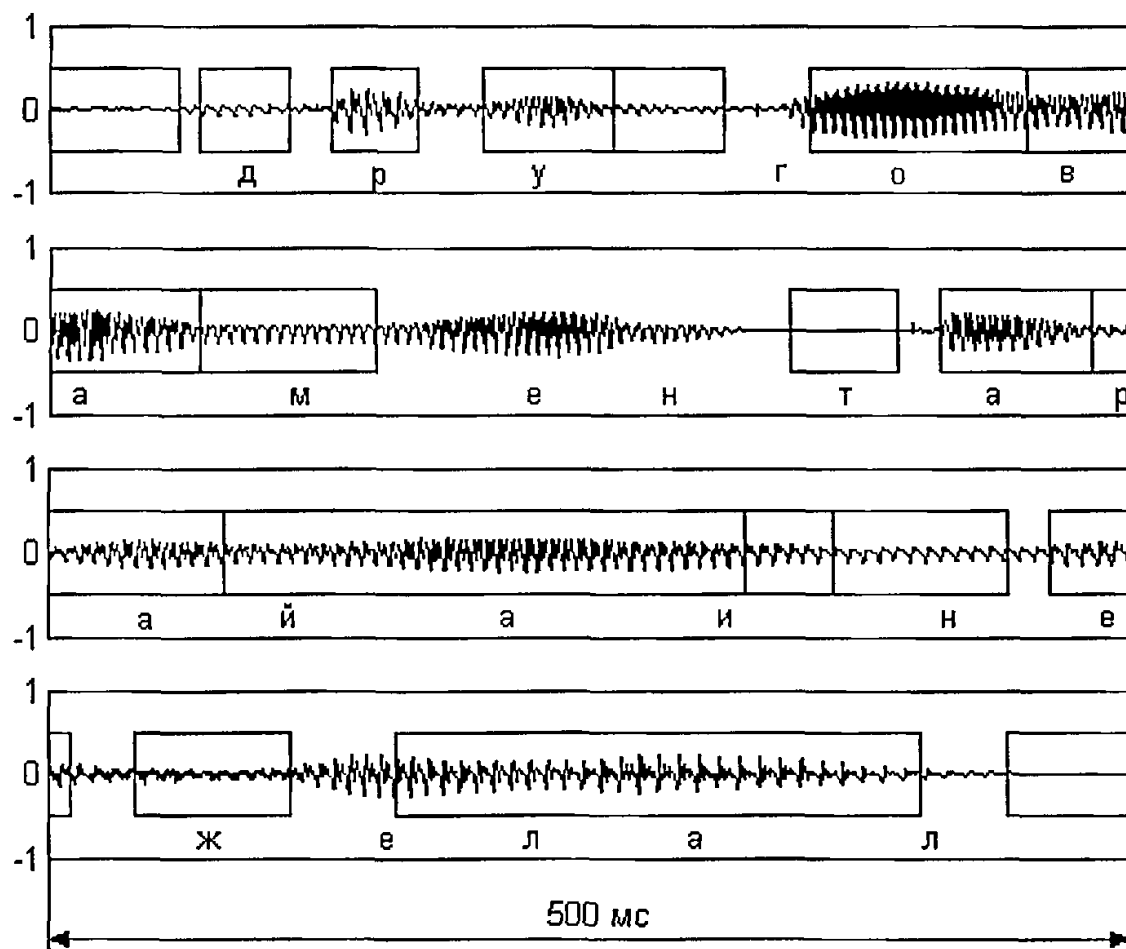
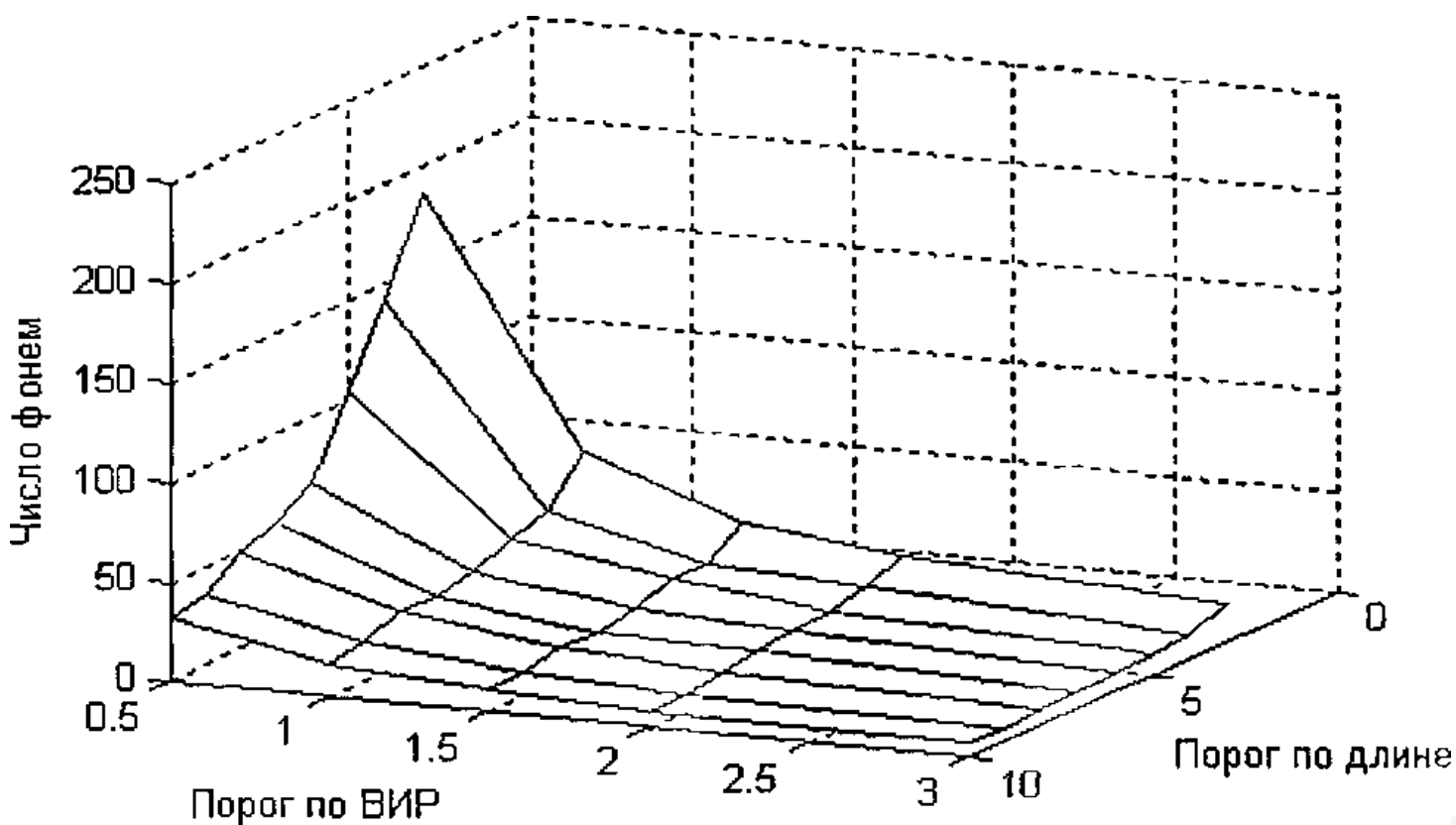
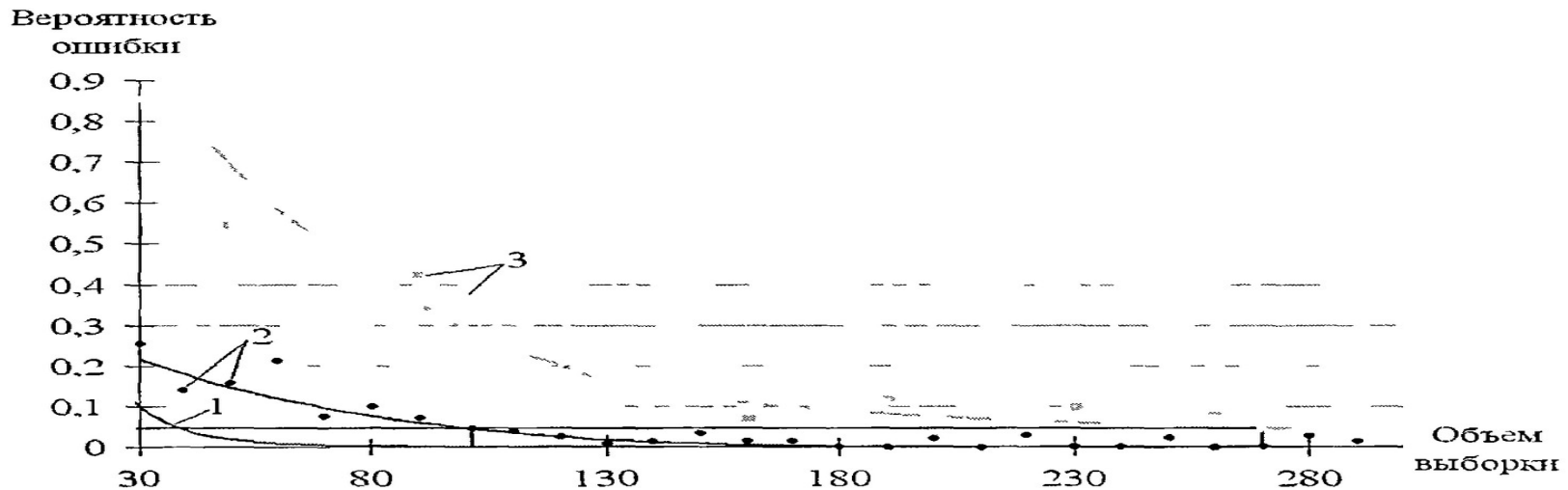


График числа выделенных фонем из речи первого диктора при различных порогах V_0 и p_0



Анализ эффективности алгоритма

Зависимость вероятности ошибки первого рода от объёма выборки в задаче анализа эффективности алгоритмов



- 1 — теоретическая зависимость для алгоритма на основе кластерной модели элементарных речевых единиц;
- 2 — экспериментальная зависимость для алгоритма на основе кластерной модели элементарных речевых единиц;
- 3 - экспериментальная зависимость для алгоритма на основе критерия согласия Пирсона.

Результаты экспериментальных исследований алгоритма фонетического анализа речи на основе адаптивной кластерной модели

Блок-схема ИС ФАР

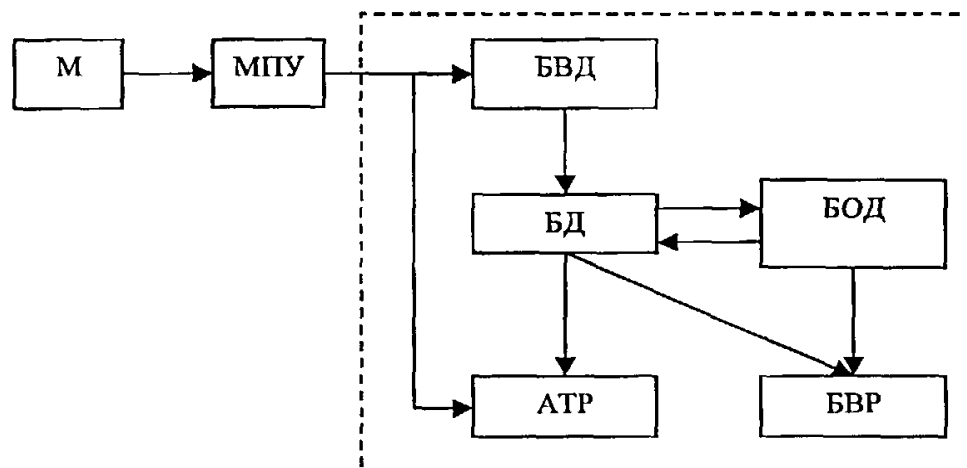
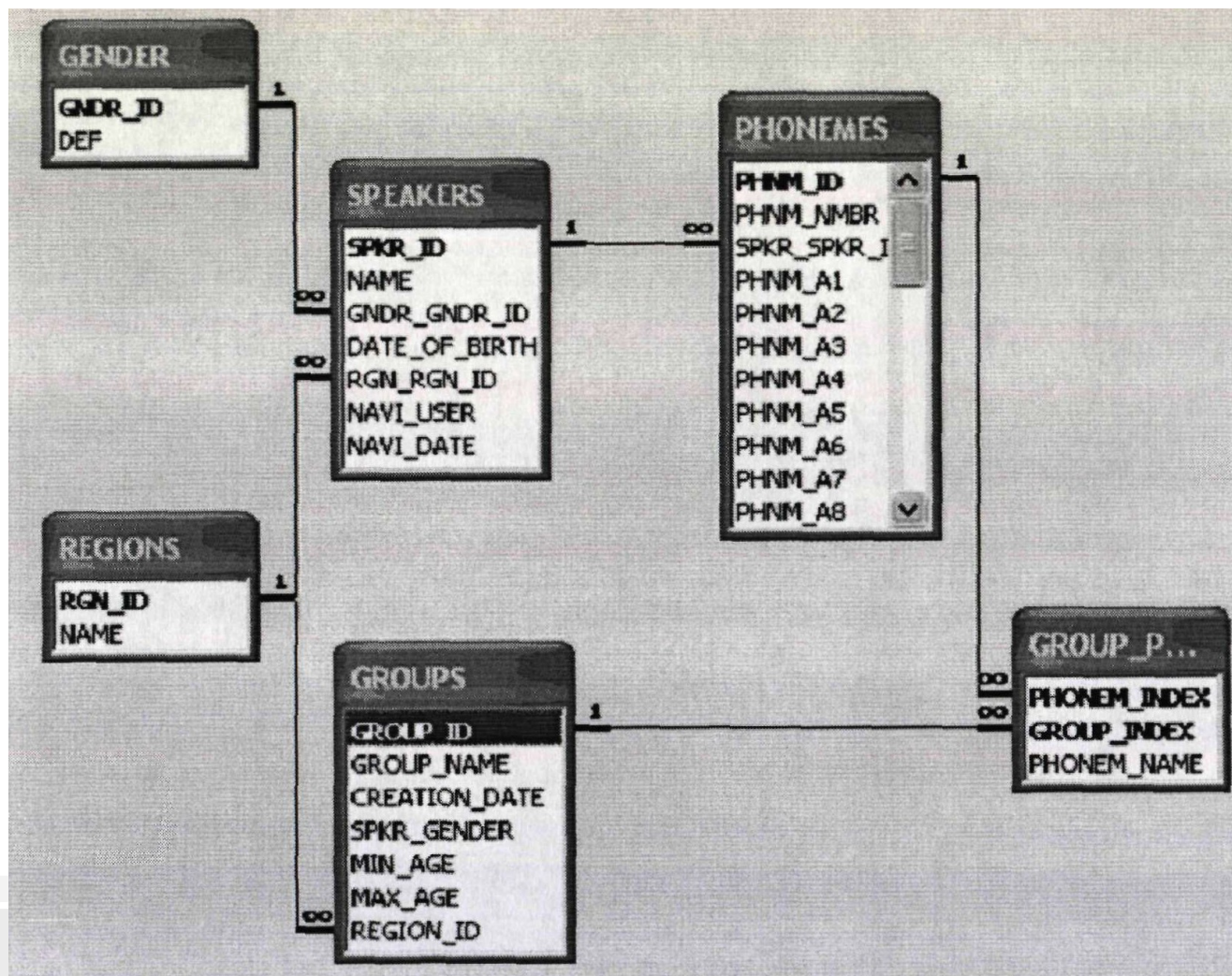


Схема базы данных ИС ФАР



Повышение точности алгоритма распознавания речи

Вероятности правильного распознавания слов для решающей статистики без ограничения и с ограничением

Слово	Алгоритм	
	Для решающей статистики (1.10)	Для решающей статистики (3.1) при пороге $U=0,5$
0	0,89	1
1	0,98	1
2	0,99	1
3	0,97	0,98
4	0,99	1
5	0,95	0,99
6	0,97	0,97
7	0,92	0,98
8	0,95	1
9	0,96	0,98
Итого	0,957	0,99

Зависимость вероятности правильного распознавания слов от величины порога U

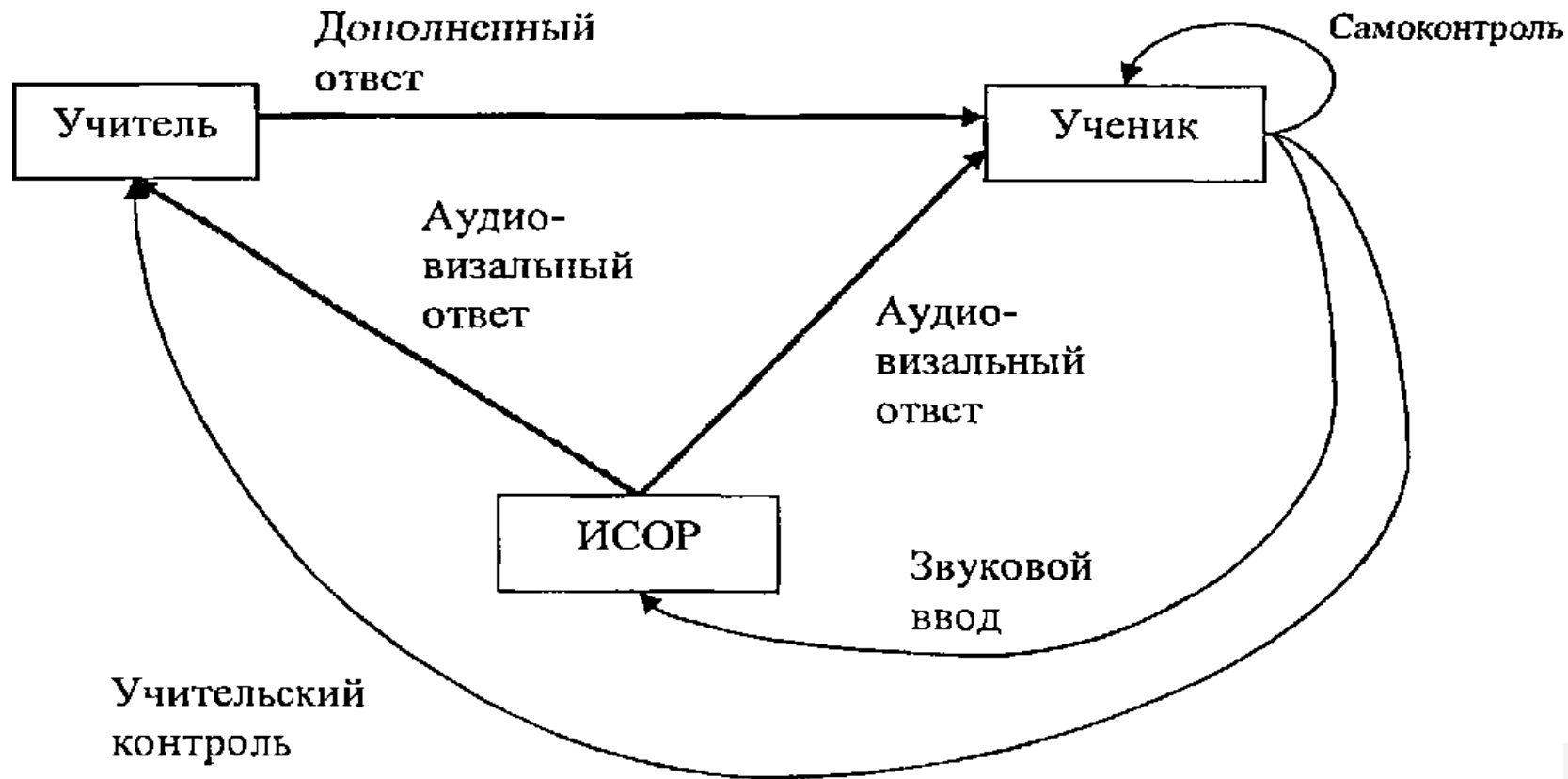
Слово	Порог						
	0,3	0,5	1	1.5	2	2,5	3
0	0,99	1	1	1	0,99	0,98	0,98
1	1	1	1	1	1	1	1
2	1	1	1	1	0,98	0,97	0,96
3	0,98	0,98	0,98	0,98	0,98	0,97	0,97
4	0,99	1	0,99	0,99	0,99	0,99	0,99
5	0,96	0,99	0,98	0,98	0,97	0,95	0,95
6	0,96	0,97	0,98	0,98	0,99	0,99	0,99
7	0,82	0,98	0,87	0,72	0,67	0,66	0,66
8	1	1	1	1	1	1	1
9	0,98	0,98	0,99	0,99	0,99	0,99	0,98
Итого:	0,968	0,99	0,979	0,964	0,956	0,95	0,948



Обзор существующих подходов к решению задачи обучения речи с визуализацией речевого сигнала

- Системы отображения речевого тракта.
- IBM SpeechViewer.
- Indiana Speech Training Aid (ISTRA).
- The Video Voice Speech Training System.
- Speech Rehabilitation Speech Training Aid for the Hearing Impaired (HARP).
- «Контроль произношения»
- "Речевой калейдоскоп" - НПО "КАСКОД" — С.Петербург;
- "Филиппок" - КУДИЦ - Москва; "КОСПР" — Минск.

Схема обучения произношению при помощи разработанной технологии





Выводы

- Использование ограничения в решающей статистике позволяет учитывать только информативные сегменты, что повышает надежность работы систем автоматического распознавания речи.
- Благодаря направленным свойствам решающей статистики МИР можно значительно сократить вычислительные затраты в системах автоматического распознавания речи.
- Разработанная адаптивная кластерная модель элементарных речевых единиц может найти широкое применение при создании новых и обновлении существующих речевых баз данных.
- Разработанная информационная система обучения речи может найти практическое применение не только при постановке речи и обучении языкам, но и в исследовательских целях, например, при сравнительном анализе качества произношения разных дикторов, анализе зависимости их произношений от внешних факторов, в частности, эмоционального состояния и т.п.

СПАСИБО ЗА ВНИМАНИЕ!