

**Лекция №2**

**ТЕХНОЛОГИИ  
АВТОМАТИЧЕСКОГО  
РАСПОЗНАВАНИЯ ОБРАЗОВ**

# OCR-СИСТЕМЫ



Методы автоматического распознавания образов и их реализация в *системах оптического чтения текстов* (OCR-системах — **Optical Character Recognition**) - одна из самых плодотворных технологий ИИ.

В приведенной трактовке **OCR** понимается как автоматическое распознавание с помощью специальных программ изображений символов печатного или рукописного текста и преобразование его в формат, пригодный для обработки текстовыми процессорами, редакторами текстов и т.д.

Сокращение OCR иногда расшифровывают как **Optical Character Reader**. В этом случае под OCR понимают устройство оптического распознавания символов или автоматического чтения текста.

# OCR-СИСТЕМЫ



Отметим следующие особенности предметной области (PrO), существенные с точки зрения OCR-систем:

- **шрифтовое и размерное разнообразие символов;**
- **искажения в изображениях символов** (разрывы образов символов, например, при увеличении изображения; слипание соседних символов и др.);
- **перекосы при сканировании;**
- **посторонние включения в изображениях;**
- **сочетание фрагментов текста на разных языках;**
- **большое разнообразие классов символов, которые могут быть распознаны только при наличии дополнительной контекстной информации.**

Автоматическое чтение печатных и рукописных текстов является частным случаем автоматического визуального восприятия сложных изображений.

Многочисленные исследования показали, что для полного решения этой задачи необходимо интеллектуальное распознавание, т.е. «распознавание с пониманием».

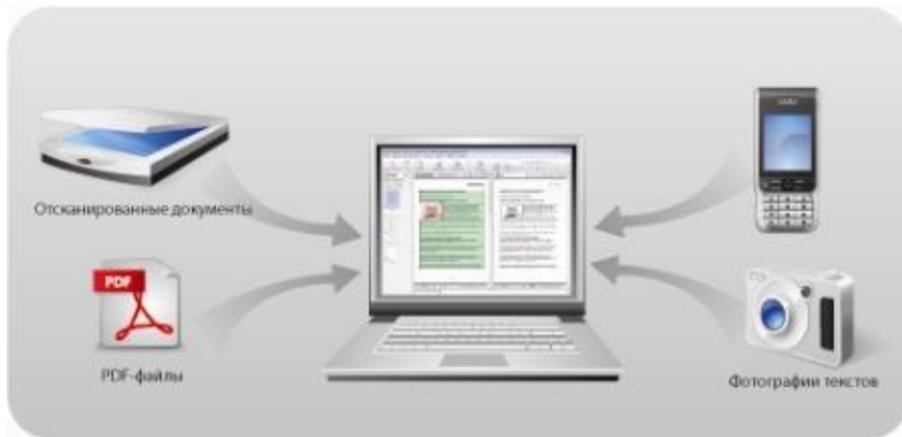
# OCR-СИСТЕМЫ

В современных OCR-системах используется технология распознавания, свойственная человеку. У человека распознавание образа является многоступенчатым.

*Три принципа, на которых основаны все OCR-системы (IPA):*

- ❑ Принцип целостности образа (integrity);
- ❑ Принцип целенаправленности (purposefulness);
- ❑ Принцип адаптивности (adaptability).

**Общая схема распознавания текста**



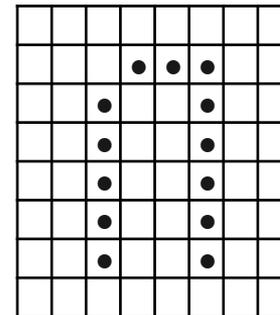
Цифровое  
представление  
текста

# OCR-СИСТЕМЫ



Графический образ символа имеет вид **шейпа**, представляющего собой матрицу из точек, которую можно редактировать поэлементно.

На рисунке приведен пример шейпа  
Буквы «Л» или «П».



При контекстной обработке для распознавания «сомнительного» шейпа привлекается информация о результатах распознавания соседних элементов текста (например, слово «**МЕ\***»).

Но эта информация не всегда достаточна для принятия решения (например, в слове «**\*ИК**» в позиции звездочки может располагаться как «Л», так и «П»). В таких случаях анализируемый контекст включает предложение или фрагмент текста и возникает необходимость решения проблемы понимания текста на естественном языке.

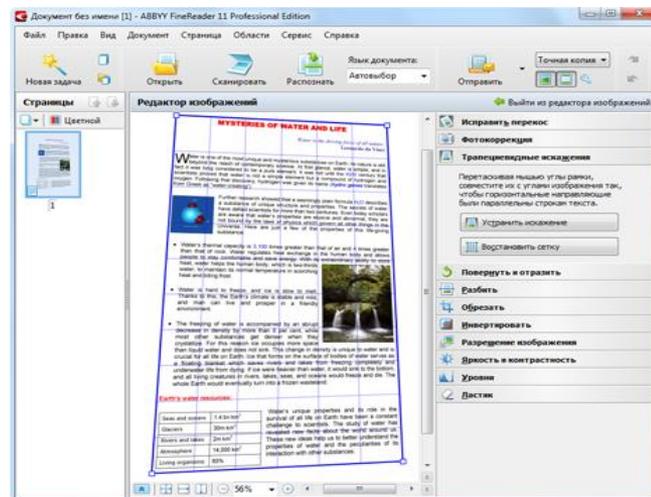
# ПРИМЕРЫ ПРОГРАММНОЙ РЕАЛИЗАЦИИ OCR-СИСТЕМ



К ведущим **российским OCR-системам** можно отнести:

❑ Линейку продуктов **FineReader** фирмы **ABBYY** (<http://www.abbyy.ru>), позволяющих распознавать как печатные, так и рукописные многоязычные тексты;

❑ Программные продукты фирмы **Cognitive Technologies** (<http://www.cognitive.ru>) OCR-систему **CuneiForm** (<http://www.cuneiform.ru>) и **Cognitive Forms**, предназначенную для массового ввода структурированных документов (например, бухгалтерских и налоговых форм отчетности, платежных документов и т.д.).



# OCR-СИСТЕМА FINEREADER



Работа системы типа **FineReader** включает два крупных этапа:

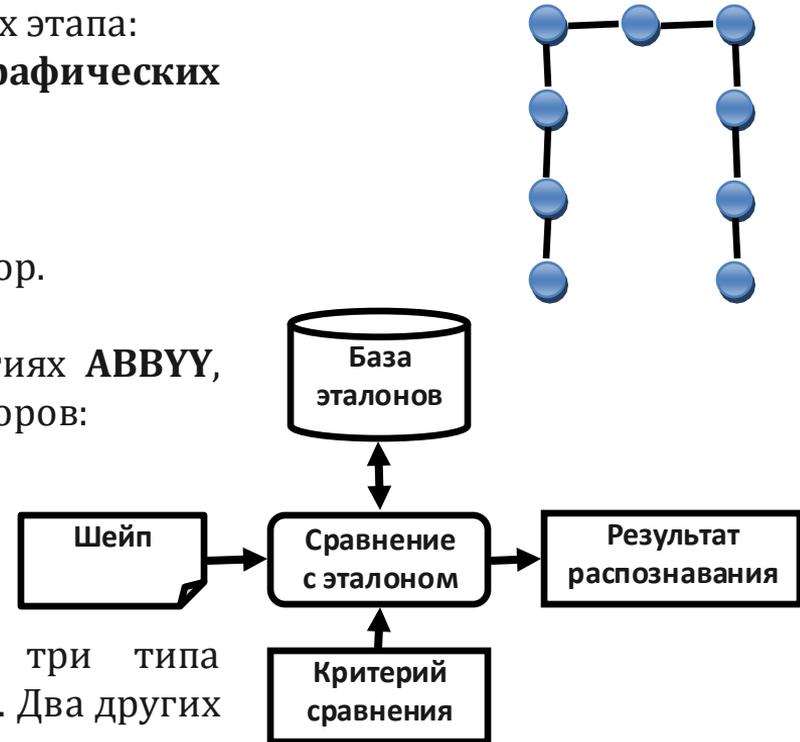
- ✓ **Анализ и предварительная обработка графических изображений;**
- ✓ **Распознавание отдельных символов.**

Системы распознавания реализуются как классификатор.

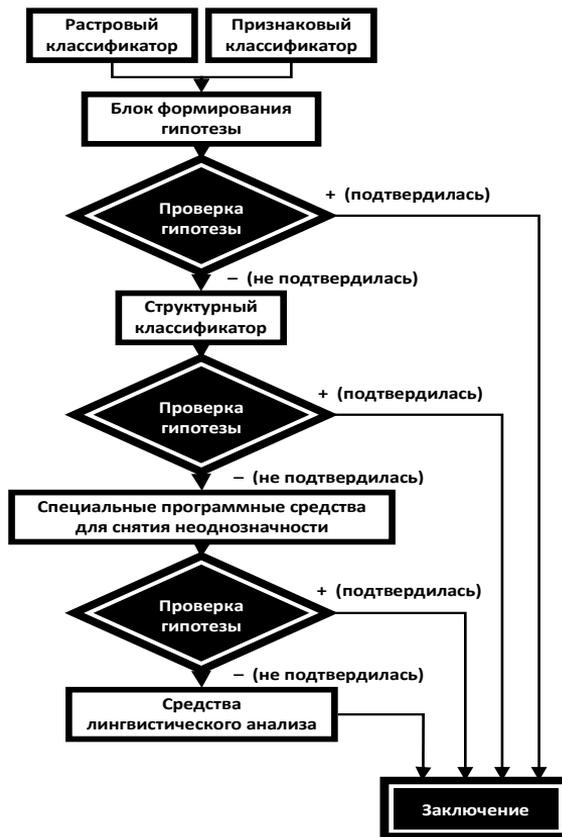
В системах распознавания, построенных на технологиях **ABBYY**, применяются следующие основные типы классификаторов:

- **растровые (шаблонные);**
- **признаковые;**
- **структурные.**

В современных OCR-системах используются все три типа классификаторов, но основным является структурный. Два других используются для ускорения и повышения качества распознавания.



# Укрупненная схема работы системы FineReader



# Особенности распознавания рукописных текстов



- Использование структурно-пятенного эталона с учетом особенностей траектории движения пишущего инструмента (выделяются кольца, дуги, точки, отрезки и другие топологические признаки).
- Основным механизмом является выдвигание и подтверждение гипотез.
- Использование методов оптимизации при управлении перебором вариантов.

# OCR-система Cognitive Forms



OCR-система **Cognitive Forms** представляет собой программный комплекс для массового ввода документов, имеющих стандартизированные формы.

Технология ввода документов в стандартизированных формах включает две стадии: **подготовительную, основную.**

**На первой стадии** создаются шаблоны документов, которые планируется вводить.

**Основная стадия** состоит из следующих этапов:

