

# FineReader 11: двойной тест

Владислав ТКАЧУК,  
Константин МАЙБОРОДА

На самом деле мы в редакции уже сделали собственный тест-обзор ABBYY Fine Reader 11. Но поскольку наши авторы решили высказать по этому поводу собственное мнение, мы решили опубликовать такой себе двойной тест. Надеемся, это позволит получить еще более объективное и всестороннее представление о продукте.

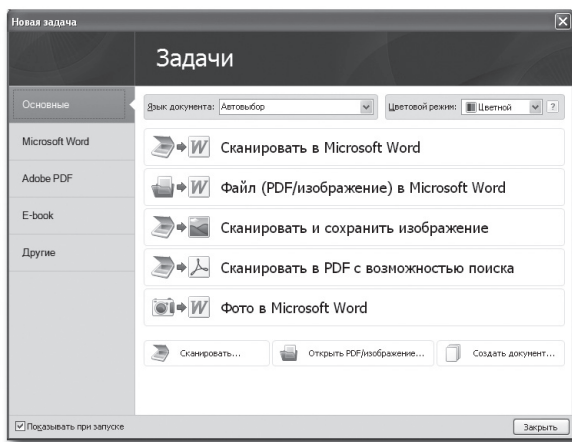
При подготовке новой версии программы упор делался на улучшение работы с типами документов, с которыми пользователи сталкиваются каждый день. В основном это книги, журналы, договоры. Анализ элементов на странице стал и быстрее, и качественнее, при этом FineReader также научился многому новому...

## СКОРОСТЬ РАСПОЗНАВАНИЯ — БЫСТРЕЕ НА 70 %

При использовании черно-белого режима распознавания текстов скорость обработки документов возрастает на 30 %, а при использовании режима быстрого распознавания и черно-белого режима скорость распознавания может увеличиться аж до 70 %. Вышеперечисленные режимы будут оптимальны при сканировании большого по объему и однотипного по содержанию документа, например, диплома, руководства, инструкции.

Отдельно стоит отметить, что при распознавании FineReader загружает каждое ядро компьютера отдельной страницей, что также приводит к увеличению скорости распознавания, в случае с 2 ядрами — в два раза, с 4 — в четыре и т. д.

В целом, благодаря совершенствованию технологии адаптивного распознавания документов (ADRT — Adaptive Document Recognition Technology), специалистам ABBYY удалось на 40 % повысить точность определения заголовков и колонтитулов, на 70 % — текста на полях, на 25 % улучшить распознавание таблиц и на 15 % — определение картинок и диаграмм. Также добавлена поддержка вертикальных колонтитулов, улучшено на 20 % качество «понимания» текстов для группы языков СJK (китайский, японский, корейский).



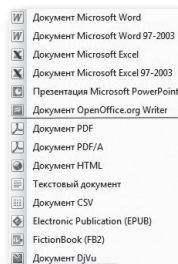
**В программе предусмотрено большое количество сценариев работы с отсканированными документами, включая создание электронных книг**

## СОЗДАНИЕ ЭЛЕКТРОННЫХ КНИГ

Разработчики программы правильно поняли тенденцию электронных книг в современном мире. В нем электронные книги прочно заняли свою нишу на рынке гаджетов и сдавать свои позиции не собираются. Именно поэтому упрощенное создание контента играет большую роль, и эта роль никак не обойдется без умения FineReader создавать электронные книги. Поддерживаемые форматы FictionBook 2.0 (FB2), Electronic Publication (ePub) и HTML позволяют создавать книги для разных типов устройств. Не забыты и устройства Kindle — созданные книги можно отправлять непосредственно на сайт kindle.com.

## ПОДДЕРЖКА ОТКРЫТЫХ ФОРМАТОВ

В 11-й версии программы была добавлена полнофункциональная поддержка открытого формата документов .odt из пакета программ OpenOffice/LibreOffice, что позволяет сохранять

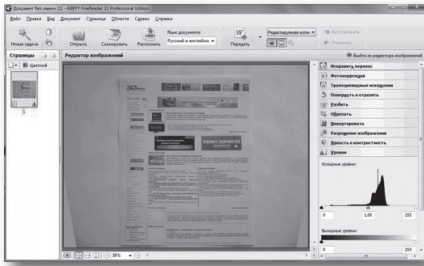


**Поддержка форматов OpenOffice является безусловным плюсом новой версии**

распознанные документы в формат .odt без потери качества и с сохранением исходного форматирования. В 11-й версии также появилась полная поддержка формата DjVu, в котором, как правило, сохраняют научную и учебную (отсканированную) документацию. Для данного формата также доступен весь перечень операций, как и для других типов документов.

## УЛУЧШЕННОЕ РАСПОЗНАВАНИЕ ФОТОГРАФИЙ И НОВЫЕ ИНСТРУМЕНТЫ ДЛЯ РЕДАКТИРОВАНИЯ ИЗОБРАЖЕНИЙ

В 11-й версии также появились новые возможности работы с полученным после сканирования документом в виде изображения. Редактировать можно настройку яркости, контрастности, интенсивность уровней света и тени. При правильном применении данных опций качество распознанных символов может увеличиваться в несколько раз. Например, если при сканировании некоторые элементы черного текста не слишком хорошо видны на белом фоне, то стоит поработать с ползунком настройки уровней света и тени, и тогда черный текст станет лучше виден.

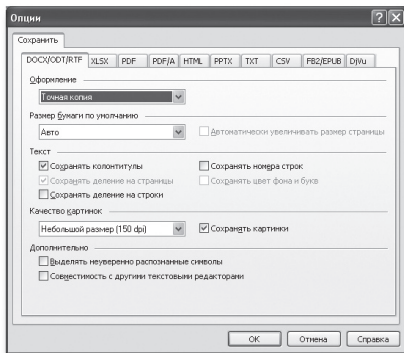


**Улучшения также коснулись работы с трапециевидными искажениями изображений с документами, полученными, например, после фотографирования документа на стенде**

а, следовательно, документ будет готов к последующей печати.

### ОПТИМАЛЬНЫЕ РЕЖИМЫ СОХРАНЕНИЯ PDF-ДОКУМЕНТОВ

В новой версии много внимания уделено работе с популярным форматом для хранения и передачи документов — PDF. Появились три новых режима сохранения документов: *Наилучшее качество*, *Небольшой размер*, *Сбалансированный режим*. Каждый из режимов поможет получить необходимое качество и размер выходного файла в зависимости от требований



**FineReader 11 поддерживает множество форматов сохранения документов и многочисленные настройки для каждого из них**

пользователя. Отдельно стоит сказать, что благодаря улучшенной технологии MRC-сжатия (Mixed Raster Content) в 11-й версии, документы при сохранении занимают в пять раз меньше места, чем при сохранении в предыдущей, 10-й версии продукта.

Также появилась возможность создавать оглавление для PDF-документа. При выборе данной опции FineReader пытается как можно точнее воссоздать структуру исходного документа, чтобы при сохранении его в формате PDF пользователю была доступна легкая навигация.

## МНЕНИЕ РЕДАКЦИИ

В жизни каждого человека рано или поздно наступает момент, когда бумажная копия документа еще есть, а ее электронного оригинала, увы, уже нет. Как правило, ситуация усугубляется еще и тем, что документ необходимо редактировать или еще как-то использовать в электронном виде. Вот тут-то и настает звездный час программ распознавания текста! Ну не набирать же его с распечатки вручную?!

Тем более что новая версия ABBYY FineReader 11 распознает текст намного быстрее предшественницы. Увеличение скорости особенно заметно в так называемом черно-белом режиме, когда программа тратит время на анализ цветных изображений. По данным производителя, прирост составляет порядка 45 %, и по нашим субъективным оценкам, эта цифра недалеко от истины.

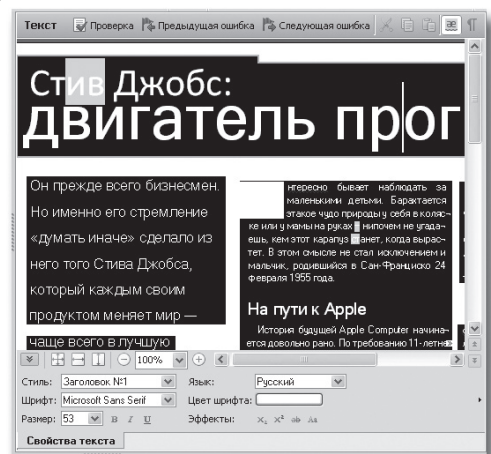
Как и в предыдущей версии OCR-системы большой упор сделан на точное воссоздание структуры документа, в частности правильное распознавание колонтитулов, заголовков, оглавлений и прочих элементов текста. Этой цели служит усовершенствованная фирменная технология ADRT, а также новый редактор стилей, позволяющий определить оформление однотипных элементов документа и упростить передачу его логической структуры. Добавьте к этому поддержку почти 190 языков (из них 40 имеют словарную поддержку) — и вы получите действительно универсальное средство для работы с отсканированными документами.

Впрочем, оперативное распознавание — далеко не единственная сфера применения подобных программ. В частности, новая версия ABBYY FineReader научилась превращать отсканированный и распознанный текст в электронные книги. Поддерживаются популярные форматы ePub и fb2. Также не проблема сохранить документ в PDF с текстовым слоем.

Единственное, чего не хватает для полноценного создания электронных книг, так это инструментов редактирования структуры. Нужно ли говорить, что интерактивное содержание и закладки существенно упрощают чтение, особенно с электронных букидеров. К другим недочетам можно отнести все еще неуверенное распознавание сложного форматирования. Например, программа упорно отказывается воспринимать двухколоночную верстку (характерна для периодических изданий) и неправильно отображает ее в итоговом документе. Также часто неправильно распознаются мелкие изображения, хотя как раз работа с картинками заявлена в числе важных сторон новой версии программы.

В целом можно сказать, что новый «файнридер» удался. Главные фишки программы — большое количество языков (включая арабский, распознавание которого непростая задача даже для человеческого глаза 😊), высокая скорость работы, качественная передача структуры и оформления документа. Все это наверняка как придется по душе новым пользователям, так и подвигнет обновиться до последней версии уже состоявшихся владельцев FineReader.

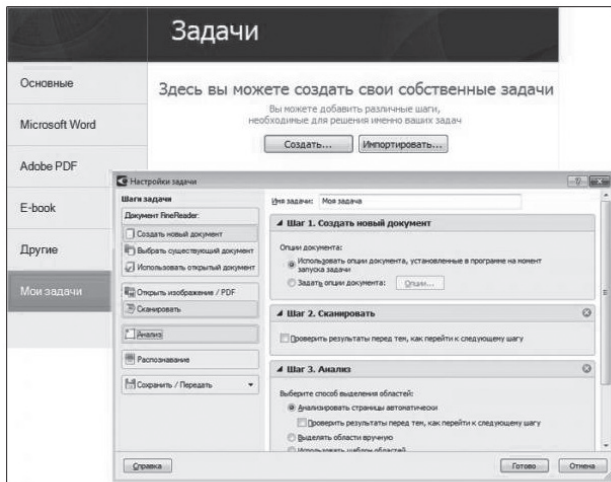
**В FineReader 11 появился новый редактор стилей, призванный улучшить передачу оформления и структуры документа**



## УЛУЧШЕННЫЙ ПОЛЬЗОВАТЕЛЬСКИЙ ИНТЕРФЕЙС

Благодаря усовершенствованному редактору стилей, теперь можно настраивать все параметры стилей в одном удобном диалоге, в то время как настройки будут сразу же применяться ко всему документу.

Теперь можно рассортировывать изображения страниц по нескольким документам для более точного сохранения оформления исходных документов, а также восстанавливать нумерацию страниц, в случае если исходный документ был напечатан с двух сторон страницы или имел тип печати «ясправа налево».



**Новой программе – новый интерфейс. Важно, что он не только обновился, но и стал удобнее, обогатился новыми пунктами и возможностью автоматизации типичных задач**

## НОВОЕ ОКНО ЗАДАЧ

В 11-й версии продукта было переработано окно задач, в которое были добавлены новые пункты, например, по созданию электронных книг, а также те пункты, которые чаще всего использует пользователь. В Corporate редакции FineReader также доступно окно *Мои задачи*, в котором можно автоматизировать процесс сканирования-распознавания-сохранения в нужный формат путем создания автоматического задания.

## ОЦИФРОВАНИЕ ВИЗИТНЫХ КАРТОЧЕК

В редакцию Corporate 11 версии была добавлена утилита для оцифровывания визитных карточек – ABVY Business Card Reader. Утилита поддерживает 25 языков распознавания, до 10 визиток на одной странице, а также экспорт оцифрованных визиток в Microsoft Outlook, vCard или отправку данных из визитки по электронной почте.

## ЗАЩИТА КОНФИДЕНЦИАЛЬНОЙ ИНФОРМАЦИИ

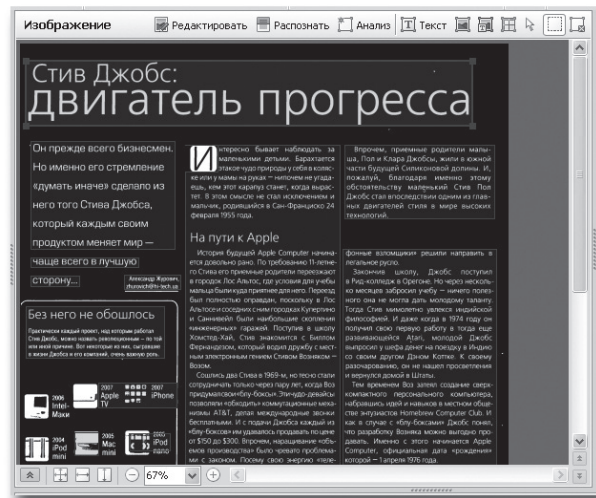
В редакцию Corporate 11-й версии также была добавлена возможность ог-

раничения доступа к некоторым частям документа, которые, по мнению автора, являются конфиденциальными. При редактировании документа такие строчки помечаются и уже при сохранении становятся недоступными (вырезаются из оригинального документа) для просмотра и копирования.

## ВЫВОДЫ

Благодаря улучшенным, обновленным и добавленным функциям, FineReader как средство повседневной работы стало намного удобнее. А благодаря добавлению функций работы с электронными книгами и свободными стандартами документов, теперь программу можно рекомендовать всем любителям читать в дороге, а также владельцам ОС Linux и бесплатных офисных пакетов. Добавление распознавания визитных карточек, несомненно, также привлечет пользователей, что позволит расширить позиции ABVY на рынке делового электронного документооборота.

В свете приближающихся праздников можно считать пакет ABVY FineReader отличным подарком тем ценителям простоты и удобства, которые постоянно работают с документами. Приобрести продукт можно по адресу: <http://abbyy.ua/products/ocr/finerader/>.



**Программа хорошо распознает оформление документов, но по-прежнему спотыкается на сложной журнальной или газетной верстке**

## ABVY FINEREADER 11: ПЛЮСЫ И МИНУСЫ

По мнению авторов:

- ⊕ ..... Более точное определение блоков на странице распознавания, работа с вертикальными колоннитулами
- ⊕ ... Поддержка форматов электронных книг
- ⊕ Поддержка открытых форматов документов
- ⊕ ..... Оцифровывание визиток
- ⊕ ... Защита конфиденциальной информации
- ⊖ ..... Стоимость лицензии (все-таки она пока высока для простого обывателя)

По мнению редакции:

- ⊕ ..... Скорость работы
- ⊕ ..... Точность воссоздания структуры и оформления документа
- ⊕ ..... Создание электронных книг
- ⊕ ..... Распознавание 189 языков
- ⊖ ..... Проблемы с воссозданием сложного форматирования

# ЧТО БЫ НИ СЛУЧИЛОСЬ С ВАМИ ЗА РУЛЕМ, ГОРЯЧАЯ ЛИНИЯ МАГНОЛИИ-ТВ ПРИДЕТ ВАМ НА ПОМОЩЬ!

«221» — ЭТО ТЕЛЕФОННАЯ ПОМОЩЬ АДВОКАТА, КОТОРЫЙ ПОДСКАЖЕТ ВАМ, КАК ОТСТОЯТЬ СВОИ ПРАВА В СЛУЧАЕ ДТП, КОНФЛИКТА С ГАИ ИЛИ СТРАХОВОЙ КОМПАНИЕЙ!

«221» — ЭТО ТЕЛЕФОННАЯ ПОМОЩЬ В СЛУЧАЕ ПОЛОМКИ В ПУТИ!  
МЫ ПОМОЖЕМ ВАМ ВЫЗВАТЬ ЭВАКУАТОР ИЛИ НАЙТИ БЛИЖАЙШУЮ СТО ПО ВСЕМ АВТОДОРОГАМ УКРАИНЫ!

«221» — ЭТО ТЕЛЕФОННАЯ ПОМОЩЬ В СЛУЧАЕ ПРОБЛЕМ СО ЗДОРОВЬЕМ В ДОРОГЕ!  
МЫ ПОМОЖЕМ ВАМ ВЫЗВАТЬ СКОРУЮ МЕДИЦИНСКУЮ ПОМОЩЬ ИЛИ НАЙТИ БЛИЖАЙШЕЕ МЕДИЦИНСКОЕ УЧРЕЖДЕНИЕ!

**ЗАПИШИТЕ «221» В СВОЙ МОБИЛЬНЫЙ!**



Для абонентів МТС, КІЇВСТАР, life:), Veeline. Вартість з'єднання стандартна, згідно з тарифним планом оператора. Вартість хвилини 8 грн. у т.ч. ПДВ, без ПФ. ТІЛЬКИ ДЛЯ ПОВНОЛІТНІХ. Телефон підтримки (044) 461-01-13. ЗАО «Украинская Мобильная Связь» - Ліцензія Мінтрансзв'язку серія АА № 720189, видана ЗАТ «УМЗ» 29.12.2004 р., ЗАО «Киевстар Дж.Эс.Эм.» - Ліцензія Державного комітету зв'язку та інформатизації України № 009503 від 12.04.2001 р., ООО «Астелит» - Ліцензія НКПРЗУ АБ № 222715 від 17.11.2005 р., ЗАО «Украинские Радиосистемы» - Ліцензія Держзв'язку серії АВ № 120093 від 16.05.2006 р.