

Система оптического распознавания текста

ABBYY® FineReader®

Версия 8.0

Руководство пользователя

Информация, содержащаяся в этом документе, может быть изменена без предварительного уведомления, и компания ABBYY не берет на себя на этот счет никаких обязательств.

Программное обеспечение, описываемое в этом документе, поставляется в соответствии с Лицензионным договором. Это программное обеспечение может быть использовано или скопировано лишь в строгом соответствии с условиями этого договора. Копирование этого программного обеспечения на какой-либо носитель информации, если на это нет специального разрешения в Лицензионном договоре или в соглашении о нераспространении, является нарушением Закона Российской Федерации «О правовой охране программ для ЭВМ и баз данных» и норм международного права.

Никакая часть настоящего Руководства ни в каких целях не может быть воспроизведена или передана в какой бы то ни было форме и какими бы, то ни было средствами, будь то электронные или механические, включая фотокопирование и запись на магнитные носители, если на то нет письменного разрешения компании ABBYY.

© 2005 ABBYY Software. Все права защищены.

© 1987–2003 Adobe Systems Incorporated. Право на использование Adobe® PDF Library предоставлено компанией Adobe Systems Incorporated.

Microsoft Reader Content Software Development Kit © 2004 Microsoft Corporation, One Microsoft Way, Redmond, Washington 98052–6399 U.S.A. Все права защищены.

Шрифты Newton, Pragmatica, Courier © 2001 ParaType, Inc. Шрифт OCR–в–ГОСТ © 2003 ParaType, Inc.

© 1999–2000 Image Power, Inc. и University of British Columbia, Canada © 2001–2002 Michael David Adams. Все права защищены.

ABBYY, логотип ABBYY, Scan&Read, ABBYY FineReader являются зарегистрированными товарными знаками или товарными знаками компании ABBYY Software Ltd.

Adobe, логотип Adobe, логотип Adobe PDF и Adobe PDF Library являются зарегистрированными товарными знаками или товарными знаками компании Adobe Systems Incorporated в Соединенных Штатах Америки и/или других странах.

Microsoft, Outlook, Excel, PowerPoint, Windows являются зарегистрированными товарными знаками или товарными знаками компании Microsoft Corporation в Соединенных Штатах Америки и/или других странах.

Unicode является товарным знаком компании Unicode, Inc.

Остальные товарные знаки являются товарными знаками или зарегистрированными товарными знаками своих законных владельцев.

Содержание

Добро пожаловать!	4
Новые возможности ABBYY FineReader 8.0.....	5
Глава 1	
Работа с программой ABBYY FineReader 8.0.....	8
Установка и запуск ABBYY FineReader 8.0.....	9
Получение и обработка изображения программой ABBYY FineReader.....	11
Анализ макета страницы.....	26
Распознавание.....	34
Проверка и редактирование текста.....	44
Сохранение во внешние редакторы и форматы.....	53
Работа с пакетами.....	69
Сценарии ABBYY FineReader	73
Глава 2	
Приложение ABBYY Screenshot Reader	84
Установка и запуск ABBYY Screenshot Reader	85
Панель ABBYY Screenshot Reader.....	85
Получение текста и таблиц с экрана компьютера	86
Получение снимка экрана	88
Дополнительные настройки приложения ABBYY Screenshot Reader	88
Глава 3	
Приложение ABBYY Hot Folder & Scheduling	90
Установка и запуск ABBYY Hot Folder & Scheduling.....	91
Главное окно ABBYY Hot Folder & Scheduling	91
Настройка задачи для папки	93
Отчет об обработке папки	96
Дополнительные настройки приложения ABBYY Hot Folder & Scheduling	96
Приложение	
Форматы, в которых ABBYY FineReader сохраняет распознанный текст	99
Поддерживаемые графические форматы	100
Горячие клавиши	101
Глоссарий.....	103

Добро пожаловать!

Благодарим за приобретение программы ABBYY FineReader!

В нашей жизни становится все больше электронных документов. Тем не менее, деловые бумаги, журналы, книги по-прежнему в ходу. Миллионы людей разных профессий используют ABBYY FineReader для перевода бумажных документов в электронный вид. Ведь в современном стремительном мире успеха добивается тот, кто умеет эффективно использовать свое время и может управлять информацией независимо от того, в каком виде она представлена. Система оптического распознавания ABBYY FineReader быстро переведет бумажные документы или PDF-файлы в удобный для редактирования формат.

ABBYY FineReader — незаменимый помощник, когда необходимо:

- быстро сделать подборку информации, подготовить сводный отчет;
- отредактировать бумажный документ или факс;
- написать реферат, статью, курсовую работу, диссертацию;
- опубликовать в интернете информацию из прессы или книг;
- извлечь информацию из PDF-файла и внести нужные изменения; и т.д.

С ABBYY FineReader очень легко работать. Даже новичок сможет получить результат за несколько минут. В то же время ABBYY FineReader отвечает самым высоким требованиям опытных пользователей; дает доступ ко множеству различных настроек. Это руководство поможет вам быстро «научить компьютер читать» и освоить все возможности системы.

К делу!

Новые возможности ABBYY FineReader 8.0

К восьмой версии в систему ABBYY FineReader внесен ряд усовершенствований. Кроме того, появились новые возможности для работы с бумажными документами, файлами изображений, PDF-файлами, а также с изображениями, полученными по факсу.

Улучшения технологии распознавания, появившиеся в новой версии, обеспечивают высокое качество распознавания документов, полученных с помощью цифровой фотокамеры, более точное распознавание факсов и бумажных документов, отсканированных с низким разрешением, а также документов сложного оформления. Добавлена возможность защиты PDF-документов с помощью паролей. Теперь в комплект поставки ABBYY FineReader входит также приложение-планировщик, приложение для распознавания снимков экрана, а также менеджер сценариев, позволяющий автоматизировать выполнение однотипных операций.

Ниже приведено описание основных новых возможностей и улучшений. Возможности, доступные только для версии ABBYY FineReader 8.0 Corporate Edition, отмечены особо.

Качество распознавания документов, отсканированных с низким разрешением, а также документов, полученных по факсу, повышено на 30%*

Система ABBYY FineReader традиционно обеспечивает высокую точность распознавания документов хорошего качества. Однако, иногда пользователю бывает необходимо распознать изображения низкого качества. Это могут быть документы, отсканированные с недостаточно высоким разрешением или пришедшие по факсу. (Изображения, получаемые по факсу, имеют разрешение ниже, чем рекомендовано для OCR-систем.) С помощью ABBYY FineReader 8.0 можно распознать и такие документы, причем по сравнению с предыдущей версией качество их распознавания улучшено на 30%.

Распознавание изображений, полученных с помощью цифровой фотокамеры

Если под рукой нет сканера, можно сделать снимок документа с помощью фотоаппарата, а затем распознать полученное изображение. Новая адаптивная технология распознавания, появившаяся в системе ABBYY FineReader 8.0, обеспечивает высокое качество обработки фотографий.

Защита PDF-файлов

В новой версии ABBYY FineReader появилась возможность защиты PDF-файлов паролем. При сохранении распознанного текста в формате PDF можно задать пароль для открытия документа, пароль для изменения прав доступа, а также выбрать уровень шифрования. В программе предусмотрены три уровня шифрования PDF-файлов: 40-битный и 128-битный на основе стандарта RC4, а также новейший 128-битный уровень, основанный на стандарте AES (Advanced Encryption Standard).

Добавление тегов в PDF-документы

Теперь ABBYY FineReader 8.0 позволяет разметить структуру документа при сохранении в формате PDF. Это обеспечивает удобство просмотра документа на экранах разного размера.

* По сравнению с ABBYY FineReader 7.0; по результатам собственных исследований.

Менеджер сценариев

Эта новая функция ABBYY FineReader 8.0 позволяет объединить в один сценарий несколько операций, выполняемых последовательно, и запускать их выполнение одним нажатием кнопки. Предусмотрен ряд готовых сценариев, имеется также возможность создавать собственные сценарии.

Распознавание гиперссылок

Новая версия системы автоматически находит в тексте гиперссылки на веб-сайты и адреса электронной почты и восстанавливает их в выходных документах. В распознанный текст также можно добавить новые гиперссылки.

Режим быстрого распознавания

Новый режим быстрого распознавания, предусмотренный в системе ABBYY FineReader 8.0, позволяет ускорить обработку документов в 2-2,5 раза. Данный режим рекомендуется для распознавания документов с простым оформлением и высоким качеством печати, отсканированных с достаточным разрешением. В большинстве случаев быстрый режим обеспечивает хорошее качество распознавания, например, при преобразовании бумажных документов в PDF-файлы с возможностью полнотекстового поиска. Однако, для более сложных документов предпочтительно использовать тщательный режим распознавания.

Сохранение в формате Microsoft Reader e-book (LIT)

В новой версии ABBYY FineReader распознанный текст можно сохранить в формате LIT, используемом для создания электронных книг. Формат LIT позволяет получать документы, пригодные для просмотра на экранах карманных компьютеров и других портативных устройств.

Добавление свойств документа

ABBYY FineReader 8.0 позволяет указать для распознанного текста дополнительные свойства, такие как описание, автор, тема, а также ключевые слова. Сохранить текст можно в файле любого из этих форматов: PDF, DOC/RTF, XLS, HTML, Word XML или LIT. Указанные свойства могут в дальнейшем использоваться для индексирования и поиска по документам.

Расширенная языковая и словарная поддержка

Общее количество языков распознавания в новой версии ABBYY FineReader 8.0 достигло 179. Словарная поддержка и функция проверки правописания доступны для 36 языков.

Словари юридических и медицинских терминов для немецкого и английского языков добавлены к остальным немецким и английским словарям. Поэтому теперь нет необходимости устанавливать специализированные языки распознавания при работе с текстами юридической и медицинской тематики.

Открытие многостраничных PDF и TIFF-файлов

Если вам нужен фрагмент многостраничного документа, а не весь документ целиком, в ABBYY FineReader 8.0 вы можете открыть и распознать только выбранные страницы.

ABBYY Screenshot Reader

(для зарегистрированных пользователей версии ABBYY FineReader 8.0 Professional Edition и всех пользователей версии ABBYY FineReader 8.0 Corporate Edition)

Это простое и удобное приложение поможет вам распознать текст с любой области экрана компьютера. Распознанный текст (как и полученный снимок экрана) может быть сохранен в файл или передан в буфер обмена.

ABBYY Hot Folder & Scheduling

(только для версии ABBYY FineReader 8.0 Corporate Edition)

По сравнению с предыдущей версией значительно расширились возможности автоматической проверки папок и обработки поступивших в них документов. Новое приложение-планировщик позволяет автоматически выполнять обработку документов в наиболее удобное время. Например, можно организовать автоматическое распознавание в ночное время, когда компьютер наименее загружен.

Глава 1

Работа с программой ABBYY FineReader 8.0

Содержание:

- Установка и запуск
- Получение и обработка изображения
- Анализ макета страницы
- Распознавание
- Проверка результатов
- Сохранение распознанного текста
- Работа с пакетами
- Сценарии ABBYY FineReader
- Коллективная работа в сети

Установка и запуск ABBYY FineReader 8.0

В данном разделе описаны основные инструкции по установке программы ABBYY FineReader 8.0: системные требования, установка на одно рабочее место и сетевая установка.

Для установки ABBYY FineReader используется специальная программа установки, входящая в состав дистрибутива. Всегда выполняйте установку с установочного компакт-диска, который вы получили при покупке.

Требования к системе

Для установки и работы программы ABBYY FineReader 8.0 необходимо:

1. ПК с процессором семейств Intel® Pentium®/Celeron®/Xeon™, AMD K6/Athlon™/Duron™/Sempron™ или совместимым с ними процессором, тактовая частота которого составляет не менее 500 МГц, или более мощным.
2. Операционная система Microsoft Windows 2003, Windows XP, Windows 2000 (для работы с локализованным интерфейсом операционная система должна обеспечивать необходимую языковую поддержку).
3. Оперативная память – 128 МБ. При работе в многопроцессорных системах требуется дополнительно 16 МБ оперативной памяти для каждого дополнительного процессора.
4. Свободное место на диске: 250 МБ для обычной установки, 100 МБ для работы системы.
5. 100% TWAIN-совместимый сканер, цифровая камера или факс-модем.
6. Видеоплата и монитор с разрешением не менее 800х600 точек.
7. Клавиатура, мышь или другое указательное устройство.

Установка программы ABBYY FineReader 8.0

Программа установки поможет вам установить ABBYY FineReader 8.0. Перед началом установки рекомендуется закрыть все другие приложения.

Чтобы установить ABBYY FineReader 8.0:

1. Вставьте установочный компакт-диск в дисковод для компакт-дисков. Программа установки запустится автоматически.
2. Следуйте инструкциям программы установки.

В случае если программа установки не запустилась автоматически:

1. Нажмите кнопку **Пуск** на Панели задач и выберите пункт **Настройка/Панель управления**.
2. Дважды нажмите на значок **Установка и удаление программ**.
3. Выберите закладку **Установка и удаление** и нажмите кнопку **Установить...**
4. Далее следуйте инструкциям программы установки.

Опции установки

В процессе установки программы ABBYY FineReader вам будет предложено выбрать одну из предложенных опций установки:



- **Обычная (рекомендуется)** – устанавливаются все компоненты программы, в том числе все языки распознавания, язык интерфейса (один) – выбранный при установке.
- **Установка по выбору** – из предложенного набора компонентов программы вы можете выбрать только те, которые нужны вам (в том числе языки распознавания).

Если при установке программы возникло сообщение об ошибке, обратитесь к файлу **readme.htm** (он находится на установочном компакт-диске ABBYY FineReader 8.0), в котором приведены рекомендации, помогающие в большинстве случаев.

Внимание! Если вы хотите применять пользовательские словари и эталоны, созданные в предыдущей версии программы, не удаляйте ее до установки новой версии. Установив новую версию, вы сможете применять эталоны и словари предыдущей версии.

Запуск программы ABBYY FineReader 8.0

Чтобы запустить программу ABBYY FineReader 8.0:

- Выберите пункт ABBYY FineReader 8.0 Professional Edition (Corporate Edition) в меню **Пуск/Программы/ABBYY FineReader 8.0**, или ,
- В приложении **Microsoft Word** нажмите кнопку , расположенную на панели инструментов.
- В **Проводнике Windows** выделите мышью файл изображения или PDF файл; в локальном меню файла выберите команду **Открыть с помощью ABBYY FineReader 8.0**.

Внимание! Перед запуском программы убедитесь, что ваш сканер подключен к компьютеру и включен в сеть. Если сканер еще не подключен, обратитесь к документации, поставляемой со сканером.

Если у вас нет сканера, вы можете использовать программу ABBYY FineReader для распознавания графических файлов (примеры таких файлов находятся в папке **ABBYY FineReader/Demo**).

Установка ABBYY FineReader Corporate Edition в локальной сети

Установка ABBYY FineReader 8.0 Corporate Edition на сетевой сервер осуществляется администратором сети. Установка осуществляется в два этапа. Сначала программа устанавливается на сервер, а затем с сервера – на рабочие станции одним из способов:

- используя Active Directory;
- с помощью Microsoft System Management Service (SMS);
- с помощью Task Scheduler;

- из командной строки;
- вручную, в интерактивном режиме.

Чтобы установить ABBYY FineReader 8.0 Corporate Edition на сервер:

1. Вставьте компакт-диск в дисковод для компакт дисков.
2. Запустите файл **Adminsetup.exe** с установочного компакт-диска ABBYY FineReader.

Подробную информацию об установке программы ABBYY FineReader на рабочие станции, о работе с Менеджером лицензий, а также об особенностях работы с программой в корпоративной сети вы можете найти в «Руководстве системного администратора», которое находится в подпапке **Administrator's Guide** папки сервера, куда был установлен ABBYY FineReader.

Получение и обработка изображения программой ABBYY FineReader

Качество распознавания во многом зависит от качества исходного изображения. В этой главе вы узнаете о том, как правильно отсканировать документ, как открыть и распознать уже имеющиеся на компьютере изображения (список поддерживаемых системой форматов приведен в разделе «Поддерживаемые графические форматы»), как можно обработать изображение и устранить некоторые его дефекты (например, мусор, возникший при сканировании) и т.д.

Сканирование

ABBYY FineReader работает со сканерами через TWAIN-интерфейс. Это единый международный стандарт, введенный в 1992 году для унификации взаимодействия устройств для ввода изображений в компьютер (например, сканера) с внешними приложениями. При этом возможно два варианта взаимодействия программы со сканерами через TWAIN-драйвер:

- через интерфейс ABBYY FineReader: в этом случае для настройки опций сканирования используется диалог программы ABBYY FineReader **Настройки сканера...**
- через интерфейс TWAIN-драйвера сканера: для настройки опций сканирования используется диалог TWAIN-драйвера сканера.

Преимущества каждого из режимов

В режиме **Использовать интерфейс TWAIN-драйвера сканера**, как правило, доступна функция предварительного просмотра изображения (preview), позволяющая точно задать размеры сканируемой области, подобрать яркость, тут же контролируя результаты этих изменений. Диалог TWAIN-драйвера сканера у каждого сканера выглядит по-своему, в большинстве случаев все надписи приводятся на английском языке. Вид этого окна и смысл опций описан в документации, прилагаемой к сканеру.

В режиме **Использовать интерфейс ABBYY FineReader** доступны такие опции, как возможность сканирования в цикле на сканерах без автоподатчика, сохранение опций сканирования в отдельный файл **набора опций** (*.fbt) и возможность использования этих опций в других пакетах.

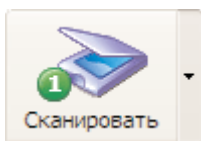
Вы можете легко переключаться между этими режимами:

- на закладке **Сканировать/Открыть** диалога **Опции** (меню **Сервис>Опции**) установите переключатель в одно из положений: **Использовать интерфейс TWAIN-драйвера сканера** или **Использовать интерфейс ABBYY FineReader**.

Замечания. 1. Для некоторых моделей сканеров опция **Использовать интерфейс ABBYY FineReader** может быть по умолчанию отключена (недоступна).
 2. Чтобы в режиме **Использовать интерфейс ABBYY FineReader** показывался диалог **Настройки сканера**, на закладке **Сканировать/Открыть** (**Сервис>Опции**) отметьте пункт **Запрашивать опции** перед началом сканирования.

Важно! Для того чтобы правильно подключить сканер, обратитесь к документации, прилагаемой к сканеру. При установке не забудьте установить **необходимое** программное обеспечение, поставляемое вместе со сканером (драйвер TWAIN и/или сканирующую программу).

Чтобы запустить сканирование:



Нажмите кнопку **1-Сканировать** или в меню **Файл** выберите пункт **Сканировать изображение**. Спустя некоторое время в **главном окне** программы ABBYY FineReader появится окно **Изображение** с «фотографией» отсканированной страницы.

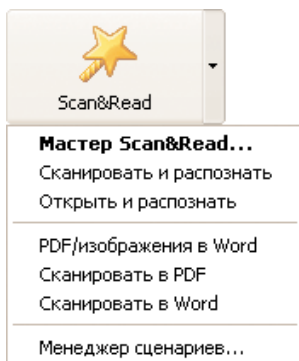
Если вы хотите отсканировать несколько страниц, то на закладке **Сканировать/Открыть** (**Сервис>Опции**) отметьте пункт **Сканировать несколько страниц**.

Замечание. Также можно открыть диалог **Опции**, выбрав пункт **Опции** в меню кнопки **1-Сканировать**.

В случае если сканирование не началось сразу, показывается один из следующих диалогов: встроенный **TWAIN-диалог** сканера. Установите параметры сканирования и нажмите кнопку **Scan** (**Final** и т.д., название кнопок зависит от конкретной модели сканера);

диалог **Настройки сканера**. Установите параметры сканирования и нажмите кнопку **Scan** (**Final** и т.д., название кнопок зависит от конкретной модели сканера).

Совет: Если вы хотите сразу запустить распознавание отсканированных страниц:



Нажмите на стрелку справа от кнопки **Scan&Read** и в локальном меню кнопки выберите пункт **Сканировать и Распознать**

ABBYY FineReader отсканирует и распознает изображения. В **главном окне** программы появятся окно **Изображение** с «фотографией» вставленного листа и окно **Текст** с результатом распознавания. Распознанный текст вы можете передать во внешние приложения или сохранить в любом из поддерживаемых форматов.

Установка параметров сканирования

Качество распознавания во многом зависит от того, насколько хорошее изображение получено при сканировании. Качество изображения регулируется установкой основных параметров сканирования: типа изображения, разрешения и яркости.

Основными параметрами сканирования являются:

- **Тип изображения** – серый (256 градаций), черно-белый или цветной. Сканирование в сером является оптимальным режимом для системы распознавания. В случае сканирования в сером режиме осуществляется автоматический подбор яркости. **Черно-белый** тип изображения обеспечивает более высокую скорость сканирования, но при этом теряется часть информации о буквах, что может привести к ухудшению качества распознавания на документах среднего и низкого качества печати. Если вы хотите, чтобы содержащиеся в документе цветные элементы (картинки, цветные буквы и цветной фон) были переданы в электронный документ с сохранением цвета, необходимо выбрать **цветной** тип изображения. В других случаях используйте **серый** тип изображения.
- **Разрешение** – используйте 300 dpi для обычных текстов (размер шрифта 10 и более пунктов) и 400 – 600 dpi для текстов, набранных мелким шрифтом (9 и менее пунктов).
- **Яркость** – в большинстве случаев подходит среднее значение яркости – 50%. На некоторых документах при сканировании в черно-белом режиме может понадобиться дополнительная настройка яркости.

Замечание. Сканирование с разрешением 400–600 dpi вместо 300 dpi или сканирование в сером или в цвете может занять существенно больше времени, чем сканирование в черно-белом режиме. На некоторых моделях сканеров сканирование с разрешением 600 dpi занимает в 4 раза больше времени, чем сканирование с разрешением 300 dpi.

Чтобы установить параметры сканирования:

- При сканировании через TWAIN с использованием **интерфейса ABBYY FineReader**: на закладке **Сканировать/Открыть** в диалоге **Опции** (меню **Сервис/Опции**) щелкните по кнопке **Настройки сканера** и в открывшемся диалоге **Настройки сканера** установите нужные опции.
- При сканировании через TWAIN с использованием **интерфейса TWAIN-драйвера сканера** для установки параметров сканирования используется диалог сканера, который открывается автоматически при нажатии на кнопку **1–Сканировать**. Опции для установки параметров сканирования могут называться по-разному, в зависимости от модели сканера. Например, яркость может называться brightness, threshold, изображаться «солнышком» или черно-белым кружком. Смысл опций описан в документации, прилагаемой к сканеру.

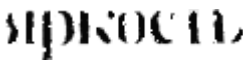

Советы по подбору яркости

Отсканированное изображение должно быть удобочитаемым (просмотрите изображение в окне **Крупный план**).

яркость

– пример хорошего (пригодного для распознавания) изображения.

Если в полученном изображении вы обнаружили множество дефектов (разрывов или склеек букв), то обратитесь к таблице, приведенной ниже. В ней указаны возможные способы устранения этих дефектов.

Особенности входного изображения:	Что сделать:
 «разорванные»; светлые, тонкие буквы	<ul style="list-style-type: none"> • Уменьшить яркость (чтобы изображение стало темнее) • Попробовать отсканировать в сером (в этом случае осуществляется автоподбор яркости)
 искаженные и залитые; склеенные символы; темные, толстые буквы	<ul style="list-style-type: none"> • Увеличить яркость (сделать изображение светлее) • Попробовать отсканировать в сером (в этом случае осуществляется автоподбор яркости)

Сканирование многостраничных документов

Для удобства сканирования большого количества страниц в программе ABBYY FineReader предусмотрен специальный режим сканирования: **Сканировать несколько страниц**. Он позволяет отсканировать несколько страниц подряд. Для того чтобы включить этот режим, отметьте опцию **Сканировать несколько страниц** на закладке **Сканировать/Открыть** диалога **Опции** (меню **Сервис>Опции**).

При этом:

- при сканировании через TWAIN с использованием **интерфейса ABBYY FineReader** по окончании сканирования страницы сканер автоматически начинает сканирование следующей;
- при сканировании через TWAIN с использованием **интерфейса TWAIN-драйвера** сканера **TWAIN-диалог** сканера не закрывается после окончания сканирования первой страницы. Вы можете положить следующую страницу в сканер и продолжить сканирование.

Вы можете отсканировать большое количество страниц двумя способами: с использованием автоподатчика (ADF) и без него.

При использовании автоподатчика (ADF)

- Интерфейс ABBYY FineReader. В диалоге **Настройки сканера** отметьте опцию **Загружать страницы из автоподатчика бумаги (ADF)** и запустите сканирование, нажав кнопку **1-Сканировать**.
- Интерфейс TWAIN-драйвера сканера. В **TWAIN-диалоге** сканера отметьте опцию использования автоподатчика (эта опция может называться по-разному в зависимости от модели сканера; см. документацию, поставляемую со сканером).

Замечание. Для того чтобы открыть диалог **Настройки сканера...**, щелкните по кнопке **Настройки сканера** на закладке **Сканировать/Открыть** диалога **Опции** (меню **Сервис/Опции**).

Без использования автоподатчика (ADF)

1. Интерфейс ABBYY FineReader

Для удобства сканирования нескольких страниц подряд на планшетном сканере без автоподатчика:

- Установите величину паузы (время от окончания сканирования одной страницы до начала сканирования следующей). Для этого в диалоге **Настройки сканера** отметьте опцию **Пауза между страницами** и установите величину паузы в секундах.
В этом случае после сканирования первой страницы сканер делает указанную паузу, во время которой вы вставляете в сканер следующую страницу. Затем сканирование автоматически продолжается.
- Выберите режим **Останавливаться между страницами**. Для этого в диалоге **Настройки сканера** отметьте опцию **Останавливаться между страницами**.
В этом случае каждый раз по окончании сканирования страницы будет появляться диалог с вопросом о продолжении сканирования. Нажмите кнопку **Да** для сканирования следующей страницы и **Нет** – для завершения сканирования.

Для завершения сканирования в меню **Файл** выберите пункт **Остановить сканирование**.

Замечание. Для того чтобы открыть диалог **Настройки сканера**, щелкните по кнопке **Настройки сканера** на закладке **Сканировать/Открыть** диалога **Опции** (меню **Сервис/Опции**).

2. Интерфейс TWAIN-драйвера сканера

- Убедитесь в том, что на закладке **Сканировать/Открыть** диалога **Опции** (меню **Сервис>Опции**) отмечен пункт **Сканировать несколько страниц**. Для начала сканирования в открывшемся TWAIN-диалоге сканера нажмите кнопку **Scan** (**Final** и т.д., название кнопки зависит от конкретной реализации TWAIN-драйвера сканера).

Для продолжения сканирования нажмите на кнопку **Scan** в **Twain-диалоге** сканера.

Для завершения сканирования в **Twain-диалоге** сканера нажмите на кнопку **Close** (или аналогичную ей по смыслу).

Совет: Чтобы вы могли контролировать результаты сканирования, в группе **Вид диалога** **Дополнительные опции** отметьте опцию **Открывать изображения по мере сканирования**. (Для того чтобы открыть диалог **Дополнительные опции**, щелкните по кнопке **Дополнительные опции** на закладке **Общие** диалога **Опции**). В этом случае по окончании сканирования страницы отсканированное изображение появится в окне **Изображение**. Если изображение отсканировано неверно, остановите сканирование (в меню **Файл** выберите пункт **Остановить сканирование**) и отсканируйте изображение заново.

Решение проблем сканирования: сканер не поддерживает работу с Twain-драйвером

Если ваш сканер не поддерживает работу с TWAIN, то вы можете работать с программой следующим образом:

1. Создайте в программе новый пакет, откройте пакет или продолжите работу в уже открытом пакете.
2. Установите параметры распознавания (язык распознавания, тип страницы, тип печати).
3. В меню **Процесс** выберите пункт **Запустить фоновое распознавание**.
4. Не закрывая программы, отсканируйте документ, который вы хотите распознать, любой графической программой и сохраните его в ту папку, в которую вы сохранили созданный пакет с именем 0001.TIF. ABBYY FineReader автоматически подхватит изображение и распознает его. Чтобы обновить список страниц, нажмите клавишу **F5**.

Замечание. Если при добавлении страниц в пакет уже были страницы, то имена добавленных файлов (также как и номера соответствующих им новых страниц пакета) будут начинаться не с единицы (1), а с номера страницы, следующей за последней страницей пакета, т.е. если в пакете было 10 страниц, то новым файлам будут присвоены следующие имена: 0011.tif, 0012.tif и т.д.

5. Отсканируйте второй документ и сохраните его с именем 0002.TIF и т.д.
6. Чтобы прекратить распознавание, в меню **Процесс** выберите пункт **Остановить фоновое распознавание**.

Все отсканированные таким образом страницы будут распознаны программой.

Открытие файлов с изображениями и PDF файлов

Если у вас нет сканера, вы можете открывать и распознавать готовые графические файлы (список открываемых форматов приведен в разделе «Поддерживаемые графические форматы»).

Чтобы открыть изображение:

- Нажмите стрелку справа от кнопки **1–Сканировать** и в локальном меню выберите пункт **Открыть изображение**. Внешний вид значка изменится; подпись **Сканировать** поменяется на **Открыть**.
- В меню **Файл** выберите пункт **Открыть PDF/изображение...**
- В **Windows Explorer**: щелкните правой кнопкой мыши на файле с изображением и в локальном меню выберите пункт **Открыть с помощью ABBYY FineReader**. Если на Вашем компьютере уже открыт ABBYY FineReader, изображение будет добавлено в текущий пакет, в противном случае перед добавлением изображения в пакет автоматически запустится ABBYY FineReader с пакетом, с которым Вы работали в последний раз.
- В **Microsoft Outlook** и/или **Windows Explorer**: Нажмите левой кнопкой мыши на файле с изображением, которое вы хотите открыть, и не отпуская кнопки, перетащите его на свернутое окно программы ABBYY FineReader. Изображение будет добавлено в текущий пакет и открыто в окне **Изображение**.

В диалоге **Открыть изображение** выберите одно или несколько изображений. Выбранные изображения появятся в окне **Пакет**, и последнее из выбранных изображений откроется в окне **Изображение** и в окне **Крупный план** ABBYY FineReader, при этом копия изображения помещается в папку пакета. Более подробно о представлении страниц в пакете и об устройстве пакета см. «Общая информация по работе с пакетом».

Совет: Если вы хотите, чтобы открытые изображения автоматически распознавались сразу после открытия:

1. Нажмите на стрелку справа от кнопки **Scan&Read** и в локальном меню кнопки выберите пункт **Открыть и Распознать**.
 2. ABBYY FineReader откроет файл и распознает изображения. В главном окне программы появятся окно **Изображение** с открытым изображением и окно **Текст** с результатом распознавания. Распознанный текст вы можете передать во внешние приложения или сохранить в любом из поддерживаемых форматов.
-

Особенности открытия PDF-файлов

Создатель PDF-файла может ограничить доступ к своему файлу, например, защитить его паролем, установить запрет на открытия файла или извлечение из него текста и графики. При открытии подобных файлов ABBYY FineReader будет запрашивать пароль, чтобы обеспечить защиту авторских прав создателя файла.

Добавление в пакет изображений со сдвоенными страницами

При сканировании книг удобнее отсканировать две страницы (книжный разворот) сразу. При этом для повышения качества распознавания такие изображения следует разделить на два, чтобы каждой странице соответствовала отдельная страница пакета (анализ и распознавание осуществляется для каждой страницы по отдельности, исправляется перекося строк).

Для этого перед сканированием или добавлением в пакет сдвоенных страниц:

- На закладке **Сканировать/Открыть** (меню **Сервис>Опции**) отметьте опцию **Делить разворот книги**.

В этом случае книжный разворот (сдвоенные страницы) будет представлен двумя страницами пакета. Подробнее о пакете см. «Общая информация по работе с пакетом».

Замечание. Если книжный разворот (сдвоенные страницы) был разделен на две страницы неудачно, то снимите отметку с опции **Делить разворот книги**, заново отсканируйте или добавьте в пакет изображение с книжным разворотом и попробуйте разделить его в диалоге **Разбить изображение** (меню **Изображение>Разбить изображение**) вручную.

Добавление в пакет изображений с визитными карточками



При вводе в компьютер информации с визитных карточек их удобнее сканировать не по одной, а сразу по несколько штук. При этом качество распознавания будет выше (в частности, благодаря исправлению перекосов) в том случае, когда каждая визитка попадет в пакет в виде отдельной страницы. Для этого в системе предусмотрены средства автоматического и ручного разделения изображений, содержащих расположенные в определенном порядке визитные карточки.

Замечание. Необходимо чтобы визитные карточки были разложены определенным образом на стекле сканера. Подробнее см. раздел «Работа с визитными карточками» в «Обучении на примерах».

Чтобы разбить изображение:

1. В окне **Пакет** выберите нужное изображение.
2. В меню **Изображение** выберите пункт **Разбить изображение**.
3. В открывшемся диалоге **Разбить изображение** нажмите кнопку **Разбить на визитки**.

Замечания.

1. Разделяемая страница удаляется из пакета; на ее место добавляются новые страницы, соответствующие разрезанным частям. Подробнее о пакете см. «Общая информация по работе с пакетом».
2. Если изображение было разрезано на визитки неверно, то попробуйте разрезать его вручную, воспользовавшись кнопками **Добавить вертикальный разделитель/Добавить горизонтальный разделитель**.
3. Чтобы удалить все разделители, нажмите кнопку **Удалить разделители**.
4. Чтобы передвинуть разделитель, перейдите в режим **Выбора объекта** (нажмите кнопку ) и переместите разделитель.
5. Чтобы удалить разделитель, перейдите в режим **Выбора объекта** (нажмите кнопку ) и переместите разделитель за пределы изображения.

Съемка текстов фотоаппаратом

Фотосъемка документов требует определенной подготовки фотографа, а также налагает ограничения на характеристики фотоаппарата и режим съемки. Данная статья поможет вам выбрать режим работы фотоаппарата и получить снимок документа, пригодный для распознавания текста. Технические параметры вашего фотоаппарата и инструкции по работе с ним приведены в описании фотоаппарата и другой документации, прилагаемой к фотоаппарату.

Перед съемкой убедитесь в том, что:

1. Страница целиком умещается в кадре, и что она целиком его занимает (не остается излишних полей).
2. Свет достаточно ровный, без теней на фотографируемом тексте.
3. Фотографируемый документ расположен перпендикулярно оптической оси объектива, т.е. фотоаппарат должен быть расположен напротив центра текста. Неровности бумаги (у корешка книги, например) нужно по возможности разгладить.

Требования к фотоаппарату

Минимальные требования

- Размер матрицы 2 млн. пикселей
- Переменная дистанция фокусировки. Не рекомендуется использовать фотоаппараты с фиксированным фокусом (сфокусированные на гиперфокальное расстояние), такие камеры обычно встречаются в сотовых телефонах и КПК.

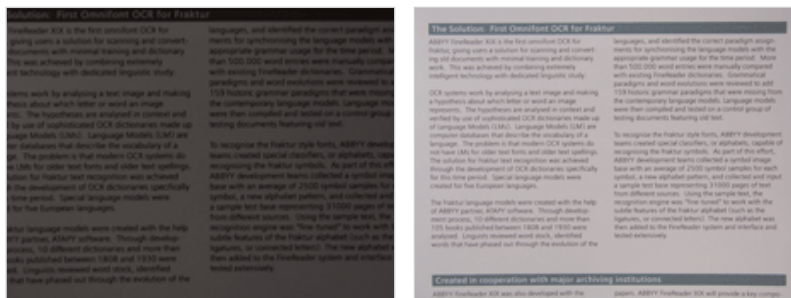
Рекомендуемые параметры фотоаппарата

- Размер матрицы 5 млн. пикселей
- Возможность отключения фотовспышки
- Возможность установки диафрагмы вручную – т.е. наличие режима приоритета диафрагмы или ручного режима
- Режим ручной фокусировки
- Объектив с оптической стабилизацией изображения, при его отсутствии рекомендуется использовать штатив
- Оптический зум

Как фотографировать тексты

Освещение

По возможности обеспечьте хорошее освещение, лучше всего дневное. Если используется подсветка лампой, то лучше использовать подсветку двумя лампами с разных сторон, чтобы избежать образования теней.



Как располагать фотоаппарат

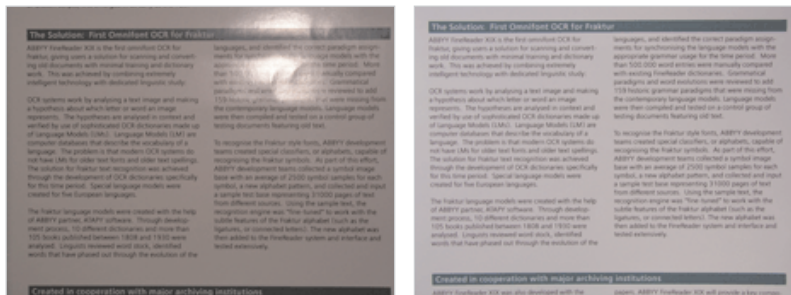
Для съемки документов рекомендуется использовать штатив. Снимать лучше всего при максимальном оптическом увеличении, при этом фотоаппарат должен быть максимально удален от текста. Объектив должен располагаться параллельно поверхности съемки и по центру фотографируемого изображения.

Фотоаппарат должен быть удален от страницы настолько, чтобы при максимальном увеличении страница полностью умещалась в кадре. Обычно это расстояние около 50 – 60 см.

Вспышка

Если освещение позволяет, вспышку лучше всего отключить, поскольку она создает пересвеченные области и резкие тени. Тем не менее, если внешнего света недостаточно, то можно использовать вспышку, если съемка ведется с достаточного расстояния (~50 сантиметров). Даже при использовании вспышки все равно желательно подсветить документ.

Внимание! Категорически не рекомендуется использовать вспышку при съемке документов, отпечатанных на глянцевой бумаге.



Выбор режима съемки

Диафрагма

При недостаточном освещении рекомендуется выбирать небольшие значения диафрагмы (~2,3 – 4,5), т.е. максимально открывать диафрагму. Если съемка ведется при ярком дневном свете, значение диафрагмы лучше увеличить, чтобы получить более резкий снимок.

Чувствительность матрицы

При недостаточном освещении рекомендуется выбрать большую чувствительность матрицы (большее значение ISO).

Фокусировка

При недостаточном освещении автоматический фокус может срабатывать плохо (аппарат не может сфокусироваться), в этом случае рекомендуется использовать ручную фокусировку.

Баланс белого

По возможности установите баланс белого по цвету бумаги. Если ваш фотоаппарат не позволяет произвольно задавать баланс белого, выберите режим, максимально отвечающий условиям съемки.



Дополнительные рекомендации

При недостаточном освещении в автоматическом режиме используются большие выдержки, что отрицательно сказывается на резкости получаемого изображения. Поэтому рекомендуется дополнительно:

- Использовать стабилизатор изображения.
- Использовать автоспуск. Это позволит избежать подрагивания камеры при нажатии на кнопку спуска. Такие проблемы могут возникать даже при наличии штатива.

Что делать, если...

Снимок слишком темный и неконтрастный.

Постарайтесь улучшить освещение. Если такой возможности нет – установите меньшее значение диафрагмы.

Снимок нерезкий.

Возможно, автофокус плохо срабатывает из-за отсутствия света. Попробуйте улучшить освещение. Если это не помогает, используйте ручную фокусировку.

Если нерезкая только часть снимка – попробуйте выставить большее значение диафрагмы. Снимайте с большего расстояния при максимальном оптическом увеличении. Наводите фокус на точку, расположенную между центром и краем изображения.

Вспышка создает пересвеченную область в центре кадра.

Отключите вспышку. Если нет возможности использовать другие источники света, снимайте с большего расстояния.

Проверка и корректирование полученного изображения

Очистить от мусора

Распознаваемое изображение может быть сильно «замусорено», т.е. содержать много лишних точек, возникших в результате сканирования документов среднего или низкого качества. Точки, близко расположенные к контурам букв, могут отрицательно сказаться на качестве распознанного текста. Чтобы уменьшить количество лишних точек, можно воспользоваться опцией **Очистить от мусора**. Для этого:

- Выберите пункт **Очистить изображение от мусора** в меню **Изображение>Обработать изображение**

Если вы хотите очистить от «мусора» отдельный блок, то:


- Выберите пункт **Очистить блок от мусора** в меню **Изображение>Обработать изображение**.

Внимание! Если исходный текст был очень светлым или в исходном тексте использовался очень тонкий шрифт, то применение функции **Очистить изображение от мусора** может привести к исчезновению точек, запятых или тонких элементов букв, что ухудшает качество распознавания.


Изменить разрешение изображения

Разрешение изображения – это параметр, определяющий, какое количество точек, составляющих изображение, приходится на единицу длины. Разрешение обычно измеряется в dpi – количестве точек, приходящихся на один дюйм. Для качественного распознавания текста с помощью системы ABBYY FineReader необходимо, чтобы разрешение изображения по вертикали и по горизонтали совпадало, и при этом разрешение находилось в пределах допустимого диапазона (50 – 3200 dpi). Рекомендованный диапазон разрешения – 200 – 600 dpi, разрешение оптимальное с точки зрения распознавания – 300 dpi.

Слишком большое или слишком маленькое разрешение может приводить к ухудшению качества распознавания. У некоторых форматов изображения разрешение отсутствует (например, у *.bmp файлов). Изображения могут также иметь нестандартное разрешение (например, 204*96 dpi), что тоже может повлиять на качество распознавание.

Система ABBYY FineReader проверяет разрешение каждого изображения и при обнаружении «подозрительного» изображения автоматически исправляет его разрешение, при этом физические размеры изображения (его длина и ширина) не изменяются. Такое изображение помечается значком  в окне **Пакет**. При наведении мыши на такое изображение возникает всплывающая подсказка.

Если качество распознавания изображения низкое, корректировка разрешения изображения может привести к улучшению результатов распознавания. Для того чтобы исправить разрешение изображения:

1. В окне **Пакет** выберите изображение со значком . Если текст всплывающей подсказки говорит о том, что изображение имеет некорректное разрешение, то в меню **Изображение** выберите команду **Исправить разрешение...**
2. В открывшемся диалоге укажите тип изображения (отсканированное изображение, изображение, полученное по факсу, или снимок экрана). Вы также можете указать точное значение разрешения в поле **Другое разрешение**.
3. Если разрешение нужно исправить только у выбранных изображений, отметьте в группе **Исправить разрешение** опцию **Выбранных изображений**. Если все изображения пакета имеют неверное разрешение, отметьте опцию **Всех изображений пакета**. Поскольку в этом случае разрешение всех изображений пакета будут приведено к одному значению, данную операцию рекомендуется применять, только если все изображения пакета получены из одного источника.

Устранить искажение строк

При сканировании книг возможно искажение строк текста в той части изображения, где страница примыкала к переплету. На изображениях, полученных с помощью фотокамеры, строки текста также могут искажаться по краям изображения. Для того чтобы устранить искажения строк:

- Отметьте опцию **Устранить искажение строк** в диалоге **Опции** на закладке **Сканировать/Открыть** (меню **Сервис>Опции**).

Замечание. Для выполнения данной операции может потребоваться значительное время.

Инвертировать изображение

Некоторые сканеры инвертируют изображения при сканировании (черный цвет переводят в белый, а белый в черный). Чтобы получить стандартное представление документа (черный шрифт на белом фоне):



- В меню **Изображение>Обработать изображение** выберите пункт **Инвертировать**.

Замечание. Если вы сканируете или открываете инвертированные изображения, то перед добавлением в пакет таких изображений отметьте пункт **Инвертировать изображения** в группе **Сканировать/Открыть** в диалоге **Дополнительные опции**. Для того чтобы открыть диалог **Дополнительные опции**, щелкните по кнопке **Дополнительные опции** на закладке **Общие** диалога **Опции** (меню **Сервис>Опции**).

Повернуть или зеркально отразить изображение


При распознавании изображение должно иметь стандартную ориентацию: текст должен читаться сверху вниз, и строки должны быть горизонтальными. По умолчанию программа при распознавании определяет и корректирует ориентацию изображения автоматически. Если ориентация изображения была определена ошибочно, то на закладке **Сканировать/Открыть** снимите отметку с пункта **Определять ориентацию страницы при распознавании** и поверните изображение вручную.

Чтобы повернуть изображение:

- на 90 градусов вправо – нажмите кнопку  или выберите в меню **Изображение>Повернуть/Отразить изображение** пункт **Повернуть по часовой стрелке**.
- на 90 градусов влево – нажмите кнопку  или выберите в меню **Изображение>Повернуть/Отразить изображение** пункт **Повернуть против часовой стрелки**.
- на 180 градусов – выберите в меню **Изображение>Повернуть/Отразить изображение** пункт **Повернуть на 180 градусов**.


Стереть участок изображения

Если вы хотите исключить какой-то участок текста из распознавания или на изображении имеются большие участки мусора, то вы можете стереть такие участки. Для этого:

- Выберите инструмент  (на панели в окне **Изображение**) и, нажав на левую кнопку мыши, выделите участок изображения, который вы хотите удалить. Отпустите кнопку, выделенная часть изображения будет удалена.



Обрезать изображение

Иногда в результате сканирования по краям изображения появляются зачерненные поля. В таком случае перед распознаванием можно выполнить обрезку изображения, удалив ненужные фрагменты. С помощью инструмента обрезки изображения можно также получить изображение стандартного размера (соответствующего одному из стандартных форматов, например, A4, A5).

1. На панели **Изображение** (в окне **Изображение**) выберите инструмент  (вы также можете воспользоваться командой **Обрезать изображение** в меню **Изображение**);
2. Изображение будет открыто в окне **Обрезать изображение**, контур изображения будет выделен черной линией. Для того чтобы:
 - Выбрать наиболее удобный режим просмотра изображения, воспользуйтесь выпадающим списком, находящимся в левом нижнем углу окна изображения;
 - Обрезать ненужные края изображения, потяните мышью черную линию, обрамляющую изображение, или маркеры, находящиеся в углах контура изображения. Часть изображения, которая будет отрезана, выделяется серым цветом. Щелкните по кнопке **Обрезать**;
 - Привести изображение к стандартному размеру, выберите нужный формат из выпадающего списка **Привести к**;
 - Не выполнять обрезку открытого изображения и перейти к следующему изображению, нажмите кнопку **Пропустить**;
 - Всегда работать только с выбранным изображением (не переходить к следующему изображению пакета после окончания работы с текущим изображением), снимите отметку с опции **Перейти к следующей странице**.

-
- Замечания.**
1. Обрезку изображения рекомендуется выполнять до того, на изображении будут выделены блоки и изображение будет распознано.
 2. Цвет рамки в окне **Обрезать изображение** можно изменить на закладке **Вид** диалога **Опции** (меню **Сервис>Опции**). В списке **Объекты** выберите пункт **Блок обрезки изображения**, затем щелкните по кнопке **Цвет** и в открывшемся диалоге выберите нужный цвет рамки.
-

Увеличить/Уменьшить масштаб изображения

- На панели **Изображение** (в окне **Изображение**) выберите инструмент  /  и щелкните мышью на изображении. Изображение увеличится/уменьшится в два раза.
- Щелкните правой кнопкой мыши на изображении и в локальном меню выберите пункт **Масштаб** и нужный вам масштаб.

Получить информацию об изображении

Вы можете получить следующую информацию об открытом изображении: ширину и высоту изображения в точках; вертикальное и горизонтальное разрешение в точках на дюйм (dpi); тип изображения. Чтобы просмотреть информацию об изображении:

- Щелкните правой кнопкой на изображении и в локальном меню выберите пункт **Свойства**. В открывшемся диалоге выберите закладку **Изображение**.

Печать изображения

Вы можете напечатать одно изображение, открытое в окне **Изображение**, несколько изображений, выделенных в окне **Пакет**, или все изображения. Для этого:

- В меню **Файл** выберите пункт **Печать>Изображение...** и в открывшемся диалоге **Печать** установите параметры печати (принтер, количество печатаемых страниц, количество копий и т.д.)

Отменить последнее действие

- Для отмены последнего действия на панели **Стандартная** нажмите кнопку **Отменить** .

Совет: Для повторного выполнения последнего отмененного действия на панели **Стандартная** нажмите кнопку **Восстановить** .

Нумерация страниц при добавлении в пакет

По умолчанию каждой сканируемой странице присваивается номер, на единицу больший номера последнего изображения в пакете.

Вы можете задать номер добавляемой страницы и вручную (например, вам нужно сохранить исходную нумерацию страниц или вы сканируете стопку отсортированных по порядку страниц). Для этого:

- на закладке **Сканировать/Открыть** (меню **Сервис>Опции**) отметьте пункт **Запрашивать номер страниц перед добавлением в пакет**.

При сканировании стопки двусторонних сортированных по порядку страниц:

1. Отметьте пункт **Запрашивать номер страницы перед добавлением в пакет** на закладке **Сканировать/Открыть (Сервис>Опции)**.
2. В диалоге **Номер страницы** укажите номер страницы, с которой начинается сканирование и выберите опцию **Через одну** в поле **Нумерация страниц**. Выберите способ нумерации страниц: по возрастанию или по убыванию. Возрастание или убывание зависит, например, от того, как вы кладете стопку в автоподатчик – находятся ли меньшие или большие номера наверху.

Опции хранения изображений в пакете

Приводить цветное/серое изображение к черно-белому

Отметьте эту опцию при сканировании через TWAIN-диалог сканера в сером режиме (с автоподбором яркости) или при сканировании в цвете, если при этом сканируемые документы не содержат цветных картинок, цветного шрифта и фона или же вам не требуется передача цвета в выходное изображение. В этом случае сохраняемые в пакет изображения будут занимать меньше места на диске.

Замечание. Данная опция устанавливается в диалоге **Дополнительные опции**. Для того чтобы открыть этот диалог, щелкните по кнопке **Дополнительные опции** на закладке **Общие** диалога **Опции** (меню **Сервис>Опции**).

Анализ макета страницы

Прежде, чем приступить к распознаванию, программа должна знать, какие участки изображения надо распознавать. Для этого проводится анализ макета страницы, во время которого выделяются блоки с текстом, картинки, таблицы и штрих-коды.

В этом разделе вы узнаете, когда может потребоваться ручной анализ макета страницы, какие типы блоков могут быть размечены на изображении, как отредактировать полученные в результате автоматического анализа блоки, а также, как упростить процесс анализа, используя шаблоны блоков.

Общая информация по анализу макета страницы

Анализ макета страницы может проводиться как автоматически, так и вручную. В большинстве случаев ABBYY FineReader сам успешно справляется с анализом сложных страниц. Автоматический анализ производится по нажатию кнопки **2-Распознать** одновременно с распознаванием текста.

Замечание. Отдельная процедура анализа макета страницы тоже доступна (меню **Процесс>Распознать>Анализ макета страницы**). Правда, при этом качество сегментации может быть ниже, т.к. при совместной процедуре распознавания и сегментации для анализа страницы используется дополнительная информация, полученная в процессе распознавания.

Ручное выделение блоков может понадобиться, если

1. вы хотите распознать часть страницы;
2. в результате автоматического анализа блоки были выделены неправильно.

-
- Совет:**
- В некоторых случаях качество автоматического анализа можно улучшить, правильно установив опции анализа макета. Проверьте установленные опции анализа (закладка **Распознать**, меню **Сервис>Опции**).
 - Если программа выделила неправильно некоторые блоки, часто оказывается быстрее исправить только их, воспользовавшись инструментами для редактирования блоков, а не выделять блоки на изображении заново вручную.
-

Типы блоков

Блоки – это заключенные в рамку участки изображения. Блоки выделяют для того, чтобы указать системе, какие участки отсканированной страницы надо распознавать и в каком порядке. Также по ним воспроизводится исходное оформление страницы. Блоки разных типов имеют различные цвета рамок. Вы можете изменить цвета рамок блоков на закладке **Вид** диалога **Опции** (меню **Сервис>Опции**) в группе **Объекты**. В поле **Объект** выберите нужный тип блока, а в поле **Цвет** – требуемый цвет.

При обработке изображений выделяют блоки следующих типов:

Зона Распознавания – блок используется для распознавания и автоматического анализа части изображения. После нажатия на кнопку **2-Распознать** выделенный блок автоматически анализируется и распознается.

Текст – блок используется для обозначения текста. Он должен содержать только одноколонный текст. Если внутри текста содержатся картинки, выделите их в отдельные блоки.

Таблица – этот блок используется для обозначения таблиц или текста, имеющего табличную структуру. При распознавании программа разбивает данный блок на строки и столбцы и формирует табличную структуру. В выходном тексте данный блок передается таблицей. Вы можете выделить и отредактировать таблицу вручную.

Картинка – этот блок используется для обозначения картинок. Он может содержать картинку или любую другую часть текста, которую Вы хотите передать в распознанный текст в качестве картинки.

Штрих-код – этот блок используется для распознавания штрих-кодов, если ваш документ содержит штрих-код, и вы хотите передать его не картинкой, а перевести его в последовательность букв и цифр, выделите штрих-код в отдельный блок и присвойте ему тип **Штрих-код**.

Замечание. Для автоматического распознавания штрих-кодов необходимо, чтобы в группе **Распознать** диалога **Дополнительные опции** был отмечен пункт **Искать штрих-коды**. Для того чтобы открыть диалог **Дополнительные опции**, щелкните по кнопке **Дополнительные опции** на закладке **Общие** диалога **Опции** (меню **Сервис>Опции**).

Типы штрих-кодов, поддерживаемых ABBYY FineReader 8.0

Code 3 of 9
Check Code 3 of 9
Code 3 of 9 without asterisk
Codabar
Code 93
Code 128
EAN 8
EAN 13
IATA 2 of 5
Interleaved 2 of 5
Check Interleaved 2 of 5
Matrix 2 of 5
Postnet
Industrial 2 of 5
UCC-128
UPC-A
UPC-E
PDF417

Опции автоматического анализа макета страницы

При автоматическом анализе макета страницы ABBYY FineReader сам выделяет блоки, содержащие тексты, таблицы, картинки и штрих-коды.

Автоматический анализ запускается по кнопке **2-Распознать** одновременно с распознаванием текста. До запуска распознавания необходимо установить опции анализа таблиц.

Опции анализа таблиц

В большинстве случаев программа делит таблицу на строки и столбцы автоматически. Дополнительные параметры анализа таблиц устанавливаются в группе **Распознать** диалога **Дополнительные опции**. (Для того чтобы открыть диалог **Дополнительные опции**, щелкните по кнопке **Дополнительные опции** на закладке **Общие** диалога **Опции**). Опции анализа таблицы рекомендуется использовать, если:

- в результате автоматического анализа макета страницы таблица была выделена и разделена на строки и столбцы неверно;
 - документ содержит много однотипных таблиц, для которых имеется дополнительная информация (например, таблица не содержит объединенных ячеек или таблица состоит из ячеек, текст в которых расположен в одну строку).
1. Опция **В каждой ячейке таблицы не более одной строки текста** используется для уверенного анализа таблиц с неполным количеством черных разделителей или без таковых, а также с ячейками, содержащими не более одной строки текста.

Километры	Мили
1	0.62
5	3.2

пример таблицы, в каждой ячейке которой текст расположен в одну строку

Физ.явление	t по Цельсию
Температура кипения воды	100
Температура замерзания воды	0

пример таблицы, ячейки которой содержат более одной строки текста

2. Опция **Таблица не содержит объединенные ячейки** используется для уверенного анализа таблиц, не содержащих объединенных ячеек.

Температура	
По Цельсию	По Кельвину
-273	0
100	373

пример таблицы с объединенными ячейками

Замечание. В таблицах общей структуры или на страницах с таблицами различной структуры отмеченные опции **В каждой ячейке таблицы не более одной строки текста** и **Таблица не содержит объединенных ячеек** могут привести к ошибкам анализа и, как следствие, понизить качество распознавания.

Выделение и редактирование блоков вручную



Чтобы создать новый блок:

1. Выберите один из инструментов:
 - выделить зону распознавания;
 - выделить текстовый блок;
 - выделить картинку;
 - выделить табличный блок.
2. Установите курсор мыши в угол предполагаемого блока. Нажмите левую кнопку мыши и, не отпуская кнопки, потяните в противоположный по диагонали угол.
3. Отпустите кнопку мыши.

Выделенная часть изображения будет заключена в рамку.

Вы можете поменять тип блока (присвоить выделенному блоку один из существующих типов: Зона распознавания, Текст, Таблица, Картинка или Штрих-код). Для этого:

- Щелкните на блоке правой кнопкой мыши и в локальном меню выберите **Изменить тип блока**, а затем – нужный вам пункт.


Редактирование формы и положения блоков

Чтобы передвинуть границу блока:


1. Установите курсор мыши на границу блока.
2. Нажмите левую кнопку мыши и потяните в нужную сторону.
3. Отпустите кнопку мыши.

Замечание. Если вы установите курсор мыши на угол блока, то при движении мыши будут одновременно изменяться вертикальная и горизонтальная границы блока.

Чтобы добавить прямоугольную часть блока:

1. Выберите инструмент .
2. Установите курсор мыши внутри блока, к которому вы хотите добавить часть. Нажмите левую кнопку мыши и, не отпуская кнопки, потяните по диагонали. Выделив нужную часть изображения, отпустите кнопку мыши. Выделенный прямоугольник будет добавлен к блоку.
3. Если нужно, передвиньте границу блока.

Чтобы удалить прямоугольную часть блока:


1. Выберите инструмент .
2. Установите курсор мыши внутри блока, там, где вы хотите вырезать часть. Нажмите левую кнопку мыши и, не отпуская кнопки, потяните по диагонали. Выделив нужную часть изображения, отпустите кнопку мыши. Выделенный прямоугольник будет удален из блока.
3. Если нужно, передвиньте границу блока.









Замечания.

1. Вы также можете изменять границы блоков, добавляя на них новые узлы (точки разбиения). Получившиеся отрезки можно перемещать мышью в любом направлении. Чтобы добавить новый узел, подведите курсор, удерживая клавишу **Shift**, к нужной точке границы (курсор при этом примет форму перекрестия) и щелкните мышью. На границе блока появится новый узел.
2. ABBYY FineReader накладывает на допустимую форму блоков некоторые ограничения, обусловленные необходимостью передавать на распознавание неразрывные текстовые строки. Для соблюдения этих ограничений ABBYY FineReader либо автоматически корректирует границы блока после операций добавления или удаления прямоугольной части блока (например, при удалении внутренней части блока снизу или сверху дополнительно


удаляется часть блока справа до границы блока), либо запрещает некоторые операции по передвижению отрезков, образующих границу блока.

Чтобы выделить один или несколько блоков:


- Выберите инструмент  и щелкните мышью по нужному блоку или, удерживая кнопку мыши, нарисуйте прямоугольник, охватывающий нужные блоки.

Замечание. Вы можете выделить один или несколько блоков, используя стандартные инструменты выделения блоков. Чтобы выбрать несколько блоков, нажмите клавишу **SHIFT** или **CTRL** (при этом должен быть выбран один из инструментов: , ,  или ) и мышью щелкните на требуемых блоках. Чтобы отменить выделение уже выбранного блока или добавить невыделенные блоки, нажмите клавишу **CTRL** (при этом должен быть выбран один из инструментов: , ,  или ) и мышью щелкните на требуемых блоках.

Чтобы передвинуть блок:


- Нажмите клавишу **ALT** (при этом должен быть выбран один из инструментов: , , , ) или ) и мышью переместите блоки.

Чтобы перенумеровать блоки:

1. Выберите инструмент .
2. Выделите блоки в том порядке, в котором Вы хотите видеть их содержимое в выходном тексте.

Замечание. Если вы перенумеровываете блоки на уже распознанном изображении, то одновременно в окне **Текст** в черновом режиме редактора происходит перегруппировка распознанного текста в соответствии с новой нумерацией.

Чтобы удалить блок:

- Выберите инструмент  и выделите блок, который Вы хотите удалить.
- Выделите блоки, которые Вы хотите удалить, и нажмите клавишу **DEL**.

Внимание! Если Вы удаляете блок с уже распознанного изображения, то одновременно с этим в окне **Текст** удаляется текст, соответствующий этому блоку.

Чтобы удалить все блоки на изображении:

- В меню **Изображение** выберите пункт **Удалить все блоки и текст**.

Внимание! Если вы удаляете блоки с уже распознанного изображения, то одновременно с этим в окне **Текст** удаляется текст, соответствующий этим блокам.

Ручной анализ таблицы

Совет: Если в результате автоматического анализа таблицы разделение на строки и столбцы произошло неверно, прежде чем анализировать таблицу вручную заново, попробуйте сначала отредактировать результаты автоматического анализа.

Чтобы отредактировать таблицу вручную:

Отредактируйте таблицу, используя инструменты (панель в окне **Изображение**):



– Добавить вертикальную линию.



– Добавить горизонтальную линию.



– Удалить линию.

Если ячейка таблицы содержит только картинку, в диалоге **Свойства блока** (меню **Вид>Свойства**) отметьте пункт **Считать ячейку картинкой**. Если же, помимо картинки, в ячейке содержится некоторый текст, то выделите картинку в отдельный блок внутри ячейки.

Чтобы объединить ячейки или строки таблицы:

В меню **Изображение>Ячейки таблицы** выберите пункты **Объединить ячейки** или **Объединить строки**. При объединении строк таблицы деление на столбцы остается.

Объединенные ячейки можно снова разделить, воспользовавшись командой **Разбить ячейки** (меню **Изображение>Ячейки таблицы**).

Замечание. Чтобы повторно не рисовать вертикальные и горизонтальные линии в таблице, выделите таблицу в отдельный блок и щелкните правой кнопкой мыши на блоке. В локальном меню выберите пункт **Анализ структуры таблицы**. А затем, используя инструменты по работе с таблицей, отредактируйте полученные результаты.

Использование шаблонов блоков

При работе с документами с одинаковым расположением текста и картинок, такими, например, как формы, бланки и т.п., вместо того чтобы анализировать макет каждой страницы, вы можете провести анализ одной из них, сохранить расположение блоков на этой странице в файл, а затем, когда потребуется, «спроецировать» эти блоки на изображение (или группу изображений) со сходным расположением текста.

Замечание. Созданный шаблон блоков можно использовать только для документов, отсканированных с тем же разрешением, что и документ, на котором данный шаблон блоков создавался.

Чтобы создать шаблон блоков:

1. Откройте изображение и выделите на нем блоки автоматически или вручную.
2. В меню **Изображение** выберите пункт **Сохранить шаблон блоков...** В открывшемся диалоге укажите имя для шаблона блоков.

Чтобы наложить шаблон блоков:

1. В окне **Пакет** выделите страницы, на которые вы хотите наложить существующий шаблон.
2. В меню **Изображение** выберите пункт **Загрузить шаблон блоков**. В открывшемся диалоге **Шаблон блоков** выберите файл (*.blk) с нужным расположением блоков.
3. В диалоге **Открыть шаблон блоков** в группе **Применить к** установите переключатель в одно из положений: **Всем страницам** (если вы хотите наложить шаблон на все страницы пакета) или **Выделенным страницам** (если вы хотите наложить шаблон только на выделенные страницы).
4. Нажмите кнопку **Открыть (Open)**.

Распознавание

Задача распознавания состоит в том, чтобы преобразовать отсканированное изображение в текст, сохранив при этом оформление страницы. Прежде чем приступить к распознаванию текста, необходимо установить язык распознавания. В этой главе описываются этот и другие параметры распознавания и приводятся описание ситуаций, в которых они используются.

Общая информация о распознавании

Внимание! Перед запуском распознавания проверьте установленные опции: язык распознавания, режим распознавания и тип печати распознаваемого текста.

Вы можете:

1. Распознать блок или несколько блоков, выделенных на изображении.
2. Распознать открытую страницу или все страницы, выделенные в окне **Пакет**.
3. Распознать все нераспознанные страницы пакета.
4. Распознать все страницы в фоновом режиме. В этом режиме возможно распознавание с одновременным редактированием уже распознанных страниц.
5. Распознать страницы в режиме распознавания с обучением. Данный режим применяется в основном для распознавания текстов, использующих декоративные шрифты, или для распознавания большого объема (более 100 страниц) документов низкого качества печати.
6. Распознать страницы одного пакета на нескольких компьютерах одновременно.

Чтобы запустить распознавание:

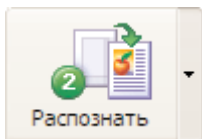
- Нажмите кнопку **2-Распознать** на панели **Scan&Read**.
- В меню **Процесс** выберите нужный вам пункт:

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

Распознать все – чтобы распознать все нераспознанные страницы пакета;

Распознать блок – чтобы распознать один или несколько блоков, выделенных 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.

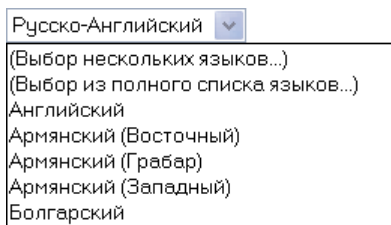
Замечание. При распознавании уже распознанной страницы перераспознаются только отредактированные и добавленные блоки.

Язык распознавания

ABBYY FineReader поддерживает распознавание как одноязычных, так и многоязычных (например, английско-французских) документов.

Чтобы указать язык распознаваемого текста:

- Выберите соответствующую строку в списке на панели **Распознавание**.



Если вы хотите распознать документ, написанный на нескольких языках:

- В списке языков на панели **Стандартная** выберите пункт **Выбор нескольких языков...**
- В открывшемся диалоге **Язык распознаваемого текста** укажите несколько языков. Для этого:
 - отметьте пункты с соответствующими названиями языков.

Совет: Если вы часто используете какую-либо комбинацию языков, то создайте новую группу, содержащую эти языки.

Замечания.

- Увеличение количества подключенных к распознаванию одного документа языков может привести к ухудшению качества распознавания. Не рекомендуется подключать более 2–3 языков.
- Перед запуском распознавания проверьте шрифты, выбранные на закладке **Сохранить** диалога **Опции**: они должны содержать все символы языка распознавания. В противном случае распознанный текст будет неправильно отображен в окне **Текст** (в словах на месте некоторых букв будут стоять значки «?» или «□»). Подробнее см. раздел «Шрифты, необходимые для корректного отображения символов поддерживаемых языков в редакторе ABBYY FineReader».

Если нужного языка нет в списке, возможны следующие варианты:

1. Данный язык не поддерживается системой ABBYY FineReader. Полный список языков приведен в разделе «Поддерживаемые языки распознавания».
2. Язык исключен из списка языков, показываемых на панели **Стандартная**. В этом случае в списке языков на панели **Стандартная** выберите пункт **Выбор из полного списка языков** и в открывшемся диалоге **Язык распознаваемого текста** укажите необходимый язык.
3. Язык был отключен при выборочной установке. Чтобы установить языки распознавания, в меню **Пуск** на **Панели Задач** выберите пункт **Настройка/Панель** управления, затем в диалоге **Панель управления** выберите пункт **Установка и удаление программ**. Из списка установленных программ выберите программу **ABBYY FineReader 8.0** и нажмите кнопку **Изменить**. В диалоге **Выборочная установка** выберите нужные языки. Далее следуйте инструкциям программы установки.

Замечание. При установке убедитесь в том, что вы указали ту же папку, в которую вы ранее установили ABBYY FineReader.

Чтобы подключить язык к списку показываемых языков (отключить от списка):

В диалоге **Редактор языков** (меню **Сервис>Редактор языков**) выберите язык, который вы хотите подключить/отключить, и отметьте пункт (снимите отметку) **Показывать в списке языков**.

Совет: Вы можете установить язык распознавания для отдельного блока. Для этого щелкните правой кнопкой мыши на блоке, для которого вы хотите установить язык распознавания, отличный от языка распознавания для всего текста, и в локальном меню выберите пункт **Свойства**. В открывшемся диалоге **Свойства** на закладке **Блок** в поле **Язык распознавания** выберите язык распознавания выделенного блока.

Тип печати входного текста

Для большинства текстов тип печати определяется автоматически. Этому соответствует значение **Авто**, установленное в группе **Тип печати** в диалоге **Дополнительные опции**. (Для того чтобы открыть этот диалог, щелкните по кнопке **Дополнительные опции** на закладке **Общие** диалога **Опции**).

При распознавании текстов, напечатанных на матричном принтере в черновом режиме или на пишущей машинке, можно добиться более высокого качества распознавания, установив правильный **Тип печати**:

- для текстов, напечатанных на пишущей машинке – значение **Пишущая машинка**;
- для текстов, напечатанных на матричном принтере – значение **Матричный принтер**.

software

Фрагмент страницы, напечатанной на матричном принтере в черновом режиме. На картинке видно, что штрихи букв состоят из отдельно стоящих точек.

software

Фрагмент страницы, напечатанной на пишущей машинке. Ширина букв одинакова (сравните, например, буквы «w» и «t»).

Чтобы поменять тип печати:

- в группе **Тип печати** в диалоге **Дополнительные опции** выберите требуемый пункт.

Замечание. После распознавания текстов, напечатанных на пишущей машинке или матричном принтере, не забудьте снова выбрать значение **Авто** при возвращении к типографскому тексту.

Другие опции распознавания

Режим распознавания

В ABBYY FineReader 8.0 предусмотрено два режима распознавания:

- **Тщательный**

Данный режим пригоден для распознавания как простых, так и сложных документов, в частности, документов, содержащих текст на цветном фоне, или растровый фон, а также для документов, содержащих таблицы, в том числе таблицы без линий сетки и таблицы с цветными ячейками.

Замечание. По сравнению с **Быстрым** режимом распознавания **Тщательный** режим требует больше времени, но обеспечивает лучшее качество распознавания.

- **Быстрый**

Данный режим рекомендуется для обработки больших объемов документов с простым оформлением и хорошим качеством печати.

Для выбора режима распознавания откройте диалог **Опции** (меню **Сервис>Опции**) и перейдите на закладку **Распознать**.

Распознавание PDF документов

Программа ABBYY FineReader извлекает текстовые данные из PDF документа и использует их при распознавании. Такой способ ускоряет обработку PDF документов в 2 – 3 раза. Однако, файлы формата PDF могут иметь нестандартные кодировки. Текст, извлеченный из таких файлов, содержит много ошибок. Если распознанный текст содержит большое количество неуверенно распознанных символов:

- на закладке **Распознать** диалога **Опции** (меню **Сервис>Опции**) в группе **Распознавание PDF документов** отметьте пункт **Распознать PDF как изображение** и распознайте документ повторно.

Распознавание штрих-кодов

Если вы хотите, чтобы ABBYY FineReader 8.0 автоматически распознавал штрих-коды, убедитесь в том, что в группе **Распознать** диалога **Дополнительные опции** отмечен пункт **Искать штрих-коды**. В этом случае штрих-коды, найденные на изображениях, будут выделены в отдельные блоки типа **Штрих-код**. При распознавании программа считает штрих-код и преобразует его в набор символов. Полный перечень штрих-кодов, поддерживаемых ABBYY FineReader 8.0, см. в разделе «Типы блоков».

Замечание. Для того чтобы открыть диалог **Дополнительные опции**, щелкните по кнопке **Дополнительные опции** на закладке **Общие** диалога **Опции** (меню **Сервис>Опции**).

Показ изображения при распознавании

Распознавание группы страниц происходит быстрее, если обрабатываемое изображение не показывается. Если вы хотите ускорить процесс распознавания:

- в группе **Вид** диалога **Дополнительные опции** снимите отметку с пункта **Показывать изображение при распознавании**. (Для того чтобы открыть этот диалог, щелкните по кнопке **Дополнительные опции** на закладке **Общие** диалога **Опции**).

Ориентация текста в блоке

Если программа некорректно распознала блоки с вертикальным направлением текста (текстовый блок, ячейки таблицы):

- Щелкните правой кнопкой мыши по блоку с вертикальным направлением текста и в локальном меню выберите пункт **Свойства**. В открывшемся диалоге **Свойства** выберите нужный вариант в списке **Направление текста** и распознajte изображение заново.



Инвертированный и отраженный текст в блоке

Если программа некорректно распознала блоки с инвертированным или отраженным текстом (текстовый блок, ячейка таблицы или вся таблица), то:

- Щелкните правой кнопкой мыши на нужном блоке и в локальном меню выберите пункт **Свойства**. В открывшемся диалоге **Свойства** выпадающем списке **Инверсия текста** выберите нужный пункт, а также отметьте опцию **Отраженный текст** и распознajte изображение заново.

Распознавание в фоновом режиме

Если вы хотите одновременно с распознаванием редактировать распознанные страницы, вы можете запустить распознавание в фоновом режиме:

- В меню **Процесс** выберите пункт **Запустить фоновое распознавание**. В строке состояния внизу экрана появится значок . Если для окна **Пакет** выбран режим показа свойства страниц (локальное меню, пункт **Окно Пакет>Со свойствами**), то напротив распознаваемой страницы в колонке **Открыта** появится значок .

В этом режиме распознавание автоматически возобновляется, как только в пакете появляются нераспознанные страницы.

Замечание. На многопроцессорных компьютерах режим «Фоновое распознавание» позволяет увеличить скорость распознавания пакетов, содержащих большое количество страниц.

Чтобы остановить Распознавание в фоновом режиме:

- В меню **Процесс** выберите пункт **Остановить фоновое распознавание**.

Внимание! В режиме работы **Распознавание в фоновом режиме** используются опции, установленные в программе до запуска фонового распознавания.

Распознавание с обучением

Как было отмечено, программа ABBYY FineReader позволяет вводить тексты разного качества, напечатанные практически любыми шрифтами. Тексты хорошего и среднего качества, а также шрифты обычного начертания распознаются без предварительного обучения.

Режим «Распознавание с обучением» используется для:

1. распознавания текстов, для набора которых использованы декоративные шрифты;
2. распознавания текстов, в которых встречаются специальные символы (например, отдельные математические символы);
3. распознавания большого объема (более 100 страниц) текста низкого качества.

В других случаях **Распознавание с обучением** использовать *не рекомендуется*, т.к. затраты на обучение будут больше, чем полученный выигрыш в качестве распознавания.

Обучение проводится при распознавании одной – двух страниц текста в специальном режиме. В результате создается эталон букв, встречающихся в тексте. Этот эталон в дальнейшем используется при распознавании основного объема текста. Некоторые пары или тройки символов в тексте могут склеиваться. Если при обучении вам не удастся переместить описывающий прямоугольник так, чтобы он заключал в себя один целый символ и не содержал при этом части соседних, то Вы можете обучить программу сочетанию символов, которые невозможно «расклеить». Такие неразделяемые сочетания двух или трех символов называются лигатурами. Это, например, такие сочетания, как ед, от, ff, ffi, ffl и другие.

-
- Внимание!**
1. Созданный эталон можно использовать только для распознавания текстов, использующих тот же шрифт и размер шрифта и отсканированных с тем же разрешением, что и документ, на котором данный эталон создавался.
 2. При удалении пакета эталон также удаляется. Вы можете сохранить созданный эталон для работы с другими пакетами. Для этого сохраните настройки пакета в файл набора опций (*.fbt).
 3. При переходе к распознаванию текстов, напечатанных другим шрифтом, не забудьте отключить эталон (на закладке **Распознать**, меню **Сервис>Опции** установите переключатель в положение **Не использовать пользовательский эталон**).
-

Последовательность действий при распознавании с обучением:

1. Установите режим **Распознавание с обучением** (на закладке **Распознать** диалога **Опции** меню **Сервис>Опции**). В строке состояния появится название эталона (по умолчанию default).
2. Нажмите кнопку **2-Распознать**.
3. Обучите эталон, распознав одну – две страницы в режиме **Распознавание с обучением**. Обучаемые символы заносятся в эталон, создаваемый системой по умолчанию. По окончании обучения система сохранит созданный эталон (default.ptn) в папке, где хранится пакет.
4. Отредактируйте эталон.
5. Отмените режим **Распознавание с обучением** (на закладке **Распознать** в группе **Обучение**

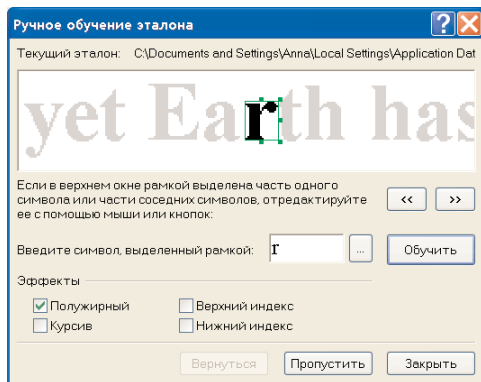
выберите пункт **Распознавание с пользовательским эталоном**).

6. Запустите распознавание основного текста, нажав на кнопку **2-Распознать**.



- Замечания.**
1. Чтобы создать несколько эталонов для одного пакета, воспользуйтесь диалогом **Редактор эталонов** (вызывается с закладки **Распознать** с помощью нажатия кнопки **Эталоны** или из меню **Сервис>Редактор эталонов**). Создайте новый эталон (нажмите в диалоге на кнопку **Новый**) и выберите его для работы (нажмите на кнопку **Выбрать**). Далее работа с созданным эталоном происходит так же, как и работа с default-эталомом, см. выше пп. 1–5.
 2. Если в процессе обучения было создано несколько эталонов, то подключается последний созданный эталон. Название подключенного эталона пишется в строке состояния. Чтобы подключить другой эталон для распознавания, в диалоге **Редактор эталонов** (меню **Сервис>Редактор эталонов**) в списке эталонов выберите эталон и нажмите кнопку **Выбрать**. На закладке **Распознать** в группе **Обучение** выберите опцию **Распознавание с пользовательским эталоном**.
 3. Если на закладке **Распознать** отмечена опция **Использовать встроенные эталоны**, то в режиме **Распознавание с обучением** программа предложит обучить только неуверенно распознанные символы. Если вы обучаете программу декоративным или нестандартным шрифтам (например, тибетскому) и при этом используете встроенные эталоны, программа может распознать часть символов неправильно, не предложив их выучить. В этом случае рекомендуется снять отметку с опции **Использовать встроенные эталоны** будет производиться обучение каждому символу.

Обучение эталона


1. Проверьте, чтобы на закладке **Распознать** (меню **Сервис>Опции**) в группе **Обучение** переключатель был установлен в положение **Распознавание с обучением**.
2. Нажмите кнопку **2-Распознать**. Программа начнет распознавание. Как только встретится символ, подлежащий обучению, откроется диалог **Ручное обучение эталона** с изображением этого символа.



Обучение символу



Описывающий прямоугольник в верхней части диалога должен содержать **один целый символ**. Если он содержит часть буквы или более одной буквы, то с помощью мыши или кнопок  и  прямоугольник можно передвинуть так, чтобы он охватывал одну целую букву.

Далее введите нужный символ и нажмите кнопку **Обучить**.

- Внимание!**
1. Обучение возможно только для символов, входящих в алфавит языка. Если вы обучаете программу символам, которые нельзя ввести с клавиатуры, то для их обозначения можно использовать комбинацию из двух символов, или вы можете скопировать требуемый символ из **Таблицы символов** (открывается при нажатии в диалоге **Ручное обучение эталона** кнопки ).
 2. Если в обучаемом тексте встречаются слова, набранные *курсивом* или **полужирным**, и вам важно сохранить гарнитуру шрифта в распознанном тексте, то при обучении таким символам в диалоге **Ручное обучение эталона** следует отметить пункты **Курсив** или **Полужирный** соответственно.
 3. При обучении следите за тем, чтобы изображениям заглавных букв соответствовали заглавные буквы, а изображениям строчных букв – строчные. Если в процессе обучения вы ошиблись, нажмите кнопку **Вернуться**, и охватывающий прямоугольник вернется на предыдущую позицию, а последняя обученная пара «изображение – символ» будет удалена из эталона. Кнопка **Вернуться** действует в пределах одного слова.

Обучение лигатурам

Лигатуры – это сочетания двух или трех символов, которые из-за особенностей их начертания невозможно разделить при обучении и которым поэтому сразу обучаются как комбинации символов. Обучение лигатурам происходит так же, как и обучение отдельным символам:

1. В строке для ввода символа введите нужное сочетание символов и нажмите кнопку **Обучить**.
2. Описывающий прямоугольник в верхней части диалога должен содержать сочетание целиком. Передвинуть прямоугольник можно с помощью мыши или используя кнопки  и .

В одном эталоне может содержаться до 1000 новых символов. Однако не следует создавать слишком много лигатур, т.к. это может отрицательно сказаться на качестве распознавания.

При обучении необходимо учесть следующие ограничения:

1. Изображения некоторых символов не различаются системой распознавания и сопоставляются с каким-то одним символом. Например, прямой (’), левый (‘) и правый (’) апострофы хранятся в эталоне как изображение прямого апострофа. Таким образом, в результате распознавания в тексте никогда не появится правый или левый апостроф, хотя при обучении были указаны именно эти символы.
2. Для некоторых изображений решение о том, какому символу в распознанном тексте сопоставить встретившееся конкретное изображение, принимается на основе общего анализа распознанного текста. Так, например, решение о том, является ли символ, обозначаемый «кружком», буквой «о» или цифрой «ноль», система принимает в зависимости от того, находятся ли рядом другие цифры или буквы.

Редактирование эталона

Прежде чем запускать распознавание с только что созданным эталоном, рекомендуется просмотреть эталон и, если потребуется, отредактировать его. Этим вы сведете к минимуму ошибки распознавания, которые могут возникнуть из-за неправильно обученного эталона.

Эталон должен содержать только целые символы или лигатуры. Символы, обрезанные по краям, и символы с неправильными подписями следует удалить из эталона.

Чтобы отредактировать эталон:

1. В меню **Сервис** выберите пункт **Редактор эталонов...**
2. В открывшемся диалоге **Редактор эталонов** выберите нужный эталон и нажмите кнопку **Редактировать...** Перед вами откроется диалог **Символы** пользовательского эталона.
3. Выбрав символ, нажмите кнопку **Свойства**, чтобы отредактировать подпись и указать правильное начертание: курсив, полужирный, верхний или нижний индексы, или нажмите кнопку **Удалить**, чтобы удалить неправильно обученные символы.

Пользовательские языки и группы языков

Вы можете использовать не только predetermined языки и группы языков, но и создать новый язык или объединить существующие языки в новую группу и при распознавании подключить именно их.

Когда требуется создание нового языка:

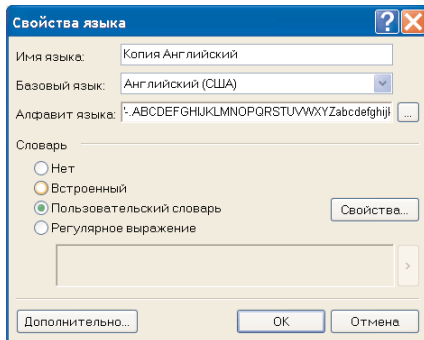
1. Для подключения пользовательского словаря. Например, необходимо распознать русский текст, содержащий аббревиатуры. Вы можете создать словарь аббревиатур и подключить его к пользовательскому языку. На основе русского языка с подключенным системным словарем и языка, созданного Вами с подключенным словарем аббревиатур, вы можете создать группу для дальнейшего ее использования при распознавании ваших текстов.
2. Для распознавания документов специального вида. Например:
 - Страница содержит перечень артикулов, состоящий из цифр и нескольких букв. Вы можете создать новый язык, включив в него минимально необходимый набор символов, и использовать его для распознавания данного типа документов.
 - Документ использует только заглавные буквы английского языка. В этом случае для повышения качества распознавания следует исключить из распознавания символы, которые заведомо не могут встретиться в тексте, в данном случае – все строчные буквы.

Когда требуется создание группы языков:

- Если вы часто используете какую-нибудь комбинацию языков.

Создать язык или группу языков можно из диалога **Редактор языков** (меню **Сервис**, пункт **Редактор языков**).


Создание нового языка



Чтобы создать новый язык распознавания:

1. В меню **Сервис** выберите пункт **Редактор языков...**
2. Нажмите кнопку **Новый**. В открывшемся диалоге установите переключатель в положение **Создать новый язык на основе существующего** и выберите язык, на основе которого вы создадите новый.
3. Перед вами откроется диалог **Свойства языка**.

При создании нового языка необходимо задать следующие параметры (все параметры задаются в диалоге Свойства языка):

1. Имя нового языка.
2. В поле **Алфавит языка** указан алфавит языка, на основе которого вы создаете новый язык. Если требуется, отредактируйте алфавит, нажав на кнопку .
3. Словарь, который будет использоваться системой при распознавании и проверке распознанного текста. Возможны следующие варианты:
 - **Нет** (не подключать словарь к языку).
 - **Встроенный словарь** (используется словарь, поставляемый с программой).
 - **Пользовательский словарь**.
Чтобы наполнить словарь, а также чтобы подключить старый пользовательский словарь или текстовый файл в Windows (ANSI)–кодировке (слова должны быть разделены пробелами или другими символами, не включенными в алфавит), нажмите кнопку **Редактировать**.

Замечание. Словарные слова пользовательского языка считаются правильными, если в тексте они встретились с той капитализацией, в которой они заданы в словаре, а также в каком-либо стандартном виде: всеми маленькими, всеми большими буквами или с большой буквы:

Вид слова в словаре:	Допустимые варианты написания слова в тексте:
abc	abc, Abc, ABC
Abc	abc, Abc, ABC
ABC	abc, Abc, ABC
aBc	aBc, abc, Abc, ABC

- **Регулярное выражение** (задается грамматика нового языка; подробнее о регулярных выражениях см. «Регулярные выражения»).

- Замечания.**
1. Нажав в диалоге **Свойства языков** на кнопку **Дополнительно**, вы можете указать дополнительные свойства нового языка, например, игнорируемые внутри слова символы или исключенные из распознавания символы и т.д.
 2. По умолчанию пользовательский язык будет сохранен в папке пакета. Версия ABBYY FineReader Corporate Edition позволяет указать другую папку. Более подробно о коллективной работе с пользовательскими языками и словарями к ним см. раздел «Коллективная работа с пользовательскими языками и пользовательскими словарями».

Проверка и редактирование текста

По завершению распознавания его результат появляется в окне **Текст**. Окно **Текст** – это встроенный редактор программы ABBYY FineReader; в нем вы можете проверить результаты распознавания и отредактировать распознанный текст.

Одна из возможностей текстового редактора ABBYY FineReader – это встроенная проверка орфографии (список языков, для которых поддерживается проверка орфографии, приведен в разделе «Поддерживаемые языки распознавания»).

Система встроенной проверки орфографии позволяет:

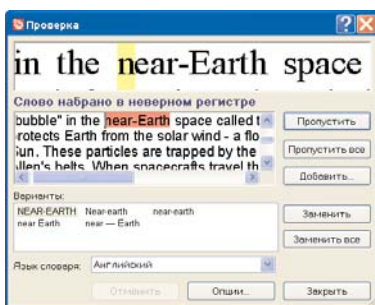
1. Находить неуверенно распознанные слова (слова, в которых есть неуверенно распознанные символы).
2. Находить орфографические ошибки (неправильно написанные слова).
3. Добавлять неизвестные системе ABBYY FineReader слова в словарь для того, чтобы впоследствии они распознавались уверенно.

Проверка распознанного текста

Неуверенно распознанные символы и слова, которых нет в словаре, выделяются различными цветами. По умолчанию для выделения неуверенно распознанных символов используется голубой цвет, для несловарных слов – розовый. Чтобы изменить цвета выделения:

- На закладке **Вид** диалога **Опции** (меню **Сервис>Опции**) в поле **Объекты** выберите пункт **Неуверенно распознанный символ (Несловарное слово)** и в поле **Цвет** – цвет подсветки.

Чтобы проверить результаты распознавания:



1. Нажмите кнопку **3-Проверить** на панели **Scan&Read** (или выберите пункт **Проверка** в меню **Сервис**).
2. Откроется диалог **Проверка**.
3. В диалоге **Проверка** три окна. Верхнее окно – аналог окна **Крупный план** программы ABBYY FineReader, в нем показано изображение слова с возможной ошибкой. Среднее окно показывает само слово с возможной ошибкой, в строке над этим окном выводится название типа ошибки. В нижнем окне, **Варианты**, предлагаются варианты замены данного

слова (если таковые имеются). Для вариантов используется словарь, указанный в поле **Язык словаря**. Вы можете использовать любой словарь из предложенного списка.

Замечание. Для удобства проверки и редактирования текста диалог **Проверка** может быть увеличен. Для этого достаточно поместить указатель мыши на границу диалога (указатель примет вид двусторонней стрелки). Перетаскивая границу, увеличьте или уменьшите размеры диалога.



4. У вас есть следующие возможности:

- Нажмите кнопку **Пропустить**, чтобы оставить слово, как есть.
- Нажмите кнопку **Пропустить все**, чтобы оставить все такие слова в распознанном тексте, как есть.

Замечание. При нажатии на кнопки **Пропустить** или **Пропустить все** со слова снимается «признак неуверенности» – считается, что слово не содержит нераспознанных или неуверенно распознанных символов; снимается также и визуальная подсказка таких символов в тексте. При экспорте в PDF в режиме **Заменять неуверенно распознанные слова их изображениями** слова, с которых при проверке был снят «признак неуверенности» (при нажатии кнопок **Пропустить** и **Пропустить все**), не будут заменены на свои изображения.

- Выберите вариант для замены и нажмите кнопку **Заменить** или **Заменить все**, чтобы заменить текущее слово или все такие слова в тексте. Если в окне **Варианты** нет правильного варианта для замены слова, отредактируйте слово или несколько соседних слов в среднем окне (обратите внимание, что при переходе в режим редактирования меняется назначение некоторых кнопок диалога и соответственно – надписи на них) и нажмите кнопку **Подтвердить**, чтобы зафиксировать сделанные изменения и перейти к следующему фрагменту текста.
- Нажмите кнопку **Добавить...**, чтобы добавить слово в словарь. В этом случае при дальнейшей проверке орфографии, если это слово (или одна из его форм) встретится в тексте, оно не будет считаться ошибочным.
- Нажмите кнопку **Опции...**, чтобы установить опции проверки распознанного текста.
- Нажмите кнопку **Заккрыть**, чтобы закрыть диалоговое окно.

Перемещение по неуверенно распознанным словам

Для более быстрой проверки результатов распознавания вы можете воспользоваться кнопками  или  для перемещения к следующему или, соответственно, предыдущему неуверенно распознанному слову.

Также для перемещения по неуверенно распознанным словам можно использовать горячие клавиши: **F4 (SHIFT F4)**.

Опции проверки и редактирования текста

Устанавливаются на закладке **Проверить** диалога **Опции** (меню **Сервис>Опции**) и в диалоге **Дополнительные опции** (эти опции отмечены особо).

Уровень выделения ошибок

В списке **Уровень выделения ошибок** возможно выбрать следующие значения:

- **Нет** – ошибки распознавания не выделяются.
- **Стандартный** – цветом выделяются нераспознанные и неуверенно распознанные символы.
- **Максимальный** – помимо нераспознанных и неуверенно распознанных символов, цветом выделяются слова, которые отсутствуют в словаре языка распознавания.

Замечание. Количество показываемых в окне **Текст** ошибок будет изменено после повторного распознавания документа.

Останавливаться на неуверенно распознанных словах

При проверке орфографии система останавливается на словах, в которых были неуверенно распознаны какие-либо буквы.

Останавливаться на несловарных словах

Эта опция позволит вам проверить слова, которых нет в словаре и которые могли быть распознаны программой неверно.

Останавливаться на сложных словах

При проверке орфографии система останавливается на словах, которых нет в словаре, но которые могут быть построены по имеющимся морфологическим моделям или которые могут быть составлены из имеющихся в словаре слов.

Замечание. Данная опция находится в диалоге **Дополнительные опции**. Для того чтобы открыть этот диалог, щелкните по кнопке **Дополнительные опции** на закладке **Общие** диалога **Опции** (меню **Сервис>Опции**).

Пропускать слова с цифрами и другими неалфавитными символами

При проверке орфографии слова, внутри которых встречаются цифры или какие-либо другие символы, не входящие в алфавит языка распознавания, не считаются ошибочными, если эти слова не содержат неуверенно распознанные символы.

Корректировать пробелы до и после знаков препинания

Если этот пункт отмечен, система не будет останавливаться на тех фрагментах, где неверно расставлены пробелы до и после знаков препинания, а исправит их автоматически.

Замечание. Данная опция находится в диалоге **Дополнительные опции**. Для того чтобы открыть этот диалог, щелкните по кнопке **Дополнительные опции** на закладке **Общие** диалога **Опции** (меню **Сервис>Опции**).

Пополнение и удаление слов из пользовательского словаря

Добавление слова в пользовательский словарь

Добавление слов в словарь – один из способов повышения качества распознавания, т.к. при распознавании система проверяет слова по словарю. В словарь имеет смысл добавлять часто встречающиеся слова (например, термины, сокращения, названия и т.д.). При добавлении слова система строит его парадигму (совокупность всех форм данного слова). Благодаря этому система может уверенно распознавать не только ту форму, которая уже однажды встретилась в тексте и была добавлена в словарь, но и все формы добавленного слова.

Чтобы добавить слово в словарь во время проверки орфографии:

- Нажмите кнопку **Добавить** в диалоге **Проверка**.

В диалоге **Начальная форма** вы должны установить следующие параметры:

1. **Часть речи** (Существительное, Прилагательное, Глагол, Неизменяемое слово).
2. Если слово всегда пишется с большой буквы, отметьте пункт **Имя собственное**. Если вы добавляете слово, являющееся сокращением, отметьте пункт **Аббревиатура**.
3. Начальная форма слова.

Нажмите **ОК**. Откроется диалог **Построение парадигмы**, в котором пользователю предлагаются вопросы, по ответам на которые строится парадигма слова. Для ответов на вопросы нажимайте кнопки **Да** или **Нет**. Если вы ошиблись при ответе на вопрос, нажмите кнопку **Сначала**, чтобы иметь возможность ответить на тот же вопрос по-другому. Система покажет построенную парадигму в диалоге **Парадигма**.

- Замечания.**
1. Если вы хотите, чтобы при добавлении слова в английский словарь парадигма не строилась (слово добавлялось бы как неизменяемое), то на закладке **Проверить** диалога **Опции** (меню **Сервис>Опции**) отметьте опцию **Добавлять слово как неизменяемое**.
 2. Вы можете добавлять слова во время просмотра списка добавленных слов. Для этого в меню **Сервис** выберите пункт **Просмотр словарей**. В открывшемся диалоге **Выбор словаря** выберите язык и нажмите кнопку **Просмотр**. Откроется словарь со списком добавленных слов. Вы можете, нажав на кнопку **Добавить**, добавить набранное слово.
 3. Построение парадигмы при добавлении слова в словарь возможно для английского, испанского, итальянского, немецкого (новая и старая орфография), русского, украинского и французского языков. Если добавляемое слово уже есть в словаре, система предупредит вас об этом. В этом случае вы можете посмотреть его парадигму. Если существующая парадигма вас не устраивает, можете создать другую (кнопка **Добавить** в диалоге **Добавить слово**).

Совет: Вы можете импортировать пользовательские словари, созданные в предыдущих версиях ABBYY FineReader (5.0, 6.0 и 7.0).

Чтобы импортировать словарь:

1. В меню **Сервис** выберите пункт **Просмотр словарей**. Далее выберите язык словаря и нажмите кнопку **Просмотр**.

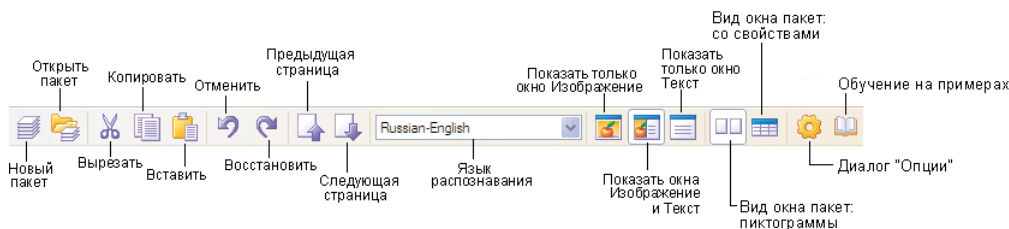
2. В открывшемся диалоге **Словарь языка** нажмите кнопку **Импорт** и выберите файлы с расширениями (*.pmd, *.txt или *.dic).

Чтобы удалить слово из словаря:

1. В меню **Сервис** выберите пункт **Просмотр словарей**. Выберите язык словаря и нажмите кнопку **Просмотр**.
2. В открывшемся диалоге выделите слово и нажмите кнопку **Удалить**.

Редактирование текста в редакторе ABBYY FineReader

Совет: Если символы в текстовом окне ABBYY FineReader отображаются некорректно (в словах на месте некоторых букв стоят значки «?» или «□»), это означает, что шрифты, выбранные на закладке **Сохранить**, не содержат всех символов, входящих в распознаваемый язык. Установите шрифт (группа **Шрифты**, меню **Сервис>Опции**), поддерживающий все символы языка документа (например, Arial Unicode MS или Bitstream Cyberbit), и распознайте документ снова. Более подробно о шрифтах для поддерживаемых языков см. раздел «Шрифты, необходимые для корректного отображения символов поддерживаемых языков в редакторе ABBYY FineReader».



Распознанный текст отображается в окне **Текст**. При сохранении текста оформление документа сохраняется в соответствии с опциями форматирования, устанавливаемыми на закладке **Сохранить** и в диалоге **Форматы**.



Неуверенно распознанные символы выделяются цветом. Вы можете отменить выделение слов с неуверенно распознанными символами: на закладке **Вид** (меню **Сервис>Опции**) снимите отметку с пункта **Выделять неуверенно распознанные символы**.

Редактор ABBYY FineReader поддерживает два режима просмотра документа: режим с полным сохранением исходного оформления документа и черновой режим просмотра.

В режиме с полным сохранением исходного оформления документа блоки с распознанным текстом, таблицами и картинками отображаются именно там, где они были в исходном изображении. Таким образом, в этом режиме сохраняется исходный вид изображения: колонки, таблицы, картинки и буквицы (большие буквы на несколько строк в абзаце). Блок, в котором находится курсор, является активным. При перемещении курсора с помощью клавиатуры последовательность обхода блоков определяется их порядковыми номерами. Если текст блока не уместается внутри его границ (например, в результате редактирования распознанного текста), то часть текста неактивного блока может быть не видна на экра-

не. Признаком этого являются специальные красные маркеры на границах блока. Когда блок становится активным, его границы расширяются таким образом, чтобы весь его текст был виден на экране целиком.

В черновом режиме редактора не отображаются: левый отступ; выравнивание абзаца (все абзацы прижаты к левому краю); цвет и фон символов. Для показа текста используется шрифт одного размера (по умолчанию 12 пт.). При этом сохраняется шрифт и форматирование текста: полужирный, курсив, подчеркнутый, верхние и нижние индексы.

Переход из одного режима в другой осуществляется нажатием кнопок  (режим отображения исходного оформления документа) и  (черновой режим редактора) в окне **Текст**.

Вы можете поменять установленный по умолчанию размер шрифта для показа в черновом режиме. Для этого:

1. В меню **Сервис** выберите пункт **Опции**.
2. На закладке **Вид** укажите нужный размер шрифта в пункте **Размер шрифта в черновом режиме**.

Редактор системы ABBYY FineReader предоставляет следующие возможности по редактированию текста:

Копирование, вырезание, вставка выделенных фрагментов текста

1. Перед применением команд копирования, вырезания или вставки выделите нужный фрагмент текста.
2. Прделайте одно из действий:

Чтобы скопировать выделенный текст:

Кнопка **Копировать** 

- Нажмите кнопку **Копировать** на инструментальной панели **Стандартная**.
- В меню **Правка** или в локальном меню выберите команду **Копировать**.
- Нажмите клавиши **ALT+C**.

Чтобы вырезать фрагмент текста:

Кнопка **Вырезать** 

- Нажмите кнопку **Вырезать** на инструментальной панели **Стандартная**.
- В меню **Правка** или в локальном меню выберите команду **Вырезать**.
- Нажмите клавиши **ALT+X**.

Чтобы вставить скопированный (вырезанный) текст:

Кнопка **Вставить** 

- Нажмите кнопку **Вставить** на инструментальной панели **Стандартная**.
- В меню **Правка** или в локальном меню выберите команду **Вставить**.
- Нажмите клавиши **ALT+V**.

Поиск и замена указанного фрагмента текста

Чтобы найти определенный фрагмент в редактируемом тексте:




1. Выполните одну из следующих операций:
 - В меню **Правка** выберите пункт **Найти**.
 - Нажмите клавиши **ALT+F**.
2. В открывшемся диалоге **Поиск** в строке **Найти** введите слово или фрагмент текста, который вы хотите найти, и установите параметры поиска.

Замечание. Чтобы повторить поиск того же слова с теми же параметрами, нажмите клавишу F3.

Чтобы найти и заменить определенный фрагмент в редактируемом тексте:





1. Выполните одну из следующих операций:
 - В меню **Правка** выберите пункт **Заменить**.
 - Нажмите клавиши **ALT+H**.
2. В открывшемся диалоге **Заменить** в строке **Найти** введите слово или фрагмент текста, который вы хотите найти, в строке **Заменить** на укажите текст замены, и установите параметры поиска.

Изменение начертания шрифта

1. Установите курсор на слово, которое вы хотите изменить, или выделите участок текста, для которого вы хотите изменить шрифт.
2. Выполните одну из следующих операций:
 - Нажмите на одну из кнопок , ,  на панели **Форматирование**.
 - Щелкните правой кнопкой мыши в окне **Текст** и в локальном меню выберите пункт **Свойства символа**. В открывшемся диалоге **Свойства**, на закладке **Шрифт** выберите название шрифта и установите нужные вам параметры шрифта.
 - Нажмите клавиши **ALT+B** – полужирный, **ALT+I** – курсив, **ALT+U** – подчеркивание.

Замечание. Устанавливаемые в диалоге **Шрифты** цвет фона блока и форматирование текста малыми прописными (отображение всех строчных букв в выделенном тексте как прописных букв уменьшенного размера) не отображаются в окне **Текст**. Вы увидите эти изменения при сохранении документа в формате приложений, поддерживающих указанные типы форматирования текста.

Изменение выравнивания текста

1. Выделите текст, который необходимо выровнять.
2. Выполните одну из следующих операций:
 - Нажмите на одну из кнопок , , , , на панели **Форматирование**.
 - Щелкните правой кнопкой мыши в окне **Текст** и в локальном меню выберите пункт **Свойства символа**. В открывшемся диалоге **Свойства** на закладке **Шрифт**, в поле **Выравнивание**

выберите нужный пункт.

Отмена и восстановление действий

Выполните одну из следующих операций:

Чтобы отменить совершенное действие:

Кнопка **Отменить** 

- Нажмите кнопку **Отменить** на инструментальной панели **Стандартная**.
- В меню **Правка** выберите пункт **Отменить**.
- Нажмите клавиши **ALT+Z**.

Чтобы восстановить отмененное действие:

Кнопка **Восстановить** 

- Нажмите кнопку **Восстановить** на инструментальной панели **Стандартная**.
- В меню **Правка** выберите пункт **Восстановить**.
- Нажмите клавиши **ALT+Y**.

Редактирование таблиц

Чтобы объединить содержимое ячеек или строк:



Удерживая клавишу **CTRL**, выделите на изображении ячейки или строки, которые вы хотите объединить. В меню **Изображение>Ячейки таблицы** выберите пункт **Объединить ячейки** или **Объединить строки**.


Чтобы разбить содержимое ячеек:

В меню **Изображение>Ячейки таблицы** выберите пункт **Разбить ячейки**.

Внимание! Команда применяется только к ранее объединенным ячейкам таблицы.

Чтобы разбить содержимое строки или столбца:

1. На панели **Изображение** выберите инструмент  или .
2. Вставьте горизонтальную/вертикальную линию в строку/столбец таблицы, содержимое которых вы хотите разделить.

Замечание. Воспользовавшись инструментом  или командой меню **Объединить строки** (меню **Изображение>Ячейки таблицы**), вы можете объединить содержимое строк (столбцов).

Чтобы удалить содержимое ячейки:

В окне **Текст** выделите ячейку (или несколько ячеек), содержимое которой вы хотите удалить, и нажмите клавишу **DEL**.

Редактирование и создание гиперссылок

Гиперссылка — это связь между веб-страницами или файлами. Гиперссылка состоит из текста, отображаемого в документе, и адреса. Объект, указанный в адресе гиперссылки, отображается в веб-обозревателе, открывается или запускается в зависимости от типа этого объекта. Часто гиперссылка указывает на веб-страницу, но может также указывать на рисунок, мультимедийный файл, документ, адрес электронной почты или программу.

При распознавании документа FineReader 8.0 находит ссылки на:

- файлы, например, file:///D:/MyDocuments/ABBYY FineReaderGuide.pdf;
- протокол передачи файлов (FTP), например, ftp://ftp.abbyy.ru/;
- сайты, например, http://www.abbyy.com;
- адреса электронной почты, например, mailto:office@abbyy.com.

Программа распознает текст и восстановит адрес гиперссылки в выходном документе. Все найденные гиперссылки выделяются синим цветом и подчеркиванием.


Замечание. Цвет выделения гиперссылок можно изменить в диалоге **Опции** на закладке **Распознать**.

При просмотре распознанного документа в окне **Текст** при наведении курсора на гиперссылку возникает всплывающая подсказка с адресом гиперссылки. Для перехода по ссылке нажмите клавишу **Ctrl** и щелкните по ссылке левой кнопкой мыши.

Редактор системы ABBYY FineReader позволяет:


- Изменить текст и адрес найденной гиперссылки,
- Вставить гиперссылку в документ,
- Удалить гиперссылку из документа.

Изменение текста и адреса гиперссылки

1. Выделите текст гиперссылки. Если текст не был выделен, в качестве текста гиперссылки используется адрес гиперссылки;
2. Нажмите кнопку , расположенную на панели инструментов окна **Текст**;
3. В открывшемся диалоге **Редактирование гиперссылки** в поле **Текст гиперссылки** укажите новый текст;
4. В этом же диалоге в группе **Связать с:** укажите новый тип адреса гиперссылки:
 - опция **Веб-страница** позволяет создать гиперссылку, указывающую на веб-страницу. Затем в поле **Адрес** укажите протокол и имя веб-страницы (например, <http://www.abbyy.com>);
 - опция **Файл** позволяет создать гиперссылку, указывающую на файл. При выборе данной опции открывается диалог **Открыть**, в котором необходимо выбрать имя файла, на который будет указывать ссылка;

- опция **Электронная почта** создает гиперссылку, позволяющую пользователю отправить письмо по заданному адресу. В поле **Адрес** укажите адрес электронной почты (например, *mailto:office@abbyy.com*).

Вставка гиперссылки в документ

1. Выделите текстовый фрагмент.
Если текст не был выделен, в качестве текста гиперссылки используется адрес гиперссылки;
2. Нажмите кнопку , расположенную на панели инструментов окна **Текст**;
3. В открывшемся диалоге **Редактирование гиперссылки** в поле **Текст гиперссылки** укажите текст гиперссылки;
4. В этом же диалоге в группе **Связать с:** укажите тип адреса гиперссылки:
 - опция **Веб-страница** позволяет создать гиперссылку, указывающую на веб-страницу. Затем в поле **Адрес** укажите протокол и имя веб-страницы (например, *http://www.abbyy.com*);
 - опция **Файл** позволяет создать гиперссылку, указывающую на файл. При выборе данной опции открывается диалог **Открыть**, в котором необходимо выбрать имя файла, на который будет указывать ссылка;
 - опция **Электронная почта** создает гиперссылку, позволяющую пользователю отправить письмо по заданному адресу. В поле **Адрес** укажите адрес электронной почты (например, *mailto:office@abbyy.com*).

Удаление гиперссылки из документа

Щелкните по ссылке в окне **Текст** правой кнопкой мыши. В открывшемся контекстном меню выберите команду **Удалить гиперссылку**.

Сохранение во внешние редакторы и форматы

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

Список поддерживаемых приложений:

Microsoft Word 6.0, 7.0, 97 (8.0), 2000 (9.0), 2002 (10.0) и 2003 (11.0); Microsoft Excel 6.0, 7.0, 97 (8.0), 2000 (9.0), 2002 (10.0) и 2003 (11.0); Microsoft PowerPoint 2002 (10.0) и 2003 (11.0); Corel WordPerfect 7.0, 8.0, 9.0, 10.0 (2002), 11.0 (2003) и 12.0; Lotus Word Pro 9.5, 97 и Millennium Edition; StarWriter 4.x, 5.x и 6.x, Adobe Acrobat/Reader (все версии).

Общая информация по сохранению распознанного текста

Вы можете:

- 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.

Чтобы сохранить распознанный текст:

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

Замечание. 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.

Опции сохранения распознанного текста

Опции сохранения распознанного текста устанавливаются на закладке **Сохранить** диалога **Опции** (меню **Сервис>Опции**). Также часть этих опций можно установить в диалогах **Мастер сохранения результатов** и **Сохранить страницы**.

Используемые шрифты (при сохранении в форматах RTE, DOC, Word XML, PPT или HTML)

При сохранении в файл или при передаче распознанного текста в приложение используются шрифты, установленные на закладке **Сохранить** в группе **Шрифты**. Вы можете изменить шрифты в окне **Текст** или, выбрав другие шрифты на закладке **Сохранить**, распознать изображение снова.

Сохранять все или только выделенные страницы

Позволяет сохранить все страницы пакета или только выделенные. При сохранении или экспорте части страниц сначала выделите их в окне **Пакет**.

Режимы передачи страниц пакета (при сохранении нескольких страниц пакета)

- **Каждая страница в отдельный файл** – позволяет сохранить каждую страницу пакета в отдельный файл. В этом случае к названию файла в конец автоматически добавляется порядковый номер страницы в пакете.
- **Сохранять с именами исходных изображений** – позволяет сохранить распознанные страницы в отдельные файлы выбранного формата, причем имена этих файлов будут совпадать с именами соответствующих файлов изображений.

Замечания.

1. Страницы, не связанные с исходными файлами изображений (например, полученные при сканировании), в этом режиме игнорируются, т.е. не сохраняются; при наличии в пакете таких страниц ABBYY FineReader выдает перед сохранением предупредительное сообщение.
2. Если несколько страниц пакета, следующие подряд в группе выбранных сохраняемых страниц, имеют в качестве исходных изображений файлы с одинаковыми именами (или один и тот же файл), то такие страницы считаются имеющими общий источник (это могут быть различные страницы многостраничного TIFF

или страницы, появившиеся в результате разрезания некоторого исходного изображения). Распознанный текст с таких страниц будет сохранен в один файл. Если же страницы с одинаковыми названиями исходных изображений в выборке следуют не подряд, такие страницы сохраняются в разные файлы, при этом к имени файла добавляется различающий индекс следующего вида: _1, _2, _3 и т.д.

- **Пофайловое деление по пустым страницам** – весь пакет рассматривается как совокупность групп страниц, разделенных пустыми страницами. Для каждой такой группы при сохранении образуется отдельный файл, в который последовательно записываются страницы группы; имена формируемых в этом режиме файлов образуются добавлением к указанному пользователем имени файла индексов следующего вида: –1, –2, –3 и т.д.
- **Все страницы в один файл** – позволяет сохранить все (или выделенные) страницы пакета в один файл.

Сохранение распознанного текста в форматах RTF, DOC, Word XML

Внимание! Сохранение в формат Word XML возможно только для Microsoft Word 2003.

Опции сохранения для форматов RTF, DOC и Word XML устанавливаются на закладке **RTF/DOC/Word XML** диалога **Форматы**. Для того чтобы открыть этот диалог, щелкните по кнопке **Форматы**, расположенной на закладке **Сохранить** диалога **Опции** (меню **Сервис>Опции**), или нажмите **CTRL+SHIFT+X**.

Замечание. При сохранении результатов распознавания в форматы RTF, DOC и Word XML ABBYY FineReader использует шрифты, выбранные на закладке **Сохранить** диалога **Опции**, или шрифты, установленные в окне **Текст**.

Следующие опции сохранения помогут вам настроить процесс сохранения и получить документы, пригодные для дальнейшего редактирования и использования. В программе ABBYY FineReader предусмотрены следующие настройки:

- Режим сохранения оформления,
- Размер бумаги,
- Параметры сохранения текста,
- Параметры сохранения иллюстраций.

Режим сохранения оформления

Режим сохранения оформления выбирается в выпадающем списке **Оформление**. Выбор режима зависит от того, как вы планируете работать с созданным документом в дальнейшем:

Оформление оригинала

Данный режим позволяет получить документ, внешний вид которого будет полностью идентичен оригиналу. *Данный режим используется по умолчанию.*

Колонки, таблицы, абзацы, шрифты

В данном режиме практически полностью сохраняется оформление оригинала, хотя внешний вид создаваемого документа может немного отличаться от оригинала.

Замечание. Данный режим сохранения рекомендуется для документов сложного оформления, например, рекламных брошюр. Электронный документ будет почти неотличим от бумажного оригинала, однако, этот режим не предполагает внесения значительных правок в текст и оформление.

Таблицы, абзацы, шрифты

Данный режим позволяет полностью сохранить распознанный текст, однако, оформление оригинала воспроизведено не будет. Этот режим рекомендуется использовать, если требуется лишь текст документа или вы планируете внести значительные изменения в оформление электронного документа.

Размер бумаги

Вы можете указать размер бумаги, который будет использован при сохранении результатов распознавания в форматах RTF, DOC и Word XML. Размер бумаги задается в выпадающем списке **Размер бумаги**.

-
- Совет.**
1. Если вы не нашли нужного размера бумаги в списке, вы можете создать свой – пользовательский – размер бумаги и использовать его для сохранения распознанного текста. Для того чтобы создать пользовательский размер бумаги, выберите пункт **Добавить пользовательский размер** и в открывшемся диалоге укажите название, высоту и ширину страницы.
 2. Для того чтобы распознанный текст полностью уместился на странице при сохранении, отметьте опцию **Автоматически увеличивать размер страницы**. В этом случае выбор размера бумаги будет выполнен автоматически.
-

Параметры сохранения текста

Замечание. Значения по умолчанию, установленные для опций сохранения текста, зависят от выбранного режима сохранения оформления.

- **Сохранять деление на строки**
Указывает, что при сохранении распознанного текста в формате RTF/DOC/Word XML разделение на строки будет таким же, как и на исходном изображении.
- **Сохранять деление на страницы**
Указывает, что при сохранении распознанного текста в формате RTF/DOC/Word XML разделение на страницы будет таким же, как и на исходном изображении.
- **Сохранять цвет текста**
Позволяет сохранить исходный цвет букв.

Замечание. Так как программа Word версий 6.0, 7.0, 97 (8.0) использует ограниченное количество цветов для выделения текста и фона, то оригинальные цвета документа могут быть

заменены на цвета, допустимые палитрой Word. Программа Word 2000 (9.0) или более поздних такого ограничения не имеет: цвета документа сохраняются полностью.

- **Удалять мягкий перенос**

Позволяет удалить символ мягкого переноса (–) из распознанного текста. При этом, если отмечена опция **Сохранять деление на строки**, символы мягкого переноса будут заменены на символы переноса (–).

- **Выделять неуверенно распознанные символы**

Отметьте эту опцию, если вы собираетесь редактировать распознанный текст в Microsoft Word, а не в окне **Текст** ABBYY FineReader. Если эта опция отмечена, все неуверенно распознанные символы будут выделены цветом в программе Microsoft Word.

Совет. Цвет выделения неуверенно распознанных символов можно изменить на закладке **Вид** в диалоге **Опции** (меню **Сервис>Опции**).

- **Совместимость с Microsoft Word 95**

Позволяет получить документ, который можно будет открывать и редактировать в программе Microsoft Word 95.

Замечание. При сохранении в формате, совместимом с Microsoft Word 95, иллюстрации сохраняются только в формате BMP.

- **Открывать окно Крупный план в Microsoft Word 2003 (только для формата Word XML)**

Указывает, что при сохранении в формате Word XML распознанное изображение можно будет просматривать в Word в окне **Крупный план**. В нем будет отображаться увеличенное изображение редактируемой строки или участка документа.

Параметры сохранения иллюстраций

Для того чтобы при сохранении распознанного текста сохранить и иллюстрации, необходимо отметить опцию **Сохранять картинки** в группе **Картинки**.

Если распознанный текст содержит большое количество иллюстраций, то размер конечного файла может быть очень большим. Размер файла документа можно уменьшить, выбрав в группе **Картинки** подходящее значение качества и формат сохраняемых иллюстраций.

Качество

Выбор уровня качества зависит от того, как вы будете использовать документ в дальнейшем. Из выпадающего списка **Качество** выберите:

- **Высокое**, если создаваемый документ предназначен для печати.
- **Среднее**, если создаваемый документ предназначен для просмотра на экране.
- **Низкое**, если создаваемый документ предназначен для размещения на веб-страницах.

Чем выше значение выбрано в списке **Качество**, тем выше будет качество иллюстраций. Размер файла также зависит от выбранного значения: чем оно выше, тем больше размер файла.

Совет. Чтобы подобрать оптимальное сочетание качества иллюстраций и размера файла, сохраните распознанный текст и иллюстрации в нескольких вариантах – с разными значениями из списка **Качество** – и откройте их в какой-нибудь графической программе.

Формат

При сохранении иллюстраций программа ABBYY FineReader автоматически подбирает наиболее подходящий формат. В этом случае в выпадающем списке **Формат** установлено значение **(Авто)**.

Если вы хотите выбрать формат самостоятельно, в выпадающем списке **Формат** выберите одно из следующих значений:

- **JPEG**, цветной (для фотографий),
- **JPEG**, серый (для фотографий),
- **PNG**, цветной (для чертежей и схем),
- **PNG**, серый (для чертежей и схем),
- **PNG**, черно-белый.

Сохранение распознанного текста в формате XLS

Опции сохранения для формата XLS устанавливаются на закладке **XLS** диалога **Форматы**. Для того чтобы открыть этот диалог, щелкните по кнопке **Форматы**, расположенной на закладке **Сохранить** диалога **Опции** (меню **Сервис>Опции**), или нажмите **CTRL+SHIFT+X**.

Следующие опции сохранения помогут вам настроить процесс сохранения и получить документы, пригодные для дальнейшего редактирования и использования. В программе ABBYY FineReader предусмотрены следующие настройки:

- **Игнорировать текст вне таблицы**
Указывает, что в файл будут сохранены только таблицы.
- **Сохранять числовые данные в формате «Цифры»**
Указывает, что при сохранении распознанного текста в формате XLS числа будут сохранены в формате «Цифры». Этот формат в Excel используется для выполнения арифметических операций над ячейками, содержащими числа.

Сохранение распознанного текста в формате PDF

Опции сохранения для формата PDF устанавливаются на закладке **PDF** диалога **Форматы**. Для того чтобы открыть этот диалог, щелкните по кнопке **Форматы**, расположенной на закладке **Сохранить** диалога **Опции** (меню **Сервис>Опции**) или нажмите **CTRL+SHIFT+X**.

Следующие опции сохранения помогут вам настроить процесс сохранения и получить документы, пригодные для дальнейшего редактирования и использования. В программе ABBYY FineReader предусмотрены следующие настройки:

- Размер бумаги по умолчанию,
- Формат сохранения,
- Теги PDF,
- Параметры сохранения иллюстраций и шрифтов,
- Параметры защиты PDF.

Размер бумаги по умолчанию

Вы можете указать размер бумаги, который будет использован при сохранении результатов распознавания в формат PDF. Размер бумаги задается в выпадающем списке **Размер бумаги по умолчанию**.

-
- Совет.**
1. Если вы не нашли нужного размера бумаги в списке, вы можете создать свой – пользовательский – размер бумаги и использовать его для сохранения распознанного текста. Для того чтобы создать пользовательский размер бумаги, выберите пункт **Добавить пользовательский размер** и в открывшемся диалоге укажите название, высоту и ширину страницы.
 2. Для того чтобы распознанный текст полностью умещался на странице при сохранении, отметьте опцию **Автоматически увеличивать размер страницы**.
В этом случае выбор размера бумаги будет выполнен автоматически.
-

Формат сохранения

Формат сохранения выбирается в выпадающем списке **Формат сохранения**. Выбор формата сохранения оформления зависит от того, как вы будете использовать созданный документ в дальнейшем:

- **Только изображение**

В данном режиме сохраняется изображение страницы. Внешний вид полученного PDF-документа будет не отличим от оригинала, однако, в таком документе невозможен поиск по тексту.

- **Только текст и картинки**

В данном режиме сохраняются распознанный текст и иллюстрации. PDF-файл такого типа имеет относительно небольшой размер, и в нем возможен полнотекстовый поиск.

- **Текст поверх изображения страницы**

В данном режиме сохраняются фон и иллюстрации исходного документа, на них накладывается распознанный текст. Обычно размер такого PDF-файла больше, чем файла, получаемого при сохранении в режиме **Только текст и картинки**. В PDF-документе этого типа есть возможность полнотекстового поиска, однако внешний вид документа может немного отличаться от оригинала из-за того, что на изображение документа накладывается распознанный текст.

- **Текст под изображением страницы**

В данном режиме сохраняется изображение страницы, распознанный текст помещается на невидимый слой под изображением. Таким образом, получается документ с возможностью поиска по тексту, при этом внешний вид PDF-документа практически не отличим от оригинала.

Теги PDF

Помимо текста и иллюстраций PDF могут содержать информацию о структуре документа: логических частях, иллюстрациях и таблицах. Эта информация хранится в виде тегов PDF. Теги PDF обеспечивают удобство просмотра документа на экранах различного размера, например, на экранах карманных компьютеров.

Для того чтобы теги PDF были добавлены в документ автоматически, отметьте опцию **Разрешить теги PDF (для Acrobat 5.0 или более поздней версии)**.

Параметры сохранения иллюстраций и шрифтов

Если распознанный текст содержит большое количество иллюстраций, то размер конечного файла может быть очень большим. Размер файла документа можно уменьшить, выбрав в группе **Картинки** подходящее значение качества и формат для сохраняемых иллюстраций.

Качество

Выбор уровня качества зависит от того, как вы будете использовать документ в дальнейшем. Из выпадающего списка **Качество** выберите:

- **Высокое**, если создаваемый документ предназначен для печати.
- **Среднее**, если создаваемый документ предназначен для просмотра на экране.
- **Низкое**, если создаваемый документ предназначен для размещения на веб-страницах.

Чем выше значение выбрано в списке **Качество**, тем выше будет качество иллюстраций. Размер файла также зависит от выбранного значения: чем оно выше, тем больше размер файла.

Совет. Чтобы подобрать оптимальное сочетание качества иллюстраций и размера файла, сохраните распознанный текст и иллюстрации в нескольких вариантах – с разными значениями из списка **Качество** – и выберите подходящий файл.

Формат

При сохранении иллюстраций программа ABBYY FineReader автоматически подбирает наиболее подходящий формат. В этом случае в выпадающем списке **Формат** установлено значение **(Авто)**.

Если вы хотите выбрать формат самостоятельно, в выпадающем списке **Формат** выберите одно из следующих значений:

- **JPEG**, цветной (для фотографий),
- **JPEG**, серый (для фотографий),
- **LZW**, цветной (без потерь),
- **LZW**, серый (без потерь),
- **ZIP**, цветной (без потерь),
- **ZIP**, серый (без потерь),
- **CCITT4**, черно-белый.

Шрифты

При сохранении распознанного текста в формат PDF ABBYY FineReader могут использовать два набора шрифтов: стандартные шрифты Acrobat и системные шрифты, установленные на компьютере пользователя. Для того чтобы указать, какой набор шрифтов должен быть использован, в выпадающем списке

Шрифт выберите:

- **Использовать стандартные шрифты**

При выборе данной опции PDF-файл содержит ссылку на стандартные шрифты Acrobat: TimesNewRoman, Arial, CourierNew.

- **Использовать системные шрифты**

При выборе данной опции PDF-файл содержит ссылку на системные шрифты, установленные на компьютере пользователя.

Параметры защиты PDF

Вы можете задать пароли, защищающие ваш PDF-документ от несанкционированного открытия, печати и редактирования.

Подробнее см. раздел «Параметры защиты PDF».

Сохранение распознанного текста в формате HTML

Опции сохранения для формата HTML устанавливаются на закладке **HTML** диалога **Форматы**.

Для того чтобы открыть этот диалог, щелкните по кнопке **Форматы**, расположенной на закладке **Сохранить** диалога **Опции** (меню **Сервис>Опции**) или нажмите **CTRL+SHIFT+X**.

Замечание. При сохранении результатов распознавания в формате HTML ABBYY FineReader использует шрифты, выбранные на закладке **Сохранить** диалога **Опции**, или шрифты, установленные в окне **Текст**.

Следующие опции сохранения помогут вам настроить процесс сохранения и получить документы, пригодные для дальнейшего редактирования и использования. В программе ABBYY FineReader предусмотрены следующие настройки:

- Режим сохранения оформления
- Формат сохранения
- Параметры сохранения текста,
- Параметры сохранения иллюстраций.
- Кодировка символов.

Режим сохранения оформления

Режим сохранения оформления выбирается в выпадающем списке **Оформление**:

- **Оформление оригинала**

Позволяет полностью сохранить оформление исходного документа.

- **Сохранить очертание и размер шрифта**

Указывает, что в распознанном тексте будут сохранены разбиение на абзацы, гарнитура и размер шрифта.

- **Не сохранять оформление**

Указывает, что форматирование текста не сохраняется. Весь текст передается одним шрифтом одного размера.

Формат сохранения

Выберите один из двух режимов создания HTML файла:

- **Полный (с использованием CSS)** – файл записывается в новом формате – HTML 4, что позволяет точно передать оформление документа, используя при этом таблицу стилей, встраиваемую в HTML-файл.
- **Простой (совместим со всеми браузерами)** – файл записывается в формате HTML 3; оформление сохраняется приблизительно (не сохраняются отступы первой строки, неточно сохраняется размер шрифта). Этот формат HTML поддерживается всеми браузерами (Netscape Navigator и Internet Explorer 3.0 и выше).

Параметры сохранения текста

- **Сохранять деление на строки**
При сохранении в формат HTML распознанный текст будет разбит на строки так же, как на исходном изображении. Если эта опция не отмечена, то при сохранении в формате HTML весь распознанный текст будет записан в одну строку.
- **Сохранять цвет текста**
Позволяет сохранить исходный цвет букв.
- **Разделять страницы сплошной линией**
Сохраненный в формате HTML распознанный текст будет разбит на страницы так же, как и исходное изображение. При этом конец страницы будет обозначен сплошной линией.

Параметры сохранения иллюстраций

Для того чтобы при сохранении распознанного текста сохранить и иллюстрации, необходимо отметить опцию **Сохранять картинки** в группе **Картинки**.

Если распознанный текст содержит большое количество иллюстраций, то размер конечного файла может быть очень большим. Размер файла документа можно уменьшить, выбрав в группе **Картинки** подходящее значение качества и формат для сохраняемых иллюстраций.

Качество

Выбор уровня качества зависит от того, как вы будете использовать документ в дальнейшем. Из выпадающего списка **Качество** выберите:

- **Высокое**, если создаваемый документ предназначен для печати.

- **Среднее**, если создаваемый документ предназначен для просмотра на экране.
- **Низкое**, если создаваемый документ предназначен для размещения на веб-страницах.

Чем выше значение выбрано в списке **Качество**, тем выше будет качество иллюстраций. Размер файла также зависит от выбранного значения: чем оно выше, тем больше размер файла.

Совет. Чтобы подобрать оптимальное сочетание качества иллюстраций и размера файла, сохраните распознанный текст и иллюстрации в нескольких вариантах – с разными значениями из списка **Качество** – и откройте получившуюся HTML-страницу.

Формат

При сохранении иллюстраций программа ABBYY FineReader автоматически подбирает наиболее подходящий формат. В этом случае в выпадающем списке **Формат** установлено значение (**Авто**).

Если вы хотите выбрать формат самостоятельно, в выпадающем списке **Формат** выберите одно из следующих значений:

- **JPEG**, цветной (для фотографий),
- **JPEG**, серый (для фотографий),
- **PNG**, цветной (для чертежей и схем),
- **PNG**, серый (для чертежей и схем),
- **PNG**, черно-белый.

Кодировка символов

Программа ABBYY FineReader автоматически подбирает кодовую страницу. Если вы хотите изменить кодовую страницу или ее тип, выбранный автоматически, выберите нужное значение из выпадающих списков в группе **Кодировка символов**.

Сохранение распознанного текста в формате PPT

Опции сохранения для формата PPT устанавливаются на закладке **PPT** диалога **Форматы**. Для того чтобы открыть этот диалог, щелкните по кнопке **Форматы**, расположенной на закладке **Сохранить** диалога **Опции** (меню **Сервис>Опции**), или нажмите **CTRL+SHIFT+X**.

Следующие опции сохранения помогут вам настроить процесс сохранения и получить документы, пригодные для дальнейшего редактирования и использования. В программе ABBYY FineReader предусмотрены следующие настройки:

- Параметры сохранения текста,
- Параметры сохранения иллюстраций.

Параметры сохранения текста

- **Сохранять деление на строки**

При сохранении в формате PPT распознанный текст будет разбит на строки так же, как на исход-

ном изображении. Если эта опция не отмечена, то при сохранении весь распознанный текст будет записан в одну строку.

- **Переносить по словам**

При сохранении деления на строки распознанный текст будет умещен по ширине текстового блока слайда.

Параметры сохранения иллюстраций

Для того чтобы при сохранении распознанного текста сохранить и иллюстрации, необходимо отметить опцию **Сохранять картинки** в группе **Картинки**.

Если распознанный текст содержит большое количество иллюстраций, то размер конечного файла может быть очень большим. Размер файла документа можно уменьшить, выбрав в группе **Картинки** подходящее значение качества и формат для сохраняемых иллюстраций.

Качество

Выбор уровня качества зависит от того, как вы будете использовать документ в дальнейшем. Из выпадающего списка **Качество** выберите: **высокое**, **среднее**, либо **низкое** качество.

Чем выше значение выбрано в списке **Качество**, тем выше будет качество иллюстраций. Размер файла также зависит от выбранного значения: чем оно выше, тем больше размер файла.

Совет. Чтобы подобрать оптимальное сочетание качества иллюстраций и размера файла, сохраните распознанный текст и иллюстрации в нескольких вариантах – с разными значениями из списка **Качество** – и выберите подходящий файл.

Формат

При сохранении иллюстраций программа ABBYY FineReader автоматически подбирает наиболее подходящий формат. В этом случае в выпадающем списке **Формат** установлено значение (**Авто**).

Если вы хотите выбрать формат самостоятельно, в выпадающем списке **Формат** выберите одно из следующих значений:

- **JPEG**, цветной (для фотографий),
- **JPEG**, серый (для фотографий),
- **PNG**, цветной (для чертежей и схем),
- **PNG**, серый (для чертежей и схем),
- **PNG**, черно-белый.

Внимание! При сохранении в формате PPT результаты распознавания сохраняются в виде специальных HTML-файлов, содержащих разные части презентации. Для того чтобы презентация была сохранена в одном файле, ее необходимо открыть и повторно сохранить в программе PowerPoint, выбрав в меню **Файл** пункт **Сохранить как** и указав в качестве формата сохранения формат PPT.

Сохранение распознанного текста в формате TXT

Опции сохранения для формата TXT устанавливаются на закладке **TXT** диалога **Форматы**. Для того чтобы открыть этот диалог, щелкните по кнопке **Форматы**, расположенной на закладке **Сохранить** диалога **Опции** (меню **Сервис>Опции**), или нажмите **CTRL+SHIFT+X**.

Следующие опции сохранения помогут вам настроить процесс сохранения и получить документы, пригодные для дальнейшего редактирования и использования. В программе ABBYY FineReader предусмотрены следующие настройки:

- Параметры сохранения текста,
- Кодировка символов.

Параметры сохранения текста

- **Сохранять деление на строки**
При сохранении в формате TXT распознанный текст будет разбит на строки так же, как на исходном изображении. Если этот пункт не отмечен, то при сохранении в формат TXT каждый абзац распознанного текста будет записан в одну строку.
- **Добавлять в конец файла**
Позволяет сохранять распознанный текст в уже существующий файл *.txt, добавив распознанный текст в конец файла (в этом случае опция **Добавлять в конец файла** должна быть отмечена).
- **Разделять страницы символом конца страницы (#12)**
При сохранении в формате TXT распознанный текст будет разбит на страницы так же, как на исходном изображении. Если этот пункт не отмечен, разбиение на страницы не сохраняется.
- **Разделять абзацы пустыми строками**
При сохранении в формате TXT абзацы распознанного текста разделяются пустыми строками.

Кодировка символов

Программа ABBYY FineReader автоматически подбирает кодовую страницу. Если вы хотите изменить кодовую страницу или ее тип, выбранный автоматически, выберите нужное значение из выпадающих списков в группе **Кодировка символов**.

Сохранение распознанного текста в формате DBF

Опции сохранения для формата DBF устанавливаются на закладке **DBF** диалога **Форматы**. Для того чтобы открыть этот диалог, щелкните по кнопке **Форматы**, расположенной на закладке **Сохранить** диалога **Опции** (меню **Сервис>Опции**), или нажмите **CTRL+SHIFT+X**.

Следующие опции сохранения помогут вам настроить процесс сохранения и получить документы, пригодные для дальнейшего редактирования и использования. В программе ABBYY FineReader предусмотрены следующие настройки:

- Параметры сохранения текста,
- Кодировка символов.

Параметры сохранения текста

- **Добавлять в конец файла**

Позволяет сохранять распознанный текст в уже существующий файл *.dbf, добавив распознанный текст в конец файла (в этом случае опция **Добавлять в конец файла** должна быть отмечена).

Кодировка символов

Программа ABBYY FineReader автоматически подбирает кодовую страницу. Если вы хотите изменить кодовую страницу или ее тип, выбранный автоматически, выберите нужное значение из выпадающих списков в группе **Кодировка символов**.

Сохранение распознанного текста в формате LIT

Опции сохранения для формата LIT устанавливаются на закладке **LIT** диалога **Форматы**. Для того чтобы открыть этот диалог, щелкните по кнопке **Форматы**, расположенной на закладке **Сохранить** диалога **Опции** (меню **Сервис>Опции**) или нажмите **CTRL+SHIFT+X**.

Следующие опции сохранения помогут вам настроить процесс сохранения и получить документы, пригодные для дальнейшего редактирования и использования. В программе ABBYY FineReader предусмотрены следующие настройки:

- Параметры сохранения текста,
- Параметры сохранения иллюстраций.

Параметры сохранения текста

- **Сохранять деление на строки**

При сохранении в формате LIT распознанный текст будет разбит на строки так же, как на исходном изображении. Если эта опция не отмечена, то весь распознанный текст будет записан в одну строку.

- **Сохранять деление на строки**

При сохранении деления на строки распознанный текст будет уменен по ширине текстового блока.

- **Преобразовывать текст в таблицы**

Если данная опция отмечена, распознанный текст будет сохранен в виде таблицы.

Параметры сохранения иллюстраций

Для того чтобы при сохранении распознанного текста сохранить и иллюстрации, необходимо отметить опцию **Сохранять картинки** в группе **Картинки**.

Если распознанный текст содержит большое количество иллюстраций, то размер конечного файла может быть очень большим. Размер файла документа можно уменьшить, выбрав в группе **Картинки** подходящее значение качества и формат для сохраняемых иллюстраций.

Качество

Выбор уровня качества зависит от того, как вы будете использовать документ в дальнейшем. Из выпадающего списка **Качество** выберите: **высокое**, **среднее**, либо **низкое**.

Чем выше значение выбрано в списке **Качество**, тем выше будет качество иллюстраций. Размер файла также зависит от выбранного значения: чем оно выше, тем больше размер файла.

Формат

При сохранении иллюстраций программа ABBYY FineReader автоматически подбирает наиболее подходящий формат. В этом случае в выпадающем списке **Формат** установлено значение (**Авто**).

Если вы хотите выбрать формат самостоятельно, в выпадающем списке **Формат** выберите одно из следующих значений:

- **JPEG**, цветной (для фотографий),
- **JPEG**, серый (для фотографий),
- **PNG**, цветной (для чертежей и схем),
- **PNG**, серый (для чертежей и схем),
- **PNG**, черно-белый.

Сохранение изображения страницы

1. Выделите страницу в пакете.
2. В меню **Файл** выберите пункт **Сохранить результаты>Сохранить изображения**.
3. В открывшемся диалоге **Сохранить изображение как** выберите диск, папку для размещения сохраняемого файла и формат.

Замечание. Вы можете ограничить сохраняемую в файле область изображения рамками выделенных на нем блоков (любого типа). Для этого выделите нужный блок или блоки, а затем в диалоге **Сохранить изображение как** отметьте пункт **Сохранить выделенные блоки**. Этот режим работает только при сохранении одиночного изображения.

4. Укажите имя файла.
5. Нажмите **Сохранить**.

Совет: Вы можете сохранить несколько изображений в один файл как многостраничный TIF. Для этого:

1. Выделите нужные изображения в окне **Пакет**.
2. В меню **Файл** выберите пункт **Сохранить результаты>Сохранить изображения**. В открывшемся диалоге выберите формат TIF и отметьте опцию **Сохранить страницы в один файл**.

Замечание. Если вы сохраняете несколько выделенных в окне **Пакет** изображений в виде отдельных файлов (т.е. не используя режим сохранения в многостраничный TIF), имена этих

файлов будут образовываться добавлением к введенному вами имени файла номера страницы в виде четырехзначного числа.

Как выбрать метод сжатия изображения

Сжатие изображений позволяет уменьшить размер файла. Методы, применяемые для сжатия изображений, различаются по степени сжатия и наличию потерь данных. На выбор метода сжатия влияют два фактора: качество получаемого изображения и размер файла. ABBYY FineReader 8.0 позволяет использовать следующие методы сжатия:

- **ZIP** – данный метод применяется для изображений, содержащих большие участки одного цвета, например, для снимков экрана (скриншотов) и изображений, созданных в графических редакторах. Также этот метод подходит для черно–белых изображений. Данный метод сжатия не приводит к потере данных.
- **JPEG** – данный метод сжатия применяется для серых и цветных изображений, например, для фотографий. Метод JPEG позволяет достичь значительного сжатия изображения, но приводит к потере данных и, как следствие, к ухудшению качества изображения (размыванию контуров и потере оттенков цветов).
- **CCITT** – этот метод применяется для черно–белых изображений, созданных в графических редакторах или полученных со сканера. Данное сжатие не приводит к потере данных. **CCITT Group 4** является широко используемым методом сжатия и применим практически ко всем изображениям. **CCITT Group 3** используется для передачи изображений по факсу.
- **Packbits** – данный метод сжатия не приводит к потере данных и подходит для отсканированных черно-белых изображений.

Добавление свойств документа

Свойства документа— это подробные сведения о файле, помогающие отличать его от других (например, описательное название, имя автора, тема и ключевые слова). Свойства документа используются для просмотра сведений о файле или при упорядочении файлов. Кроме того, можно выполнять поиск документов по их свойствам. Для задания или изменения свойств документа:

- Выберите документ в окне **Пакет**, затем в меню **Файл** выберите команду **Свойства документа**, или
- Щелкните правой кнопкой мыши по документу в окне **Пакет** и выберите из контекстного меню команду **Свойства**.

В открывшемся диалоге укажите название, автора, тему и ключевые слова документа.

Замечание. Не все форматы файлов поддерживают свойства документа. Свойства, указанные вами, будут добавлены только для файлов формата DOC, XLS, HTM, PDF, и LIT.

Работа с пакетами

В пакете хранятся как исходные изображения, так и соответствующий им распознанный текст. Большинство установок ABBYY FineReader сохраняются на пакет (опции сканирования, распознавания, сохранения, а также созданные в процессе работы пользовательские эталоны, языки и группы языков). При создании нового пакета могут использоваться настройки по умолчанию, настройки текущего пакета или настройки из файла с расширением *.fbt.

Общая информация по работе с пакетом

При запуске программы ABBYY FineReader по умолчанию открывается новый пакет. Вы можете работать с этим пакетом или создать новый. В пакете может содержаться до 9999 страниц. В один пакет для удобства работы рекомендуется объединять изображения, логически связанные между собой (например, страницы одной книги, тексты на одном языке или изображения с однотипным расположением текста и т.д.).

В окне **Пакет** вы можете видеть список страниц, входящих в открытый пакет. Чтобы просмотреть страницу, нужно щелкнуть мышью по ее значку или дважды щелкнуть по ее номеру. Откроются файлы, которыми данная страница представлена в пакете: текстовый (если данная страница была распознана) и графический, – в окнах **Текст** и **Изображение** соответственно.

Существует два основных способа представления страниц в окне **Пакет**:



Пиктограммы.

Страницы пакета отображаются пиктограммами – своими сильно уменьшенными изображениями. По мере обработки изображения на пиктограмме появляются дополнительные специальные значки, отображающие действия, произведенные над страницей. Этот способ представления страниц пакета удобно использовать, например, для открытия необходимой страницы пакета: страница представлена своим миниатюрным изображением, и вам не надо запоминать номер, под которым она была отсканирована. Чтобы открыть изображение, щелкните мышью по его пиктограмме.

Вид со свойствами.

В этом случае в окне **Пакет** отображается дополнительная информация о странице и возможна сортировка по выбранному признаку. **Вид со свойствами** удобен при обработке пакета, содержащего большое количество страниц, т.к. на экране помещается большее (чем в режиме **Пиктограммы**) количество страниц. Чтобы открыть изображение, дважды щелкните на соответствующий ему значок в окне **Пакет**.

Чтобы выбрать вид страниц в окне Пакет:

- Нажмите на соответствующую кнопку (/) на панели инструментов **Стандартная** или
- Щелкните правой кнопкой мыши в окне **Пакет** и в открывшемся локальном меню выберите пункт **Окно Пакет>...**

Помимо выбора способа представления изображений в окне **Пакет**, вы можете выбрать показываемые свойства страницы, способ сортировки страниц пакета и т.д. Для этого:


- Щелкните правой кнопкой мыши в окне **Пакет** и в открывшемся локальном меню выберите пункт **Окно Пакет>Настройка**. В открывшемся диалоге на закладках **Пиктограммы** и **Вид со свойствами** установите нужные вам опции.

В одном пакете можно выделить несколько страниц подряд или выборочно:

- **Чтобы выделить несколько страниц подряд:** удерживая клавишу **SHIFT**, укажите с помощью мыши первую, а затем последнюю страницу из выборки.
- **Чтобы выделить несколько страниц выборочно:** удерживая клавишу **CTRL**, укажите на них мышью.
- **Чтобы выделить все страницы пакета:** при активном окне пакета из меню **Правка** выберите пункт **Выделить все** или нажмите **ALT+A**.

Создание нового пакета

Чтобы создать новый пакет:


- В меню **Файл** выберите пункт **Новый пакет** или
- Нажмите кнопку **Новый пакет** ()

Открытие пакета


При запуске ABBYY FineReader по умолчанию открывает новый пакет.

Замечание. Если вы хотите, чтобы при запуске программа открывала последний пакет, с которым вы работали, отметьте опцию **Открывать последний пакет при запуске ABBYY FineReader** на закладке **Общие** диалога **Опции** (меню **Сервис>Опции**).

Чтобы открыть другой пакет:

1. В меню **Файл** выберите пункт **Открыть пакет** или нажмите кнопку **Открыть пакет** ().
2. В открывшемся диалоге **Открыть пакет** выберите папку, где находится нужный вам пакет. В случае, если вы работали с новым пакетом, и он вам понадобится для дальнейшей работы, сохраните его перед выходом из программы.

Вы можете открыть пакет из Windows Explorer. Для этого:

- Щелкните правой кнопкой мыши по папке пакета (отображается значком ) и в локальном меню выберите пункт **Открыть с помощью ABBYY FineReader**. Запустится новая копия ABBYY FineReader с выбранным пакетом.

Добавление изображений в пакет

- В меню **Файл** выберите пункт **Открыть PDF/изображение**. Клавиатурная команда: **CTRL+O**.
- В диалоге **Открыть изображение** выберите одно или несколько изображений. Изображение будет добавлено в открытый пакет, и его копия будет сохранена в папке пакета.

Замечание. Вы можете добавить изображение из Windows Explorer. Для этого:

1. В Windows Explorer выберите один или несколько файлов с изображениями.
2. Щелкните правой кнопкой мыши и в локальном меню выберите пункт **Открыть с помощью ABBYY FineReader**. Если на вашем компьютере уже

открыт ABBYY FineReader, выбранные файлы будут добавлены в текущий пакет, в противном случае автоматически запустится ABBYY FineReader и файлы будут добавлены в новый пакет. Данный пункт локального меню доступен только для файлов, формат которых поддерживается ABBYY FineReader 8.0.

Нумерация страниц пакета

Все страницы в пакете пронумерованы. Один пакет может содержать до 9999 страниц (четырёхзначный номер). Номер страницы указывается непосредственно в самом пакете, около условного обозначения страницы.

Страницы можно перенумеровать непосредственно в окне **Пакет** или через диалог **Перенумеровать страницы**.

Чтобы перенумеровать страницы в окне **Пакет**:

1. Выделите страницу в окне **Пакет** и щелкните мышью по ее номеру или нажмите **F2**.
2. Введите новый номер страницы.

После изменения номера положение страницы в окне **Пакет** изменится в соответствии с введенным номером.

Замечание. Двойной щелчок мыши по номеру страницы открывает эту страницу в программе.

Чтобы перенумеровать страницы через диалог **Перенумеровать страницы**:

1. Выделите страницу или несколько страниц.
2. В меню **Пакет** выберите пункт **Перенумеровать страницы...**
3. Введите новый номер для первой страницы из выборки (страницы с наименьшим номером).

-
- Замечания.**
1. Если вы хотите перенумеровать все страницы пакета, то в диалоге **Перенумеровать страницы** выберите пункт **Перенумеровать все страницы**.
 2. Если вы хотите перенумеровать часть страниц пакета, то
 - В окне **Пакет** выделите те страницы, которые вы хотите перенумеровать.
 - В диалоге **Перенумеровать страницы** выберите пункт **Перенумеровать выделенные страницы**.
 3. Если вы хотите, чтобы выделенные страницы были перенумерованы по порядку, начиная с указанного номера, то отметьте опцию **Сплошная нумерация страниц**. Например, перенумеровываются страницы 2, 5, 6; в качестве начального номера был указан 1. Новые номера страниц будут: 1, 2, 3. В противном случае (опция **Сплошная нумерация страниц** не отмечена) новыми номерами страниц будут: 1, 4, 5. Т.е. первая из участвующих в операции страниц получит указанный вами начальный номер, а остальные будут перенумерованы таким образом, чтобы сохранить интервалы нумерации между страницами.
 4. Если перенумерация затрагивает не все страницы пакета, то возможна ситуация, когда в соответствии с заданными вами параметрами некоторая страница должна получить новый номер, совпадающий с номером одной из страниц,

не участвующих в операции. В этом случае ABBYY FineReader выдаст соответствующее предупреждение, и вся операция будет отменена.

Сохранение пакета

Чтобы сохранить пакет:

1. В меню **Файл** выберите пункт **Сохранить пакет как**.
2. В открывшемся диалоге **Сохранить пакет как** укажите название пакета и путь, по которому он будет храниться.

Заккрытие страницы пакета или всего пакета

Чтобы закрыть страницу пакета:

- В меню **Пакет** выберите пункт **Заккрыть текущую страницу**.

Чтобы закрыть пакет:

- В меню **Файл** выберите пункт **Заккрыть пакет**.

Удаление пакета

Внимание! При удалении пакета удаляются все его страницы (изображения и текст) и дополнительные файлы, созданные в процессе работы с этим пакетом: эталон, пользовательские языки, т.е. удаляется все содержимое папки, соответствующей пакету.

- Чтобы удалить пакет, в меню **Пакет** выберите пункт **Удалить пакет**.

Чтобы удалить одну или несколько страниц в пакете:

1. Выделите нужные страницы в окне пакета.
2. В меню **Пакет** выберите пункт **Удалить страницу**. Клавиатурная команда: **DEL**

Настройки пакета

Чтобы сохранить настройки пакета в отдельный файл:

1. На закладке **Общие** диалога **Опции** (меню **Сервис>Опции**) нажмите кнопку **Сохранить опции**.
2. В открывшемся диалоге укажите имя файла.

В этот файл будут сохранены значения опций с закладок **Сканировать/Открыть**, **Распознать**, **Проверить** и **Сохранить**; а также значения опций с закладок диалога **Форматы** и диалога **Дополнительные опции**; путь к папке, в которой хранятся пользовательские языки и их словари; а также пользовательские языки, группы языков, пользовательские словари и эталоны.

Чтобы вернуться к опциям, устанавливаемым системой по умолчанию:

- На закладке **Общие** нажмите кнопку **Восстановить умолчания**.

Чтобы загрузить опции:

- На закладке **Общие** нажмите кнопку **Загрузить опции** выберите файл набора опций ABBYY FineReader (*.fbi), содержащий требуемые опции.

Расширенный поиск

Внимание! Для работы данной опции необходим Internet Explorer 4.0 или выше.

Вы можете задать поиск текстового фрагмента по всем страницам пакета, содержащим распознанный текст. В качестве искомого текстового фрагмента может выступать одно слово или несколько слов. При этом слово или слова могут быть в любой форме (для языков со словарной поддержкой); слова из запроса могут находиться в тексте страницы на произвольном расстоянии друг от друга и располагаться в произвольном порядке.

Для этого:

1. В меню **Правка** выберите пункт **Расширенный поиск** или нажмите **ALT+F3**.
2. Откроется окно **Расширенный поиск**.
3. В поле **Найти** введите, вставьте из буфера обмена или выберите из списка ранее введенных вами запросов искомый текст.
4. Нажмите кнопку **Найти**.

В окне **Поиск** появится список номеров страниц пакета, на которых были найдены ВСЕ искомые слова. Дополнительно для каждой найденной страницы указывается дата ее последней модификации и ближайший к началу страницы фрагмент ее текста, включающий одно или несколько из искомых слов (они выделены цветом). Щелчок по номеру страницы открывает ее в окнах **Изображение**, **Текст** и **Крупный план**; при этом все найденные слова из запроса выделяются цветом фона во всех трех окнах.

Замечание. Расширенный поиск не позволяет осуществлять поиск специальных символов (например, символа конца строки или конца абзаца).

Сценарии ABBYY FineReader

Процесс распознавания документов очень часто состоит из одного и того же набора операций – например, сканирование документов, их распознавание и сохранение распознанного текста в определенном формате. Автоматизировать однотипные действия вам помогут сценарии ABBYY FineReader.

Сценарий ABBYY FineReader – это набор последовательных шагов, каждый из которых соответствует одному этапу обработки документа. Запуск выполнения сценария осуществляется из меню кнопки **Scan&Read**. Переход от одного шага сценария к следующему выполняется автоматически.

В программе предусмотрено три встроенных сценария, полностью готовых к использованию. С помощью **Менеджера сценариев** можно также создать свои – пользовательские сценарии, состоящие из шагов, необходимые для решения ваших задач.

Запуск и выполнение сценария

Для того чтобы запустить сценарий:

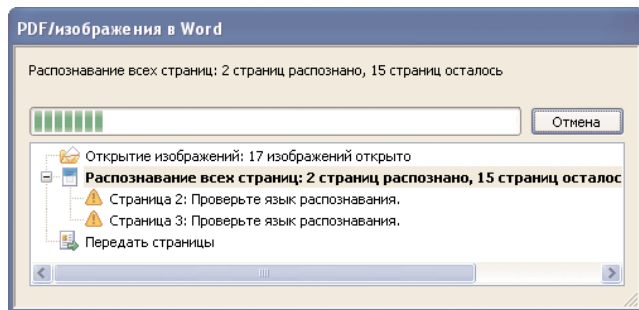
- выберите сценарий из выпадающего списка многофункциональной кнопки **Scan&Read**, или
- воспользуйтесь горячими клавишами **Ctrl+W**, если на кнопке **Scan&Read** выбран нужный сценарий, или
- в диалоге **Менеджера сценариев** выберите сценарий из списка и нажмите кнопку **Запуск**.

Выполнение сценария

Сценарии ABBYY FineReader выполняются автоматически, с использованием настроек, выбранных в диалоге **Опции** (для того чтобы открыть диалог **Опции**, воспользуйтесь командой **Опции...** в меню **Сервис**).

- Замечания.**
1. Если при выполнении сценария изображения добавляются в пакет, уже содержащий изображения, то будет выполнена обработка только добавленных изображений. Если же в ходе выполнения сценария должен быть обработан пакет, то будут обработаны все изображения из этого пакета.
 2. Если в ходе сценария должен быть обработан пакет ABBYY FineReader, в котором есть распознанные страницы, и страницы, на которых уже была выполнена разметка блоков, то распознанные страницы повторно обрабатываться не будут, страницы с размеченными блоками будут распознаны.

После запуска сценария на экране появляется **панель выполнения сценария**, содержащая индикатор выполнения сценария, список шагов сценария, а также подсказки и предупреждения.



- Совет.** Если для сценария необходимы значения настроек, которые вы обычно не используете, то можно создать **набор настроек** и загружать его перед выполнением сценария, при этом вам не придется изменять значение настроек вручную перед каждым выполнением сценария. Для того чтобы создать набор настроек, установите в диалоге **Опции** значения настроек и сохраните набор настроек (кнопка **Сохранить настройки** на закладке **Общие** в диалоге **Опции**). Затем перед выполнением сценария загрузите набор настроек, требующийся для выбранного сценария, (кнопка **Загрузить настройки**) и запустите выполнение сценария.

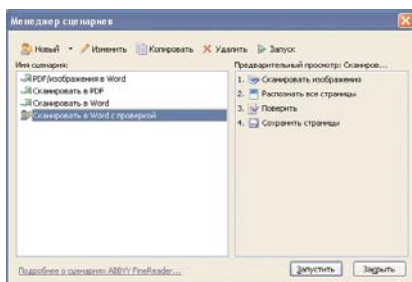
Управление сценариями

Управление сценариями – это запуск выполнения сценариев, создание и редактирование пользовательских сценариев, а также удаление ненужных пользовательских сценариев.




Управление сценариями ABBYY FineReader осуществляется с помощью **Менеджера сценариев**.

Для того чтобы открыть диалог Менеджера сценариев:

- выберите команду **Менеджер сценариев** из списка команд кнопки **Scan&Read**, или
- в меню **Процесс** выберите команду **Сценарии ABBYY FineReader>Менеджер сценариев**, или
- в меню **Сервис** выберите команду **Менеджер сценариев**.








На панели инструментов диалога **Менеджера сценариев** расположены кнопки, позволяющие создать, изменить, копировать, удалить и запустить выполнение сценария.

В левой части окна **Менеджера сценариев** располагается список сценариев. Встроенные сценарии ABBYY FineReader отмечены значком , пользовательские сценарии отмечены значком . Сценарии, которые невозможно выполнить на данном компьютере, отмечаются значком .

При выборе сценария из списка в правой части окна появляется список шагов сценария.

Замечание. Для передачи распознанного текста в какое-либо внешнее приложение, требуется наличие этого приложения на компьютере. Поэтому сценарии, предусматривающие передачу распознанного текста в приложения, отсутствующие на компьютере, не будут выполняться на данном компьютере. Такие сценарии не отображаются в списке сценариев кнопки **Scan&Read**.

Нажмите кнопку	Для того чтобы...
 Новый	Создать новый сценарий. Мастер сценариев поможет вам выбрать шаги сценария и полностью их настроить.
Экспорт...	<p>Экспортировать пользовательский сценарий. Файл сценария затем можно передать другому пользователю.</p> <p>В открывшемся диалоге Экспортировать сценарий укажите имя файла, в который будет экспортирован сценарий (файлы пользовательских сценариев имеют расширение *.fta).</p> <p>Замечание. По умолчанию файлы сценариев сохраняются в %UserProfile%\Local Settings\Application Data\ABBYY\FineReader\8.00</p>

	Импорт...	Импортировать сценарий, созданный другим пользователем. В открывшемся диалоге Импорт сценария выберите имя файла пользовательского сценария (файл *.fta) и папку, в которую он будет сохранен.
	Изменить	Отредактировать сценарий. Мастер сценариев поможет вам выбрать шаги сценария и полностью их настроить. Замечание. Встроенные сценарии ABBYY FineReader нельзя редактировать, можно создать копию такого сценария, а затем внести в нее все необходимые изменения.
	Копировать	Создать копию сценария. Копия сценария будет добавлена в список сценариев сразу после оригинала.
	Удалить	Удалить сценарий. Замечание. Встроенные сценарии ABBYY FineReader удалить нельзя.
	Запустить	Начать обработку документов в соответствии с выбранным сценарием.

Совет. Изменить название пользовательского сценария можно, выбрав команду **Переименовать** в контекстном меню сценария.

Встроенные сценарии ABBYY FineReader

Встроенные сценарии ABBYY FineReader содержат все шаги, необходимые для того, чтобы получить изображение документа, распознать изображение и сохранить распознанный текст. В состав встроенных сценариев не входит шаг проверки результатов распознавания.

В программе предусмотрены следующие готовые сценарии:



Сценарий PDF/изображения в Word



Последовательность шагов сценария:

1. Открытие файлов изображений или PDF-файлов.
При запуске сценария открывается диалог **Открыть**, в этом диалоге укажите файлы документов, подлежащих обработке. Выбранные изображения будут добавлены в текущий пакет ABBYY FineReader.
2. Распознавание изображений.
3. Сохранение распознанного текста в Microsoft Word.
В программе Microsoft Word будет открыт новый документ, содержащий распознанный текст.

Замечание. Для данного сценария необходимо, чтобы на вашем компьютере была установлена программа Microsoft Word.

Сценарий Сканировать в Word

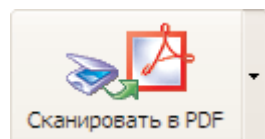


Последовательность шагов сценария:

1. Сканирование бумажных документов.
Полученные со сканера изображения будут добавлены в текущий пакет ABBYY FineReader.
2. Распознавание изображений.
3. Сохранение распознанного текста в Microsoft Word.
В программе Microsoft Word будет открыт новый документ, содержащий распознанный текст.

Замечание. Для данного сценария необходимо, чтобы на вашем компьютере была установлена программа Microsoft Word.

Сценарий Сканировать в PDF



Последовательность шагов сценария:

1. Сканирование бумажных документов.
Полученные со сканера изображения будут добавлены в текущий пакет ABBYY FinaReader.
2. Распознавание изображений.
3. Сохранение распознанного текста в формате PDF.
Все документы, обработанные в ходе выполнения сценария, сохраняются в файл формата PDF.

Пользовательские сценарии

Пользовательские сценарии предоставляют большее разнообразие шагов по сравнению со встроенными сценариями. В отличие от встроенных сценариев пользовательские сценарии могут содержать шаг проверки результатов распознавания, а также несколько шагов сохранения документа, что позволяет создавать копии одного и того же документа в различных форматах.

Последовательность шагов сценария должна соответствовать этапам обработки документа в программе ABBYY FineReader: *получение изображения – распознавание – проверка результатов – сохранение результатов распознавания*.

При создании сценария шаги добавляются со свойствами по умолчанию. Вы можете изменить свойства шагов и настроить процесс обработки документов наиболее удобным образом.

Пользовательский сценарий может включать в себя основные и дополнительные шаги.

Основные шаги пользовательского сценария

В сценарий можно добавить только один шаг получения изображения, один шаг распознавания и несколько шагов сохранения.

Получение изображения

Данный шаг всегда первый в сценарии, его цель – получение изображения для дальнейшей обработки.

Название шага	Свойства шага	Описание
Сканировать изображения		Сканирование выполняется с текущими настройками ABBYY FineReader
Открыть изображения	Запрашивать файлы изображений при запуске сценария (Свойство по умолчанию)	При запуске сценария программа запрашивает файлы изображений и добавляет их в текущий пакет ABBYY FineReader. В диалоге Открыть изображение укажите файлы, подлежащие обработке.
	Обработать изображения из папки	При запуске сценария программа откроет указанную папку, и добавит изображения, найденные в ней, в текущий пакет ABBYY FineReader.
Обработка изображений из пакета ABBYY FineReader	Запрашивать имя пакета при запуске сценария (Свойство по умолчанию)	При запуске сценария программа запрашивает пакет ABBYY FineReader. В открывшемся диалоге Открыть изображение укажите пакет ABBYY FineReader, изображения которого должны быть обработаны. При выборе данной опции необходимо указать папку, содержащую изображения. Для того чтобы при выполнении сценария были обработаны изображения из указанной папки и всех вложенных папок, отметьте опцию Обработать изображения из вложенных папок .
	Использовать текущий пакет	При запуске сценария начинается обработка изображений из текущего пакета.
	Использовать указанный пакет	При запуске сценария начинается обработка изображений из указанного пакета. Укажите путь к файлу пакета ABBYY FineReader, изображения из которого должны быть обработаны.

Анализ макета страницы

Название шага	Свойства шага	Описание
Загрузить шаблон блоков	Запрашивать файл шаблона блоков при запуске сценария (Свойство по умолчанию)	При выполнении сценария в диалоге Открыть укажите файл шаблона блоков.
	Использовать шаблон блоков из файла	Укажите путь к файлу шаблона блоков.
Анализ макета страницы	Анализировать страницы автоматически, а затем вносить исправления вручную (Свойство по умолчанию)	После получения изображения программа ABBYY FineReader выполнит анализ изображений и разметит на них блоки.
	Нарисовать блоки вручную	После получения изображения пользователь выполняет разметку блоков.

Распознавание

Название шага	Свойства шага	Описание
Распознать все страницы	Распознавание выполняется с текущими настройками ABBYY FineReader 8.0	Автоматическое распознавание изображений из указанного пакета или папки.

Проверка результатов распознавания

Название шага	Свойства шага	Описание
Проверить	Проверка орфографии (Свойство по умолчанию)	При переходе к этому шагу автоматически открывается диалог Проверка .
	Просмотр текста без проверки распознавания без проверки орфографии	Распознанный документ будет открыт в окне Текст .

Сохранение

Название шага	Свойства шага	Описание
Сохранить страницы	Запрашивать имена файлов при сохранении. (Свойство по умолчанию)	В этом диалоге необходимо указать имя файла, в который будет сохранен распознанный текст, а также можно выбрать опции сохранения. При переходе к шагу сохранения открывается диалог Сохранить текст как...
	Сохранить с указанными именами в указанную папку	<p>При выборе данного свойства необходимо указать:</p> <ol style="list-style-type: none"> 1. папку, в которую будет помещены файлы с распознанным текстом; Для того чтобы при каждом выполнении вам не приходилось задавать новую папку (т.е. перенастраивать сценарий), отметьте опцию при каждом запуске сценария сохранять файлы в новую папку, к имени которой добавлять время и дату; 2. формат файла; 3. режим сохранения страниц: <ul style="list-style-type: none"> • Каждая страница в отдельный файл – позволяет сохранить каждую страницу в отдельный файл; • Сохранять с именами исходных изображений – позволяет сохранить распознанные страницы в отдельные файлы выбранного формата, причем имена этих файлов будут совпадать с именами соответствующих файлов изображений; • Пофайловое деление по пустым страницам – все страницы рассматривается как совокупность групп страниц, разделенных пустыми страницами. Для каждой такой группы при сохранении образуется отдельный файл, в который последовательно записываются страницы группы; имена формируемых в этом режиме файлов образуются добавлением к указанному пользователем имени файла индексов следующего вида: –1, –2, –3 и т.д.; • Все страницы в один файл – позволяет сохранить все (или все выделенные) страницы пакета в один файл; 4. имя файла.

Сохранить изображения	Запрашивать имя файла при сохранении изображения (Свойство по умолчанию)	При переходе к шагу сохранения открывается диалог Сохранить изображение как... В этом диалоге необходимо указать имя файла, в который будет сохранен распознанный текст, а также можно выбрать опции сохранения.
	Сохранять изображения с указанными именами и в указанную папку	При выборе данного свойства необходимо указать: <ol style="list-style-type: none"> 1. папку, в которую будет помещены файл(ы) с изображениями; 2. формат файла; для того чтобы сохранить все изображения в один файл, отметьте опцию Сохранить в один многостраничный файл (только для форматов TIFF и PDF); 3. имя файла.

Дополнительные шаги пользовательского сценария

Дополнительные шаги сценария позволяют передать распознанный текст во внешнее приложение, создать письмо электронной почты с прикрепленным файлом, в которой можно поместить полученное изображение или распознанный текст и создать копию пакета ABBYY FineReader.

Передать

Название шага	Свойства шага	Описание
Передать во внешнее приложение	Мастер сохранения (Свойство по умолчанию)	Выберите нужное приложение из выпадающего списка. При выполнении сценария распознанный текст будет помещен в новый файл, открытый в выбранном приложении.

Отправка распознанного текста и изображения по электронной почте

Название шага	Свойства шага	Описание
Отправить страницы по электронной почте	Отправить как	Из выпадающего списка выберите нужный формат файла. Полный список поддерживаемых форматов см. в разделе «Форматы, в которых ABBYY FineReader сохраняет распознанный текст».
	Опции файла	При выборе данного свойства необходимо указать режим сохранения страниц:

		<p>Каждая страница в отдельный файл – позволяет сохранить каждую страницу в отдельный файл;</p> <p>Сохранять с именами исходных изображений – позволяет сохранить распознанные страницы в отдельные файлы выбранного формата, причем имена этих файлов будут совпадать с именами соответствующих файлов изображений;</p> <p>Пофайловое деление по пустым страницам – все страницы рассматривается как совокупность групп страниц, разделенных пустыми страницами. Для каждой такой группы при сохранении образуется отдельный файл, в который последовательно записываются страницы группы; имена формируемых в этом режиме файлов образуются добавлением к указанному пользователем имени файла индексов следующего вида: -1, -2, -3 и т.д.;</p> <p>Все страницы в один файл – позволяет сохранить все (или все выделенные) страницы пакета в один файл.</p>
--	--	---

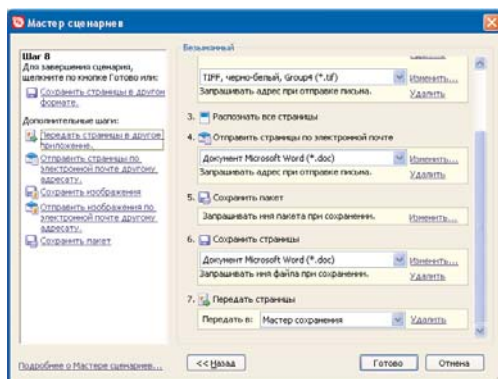
Сохранение пакета ABBYY FineReader

Название шага	Свойства шага	Описание
Сохранить пакет	Запрашивать имя пакета при сохранении (Свойство по умолчанию)	При переходе к данному шагу открывается диалог Сохранить пакет как... , в данном диалоге укажите имя пакета и папку, в которую будет сохранен пакет.
	Сохранить пакет в папку	Укажите имя файла пакета ABBYY FineReader.

Создание пользовательского сценария

- Запустите **Менеджер сценариев**. Для этого:
 - выберите команду **Менеджер сценариев** из списка команд кнопки **Scan&Read**, или
 - нажмите **Ctrl+T**,
 - в меню **Процесс** выберите команду **Сценарии ABBYY FineReader>Менеджер сценариев**, или
 - в меню **Сервис** выберите команду **Менеджер сценариев**.
- На панели инструментов **Менеджера сценариев** нажмите кнопку **Новый**.

3. В открывшемся диалоге введите имя сценария. Нажмите **ОК**.
4. Откроется окно **Мастера сценариев**. Мастер сценариев поможет вам выбрать шаги сценария и их свойства.



Диалог **Мастера сценариев** состоит из двух панелей. Левая панель содержит список шагов сценария, доступных для выбора. Состав списка зависит от уже выбранных шагов. На правой панели отображается список выбранных шагов сценария, который вы создадите, и их свойства.

5. На левой панели выберите один из доступных шагов. Выбранный шаг будет помещен на правую панель.
6. При добавлении в сценарий шаги имеют свойства по умолчанию. Значение свойства, с которым шаг был добавлен в сценарий, отображается в подсказке, расположенной под названием шага. Рядом с подсказкой находится кнопка **Изменить**. Щелкните по этой кнопке, чтобы увидеть все возможные свойства данного шага и выберите наиболее подходящие.
7. В строке подсказки к шагу сценария также находится кнопка **Удалить**, с помощью которой можно удалить из сценария ненужный шаг.

Поскольку выбор каждого следующего шага сценария зависит от шагов выбранных ранее, не все шаги сценария можно удалить по отдельности. Если в сценарий добавлен шаг **Распознавание**, то из сценария уже невозможно удалить шаг анализа макета страницы. В этом случае для удаления шага вам придется воспользоваться кнопкой **Назад** и вернуться к тому шагу, который вы хотите удалить.

8. После того, как вы добавили в сценарий все необходимые шаги и выбрали их настройки, нажмите кнопку **Готово**.

Созданный сценарий будет добавлен в список сценариев **Менеджера сценариев** и в список сценариев кнопки **Scan&Read**.

Глава 2

Приложение **ABBYY Screenshot Reader**

ABBYY Screenshot Reader – простое, удобное приложение, с помощью которого можно не только сделать снимок экрана («скриншот» от англ. «screenshot»), но и распознать текст с экрана компьютера.

Возможности ABBYY Screenshot Reader:

- Распознавание текста с любой части экрана.
- Распознавание таблицы с любой части экрана.
- Снимок любой области экрана.
- Сохранение результата в файл различного формата, буфер обмена или передача во внешнее приложение.
- **ABBYY Screenshot Reader** обладает простым и понятным интерфейсом, поэтому работа с ним не требует специальных знаний и усилий. На экране вашего компьютера может быть открыто любое приложение, интернет-сайты, фотографии вашего отпуска, **ABBYY Screenshot Reader** поможет вам сделать снимок любого участка экрана, распознать текст и сохранить изображение и текст.

Замечание. Воспользоваться приложением **ABBYY Screenshot Reader** могут все пользователи программы **ABBYY FineReader Corporate Edition** и зарегистрированные пользователи **ABBYY FineReader Professional Edition**.

Содержание

- Установка и запуск приложения
- Панель ABBYY Screenshot Reader
- Получение текста и таблиц с экрана компьютера
- Получение снимка экрана
- Дополнительные возможности приложения ABBYY Screenshot Reader

Установка и запуск ABBYY Screenshot Reader

Установка приложения ABBYY Screenshot Reader

Приложение **ABBYY Screenshot Reader** устанавливается вместе с программой ABBYY FineReader 8.0 по умолчанию. Если при выборочной установке ABBYY FineReader 8.0 приложение **ABBYY Screenshot Reader** не было включено в набор устанавливаемых компонентов, то для того чтобы установить **ABBYY Screenshot Reader** позднее:

1. Нажмите кнопку **Пуск** на **Панели задач** и выберите пункт **Настройка/Панель управления**.
2. Дважды щелкните мышью по иконке **Установка и удаление программ**.
3. Из списка установленных программ выберите программу **ABBYY FineReader 8.0** и нажмите кнопку **Изменить**.
4. В диалоге **Выборочная установка** выберите нужные компоненты.
5. Далее следуйте инструкциям программы установки.

Активация ABBYY Screenshot Reader

Замечание. Активация необходима только для пользователей **ABBYY FineReader 8.0 Professional Edition**.

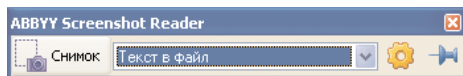
Для того чтобы воспользоваться программой **ABBYY Screenshot Reader**, вам необходимо стать зарегистрированным пользователем программы **ABBYY FineReader**. После регистрации вам будет выслан код программы **ABBYY Screenshot Reader**, который позволит вам активировать приложение **ABBYY Screenshot Reader**. Получив код **Screenshot Reader**, введите в соответствующее поле в диалоге, возникающем при запуске приложения **ABBYY Screenshot Reader**.

Запуск приложения Screenshot Reader

Для того чтобы запустить приложение ABBYY Screenshot Reader, в меню **Пуск>Программы>ABBYY FineReader 8.0** выберите пункт **ABBYY Screenshot Reader**.

Панель ABBYY Screenshot Reader

После запуска приложения **ABBYY Screenshot Reader** на экране появляется панель **ABBYY Screenshot Reader**.



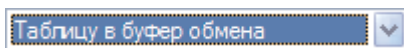
Панель **ABBYY Screenshot Reader** содержит инструменты для распознавания текста и таблиц с экрана компьютера, получения снимка выбранной области экрана, выбора дополнительных настроек приложения **ABBYY Screenshot Reader**.

Кнопка Снимок.

При нажатии на данную кнопку курсор меняет форму и включается инструмент, позволяющий выде-

лить область экрана. Выделенная часть изображения будет заключена в рамку, и автоматически начнет-ся распознавание текста, или будет сделан снимок экрана. Распознанный текст, таблица или снимок экрана будет сохранен в файл, помещен в буфер обмена или передан во внешнее приложение согласно пункту, выбранному из выпадающего списка, приведенного ниже.

Выпадающий список




Данный список позволяет выбрать объект для извлечения и способ его сохранения.

Кнопка .

Открывает диалог **Опции – ABBYY Screenshot Reader**, в котором можно выбрать язык распознавания, режим отображения панели **ABBYY Screenshot Reader** и сообщение или звуковой сигнал, выдаваемый приложением **ABBYY Screenshot Reader** при завершении копирования в буфер обмена.


Кнопка / .

Данная кнопка позволяет управлять отображением панели **ABBYY Screenshot Reader**. Если выбран режим , панель **ABBYY Screenshot Reader** будет располагаться поверх всех окон при переключении в другие приложения.

Получение текста и таблиц с экрана компьютера

Приложение **ABBYY Screenshot Reader** позволяет распознать текст или таблицу с экрана компьютера и сохранить результаты распознавания в файл выбранного формата, буфер обмена или передать во внешнее приложение.

Для того чтобы распознать текст с экрана:


- Из выпадающего списка панели **ABBYY Screenshot Reader** выберите один из пунктов:
 - **Текст в буфер обмена**,
 - **Таблицу в буфер обмена**,
 - **Текст в Microsoft Word**,
 - **Таблицу в Microsoft Excel**,
 - **Текст в файл**,
 - **Таблицу в файл**.
- Нажмите кнопку **Снимок**, курсор изменится на .
- Установите курсор мыши в угол интересующей вас области экрана. Нажмите левую кнопку мыши и, не отпуская кнопки, потяните в противоположный по диагонали угол. Отпустите кнопку мыши. Выделенная часть изображения будет заключена в рамку, и автоматически начнется распознавание текста.
- После того, как программа завершит распознавание текста, распознанный текст будет помещен согласно выбору, сделанному в п.1: в буфер обмена (*Текст в буфер обмена* или *Таблицу в буфер обмена*), в новый файл (*Текст в Microsoft Word* или *Таблицу в Microsoft Excel*). Если была выбрана опция *Текст в файл* или *Таблицу в файл*, то откроется диалог **Сохранить текст как**. В этом диалоге укажите имя и формат файла, в который будут сохранен распознанный текст.

Замечание. Для отмены распознавания нажмите кнопку **Stop** на индикаторе хода распознавания или на клавишу **Esc**.

Что делать, если...

Вы работаете с текстами, написанными на разных языках

Перед запуском процесса извлечения текста убедитесь в том, что язык распознавания, выбранный в диалоге **Опции – ABBYY Screenshot Reader**, соответствует языку, интересующего вас текста. Если необходимо, выберите другой язык распознавания.

Замечание. Для того чтобы открыть диалог **Опции – ABBYY Screenshot Reader**, нажмите кнопку , в открывшемся диалоге **Опции – ABBYY Screenshot Reader** из списка **Язык распознавания**, выберите язык, на котором написан интересующий вас текст.

Текст на экране написан на нескольких языках?

В диалоге **Опции – ABBYY Screenshot Reader** из списка **Язык распознавания** выберите пункт **(Выбор нескольких языков...)**, в открывшемся диалоге отметьте нужные языки.

Внимание! Увеличение количества подключенных к распознаванию одного документа языков может привести к ухудшению качества распознавания. Не рекомендуется подключать более 2–3 языков.

Вы не нашли нужный язык в списке языков распознавания

В диалоге **Опции – ABBYY Screenshot Reader** из списка **Язык распознавания** выберите пункт **(Выбор из полного списка языков...)**, в открывшемся диалоге отметьте нужные языки.

Если же нужного языка нет и в полном списке языков распознавания, возможно:

- Данный язык не поддерживается системой ABBYY FineReader. Полный список языков приведен в разделе «Поддерживаемые языки распознавания».
- Язык был отключен при выборочной установке. Чтобы установить языки распознавания:
 1. В меню **Пуск** на **Панели Задач** выберите пункт **Настройка/Панель управления**.
 2. В диалоге **Панель управления** выберите пункт **Установка и удаление программ**.
 3. Из списка установленных программ выберите программу **ABBYY FineReader 8.0** и нажмите кнопку **Изменить**.
 4. В диалоге **Выборочная установка** выберите нужные языки.
 5. Далее следуйте инструкциям программы установки.


Область экрана содержит и текст и иллюстрации

На панели **ABBYY Screenshot Reader** из выпадающего списка выберите пункт *Изображение в ABBYY FineReader*. Изображение выбранной области экрана будет добавлено в пакет ABBYY FineReader. Далее обработайте изображение в программе ABBYY FineReader. Если на вашем компьютере уже открыт ABBYY FineReader, изображение будет добавлено в текущий пакет, в противном случае изображение будет добавлено в новый пакет.

Получение снимка экрана

Приложение **ABBYY Screenshot Reader** позволяет сделать снимок выбранной области экрана и сохранить его в файл выбранного формата, буфер обмена или передать в программу ABBYY FineReader.

Для того чтобы получить снимок выбранной области экрана:

1. Из выпадающего списка панели ABBYY Screenshot Reader выберите один из пунктов:
Изображение в буфер обмена,
Изображение в ABBYY FineReader – выберите эту опцию, если интересующая вас область экрана, содержит иллюстрации и текст, программа ABBYY FineReader поможет вам обработать сложное изображение (если на Вашем компьютере уже открыт ABBYY FineReader, изображение будет добавлено в текущий пакет, в противном случае изображение будет добавлено в новый пакет),
Изображение в файл.
2. Нажмите кнопку **Снимок**, курсор изменится на .
3. Установите курсор мыши в угол интересующей вас области экрана. Нажмите левую кнопку мыши и, не отпуская кнопки, потяните в противоположный по диагонали угол. Отпустите кнопку мыши. Выделенная часть изображения будет заключена в рамку.
4. Готовое изображение будет помещено согласно выбору, сделанному в п.1: в буфер обмена (*Изображение в буфер обмена*) или добавлено в пакет ABBYY FineReader (*Изображение в ABBYY FineReader*). Если была выбрана опция *Изображение в файл*, то откроется диалог **Сохранить текст как**. В этом диалоге укажите имя и формат файла, в который будет сохранено полученное изображение.

Замечание. Для отмены снимка выбранной области экрана нажмите на клавишу **Esc**.

Дополнительные настройки приложения ABBYY Screenshot Reader

Настройками приложения **ABBYY Screenshot Reader** выбираются в диалоге **Опции – ABBYY Screenshot Reader**. Для того чтобы открыть этот диалог, нажмите кнопку  на панели **Screenshot Reader**. В этом диалоге вы можете:

- Выбрать язык распознавания, соответствующий языку текста, нужной области экрана.
- Отметить опцию **Поверх других окон**, для того чтобы панель **ABBYY Screenshot Reader** распо-

лагалась поверх всех окон при переключении в другие приложения.

- Отметить опцию **Выдавать звуковое сообщение о том, что текст или изображение были помещены в буфер обмена**. Завершив копирование в буфер обмена, ABBYY Screenshot Reader сообщит вам об этом.
- Отметить опцию **Показывать сообщение о том, что текст или изображение были помещены в буфер обмена**. Завершив копирование в буфер обмена, ABBYY Screenshot Reader сообщит вам об этом.

Глава 3

Приложение **ABBYY Hot Folder & Scheduling**

Версия ABBYY FineReader 8.0 Corporate Edition включает в себя приложение **Hot Folder & Scheduling**. Это приложение-планировщик, с помощью которого можно назначить время обработки документов из выбранной папки. Открытие файлов и распознавание изображений может быть выполнено в то время, когда компьютер наименее загружен, например, ночью.

Для автоматической обработки изображений необходимо настроить задачу для выбранной папки. Настройки задачи включают в себя все необходимые параметры открытия изображений, распознавания и сохранения распознанного текста, режим проверки папки (однократно или постоянно), время начала обработки папки и периодичность проверки папки.

Содержание:

- Установка и запуск приложения **ABBYY Hot Folder & Scheduling**
- Главное окно приложения
- Настройка задачи для папки
- Отчет об обработке папки
- Дополнительные возможности **ABBYY Hot Folder & Scheduling**

Установка и запуск ABBYY Hot Folder & Scheduling

Приложение **ABBYY Hot Folder & Scheduling** устанавливается вместе с программой ABBYY FineReader 8.0. Если при выборочной установке ABBYY FineReader 8.0 приложение **ABBYY Hot Folder & Scheduling** не было включено в набор устанавливаемых компонентов программы ABBYY FineReader 8.0, то для того чтобы установить **ABBYY Hot Folder & Scheduling** позднее:

1. Нажмите кнопку **Пуск** на Панели задач и выберите пункт **Настройка/Панель управления**.
2. Дважды щелкните мышью по иконке **Установка и удаление программ**.
3. Из списка установленных программ выберите программу **ABBYY FineReader 8.0** и щелкните по кнопке **Изменить**;
4. В диалоге **Выборочная установка** выберите компонент **ABBYY Hot Folder & Scheduling**;
5. Далее следуйте инструкциям программы установки.

Запуск приложения ABBYY Hot Folder & Scheduling

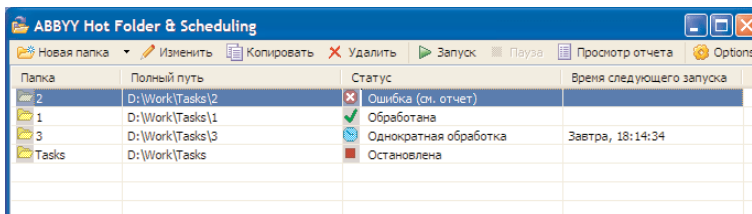
Для того чтобы запустить приложение **ABBYY Hot Folder & Scheduling**:

- В меню **Пуск>Программы>ABBYY FineReader 8.0** выберите пункт **ABBYY Hot Folder & Scheduling**, или
- В меню **Сервис** программы **ABBYY FineReader 8.0** выберите команду **ABBYY Hot Folder & Scheduling**, или
- Щелкните по иконке приложения, расположенной на панели задач Windows (данная иконка появляется после того, как на компьютере была настроена хотя бы одна задача).

Замечание. Если вам не удастся запустить задачу для выбранной папки, убедитесь в том, что на вашем компьютере включена служба планировщика заданий, или обратитесь к системному администратору.

Главное окно ABBYY Hot Folder & Scheduling

При запуске приложения **ABBYY Hot Folder & Scheduling** открывается главное окно.










На панели инструментов **ABBYY Hot Folder & Scheduling** расположены кнопки, позволяющие настраивать, копировать, удалять задачи для папок, а также запускать обработку папок и просматривать отчет об обработке папки.

Нажмите кнопку	Для того чтобы...
 Новая	Настроить задачу. Мастер Hot Folder & Scheduling поможет настроить задачу для выбранной папки.
Экспорт...	Экспортировать файл задачи. Этот файл затем можно передать другому пользователю. В открывшемся диалоге укажите имя файла, в который будут экспортированы настройки задачи (файлы задач имеют расширение *.hft). Замечание. По умолчанию файлы задач сохраняются в папку %UserProfile%\Local Settings\Application Data\ABBYY\HotFolder\8.00
Импорт...	Импортировать файл задачи. В открывшемся диалоге Импортировать файл задачи укажите путь файлу задачи. Замечание. Для выполнения импортированной задачи в системе должна существовать папка, путь к которой указан в настройках задачи, а также должны быть установлены все необходимые языки распознавания.
 Изменить	Изменить задачу. Мастер Hot Folder & Scheduling поможет вам выбрать новые настройки.
 Копировать	Создать копию задачи. Копия задачи будет добавлена в список задач сразу после оригинала. Копия задачи имеет статус Остановлена . На основе копии задачи можно создать новую задачу, выбрав для нее нужные настройки.
 Удалить	Удалить ненужную задачу.
 Запустить	Начать обработку документов согласно выбранному заданию.
 Пауза	Позволяет остановить выполнение задачи.
 Просмотр отчета	Позволяет просмотреть отчет об обработке выбранной папки, в отчет вносятся все события процесса обработки.
 Опции	Выбрать дополнительные настройки приложения Hot Folder & Scheduling .

В окне **ABBYY Hot Folder & Scheduling** располагается список настроенных задач. Для каждой задачи указывается полный путь к обрабатываемой папке, статус папки и время следующего сеанса обработки папки.

Статусы задач

Статусы задач отражают ход выполнения обработки папки. В процессе выполнения задача может иметь следующие статусы:

Статус	Описание
 Обработка	Выполняется обработка изображений, находящихся в папке.
 Ожидание	Время обработки папки наступило, но не закончена обработка предыдущего задания.
 Однократная обработка	Для данной папки выбран режим разовой обработки, момент начала обработки указан в столбце Следующий сеанс обработки.
 Проверяемая	Изображения из данной папки будут обрабатываться по мере их поступления в папку.
 Остановлена	Обработка папки приостановлена.
 Обработана	Обработка папки завершена.
 Ошибка	В процессе обработки папки произошла ошибка, и обработка изображений не была завершена. Описание ошибок приводится в отчете об обработке папки. Для того чтобы установить причину ошибки, выберите папку и нажмите кнопку Просмотр отчета, расположенную на панели инструментов.

Настройка задачи для папки

Мастер Hot Folder & Scheduling поможет вам настроить задачу для папки. Для того чтобы открыть окно мастера щелкните по кнопке Новая на панели инструментов приложения ABBYY Hot Folder & Scheduling.

Процесс настройки задачи для папки состоит из трех шагов.

Шаг 1. Получения изображения

На данном шаге необходимо указать папку, из которой будут браться изображения. Также на этом шаге необходимо выбрать режим проверки папки: однократная проверка или многократная проверка.

Выбор настроек шага

1. В выпадающем списке **Выбор типа папки** укажите тип папки:

- Пункт **Открыть изображение из папки** позволяет обрабатывать изображения из локальных и сетевых папок.
- Пункт **Открыть изображение из FTP папки** позволяет обрабатывать изображения из папок, расположенных на FTP сервере.
Если для получения доступа к FTP папке необходимо указать логин и пароль, введите их в поля **Логин FTP** и **Пароль FTP**. Если же для доступа к FTP папке не требуются логин и пароль, отметьте опцию **Анонимное подключение**.

2. В поле **Папка** укажите полный путь к обрабатываемой папке.

3. В поле **Тип файлов** укажите формат файлов изображений.
4. Отметьте опцию **Делить разворот книги**, для того чтобы разделять изображения, содержащие разворот книги (сдвоенные страницы), на две страницы пакета. Это позволяет улучшить качество распознавания, а также сохранить каждую страницу в отдельный файл.
5. В группе **Обработанные изображения** задания укажите, нужно ли удалять изображения из папки входящих изображений, если задача была выполнена успешно.
 - Если выбрана опция **Не удалять входящие изображения из папки**, после завершения сеанса обработки обработанные изображения будут оставлены в папке входящих изображений.
 - Если выбрана опция **Удалять изображения из папки**, после успешного завершения сеанса обработки из папки входящих изображений будут удалены обработанные изображения.
6. Группа **Режим проверки папки** позволяет выбрать режим обработки папки:
 - Если выбрана опция **Проверять однократно при запуске**, при запуске задачи будут обработаны все изображения, находящиеся на этот момент в папке. На этом обработка папки закончится.
 - Опция **Проверять постоянно** позволяет обрабатывать папки, в которые постоянно добавляются изображения. Интервал проверки папки – 1 минута. При обнаружении новых файлов выполняется их обработка.
7. В поле **Время запуска** укажите момент начала обработки папки. Выберите один из пунктов:
 - **Начать сейчас** – при выборе данного пункта выполнение задачи начнется сразу же после того, как задача будет настроена.
 - **Начать в указанное время**. Если выбран данный пункт, укажите дату и время начала обработки.
 - **Отклонить** – при выборе данного пункта выполнение задачи будет отложено.

Шаг 2. Распознавание

На данном шаге необходимо выбрать настройки распознавания.

Выбор настроек шага

1. В выпадающем списке **Язык распознавания** выберите язык, соответствующий тексту на изображениях, которые вы собираетесь распознать.

Замечание. Возможен выбор нескольких языков распознавания.

2. Группа **Режим распознавания** позволяет выбрать режим распознавания:
 - Режим **Тщательный** позволяет распознавать изображения даже низкого качества.
 - Режим **Быстрый** рекомендуется только для изображений хорошего качества.
3. В группе **Распознавание PDF документов** выберите:

- **Извлечь текст из PDF документа**

Если данный пункт отмечен, то программа ABBYY FineReader извлекает текстовые данные из PDF документа и использует их при распознавании. Такой способ ускоряет обработку PDF документов в 2 – 3 раза.

- **Распознавать PDF как изображение**

Данную опцию следует применять при обработке PDF файлов с нестандартной кодировкой. Текст, извлеченный из таких файлов, содержит много ошибок, и его необходимо распознавать.

5. В группе **Гиперссылки**:

- отметьте опцию **Выделять гиперссылки**. Для того, чтобы выбрать цвет выделения гиперссылок, щелкните по кнопке **Цвет** и в открывшемся диалоге выберите нужный цвет.

6. В группе **Обучение** выберите одну из опций:

- **Не использовать пользовательский эталон**. Данная опция отмечена по умолчанию.
- **Использовать пользовательские эталоны**. Если выбрана данная опция, при распознавании будут использованы пользовательские эталоны. Подробнее о пользовательских эталонах см.

Шаг 3. Сохранение распознанного документа и изображения

На данном шаге необходимо указать папку сохранения результатов распознавания и выбрать формат сохранения.

Выбор настроек

1. В поле **Сохранить в папку** распознавания укажите путь к папке.

2. Группа **Режим сохранения** позволяет выбрать режим сохранения страниц и имя файла.

В выпадающем списке **Опции файла** выберите один из пунктов:

- Пункт **Сохранять каждую страницу в отдельный файл** – позволяет сохранить каждую страницу в отдельный файл;
- Пункт **Именовывать файлы по исходным изображениям** – позволяет сохранить распознанные страницы в отдельные файлы выбранного формата, причем имена этих файлов будут совпадать с именами соответствующих файлов изображений;
- Пункт **Пофайловое деление по пустым страницам** – все страницы рассматривается как совокупность групп страниц, разделенных пустыми страницами. Для каждой такой группы при сохранении образуется отдельный файл, в который последовательно записываются страницы группы; имена формируемых в этом режиме файлов образуются добавлением к указанному пользователем имени файла индексов следующего вида: -1, -2, -3 и т.д.;
- Пункт **Сохранять все страницы в один файл** – позволяет сохранить все (или все выделенные) страницы в один файл.
- В поле **Имя файла** укажите имя, которое будет присвоено файлам, содержащим распознанный текст.

3. Отметьте опцию **Сохранить пакет**, для того чтобы сохранить пакет ABBYY FineReader. Для того чтобы указать место расположения сохраненного пакета, нажмите кнопку **Обзор**, и в открывшемся диалоге укажите путь к папке, в которую будет сохранен пакет.

Совет. Вы можете включить в процесс обработки папки несколько шагов сохранения, и создать копии документов в различных форматах. Для этого в выпадающем списке **Сохранить как** выберите пункт **Выбор нескольких форматов**. В открывшемся диалоге отметьте нужные форматы.

Отчет об обработке папки

Обработка изображений из папки, для которой была настроена задача обработки, выполняется автоматически. В ходе обработки приложение **Hot Folder & Scheduling** фиксирует все события процесса обработки, и регистрирует их в отчете об обработке папки в процессе выполнения задачи.

Данный отчет содержит:

- Название задачи и перечень настроек задачи,
- Список ошибок и предупреждений,
- Статистику выполнения задачи (количество обработанных страниц, количество ошибок и предупреждений, выявленных в ходе обработки, количество неуверенно распознанных символов).

Просмотр отчета об обработке папки

Для того чтобы просмотреть отчет:

1. В списке задач главного окна приложения **Hot Folder & Scheduling** выберите задачу
2. На панели инструментов нажмите кнопку **Просмотр отчета**.

Дополнительные настройки приложения ABBYY Hot Folder & Scheduling

Дополнительные настройки приложения ABBYY Hot Folder & Scheduling выбираются в диалоге **Опции**. Для того чтобы открыть этот диалог, нажмите кнопку **Опции** на панели **ABBYY Hot Folder & Scheduling**. В этом диалоге вы можете:

- Отметить **Сохранять отчет в папке сохранения распознанного текста**; отчет об обработке папки будет сохранен в файл формата TXT в той же папке, что и файлы с распознанным текстом.
- Отметить опцию **Выдавать сообщение при завершении задачи**. Сообщение о завершении задачи отображается при наведении курсора мыши на иконку приложения, расположенную на панели задач рабочего стола Windows.

Приложение

Содержание

- Форматы, в которых ABBYY FineReader сохраняет распознанный текст
- Поддерживаемые графические форматы
- Горячие клавиши
- Глоссарий

Форматы, в которых ABBYY FineReader сохраняет распознанный текст

ABBYY FineReader сохраняет результаты распознавания в следующих форматах:

Microsoft Word Document (*.DOC)

Rich Text Format (*.RTF)

Microsoft Word XML Document (*.XML) (только для Microsoft Office Professional Edition 2003)

Adobe Acrobat Format (*.PDF)

HTML

Microsoft PowerPoint Format (*.PPT)

Comma Separated Values File (*.CSV)

Простой текст (*.TXT). ABBYY FineReader поддерживает различные кодовые страницы (Windows, DOS, Mac, ISO) и кодировку Unicode

Microsoft Excel Spreadsheet (*.XLS)

DBF

Microsoft Reader eBook (*.LIT)

Поддерживаемые графические форматы

ABBY FineReader открывает файлы следующих форматов:

PDF: Файлы в формате PDF, версии 1.6 или более ранней

BMP:

2-битный – черно-белый

4- и 8-битный – Palette

16-битный

24-битный – Palette

TrueColor 32-битный

PCX, DCX:

2-битный – черно-белый

4- и 8-битный – Palette

24-битный – TrueColor

JPEG: серый и цветной

JPEG 2000: серый и цветной

TIFF:

черно-белый – несжатый, CCITT3, CCITT3FAX, CCITT4, Packbits, ZIP, LZW

серый – несжатый, Packbits, JPEG, ZIP, LZW

цветной – несжатый, JPEG, ZIP, LZW

Palette – несжатый, Packbits, ZIP, LZW

многостраничный TIFF

PNG: черно-белый, серый, цветной

ABBY FineReader сохраняет изображения в следующих форматах:

BMP: черно-белый, серый, цветной

PCX: черно-белый, серый

JPEG: серый, цветной

JPEG 2000: серый и цветной

TIFF:

черно-белый – несжатые, CCITT3, CCITT4, Packbits, ZIP

серый – несжатый, Packbits, JPEG, ZIP

цветной – несжатый, JPEG, ZIP

многостраничный TIFF

PNG: черно-белый, серый, цветной

Горячие клавиши

Меню	Команда	Нажмите:
Файл	Открыть PDF/изображение	CTRL+O
	Сканировать изображение	CTRL+K
	Остановить сканирование	Esc
	Новый пакет	CTRL+N
	Открыть пакет	CTRL+SHIFT+N
	Сохранить страницы	CTRL+S
	Отправить страницы по электронной почте	CTRL+M
	Сохранить изображения	CTRL+ALT+S
	Отправить изображения по электронной почте	CTRL+ALT+M
Правка	Отменить	CTRL+Z
	Восстановить	CTRL+Enter
	Вырезать	CTRL+X
	Копировать выделенный фрагмент в буфер обмена	CTRL+INS или CTRL+C
	Вставить содержимое буфера обмена	CTRL+V или SHIFT+INS
	Выделить весь текст в окне Текст , все страницы в окне Пакет или все блоки на открытой странице	CTRL+A
	Найти указанный текст	CTRL+F
	Найти следующее вхождение указанного текста	F3
	Найти и заменить указанный текст	CTRL+H
	Расширенный поиск	ALT+F3
Вид	Развернуть окно Пакет	CTRL+0
	Показать окно Изображение	CTRL+F2
	Увеличить изображение в окне Изображение	CTRL+SHIFT+NUM+
	Уменьшить изображение в окне Изображение	CTRL+SHIFT+NUM-
	Показать активный блок	CTRL+SHIFT+NUM*
	Показать окно Текст	CTRL+F3
	Показывать окно Крупный план	CTRL+F5
	Свойства	ALT+ENTER
Пакет	Открыть следующую страницу пакета	CTRL+NUM+

	Открыть предыдущую страницу пакета	CTRL++NUM–
	Открыть страницу с номером	CTRL+G
	Закрыть текущую страницу	CTRL+F4
Изображение	Изменить тип блока на тип Зона распознавания	CTRL+1
	Изменить тип блока на тип Текст	CTRL+2
	Изменить тип блока на тип Таблица	CTRL+3
	Изменить тип блока на тип Картинка	CTRL+4
	Изменить тип блока на тип Штрих-код	CTRL+5
	Удалить все блоки в окне Изображение и распознанный текст в окне Текст	CTRL+DEL
	Удалить распознанный текст в окне Текст	CTRL+SHIFT+DEL
	Разбить изображение	CTRL+SHIFT+I
	Обрезать изображение	CTRL+SHIFT+C
	Исправить разрешение	CTRL+SHIFT+T
Процесс	Сканировать и распознать изображение	CTRL+D
	Открыть и распознать изображение или PDF-файлы	CTRL+ALT+D
	Открыть PDF/изображение	CTRL+O
	Сканировать изображение	CTRL+K
	Остановить сканирование	Esc
	Распознать открытые или выбранные страницы	CTRL+R
	Распознать все страницы пакета	CTRL+SHIFT+R
	Распознать выделенный или выбранный блок	CTRL+SHIFT+B
	Анализ макета страницы	CTRL+E
	Анализ всех страниц пакета	CTRL+SHIFT+E
	Проверить распознанный текст	F7
	Перейти к следующей ошибке	F4
	Перейти к предыдущей ошибке	SHIFT+F4
	Сохранить страницы	CTRL+S
	Отправить страницы по электронной почте	CTRL+M
	Сохранить изображения	CTRL+ALT+S
	Отправить изображения по электронной почте	CTRL+ALT+M

	Открыть Менеджер сценариев	CTRL+T
Сервис	Запустить приложение ABBYY Hot Folder & Scheduling	CTRL+SHIFT+H
	Поверить распознанный текст	F7
	Просмотреть и отредактировать словарь	CTRL+SHIFT+D
	Открыть диалог Редактор языков , в котором вы можете создать и отредактировать язык или группу языков	CTRL+SHIFT+L
	Открыть диалог Редактор эталонов , в котором вы можете создать и отредактировать пользовательские эталоны	CTRL+SHIFT+A
	Открыть диалог Опции	CTRL+SHIFT+O
Окна	Открыть следующее окно	CTRL+F6
	Открыть предыдущее окно	CTRL+SHIFT+F6
	Открыть окно Пакет	ALT+1
	Открыть окно Изображение	ALT+2
	Открыть окно Текст	ALT+3
	Открыть окно Крупный план	ALT+4
	Перейти в диалог Поиск	ALT+5
	Открыть диалог Поиск	ALT+F3
Справка	Открыть справку	F1
Общие	Отметить выделенный фрагмент текста как полужирный	CTRL+B
	Отметить выделенный фрагмент текста как курсив	CTRL+I
	Подчеркнуть выделенный фрагмент текста	CTRL+U
	Перейти к следующей ячейке таблицы	Стрелки влево, вправо, вниз и вверх

Глоссарий

А

ABBYY Hot Folder & Scheduling – приложение-планировщик, с помощью которого можно назначить время обработки документов из выбранной папки. Открытие файлов и распознавание изображений может быть выполнено в то время, когда компьютер наименее загружен, например, ночью.

ABBYY Screenshot Reader – приложение, позволяющее сделать снимок выбранной области экрана и распознать текст, содержащийся на нем.

ADF (Automatic Document Feeder) – устройство автоматической подачи бумаги, позволяющее отсканировать большое количество документов без ручного вмешательства. ABBYY FineReader поддерживает сканирование многостраничных документов.

dpi (Dots per Inch) – количество точек на дюйм; единица измерения разрешения.

Scan&Read – главная кнопка программы, по нажатию на которую система сканирует изображение и распознает (анализирует макет страницы и распознает документ).

TWAIN, TWAIN-диалог – диалог сканера.

Unicode – стандарт, разработанный концерном Unicode. Управляет процессом кодирования символов и предоставляет шестнадцатибитную международную расширяемую систему кодировки символов для обработки информации на большинстве языков народов мира. Стандарт Unicode определяет кодировку символа, а также свойства и алгоритмы, используемые при реализации этого процесса.

Аббревиатура – сокращение, образованное из начальных букв слов. Например, MY, MS-DOS и т.д.

Активация – процесс получения пользователем в компании ABBYY специального кода, необходимого для обеспечения работы программы в полнофункциональном режиме на конкретном компьютере.

Активационный код – код, который выдается пользователю после прохождения процедуры активации (для версии Professional Edition). Предназначен для активации продукта на той машине, на которой был получен Installation ID.

Активационный файл – файл, который выдается пользователю после прохождения процедуры активации (для версии Corporate Edition). В нем содержится информация, необходимая для активации сервера или компьютера (в случае установки на отдельное рабочее место). Рабочие станции активируются сервером.

Автоподбор яркости – автоматический подбор яркости, производится сканером или системой ABBYY FineReader. Автоподбор позволяет подбирать яркость для каждого участка изображения в отдельности.

Активный блок – блок, над которым выполняется большинство производимых действий (например, удаление, изменение типа и т.д.). На углах такого блока имеются «квадратики».

Анализ макета страницы (выделение блоков) – процесс выделения блоков на изображении. Блоки могут быть разных типов. Анализ макета может проводиться автоматически одновременно с распознаванием при нажатии кнопки **2-Распознать** или вручную до распознавания.

Б

Блок – участок изображения, выделенный в рамку.

Д

Драйвер – программа, управляющая устройством.

З

Запрещенные символы – в качестве запрещенных символов для текущего языка указываются те, которые заведомо не могут встречаться в текстах, распознаваемых с подключением данного языка. Указание таких символов может существенно увеличить скорость и надежность распознавания. Например, при распознавании текстов, в которых встречаются только заглавные буквы, в качестве запрещенных следует указать все строчные буквы.

Зона распознавания – блок используется для распознавания и автоматического анализа части изображения. После нажатия на кнопку **2-Распознать** выделенный блок автоматически анализируется и распознается.

И

Игнорируемые символы – в качестве игнорируемых символов указываются те, которые могут встречаться внутри слова, например, знаки слогоделения или ударения в словарях. При проверке по словарю программа не учитывает эти знаки. В распознанном тексте эти символы сохраняются, но при проверке орфографии не учитываются.

Инвертированное изображение – изображение, текст которого напечатан на темном фоне светлыми буквами.

К

Картинка – блок, используемый для выделения картинок. Он может содержать картинку или любую другую часть текста, которую вы хотите передать в распознанный текст в качестве картинки.

Кодовая страница – таблица, в которой задано отношение между кодами символов и их начертаниями. В системе ABBYY FineReader кодовая страница представлена как набор символов, из которого можно выбрать нужные символы.

Л

Лигатура – пара или тройка букв, которые из-за особенностей начертания склеиваются между собой. Например, ge, go, fe, ffi, tt и т.п.

М

Мастер Scan&Read – специальный режим сканирования и распознавания, в котором система контролирует действия пользователя и подсказывает ему, что делать, чтобы получить тот или иной результат.

Макет страницы – расположение текста, таблиц и картинок в документе, разбиение на абзацы, гарнитура и размер шрифта, колонки, направление текста, цвет букв и фон текста.

Менеджер лицензий – приложение, в котором осуществляется управление лицензиями и активация ABBYY FineReader 8.0 Corporate Edition.

Менеджер сценариев – встроенный менеджер, позволяющий запускать выполнение сценариев, создавать сценарии и настраивать их, а также удалять неиспользуемые пользовательские сценарии.

Моноширинный шрифт – любой шрифт, буквы которого имеют одинаковую ширину (моно). Например, шрифт Courier New. Для улучшения качества распознавания моноширинных шрифтов на закладке **Распознать** в пункте **Тип печати** установите переключатель в положение **Пишущая машинка**.

Мягкий перенос – перенос (¬) показывает, в каком именно месте должно быть разорвано слово или словосочетание (например, «Автоформат»), если оно попадает на конец строки (например, «Авто-формат»). Все переносы в словарных словах ABBYY FineReader заменяет на мягкий перенос.

Н

Набор опций – совокупность значений опций, расположенных на закладках **Сканировать/Открыть**, **Распознать**, **Проверить** и **Сохранить** диалога **Опции**, а также в диалогах **Форматы** и **Дополнительные опции**. В набор опций также входят пользовательские языки и эталоны. Набор опций можно сохранить и затем использовать (загружать) в другие пакеты ABBYY FineReader.

Начертание шрифта – способ выделения в тексте (полужирный, наклонный, с подчеркиванием, перечеркнутый, верхний индекс, нижний индекс, малые прописные).

Начальная форма – форма, в которой слова даются в словаре. Для существительных это форма именительного падежа единственного числа; для прилагательных – форма мужского рода единственного числа именительного падежа; для глаголов, причастий и деепричастий – это неопределенная форма (инфинитив).

О

Обучение – создание пар «растровое изображение – название символа».

Омнифонтовая система – система распознавания, которая распознает символы практически любых размеров и начертаний.

Очистка изображения – удаление отдельных мелких черных точек на изображении.

П

Пакет — это папка, в которой хранятся изображения и рабочие файлы программы. В пакете может содержаться до 9999 страниц. В один пакет для удобства работы рекомендуется объединять изображе-

ния, логически связанные между собой (например, страницы одной книги, тексты на одном языке или изображения с однотипным расположением текста и т.д.).

Парадигма – совокупность всех грамматических форм слова.

Параметры защиты PDF – ограничения на открытие, редактирование копирование и печать PDF документа. В эти параметры входят пароль открытия документа, пароль для изменения прав доступа и уровень шифрования.

Пароль для изменения прав доступа – пароль, присваиваемый PDF документам. Пользователи могут распечатать или внести изменения в PDF документ только после ввода пароля, заданного автором документа. При использовании защиты PDF документа пользователи также должны будут указать этот пароль, для того чтобы изменить параметры защиты PDF.

Пароль открытия документа – пароль, присваиваемый PDF документам. Пользователи могут открыть PDF документ только после ввода пароля, заданного автором документа.

Префиксы – небуквенные символы, которые могут встречаться в абсолютном начале слова.

Принцип одной кнопки – принцип, когда при нажатии на одну кнопку (Scan&Read) система сканирует и распознает документ.

Р

Разделители – символы, которые могут разделять слова, например, /, \, тире и т.п. и которые пишутся отдельно от слов.

Разрешение – параметр сканирования, показывает кол-во точек на единицу длины. 300 dpi (размер шрифта 10 и более пунктов), 400–600 dpi для текстов, набранных мелким шрифтом (9 и менее пунктов).

С

Свойства документа— свойства, присваиваемые документу и позволяющие выполнять поиск документов и их сортировку. Свойства документа включают описательное название, имя автора, тему и ключевые слова.

Сканер – устройство, предназначенное для ввода изображений в компьютер.

Сложное слово – слово, отсутствующее в словаре, но которое может быть образовано из имеющихся в словаре слов.

Сценарий ABBYY FineReader – это набор последовательных шагов, каждый из которых соответствует одному этапу обработки документа. В программе ABBYY FineReader 8.0 предусмотрено три встроенных сценария, полностью готовых к использованию, также можно создавать собственные – пользовательские сценарии. Запуск выполнения сценария осуществляется по кнопке **Scan&Read**.

Суффиксы – небуквенные символы, которые могут встречаться в абсолютном конце слова.

Т

Таблица – блок, используемый для выделения таблиц или текста, имеющего табличную структуру. При распознавании программа разбивает данный блок на строки и столбцы и формирует табличную структуру. В выходном тексте данный блок передается таблицей.

Теги PDF – специальные теги, используемые для выделения логических частей и разметки таблиц и картинок в PDF документе. Теги, встраиваемые в PDF документ, обеспечивают удобство просмотра документа на экранах разного размера, например, экранах карманных компьютеров.

Текст – блок, используемый для выделения текстовых областей. Он должен содержать только одноколоночный текст.

Тип изображения – параметр сканирования; изображение может быть черно-белым, серым или цветным.

Тип (печати) входного текста – особенность в начертании символов входного текста, в зависимости от того, каким образом он напечатан (в типографии, на матричном принтере в черновом режиме, на пишущей машинке). Для типографского текста следует устанавливать режим **Авто**, для машинописного – режим **Пишущая машинка**, для текста, напечатанного на матричном принтере в черновом режиме – режим **матричный принтер**.

Тип блока – блок может быть Зоной распознавания, Текстовым, Картинкой, Таблицей, Штрих-кодом.

Ф

Фоновое распознавание – режим, позволяющий одновременно с распознаванием редактировать и сохранять распознанные страницы.

Ш

Шаблон блоков – в шаблоне описано положение и размеры блоков на странице.

Штрих-код – блок, используемый для участков изображения, содержащих штрих-код.

Э

Эталон – набор пар «усредненное точечное изображение символа – его название», который создается в процессе обучения системы на конкретном тексте.

Я

Яркость – параметр сканирования, определяет контрастность, т.е. различие между черными и белыми участками текста. При правильной настройке яркости качество распознавания возрастает.