

УВАЖАЕМЫЕ СТУДЕНТЫ! Законспектируйте в своей рабочей тетради по дисциплине приведенную лекцию (объемом 4-5 страницы), ответьте письменно на контрольные вопросы.

Результаты работы, фотоотчет, предоставить преподавателю на e-mail: igor-gricenko-95@mail.ru **в течении ТРЕХ дней.**

При возникновении вопросов по приведенному материалу обращаться по следующему номеру телефона: (072)132-63-42

ВНИМАНИЕ!!! При отправке работы, не забывайте указывать **ФИО студента, наименование дисциплины, дата проведения занятия (по расписанию).**

Лекция 8.1 Знакомство с работой сканера. Устройство и принцип работы сканера. ABBYY FineReader. Настройка интерфейса, режимов сканирования, сохранение во внешнее приложение

Цель: изучить устройство и принцип работы сканера, интерфейс и порядок работы с программой ABBYY FineReader.

План

1. Знакомство с работой сканера. Устройство и принцип работы сканера.
2. ABBYY FineReader. Настройка интерфейса, режимов сканирования, сохранение во внешнее приложение.

Сканер - устройство, которое создает цифровое изображение сканируемого объекта. Полученное изображение может быть сохранено как графический файл, или если оригинал содержит текст, распознано посредством программы распознавания текста и сохранено как текстовый файл. Изобретателем сканера считается Рудольф Хелл. В конце 20-х годов он создал сканер - факс, который «научил» прототип современного сканер - факса переводить текст и изображения в точки и линии и обратно.

Виды сканеров:

- Ручные сканеры обрабатывают полосы документа шириной около 10 см и представляют интерес прежде всего для владельцев мобильных ПК. Они медлительны, имеют низкие оптические разрешения, часто обуславливают перекосы отсканированного изображения, но в то же время недороги и компактны.
- Листопротяжные: принцип как в факсимильном аппарате, страницы документа при считывании пропускаются через специальную щель с помощью направляющих роликов. Возможности ограничены.

- Планшетные сканеры напоминают верхнюю часть копировального аппарата. Принцип действия рассмотрим ниже.
- Барабанные сканеры по светочувствительности значительно превосходящие потребительские планшетные устройства, применяются исключительно в полиграфии. В барабанных сканерах оригиналы размещаются на внутренней или внешней стороне прозрачного цилиндра, который называется барабаном. Чем он больше, тем больше площадь поверхности, на которую монтируется экран, тем больше площадь сканирования.

Принцип действия планшетных сканеров. Современный планшетный сканер функционально состоит из трех частей:

- собственно сканирующего механизма,
- аппаратных средств, расположенных в самом сканере и на плате (карте) в компьютере,
- программной части (TWAIN-модуль, программа управления цветом, градацией и т.д.).

Сканируемое изображение (оригинал) располагается на прозрачном неподвижном стекле, вдоль которого, в корпусе сканера, передвигается сканирующая головка с источником света, оптикой и фотоприемниками. На сканирующей головке располагаются также электронные элементы аналогового тракта и аналого-цифровой преобразователь (АЦП). Многие планшетные сканеры позволяют сканировать непрозрачные и прозрачные оригиналы (слайды), используя один сканирующий механизм. Переход к режиму сканирования слайда возможен при наличии в сканере слайд-модуля. Слайд-модуль располагается поверх стекла сканера, на месте прижимной крышки для непрозрачных оригиналов, и содержит дополнительный источник света для сканирования изображения на просвет.

Оптическая система сканера, состоящая из объектива и зеркал, проецирует световой поток от сканируемого оригинала на трехстрочную матрицу фотоэлектрических преобразователей, например, фотодиодов. Каждая строка матрицы имеет свой спектральный фильтр, установленный при ее изготовлении, а сама матрица оформлена в виде интегральной схемы с прозрачным окном на поверхности корпуса. Проходя через спектральные фильтры световой поток разделяется на базовые цветовые компоненты.

В результате преобразования света получается электрический сигнал, содержащий информацию об активности цвета в исходной точке сканируемого изображения. После оцифровки аналогового сигнала в АЦП цифровой сигнал через аппаратный интерфейс сканера идет в компьютер, где его получает и анализирует программа для работы со сканером. После

окончания одного такого цикла (освещение оригинала — получение сигнала — преобразование сигнала — получение его программой) источник света и приемник светового отражения перемещается относительно оригинала. Основной деталью планшетного сканера является считывающая головка,двигающаяся вдоль сканируемого изображения. Важнейшей частью считывающей головки является фотоприемник. На сегодняшний день наиболее распространены два типа фотопринимающей матрицы: ПЗС-матрица (прибор с зарядовой связью, в английских обозначениях — CCD, Couple-Charged Device) и КДИ-матрица (контактный датчик изображения, в английских обозначениях — CIS, Contact Image Sensor). На качество изображения получаемого в результате сканирования влияет источник света. Используются следующие источники света:

- ксеноновые лампы отличаются малым временем прогрева, высокой стабильностью излучения;

- люминесцентные лампы:

- с горячим катодом - малое время прогрева, но короткий срок службы.

- с холодным катодом - служат в 10 раз дольше, время прогрева большое, имеют низкую рабочую температуру.

- светодиоды - применяются сканером, имеют маленькое время прогрева, небольшие габариты. У сканеров с таким источником света страдает качество цветопередачи.

Сканеры также делятся на:

- Сканеры изображений - привычное нам устройство;

- 3D-сканер - устройство анализирующее физический объект на основе полученных данных создающее его 3D-модель.

- Биометрический сканер – используется для целей идентификации личности.

- Сканер штрих-кода – устройство для считывания информации, представленной в виде штрих-кода.

ABBYY FineReader – это система оптического распознавания текстов (Optical Character Recognition, OCR). Она предназначена для того, чтобы быстро перевести отсканированные бумажные документы, а также PDF-файлы в удобный для дальнейшего использования электронный формат – файл Word или Excel, HTML-страничку, презентацию PowerPoint или в любой другой из поддерживаемых форматов, при этом полностью сохранив оформление документа. FineReader удовлетворяет самым высоким требованиям профессиональных пользователей благодаря большому количеству различных опций и настроек.

Процесс ввода документов в компьютер состоит из четырех этапов: сканирование, распознавание, проверка и сохранение результатов распознавания. В результате сканирования появится окно Изображение, содержащее «фотографию» страницы. Затем программа попросит Вас установить параметры распознавания и приступит к распознаванию изображения, одновременно анализируя его. Обработанные участки изображения закрашиваются голубым цветом.

Результат распознавания Вы увидите в окне Текст. В этом же окне Вы можете проверить и отредактировать распознанный текст. Следуя далее указаниям Мастера Scan&Read, Вы можете либо передать распознанный текст в выбранное Вами приложение или сохранить его на диск, либо продолжить обработку следующих изображений.

Работая с программой FineReader, пользователь всегда имеет дело с некоторым пакетом.

Пакет – это папка, в которой хранятся изображения и рабочие файлы программы. Каждое отсканированное изображение записывается как отдельная страница пакета. При запуске программы FineReader по умолчанию открывается новый пакет. Вы можете продолжить работу с новым пакетом или же открыть уже существующий пакет. Вы можете продолжить работу с пакетом по умолчанию или создать новый пакет. Более подробно о работе с пакетом см. Общие сведения по работе с пакетом.

Вверху Главного окна FineReader находится меню системы, под ним – инструментальные панели. В программе их четыре: Стандартная, Форматирование, Изображение и Scan&Read.

Спрятать или показать инструментальные панели на экране можно через меню Вид (пункт Панели инструментов) или через локальное меню. Чтобы открыть локальное меню, щелкните правой кнопкой мыши на одной из инструментальных панелей. Панели, которые видны на экране, отмечены галочкой. Выберите панель, которую Вы хотите спрятать или, наоборот, показать.

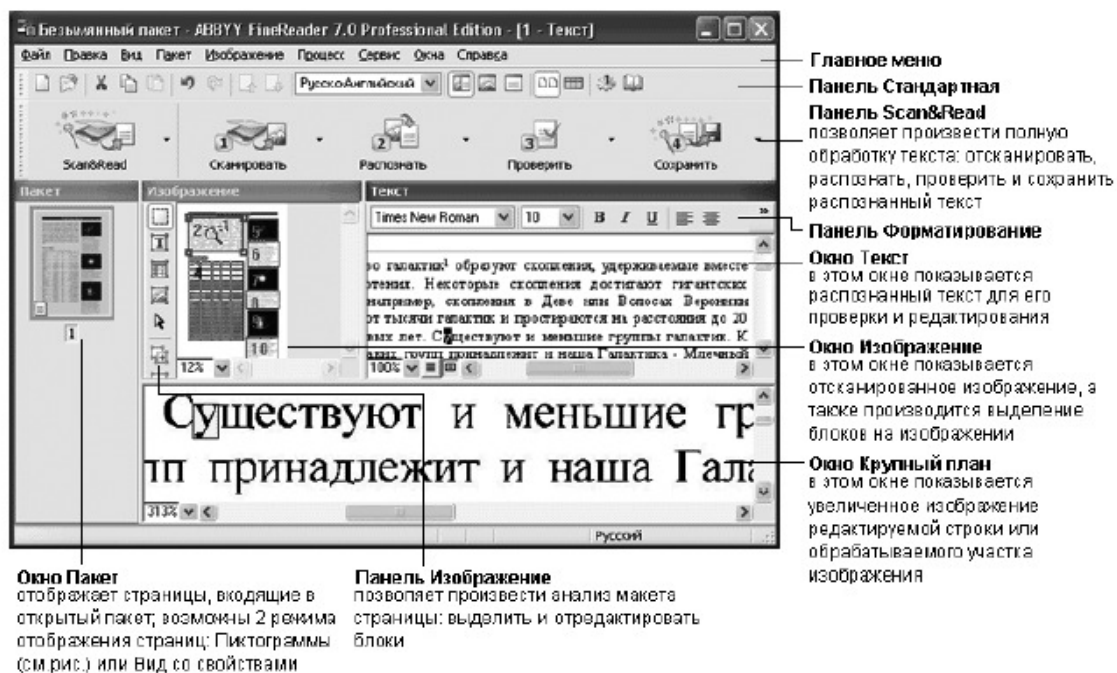


Рисунок 1 - Главное окно программы FineReader

Чтобы выбрать вид окна Пакет:

- Нажмите на соответствующую кнопку (/) на панели инструментов Стандартная, или Нажмите правой кнопкой мыши на окне Пакет и выберите в локальном меню пункт Вид...

- Внизу окна расположена информационная панель (еще ее называют строкой состояния), где Вы можете получить сведения о состоянии системы и производимых ею операциях, а также краткую справку о выбираемых Вами пунктах меню и кнопках. Остальное пространство Главного окна занимают по мере своего появления рабочие окна системы: Пакет, Изображение, Крупный план и Текст.

Окна Изображение, Крупный план и Текст связаны между собой: при двойном щелчке на изображении в окне Изображение курсор в окнах Крупный план и Текст (при наличии распознанного текста) переместится на ту же позицию, что и в окне Изображение.

Используется при редактировании большого количества страниц пакета.

Окно Пакет сверху; тип отображения страниц пакета: Вид со свойствами; Окна Текст и Крупный план.

Используется при большом количестве страниц пакета.

Окно Пакет сверху; тип отображения страниц пакета:

Вид со свойствами; Окна Изображение, Текст и Крупный план.

Используется при малом количестве страниц пакета.

Окно Пакет слева; тип отображения страниц пакета:

Пиктограммы; Окна Изображение, Текст и Крупный план Применение:
Вариант расположения окон:

Возможности программы ABBYY FineReader

FineReader позволяет ввести документ одним нажатием на кнопку Scan&Read, не вдаваясь в подробности работы программы. Распознанный текст можно передать в текстовый редактор или электронную таблицу, сохранить в форматах RTF, DOC, Word XML (при сохранении в Microsoft Office 2003), PDF и HTML с полным сохранением оформления документа или сохранить в базе данных.

Форматы, в которых FineReader сохраняет распознанный текст
FineReader сохраняет результаты распознавания в следующих форматах:

- Microsoft Word Document (*.DOC)
- Rich Text Format (*.RTF)
- Microsoft Word XML Document (*.XML) (только для Microsoft Word 2003)
- Adobe Acrobat Format (*.PDF)
- HTML
- Microsoft PowerPoint Format (*.PPT)
- Comma Separated Values File (*.CSV)
- Простой текст (*.TXT). FineReader поддерживает различные кодовые страницы (Windows, DOS, Mac, ISO) и кодировку Unicode.
- Microsoft Excel Spreadsheet (*.XLS)
- DBF.

Задача распознавания состоит в том, чтобы преобразовать отсканированное изображение в текст, сохранив при этом оформление страницы. Прежде чем приступить к распознаванию текста, необходимо установить основные параметры распознавания: язык распознавания, тип печати распознанного текста и тип страницы. В этой главе описываются эти и другие параметры распознавания и приводятся ситуации, в которых они используются.

Перед запуском распознавания проверьте установленные опции: язык распознавания, тип печати распознаваемого текста и тип страницы.

Вы можете:

- Распознать блок или несколько блоков, выделенных на изображении.
- Распознать открытую страницу или все страницы, выделенные в окне Пакет.
- Распознать все нераспознанные страницы пакета.

- Распознать все страницы в фоновом режиме. В этом режиме возможно распознавание с одновременным редактированием уже распознанных страниц.

- Распознать страницы в режиме распознавание с обучением. Данный режим применяется в основном для распознавания текстов, использующих декоративные шрифты, или для распознавания большого объема (более 100 страниц) документов плохого качества печати.

- Распознать страницы одного пакета на нескольких компьютерах одновременно.

Чтобы запустить распознавание:

- Нажмите кнопку 2–Распознать на панели Scan&Read.

- В меню Процесс выберите нужный Вам пункт:

Распознать – чтобы распознать открытую страницу или все страницы, выделенные в окне Пакет;

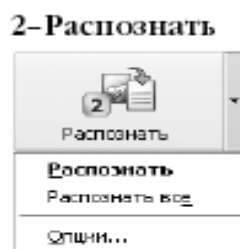


Рисунок 2 – Окно распознавания

Распознать все – чтобы распознать все нераспознанные страницы пакета;

Распознать Блок – чтобы распознать блок или несколько блоков, выделенные на изображении;

Запустить фоновое распознавание – чтобы запустить распознавание в фоновом режиме.

Кнопка 2–Распознать запускает распознавание открытого изображения. Чтобы изменить режим кнопки, нажмите на стрелку справа от нее и из открывшегося меню выберите нужный пункт.

Результаты распознавания можно сохранить в файл, передать во внешнее приложение, не сохраняя на диск, скопировать в буфер обмена или отправить по электронной почте в любом из поддерживаемых программой FineReader форматов сохранения. Сохранить можно все страницы или только выбранные.

Сохранение распознанного текста:

- Сохранить распознанный текст, используя Мастер сохранения результатов.

- Сохранить открытую или выделенные в окне Пакет страницы в файл или во внешнее приложение.

- Сохранить все страницы пакета в файл или во внешнее приложение.

- Сохранить изображение страницы.

Кнопка 4–Сохранить позволяет передать результаты распознавания в выбранное приложение или сохранить их в файл. Внешний вид значка меняется в зависимости от выбранного режима сохранения; подпись Сохранить меняется на название выбранного приложения.

Чтобы сохранить распознанный текст:

- Нажмите стрелку справа от кнопки 4–Сохранить и в локальном меню выберите необходимый пункт.

При сохранении нескольких страниц сначала выделите их в окне Пакет.

После того как Вы экспортировали распознанный текст в выбранное Вами приложение, отправили его по электронной почте, передали в буфер или сохранили в файл, “информация” об этом действии отразится на значке кнопки 4–Сохранить. Поэтому для того, чтобы повторить ту же операцию для другого изображения, Вам достаточно нажать на этот значок.

Контрольные вопросы

1. Что такое сканер, его назначение?
2. Перечислите виды сканеров.
3. Перечислите назначение и возможности программы ABBYY FineReader.