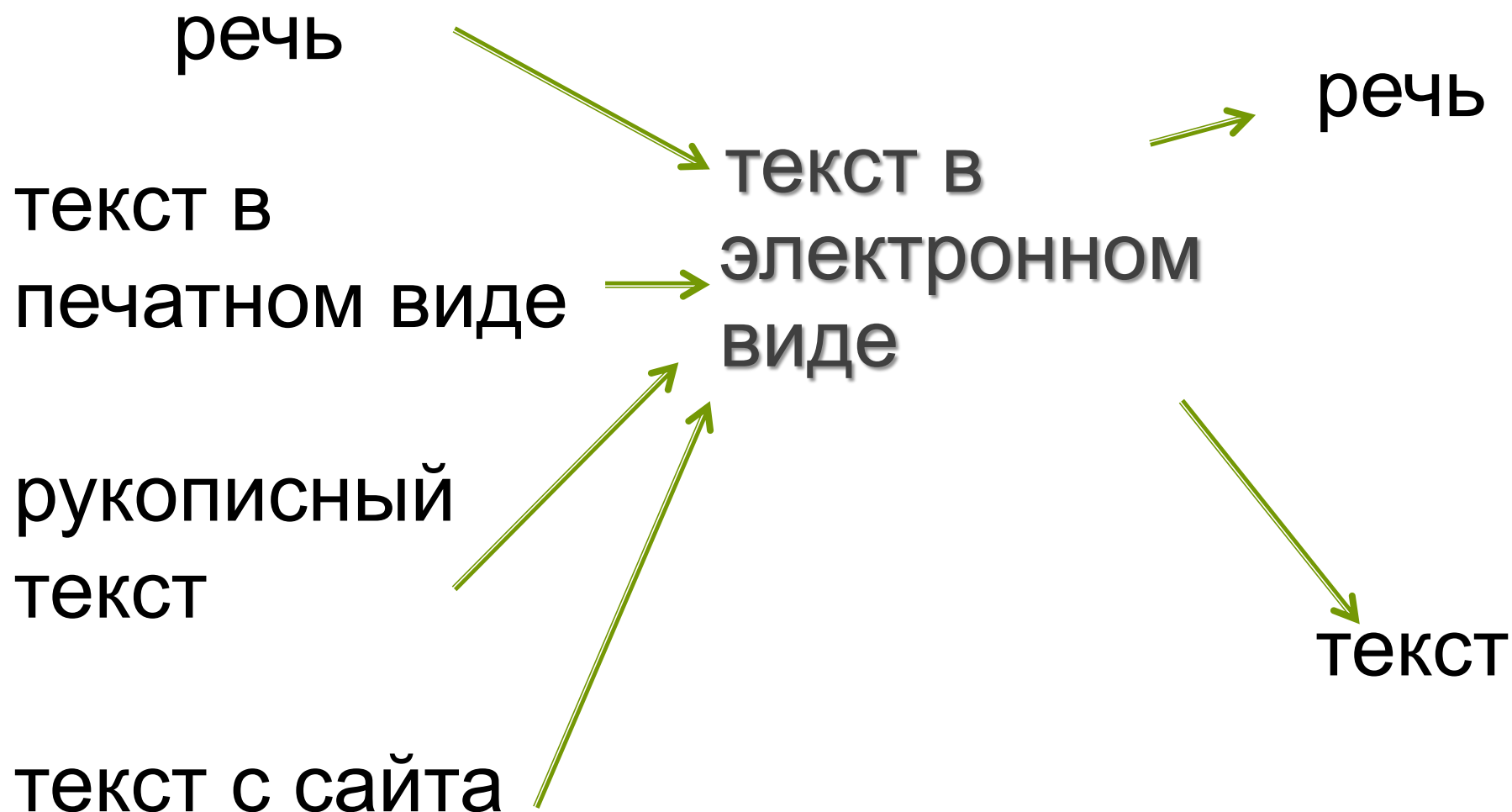




Распознавание текста Генерация и синтез речи

Ефремова Наталья Эрнестовна
Грацианова Татьяна Юрьевна
Большакова Елена Игоревна

Содержание



Распознавание текста – определение, история



Перевод изображений рукописного, машинописного или печатного текста в текстовые данные, использующихся для представления символов в компьютере (например, в текстовом редакторе).

Оптическое распознавание текста является исследуемой проблемой в областях распознавания образов, искусственного интеллекта и компьютерного зрения.

30-е годы прошлого века – оптическое распознавание символов

50-е годы – практическое использование (почта, кредитные карты, купоны)

Распознавание текста - классификация



Форма текста

- Печатный (известный шрифт)
- рукописный

Статистическое и динамическое распознавание

- Статическая форма текста (оф-лайн)
- он-лайн распознавание непосредственно во время написания

Цель распознавания

- Простое конвертирование документа в электронный вид
- извлечение необходимой информации

Распознавание текста: языковая модель



После распознавания может выполняться дополнительная коррекция, позволяющая увеличить качество распознавания спорных символов (то есть символов у которых есть несколько кандидатов с приблизительно одинаковой оценкой степени соответствия нескольким эталонам) на основе:

- анализа буквосочетаний, характерных для языка;
- словаря языка;
- грамматического анализа;
- и других методов.

Распознавание речи.

Проблемы



- Посторонние шумы
- Вариативность произношения (изменение тембра голоса, акцент)
- Слитность речи (границы слов и предложений), «смазываются» концы слов и фраз
- Различное написание и произношение
- Случайные паузы, повторы, слова-паразиты
- Знаки препинания

История

- Первая экспериментальная система 1952 г.
- коммерческие системы – 70-е годы, часто неудачные
- активные разработки, использование – начало 21 века



Распознавание речи. Классификация

- Дикторозависимая/независимая, адаптивная
- Изолированные команды или связная речь, спонтанная разговорная речь, ключевые слова
- Размер словаря
- Специализация

Распознавание речи. Этапы



1. Выделение последовательностей звуков, которые могут быть интерпретированы как слова (несколько гипотез)
2. Проверка гипотез, выбор нужной

Языковая модель: N-граммы, согласование слов, словарь (машинное обучение, нейронные сети, специализированная тематика)

ограничения на комбинации фонем

составление слогов

лексический уровень

синтаксический и семантический уровень
(сочетаемость слов)

Синтез речи. Виды



Компилятивный (компиляционный), синтез на основе конкатенации

По правилам (параметрический)

форматный

статистический

артикуляционный

Компилятивный (компиляционный), синтез на основе конкатенации



- единицы словаря: фразы, слова

- единицы — слоги, звуки (аллофоны), дифоны (речевой фрагмент, начинающийся в середине одного звука и заканчивающийся в середине следующего звука)

выборочная интонационная и темпоральная модификация сегментов

Селективный метод (технология выбора образца) — поиск единиц разного уровня в большой базе

Форматный синтез



Порядка 60 параметров: частота, амплитуда, ширина форманты..

Генерация цифровых сигналов, имитирующих результат работы источников звука

Недостатки: роботоподобность, механические артефакты

Достоинства: генерация любого голоса, манеры произнесения, темпа речи и т.п.

небольшой размер

Статистический параметрический синтез речи



Обучение на естественной речи диктора

клонирование акустических параметров (частота, тембр и
т.п.)

Артикуляционный синтез



Моделирование процесса речеобразования (геометрия речевого тракта человека, положения и движения органов речи)

Исторически первый метод

Недостатки: ориентация на субъективные представления экспертов

небольшая скорость

среднее качество

Лингвистическая обработка текста для синтеза речи



Преобразование орфографической записи в фонетическую транскрипцию

- Определение структуры и разметки документа (абзацы, знаки препинания), отбрасывание метаинформации (номера страниц)
- Нормализация текста, расшифровка аббревиатур и спецзнаков, исправление ошибок; остаются только значимые пунктуационные знаки
- Лингвистический анализ
 - Морфологический анализ (ударения)
 - Синтаксический, семантический анализ (правильная акцентировка)
- Фонетический анализ (преобразования от буквы к звуку)
- Просодическая параметризация



Спасибо за внимание