

Системы оптического распознавания документов

Презентацию выполнила: Буйлова Татьяна
ученица 10 «А» класса

Проверила Пospelова Г. В.

МБОУ «СОШ №20»

г. Новомосковска



Содержание

Системы оптического распознавания символов

Растровый метод распознавания текста

Структурный метод распознавания

Системы оптического распознавания форм

Системы распознавания рукописного текста

Источники информации

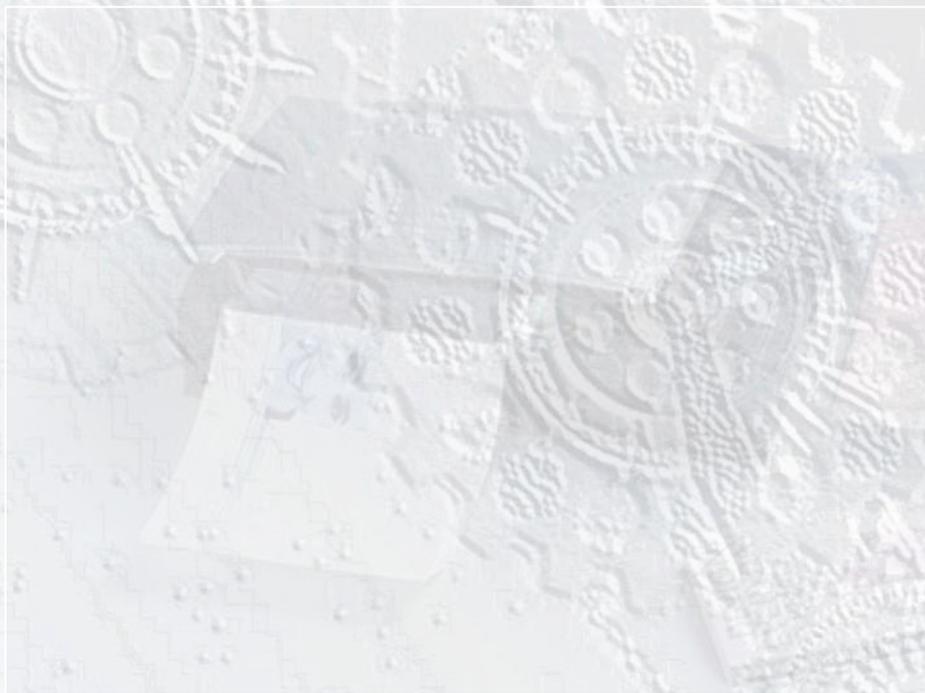
ВЫХОД

Системы оптического распознавания символов

Оптическое распознавание символов — механический или электронный перевод изображений рукописного, машинного или печатного текста в текстовые данные используются для представления символов в компьютере (например, в текстовом редакторе).



Системы оптического распознавания символов используются при создании электронных библиотек и архивов путем перевода книг и документов в цифровой компьютерный формат.



- с помощью сканера получают изображение страницы текста в графическом формате.
- проводят распознавание текста, т. е. преобразовывают элементы графического изображения в последовательность текстовых СИМВОЛОВ.



Схема распознавания символов



Растровый метод распознавания текста

Для отсканированных документов типографского качества (достаточно крупный шрифт, отсутствие плохо напечатанных символов или исправлений) распознавание символов проводится путем их сравнения с растровыми шаблонами.



Растровый метод распознавания текста

Растровое изображение каждого символа последовательно накладывается на растровые шаблоны символов, хранящиеся в памяти системы оптического распознавания. Результатом распознавания является символ, шаблон которого в наибольшей степени совпадает с изображением



Структурный метод распознавания

Векторный метод используется при распознавании документов с низким качеством печати.

В изображении символа выделяются геометрические примитивы и сравниваются с векторными шаблонами символов.

Выбирается символ, для которого совокупность всех геометрических примитивов и их расположение больше всего соответствует распознаваемому символу.



Бланк - стандартный лист бумаги, на котором размещается постоянная информация и отведено место для переменной.

Необходимо распознать написанные от руки символы, довольно сильно различающиеся у разных людей.



Системы оптического распознавания форм

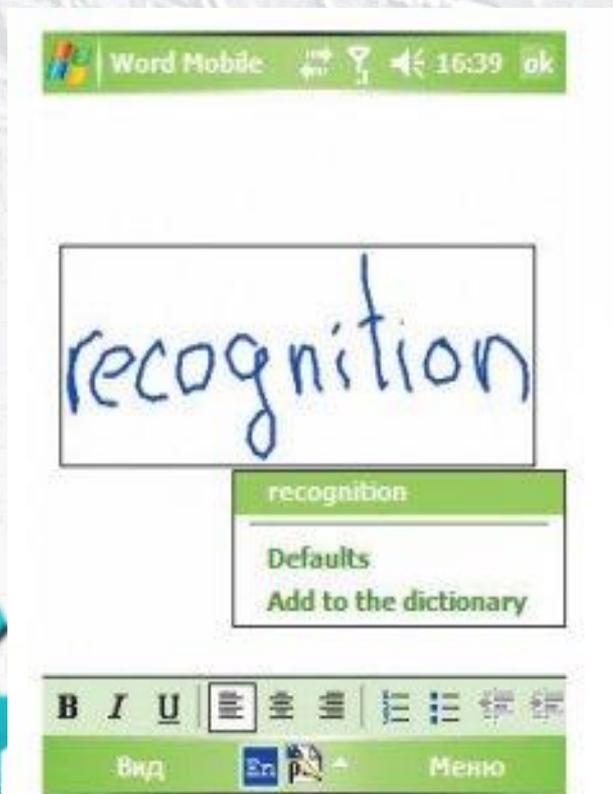
Для обработки бланков предназначено приложение *FineReader Forms*.

- создается шаблон формы.
- накладывается шаблон.
- корректируется положение шаблона.
- распознается содержимое бланка.



Системы распознавания рукописного текста

Системы распознавания рукописного текста преобразуют текст, написанный на экране карманного компьютера специальной ручкой, в текстовый компьютерный документ.



Источники информации

- Угринович Н. Д. Информатика и ИКТ. Базовый уровень: учебник для 10 класса / Н.Д. Угринович . 9-е издание : БИНОМ. Лаборатория знаний, 2012 - 213с.
- http://www.ci.ru/inform16_02/p_22text.htm
- http://dig-vid.blogspot.ru/p/blog-page_9552.html
- http://danilova.biz/index.php?option=com_content&task=view&id=329&Itemid=120

