**Лекция: *Сканеры. Сканирование текстовых и графических материалов. Распознавание сканированных текстов***

Сканер: устройство или программа, осуществляющие *сканирование*, то есть исследование объекта, наблюдение за ним или считывание его параметров.

Виды сканеров:

* Сканер изображений — устройство для считывания двумерного изображения и представления его в растровой электронной форме
* 3D-сканер — устройство для считывания формы объёмного объекта
* Биометрические сканеры используются для целей идентификации личности
* Устройства автоматизированного считывания служебной информации:
	+ Сканер штрихкода — устройство для считывания информации, представленной в виде штрих-кода
	+ Считыватель RFID-меток

В настоящее время для ввода текстов и графической информации используют следующие типы сканеров:

Планшетные сканеры: наилучшее качество и максимальное удобство при работе. Под крышкой планшета располагается прозрачное основание, на которое кладется документ. Блок сканирования перемещается вдоль документа внутри корпуса сканера. Продолжительность сканирования листа от одной до нескольких секунд.

Ручные сканеры: наиболее простой вид, дающий наименее качественное изображение. Сканирование производится путем перемещения сканера по документу вручную. Пригодны для сканирования текстов.

Барабанные сканеры: очень высокое разрешение сканирования слайдов, негативов**и т.п.**Сканирующее устройство закреплено неподвижно, а изображение закрепляется на барабане и вращается при сканировании.

Листовые сканеры: позволяют за одну операцию сканировать стандартный лист бумаги (A4). Бумага протягивается мимо сканирующего блока при помощи валиков (как в принтере). Хорошее качество сканирования, но невозможность сканирования книжной страницы и журнала.

Сканеры форм: для ввода информации с заполненных бланков (разновидность листовых). Очень высокое быстродействие (для анкет, опросных листов, избирательных бюллетеней).

Штрих-сканеры: разновидность ручных. Для считывания штрих-кодов с маркировки товаров в магазинах. Позволяют автоматизировать процесс подсчета стоимости покупок.

**Технические характеристики сканеров**

Для того чтобы правильно выбрать тип сканера, а соответственно максимально использовать его возможности для реализации поставленных задач, необходимо хорошо разбираться в технических характеристиках сканеров и существующей терминологии.

*Разрешающая способность (разрешение).*

Максимальное разрешение сканера говорит о степени детализации при сканировании. Чем больше разрешение, тем больше информации об оригинале может быть введено в компьютер и подвергнуто дальнейшей обработке.

*Производительность*

Количество времени, затраченное на предварительный просмотр и собственно сканирование, может значительно влиять на производительность. Скорость сканирования зависит от таких факторов, как разрешение, размер и тип изображения (например, черно-белая или цветная фотография или же текст).

*Глубина цвета*

Глубина цвета – это характеристика, показывающая, насколько точна информация о цвете каждой точки отсканированного изображения, и отражающая разрядность аналого-цифрового преобразователя.

Сканеры подключаются к ПК

* С параллельным или последовательным интерфейсом
* С интерфейсом USB

С интерфейсом Fire Wire

**Программное обеспечение сканера**

Программное обеспечение сканера состоит из двух частей: программного интерфейса и пакета прикладных графических программ. Программный интерфейс обеспечивает управление сканером, а также его связь с графическими программами сторонних производителей. Это так называемый TWAIN-модуль или драйвер сканера. Стандарт TWAIN поддерживают практически все графические программы.

TWAIN-модуль имеет пользовательский интерфейс (диалоговое окно), с помощью которого можно настроить параметры сканирования. Внешний вид и состав параметров этого модуля могут быть различными, поскольку производители программного обеспечения сканера ограничены только собственно стандартом TWAIN, а совершенствовать пользовательский интерфейс им никто не мешает. Вместе с тем, существует стандартный набор параметров, которые присутствуют во всех интерфейсах: выбор режима и области сканирования, разрешения, контрастности, яркости и т. д.

Кроме TWAIN-модуля в программное обеспечение сканера обычно входит какой-нибудь, обычно весьма скромный по возможностям, графический редактор и, возможно, программа оптического распознавания символов (OCR).

**Сканирование текстовых и графических материалов. Распознавание сканированных текстов**

Цели и задачи распознавания документов многообразны, как и сами документы. Их решение строится на основе анализа изображения документа, полученного со сканера, видеокамеры или другого сканирующего устройства. Процесс ввода документа при этом называется *оптическим вводом.* После такого ввода документ представляет собой совокупность черных и белых (либо цветных) точек (растров) - графическое изображение.Поэтому общая задача распознавания графического изображения - преобразование графической формы в другую, воспринимаемую человеком и программой. Эта задача часто может быть разбита на подзадачи распознавания компонент документа:

•Фотографии, рисунки и диаграммы часто сохраняют в графическом (растровом) формате.

•Контурное изображение (печати, рамки таблиц и т.д.) удобно хранить в векторном, а не растровом виде.

•Распознавание текста - это перевод его из графической формы в символьную. Ведется распознавание и рукописного текста, включающее задачи психологического характера (медицина, психология, судмедэкспертиза).

•Анализ структуры и анализ реквизитов документов проводится для подтверждения подлинности документа (ценные бумаги и т.п.) и распознавания его компонент.

Хранение текстовых документов в растровом формате нецелесообразно,т.к. растровая графика занимает много дисковой памяти и не подлежит редактированию в редакторах текста. Создание электронной формы документа предполагает:

-детальный анализ структуры (сегментирование и анализ блоков: текст, рисунок, таблица);

-распознавание текста;

-извлечение из него графических компонент;

-компоновку в среде текстового процессора или издательской системы.

**Программы распознавания текстов**

В настоящее время лидерами программ распознавания текста и создания электронного документа являются программы CuneiForm и FineReader.

Пакет FineReader предназначен для распознавания текстов на русском, английском, немецком, французском и других языках, а также двуязычных текстов. Позволяет объединять сканирование и распознавание в одну операцию, работать с пакетами документов и бланками. Программу можно обучать для повышения качества распознавания.

**Сканирование и корректировка текстового документа**

1 Запустить программу Fine Reader

2 Сканировать текстовый документ:

2.1 Поместить документ в сканер, подключить сканер к сети

2.2 Выполнить сканирование, щелкнув по пиктограмме, изображающей сканер на панели инструментов. В образовавшемся окне сканировать текстовый документ:

2.2.1 В окне программы ScanExpress 6000 P выбрать параметрысканирования и щелкнутьno кнопке "PreScan", выполняя предварительное сканирование;

2.2.2 Выделить область сканирования рамкой и сканировать выделенный участок, щелкнув по кнопке "Sсan";

2.2.3 Выйти из окна сканирования, щелкнув по кнопке "Exit" или по кнопке «Закрыть»

3 Произвести преобразование бумажного документа в электронный

Выполнить распознавание текста, щелкнув по кнопке "Распознать". Автоматически распознанный текст появляется в специальном окне. Он теряет связь с исходным изображением и может редактироваться и форматироваться независимо от него.

4 Полученный документ необходимо сохранить, использовав меню ФАЙЛ -СОХРАНИТЬ КАК... Сохраненный документ может быть вызван для редактирования в текстовом процессоре WORD или в табличном процессоре EXCEL

**Сканирование графического изображения**

Включить сканер и запустить программу графического редактора Adobe PhotoShopОткрыть меню ФАЙЛ->ИМПОРТ - > TWAIN-32, выполнить сканирование графического изображения:

В окне программы ScanExpress выбрать параметры сканирования и щелкнуть по кнопке"PreScan". выполняя предварительное сканирование:

Выделить область сканирования и сканировать выделенный участок, щелкнув по кнопке"Sсan";

Выйти из окна сканирования, щелкнув по кнопке "Exit" или по кнопке «Закрыть»;

В открывшемся окне графического редактора Adobe PhotoShop произвести редактирование документа

Кадрировать изображение: щелкнуть по кнопке "Рамка" на панели инструментов, растянуть выделениена нужный участок изображения и выбрать в меню ИЗОБРАЖЕНИЕ -> ОБРЕЗАНИЕ.

Откорректировать цветовую гамму изображения через меню ИЗОБРАЖЕНИЕ-КОРРЕКЦИЯ - ЯРКОСТЬ/КОНТРАСТНОСТЬ, перетащив указатели в нужном направлении и закрыть окно.

Полученное изображения можно сохранить, используя меню ФАЙЛ->СОХРАНИТЬ КАК…или щелкнуть по кнопке "Сохранить" для сохранения в стандартном формате с порядковым номером файла.

Полученное изображение можно вывести на принтер, используя меню ФАЙЛ-ПЕЧАТЬ