

Лабораторная работа 3.

ВОССТАНОВЛЕНИЕ ДОКУМЕНТА ПО СОХРАНИВШИМСЯ ФРАГМЕНТАМ

Цель работы:

- *изучить технические характеристики сканеров;*
- *освоить программу FineReader, получить навыки сканирования и распознавания документов в различных форматах.*

Задание на выполнение лабораторной работы

1. Изучить техническое описание сканера.
2. Запустить программу FineReader 7.0. В центральном окне нажать кнопку «Обучение на примерах». В окне «Справка» нажать строку «Примеры» (левая колонка). Изучить сканирование и распознавание следующих типов документов:
 - 1) Простое письмо
 - 2) Документ - текст в одну колонку
 - 3) Книжный разворот
 - 4) Факсы и другие документы с низким качеством печати
 - 5) Сложная журнальная страница
 - 6) Газетная страница
 - 7) PDF документ
 - 8) Работа с таблицами
 - a. Простая таблица
 - b. Таблица с неполным количеством черных разделителей
 - c. Сложная таблица
3. Восстановить документ – стандарт ГОСТ 6.10.3–83 по фрагментам, причем: 1-й фрагмент – изображение титульных страниц в формате tif, файл **titl.tif**; 2-й фрагмент – страницы 1 и 2 распечатаны на бумажном носителе; 3-й фрагмент – страницы 3...8 в формате pdf, файл **3_8.pdf**.
4. Распознавание первого фрагмента.
 - 4.1. В программе FineReader загрузить изображение титульных страниц: **Файл=>Открыть изображение**. В открывшемся локальном окне открытия изображения выбрать файл titl.tif.

4.2. На панели «Пакет» отметить первую страницу (щелкнуть левой клавишей мыши на изображении первой страницы). Распознать выделенную страницу – на панели инструментов клавиша «**2 Распознать**».

4.3. На панели «Изображение» проверить выделенные поля. Текстовые фрагменты выделяются зеленой рамкой, табличные – синей, графические изображения – красной. Ненужные поля: печати, подписи, помарки – удалить. Если какие-то фрагменты текста, таблиц остались вне выделенных полей, их нужно либо ввести в имеющиеся поля путем расширения и выравнивания границ поля, либо прочертить новое поле, предварительно выбрав тип поля на левой колонке панели «Изображение» («Выделить текстовый блок», «Выделить табличный блок», «Выделить картинку»). Если границы полей корректировались, повторно запустить процесс распознавания.

4.4. На панели «Текст» проверить распознанный текст. Зоны неуверенного распознавания отмечены голубым фоном. Исправить выявленные ошибки.

4.5. Аналогично распознать вторую страницу.

4.6. Сохранить текст в формате Microsoft Word: **Файл=>Сохранить текст как**. В открывшемся локальном окне выбрать папку, где будем хранить файл, **Тип файла** (Документ Word (*.doc)), **Имя файла** и щелкнуть **Сохранить**.

4.7. Открыть сохраненный файл в программе Microsoft Word.

4.8. В программе FineReader закрыть пакет: **Файл=>Закреть пакет**. В открывшемся локальном окне «*Вы хотите сохранить изменения в пакете?*» выбрать опцию **Нет**.

5. Распознавание второго фрагмента.

5.1. На сервере запустить FineReader. Открыть сканер и положить на сканируемую поверхность первый лист лицевой стороной вниз. Отсканировать первый лист: **Файл=>Сканировать изображение** (или треугольник на клавише **Открыть =>Сканировать изображение**).

5.2. Отсканировать второй лист.

5.3. Распознать и проверить тексты.

5.4. Сохранить текст в формате Word.

5.5. Открыть папку с сохраненным файлом для доступа по сети: на имени папки нажать правую клавишу мыши до появления всплывающего меню, если Вы работаете с TotalComander, или просто щелкнуть правой клавишей, если Вы работаете с Проводником. В появившемся меню выбрать опцию «Общий доступ и безопасность...» или опцию «Свойства». В

появившемся локальном окне открыть панель «Доступ» и отметить галочкой «Открыть общий доступ к этой папке», щелкнуть «Применить» и «ОК».

5.6. На своем рабочем месте по сети открыть нужную папку на сервере (сетевое имя сервера – Server DVU) и скачать файл с текстом на свой компьютер в папку, где был сохранен первый фрагмент документа.

5.7. Открыть файл с текстом второго фрагмента в программе Microsoft Word. Скомпоновать фрагменты в один.

6. Распознавание третьего фрагмента.

6.1. Открыть файл с третьим фрагментом: клавиша **Открыть**. В открывшемся локальном окне выбрать файл **3_8.pdf**.

6.2. Распознать и проверить все страницы фрагмента.

6.3. Передать полученный текст в программу Word: **Файл=>Передать все страницы в=>Microsoft Word**.

6.4. Скомпоновать текст третьего фрагмента с первыми.

6.5. Проверить документ в целом.

Задание для домашней работы

1. Изучить основные типы сканеров и технологию сканирования документов.
2. Подготовить ответы на контрольные вопросы.

Контрольные вопросы

1. В чем специфика распознавания различных видов документов?
2. Какие основные типы сканеров Вы знаете?
3. Какие программные средства позволяют преобразовать документ в pdf формат?

Литература.

1. Бройдо В.Л. Офисная оргтехника для делопроизводства и управления. – М., 1998.
2. Степаненко О.С. Сканеры и сканирование. – М.: Издательский дом «Вильямс», 2005. – 288 с.