Руководство пользователя

OMNIPAGE 18





Официальное уведомление

© Nuance Communications, Inc., 2011. Все права защищены. Никакая часть настоящей публикации не может передаваться, транскрибироваться, воспроизводиться, сохраняться в любой системе поиска или переводиться на любой язык или компьютерный язык в любом виде и любыми средствами, включая механические, электронные, магнитные, оптические, химические, ручные или иные, без письменного разрешения компании Nuance Communications, Inc., 1 Wayside Road, Burlington, Massachusetts 01803-4609.

Программное обеспечение, описанное в этом документе, предоставляется по лицензии, поэтому его можно использовать или копировать только в соответствии с условиями этой лицензии.

Важное замечание

Компания Nuance Communications, Inc. предоставляет эту публикацию «как есть» без каких-либо гарантий, выраженных явно или подразумеваемый, включая все без ограничения подразумеваемые гарантии гарантий пригодности для продажи или определенной цели. В некоторых областях не разрешен отказ от оговоренных явно или подразумеваемых гарантий при определенных сделках, поэтому данное положение, возможно, к вам не относится. Компания Nuance оставляет за собой право исправлять настоящую публикацию, а также время от времени вносить изменения в ее содержание без обязательства уведомления о подобных исправлениях.

Товарные знаки и их владельцы

Nuance, ScanSoft, OmniPage, PaperPort, True Page, Direct OCR, Logical Form Recognition, RealSpeak являются товарными знаками Nuance Communications, Inc., зарегистрированными в США и/или других странах. Все остальные упомянутые здесь названия компаний или продуктов могут являться товарными знаками соответствующих владельцев.

Сторонние лицензии и уведомления

См. официальные заявления/уведомления в конце данного руководства.

Nuance Communications, Inc.

1 Wayside Road Burlington, MA 01803-4609 CIIIA

Nuance Communications International BVBA International Headquarters

Guldensporenpark 32 Building D BE-9820 Merelbeke Бельгия

СОДЕРЖАНИЕ

Введение	5
Новые возможности программы OmniPage 18	7
Новые возможности программы OmniPage 17	8
Основные возможности решения OmniPage Professional	11
Установка и настройка	12
Требования к системе	12
Установка программы OmniPage	13
Настройка сканера для работы с программой OmniPage	14
Запуск программы	16
Регистрация программы	17
Активация программы OmniPage	18
Удаление программы	18
Использование программы ОмніРаде	19
Локументы OmniPage	19
Рабочий стол и вилы программы OmniPage	20
Основные этапы обработки	26
Использование программы OmniPage с компонентом PaperPort	28
Обработка документов	29
Способы обработки	29
Определение источника изображений страницРисунок	33
Описание макета документа	39
Предварительная обработка изображений	41
Области и фон	49
Проверка и релактирование	57
Отображение в редакторе и уровни форматирования	57
Проверка результатов распознавания	58
Верификация текста	60
Таблица символов	60
Пользовательские сповари	61
Языки	62
Эталон	65
Редактирование текста и изображения	67
Мгновенные изменения	69
Пометка и редактирование	70
Озвучивание текста	70
Создание и редактирование форм	72

Сохранение и экспорт	76
Сохранение и экспорт	76
Сохранение исходных изображений	77
Сохранение результатов распознавания	77
Отправка страниц по электронной почте	83
Отправка в Kindle	84
Другие варианты экспорта	85
Рабочие процессы	87
Помощник рабочего процесса	89
Диспетчер пакетной обработки	92
Создание новых заданий	92
Контролируемые папки	97
Контролируемые почтовые ящики	98
Обработка с помощью штрихкода	99
Помощник подшивки	101
Техническая информация	102
Поиск и устранение неполадок	102
Поддерживаемые типы файлов	106
Указатель	107

Введение

Благодарим вас за приобретение программы распознавания текста OmniPage[®] 18! Следующая документация позволяет приступить к работе с программой и дает общее представление о ней.

Руководство пользователя

Руководство познакомит вас с программой OmniPage 18. В нем содержатся указания по установке, описание команд и рабочих областей программы, инструкции по выполнению определенных задач, описание способов настройки обработки и управления ей, а также технические данные. Все описания приведены для операционной системы Windows 7[™].

В соответствии с экологической политикой компании Nuance руководство пользователя поставляется только в формате PDF. Для получения печатной копии на бумаге нормального размера рекомендуется использовать двухстороннюю печать с двумя страницами на одном листе.

Руководство рассчитано на пользователей, знакомых со средой Microsoft Windows. Сведения о диалоговых окнах, командах меню, полосах прокрутки, операциях перетаскивания, контекстных меню и т. п. см. в документации по операционной системе Windows.

Помимо этого, предполагается, что пользователь знаком со своим сканером и программным обеспечением для работы с ним. Перед установкой программы OmniPage 18 следует установить сканер и убедиться в том, что он правильно работает. При необходимости см. документацию к сканеру.

Мини-руководства с указаниями



Мини-руководства с указаниями можно найти с помощью меню «Справка». Это набор мини-руководств, которые позволяют быстрее начать работу. В них приводится краткий обзор основных областей программ, например: получение данных, оптимизация изображения, создание областей, распознавание, редактирование, проверка, новые функции и другие.

Электронная справка

В справке программы OmniPage представлены сведения о функциях, параметрах и процедурах. Кроме того, в нее включен полный словарь терминов с собственным алфавитным указателем и оглавлением. HTML-справка предназначена для простого и быстрого поиска сведений. Ее можно использовать после установки программы OmniPage.

Полная контекстная справка содержит достаточно сведений, чтобы сразу приступить к работе. Она открывается из диалоговых окон. Чтобы открыть контекстную справку, нажмите клавишу F1 или щелкните кнопку справки, если она есть в диалоговом окне.

Файл Readme

В файле Readme содержится самая свежая информация о программе. Ознакомьтесь с ней, прежде чем приступить к работе с программой OmniPage. Чтобы открыть этот HTML-файл, выберите файл Readme в программе установки OmniPage или воспользуйтесь меню справки после установки программы.

Сканирование и прочие сведения

На веб-сайте компании Nuance[®] по адресу www.nuance.com можно найти актуальные сведения о программе. Руководство по сканированию (http://www.nuance.com/scannerguide/) содержит последние данные о поддерживаемых моделях сканеров и способах решения связанных с ними проблем. Компания Nuance проводит тестирование 25 наиболее популярных моделей сканеров. Веб-сайт компании Nuance можно открыть из установщика OmniPage 18 или с помощью меню справки после установки программы.

Техническая информация

Веб-сайт www.nuance.com содержит технические примечания о часто возникающих проблемах при использовании программы OmniPage. Кроме того, на веб-сайт можно обратиться за помощью в установке программы и решении проблем.

Новые возможности программы OmniPage 18

Ниже перечислены различия по сравнению с 17 версией программы. Чтобы получить дополнительные сведения, перейдите по представленным ниже ссылкам.

- **Первая страница.** При запуске OmniPage на этой странице отображаются команды, позволяющие открыть или отсканировать документы, а также получить доступ к проектам OmniPage. Кроме того, здесь можно найти предварительно настроенные рабочие процессы для быстрого преобразования документов из одного формата в другой.
- Помощник eDiscovery для PDF с возможностью поиска. Этот помощник разработан специально для создания PDF-файлов с возможностью поиска из PDF-файлов, содержащих только изображения либо некоторые текстовые элементы или страницы, без изменения существующего текста или его распознавания. Все текстовые элементы (включая метаданные документа, примечания, разметку для редактирования, штампы и т. д.) в PDF-файлах остаются без изменений. Процесс может выполняться автоматически или взаимодействовать с пользователем при разметке областей или проверке орфографии. См. раздел «Помощник eDiscovery для PDF с возможностью поиска» на стр. 82.
 - Подключение к облачным хранилищам. Теперь можно загружать входные файлы из веб-хранилищ и отправлять в них результаты распознавания. В OmniPage реализована интеграция со службами Evernote и Dropbox. Кроме того, компонент Nuance Cloud Connector обеспечивает доступ к различным облачным службам, включая Microsoft Live SkyDrive, Документы Google, Box.net и т. д. Преимущество компонента Nuance Cloud Connector заключается в том, что он тесно интегрируется в систему Microsoft Windows, позволяя перетаскивать файлы непосредственно из облачных служб. Компонент Nuance Cloud Connector также можно обновить до более функционального решения под названием Gladinet Cloud Desktop Pro. В этой улучшенной версии доступны дополнительные функции, позволяющие выполнять автоматическое резервное копирование и синхронизацию файлов с помощью облачных служб. См. разделы «Получение из облачного хранилища» на стр. 34 и «Другие варианты экспорта» на стр. 85.
- Новые инструменты оптимизации изображений. Усовершенствованы алгоритмы удаления дефектов и точек с изображений, что повышает точность распознавания слов. Кроме того, теперь можно выбрать способ удаления дефектов

(«Обычный», «Полутон», «Соль и перец»). Реализованы функции автоматической обрезки страниц для обнаружения и уменьшения полей. Кроме того, поддерживается удаление следов дырокола и очистка краев от теней сканирования и ненужных заметок. Теперь можно улучшить качество текста и диаграмм на белой доске, сфотографированных цифровой камерой. Это повышает их удобочитаемость. См. раздел «Инструменты оптимизации изображения» на стр. 43.

- Улучшенное определение пустых страниц. Благодаря новым параметрам чувствительности повышена точность определения пустых страниц, которые могут быть отсканированы как светло-серые или цветные. Для этого используется регулировка порогового значения черного цвета. Таким образом повышается эффективность двух элементов управления OmniPage: нового параметра предварительной обработки «Пропускать пустые страницы» и существующего параметра «Создавать новый файл для каждой пустой страницы».
- Автоматическое определение языка. Программа может самостоятельно назначать каждой входной странице один язык распознавания во время автоматической обработки. См. раздел «Распознавание текста на азиатских языках» на стр. 63.
- Выбор предлагаемых вариантов во время проверки с помощью сочетаний клавиш. Предлагаемые варианты в диалоговом окне «Программа проверки распознавания» пронумерованы. Вариант можно выделить указателем мыши и принять, нажав кнопку «Изменить», а также можно ввести номер варианта, удерживая нажатой клавишу CTRL. См. раздел «Проверка и редактирование» на стр. 57.
- Сканеры ISIS Сканеры с поддержкой драйверов ISIS теперь можно использовать непосредственно в OmniPage

Новые возможности программы OmniPage 17

Ниже перечислены различия по сравнению с 16 версией программы. Чтобы получить дополнительные сведения, перейдите по представленным ниже ссылкам.

- Распознавание текста на азиатских языках. Программа позволяет распознавать вертикальный и горизонтальный текст на японском, корейском и китайском (традиционное и упрощенное письмо) языках даже с небольшими вставками английского текста. Результаты можно просмотреть и проверить в текстовом редакторе. См. раздел «Распознавание текста на азиатских языках» на стр. 63.
- Вертикальный неазиатский текст. Программа поддерживает автоматическое определение вертикального текста в ячейках таблицы и на страницах PDF- или XPS-файлов, а также на изображениях других типов (при определенных условиях). Новые инструменты позволяют вручную выделять области вертикального текста. Вертикально отображаемый текст можно отредактировать в текстовом редакторе с помощью уровня форматирования True Page[®] («Фактическая страница»). На других уровнях текст отображается горизонтально. См. раздел «Автоматическая разметка областей» на стр. 49. и См. раздел «Типы и свойства областей» на стр. 51.
- Простой загрузчик. Файловая система теперь отображается в стиле проводника Windows в одном из окон OmniPage. Эта функция позволяет всегда иметь под рукой список файлов во время работы и пользоваться всеми возможностями проводника, что упрощает быстрое выделение файлов. Диалоговое окно с возможностью блокировки позволяет составить набор файлов перед началом загрузки. В виде быстрого преобразования простой загрузчик дает возможность не только быстро загрузить файлы, но и выполнить обработку нажатием всего одной кнопки: загрузить > распознать > сохранить. См. раздел «Получение через простой загрузчик» на стр. 34.
- Расширенная поддержка систем ECM. Доступны ссылки на системы Hummingbird (OpenText) и iManage (Interwoven). При работе с системой SharePoint сведения о сервере, имени пользователя и пароле требуется ввести только один раз в течение сеанса. После этого они сохраняются и подставляются автоматически.
- **Поддержка приложений Office 2007 и 2010.** Кнопки Direct OCR отображаются на отдельной вкладке Nuance OCR, а не вместе с другими надстройками.
- Более удобная пакетная обработка. Диспетчер пакетной обработки автоматически, не останавливая основной рабочий процесс, пропускает файлы (в том числе защищенные паролем), которые нельзя обработать. Эти файлы затем указываются в окне с результатами работы.

- **Быстродействие.** Увеличена скорость запуска программы и ее быстродействие на компьютерах, оснащенных многоядерными процессорами. Реализована поддержка четырехъядерных процессоров.
- **Привязка рабочих процессов к кнопкам сканера.** Функции и рабочие процессы OmniPage можно связать с кнопками сканера. Благодаря этому предварительную обработку, распознавание и сохранение документов можно запустить прямо со сканера. См. раздел «Сканирование с помощью OmniPage и рабочих процессов» на стр. 38.
- Отправка результатов в электронную книгу Kindle. Помощник Kindle позволяет создавать рабочие процессы для отправки результатов распознавания в учетную запись Kindle на сервере Amazon с их последующим отображением на устройстве Kindle, зарегистрированном для этой учетной записи. См. раздел «Отправка в Kindle» на стр. 84.
- Другие улучшения. Усовершенствования в области предварительной обработки изображения обеспечивают более качественную разметку и повышают общую точность распознавания, особенно для XPS-файлов и при преобразовании документов из одного формата в другой. Поддерживается загрузка фотографий высокой четкости (JPEG XR). Улучшена интеграция с Microsoft Word, Excel и PowerPoint. Можно создавать линеаризованные PDF-файлы, оптимизированные для ускоренного просмотра в Интернете.

Основные возможности решения OmniPage Professional



Этот значок используется в руководстве для обозначения функций, которые доступны только в OmniPage Professional 18.

- **Извлечение** данных из заполненных форм: Этап рабочего процесса позволяет извлекать данные из наборов форм и затем экспортировать их в базы банных с помощью шаблона PDF-форм. Это могут быть активные формы PDF, статические формы в различных форматах изображений или сканированные бумажные формы.
- **Пометка и редактирование**: текст можно выделять, зачеркивать и редактировать (делать нечитаемым) в текстовом редакторе. Функция редактирования удобна для юридических документов, а также для документов с конфиденциальным содержанием.
- Помощник подшивки: более эффективная помощь при создании и использовании рабочих процессов титульных страниц со штрихкодом.
 Эти страницы позволяют выполнять автоматическую обработку и сохранение документов. Для запуска достаточно нажать только одну кнопку сканера.

Более подробный список функций и различий между версиями OmniPage см. в справке.

Программа OmniPage 18 поставляется в корпоративных версиях для работы в сети. Кроме того, она поставляется в специальных редакциях для производителей сканирующих устройств и других торговых посредников. Наборы функций в этих изданиях могут отличаться в соответствии с требованиями заказчика.

Установка и настройка

Эта глава посвящена установке и запуску программы OmniPage.

Требования к системе

Минимальные требования для установки и запуска программы OmniPage 18:

- Компьютер с процессором Intel[®] Pentium[®] с тактовой частотой не менее 1 ГГц (или аналогичным). Рекомендуется двух- или четырехъядерный процессор.
- 32-разрядная версия Microsoft Windows[®] ХР[™] с пакетом обновления 3 (SP3) процессор с частотой 400 МГц; 32-разрядная версия Windows[®] Vista[™]с пакетом обновления 2 (SP2), 64-разрядная версия Windows[®] Vista[™]с пакетом обновления 2 (SP2), 32-разрядная версия Microsoft Windows[®] 7[™] процессор с частотой 1 ГГц.
- 512 МБ оперативной памяти. Для более высокой производительности рекомендуется 1 ГБ;
- 250 МБ свободного места на жестком диске для приложения и образцов изображений, а также 100 МБ рабочего пространства во время установки. Дополнительно:
 - 230 МБ для всех модулей RealSpeak[®] (90 МБ для модуля RealSpeak[®] Solo для американского варианта английского языка, а также по 10—15 МБ для модулей RealSpeak Solo для остальных языков).
 - 30 МБ для Nuance Cloud Connector.
 - 150 МБ для компонента Nuance PDF Create (поставляется только в комплекте с OmniPage Professional).
 - 500 КБ для компонента PaperPort[®] (поставляется только в комплекте с OmniPage Professional OmniPage Professional).
- Цветной монитор с разрешением 1024х768 точек и видеокартой, поддерживающей цветовую палитру 16 бит или более.
- Дисковод компакт-дисков для установки программы с диска или доступ к Интернету для ее загрузки.
- Звуковая карта и динамики для чтения текста вслух.

- Указывающее устройство, совместимое с Windows.
- Цифровая камера с автофокусировкой и разрешением не менее 2 мегапикселей для съемки текста. Подробные сведения см. в справке.
- Совместимый с программой сканер с собственным драйвером и программным обеспечением для сканирования документов (драйверы сканера WIA, TWAIN или ISIS). Список поддерживаемых сканеров см. в руководстве к сканерам на вебсайте Nuance (www.nuance.com).
- Подключение к Интернету для активации, регистрации и обновления продукта, работы компонента Nuance Cloud Connector, а также для обновления баз данных мастера настройки сканера.
- Для просмотра документов на японском, китайском или корейском языке необходимо включить поддержку восточноазиатских языков в операционной системе («Панель управления» > «Язык и региональные стандарты»).

Установка программы OmniPage

В каждом окне программы установки OmniPage 18 приводятся подробные указания по установке.

Действия перед установкой программы OmniPage

- Закройте все приложения, в особенности антивирусные программы.
- Для установки необходимо иметь права администратора.
- При наличии предыдущей версии OmniPage, а также при обновлении демонстрационной версии или специального издания OmniPage необходимо сначала удалить этот продукт.

Установка OmniPage

- 1. Загрузите файл программы и выберите команду «Выполнить» по окончании загрузки или вставьте компакт-диск OmniPage в дисковод. Программа установки запустится автоматически. Если это не произойдет, найдите дисковод в проводнике Windows и дважды щелкните программу Autorun.exe в корневом каталоге компакт-диска.
- 2. Выберите язык, который будет использоваться в процессе установки. Примите условия лицензионного соглашения и введите серийный номер, полученный по электронной почте или указанный на упаковке компакт-диска.

- 3. Выберите полную или выборочную установку. При полной установке устанавливаются речевые модули RealSpeak[®] для всех языков (сейчас их девять). При выборочной установке модули можно исключать или добавлять. Чтобы исключить модуль, щелкните стрелку вниз рядом с этим модулем и выберите параметр «Этот компонент будет недоступен».
- 4. Для установки программного обеспечения следуйте инструкциям на каждом экране. Все необходимые для сканирования файлы автоматически копируются в процессе установки.

После установки OmniPage Professional запустится установка компонента Nuance PDF Create 7, если это не было отменено во время установки OmniPage. Функция преобразования документов из одного формата в другой доступна, только если на компьютере установлен компонент PDF Create.

Продукт OmniPage Professional поставляется с бесплатной копией Nuance PaperPort[®] — приложения управления документами. Его необходимо устанавливать отдельно, и он предъявляет отдельные требования к системе.

Настройка сканера для работы с программой OmniPage

Все необходимые для установки и поддержки работы сканера файлы копируются автоматически во время установки программы. Однако настройка сканера при этом не производится. Перед работой с программой OmniPage необходимо установить сканер и его драйверы, а также убедиться в том, что он правильно функционирует. Драйверы сканера не входят в комплект поставки OmniPage.



Настройка сканера выполняется с помощью мастера настройки сканера. Запустите его так, как описано ниже. Если этого не сделать, он запускается при первой попытке сканирования. Выполните следующие действия.

Последовательно выберите элементы «Пуск» > «Программы» > Nuance > OmniPage 18 > «Мастер настройки сканера»;
 или нажмите кнопку «Настройка» на панели «Сканер» в диалоговом окне «Параметры»;
 или выберите пункт «Сканирование» в раскрывающемся списке «Получить страницу» на панели инструментов, а затем нажмите кнопку «Получить страницу».

- Откроется окно мастера настройки сканера. При наличии соединения с Интернетом в первую очередь будет предложено обновить базы данных по сканерам. Выберите «Да» или «Нет» и нажмите кнопку «Далее».
- Выберите «Выбрать и протестировать устройство» и нажмите кнопку «Далее». Установленный сканер выводится в списке вместе со всеми ранее настроенными для OmniPage сканерами. Если необходимый сканер отсутствует в списке, нажмите кнопку «Добавить сканер...».
- Откроется список всех обнаруженных драйверов сканеров в отмеченных категориях. В список могут быть включены и сетевые устройства. Выберите нужное устройство и нажмите кнопку «ОК». Для установки второго устройства запустите мастер сканера еще раз.
- Мастер сообщает, имеются ли в базе данных параметры для выбранной модели сканера. Если параметры модели включены в базу данных, то тестирование не потребуется. В противном случае следует протестировать устройство. Нажмите кнопку «Далее».
- Если тестирование не требуется, нажмите кнопку «Готово».
 Чтобы протестировать подключение сканера, нажмите кнопку «Далее».
 При правильном подключении откроется меню дальнейшего тестирования.
 Выберите тесты на свое усмотрение. Рекомендуется выполнить основное контрольное сканирование.
- По умолчанию программа OmniPage использует собственный интерфейс: диалоговое окно «Параметры» с панелью «Сканер». Для настройки параметров собственными средствами сканера, нажмите кнопку «Дополнительные параметры...» и выберите нужный интерфейс. Кнопки «Редактор настроек» и «Изменить настройки» следует использовать только опытным пользователям сканеров или по рекомендации службы технической поддержки.
- Нажмите кнопку «Далее» для проведения тестирования. Для выполнения основного контрольного сканирования вставьте в сканер тестовую страницу. Мастер выполнит сканирование с использованием программного обеспечения производителя сканера. Нажмите кнопку «Далее». Откроется программа, поставляемая производителем сканера.
- Нажмите кнопку «Сканировать», чтобы начать пробное сканирование.
- При необходимости нажмите кнопку «Потери изображения...» или «Неправильная ориентация...» и выберите необходимые параметры.
- Как только изображение в окне будет приемлемым, нажмите кнопку «Далее».

- Для проведения остальных тестов следуйте инструкциям на экране.
- Если все предложенные тесты выполнены успешно, мастер сканера сообщит об этом и предложит нажать кнопку «Готово».
- Настройка сканера для работы с программой OmniPage 18 успешно завершена. Чтобы впоследствии удалить, настроить сканер или изменить его параметры, откройте окно мастера настройки сканера с помощью меню «Пуск» системы Windows или панели «Сканер» в диалоговом окне «Параметры».

Для тестирования и ремонта неисправного сканера откройте окно мастера и выберите «Выбрать и протестировать устройство» на второй панели. Выполните описанную выше процедуру. При необходимости обратитесь за консультацией в службу технической поддержки.

Чтобы указать в качестве используемого по умолчанию другой сканер, откройте в окне мастера список всех настроенных сканеров. Выделите нужный сканер и закройте окно мастера, нажав кнопку «Готово».

Для получения обновленных параметров используемого сканера откройте окно мастера, загрузите новую базу данных в первом окне и выберите «Использовать настройки устройства». Нажмите кнопку «Далее», затем кнопку «Готово».

Запуск программы

Запустить программу OmniPage 18 можно несколькими способами.

 На панели задач Windows нажмите кнопку «Пуск» и последовательно выберите элементы «Все программы» > Nuance > OmniPage 18 > OmniPage [Professional] 18.



• Дважды щелкните значок OmniPage в папке установки программы или на рабочем столе Windows (если он добавлен туда).

• Дважды щелкните значок «Документ OmniPage (OPD)» или имя файла. Выбранный документ будет загружен в программу. См. раздел «Документы OmniPage» в следующей главе.

• Щелкните правой кнопкой мыши один или несколько выделенных файлов изображений. Откроется контекстное меню. Выберите пункт «Открыть с помощью...», а потом – приложение OmniPage. Выделенные файлы будут загружены в программу.

Во время запуска на экране появляется заставка программы OmniPage, а затем открывается панель выбора вида. OmniPage содержит виды трех основных типов.

Подробное описание см. в разделе «Рабочий стол и виды программы OmniPage» в следующей главе. В этой главе приведено описание основных рабочих областей программы.

Существует несколько способов запуска программы с ограниченным интерфейсом.

- Воспользуйтесь диспетчером пакетной обработки. Нажмите кнопку «Пуск» на панели задач Windows и последовательно выберите элементы «Все программы» > Nuance > OmniPage 18 > «Диспетчер пакетной обработки OmniPage». См. главу «Рабочие процессы».
- В меню приложения «Файл» выберите пункт «Получить текст» (приложение должно быть настроено для работы с Direct OCRTM). См. раздел «Настройка Direct OCR» в главе «Обработка документов».
- Щелкните правой кнопкой мыши один или несколько выделенных файлов изображений в проводнике Windows. Откроется контекстное меню. Выберите пункт OmniPage 18, а затем в его подменю — конечный формат, элемент «Мастер преобразования» или рабочий процесс. Файлы будут обрабатываться в соответствии с инструкциями для рабочего процесса. См. главу «Рабочие процессы».
- Щелкните значок агента OmniPage 🧖 на панели задач. Выберите рабочий процесс запуска программы и запустите его.
- Для добавления служб распознавания используйте OmniPage 18 с программой управления документами Nuance PaperPort. См. раздел «Использование программы OmniPage с компонентом PaperPort» в главе «Использование программы OmniPage».

Регистрация программы

В конце установки запускается процедура интерактивной регистрации Nuance. Обеспечьте доступ к сети Интернет и заполните простую электронную форму. На это потребуется не более пяти минут. После заполнения формы нажмите кнопку «Отправить». Если процедуру регистрации не пройти при установке программы, регистрация будет периодически предлагаться в дальнейшем. Регистрацию можно выполнить на веб-сайте www.nuance.com. Нажмите кнопку «Поддержка» и выберите в левой части экрана пункт «Зарегистрироваться». Чтобы прочитать заявление об использовании регистрационных данных, просмотрите политику конфиденциальности Nuance.

Активация программы OmniPage

В конце установки будет предложено активировать продукт. (для этого необходимо соеинение с Интернетом). Если серийный номер введен без ошибок и сохранен там, где следует, процесс проходит без участия пользователя и передачи персональных данных. Если продукт не был активирован при установке, предложение об активации поступает при каждом запуске программы. Без активации программу OmniPage 18 можно запустить ограниченное количество раз. Рекомендуется выполнять автоматическую активацию.

Удаление программы

Иногда возникшую проблему можно решить путем удаления и повторной установки программы OmniPage. Программа удаления OmniPage не удаляет файлы с результатами распознавания и следующие пользовательские файлы:

шаблоны областей (ZON) шаблоны оптимизации изображений (IPP) файлы эталонов (OTN) пользовательские словари (UD) документы OmniPage (OPD) файлы заданий (OPJ) файлы рабочих процессов (XWF)

Для удаления программы необходимо иметь права администратора.

Удаление или переустановка программы OmniPage

- Закройте программу OmniPage.
- Нажмите кнопку «Пуск» на панели задач Windows и последовательно выберите пункты «Панель управления» и «Удаление программ» (в более ранних версиях Windows: «Установка и удаление программ»).
- Выберите пункт OmniPage SE 16 и нажмите кнопку «Удалить».
- В открывшемся диалоговом окне подтвердите удаление, нажав кнопку «Да».
- Выберите «Да», если требуется перезагрузить компьютер немедленно, или «Нет», чтобы сделать это позднее.
- Следуйте инструкциям до завершения процесса.

При удалении программы OmniPage также удаляются настройки для работы со сканером. После повторной установки программы OmniPage необходимо заново выполнить установку сканера. Все модули RealSpeak[®], которые были установлены вместе с программой, также будут удалены. Такие компоненты продукта OmniPage 18 Professional, как Nuance PDF Create 7 и PaperPort, необходимо удалять отдельно.

Использование программы OmniPage

С помощью технологии оптического распознавания символов программа OmniPage 18 преобразует отсканированные документы или файлы изображений в редактируемый текст, с которым можно работать в любых приложениях.

Кроме того, в процессе распознавания текста программа OmniPage сохраняет следующие элементы и атрибуты документа:

изображения (фотографии, логотипы);

элементы формы (флажки, переключатели, текстовые поля);

форматирование текста (символы и абзацы);

форматирование страницы (структуры столбцов, форматы таблиц, заголовки, размещение изображений).

Документы в программе OmniPage

Документ в программе OmniPage состоит из набора изображений, каждый из которых соответствует отдельной странице. После распознавания документ также содержит распознанный текст. Он отображается в окне текстового редактора вместе с изображениями, таблицами и элементами форм.

Документы OmniPage



Документ OmniPage (OPD) содержит изображения страниц оригинала (которые могут быть при необходимости обработаны) с выделенными областями. После распознавания документ содержит также результаты распознавания.

Документ OmniPage может содержать встроенный словарь пользователя, файл эталона, файл шаблона областей или файл шаблона оптимизации изображения. Это может значительно увеличить размер файла, однако делает его более удобным для переноса. Чтобы встроить файл, откройте соответствующее диалоговое окно с помощью меню «Сервис», выберите необходимый файл и нажмите кнопку «Встроить». Чтобы получить локальную копию встроенного файла, находящегося внутри полученного файла OPD, нажмите кнопку «Извлечь».

При открытии документа OmniPage вместо текущих параметров программы применяются параметры открываемого документа.

Рабочий стол и виды программы OmniPage

OmniPage предлагает три различных вида, которые можно использовать для выполнения задач.

- Классический вид этот вид аналогичен тому, который использовался в предыдущих версиях программы OmniPage.
- **Гибкий вид** в этом виде функциональные панели OmniPage расположены в виде отдельных вкладок, что позволяет увеличить размер каждой панели.
- **Вид быстрого преобразования** этот вид предназначен для быстрого и простого преобразования документов без детального изучения функций программы. Самые важные параметры преобразования четко видны на одном экране.

Меню «Окно» позволяет переключаться между видами и сохранять их (см. ниже). При запуске нового сеанса восстанавливаются вид и расположение окон, которые использовались при последнем закрытии программы.

Для всех трех видов можно восстановить значения по умолчанию с помощью команды «Сбросить текущий вид» в меню «Окно».

Панели программы

В программе OmniPage есть набор панелей, которые можно закреплять (в виде вкладок или мозаики), делать плавающими, сворачивать или восстанавливать по отдельности. Кроме того, можно изменять их размер. Вот список этих панелей: «Эскизы», «Изображение страницы», «Текстовый редактор», «Диспетчер документов», «Простой загрузчик», «Состояние рабочего процесса» и «Справка». Чтобы сделать панель плавающей, дважды щелкните ее заголовок или вкладку. Чтобы восстановить последнее закрепленное состояние плавающей панели, дважды щелкните ее заголовок или вкладку. Чтобы закрепить панель в новом месте, перетащите ее туда. Позиция закрепления будет обозначена цветным прямоугольником. Отпустите кнопку мыши, чтобы закрепить там панель. Чтобы увидеть все возможные позиции закрепления друг за другом (в виде вкладок или мозаики), перетащите панель в главное окно OmniPage, удерживая нажатой левую кнопку мыши. Чтобы переместить плавающую панель без закрепления, удерживайте нажатой клавишу СTRL во время перетаскивания.

Классический вид

В классическом виде на рабочем столе программы OmniPage по умолчанию мозаикой отображаются четыре основные рабочие области, отделенные друг от друга разделителями: «Диспетчер документов», «Изображение страницы», «Эскизы» и «Текстовый редактор». Область «Изображение страницы» содержит панель инструментов «Изображение», а текстовый редактор — панель инструментов форматирования.



Панель инструментов OmniPage: содержит набор инструментов для выполнения необходимой обработки.

Панель эскизов: на этой панели отображаются эскизы страниц.

Диспетчер документов: представляет сведения о документе в виде таблицы. Каждой странице соответствует одна строка. В столбцах отображаются статистические данные

и информация о состоянии каждой страницы, а также (если возможно) документа в целом.

Изображение страницы: изображение текущей страницы вместе с ее областями. Когда отображается страница, становится доступной панель инструментов «Изображение».

Текстовый редактор: здесь отображаются результаты распознавания текущей страницы.

Панели можно свободно перемещать по горизонтали и вертикали. Используйте меню «Окно», чтобы открыть панели «Простой загрузчик», «Состояния рабочего процесса» или «Справка». Панели можно свернуть или закрыть, но не закрепить как вкладки. Чтобы восстановить классический вид, выберите в меню «Окно» команду «Сбросить текущий вид».

Гибкий вид

Этот вид используется для настройки рабочего пространства OmniPage с целью оптимизации работы. По умолчанию отображаются все панели. Есть пять вкладок: «Изображение страницы» (включая «Эскизы»), «Текстовый редактор», «Текстовый редактор», «Простой загрузчик», «Состояние рабочего процесса» и «Справка». Диспетчер документов отображается на горизонтальной панели у основания рабочей области. Панели можно откреплять, перемещать, сворачивать, группировать или закрывать, как уже было описано выше. Чтобы преобразовать вкладку в обычную расположенную мозаикой панель, просто перетащите ее в рабочую область. Чтобы снова сделать панель вкладкой, перетащите ее в строку вкладок или используйте пробел, как было описано выше. Если панели сгруппированы в одной вкладке, в ее подписи отображается название активной панели Чтобы восстановить исходное состояние гибкого вида, в меню «Окно» выберите команду «Сбросить текущий вид».

Простой загрузчик, как и проводник Windows, отображает список файлов и предлагает те же функции. При этом он может оставаться открытым во время работы, позволяя быстро выбирать файлы и собирать их (см. Глава 4, стр. 34).

Предлагаемые сценарии:

Развертывание рабочего пространства (один экран)

Загрузите документ. Откройте панели, которые необходимо использовать. Перетащите их по одной за заголовки так, чтобы они прикрепились рядом с активной в виде вкладок. Справку также можно прикрепить, чтобы не переключаться между двумя окнами.

Работа с результатами распознавания (один экран)

Загрузите документ и выполните его распознавание. Закройте все панели, кроме диспетчера документов и текстового редактора. Разверните обе панели по горизонтали, уменьшите панель диспетчера документов и прикрепите сверху или снизу. Теперь можно просматривать страницы, дважды щелкая по очереди каждую из них на панели «Диспетчер документов», и результаты распознавания на панели «Текстовый редактор».

Количество сомнительных слов и нераспознанных символов в диспетчере документов позволяет определить страницы, на которых возникли сложности с распознаванием.

Обработка больших документов (два экрана)

Загрузите необходимый документ. Переместите его вид эскизов на второй монитор, а затем разверните для просмотра документа с увеличенным масштабом и получения дополнительного пространства для операций с эскизами.

Проверка (два экрана)

Поместите панель «Изображение страницы» на один экран, а «Текстовый редактор» на другой. Благодаря этому обеспечивается дополнительное пространство для редактирования и проверки.

Панель «Изображение страницы» всегда доступна для проверки распознавания, а также выполнения мгновенного выделения областей и редактирования.









Приведенные ранее сценарии являются только примерами, дающими представление о возможных операциях в режиме гибкого вида.

Вид быстрого преобразования

Вид быстрого преобразования используется для быстрого распознавания и сохранения. В вид быстрого преобразования можно переключаться только в том случае, если не открыты документы. В этом виде можно обрабатывать только один входной и один выходной документы. На изображении представлен внешний вид программы по умолчанию.



Простой загрузчик по умолчанию представляет собой вкладку, на которую можно перейти из панели параметров быстрого преобразования. Можно также добавить панель «Справка», но другие панели не будут видны в этом виде. Вкладки можно превратить в отдельные панели и свернуть их, как в других видах.

После загрузки файла его необходимо преобразовать перед загрузкой следующего файла. После преобразования изображения его не нужно закрывать. Просто загрузите новый файл.

Простой загрузчик в быстром виде имеет дополнительную возможность — обработку одним нажатием кнопки. В меню «Процесс» выберите пункт «Простой загрузчик», а затем — команду «Загрузить файлы» или «Получить и преобразовать». С помощью последней команды можно выбрать несколько файлов, которые будут загружены, распознаны и сохранены с использованием текущих параметров. Для этого установите одинаковые имена для выходных и входных файлов. Подробные сведения см. в Глава 4, стр. 34 и в справке.

На панели «Изображение» в виде быстрого преобразования есть панель инструментов «Быстрое преобразование», содержащая наиболее удобные команды обработки изображений. Для вызова дополнительных команд, таких как сохранение файлов изображений, оптимизация изображений, мгновенная разметка областей и ручная разметка областей для вертикального текста, необходимо использовать другой вид.

Пользовательские виды

Чтобы создать пользовательский вид, расположите панели и панели инструментов так, как необходимо, а затем выберите в меню «Окно» > «Пользовательские виды» > «Управление». Нажмите кнопку «Добавить» и укажите имя вида. Компоновки экрана отобразятся в подменю «Пользовательские виды», а активный вид будет отмечен галочкой. В пользовательских видах невозможно восстановить используемые по умолчанию параметры.

изменение видов

Для переключения между видами используйте меню «Окно». Панели отображаются или скрываются и размещены так же, как при последнем использовании выбранного вида. Раздел справки остается на экране без изменений независимо от вида. Вне зависимости от вида в простом загрузчике сохраняется расположение файлов, а на панели «Состояние рабочего процесса» продолжают отображаться сведения о последнем запущенном рабочем процессе. После перезапуска программы на панели «Справка» отображается раздел «Введение», а путь к папке по умолчанию в простом загрузчике и панель «Состояние рабочего процесса» процесса» пусты.

Панели инструментов

В программе имеется одиннадцать основных панелей инструментов. Включить, отключить или настроить их можно с помощью меню «Вид». В строке состояния в нижней части окна программы OmniPage приводится описание назначения всех инструментов.

Стандартная панель инструментов: предназначена для выполнения основных функций.

Панель инструментов «Изображение»: позволяет выполнять операции с изображениями, областями и таблицами. Тремя группами инструментов этой панели (мини-панелями) можно управлять отдельно.

- Панель инструментов «Области»: содержит инструменты для работы с областями.
- Панель инструментов «Поворот»: содержит инструменты поворота.
- Панель инструментов «Таблица»: вставка, перемещение и удаление разделителей строк и столбцов.

Панель инструментов «Форматирование»: позволяет форматировать распознанный текст в текстовом редакторе.

Панель инструментов «Верификатор»: позволяет определять местоположение и вид верификатора.

Панель инструментов «Переопределение порядка»: служит для изменения порядка элементов на распознанных страницах.

Панель инструментов «Пометка текста»: содержит инструменты для пометки и редактирования текста.

Панель инструментов «Создание формы»: предназначена для создания новых элементов форм.

Панель инструментов «Расположение формы»: служит для упорядочивания и выравнивания элементов формы.

Все панели инструментов можно перемещать и настраивать на каждом виде с учетом потребностей, включая использование дополнительного монитора.



Панель инструментов «Форма» и «Пометка текста» (см. подробное описание в Глава 4, стр. 70) отображаются только в OmniPage Professional 18.

Основные этапы обработки

Существуют три способа обработки документов: автоматически, вручную или с помощью очереди документов. Основные этапы для всех способов практически не отличаются.



1. Загрузите набор изображений в программу OmniPage. Можно отсканировать бумажный документ с помощью устройства автоматической подачи документов либо загрузить один или несколько файлов изображений из локальной файловой системы, из облачного хранилища или с FTP-сайта.



2. Выполните распознавание, чтобы получить редактируемый текст. После распознавания можно проверить текст и исправить ошибки с помощью функции проверки распознавания, а также отредактировать документ в текстовом редакторе.



3. Экспортируйте документ в нужную папку. Документ можно сохранить с указанным именем и типом, поместить в буфер, отправить в качестве вложения в сообщении электронной почты или опубликовать. Документы можно сохранять многократно, в файлах разного типа, с разными настройками и уровнями форматирования.

При работе с программой OmniPage можно обрабатывать изображения следующими способами: автоматический, ручной, комбинированный или с помощью рабочего процесса. Можно запускать распознавание из других приложений, с помощью функции Direct OCR, а также планировать запуск обработки на более позднее время. Методы обработки подробно описываются в следующей главе и в справке программы.

Параметры



Большая часть параметров программы OmniPage настраивается в диалоговом окне «Параметры». Это окно можно открыть с помощью панели инструментов «Стандартная» или меню «Сервис». Сведения о каждом параметре можно найти в контекстной справке.

Использование программы OmniPage с компонентом PaperPort

PaperPort[®] — это программа компании Nuance для управления документооборотом. Она позволяет связывать страницы с соответствующими приложениями. Страницы могут содержать текст, рисунки или то и другое одновременно. При установке пакета PaperPort на компьютер с программой OmniPage возможности PaperPort расширяются, поскольку становятся доступными функции распознавания текста. Чтобы выбрать программу распознавания, щелкните ссылку PaperPort в текстовом приложении, выберите пункт «Параметры», а затем выберите OmniPage 18 в качестве пакета распознавания. Параметры распознавания текста будут такими же, как в функции Direct OCR.



Программа PaperPort предоставляет максимально простой способ преобразования бумажных документов в цифровые документы, которыми смогут пользоваться другие сотрудники в офисе. PaperPort работает со сканерами, многофункциональными принтерами и сетевыми цифровыми копирами, с помощью которых выполняется преобразование бумажной документации в цифровые документы. Кроме того, программа помогает организовать все электронные документы с помощью удобной и простой в использовании файловой системы документов.

Крупные и четкие эскизы PaperPort позволяют визуально упорядочивать, искать и использовать сканированные документы, включая файлы Word, электронные таблицы, файлы PDF и даже цифровые фотографии. Благодаря инструментам оптимизации сканирования из пакета PaperPort сканированные документы отлично выглядят, а инструменты для создания комментариев позволяют добавлять примечания и выделять фрагменты в любом сканированном изображении.



Программа PaperPort входит в состав пакета OmniPage Professional. Для получения сведений о приложении см. документацию по программе PaperPort. PaperPort необходимо устанавливать и удалять отдельно от программы OmniPage.

Если программа PaperPort установлена, ее структура папок отображается в диалоговых окнах «Загрузить из файла» и «Сохранить в файл» программы OmniPage.

Обработка документов

В данной главе описываются различные способы обработки документов и приводятся сведения о ее ключевых этапах.

Способы обработки

При работе с программой OmniPage можно обрабатывать изображения следующими способами:



Автоматически

Самый быстрый и простой способ обработки документов — предоставить программе OmniPage возможность обработать их автоматически. Выберите необходимые параметры в диалоговом окне «Параметры» и в раскрывающихся списках панели инструментов OmniPage, а затем нажмите кнопку *Начать*. Программа выполнит весь процесс обработки с начала и до конца для каждой страницы документа, по возможности обрабатывая их параллельно. Области на страницах обычно выделяются автоматически.

Вручную



Обработка вручную обеспечивает более точный контроль над процессом обработки страниц. Обрабатывая документ постранично, для каждой страницы можно задавать новые параметры. Программа останавливается перед каждым следующим шагом: получением изображений, выполнением распознавания, экспортом. Это позволяет, например, разметить вручную области или изменить языки

распознавания. Каждый этап начинается с нажатия трех кнопок на панели инструментов OmniPage.

- 5. Первая кнопка используется для получения набора изображений.
- 6. Ручное определение областей используется в тех случаях, когда требуется обработать только часть страницы или предоставить точные указания по выделению областей. Для исключения определенных частей из обработки выберите пропуск

фона или включите эти части в пропускаемые области. Для автоматического разбиения определенных частей на области выберите обрабатываемый фон или области.

- 7. Вторая кнопка используется для запуска распознавания страниц.
- 8. При необходимости выполните проверку и редактирование.
- 9. Третья кнопка используется для сохранения результатов.

По умолчанию при обработке вручную автоматически становятся выбранными все загруженные страницы. Это позволяет распознать все новые страницы с помощью одного щелчка мыши. Отменить эту настройку по умолчанию можно на панели «Обработка» в диалоговом окне «Параметры».

Комбинированная обработка

Документ можно обработать автоматически и просмотреть результаты в текстовом редакторе. Если большинство страниц обработаны правильно, но на некоторых страницах результаты обработки неудовлетворительные, можно переключиться в режим обработки вручную, изменить параметры и распознать неудачные страницы повторно. Можно также получить изображения с обработкой вручную, выделить области на всех изображениях или только на некоторых из них, а затем выполнить автоматическую обработку всех страниц, нажав кнопку «Начать» и выбрав обработку существующих страниц.

Рабочий процесс

Рабочий процесс включает в себя последовательность этапов и соответствующие параметры. Обычно (но не обязательно) он включает в себя этап распознавания. Он не обязательно должен соответствовать шаблону 1-2-3 стандартной обработки. Стандартные и созданные пользователем рабочие процессы перечислены в раскрывающемся списке «Рабочий процесс». Рабочие процессы позволяют более эффективно выполнять повторяющиеся задачи, поскольку все их этапы и параметры предварительно определены. Можно разместить значок агента OmniPage на панели задач. В его контекстном меню будут перечислены рабочие процессы. Чтобы запустить программу OmniPage и запустить рабочий процесс, щелкните его мышью.

Для создания нового рабочего процесса воспользуйтесь помощником рабочего процесса. В нем предлагаются варианты действий и необходимые параметры.

После каждого этапа для добавления следующего этап необходимо нажимать кнопку «Далее». В нем можно также найти дополнительные рекомендации по автоматической обработке. См. раздел «Помощник рабочего процесса» в Глава 4, стр. 89.

В указанное время



В диспетчере пакетной обработки OmniPage можно планировать автоматическое выполнение заданий распознавания и других видов обработки в любое время, даже когда пользователь будет отсутствовать за компьютером.

Это делается с помощью диспетчера пакетной обработки. При этом компьютер можно выключать после настройки задания. Главное, чтобы он был включен в заданное время запуска задания. В случае сканирования страниц на момент запуска задания должен быть включен сканер, а в устройство автоматической подачи загружены страницы.

При создании нового задания сначала открывается окно мастера заданий, а потом окно помощника рабочего процесса, в котором содержится несколько измененный набор вариантов действий и параметров. На первом экране мастера заданий укажите тип и имя задания. Далее определите время начала выполнения и инструкции для повторяющегося задания или задания на отслеживание папки.

Задание будет включать в себя рабочий процесс с инструкциями о времени выполнения. См. раздел «Диспетчер пакетной обработки» в Глава 4, стр. 92.

Обработка из других приложений

Функцию Direct OCR[™] можно использовать для вызова служб распознавания программы OmniPage во время работы в следующих приложениях: Microsoft Office XP или более поздних версий, Corel WordPerfect 12 или X3. В первую очередь необходимо установить флажок «Включить Direct OCR», последовательно выбрав пункты «Сервис» > «Параметры» > «Общие». После этого можно будет с легкостью получить доступ к функциям распознавания с помощью двух кнопок на вкладке Nuance OCR в Office 2007 или 2010 либо на панели инструментов OmniPage.

Настройка Direct OCR

Запустите приложение, которое нужно подключить к программе OmniPage. Запустите программу OmniPage, откройте диалоговое окно «Параметры», выберите панель «Общие», а затем параметр *Включить Direct OCR*. В конечном приложении нажмите кнопку 💽 «Параметры получения текста» на панели инструментов OmniPage (в Office 2007 или 2010 откройте вкладку Nuance OCR). Настройте параметры на указанных ниже панелях.

- Распознавание: языки, словари, макет, шрифты;
- **Обработка:** предварительная обработка изображений, способы открытия PDF-файлов, сохранение;
- Формат вывода: выбор уровня форматирования;
- Direct OCR: автоматическая и ручная разметка областей, выполнение или пропуск проверки, источник изображений;
- Сканер: настройка или изменение параметров сканера.

Эти настройки будут использовать для всех последующих операций Direct OCR до тех пор, пока они снова не будут изменены; они не используются, когда программа OmniPage используется отдельно.

Использование функции Direct OCR

- 1. Откройте приложение и необходимый документ. Для получения результатов распознавания отсканированных страниц, поместите страницы в сканер.
- 2. Чтобы просмотреть параметры распознавания, нажмите кнопку

«Параметры получения изображений» 🔊 на панели инструментов OmniPage или одноименный пункт в меню «Файл» конечного приложения (или на вкладке Nuance OCR в Office 2007 и 2010). На панели Direct OCR можно выбрать источник изображений: сканер, файлы изображений или файлы с цифровой камеры.

3. Для получения изображений из указанного источника используйте кнопку

«Получить текст» (м) на панели инструментов OmniPage или одноименный пункт в меню «Файл» (или на вкладке Nuance OCR в Office 2007 или 2010).

- 4. Если на панели Direct OCR диалогового окна «Параметры», которое открывается при нажатии кнопки «Параметры получения текста», установить флажок *Рисовать области автоматически*, то распознавание начнется автоматически.
- 5. Если флажок *Рисовать области автоматически* не установлен, то будет отображаться изображение каждой страницы, на котором можно вручную отметить области. Для перехода к процедуре распознавания нажмите кнопку «Распознавание».

 Если включена проверка, она будет выполнена после распознавания.
 Затем распознанный текст помещается в приложение на место, указанное курсором. Уровень форматирования задается в разделе «Формат вывода» диалогового окна «Параметры получения текста».

Определение источника изображений страницРисунок

Существует три возможных источника изображений: файлы изображений, цифровая камера и сканер. Существует два основных типа сканеров: планшетный сканер и сканер протяжного типа. Сканер может иметь встроенный или дополнительный автоподатчик, что облегчает сканирование многостраничных документов. Изображения отсканированных документов могут передаваться непосредственно в программу OmniPage или сохраняться с помощью собственного программного обеспечения сканера в файлы изображений, которые позднее можно открыть в программе OmniPage. Минимальная ширина или высота для файла изображения составляет 16 на 16 пикселей; а максимальная – 8400 пикселей (71 см или 28 дюймов при разрешении от 201 до 600 точек на дюйм). Сведения о предельно допустимом количестве пикселей см. в справке.

Способы открытия PDF-файлов настраиваются в меню «Сервис» > «Параметры» > «Обработка». Здесь можно настроить, будут ли PDF-файлы открываться с текстовым слоем или как изображение, нужно ли импортировать данные тегов, чтобы сохранить макет, а также будут ли использоваться шрифты PDF-файлов или системные шрифты. Сведения о добавлении возможности поиска в PDF-файлы, содержащие только изображения, см. в разделе Помощник eDiscovery для PDF с возможностью поиска.

Использование файлов изображений

Можно создавать файлы изображений с помощью сканера либо получать их по электронной почте или в виде файлов факса. В программе OmniPage можно открывать файлы изображений различный типов. В раскрывающемся списке «Получить страницы» выберите пункт *Загрузить файлы*. Файлы указываются в диалоговом окне «Загрузить файлы». Оно открывается при запуске автоматической обработки. При обработке вручную нажмите кнопку «Получить страницу» или воспользуйтесь меню «Обработка». В нижней части диалогового окна отображаются дополнительные параметры. Их можно отобразить или скрыть.

Получение из облачного хранилища

В раскрывающемся списке «Получить страницы» доступны прямые подключения к следующим веб-хранилищам: Evernote и Dropbox.



OmniPage 18 поставляется с компонентом Nuance Cloud Connector, который можно легко настроить, выбрав его в группе OmniPage в меню «Пуск» системы Windows. Укажите сайты облачных хранилищ и FTP-сайты, с которых будут загружаться входные файлы.

При загрузке файлов из облачного хранилища может потребоваться ввести учетные данные.



В программе OmniPage Professional файлы также можно импортировать из служб Microsoft SharePoint 2003, 2007 и 2010, Hummingbird, iManage или ODMA-совместимых систем управления корпоративным содержимым.

Получение с цифровой камеры



Можно перенести фотографии документов с цифровой камеры в программу OmniPage для распознавания. Сначала убедитесь в том, что драйвер устройства установлен правильно. Затем подсоедините

камеру и загрузите изображения. В раскрывающемся списке «Получить страницы» выберите «Загрузить файлы цифровой камеры». При выборе этого параметра для изображений автоматически выполняется 3D-выравнивание, увеличение разрешения и выпрямление текстовых строк. Можно также выполнить 3D-выравнивание (см. раздел Инструменты оптимизации изображения ниже в этой главе).

Чтобы получить фотографии цифровой камеры, содержащие текст, с помощью функции Direct OCR или PaperPort, установите флажок «Загрузить как изображение цифровой камеры». При этом будут выполняться указанные выше автоматические функции оптимизации.

Советы и рекомендации по работе с изображениями с цифровой камеры см. в инструкциях и справке.

Получение через простой загрузчик

Простой загрузчик — это похожий на проводник Windows интерфейс, который можно использовать в окне OmniPage. По умолчанию загрузчик отображается в гибком виде

и виде быстрого преобразования. Чтобы добавить загрузчик в классический вид либо отобразить (или скрыть) его в других видах, выберите пункт «Простой загрузчик» в меню «Окно». Загрузчик является альтернативой диалоговому окну «Открыть файл», позволяя перемещаться по всей файлой системе и быстро выбирать файлы для загрузки в OmniPage. Чтобы просматривать и упорядочивать файлы в виде списка, эскизов, мозаики, таблицы или значков (как в проводнике Windows), выберите пункт «Процесс» > «Простой загрузчик» > «Папка». Окно загрузчика может отображаться во время работы.

Управление простым загрузчиком производится через меню «Процесс». Для выбора и отправки файлов в OmniPage можно открыть диалоговое окно очереди с блокировкой. Включите блокировку, чтобы составить список файлов (которые могут быть из разных папок) или изменить его порядок. Блокируются все файлы, собранные для добавления в открытый документ. После составления списка снимите блокировку, чтобы начать загрузку. Если блокировка выключена с самого начала, файлы добавляются в список только в том случае, если они были выбраны быстрее, чем OmniPage может загрузить их. В простом загрузчике можно загрузить несколько файлов и отправить их на распознавание, во время которого можно продолжить составление списка остальных входных файлов.

Если в меню установить флажок *Автоматически показать или скрыть окно очереди*, окно очереди автоматически откроется вне зависимости от наличия в списке файлов, которые не были загружены, и закроется после того, как список опустеет.

Простой загрузчик можно использовать в классическом и гибком видах для сборки файлов для нескольких документов. Включите блокировку, сделайте документ 1 активным и соберите файлы. Затем сделайте активным документ 2 и соберите файлы для него и т. д. По окончании снимите блокировку. Для каждого документа устанавливается независимая блокировка, но в меню «Процесс» есть команды Заблокировать все и Разблокировать все для блокировки и разблокировки всех файлов всех документов. Можно удалить выбранные файлы командой Удалить либо удалить все файлы в текущем документе командой Удалить все используется для удаления всех файлов, предназначенных для всех открытых документов. Сведения о загрузке файлов для нескольких документов см. в справке.

В виде быстрого преобразования простой загрузчик доступен как панель. В меню «Процесс» есть две команды, которые отображаются только в этом виде.

- Команда Получить и преобразовать обеспечивает обработку одним нажатием кнопки: файлы загружаются, распознаются и сохраняются с использованием существующих параметров. Только в этом случае в виде быстрого преобразования можно выбрать несколько файлов. Результатом будет один выходной документ для каждого входного файла. Перед началом в разделе «Имя выходного файла» нужно установить переключатель Аналогично имени исходного файла.
- Команда Загрузить файлы загружает файлы без распознавания (так же, как и в других видах). Простой загрузчик позволяет загружать только по одному файлу за раз. Каждый входной файл должен быть обработан перед выбором нового. В этом случае окно очереди и его блокировка не играют особой роли.

Простой загрузчик поддерживает работу с изображениями с цифровой камеры. Эту функцию можно настроить на панели параметров быстрого преобразования перед запуском простого загрузчика. Если в качестве источника выбран *сканер*, эта функция временно игнорируется, а страницы загружаются как обычные изображения (не с камеры).

В простом загрузчике доступны все функции проводника Windows. Например, можно выбрать файлы и с помощью контекстного меню OmniPage 18 в фоновом режиме открыть их в формате MS Excel, MS Word, PDF, RTF, TXT и WordPerfect. При этом используются существующие параметры и созданные по умолчанию файлы помещаются в исходную папку. В мастере преобразования доступны основные параметры, в том числе позволяющие выбрать, нужно ли отображать результаты в конечном приложении. Этот мастер позволяет выполнять немедленное преобразование или вызывать помощник рабочего процесса для доступа ко всем параметрам, например для изменения имени конечного файла и папок назначения. Через этот раздел контекстного меню также можно получить доступ ко всем рабочим процессам, которые используют файлы изображений в качестве источников.

Получение со сканера

Для этого требуется исправный поддерживаемый сканер, который правильно установлен в программе OmniPage 18. Можно выбрать один из нескольких режимов сканирования. Выбирая режим, следует учесть два фактора.

- Необходимый тип экспортируемого документа.
- Режим, который обеспечит наивысшую точность распознавания.


Черно-белое сканирование

Выберите этот параметр для выполнения сканирования в черно-белом режиме. Черно-белые изображения сканируются и загружаются быстрее цветных и занимают меньше места на диске.



Сканировать в оттенках серого

Выберите этот параметр для сканирования с оттенками серого. Используйте его для наиболее качественного распознавания страниц с переменной или низкой контрастностью и страниц с цветным или темным фоном.



Сканировать в цвете

Выберите этот параметр для выполнения цветного сканирования. Работа в этом режиме возможна только при наличии цветного сканера. Выбирайте этот режим для загрузки в документ цветной графики или текста. Этот режим менее полезен с точки зрения качества распознавания, чем режим с оттенками серого, а кроме того, он требует больших системных ресурсов и дискового пространства.

Яркость и контрастность

Правильные установки яркости и контрастности очень важны для точного распознавания. Эти параметры задаются на панели «Сканер» диалогового окна «Параметры» или в интерфейсе сканера. Загрузив изображение, проверьте его внешний вид. Если символы жирные и расположены близко друг к другу, увеличьте яркость. Если символы тонкие и неровные, уменьшите яркость. После изменения яркости повторно отсканируйте страницу. Если результаты сканирования по-прежнему неудовлетворительные, откройте сканированное изображение в окне «Оптимизация изображения» и отредактируйте его с помощью имеющегося набора инструментов.

Сканирование с использованием устройства автоматической подачи

Многостраничные документы лучше всего сканировать с помощью устройства автоматической подачи. Просто загрузите страницы в нужном порядке в лоток автоподачи. С помощью устройства автоматической подачи можно сканировать двусторонние документы. Двусторонний сканер делает это автоматически.

Сканирование без использования устройства автоматической подачи

Интерфейс OmniPage позволяет эффективно выполнять сканирование многостраничных документов с помощью планшетного сканера даже без использования устройства автоматической подачи. В диалоговом окне «Параметры» на панели «Сканер» выберите параметр *Сканировать автоматически* и укажите значение паузы в секундах. После этого сканирование будет осуществляться в автоматическом режиме. Между сканированием отдельных страниц сканер будет выдерживать паузу указанной длины. Во время этой паузы помещайте в сканер следующую страницу.

Сканирование с помощью OmniPage и рабочих процессов

Перейдите в раздел «Сервис» > «Параметры» > «Сканеры», чтобы выбрать действие, выполняемое при нажатии кнопки на сканере. Это может быть простое сканирование с загрузкой изображений в OmniPage. Можно также выбрать один из созданных рабочих процессов для работы со сканером или настроить отображение запроса на выбор рабочего процесса при нажатии кнопки. С помощью кнопки «Панель управления» приложение OmniPage можно связать с событием сканера (нажатием кнопки на сканере). После этого при нажатии кнопки будет открываться OmniPage и запускаться рабочий процесс, а результаты будут отправляться в указанное целевое приложение. При этом программа может взаимодействовать с пользователем или выполнять все этапы автоматически.

В OmniPage Professional эту возможность можно также использовать для запуска рабочих процессов обработки штрихкодов (см. Глава 4, стр. 93).

Преобразование документов из одного формата в другой



В программе OmniPage Professional 18 можно открывать не только файлы изображений, но также и документы, созданные с помощью текстовых редакторов и других

подобных приложений. Поддерживаются следующие типы файлов: .doc, .xls, .ppt, .rtf, .wpd и другие. На панели инструментов OmniPage нажмите кнопку «Загрузить файлы» или в меню «Файл» выберите пункты «Получить страницу», а затем «Загрузить файлы». В окне «Загрузить файлы» выберите пункт «Документы». После завершения можно выбрать различные типы файлов документов для сохранения. Для этих преобразований должен быть установлен компонент Nuance PDF Create.

Описание макета документа

Перед запуском распознавания необходимо описать макет исходных страниц для упрощения процесса автоматического распознавания. При автоматической обработке разметка областей выполняется автоматически, кроме случаев, когда задан шаблон, не содержащий обрабатываемых областей и фона. Иногда автоматическая разметка областей выполняется и при ручной обработке. См. раздел справки *Когда запускается автоматическая разметка областей*? Далее приведены варианты описания исходных файлов:



Автоматически

Выберите этот параметр, чтобы программа автоматически выделила все области. В этом случае программа определит, разбит ли текст на колонки, является ли элемент изображением или текстом, который нужно распознать, и имеются ли в тексте таблицы.

Одна колонка без таблиц

Выберите этот вариант, если страницы содержат по одной колонке текста и не содержат таблиц. Обычно это относится к деловым письмам или страницам книги.

Несколько колонок без таблиц

Выберите этот вариант, если на некоторых страницах текст разбит на колонки, и это разбиение нужно убрать или сохранить отдельные столбцы, как в исходном макете.



लो

Одна колонка с таблицей

Выберите этот вариант, если на странице имеется только одна колонка текста и таблица.

Электронная таблица

Выберите этот вариант, если вся страница представляет собой таблицу, которую нужно экспортировать в редактор электронных таблиц или просто обрабатывать как единую таблицу.



Форма

Выберите этот вариант, если вся страница представляет собой форму, и требуется автоматически распознать элементы формы. После распознавания можно изменить свойства элементов формы, добавить новые элементы и отредактировать макет. Этот вариант доступен только в OmniPage Professional 18.



Юридические документы

Выберите этот параметр для распознавания юридических документов. При этом распознаются и удаляются колонтитулы юридических документов. Выберите сохранение или удаление номеров юридических документов.

Настройка

Выберите этот вариант, чтобы получить максимальный контроль над автоматическим выделением областей. В этом случае можно будет включить или выключить определение столбцов, изображений и таблиц. Измените параметры на панели «Распознавание» в диалоговом окне «Параметры».



2

Шаблон

Выберите файл шаблона области, если требуется применять значение его фона, области и свойства ко всем получаемым страницам. Разметка областей шаблона применяется также к текущей странице, заменяя существующую.

Если результаты распознавания после автоматической разметки неудовлетворительны, разметьте страницы вручную и повторите их распознавание.

Предварительная обработка изображений

SET

Для улучшения результатов распознавания изображение перед разметкой

и распознаванием можно подготовить с помощью инструментов оптимизации изображения.

Чтобы открыть окно «Оптимизация изображения», нажмите кнопку «НАСТРОИТЬ — Оптимизировать изображение» на панели инструментов «Изображение». В этом окне есть панель начального изображения (1) и панель результатов (2). Выберите инструмент (см. разделы ниже), а затем с помощью бегунков настройте элементы управления (3). Получив приемлемый результат, нажмите кнопку «Применить» (4). Чтобы отменить действия, выберите команду «Отменить последнее изменение» (5) или «Отменить все изменения» (6). При нажатии кнопки «Применить» изображение перемещается с левой панели на правую и становится новым начальным изображением для дальнейшей оптимизации. Изменения отображаются на панели журнала (7). Внеся все необходимые изменения, нажмите кнопку «Страница готова» (8) для загрузки страницы или «Документ готов» (9) для завершения оптимизации.



Необходимо различать три типа изображений.

Исходное изображение. Изображение, которое было создано с помощью сканера или содержалось в файле, прежде чем оно попало в программу.

Базовое изображение. Состояние исходного изображения после его загрузки в программу OmniPage, возможно, измененное в результате операций автоматической или ручной предварительной обработки.

Изображение для распознавания. Черно-белое изображение, полученное из базового изображения и оптимизированное для правильного распознавания.

Оптимизацию можно применять только к базовому изображению.



Этот инструмент позволяет переключаться между базовым изображением и изображением для распознавания.

Обработка базового изображения и изображения для распознавания выполняется разными инструментами. Обращайте внимание на тип редактируемого изображения.

Правильные установки яркости и контрастности очень важны для точного распознавания. Эти параметры задаются на панели «Сканер» диалогового окна «Параметры» или в интерфейсе сканера. На рисунке показана оптимальная яркость изображения. Загрузив изображение, проверьте его внешний вид. Если символы жирные и расположены близко друг к другу, увеличьте яркость. Если символы тонкие и неровные, уменьшите яркость. Установить оптимальный уровень яркости можно с помощью инструмента «Яркость распознанного».

extensive research extensive research

Неудовлетворительно
Удовлетворительно
Хорошо
Оптимально
Хорошо
Удовлетворительно
Неудовлетворительно

Инструменты оптимизации изображения

Инструменты оптимизации изображения можно применять также для редактирования базовых изображений с целью их сохранения и дальнейшего использования в качестве файлов изображений. На панели инструментов расположены указанные ниже инструменты. Особенности их использования описаны ниже.

Р — влияет только на базовое изображение.

О — влияет только на изображение для распознавания.

PO — применяется к базовому изображению либо к изображению для распознавания (или к обоим).

P+O — одно действие применяется и к базовому изображению, и к изображению для распознавания.

Р/О — влияет на оба изображения.

WH — влияет только на изображение целиком.

AR — применяется к выделенным областям изображения.

- Указатель (F5) это нейтральный инструмент, который в разных случаях выполняет различные операции (например, используется для выбора цвета при заливке или выборе линии выравнивания). РО.
- Масштаб (F6) выберите этот инструмент, затем щелкните по изображению левой кнопкой мыши, чтобы увеличить, или правой кнопкой мыши, чтобы уменьшить его. Для изменения масштаба можно использовать и колесико мыши, причем даже при неактивном окне оптимизации изображения. В активном окне ту же функцию выполняют клавиши «+» и «-». Р+О. WH.

Выделить область (F7) — нажмите эту кнопку, а затем выберите инструмент, который работает с областью страницы (с пометкой AR), и выделите область на изображении. По умолчанию инструменты оптимизации изображения применяются ко всей странице. Существует три режима выделения (режим выбирается в меню «Вид»): «Обычный», «Объединяющий» и «Вычитающий». PO. AR.

Базовое изображение/изображение для распознавания — этот инструмент позволяет переключаться между базовым изображением и изображением для распознавания на активной панели. Базовые изображения могут быть любыми (цветными, в оттенках серого или черно-белыми). Изображение для распознавания – это черно-белый вариант базового изображения, созданный специально для распознавания. Р/О. WH.

Синхронизовать панели — выберите этот инструмент, чтобы увеличить и прокрутить неактивный вид, выбрав то же значение масштаба и положение прокрутки, что и на активном виде. Чтобы изображения на активной и неактивной панелях автоматически синхронизировались, в меню «Вид» или в контекстном меню выберите пункт «Синхронизировать всегда». РО. WH.

Перечисленные ниже инструменты оптимизации позволяют изменять содержимое изображений.



Яркость и контрастность — выберите этот инструмент, чтобы настроить яркость и контрастность базового изображения или его выделенной части. Значения параметров можно задать с помощью бегунков. Р. AR.



Оттенок/насыщенность/освещенность — выберите этот инструмент, чтобы изменить оттенок, насыщенность и освещенность базового изображения с помощью бегунков. Р. AR.

Обрезка — позволяет обрезать часть изображения. Выберите инструмент «Выделить область», затем — «Обрезка» и выделите область, которую необходимо оставить. Остальная часть изображения будет удалена. P+O. WH > AR.



Поворот — выберите этот инструмент, чтобы повернуть (на 90, 180 или 270 градусов) или отразить все изображение. Р+О. WH.



Удаление дефектов — выберите этот инструмент, чтобы удалить случайные точки на изображении для распознавания. Возможны 4 уровня удаления дефектов на изображении для распознавания. Этот инструмент можно также использовать для того, чтобы подчеркнуть контуры букв. Для этого установите флажок «Инвертировать удаление дефектов». О. AR.



Яркость распознавания — используйте этот инструмент, чтобы задать яркость и контрастность изображения для распознавания. См. диаграмму, на которой показана оптимальная яркость, в разделе «Предварительная обработка изображений» выше. О. AR.



Отбросить цвет — щелкните этот инструмент, а затем выберите красный, зеленый, синий или другой цвет на базовом изображении с помощью инструмента «Выбрать область». Участки отсканированного изображения, окрашенные в выбранный цвет, станут прозрачными. Инструмент применяется к изображению для распознавания. Эта функция позволяет исключить выбранный цвет при сканировании или загрузке цветных форм. Затем определенный текст, рамки и другие элементы можно удалить, оставив видимыми только необходимые для распознавания. Р/О. WH.



Разрешение — используйте этот инструмент для уменьшения разрешения базового изображения в процентах. Примечание: невозможно установить разрешение выше изначального. Р. WH.



Выравнивание — иногда страницы сканируются с перекосом. Чтобы выпрямить строки текста вручную, используйте инструмент «Выравнивание». Автоматическое выравнивание можно установить на панели обработки в окне «Параметры». Р+О. WH.



3D-выравнивание — используйте этот инструмент для устранения искажения перспективы на изображениях с цифровой камеры. Это особенно полезно, когда требуется проверить результаты автоматического 3D-выравнивания или необходимо выполнить 3D-выравнивание вручную после этапа загрузки файлов. Р+O. WH.

3D-выравнивание осуществляется посредством привязки искаженного изображения к сетке. Пользователю достаточно вручную выпрямить сетку. Координаты изображения изменятся соответствующим образом. См. рисунок ниже (до и после 3D-выравнивания).





Заливка — этот инструмент позволяет применить однородный цвет к изображению или его выделенной части. РО. AR.



Автообрезка — этот инструмент автоматически определяет поля на странице и уменьшает их до минимума. Таким образом можно сделать одинаковыми поля на большом количестве страниц с областями текста различного размера. P+O. WH. > AR



Удалять границы — удаление теней сканирования, точек и ненужных заметок на полях. Р+О. WH. Применимо только к области по краям.



Удаление отверстий от дырокола — закрашивание отверстий от дырокола цветом фона страницы. Р+О. WH. Применимо только к области по краям.

Улучшать фотографии белой доски — позволяет с помощью бегунка повысить удобочитаемость текста и диаграмм на белых или черных досках, снятых цифровой камерой. На рисунках ниже показана возможная разница при использовании этого инструмента вместе с 3D-выравниванием.

Ниже представлена типичная цифровая фотография белой доски, снятая сбоку и с низкой контрастностью.



А это пример использования 3D-выравнивания и его результат (справа).



С помощью бегунка инструмента «Улучшать фотографии белой доски» можно повысить контрастность изображения. Слева показано исходное изображение, а справа — результат.

4 4 4 4 4 4 4 4 4 4 4 4 4 4	ydney ndon erlin udapest	- Sydney - London - Berlin - Budopest
	*	

Некоторые из этих инструментов также можно использовать для автоматической предварительной обработки всех входных изображений. Они отображаются на панели «Обработка» диалогового окна «Параметры».

Использование журнала оптимизации изображения

Чтобы принять или отменить изменения изображения (по одному или все вместе), используйте панель журнала в окне «Оптимизация изображения». Изменения, примененные к исходному изображению, отображаются в окне результата.

D

Чтобы подтвердить изменения, нажмите кнопку «Применить» справа от списка журнала. Изменения, не добавленные в журнал изменений с помощью кнопки «Применить», не применяются.

- Чтобы отменить внесенные изменения перед их применением, нажмите кнопку «Сброс».
- Чтобы восстановить вид изображения на момент начала текущего сеанса оптимизации, нажмите кнопку «Отменить все изменения».

Результат применения каждого этапа преобразований можно посмотреть в любой момент, дважды щелкнув по соответствующему пункту в списке журнала. На экране будут показаны результаты действия и убраны все последующие действия. Если применить новое изменение к показанному изображению, будут заменены все изменения, сделанные в списке журнала после выбранного изменения.

Сохранение и применение шаблонов

Для оптимизации нескольких похожих изображений можно составить список этапов оптимизации и применить его ко всем изображениям.

Для создания и сохранения шаблона оптимизации поместите файл изображения в окно «Оптимизация изображения», выполните необходимые этапы оптимизации и добавьте их в журнал, нажав кнопку «Применить». Затем в окне «Оптимизация изображения» в меню «Файл» выберите пункт «Сохранить шаблон оптимизации». Выберите папку и сохраните файл шаблона (с расширением .ipp).

Для применения преобразований, сохраненных в файле шаблона, к другому изображению откройте в окне «Оптимизация изображения» новое изображение и выберите в меню «Файл» пункт «Загрузить шаблон оптимизации».

Оптимизация изображений в рабочих процессах



Чтобы выполнить оптимизацию изображений с помощью рабочего процесса, выберите соответствующий значок в помощнике рабочего процесса.

При этом можно настроить следующие параметры.

Открыть изображения для оптимизации вручную. При выполнении рабочего процесса каждое загружаемое изображение для оптимизации вручную будет отображаться на дисплее.

Применить шаблон оптимизации. При выполнении рабочего процесса к изображению автоматически применяется заранее сохраненный шаблон оптимизации.

Применить шаблон оптимизации и вывести на экран. Рабочий процесс применяет выбранный шаблон оптимизации и выводит изображение для дальнейшего редактирования.

Области и фон

Областями называются фрагменты страницы, которые могут обрабатываться или пропускаться программой. Области бывают прямоугольной и неправильной формы с вертикальными и горизонтальными сторонами. Для изображений страниц указывается значение параметра обработки фона: обработать или пропустить (обычно последнее). Значение параметра можно изменить с помощью соответствующих инструментов. Области можно размечать на фоне страниц с помощью инструментов, представленных на вкладках «Типы областей» и «Свойства».

Обрабатываемые участки областей или фона автоматически размечаются перед распознаванием.

Пропускаемые участки областей или фонов не обрабатываются. Текст не распознается, и изображения не переносятся.

Автоматическая разметка областей

При автоматической разметке программа сама выделяет блоки текста, колонтитулы, рисунки и другие элементы страницы и размечает содержащие их области.

Автоматически можно разметить как всю страницу, так и ее часть. Области, размеченные автоматически или с применением шаблона, имеют сплошные границы. Измененные области и области, выделенные вручную, имеют пунктирные границы.

Автоматическая разметка областей на фоне

Получите страницу. Фон в ней будет обрабатываемым. Выделите область. Фон станет необрабатываемым. Выделите текстовые, табличные или графические области внутри области, размечаемой вручную. Выберите инструмент обработки фона (показан на рисунке), чтобы сделать фон обрабатываемым. Выделите пропускаемые области на тех участках страницы, которые не нужно обрабатывать. После распознавания страница отобразится с пропускаемым фоном и новыми областями, включающими все обнаруженные на фоне элементы.

Автоматическая разметка вертикального текста

При распознавании японского, корейского или китайского текста функция автоматической разметки областей сама находит текстовые блоки и определяет в них направление текста. Вертикальный азиатский текст отображается в текстовом редакторе горизонтально, но его можно экспортировать как вертикальный (см. Глава 4, стр. 62).

При автоматической разметке областей в тексте на неазиатских языках выявляются вертикальные фрагменты в ячейках таблиц и любых частях PDF- или XPS-страниц. В таких случаях могут быть выявлены сразу несколько строк.

Для PDF и XPS-документов, содержащих только изображения, а также для файлов изображений и результатов сканирования автоматическая разметка работает только при соблюдении указанных ниже условий.

- В файле должна быть только одна строка текста.
- Текст должен располагаться слева или справа от диаграммы или изображения.
- Текст должен располагаться у правого или левого края страницы, а его высота должна быть меньше общей высоты страницы.

Вертикальный текст за пределами таблиц необходимо разметить на области вручную. Это позволит правильно обработать вертикальные строки.

Чтобы вертикальный текст можно было просматривать и изменять в текстовом редакторе с помощью вертикального курсора, рекомендуется использовать уровень форматирования «Фактическая страница». При использовании других уровней форматирования текст располагается горизонтально.

Типы и свойства областей

Каждая область относится к определенному типу. У областей, содержащих текст, может также присутствовать параметр «Содержимое области», имеющий два значения: буквенно-цифровое или числовое. Тип и содержимое области составляют свойства области. Свойства области можно изменить в контекстном меню, которое открывается целчком правой кнопки мыши. Можно выбрать несколько областей, удерживая клавишу Shift при выборе областей левой кнопкой мыши, и изменить их свойства одновременно.

Панель инструментов «Изображение» содержит инструменты разметки областей, по одному для каждого типа.



Обрабатываемая область

Используйте этот инструмент, чтобы определить область на странице, где будет выполняться автоматическая разметка областей. После распознавания такая область заменяется несколькими областями, тип которых определяется автоматически.



Пропускаемая область

Используйте этот инструмент для определения области пропуска, то есть области страниц, которая не передается в текстовый редактор.



Область текста

Используйте этот инструмент, чтобы определить область текста. Обычно выделяется отдельный блок текста. Содержимое области будет обрабатываться как поточный текст без разбиения на колонки. Эта область используется для текста на основе латиницы, кириллицы и греческого алфавита, а также для горизонтального текста на азиатском языке.

Область вертикального азиатского текста Используйте этот инструмент, чтобы разметить область для вертикального текста на японском или китайском языке. Области должны быть прямоугольной формы.

Область вертикального текста, повернутого влево Используйте этот инструмент, чтобы разметить область вертикального текста, повернутого влево (только для неазиатских языков). Области должны быть прямоугольной формы.



Область вертикального текста, повернутого вправо

Используйте этот инструмент, чтобы определить область вертикального текста, повернутого вправо (только для неазиатских языков). Области должны быть прямоугольной формы.



Область таблицы

Используйте этот инструмент, чтобы считать содержимое области таблицей. Сетки таблицы могут определяться автоматически или размещаться вручную. Области таблицы должны быть прямоугольной формы. Вертикальный текст в таблицах нельзя разбить на области вручную. Он автоматически размечается в таблице с сеткой.



Область изображения

Используйте этот инструмент, чтобы выбрать изображение, диаграмму, рисунок, подпись или любой другой объект, который необходимо перенести в текстовый редактор в виде встроенного изображения, а не распознаваемого текста.



Область формы



Используйте этот инструмент, чтобы выбрать область документа, содержащую элементы формы, например флажок, переключатель, текстовое поле или любой другой объект, который необходимо перенести в текстовый редактор в виде элемента формы. На уровне форматирования «Фактическая страница» можно редактировать макет формы и изменять свойства элементов формы. Области формы доступны только в версии OmniPage Professional.

Работа с областями



Панель инструментов «Изображение» содержит инструменты для редактирования областей. Сгруппированные инструменты можно откреплять и делать плавающими или повторно закреплять в виде отдельной мини-панели инструментов. Один из них всегда активен.

Если этот инструмент больше не нужен, щелкните любой другой. Некоторые инструменты объединены в группы. В случае закрепления в виде одного инструмента то остается видимым только последний выбранный инструмент в группе. Чтобы выбрать видимый инструмент, достаточно щелкнуть его.

Для выделения одной области выберите инструмент для обозначения области нужного типа и обведите курсором нужную область, удерживая нажатой левую кнопку мыши.

Для изменения размеров области выделите область щелчком мыши, наведите курсор на сторону или угол области и растяните ее до нужных размеров. Области не должны перекрываться.

Для создания области неправильной формы путем добавления выделите частично перекрывающуюся область того же типа.

Чтобы объединить две области одного типа, выделите перекрывающуюся область того же типа (новые области слева, итоговые области справа).



Для создания области неправильной формы путем вырезания выделите перекрывающуюся область того же типа, что и фон.

Для разделения области нужно создать разделяющую область одного типа с фоном.

Полный набор изображений областей приводится в справке программы.

Если новая область при выделении частично перекрывает уже существующую область другого типа, то перекрываемая часть существующей области будут заменена новой областью.

Следующие типы областей не допускаются:







Быстрая разметка областей позволяет быстро выполнить разметку областей вручную. Включите курсор выбора области и выделите им изображение на странице. Участки с автоматически размеченными областями будут помечены затенением. Для преобразования помеченных участков в области дважды щелкните их.

Сетки таблиц на изображении

- После автоматической обработки на страницу могут помещаться областитаблиц.
- F Они обозначаются соответствующим значком в левом верхнем углу области.
- С помощью контекстного меню можно преобразовать прямоугольную область
- в таблицу и наоборот. Области таблиц можно размечать вручную, но они должны ₽ сохранять прямоугольную форму. ×
 - Чтобы определить сетку таблицы в текстовом редакторе, можно создавать
- X или перетаскивать разделительные линии столбцов или строк. Можно создать область таблицы или изменить ее размер (при условии, что она останется прямоугольной), чтобы удалить ненужные внешние столбцы или строки таблицы.

С помощью инструментов для работы с таблицами можно вставлять, перемещать и удалять разделители строк и столбцов. Для автоматического определения и установки разделителей нажмите кнопку «Добавить/удалить разделители».

Форматы границ и сетки можно выбирать в контекстном меню. Еще больший выбор границ и видов заливки предоставляется в текстовом редакторе после распознавания.

Использование шаблонов областей

Файл шаблона содержит значение параметра обработки фона и набор областей с их свойствами. Можно загрузить файл шаблона области, чтобы во время распознавания использовались области шаблона. Файл шаблона можно загрузить из списка «Описание макета» или из меню «Сервис». Кроме того, можно загрузить из сетевой папки шаблон, созданный другим пользователем.

При загрузке шаблона его фон и области применяются следующим образом:

- к текущей странице с заменой всех имеющихся областей; ٠
- к следующим страницам; ٠
- к уже существующим страницам, отправленным на распознавание без разметки.

В двух первых случаях при обработке вручную шаблоны областей перед распознаванием можно просматривать и изменять.

При автоматической обработке шаблоны областей можно просматривать и изменять только после распознавания.

При обработке с использованием рабочего процесса включите в процесс этап выделения областей на изображении. Он сочетает в себе два действия: загрузка шаблонов и разметка областей вручную. Чтобы использовать шаблон области, нажмите кнопку «Добавить» на соответствующей панели помощника рабочего процесса и выберите файл шаблона. Затем выберите один из вариантов: открыть изображения для разметки вручную, применить шаблон или применить шаблон и открыть изображение.

В шаблонах указываются пропускаемые и обрабатываемые области, а также фон. Поэтому они могут быть полезны при определении того, какие части страниц следует разделять на области автоматически, а какие – не обрабатывать. Обрабатываемые области или области обрабатываемого фона из шаблона могут быть заменены при распознавании набором областей меньшего размера. Каждой области будет назначен соответствующий тип.

Сохранение шаблона области

Выберите параметры обработки фона и подготовьте области страницы. Проверьте их местоположение и свойства. Выберите пункт «Шаблон областей...» в меню «Сервис». В диалоговом окне выберите пункт [Области на странице], нажмите кнопку «Сохранить», затем назначьте имя и другой путь (необязательно). Чтобы совместно использовать шаблон, поместите его в сетевую папку. Нажмите кнопку «ОК». Новый шаблон области останется открытым.

Изменение шаблона области

Загрузите шаблон и получите нужное изображение посредством обработки вручную. Появятся области из шаблона. Измените области и/или свойства, как необходимо. Откройте диалоговое окно «Файлы шаблонов областей». Текущий шаблон уже выбран. Нажмите кнопку «Сохранить», а затем «Закрыть».

Выгрузка шаблона

Выберите в списке «Описание макета» вариант, не связанный с шаблоном. Области шаблона не удаляются с текущих и остальных страниц, однако при дальнейшей обработке они не будут использоваться. Можно также открыть диалоговое окно «Файлы шаблонов областей», выбрать пункт [HeT] и нажать кнопку «Сделать текущим». В этом случае описание макета вновь станет определяться автоматически.

Замена одного шаблона другим

Выберите другой шаблон в списке «Описание макета» или откройте диалоговое окно «Файлы шаблонов областей», выберите нужный шаблон и нажмите кнопку «Сделать текущим». Области нового шаблона применяются к текущей странице, заменяя все существующие области. Они применяются точно так же, как описано ранее.

Удаление файла шаблона

Откройте диалоговое окно «Файлы шаблонов областей». Выберите шаблон и нажмите кнопку «Удалить». Области, выделенные в соответствие с этим шаблоном, не удаляются. Файл шаблона можно удалить только из операционной системы.

Включение файла шаблона в документ OmniPage

Откройте документ, затем выберите в меню «Сервис» пункт «Шаблон области». Выберите шаблон, который необходимо включить, а затем нажмите кнопку «Встроить». Затем сохраните документ в формате OPD. Это означает, что шаблон будет перемещаться вместе с документом OPD в случае его отправки в новое местоположение. Если затем открыть OPD-файл, включенный шаблон области отобразится в диалоговом окне «Файлы шаблонов областей» как [ВСТРОЕННЫЙ]. Его можно сохранить в новом файле шаблона в новом месте с помощью кнопки «Извлечь».

Проверка и редактирование

Результаты распознавания отображаются в текстовом редакторе. Это могут быть тексты, таблицы, формы и встроенные изображения. В данной главе подробно описывается текстовый редактор, работающий по принципу WYSIWYG («что видите, то и получаете»). Обработка азиатского текста несколько отличается от обработки других языков. См. раздел «Распознавание текста на азиатских языках» на стр. 63.

Отображение в редакторе и уровни форматирования

В текстовом редакторе отображается распознанный текст, а также могут отмечаться слова, вызвавшие сомнения во время распознавания, красным цветом, подчеркиванием или волнистой линией. В программе проверки распознавания такие слова выделяются красным цветом.

Слово считается сомнительным, если оно не обнаружено ни в одном из активных словарей (стандартном, пользовательском или профессиональном). Однако даже обнаруженное в словаре слово может считаться сомнительным по результатам распознавания. Символы, отличающиеся от стандартных, выделяются желтым маркером как в редакторе, так и в программе проверки распознавания.

На панели «Проверка» диалогового окна «Параметры» нужно указать, следует ли отмечать слова, отсутствующие в словаре. С помощью флажков на панели текстового редактора в диалоговом окне «Параметры» можно показать или скрыть все отметки. Можно также отобразить или скрыть непечатаемые символы, а также индикаторы колонтитулов. На панели текстового редактора можно задать единицы измерения и настроить перенос по словам. Эти настройки будут применяться на всех уровнях форматирования текстового редактора, за исключением уровня «Простой текст».

В программе OmniPage могут отображаться три уровня форматирования. Их можно переключать тремя кнопками в левом нижнем углу текстового редактора или с помощью меню «Вид».



🔳 Простой текст

Используется для простого вывода текста, выровненного по левому краю, без колонок, с применением одного типа и размера шрифта, а также разрывов строк, аналогичных исходным.



Форматированный текст

Используется для вывода текста с применением стилей абзацев и шрифтов, без колонок.



Фактическая страница

На уровне форматирования True Page[®] («Фактическая страница») сохраняется максимально возможный объем форматирования исходного документа, включая стиль символов и абзацев. Порядок обхода обозначается стрелками.

Проверка результатов распознавания

После распознавания страницы результаты распознавания выводятся в текстовом редакторе. Проверка начинается автоматически, если она включена на панели проверки диалогового окна «Параметры». Проверку вручную можно начать в любое время. Выполните следующие действия. AB.

- 1. Выберите инструмент «Проверить распознанное» на стандартной панели инструментов или выберите в меню «Сервис» пункт «Проверить распознанное...».
- 2. Проверка начнется с текущей страницы. Уже проверенный текст пропускается. При обнаружении возможной ошибки программа проверки распознавания обозначает сомнительное слово цветом, выделяет символы желтым и показывает вид слова на исходном изображении. В описании указывается «Сомнительное слово» или «Отсутствующее в словаре слово».
- 3. Если слово распознано правильно, нажмите кнопку «Пропустить» или «Пропустить все», чтобы перейти к следующему сомнительному слову. Чтобы добавить слово в текущий пользовательский словарь и перейти к следующему сомнительному слову, нажмите кнопку «Добавить».
- 4. Если в распознанном слове содержится ошибка, измените его на панели правки или выберите вариант из словаря. Нажмите кнопку «Изменить» или «Изменить все», чтобы применить изменение и перейти к следующему сомнительному слову. Чтобы добавить измененное слово в текущий пользовательский словарь и перейти к следующему сомнительному слову, нажмите кнопку «Добавить».

- 5. Вариант можно выделить указателем мыши и принять, нажав кнопку «Изменить», а также можно ввести номер варианта, удерживая нажатой клавишу CTRL.
- 6. После проверки слов в текстовом редакторе цветное выделение снимается. Во время проверки можно перейти в текстовый редактор и делать исправления там. Чтобы возобновить проверку, нажмите кнопку «Возобновить». Нажмите кнопку «Страница готова», чтобы перейти к следующей странице, и «Документ готов» или «Закрыть», чтобы прекратить проверку до достижения конца документа.
- Если проверка выполнена до конца страницы, на ее эскиз и в диспетчере документов добавляется значок выполненной проверки 🖾. Выберите в меню «Сервис» параметр «Перепроверить страницу...», чтобы выполнить повторную проверку страницы.

Верификация текста

Выполнив распознавание, можно сравнить любую часть распознанного текста с соответствующей частью исходного изображения и проверить, правильно ли распознан текст.



Инструмент верификатора находится на панели инструментов форматирования. Управлять верификатором можно также с помощью меню «Сервис». Панель инструментов верификатора отображается при наведении указателя на окно верификатора. Ниже показано применение верификатора.



Чтобы включить верификатор, выберите инструмент «Верификатор» или нажмите клавишу F9. Чтобы его отключить, нажмите кнопку «Верификатор» еще раз, еще раз нажмите клавишу F9 или нажмите клавишу Esc.

Полный перечень сочетаний клавиш верификатора приводится в справке по программе.

Таблица символов



Таблица символов — это прикрепляемый инструмент, который помогает выполнять проверку. Она применяется преимущественно с двумя целями.

- Для вставки во время проверки и редактирования символов, которые отсутствуют на клавиатуре или являются труднодоступными. В этом отношении она очень похожа на таблицу символов системы Windows.
- Для демонстрации всех допустимых символов текущих языков распознавания.

Для получения доступа к таблице символов нажмите соответствующую кнопку на панели инструментов «Форматирование» или выберите в меню «Вид» пункт «Таблица символов» и нажмите кнопку «Показывать».

В меню «Таблица символов» можно также выбрать отображение только недавно использовавшихся символов или другие наборы символов (по умолчанию отображаются только два набора). Азиатские символы не поддерживаются.

Таблицу символов можно открыть и другими способами.

- Выберите «Сервис» > «Параметры», а затем откройте вкладку «Распознавание». Нажмите кнопку «Дополнительные символы» и выберите символы, которые следует использовать при проверке. Аналогичным образом с помощью таблицы символов можно изменять нераспознанные символы.
- Выберите в меню «Сервис» пункт «Эталон символа». Нажмите кнопку (...) рядом с полем «Правильно».
- В текстовом редакторе в контекстном меню сомнительного слова или слова, отсутствующего в словаре, выберите пункт «Эталон символа».

Пользовательские словари

В программе имеются встроенные словари для многих языков. Они оказывают помощь при распознавании и могут предлагать возможные варианты при проверке. Эти словари можно дополнять пользовательскими словарями. Можно сохранять любое количество пользовательских словарей, однако загрузить несколько словарей одновременно нельзя. В качестве словаря, применяемого по умолчанию в программе Microsoft Word, используется словарь «Пользовательский».

Создание пользовательского словаря

В диалоговом окне программы проверки распознавания нажмите кнопку «Добавить», если пользовательский словарь не загружен, или откройте диалоговое окно «Файлы пользовательских словарей» из меню «Сервис» и нажмите кнопку «Создать».

Загрузка и выгрузка пользовательского словаря

Эта операция выполняется с помощью панели «Распознавание» диалогового окна «Параметры» или в диалоговом окне «Файлы пользовательских словарей».

Редактирование или удаление пользовательского словаря

Чтобы добавить слова, загрузите пользовательский словарь, а затем нажмите кнопку «Добавить» в диалоговом окне «Программа проверки распознавания». Слова можно добавлять и удалять, нажав кнопку «Изменить» в диалоговом окне «Файлы пользовательских словарей». Слова можно также импортировать из пользовательских словарей программы OmniPage (*.ud). Во время редактирования пользовательского словаря можно импортировать список слов из текстового файла, чтобы быстро добавить слова в словарь. Каждое слово в тексте должно располагаться на отдельной строке и не должно сопровождаться знаками препинания. С помощью кнопки «Удалить» можно удалить выбранный пользовательский словарь из списка. Чтобы встроить пользовательский словарь в документ OmniPage, загрузите входной файл, выберите «Сервис» > «Пользовательский словарь», выберите пользовательский словарь, который необходимо использовать, нажмите кнопку «Встроить» и введите необходимо имя. Затем сохраните в файл типа документа OmniPage.

Языки

Программа поддерживает более 120 языков на основе различных алфавитов: латиницы и кириллицы, а также греческого, китайского, японского и корейского. Полный список языков можно просмотреть на панели «Распознавание» в диалоговом окне «Параметры». Там указано, для каких языков имеются словари. Выберите язык или языки, которые будут в распознаваемых документах. Если указать большое число языков, точность распознавания может снизиться.

Список языков также можно найти на веб-сайте Nuance.

Параметр Автоматически определять один язык отменяет необходимость выбирать языки. Он предназначен для автоматической обработки документов или форм, которые могут быть на разных языках. При этом проверяется каждая входная страница и ей назначается один язык распознавания. Следовательно, эта функция не подходит для страниц, содержащих текст на нескольких языках.

Языки выбираются из списка языков с поддержкой словаря на основе латиницы (т. е. русский и греческий языки исключаются), включая дополнительные азиатские языки. Выбирать можно из трех групп языков:

- на основе латиницы (выберите этот вариант, чтобы увидеть все включенные языки);
- азиатские языки (китайский традиционный, китайский упрощенный, японский и корейский);
- на основе латиницы и азиатские языки.

Если этот параметр включен, выбрать язык вручную нельзя, а параметр *Проверять выбранные языки* (см. ниже) будет недоступен.

В дополнение к пользовательским словарям представлены специализированные профессиональные словари для некоторых языков (в настоящее время медицинский,

юридический и финансовый). Просмотреть список и выбрать нужные словари можно на панели распознавания в диалоговом окне «Параметры».

Распознавание текста на азиатских языках

Приложение поддерживает четыре азиатских языка: японский и корейский, а также традиционное и упрощенное письмо китайского языка. Идеальный размер шрифта для основного текста составляет 12 пт, а разрешение сканирования — 300 тчк/дюйм. При этом полученные символы будут иметь размер около 48 x 48 пикселей. Минимальный размер символов — 30 x 30 пикселей (10,5 пт при разрешении 300 тчк/дюйм). Чтобы получить символы меньшего размера, используйте разрешение 400 тчк/дюйм. Азиатский текст может быть горизонтальным (слева направо) или вертикальным (сверху вниз, справа налево). Операционные системы, поддерживаемые приложением OmniPage 18, могут работать с азиатскими языками. Однако если во время установки системы не была выбрана поддержка восточноазиатских языков, ее необходимо включить, установив флажок «Панель управления» > «Язык и региональные стандарты» > «Языки» > «Дополнительная языковая поддержка» > «Установить поддержку языков с письмом иероглифами». Может потребоваться вставить установочный диск Windows.

Четыре азиатских языка перечислены в алфавитном порядке вместе с другими языками на панели «Параметры» > «Распознавание». Эти языки рекомендуется выбирать по отдельности, не сочетая их с другими языками. При обработке текста на азиатских языках могут быть распознаны небольшие фрагменты английского текста без явного указания этого языка. Однако данная функция не предназначена для больших текстов на английском или других западных языках. Вертикальный текст на японском и китайском языке может содержать фрагменты на английском языке в различных ориентациях. Программа может обрабатывать все эти варианты. На выходе будет получен текст, повернутый вправо.

Кроме отображения списка языков, функция *Проверить выбранные языки* запускает автоматическую проверку языка, которая предупреждает о различиях между определенным и указанным языками. Эта функция работает на уровне страницы и может определить четыре категории языков: японский, китайский, корейский и неазиатский. Однако она не различает традиционное и упрощенное письмо китайского языка, а также неазиатские языки. Функция относит язык к четвертой категории, если ей не удается найти в тексте японские, китайские или корейские иероглифы. Проверка выполняется во время предварительной обработки изображения, поэтому нужный язык распознавания необходимо задать до загрузки изображения.

Для страниц на азиатских языках рекомендуется использовать автоматическое определение макета и автоматическую разметку областей. При этом весь найденный текст разделяется на области. Если выбрать азиатский язык, функция распознавания текста будет искать азиатский текст в этих областях, автоматически определяя и передавая направление текста, а также обрабатывая смешанные области горизонтального и вертикального текста на странице.

Однако инструмент разметки областей 🕅 позволяет принудительно распознать вертикальный азиатский текст путем разметки вручную. Просто выделите прямоугольные области с помощью этого инструмента. Чтобы вручную разметить области в горизонтальном тексте на азиатском языке, используйте тип областей, предназначенный для обычного текста. Не используйте два других инструмента для вертикального текста на азиатских языках. Выделение области вертикального азиатского текста не приводит к автоматическому включению азиатского языка, а также никак не влияет на автоопределение языка.

Для азиатских языков можно использовать изображения с цифровых фотоаппаратов. Однако в таких случаях алгоритм автоматического трехмерного выравнивания чаще всего оказывается бесполезным, особенно для вертикального текста. Лучше использовать обычную команду загрузки изображений и при необходимости вручную выполнить трехмерное выравнивание с помощью соответствующего инструмента оптимизации изображений. Как правило, инструменты оптимизации изображений можно использовать для изображений с текстом на азиатских языках.

Страницы распознанного азиатского текста отображаются в текстовом редакторе (всегда в горизонтальном направлении) при условии, что система поддерживает восточноазиатские языки. Нет необходимости указывать азиатские шрифты в разделе «Параметры» > «Распознавание текста», поскольку автоматически применяется шрифт по умолчанию (как правило Arial Unicode MS). В текстовом редакторе можно выбрать другие азиатские шрифты, доступные в системе. В редакторе поддерживается просмотр и проверка текста. В качестве уровня форматирования рекомендуется «Форматированный текст». Внесение существенных изменений и проверку орфографии рекомендуется выполнять в конечном приложении. Проверка, а также эталоны и словари не поддерживаются для азиатского текста. Поэтому перед распознаванием азиатского текста откройте панель «Проверка» в разделе «Параметры» и отключите пометку слов, которых нет в словаре, автоматическую проверку правописания и функцию IntelliTrain, а также следите за тем, чтобы не загружался файл эталона. Азиатский текст можно редактировать с помощью выделения или поиска.

К страницам на азиатских языках не следует применять этап рабочего процесса «Извлечение данных формы».

Обычно азиатский текст может быть выведен в формате RTF, Microsoft Word, PDF с возможностью поиска или XPS. Направление текста будет определено во время предварительной обработки. Экспортируются и изменения, сделанные в текстовом редакторе (где текст является горизонтальным), равно как и для вертикального текста. Кроме того, поддерживается преобразование в простой текст (Юникод, формат Блокнота), но при этом всегда используется горизонтальное направление.

Эталон

Создание эталона – это процесс изменения распознанных символов или их комбинаций, назначенных символам в изображении. Это полезно при работе с документами низкого качества или при использовании нестандартных шрифтов. В программе OmniPage эталоны можно создавать двумя способами: вручную и автоматически (IntelliTrain). Данные, полученные в результате создания эталонов, можно комбинировать и сохранять в файл эталона.

При выходе со страницы, на которой создавались эталонные данные, появляется диалоговое окно с вопросом, как применять эти данные к другим страницам документа.

Создание эталонов вручную

Для сохранения эталонов вручную поместите курсор перед нужным символом или выберите группу символов (не больше одного слова). Выберите в меню «Сервис» или в контекстном меню пункт «Создать эталон...». Символы, для которых создается эталон, отобразятся в увеличенном виде вместе с результатом их распознавания. Внесите нужные изменения и нажмите кнопку «ОК». Программа примет изменения и проверит оставшуюся часть страницы. Если будут найдены слова с похожими символами, они отобразятся в диалоговом окне «Проверка эталонов». Перед принятием списка нужно заново создать эталон для неправильно распознанных слов.

IntelliTrain

IntelliTrain представляет функцию автоматического создания эталонов. Она основывается на исправлениях, сделанных вручную в процессе проверки. При внесении изменений происходит запоминание очертаний символов и внесенных изменений. Осуществляется поиск похожих символов в документе и особенно в сомнительных словах. Оценивается необходимость применения исправлений, сделанных пользователем.

Функцию IntelliTrain можно включать и отключать на панели «Проверка» в диалоговом окне «Параметры». Программа IntelliTrain запоминает собранные эталонные данные и дополняет ими процесс создания эталонов вручную. Эталоны могут быть сохранены в файл эталонов для дальнейшего использования в похожих документах.

Примеры интеллектуальной проверки IntelliTrain приводятся в справке по программе.

Файлы эталонов

При каждом закрытии документа или переключении в другой документ, когда имеются несохраненные данные эталонов, отображается диалоговое окно, позволяющее сохранить эти данные. Чтобы сохранить файл эталона в документе OPD, загрузите его с помощью команд меню «Сервис» > «Файл эталона», выберите «Встроить» и сохраните в файл типа документ OmniPage.

Сохранение данных эталона в файл, загрузка, редактирование и выгрузка файлов эталона выполняется в диалоговом окне «Файлы эталонов».

Несохраненные файлы эталонов можно редактировать в диалоговом окне «Изменение эталона», вместо имени файла эталона в строке заголовка отображается звездочка. Их можно сохранять в диалоговом окне «Файлы эталонов».

Файл эталона также можно изменять. Его имя отображается в строке заголовка. Если в файле присутствуют несохраненные эталоны, после имени файла отображается звездочка. Все несохраненные или измененные эталоны сохраняются при закрытии диалогового окна.

В диалоговом окне «Изменение эталона» отображаются рамки с изображением символа и соответствующий результат распознавания. Выделите рамку. Теперь можно удалить ее с помощью клавиши Delete или изменить соответствие. Для перехода к следующей или предыдущей рамке используйте клавиши со стрелками.



Редактирование текста и изображения

В программе OmniPage имеется текстовый редактор WYSIWYG, располагающий большим набором инструментов для редактирования. Его функции во многом напоминают функции ведущих текстовых процессоров.

Изменение атрибутов символов

На всех уровнях форматирования, кроме «Простой текст», можно изменять тип, размер и атрибуты (полужирный, курсив, подчеркнутый) шрифта выделенного текста.

Изменение атрибутов абзацев

На всех уровнях форматирования (кроме «Простой текст») для выделенных абзацев можно задать выравнивание или оформить их в виде маркированных списков.

Стили абзацев

Стиль каждого абзаца при распознавании определяется автоматически. Список стилей создан и представлен в поле выбора в левой части панели «Форматирование». Он используется для применения стиля к выделенным абзацам.

Изображения

Можно изменять содержимое выбранного графического изображения, если на компьютере имеется редактор графических изображений. Выберите в меню «Формат»

пункт «Редактировать изображение...». Выберите в системе Windows редактор изображений, связанный с файлами ВМР, и загрузите изображение. Можно также выбрать пункт «Выбрать программу...» и указать другую программу. Эта программа будет использоваться по умолчанию вместо текущего редактора изображений. Отредактируйте изображение и закройте редактор изображений, чтобы вернуть изображение в текстовый редактор. Не изменяйте размеры, разрешение или тип изображения. Это может помешать возврату изображения в текстовый редактор. Изображения можно редактировать перед распознаванием с помощью инструментов оптимизации изображения.

таблицы

Таблицы отображаются в текстовом редакторе с сетками. Переместите указатель внутрь таблицы. Вид указателя изменится. Теперь линии сетки можно изменять. При изменении таблицы можно также пользоваться линейками текстового редактора. Изменить положение текста в ячейке таблицы можно, используя кнопки выравнивания на панели инструментов «Форматирование» или маркеры табуляции на линейке.

Гиперссылки

Адреса веб-страниц и электронной почты могут распознаваться и размещаться в виде ссылок в распознанном тексте. Для редактирования существующих ссылок и создания новых выберите в меню «Формат» пункт «Гиперссылка...».

Редактирование в режиме «Фактическая страница»

Элементы страницы содержатся в текстовых блоках, блоках изображений и таблиц. Обычно блоки соответствуют текстовым, табличным и графическим областям на изображении. Щелкните внутри элемента для отображения его границ. Они окрашены так же, как границы соответствующих областей. В разделе справки Фактическая страница можно найти подробное описание операций, приведенных здесь.

Рамки имеют серые границы и включают один или несколько блоков. Рамка устанавливается при обнаружении видимых границ в исходном изображении. Форматирование рамок и границ таблиц, а также заливка осуществляются с помощью контекстного меню или с помощью пункта «Таблица...» в меню «Формат». Заливку текстового блока можно выбрать в контекстном меню.

Области с несколькими колонками имеют оранжевые границы и содержат один или несколько блоков. Они определяются автоматически и показывают текст, который будет считаться поточным в колонке при экспорте с уровнем форматирования «Поточная страница».

- При этом можно отобразить и изменить порядок обхода. Выберите на панели инструментов «Форматирование» пункт «Показать/скрыть порядок обхода», чтобы обозначить порядок обхода стрелками. Чтобы скрыть стрелки, нажмите эту кнопку еще раз.
- Выберите инструмент «Изменить порядок обхода», чтобы отобразить набор кнопок изменения порядка обхода вместо панели инструментов форматирования. Измененный порядок применяется на уровне форматирования «Простой текст» и «Форматированный текст». Новый порядок изменяет последовательность перемещений курсора по странице при экспорте в режиме True Page.

Мгновенные изменения

Эта функция позволяет изменять распознанные страницы, редактируя разметку областей, без повторной обработки всей страницы. Если включено мгновенное изменение областей, изменения (удаление, создание, изменение размеров или типа) сразу же применяются к распознанной странице. И наоборот, изменения элементов в текстовом редакторе на уровне форматирования «Фактическая страница» автоматически вызывают изменения областей на странице.

При мгновенной обработке используются два связанных инструмента на панели инструментов «Изображение». Если не происходит распознавания, один из этих инструментов всегда активен.

- Нажмите эту кнопку, чтобы включить функцию мгновенного редактирования. Красный кружок означает, что сохраненных изменений в областях нет.
- Нажмите эту кнопку, чтобы отключить функцию мгновенного редактирования. Изменения областей сохраняются; на инструменте мгновенного редактирования отображается зеленый сигнал, который показывает, что имеются сохраненные изменения. Изменения можно применить несколькими способами.
- Щелкните инструмент мгновенного редактирования с зеленым сигналом. В текстовом редакторе появятся изменения, соответствующие сохраненным изменениям в областях.



Нажмите кнопку «Распознавание» для повторного распознавания всей страницы, включая измененные области.

Для получения подробных сведений о том, как обрабатываются изменения при мгновенной разметке областей и их влияние на виды текстового редактора см. раздел *Мгновенная обработка* в справке.

Пометка и редактирование





Панель инструментов «Пометка текста» содержит инструменты, позволяющие выполнять пометку (выделение или зачеркивание) и редактирование текста. Чтобы отобразить эту панель инструментов, используйте меню «Вид». Эту группу инструментов можно закреплять или делать плавающей. Для каждого инструмента имеется эквивалентный пункт в меню «Форматирование» или контекстном меню текстового редактора.

Под редактированием понимается закраска черным цветом конфиденциальной информации. Текст становится нечитаемым и недоступным для поиска. Чтобы пометить и отредактировать текст вручную, выберите инструмент «Пометить для редактирования» и с помощью курсора выберите все части текста, которые необходимо редактировать. Эти части выделяются серым цветом. После завершения нажмите кнопку «Редактировать документ». Выберите этот параметр для редактирования копии (более безопасный вариант) или оригинала документа. Если выбрано редактирование копии, то в программе OmniPage останутся открытыми и копия, и оригинал, готовые для сохранения.

ПРЕДУПРЕЖДЕНИЕ. В случае редактирования оригинала документа получение информации, закрашенной черным цветом, будет невозможным.

Чтобы найти и отредактировать текст с помощью функции поиска, в меню «Правка» выберите пункт «Найти и пометить текст», чтобы отобразить диалоговое окно «Найти и заменить». Выполните поиск текста, который необходимо пометить для редактирования. Просмотрите все найденные варианты и определите, какие из них следует отредактировать немедленно, а какие следует пометить для последующего редактирования. В последнем случае выполните редактирование, выбрав «Закрыть и редактировать документ» в диалоговом окне «Пометка текста» или позже нажмите кнопку «Редактировать документ».

Можно также выделять и вычеркивать выбранный текст или использовать при этом функцию поиска.

Озвучивание текста

Голосовая функция Nuance RealSpeak[®] предназначена для людей с ослабленным зрением, однако ее также можно использовать во время проверки и верификации.

Можно управлять озвучиванием, перемещая курсор в текстовом редакторе с помощью мыши или клавиатуры.

Прослушивание текста	Клавиши
Посимвольно, вперед или назад	Стрелки вправо или влево. При этом произносятся названия букв, цифр и знаков препинания.
Текущее слово	Ctrl + 1 на цифровой клавиатуре
Одно слово справа от курсора	Ctrl + стрелка вправо
Одно слово слева от курсора	Ctrl + стрелка влево
Одна строка	Установите курсор на нужную строку
Следующая строка	Стрелка вниз
Предыдущая строка	Стрелка вверх
Текущее предложение	Ctrl + 2 на цифровой клавиатуре
От курсора до конца предложения	Ctrl + 6 на цифровой клавиатуре
С начала предложения до курсора	Ctrl + 4 на цифровой клавиатуре
Активная страница	Ctrl + 3 на цифровой клавиатуре
С начала страницы до курсора	Ctrl + Home
От курсора до конца текущей страницы	Ctrl + End
Предыдущая, следующая или какая-либо страница	Ctrl + PgUp, PgDown или клавиши навигации
Введенные символы	Каждый введенный символ произносится отдельно.

Функция преобразования текста в речь включается и отключается с помощью команды «Речевой режим» в меню «Сервис» или клавиши F10. Второй пункт меню «Настройки голоса...» позволяет выбрать голос (например, мужской или женский), скорость чтения и громкость. Проверьте, правильно ли выбран язык прослушивания текста. Кроме того, можно пользоваться следующими сочетаниями клавиш:

Действие	Сочетание клавиш
Пауза/продолжение	Ctrl + 5 на цифровой клавиатуре
Увеличение скорости	Ctrl + «+» на цифровой клавиатуре
Уменьшение скорости	Ctrl + «-» на цифровой клавиатуре
Восстановление скорости	Ctrl + «*» на цифровой клавиатуре

При полной установке с программой OmniPage устанавливаются все системы преобразования текста в речь. При выборочной установке можно установить компоненты только для необходимых языков.

Создание и редактирование форм



В программу OmniPage Professional можно загружать бумажные или статические электронные формы (распространяемые в основном в офисной **PROFESSIONAL** среде в виде файлов PDF), распознавать их редактировать содержимое,

макет или и то, и другое на уровне форматирования «Фактическая страница». Выделите области формы в соответствующих областях изображения перед распознаванием или выберите макет распознавания «Форма», а затем воспользуйтесь панелями инструментов «Создание формы» и «Расположение формы». С их помощью можно внести изменения, создать заполняемую форму и сохранить ее в форматах PDF, RTF или XSN (формат Microsoft Office InfoPath 2003). Статические формы можно сохранять в формате HTML. В OmniPage Professional используется технология Logical Form RecognitionTM, позволяющая создавать заполняемые формы из статических.

Программа OmniPage поддерживает создание и редактирование форм, однако заполнение форм в программе OmniPage не предусмотрено.

Панель инструментов «Создание формы»

Это плавающая панель, отображаемая в текстовом редакторе. Она позволяет создать большое количество элементов формы с помощью следующих инструментов:

- Выделение: выберите этот инструмент для выделения, перемещения или изменения R размеров элементов формы.
- Текст: этот инструмент используется для добавления в форму фиксированных Α текстовых описаний, например названий, колонтитулов и меток.
Линия: инструмент «Линия» используется в основном при разработке макетов. Выберите этот инструмент и проведите линии для выделения разделов в форме. Прямоугольник: используется для создания прямоугольников в формах. Изображение: этот инструмент используется для указания, какие области формы считать изображениями. Поле ввода: используется для создания заполняемых текстовых полей, то есть la полей, предназначенных для ввода текста. Составное поле: с помощью этого инструмента можно создать текстовое поле, 10001 которое состоит из блоков. Обычно используется для ввода информации, например почтовых индексов. Флажок: используется для создания флажков. Применение: вопросы типа «да-нет» R или выбор одного или нескольких пунктов. Текст в кружочке: по своему назначению напоминает флажок (выше). Позволяет создать элемент, который при выборе обводится кружочком. Таблица: этот инструмент предназначен для создания таблиц в форме.

Элементы форм можно создать, щелкнув правой кнопкой мыши существующий распознанный элемент формы и выбрав пункт меню «Вставить объект формы».

Панель инструментов «Расположение формы»

С помощью инструментов этой панели можно выравнивать элементы формы и выбирать, какой из них должен отображаться над другими, если элементы перекрываются. Последняя функция полезна, когда, например, нужно создать фоновую графику для формы.



Чтобы задать последовательность для перекрывающихся элементов, используйте кнопки На передний план и На задний план.

Для выравнивания правой и левой, верхней и нижней сторон или центральных частей выбранных элементов форм применяются параметры:



по горизонтали — использование инструментов для выравнивания горизонтали

по вертикали — использование инструментов для выравнивания по вертикали

Команды, представленные на панели инструментов «Расположение формы», доступны также в контекстном меню каждого элемента формы.

Редактирование свойств объектов формы

Выберите объект формы для редактирования напрямую и щелкните его правой кнопкой мыши, чтобы открыть контекстное меню. С его помощью можно редактировать вид или свойства любого элемента формы. Воспользуйтесь следующими командами.

Вид объектов формы — использование вкладок «Границы», «Заливка» и «Тень» для определения внешнего вида элементов формы аналогично тому, как это делается в приложении текстового редактора.

Свойства объектов формы — эта команда предоставляет доступ к свойствам элемента, таким как размер, положение и имя. Свойства динамически меняются в зависимости от выбранного типа элемента.

Извлечение данных формы



Извлечение данных формы (FDE) — это этап рабочего процесса. Данные извлекаются из таких элементов, как заполняемые поля, флажки и кнопки выбора. FDE — это упрощенная реализация технологии Logical Form Recognition.

Чтобы создать рабочий процесс, содержащий извлечение данных формы, выполните следующие действия.

- Определите обработку входных данных и их параметры. Типы входных данных включают: изображения в формате PDF, формы в формате PDF, а также файлы изображений форм, отсканированных с бумаги.
- Выберите «Извлечь данные формы вместо распознавания» и укажите параметры. Сюда также входит выбор языка. Параметр *Автоматически определять один язык* может быть полезен для автоматической обработки форм, когда язык, используемый для заполнения каждой из форм, не может быть определен заранее. См. раздел «Языки».
- Задайте активную форму PDF в качестве шаблона. Это может быть одностраничная или многостраничная форма, заполняемая или незаполняемая. Программа определяет местоположение и тип полей формы в зависимости от данного шаблона формы.
- Завершите рабочий процесс этапом сохранения.

Программа OmniPage извлечет данные из входных форм, используя указанный шаблон. Для экспорта будет выбран текстовый файл со значениями, разделенными запятыми (.csv), который можно загружать в электронную таблицу.

После выбора извлечения данных формы в рабочем процессе остаются только этапы сохранения.

Сохранение и экспорт

После получения хотя бы одного изображения для документа эти изображения можно экспортировать в файл. После распознавания хотя бы одной страницы результаты распознавания можно экспортировать. После дальнейшего распознавания одну страницу, выбранные страницы или весь документ можно сохранить файл, скопировать в буфер обмена или отправить по электронной почте. Кроме того, файл всегда можно сохранить в формате документа OmniPage. Программа OmniPage полностью поддерживает приложения и форматы Office 2007 и 2010.

Документ сохраняется в программе OmniPage после экспорта. Это позволяет неоднократно сохранять, копировать или пересылать документы, изменяя уровни форматирования, типы файлов, имена или места размещения. Кроме того, можно добавлять или повторно распознавать страницы, а также изменять распознанный текст.

При автоматической обработке и в диспетчере пакетной обработки место сохранения указывается до начала обработки.

Рабочий процесс может включать один или несколько этапов сохранения, в том числе сохранение для различных целей (например, в файл и для отправки по почте). Задание в диспетчере пакетной обработки должно включать хотя бы одно сохранение. См. Глава 6, стр. 92, раздел «Рабочие процессы».

Сохранение и экспорт

Чтобы иметь возможность работать с документом во время следующего сеанса, документ следует сохранить в формате OmniPage. Это специальный тип конечного файла. В нем сохраняются одновременно исходные изображения, результаты распознавания, параметры и эталоны.

Экспорт осуществляется при нажатии кнопки 3 на панели инструментов OmniPage. При этом открывается список возможных назначений экспорта. Некоторые отображаются только при определении на компьютере доступа к конечной папке. Чтобы начать экспорт, выберите нужный вариант и нажмите кнопку «Экспорт результатов». Экспорт также можно выполнить с помощью меню «Обработка».

Сохранение исходных изображений

Оригиналы изображений можно сохранять на диск в виде файлов различных типов, как с оптимизацией изображения (с помощью инструментов оптимизации изображений), так и без него.

1. В раскрывающемся списке «Экспорт результатов» выберите пункт «Сохранить в файлы». В отобразившемся диалоговом окне в поле «Сохранить как» выберите пункт «Изображение».



- 2. Выберите папку назначения и тип файла. Введите имя файла.
- 3. Выберите, что следует сохранить: только изображения выделенных областей, изображение текущей страницы или все изображения. Несколько страниц или областей можно сохранить в отдельном многостраничном файле изображения, выбрав в качестве типа файла формат TIFF, MAX, DCX, JB2, PDF или XPS (только изображение). В противном случае каждое изображение помещается в отдельный файл. Программа OmniPage присваивает файлам уникальные имена, добавляя к имени файла порядковые индексы.
- Нажмите кнопку «Параметры», если требуется указать режим сохранения (черно-белый, оттенки серого, цветной или «как есть»), максимальное разрешение и другие параметры. Для файлов TIFF укажите метод сжатия.
- 5. Для сохранения изображения с указанными настройками нажмите кнопку «ОК». Области и распознанный текст не сохраняются вместе с файлом.

Сохранение результатов распознавания

Распознанные страницы можно сохранять на диске в файлах различных типов.

- 1. В меню «Файл» выберите раздел «Экспорт результатов» или нажмите кнопку «Экспорт результатов» на панели инструментов OmniPage, выбрав в раскрывающемся списке пункт «Сохранить в файлы».
- 2. Откроется окно «Сохранить в файлы». В поле «Сохранить как» выберите «текст».
- Выберите папку и тип файла для сохранения документа. Выберите диапазон страниц, параметры файла, параметры присвоения имени и уровень форматирования документа. См. раздел «Выбор уровня форматирования» ниже.

- 4. Введите имя файла. Если нужно задать параметры экспорта, нажмите кнопку «Параметры». См. раздел «Выбор параметров конвертера» далее в этой главе.
- 5. Нажмите кнопку «ОК». После этого документ будет сохранен на диске. Если выбран параметр *Просмотреть результат*, экспортированный файл откроется в конечном приложении, т. е. в приложении, которое связано с выбранным типом файлов в системе Windows или в расширенных параметрах сохранения для конвертера выбранного типа файлов.

Выбор уровня форматирования

Уровень форматирования экспортируемого файла выбирается во время экспорта в соответствующем диалоговом окне («Сохранить в файлы», «Копировать в буфер обмена», «Отправить по почте» и др.). Три уровня соответствуют форматированию в одноименных режимах просмотра текста в текстовом редакторе. Однако уровень форматирования файла при сохранении не зависит от формата изображения в текстовом редакторе. При экспорте в файл или электронную почту укажите в первую очередь тип файла. Он определяет возможные уровни форматирования.

Существуют следующие уровни форматирования.



Простой текст

Экспортируется текст, выровненный по левому краю, без разбиения на колонки, с применением одного типа и размера шрифта. При экспорте в файлы типов «Текст» или «Юникод» графика и таблицы не поддерживаются. Простой текст можно экспортировать почти во все типы файлов и конечные приложения с сохранением графики, таблиц и списков.



Форматированный текст

Экспортируется текст без разбиения на колонки с сохранением стиля абзацев и шрифтов. Может содержать графику и таблицы. Этот режим может использоваться почти для всех типов файлов.



Поточная страница

Сохраняется исходный макет страниц, включая колонки. Применяется для колонок и отступов, но не для текстовых блоков или рамок. Текст может перетекать из одной колонки в другую, что невозможно при использовании текстовых блоков.



Фактическая страница

Сохраняется исходный макет страниц, включая колонки. Применяется для текстовых, графических и табличных блоков, а также рамок. Экспортируется только в приложение, поддерживающее обработку блоков и рамок. Форматирование в режиме True Page — это единственный вариант, подходящий для экспорта в XML- и PDF-файлы (кроме типа «PDF, отредактированный»).



Электронная таблица

Экспортируются результаты распознавания в табличной форме. Применяется в электронных таблицах. В этом формате каждая страница документа располагается на отдельном листе.

Формат «Электронная таблица» удобен для сохранения и экспорта в Microsoft Excel таблиц, занимающих страницу целиком. Для таблиц меньшего размера предпочтительнее формат «Форматированный текст»: каждая таблица будет размещена на отдельном листе, а части документа, не входящие в таблицы, – на листе указателя со ссылками на соответствующие листы таблиц.

Выбор параметров конвертера

Для более точной настройки экспорта нажмите кнопку «Параметры» в диалоговом окне сохранения. Откроется окно с именем конвертера, связанного с типом выбранного файла. В нем представлен набор параметров, применимый к данному типу файла. В первую очередь подтвердите или измените уровень форматирования, поскольку он влияет на список остальных параметров. Выберите нужные параметры. Подробные инструкции см. в справке.

Чтобы изменения применялись ко всем будущим операциям экспорта с помощью данного конвертера, установите флажок *Сделать изменения постоянными*. Если не установить его, изменения будут применены только к текущему экспорту и не сохранятся для дальнейшего использования. Параметры экспорта можно изменить и сохранить без сохранения документа. Для этого выберите команду «Сервис» > «Сохранение параметров».

Использование множественных конвертеров

Множественные конвертеры позволяют выполнять экспорт сразу в файлы нескольких типов. В диалоговом окне сохранения выберите параметр «Различные форматы».

<u>т</u>екст <u>и</u>зображение <u>ора</u>зличные форматы

Чтобы создать собственный множественный конвертер, откройте диалоговое окно «Сохранение параметров» через меню «Сервис». Выберите параметр «Множественные конвертеры». Выберите конвертер и нажмите кнопку «Создать из...» . При этом будет создана копия выбранного конвертера, которую можно смело изменять, сохраняя оригинальный.

В списке появится новый конвертер. Выберите его и нажмите кнопку «Параметры» для задания параметров. Откроется список всех конвертеров текстовых файлов, а затем конвертеров файлов изображений. Отметьте нужные конвертеры. Дополнительно укажите путь к папке для каждого типа файла.

Можно сохранять страницы с разными параметрами или уровнями форматирования в файлы разных типов. Сохранение будет произведено в соответствии с настройками простых конвертеров для этих типов файлов. Некоторые операции сохранения с применением множественных конвертеров невозможны. К ним относятся следующие.

Сохранение документов OmniPage

Рабочие процессы OmniPage нельзя сохранить с помощью множественных конвертеров. Используйте меню «Файл» или рабочий процесс с этапом «Сохранить в документ OmniPage».

Сохранение в разные объекты

К примеру, нельзя использовать множественный конвертер для одновременного сохранения документа в файл и отправки его по электронной почте. Используйте рабочий процесс, включающий два этапа сохранения, или выполните операцию сохранения дважды.

Сохранение разных диапазонов страниц

Невозможно сохранить разные диапазоны страниц в разные файлы, поскольку во время сохранения может существовать только один набор выбранных страниц. По той же причине нельзя использовать и один рабочий процесс. Выполните операцию сохранения дважды или примените два рабочих процесса.

Сохранение в формате PDF

Для сохранения документов в формате PDF (Portable Document Format) можно выбрать один из пяти вариантов. Первые четыре варианта представлены в списке конвертеров текстовых файлов, а пятый относится к конвертерам файлов изображений.

PDF (обычный)

Страницы экспортируются в том виде, как они отображаются в текстовом редакторе в режиме просмотра True Page. Такой PDF-файл можно просматривать, редактировать в PDF-редакторе и осуществлять в нем поиск.

PDF, отредактированный

Если после распознавания файла произведены большие изменения, выберите этот вариант. При этом можно выбрать три варианта форматирования, включая режим True Page. Такой PDF-файл можно просматривать, редактировать и осуществлять в нем поиск.

Изображение PDF с возможностью поиска;

Такой PDF-файл можно только просматривать. Редактировать его нельзя. Изображения с оригинальной страницы экспортируются, однако они связаны с текстом, на котором расположены, и поэтому возможен поиск по тексту. Найденное слово выделяется на изображении цветом.

PDF с подставленными изображениями

Аналогичен обычному формату PDF, однако для слов, содержащих нераспознанные и сомнительные символы, используются наложенные изображения, так что эти сомнительные слова отображаются так же, как в исходном документе. Такой PDF-файл можно просматривать, редактировать и осуществлять в нем поиск.

Изображение в формате PDF

Исходные изображения экспортируются. Такой PDF-файл можно только просматривать. Редактирование и поиск по тексту недоступны.

Кроме указанных выше разновидностей можно использовать другие параметры при определении выходного PDF-файла с помощью команды «Параметры»:

PDF 1.6 или 1.7

Сохранение в файл PDF версии 1.6 или 1.7 для обеспечения повышенной безопасности, возможности использования функций пометки и встраивания вложений.

PDF/A

Выберите этот параметр для создания файлов, совместимых с PDF-A, чтобы они отображались строго одинаково независимо от среды компьютера и оставались читаемыми даже после многих лет технологической эволюции.

Размеченный файл PDF

Создание размеченного файла PDF для сохранения его структуры. При этом обеспечивается логическая последовательность чтения, правильная структура таблиц и другое.

PDF MRC

Используйте эту технологию сильного сжатия для получения *хорошего качества* при небольшом размере файла. Этот параметр доступен для цветных изображений и изображений в оттенках серого в формате, а также для изображений PDF с поддержкой поиска.

Линеаризованный PDF

Выберите этот вариант для создания PDF-файлов, оптимизированных для быстрой загрузки и отображения на веб-сайтах.

Защита паролем

В OmniPage Professional для PDF-файлов можно установить вид и уровень шифрования, а затем — пароль на открытие и (или) пароль на изменение разрешений.

При сохранении в XPS-файлы список вариантов более ограничен.

Преобразование из формата PDF

Чтобы извлечь текстовое содержимое из файла PDF, загрузите его в программу OmniPage, выполните распознавание и сохраните результаты в текстовом формате.

В контекстном меню файла PDF доступны различные варианты вывода: Word, Excel, RTF, WordPerfect или текстовый. Для получения доступа к дополнительным вариантам используйте мастер преобразования.

Помощник eDiscovery для PDF с возможностью поиска

Этот помощник вызывается через меню «Сервис» или контекстное меню PDF-файла в проводнике Windows. Этот помощник разработан специально для создания PDF-файлов с возможностью поиска из PDF-файлов, содержащих только изображения либо некоторые текстовые элементы или страницы, без изменения существующего текста или его распознавания. Другими словами, помощник обрабатывает только изображения в PDF-файлах. Все текстовые элементы (включая метаданные документа, примечания, разметку для редактирования, штампы и т. д.) в PDF-файлах остаются без изменений. Процесс может выполняться автоматически или взаимодействовать с пользователем при разметке областей или проверке орфографии. Помощник загружает файлы, которые были выбраны в файловой системе, и возвращает результаты в исходную папку. При этом можно выбрать, будут ли оригинальные файлы перезаписываться или сохраняться в качестве резервных копий. Разметка областей и проверка орфографии выполняется в отдельных окнах без взаимодействия с документами, открытыми в OmniPage.

Создание PDF-файлов из других приложений

Компонент Nuance PDF Create, поставляемый вместе с OmniPage Professional, позволяет создавать обычные PDF-файлы из документов в любом приложении с поддержкой печати. Выберите команду «Файл» > «Печать» и выберите принтер **ScanSoft PDF Create!**. Настройте нужные свойства, нажмите кнопку «ОК» и укажите имя и местоположение файла. Если включить параметр *Просмотреть выходной PDF-файл*, результат будет показан в используемой по умолчанию программе просмотра PDF-файлов.

Отправка страниц по электронной почте

Можно отправлять изображения страниц или распознанные страницы в виде одного или нескольких вложений в сообщения электронной почты, если на компьютере установлено MAPI-совместимое почтовое приложение, например Microsoft Outlook. Чтобы отправить страницы по

электронной почте, выполните указанные ниже действия.

При автоматической обработке выберите параметр «Отправить по почте» в раскрывающемся списке «Экспорт результатов» на панели инструментов OmniPage. Сразу после распознавания и проверки последней страницы документа появится диалоговое окно «Параметры экспорта». После настройки параметров экспорта появится пустое сообщение электронной почты с вложенными файлами. Добавьте получателей и текст сообщения.

- При обработке вручную выберите параметр «Отправить по почте» в раскрывающемся списке «Экспорт результатов» и нажмите соответствующую кнопку. Откроется диалоговое окно «Параметры экспорта», а затем — почтовое сообщение с вложенными файлами.
- Рабочие процессы и задания могут включать этап отправки по почте, но для него необходимо указать получателей и ввести текст сообщения как параметры рабочего процесса, чтобы рабочий процесс мог быть выполнен без участия пользователя.

Отправка в Kindle

Kindle — это устройство для чтения электронных книг от Amazon. Помощник Kindle в меню «Сервис» позволяет создать простой рабочий процесс, который будет отправлять результаты распознавания в учетную запись Kindle на сервере Amazon. Затем эти результаты оптимизируются для экрана устройства и передаются в Kindle, для которого зарегистрирована эта учетная запись.

Подготовка рабочего процесса Kindle.

- 1. Держите под рукой устройство Kindle и связанный с ним адрес электронной почты.
- 2. В меню «Сервис» выберите пункт «Помощник Kindle».
- 3. Введите имя нового рабочего процесса.
- 4. Укажите источник документа: «Сканировать», «Загрузить файлы» или «Загрузить файлы с цифровой камеры». Если нужно использовать файлы, при запуске рабочего процесса будет показан запрос на выбор файлов.
- 5. Введите адрес электронной почты, который привязан к устройству Kindle.
- 6. Введите имя выходного файла. Все результаты распознавания будут сохранены в один файл.
- 7. Чтобы сохранить рабочий процесс для использования в будущем, нажмите кнопку *Сохранить*. Кроме того, можно нажать кнопку *Сохранить и запустить*, чтобы немедленно запустить рабочий процесс и передать его результаты на устройство Kindle.

Этот простой рабочий процесс состоит из трех этапов: получение изображений, распознавание текста и отправка его в Kindle. Распознавание выполняется только на английском языке. Все остальные параметры имеют значения по умолчанию или значения, оптимизированные для Kindle.

При пером запуске помощника Kindle создается настраиваемый формат вывода под названием «Документ Kindle». Он преобразует цветные элементы в оттенки серого, понижает разрешение изображений до 72 точек на дюйм и устанавливает уровень форматирования «Форматированный текст», чтобы удалить все колонки. Этот формат затем доступен для дальнейшей обработки с помощью рабочих процессов и без них. Рабочий процесс Kindle можно изменить в помощнике рабочего процесса, например, для добавления других этапов или изменения параметров. Можно указать диапазон страниц или добавить дополнительные этапы сохранения, чтобы файл не только отправлялся в Kindle, но и сохранялся с использованием различных параметров (например, с уровнем форматирования «Поточная страница» и сохранением цвета). Не рекомендуется вносить изменения, которые не подходят для Kindle, например создавать несколько выходных файлов, использовать неподдерживаемые языки и т. п.

В помощнике рабочего процесса можно также настроить различные варианты вывода рабочего процесса. Для этого выберите этап «Отправить по электронной почте», укажите формат вывода «Документ Kindle» и введите адрес электронной почты. То же самое можно сделать без использования рабочего процесса, выбрав команду «Отправить по почте» в раскрывающемся списке «Экспорт результатов».

Следует учесть, что на данный момент (май 2011 г.) эта услуга Kindle от Amazon доступна только в США. Поэтому помощник Kindle отображается только, если для интерфейса программы выбран английский язык.

Другие варианты экспорта

С помощью модуля Nuance RealSpeak можно преобразовать распознанный текст в звуковой WAV-файл для дальнейшего прослушивания В режиме множественных конвертеров можно одновременно сохранить документ в файл и создать звуковой файл. Для звукового файла в настройках конвертера необходимо указать язык чтения.



OmniPage 18 поставляется с компонентом Nuance Cloud Connector, который можно легко настроить, выбрав его в группе OmniPage в меню «Пуск» системы Windows. Укажите сайты облачных хранилищ и FTP-сайты, на которых будут сохраняться файлы. После установки хотя бы одной связи

компонент Nuance Cloud Connector становится доступен в раскрывающемся списке «Экспорт результатов».

В этом списке также можно найти прямые ссылки на два веб-хранилища, которые недоступны через Nuance Cloud Connector: Evernote и Dropbox. У некоторых служб облачных хранилищ могут быть свои ограничения. Например отправлять файлы изображений могут только пользователи службы Google Apps Premier.



В OmniPage Professional можно выбрать другие варианты для экспорта. Файлы можно сохранять в системе SharePoint 2003, 2007 или 2010, Hummingbird (Open Text) и iManage (Interwoven). Выбор параметров экспорта осуществляется в диалоговом окне «Параметры экспорта».

При нажатии кнопки «ОК» может быть выведен запрос на ввод учетных данных или указания необходимого пути.

При работе с системой SharePoint сведения о сервере, имени пользователя и пароле нужно ввести только один раз в течение сеанса. После этого они сохраняются и подставляются автоматически.

При поддержке интерфейса ODMA можно пользоваться системой управления документами (DMS). Если возможен доступ к нескольким таким системам, используется заданная по умолчанию. Чтобы сервер ODMA принимал типы файлов, экспортируемые из OmniPage Professional, его необходимо предварительно настроить.

Дополнительные сведения об этих системах см. в справке.

Рабочие процессы

Рабочий процесс включает последовательность этапов обработки и соответствующие параметры. Ее можно сохранить, чтобы применить впоследствии при решении аналогичной задачи. Обычно рабочий процесс начинается с этапа сканирования или загрузки, но может начинаться, когда документ OmniPage уже открыт. Рабочий процесс не обязательно должен соответствовать обычному шаблону обработки 1-2-3. Обычно рабочий процесс включает в себя этап распознавания (но не обязательно). Можно, например, сохранить изображения страницы в файлы изображений различных типов или в документ OmniPage. Независимо от того, используете ли вы распознавание или нет, количество этапов сохранения, каждый из которых имеет свои параметры экспорта, может быть любым. Кроме того, файлы назначения для каждого из этапов могут различаться.

Рабочие процессы предназначены для эффективной обработки документов целиком. Они также позволяют выполнять распознавание или сохранение одной или выбранных страниц документа.

Некоторые рабочие процессы выполняются без участия пользователя. Рабочие процессы, требующие взаимодействия, – это процессы, включающие этап оптимизации изображения вручную, этап разметки областей вручную, этап проверки и редактирования, процессы, в которых отображается запрос имен входных и выходных файлов и путей, а также рабочие процессы сканирования, в которых отображаются запросы на загрузку дополнительных страниц.

Задания диспетчера пакетной обработки тесно связаны с рабочими процессами. Задания создаются в окне мастера заданий, который в процессе работы обращается к помощнику рабочего процесса. Задания запускают рабочие процессы в соответствии с указанными параметрами (обычно инструкции по времени). Наиболее типичным является их выполнение без вмешательства пользователя.

Чтобы увидеть этапы и параметры рабочего промесса, нажмите кнопку

«Помощник рабочего процесса» 🔁 на стандартной панели инструментов.

Выполнение рабочих процессов

Выполнение стандартного или вновь созданного рабочего процесса:

- 1. Если предполагается сканирование, поместите документ в устройство автоматической подачи или его первую страницу на планшет сканера.
- 2. Выберите нужный рабочий процесс в раскрывающемся списке «Рабочий процесс».
- Нажмите кнопку «Начать». На панели инструментов OmniPage выводятся этапы рабочих процессов, она играет роль индикатора выполнения процесса. На панели «Состояние рабочего процесса» более подробно отображается ход выполнения процессов. Чтобы остановить выполнение рабочего процесса, нажмите кнопку «Прервать».
- 4. Если указаны выбор входных данных во время выполнения, то программа ожидает выбора файлов в диалоговом окне «Загрузить файлы».
- 5. Если запрошенный этап требует вмешательства пользователя (оптимизация изображения, разметка областей вручную или проверка), программа отобразит соответствующие страницы.
- Во время оптимизации страницы, разметки областей или проверки нажмите кнопку «Страница готова» на панели инструментов или в соответствующем диалоговом окне, чтобы перейти к следующей странице.
- 7. Во время оптимизации или проверки последней страницы, разметки на ней областей, а также если больше не требуется выполнять разметку областей или проверку, нажмите соответствующую кнопку «Документ готов» на панели задач. Все страницы с неразмеченными областями будут размечены автоматически.
- 8. В подменю «По окончании» меню «Обработка»/«Рабочие процессы» можно задать один из трех вариантов действий по окончании выполнения рабочего процесса: закрыть документ, закрыть программу OmniPage или выключить компьютер. Эти параметры обычно применяются в случае выполнения рабочего процесса без участия пользователя. Если используется подобный рабочий процесс, не забудьте добавить этап сохранения.



>

Рабочие процессы можно также запускать с помощью значка агента OmniPage на панели задач Windows. Щелкните правой кнопкой мыши значок агента. Откроется контекстное меню со списком рабочих процессов. Выберите и запустите нужный процесс. При необходимости запустится программа OmniPage. Если процесс выполняется с загруженным документом, отобразится диалоговое окно «Начало рабочего процесса», в котором показано, где можно выбрать то, что должно обрабатываться из текущего документа: только страницы, определенные в рабочем процессе, все страницы, выбранные страницы или текущую страницу.

Если значок агента OmniPage не отображается, включите его на панели «Общие» диалогового окна «Параметры» или выберите «Пуск» > «Все программы» > Nuance OmniPage 18 > «Агент OmniPage».

Некоторые рабочие процессы можно запускать с рабочего стола, из проводника Windows или из простого загрузчика. Щелкните правой кнопкой мыши значок файла изображения или имя файла. Откроется контекстное меню. Можно выбрать несколько файлов одновременно. Выберите пункт OmniPage 18 и имя рабочего процесса в подменю. Данное подменю также предоставляет быстрый доступ к шести конечным форматам с использованием настроек по умолчанию: Word, Excel, PDF, RTF, TXT и WordPerfect. Чтобы настроить список отображаемых в этом подменю рабочих процессов, выберите пункт меню «Добавление и удаление рабочих процессов». Здесь отображаются только рабочие процессы, во время выполнения которых отображается запрос на входные данные.

При нажатии кнопки «Прекратить» во время выполнения рабочего процесса процесс приостанавливается. Для продолжения обработки нажмите кнопку «Начать». Если приостановить рабочий процесс (возможно, чтобы выполнить некоторую обработку вручную), а затем сохранить документ как OmniPage, то при последующем открытии этого документа OmniPage выполнение прерванного рабочего процесса будет возобновлено.

Помощник рабочего процесса

Позволяет создавать и изменять рабочие процессы. Кроме того, помощник используется мастером заданий для создания или изменения рабочего процесса. Дополнительные сведения см. в следующем разделе. Помощник предлагает один или несколько этапов, каждый из которых содержит раскрывающийся список. На левой

панели диалогового окна «Помощник рабочего процесса» можно создать собственный рабочий процесс.



В любой момент во время выполнения этого процесса в раскрывающемся меню помощника отображаются все этапы, которые являются логически возможными в данный момент.

В OmniPage Professional имеются дополнительные этапы: «Извлечь данные формы» и «Пометить текст».

Создание рабочих процессов

Выберите пункт «Создать рабочий процесс...» в раскрывающемся списке «Рабочий процесс» или в меню «Обработка». Можно также щелкнуть значок помощника рабочего процесса на стандартной панели инструментов, если рабочий процесс не выбран. В открывающейся панели помощника предлагаются две начальные точки.

Выберите **Начать заново**, чтобы начать рабочий процесс без каких-либо этапов на диаграмме рабочего процесса справа. Примите или измените имя рабочего процесса по умолчанию. Затем нажмите кнопку «Далее» и выберите первый этап. Выберите этап загрузки изображения, в которым входные данные получаются из файла, со сканера или из файлов цифровой камеры. Укажите параметры справа. Затем продолжите создание рабочего процесса. Он может включать множество различных этапов. По завершении нажмите кнопку «Готово».

Выберите **Существующие рабочие процессы**, чтобы просмотреть список существующих рабочих процессов. Этот список содержит все стандартные и созданные пользователем рабочие процессы. Выберите один из них в качестве шаблона. Этапы выводятся на диаграмме рабочего процесса справа. Введите имя для нового рабочего процесса. Нажмите кнопку «Далее», чтобы продолжить, измените этапы и параметры рабочего процесса, как описано в следующем разделе. Измененные параметры применяются только к новому рабочему процессу, и не записываются в рабочий процесс, который использовался в качестве шаблона. Все изменения параметров применяются к созданному рабочему процессу, но не влияют на параметры программы. Наконец нажмите кнопку «Готово», чтобы завершить создание нового рабочего процесса.

Изменение рабочих процессов

В раскрывающемся списке «Рабочий процесс» выберите рабочий процесс, который требуется изменить, и нажмите кнопку «Помощник рабочего процесса» на стандартной панели инструментов. Кроме того, можно выбрать в меню «Сервис» пункт «Рабочие процессы...», указать нужный процесс и нажать кнопку «Изменить». Появится первая панель помощника рабочего процесса с загруженным рабочим процессом. На диаграмме рабочих процессов щелкните значок этапа, который требуется изменить. Щелкните стрелку вниз под значком, чтобы заменить этот этап на другой. Продолжите изменение этапов и/или параметров. Помните, что удаление или изменение этапа может привести к последующему удалению зависимых этапов. Нажмите кнопку «Далее», чтобы заменить удаленные этапы или добавить новые. Нажмите кнопку «Готово», чтобы подтвердить изменения рабочего процесса.

После создания или изменения рабочего процесса необходимо либо запустить рабочий процесс, либо выбрать пункт 1-2-3 в раскрывающемся списке «Рабочий процесс», чтобы вернуться к нормальной обработке.

Рабочий процесс для Kindle

Помощник Kindle в меню «Сервис» позволяет создать простой рабочий процесс, который будет получать входные файлы, распознавать их и отправлять результаты распознавания в соответствующем формате в учетную запись Kindle на сервере Amazon. Затем эти результаты передаются в Kindle, для которого зарегистрирована эта учетная запись. См. раздел «Отправка в Kindle» на стр. 84.

Диспетчер пакетной обработки



Диспетчер пакетной обработки — это отдельная интегрированная программа, которая позволяет создавать задания для немедленной обработки или обработки через некоторое время. Правильно выбирая этапы, можно настроить задание, которое будет выполняться без вмешательства

пользователя. Рабочие процессы выполняются в соответствии с указанными параметрами заданий. Задания создаются в окне мастера заданий.



В программе OmniPage Professional имеются следующие дополнительные функции диспетчера пакетной обработки:

- настройка расписания и периодичности выполнения задания;
- отслеживание входящих файлов изображений в указанных папках;
- отслеживание входящих вложений в почтовых ящиках (Outlook и Lotus Notes);
- отправка уведомлений о завершении задания по указанным адресам электронной почты;
- управление рабочими процессами с помощью штрихкодов.

Создание новых заданий

Откройте диспетчер пакетной обработки с помощью меню «Процесс» или в операционной системе, выбрав в меню команды «Пуск» > «Все программы» > Nuance OmniPage 18 > «Диспетчер пакетной обработки OmniPage» или в агенте OmniPage на панели задач.



Создание задания в основном заключается в определении времени выполнения рабочего процесса. Откройте окно диспетчера пакетной обработки, как описано

выше, и нажмите кнопку «Создать задание» или выберите в меню «Файл» пункт «Создать задание».

Откроется окно мастера заданий. В первую очередь определите тип задания. Можно создать пять различных типов заданий в двух основных категориях: обычные задания и задания на отслеживание.

Как обычные задания, так и задания на отслеживание могут содержать повторяющуюся последовательность действий. Задания на отслеживание связаны с передачей файлов, поступающих в определенные папки или почтовые ящики, на обработку в программу OmniPage. Отдельным типом заданий в данной категории являются задания с титульными страницами со штрихкодом, в которых для определения выполняемого рабочего процесса используются титульные страницы со штрихкодом.



Обычное задание: задайте время начала выполнения и выберите или создайте рабочий процесс для выполнения. Если выбран параметр «Не запускать сейчас», используйте для запуска кнопку «Включить» в диспетчере пакетной обработки.



Типы заданий, доступных только в программе OmniPage Professional



Задание титульной страницы со штрихкодом: это специальный тип задания на отслеживание папок (см. далее). В нем отслеживается появление входящих страниц со штрихкодами в папке, а затем выполняется обработка последующих входящих изображений с помощью рабочего процесса, определенного в штрихкоде. Для получения подробных сведений см. описание обработки штрихкодов далее в этой главе.



Задание контроля за папками: для этого типа заданий нужно указать папки для контроля входящих файлов изображений.



Задание контроля за почтовым ящиком Outlook: контроль поступления в почтовый ящик Outlook вложенных файлов изображений определенных типов.



Задание контроля за почтовым ящиком Lotus Notes: то же самое для почтового ящика Lotus Notes.

Введите имя задания и нажмите кнопку «Далее».

Откроется панель параметров запуска и остановки. Выберите время начала и завершения задания и укажите, следует ли удалять входящие файлы по завершению работы. Чтобы разрешить участие пользователя для задания, установите флажок *Автоматически выполнять задание*. Использование этого флажка позволяет выполнять задание двумя способами, не создавая отдельные задания. Если вы будете за компьютером во время выполнения задания, снимите этот флажок. Если задание должно быть выполнено без участия пользователя, установите флажок. При этом будет выполнено только автоматическая оптимизация изображения, ручная разметка областей будет заменена на автоматическую, а проверка будет пропущена. В этом случае необходимо заранее указать имена входных и сохраняемых файлов, а также выбрать соответствующие папки.

В OmniPage Professional можно установить периодичность и запросить уведомление по электронной почте после завершения задания.

Начиная со следующей панели, можно создавать задания (кроме заданий титульных страниц со штрихкодом) как обычно при использовании рабочих процессов. Задайте начальную точку (Начать заново или Существующие рабочие процессы) и выполните действия, описанные в разделе «Рабочие процессы».

Откройте диалоговое окно «Параметры» диспетчера пакетной обработки в меню «Сервис». На панели «Общие» имеется параметр *Включать агент OmniPage на панели задач при запуске системы*. По умолчанию он включен. Для выполнения заданий в указанное время параметр следует оставить включенным. Данная функция позволяет отказаться от выполнения всех заданий, чтобы не отключать их по отдельности. Состояние этого параметра также управляет выполнением заданий с титульными страницами со штрихкодом.

На панели «Общие» можно огранить количество допустимых страниц в выходном документе, даже если выбран параметр файла *Один файл для всех страниц*. По достижении максимального количества страниц создается новый файл, отличающийся порядковым индексом.

После создания задания нажмите кнопку «Готово».

Изменение заданий

1

Изменять можно только неактивные задания. Выберите задание на левой панели диспетчера пакетной обработки, а затем в меню «Правка» выберите команду «Изменить» или нажмите кнопку «Изменить задание». Сначала изменить

инструкции по времени выполнения, как необходимо. Затем отобразится помощник

рабочего процесса с загруженными этапами рабочего процесса и параметрами. Внесите изменения так же, как и для рабочих процессов. См. раздел «Изменение рабочих процессов» выше.

Выполнение заданий и управление ими

Эти функции выполняются с помощью диспетчера пакетной обработки. Он состоит из двух панелей. В левой панели перечислены все задания, для каждого указано время следующего запуска, состояние задания и его журнал. Возможные состояния

Ожидание:	время запуска задания еще не наступило.
Выполнение:	в настоящий момент идет обработка.
Слежение:	происходит отслеживание. Обработка не производится.
Неактивное:	задания, созданные с параметром <i>Не запускать сейчас</i> , или отключенные задания.
Просроченное:	время запуска задания уже прошло.
Сбор файлов:	производится слежение, задание ждет поступления файлов.
Приостановлено:	пользователь приостановил выполнение задания и не возобновил его.
Завершение:	выполняется сохранение результата для задания слежения.
Запуск:	состояние перехода к выполнению. Это состояние отображается в момент запуска задания или когда диспетчер пакетной обработки не может запустить все задания одновременно.

Щелкните задание для вывода в правой панели поэтапного анализа всех его страниц. В анализе указывается источник входных файлов, состояние страницы и расположение выходных файлов. Щелкните значок плюса для получения дополнительных сведений о странице. Щелкните на значке минуса, чтобы скрыть дополнительные сведения. Для заданий, находящихся в состоянии «ошибка» или «предупреждение», выводится перечень ошибок или информация о том, при обработке каких страниц произошел сбой.

Команда «Сделать задание активным» в меню «Файл» позволяет мгновенно сделать неактивное задание активным.



Команда «Сделать задание неактивным» в меню «Файл» позволяет сделать активное задание неактивным. Если это задание выполняется, перед переключением в неактивное состояние оно будет остановлено. Чтобы быстро закрыть задание на отслеживание папок и сохранить результат, сделайте его неактивным.

E			<u>D</u>	
Ľ		12		
E.	D		1.	

Команда «Остановить задание» в меню «Файл» позволяет остановить выполнение задания, находящегося в состоянии «Запуск», «Выполнение» или «Приостановлено».



Приостановить задание отображается для заданий, находящихся в состоянии «Выполнение» или «Запуск». Для изменения параметров запуска по времени у такого задания его необходимо сначала прервать.



Команда **Возобновить задание** позволяет продолжить выполнение задание в том состоянии, в котором оно было приостановлено.



a

Команда **«Удалить задание»** в меню «Правка» позволяет удалить выбранное задание. Можно удалять только неактивные задания.

Команда «Переименовать задание» используется для изменения имени любого задания.

Меню «Правка» используется для отправки копии отчета о состоянии задания в буфер обмена.

Пункт «Сохранить документ OmniPage как...» меню «Файл» позволяет сохранить промежуточный результат приостановленного задания в файл формата OPD.

Чтобы удалить файлы данных, в меню «Правка» выберите команду Удалить экземпляр. Она удаляет файлы данных, в которых хранятся сведения о выполнении экземпляра текущего задания. Команда Удалить все экземпляры позволяет удалить данные всех экземпляров выбранного задания. Эти два параметра помогают освободить дисковое пространство, однако используйте их с осторожностью, поскольку удаленные записи нельзя будет просмотреть.

Окно просмотра рабочих процессов

Окно просмотра рабочих процессов встроено в диспетчер пакетной обработки и отображается справа от списка заданий. В нем предоставляются подробные сведения о выполнении каждого задания. Весь процесс представляется поэтапно в соответствии с этапами рабочего процесса. Для каждого этапа приводятся входные и выходные данные, позволяя быстро просмотреть любую страницу. Результаты заданий отмечаются значками. Сведения об этапах обработки содержатся в раскрывающемся списке.

Контролируемые папки



В программе OmniPage Professional можно указывать контролируемые папки и почтовые ящики (Outlook и Lotus Notes) в качестве входных данных для **релерации** заданий. Это позволяет автоматически начинать обработку файла изображений сразу после его поступления в определенную папку или почтовый яшик.

Это удобно при поступлении набора файлов с известным содержимым из удаленных местоположений, которые обрабатываются автоматически по мере поступления, даже если никто лично не контролирует подобную обработку. Чаще всего это отчеты или документы в виде форм, передаваемые с определенным интервалом (например, раз в неделю или в месяц).

Чтобы воспользоваться этой возможностью, выберите папки (в том числе почтовые) для отслеживания. Эти папки не следует использовать ни в каких других целях, даже для заданий титульной страницы со штрихкодом. При настройке такого задания выберите задание на отслеживание папок, введите имя задания и нажмите кнопку «Далее». В открывшемся диалоговом окне выберите папки.

Входящие файлы удаляются из контролируемых папок после их передачи в OmniPage на обработку. Поэтому для хранения входящих файлов необходимо выделить дополнительное место на диске.



Добавьте нужные папки и типы файлов (один или все типы). Установите флажок перед выбранной папкой, если нужно отслеживать вложенные папки. Для подключения нескольких типов файлов добавьте папку несколько раз, по одному на каждый тип файла. Установите флажок для отслеживания вложенных папок.

После отображения следующей панели мастера заданий необходимо указать время обработки: время запуска и остановки отслеживания. Можно указать периодичность повторения, например проверку папок только в обеденный перерыв (начало в 12.15, окончание в 13.05) каждый понедельник, среду и пятницу или ночью в последние три дня каждого месяца, когда компьютер оставляется включенным для сбора и обработки поступающих месячных отсчетов.

Время ожидания после поступления файлов в папку определяется в окне «Параметры» диспетчера пакетной обработки. Все поступившие за это время файлы обрабатываются вместе. Когда файлы перестанут поступать, запускается процесс их обработки.

Чтобы закончить отслеживание раньше указанного срока, выберите команду «Сделать задание неактивным». После этого задание можно будет изменить.

Контролируемые почтовые ящики



В программе OmniPage Professional можно указать контролируемые почтовые ящики в качестве входных данных для задания. Как только файл изображения определенного типа поступает в выбранную папку почтовой



системы, задание автоматически запускает его обработку. Это удобно при поступлении набора файлов с известным содержимым, которые обрабатываются автоматически по мере поступления, даже если никто лично не контролирует подобную обработку.

Программа поддерживает просмотр почтовых ящиков Microsoft Outlook и Lotus Notes.

Обработка с помощью штрихкода



В программе OmniPage Professional можно запускать рабочие процессы (последовательности этапов с определенными параметрами) с помощью титульных страниц со штрихкодом, определяющих нужный процесс. Штрих-код на титульной странице определяет рабочий процесс посредством идентификатора, имени и этапов процесса, а также содержит имя создателя процесса, дату создания и т.д. Следует иметь в виду, что задания обработки с использованием штрихкодов не могут быть повторяющимися.

Существует два варианта управляемой штрихкодом обработки.

Ввод со сканера:

управление обработкой с помощью рабочего процесса выполняется путем размещения титульного листа сверху сканируемого документа, после нажатия кнопки «Начать» на сканере.

Ввод из графического файла:

для запуска обработки задания необходимо скопировать изображение титульной страницы со штрихкодом в контролируемую папку, в которую поступают изображения документов для обработки.

Для ввода со сканера необходимо выполнить следующие действия.

- 1. Создать рабочий процесс, содержащий этапы обработки отсканированных изображений.
- 2. Напечатать страницу со штрихкодом, идентифицирующим рабочий процесс.
- 3. Начать обработку с помощью штрихкода со сканера.

Сканирование по штрихкоду:

- 1. Поместите титульную страницу со штрихкодом поверх документа в устройство автоматической подачи.
- 2. Нажмите кнопку запуска на сканере.
- 3. На вкладке «Сканер» диалогового окна «Параметры» выберите «Рабочий процесс титульной страницы со штрихкодом» в качестве операции кнопки сканера по умолчанию. Можно также выбрать значение Запрос рабочего процесса. В этом случае откроется диалоговое окно с доступными параметрами. Сканирование, Рабочий процесс титульной страницы со штрихкодом и все рабочие процессы сканирования.

Все имеющиеся страницы будут обработаны с использованием указанного рабочего процесса. Выполнение процесса может быть прервано, если в процессе обработки встретится новая страница со штрихкодом. Результат будет сохранен, как указано в рабочем процессе.

Для ввода изображения необходимо создать задание титульной страницы со штрихкодом.

Для задания титульной страницы со штрихкодом используется особый тип отслеживаемой папки. Всегда используйте для обработки с помощью штрихкода отдельную папку. Временем начала выполнения рабочего процесса считается момент, когда титульная страница со штрихкодом попадает в отслеживаемую папку.

Для обработки задания титульной страницы со штрихкодом необходимо выполнить следующие действия.

- 1. Создать рабочий процесс, содержащий нужные этапы обработки. Выберите в качестве входных данных параметр «Загрузить файлы», установив флажок «Запрашивать файлы для загрузки при каждом запуске этого рабочего процесса».
- 2. Сохранить титульную страницу со штрихкодом, который определяет рабочий процесс.
- 3. Определить временные параметры для отслеживания соответствующей папки в диспетчере пакетной обработки путем создания задания титульной страницы со штрихкодом.

Чтобы обработать задание титульной страницы со штрихкодом, выполните следующие действия.

- 1. Убедитесь, что задание запускается в нужное время.
- 2. Как только титульная страница со штрихкодом будет помещен в отслеживаемую папку, начнется выполнение рабочего процесса.
- 3. Рабочий процесс будет применен к файлам, поступающим в папку после титульной страницы.
- 4. Обработка рабочих процессов будет завершена в заданный момент времени или по факту появления в папке новой титульной страницы.

Титульную страницу со штрихкодом и файлы изображений в обрабатываемую папку может скопировать любой пользователь, имеющий доступ к этой папке. Кроме того, существует возможность поместить в папку только титульную страницу и дать сетевому сканеру команду на обработку страниц и сохранение файлов изображений в этой папке.

Помощник подшивки

Помощник подшивки позволяет создавать рабочие процессы сканирования для повторяющихся заданий преобразования документов. Помощник для заданий сканирования, не требующих вмешательства пользователя во время обработки. Обычно операторы сканеров подготавливают документы, применяют соответствующую титульную страницу штрихкода для каждого из них и обычно не задумываются о последующей обработке или папке назначения для документа, поскольку все эти настройки уже определены до этого. Установите связь программы OmniPage с кнопкой сканера (см. главу 3 в разделе «Сканирование») и напечатайте титульную страницу со штрихкодом для идентификации рабочего процесса. После этого можно будет сканировать, выполнять преобразование и сохранять документ одним нажатием кнопки сканера.

Создание рабочего процесса

- 1. Выберите помощник подшивки в меню «Сервис».
- 2. Укажите имя рабочего процесса, а также тип выходного файла, его местоположение и имя.
- 3. Проверьте и при необходимости измените параметры рабочего процесса.
- 4. Напечатайте титульную страницу со штрихкодом.
- 5. Установите связь программы OmniPage с кнопкой сканера (выполняется однократно) на панели управления. См. раздел «Сканирование с помощью OmniPage и рабочих процессов» на стр. 38.

Применение рабочего процесса

- 1. Поместите титульную страницу со штрихкодом поверх документа в сканере.
- 2. Нажмите кнопку сканера, связанную с программой OmniPage. Документ будет преобразован с использованием этапов и параметров соответствующего рабочего процесса, а затем отправлен в указанное местоположение.

Для запуска заданий из отслеживаемых папок можно использовать титульные страницы со штрихкодом, сохраненные как файлы изображений. Подобные задачи допускают наличие интерактивных этапов, например разметку областей вручную и проверку, которые невозможны при использовании помощника подшивки.

Техническая информация

Эта глава посвящена поиску и устранению неполадок, кроме того, она содержит другую техническую информацию об использовании программы OmniPage. Прочтите также файл сведений и другие разделы справки или посетите веб-сайт Nuance.

Поиск и устранение неполадок

Несмотря на то, что программа OmniPage проста в применении, иногда возникают некоторые проблемы. Многие сообщения об ошибках сопровождаются описанием необходимых действий (проверить соединение, закрыть другие приложения, чтобы освободить память, и т. д.)

Дополнительные сведения об оптимизации операционной системы можно найти в справке программы OmniPage или в документации к системе Windows.

Здесь перечислены поддерживаемые форматы. В справке даны более подробные сведения.

Действия, которые следует предпринять в первую очередь

Далее описаны пути решения наиболее часто возникающих проблем запуска программы OmniPage и работы с ней.

- Убедитесь, что система удовлетворяет всем перечисленным требованиям. См. главу «Установка и настройка».
- Убедитесь, что сканер включен и все кабели надежно подключены.
- Посетите раздел поддержки на веб-сайте компании Nuance по адресу: www.nuance.com. Он содержит технические примечания о часто возникающих проблемах при использовании программы OmniPage. Кроме того, на веб-узле можно найти советы по установке программы и решению проблем.
- Перед запуском программы OmniPage проверьте работоспособность сканера с помощью программного обеспечения сканера, прилагаемого к сканеру.
- Убедитесь, что на компьютере установлены соответствующие драйверы сканера, принтера и видеокарты. Для получения дополнительных сведений посетите вебстраницу Nuance, используя меню «Справка» и обратитесь к разделу, посвященному сканеру.

- Выполните дефрагментацию жесткого диска. Дополнительные сведения см. в справочной системе Windows.
- Удалите и повторно установите программу OmniPage, как описано в разделе «Удаление программы» главы «Установка и настройка».

Тестирование программы OmniPage

Перезапуск системы Windows в безопасном режиме позволяет протестировать программу OmniPage в упрощенной системе. Это рекомендуется сделать при решении проблем, связанных с аварийным завершением и невозможностью запуска OmniPage. Дополнительные сведения см. в справочной системе Windows.

Тестирование программы OmniPage в безопасном режиме

- 1. Перезагрузите компьютер в безопасном режиме, нажав на клавишу F8 сразу после появления сообщения «Запуск Windows».
- 2. Запустите программу OmniPage и попробуйте выполнить распознавание изображения. Используйте проверенный графический файл, например один из файлов образцов изображений.
 - Если программа OmniPage не запускается или неудовлетворительно работает в безопасном режиме, проблемы могут быть связаны с установкой. Удалите и переустановите программу OmniPage, а затем запустите ее в безопасном режиме Windows.
 - Если программа OmniPage работает в безопасном режиме, значит, работе программы OmniPage может мешать один из драйверов устройств, установленных в системе. Попытайтесь устранить проблему, перезагрузив Windows в режиме загрузки с пошаговым подтверждением. Дополнительные сведения см. в справочной системе Windows.

Текст распознается неправильно

Если какая-либо часть исходного документа в процессе распознавания неправильно преобразуется в текст, воспользуйтесь следующими рекомендациями.

 Откройте изображение всей страницы и убедитесь в том, что все текстовые участки включены в текстовые области. Если участок не заключен в область, то в процессе распознавания он, как правило, пропускается. См. раздел, посвященный созданию и изменению областей, в главе «Обработка документов».

- Проверьте, правильно ли выделены текстовые области. При необходимости заново определите типы и содержимое областей и повторите распознавание документа. См. раздел «Типы и свойства областей» в главе «Обработка документов».
- Убедитесь, что по ошибке не загружен неподходящий шаблон. Если границы областей в тексте обрезаны, возможны ошибки распознавания.
- Отрегулируйте яркость и контрастность на панели «Сканер» в диалоговом окне «Параметры». Для достижения желаемых результатов можно поэкспериментировать с различными комбинациями параметров.
- Оптимизируйте распознаваемое изображение с помощью инструментов оптимизации.
- Проверьте разрешение исходного изображения. Переместите курсор на миниатюру страницы для вызова всплывающего окна. Если разрешение значительно отличается от 300 точек на дюйм, вероятнее всего, это отрицательно скажется на распознавании.
- Обязательно выберите необходимые языки документа на панели «Распознавание» в диалоговом окне «Параметры». Следует выбирать только те языки, которые включены в документ. В частности, установка азиатского языка для текста не на азиатском языке приведет у получению непригодного результата.
- Результаты распознавания на японском, корейском и китайском языках можно просмотреть и сохранить, только если в системе включена поддержка восточноазиатских языков. См. раздел «Распознавание текста на азиатских языках» на стр. 63.
- Включите функцию IntelliTrain и внесите правки. В основном это помогает при работе с художественными шрифтами или с документами, отличающимися более низким качеством. Если функция IntelliTrain активирована, попробуйте выключить ее, так как для документов некоторых типов с ухудшенными характеристиками она может не работать. См. раздел «IntelliTrain» на стр. 65.
- Создайте вручную несколько эталонов или измените существующие, чтобы удалить те из них, которые оказались неудачными.
- Если в текстовом редакторе или для экспорта используется уровень форматирования «Фактическая страница», то распознанный текст будет помещен в текстовые блоки или рамки. Если текстовый блок будет слишком

маленьким, какая-то часть текста может оказаться скрытой. Чтобы просмотреть текст, поместите курсор в текстовый блок и с помощью клавиш со стрелками на клавиатуре прокрутите содержимое блока вверх, вниз, влево или вправо.

• Проверьте стекло, зеркальные поверхности и линзы сканера на наличие пыли, загрязнений и царапин. При необходимости очистите их.

Проблемы с распознаванием факсов

Советы из этого раздела помогут повысить точность распознавания изображений, полученных по факсу:

- Просите отправителей присылать по возможности чистые оригиналы документов.
- Просите отправителей выбирать высокое или наивысшее качество передачи факса. Таким образом, полученное изображение будет иметь разрешение 200х200 точек на дюйм.
- Просите отправителей присылать изображения не по факсу, а по факс-модему напрямую на компьютер, если это возможно. В этом случае вы сможете сохранить факс в файл изображения и загрузить его в OmniPage. См. раздел «Использование файлов изображений» в главе «Обработка документов».

системные ошибки и проблемы производительности при распознавании

Советы из этого раздела относятся к ситуациям, связанным с аварийными отказами при распознавании, а также с очень медленным распознаванием.

- Проверьте качество изображения. Найдите в документации сканера способы повышения качества сканирования.
- Разделите сложные изображения страниц (содержащие много текста и графики или сложное форматирование) на более мелкие задания. Разметьте области вручную или измените выделенные автоматически и выполните распознавание каждой области по очереди. См. раздел «Работа с областями» в главе «Обработка документов».
- Перезапустите систему Windows XP или Vista в безопасном режиме и протестируйте программу OmniPage, выполнив распознавание прилагаемых файлов образцов изображений.

При одновременном выполнении нескольких задач, например задач распознавания и печати, распознавание может выполняться дольше.

Поддерживаемые типы файлов

Поддерживаемые форматы файлов изображений для загрузки: TIFF, PCX, DCX, BMP, JPEG, JB2, JP2, GIF, PNG, XIFF, MAX, PDF, XPS и HD Photo.

Форматы файлов, поддерживаемые при сохранении результатов распознавания в виде текста:

- HTML 3.2, 4.0
- Microsoft Excel 97, 2000, XP, 2003, 2007
- Microsoft PowerPoint 97
- Microsoft Publisher 98
- Microsoft Word 97, 2000, XP, 2003 (WordML), 2007
- Документы OmniPage
- PDF* (обычный), отредактированный, с изображениями на тексте, с подставленными изображениями
- RTF Word 6.0/95, RTF Word 97, RTF Word 2000, RTF 2000 ExactWord
- WordPad
- WordPerfect 12, X3
- Текст, текст с разбиением на строки, форматированный текст, текст, разделенный запятыми
- Текст Юникод, текст Юникод с разбиением на строки, форматированный текст Юникод, текст Юникод с разделителями-запятыми
- Аудиоконвертер (для сохранение озвучиваемого распознанного текста). Программа OmniPage Professional также поддерживает форматы
- eBook, Microsoft InfoPath (для форм), Microsoft Reader и XML.

Указатель

Чтобы перейти к упоминаемому пункту, щелкните номер страницы.

Цифры

3D-выравнивание 45 3D-выравнивание вручную 45

Α

Amazon Kindle 84 Arial Unicode MS 64

В

Box Net 34, 85

С

Cloud Connector 34, 85

D

Direct OCR 31 Dropbox 34, 85

Ε

Evernote 34, 85 Excel 2007 (XLSX) 106

I

IntelliTrain 65

Κ

Kindle 10, 84

L

Lotus Notes 92, 93, 98

Μ

Microsoft Live SkyDrive 34, 85 Microsoft Outlook 83, 92, 93, 98 Microsoft Word, открытие файлов PDF в 82

Ν

Nuance Cloud Connector 34, 85

0

OmniPage активация 18 документы в 19 запуск 14 переустановка 18 предыдущие версии 13 связывание с кнопками сканера 38 тестирование 103 удаление 18 установка 13 Outlook 92, 93, 98

Ρ

РарегРогt 18, 28 PDF в MS Word 83 PDF с возможностью поиска 81, 82 PDF с возможностью полнотекстового поиска 82 PDF, оптимизированный для отображения в Интернете 82 PDF, отредактированный 81 PDF, преобразование из/в 82 PowerPoint 2007 (PPTX) 106

Т

True Page 58

U

URL-адреса 68

W

Word 2007 (DOCX) 106

Х

XPS 82 XPS; 106

Α

абзац изменение атрибутов 67 стили 67, 78 автоматическая отправка по электронной почте 83 автоматическая разметка 39 автоматическая разметка вертикального текста 50 автоматическое определение макета 39 автоматическое создание эталонов 65 Агент OmniPage 17, 89 Агент запуска OmniPage 17 агент запуска OmniPage 89 азиатский текст, вертикальный 50 активация программы OmniPage 18 атрибуты символов 67

Б

базовое изображение 42

базовое изображение/изображение для распознавания (E) 43 безопасный режим 103 блоки 68 буфер отправка результатов распознавания 76 быстрая разметка областей 53

быстрое распознавание и сохранение 24

В

верификация текста 60 вертикальные словари 62 вертикальный текст 63 вертикальный текст, автоматическая разметка 50 Вид быстрого преобразования 20, 24 вид быстрого преобразования с простым загрузчиком 25, 35 видов виды 25 Виды Быстрое преобразование 24 Гибкий 22 изменение 25 использование меню «Окно» 22 Классический 21 Настройка 25 сброс 22 виды 20 виды рабочего стола OmniPage 20 включение значка OmniPage на панели задач 89 вложение в почтовые сообшения 83 возможности, новые 7 время выполнения заданий 98 встраивание шаблонов в файлы OPD 54 встраивание элементов в документы OPD 19 Входные файлы Word 39 выбор оптимальной яркости 37 выбрать область (Е) 43 выгрузка пользовательские словари 61 файлы эталонов 66 шаблоны областей 55 выделение области для вычитания (Е) 43 выделение области для добавления (Е) 43 выделение текста 70 выполнение задания без участия пользователя 94 задания диспетчера пакетной обработки 94 рабочие процессы 88 выравнивание (Е) 45

выравнивание вручную 45 выравнивание изображений с цифровой камеры 45 выравнивание страниц 45

Г

гибкий вид 20, 22 гиперссылки 68 группировка элементов 68

Д

данные формы, извлечение 74 два экрана 23 двусторонние документы 37 действия при медленном распознавании 102 диалоговое окно «Параметры» 27 динамический верификатор 60 дисковое пространство 12 диспетчер документов 20, 21 диспетчер пакетной обработки 92 добавление в области 53 вложения в почтовые сообщения 83 данных эталонов в файлы эталонов 66 слов в пользовательский словарь 58 этапы рабочего процесса 91 добавление возможности поиска в PDF 82 документы в программе OmniPage 19 двусторонние 37 описание макета 39 отправка в буфер обмена 76 с разным макетом 39 сохранение 76 экспорт 76 Документы Google 34, 85 документы OmniPage 19 сохранение как 76 драйверы сканера TWAIN 15 дуплексные сканеры 37

Ж

журнал оптимизации изображения 48

3

завершение проверка в рабочем процессе 88 рабочие процессы 91 разметка областей в рабочем процессе 88 загрузка
изображений с помощью простого загрузчика 25.35 изображения из проводника Windows 36 пользовательские словари 61 файлы изображений 33, 34 файлы эталонов 66 шаблоны областей 40, 54 шаблоны оптимизации изображений 48 задания выполнение 95, 97 выполнение без участия пользователя 94 запрет 94 изменение 94 инструкции по времени выполнения 98 ограничение страниц 94 повторяющиеся 98 сообщения об ошибках 95, 97 состояние 95, 97 уведомление о завершении задания 92 управление 95, 97 закраска черным цветом конфиденциальных слов 70 закрепление и открепление панелей инструментов 60 закрепление панелей 20 заливка (Е) 45 заливка цветом области изображения 45 замена шаблонов областей 55 запрет выполнения задания 94 запрещенные формы областей 53 запуск диспетчера пакетной обработки 92 запуск нового рабочего процесса 91 запуск программы 14 зачеркивание текста 70 зеркальное отражение изображений 44 значок рабочего процесса на панели задач 89

И

извлечение данных формы 74 Извлечение текста из PDF-файлов 83 извлечение элементов из документов OmniPage 19 изменение видов 20 задания 94 качество изображения 41 области 52 порядок обхода 69 таблицы 54, 68 часть страницы 69 шаблоны областей 55 изменение масштаба 21, 60 Изменение рабочих процессов 91 изменение рабочих процессов 91 изображения в данных экспорта 78 выравнивание 45 зеркальное отражение 44 качество 37 обрезка 44 оттенки серого 77 поворот 44 подставленные изображения в файле PDF 81 разрешение 45, 77, 104 редактирование 67 сохранение 77 фон 49 цветные 77 черно-белые 77 индикаторы колонтитулов 57 инструмент выделения (F) 72 инструмент выноса на передний план (F) 73 инструмент выравнивания по вертикали (F) 73 инструмент выравнивания по горизонтали (F) 73 инструмент изображения (F) 73 инструмент линии (F) 73 инструмент перемещения на задний план (F) 73 инструмент поля ввода (F) 73 инструмент прямоугольника (F) 73 инструмент составного поля (F) 73 инструмент таблицы (F) 73 инструмент текста (F) 72 инструмент текста в кружочке (F) 73 инструмент флажка (F) 73 Инструменты оптимизации изображений определение области 43 инструменты оптимизации изображений 43 интеллектуальная проверка IntelliTrain 104 интеллектуальные папки 97, 98 интерактивные этапы задания 94 исходное изображение 42

К

качество изображений 37 китайский 63 классический вид 20, 21 кнопка «Документ готов» 88 кнопка «Страница готова» 88 кнопка «Экспорт результатов» 77 конвертеры для экспорта 79 конвертеры множественные 80 контрастность 37, 104 Контролируемые папки 98 копирование в буфер обмена 76 корейский 63

Л

Линеаризованный PDF 82 линеаризованный PDF 82

Μ

макет документа, форма 40 макет, автоматическое определение 39 Мастер выполнения рабочих процессов OmniPage 17 мастер выполнения рабочих процессов OmniPage 89 мастер для прямого преобразования 36, 83 мастер настройки сканера 14 мастер преобразования 36, 82, 83 масштаб (Е) 43 мгновенная разметка областей 69 мгновенные изменения 69 медицинские словари 58, 62 медленное распознавание 105 меню «Окно» для управления видами 22 минимальные требования к системе 12 многостраничные файлы изображений 77 множественные конвертеры 80

Н

назначение рабочих процессов 87 назначение эталонов 65 направление текста 50, 63 настройка Direct OCR 32 настройка конвертеров для экспорта 79 настройка сканера 14 непечатаемые символы 57 Новые возможности 7 новые возможности 7 номера юридических документов 40

0

обзор этапов обработки 20 области 51 буквы и цифры 51 в Direct OCR 33 вертикальный азиатский текст 63 вручную 49, 103, 105 выгрузка шаблонов 55 добавление в 53 замена шаблонов 54 запрещенные формы 53 изменение типов 52

изменение шаблонов 54 изображение 52 неправильная форма 53 обработка 52 объединение 53 пропустить 52 работа с 52 свойства 51 сохранение шаблонов 54 таблица 52, 54 типы 51, 104 удаление шаблонов 54 цифры 51 шаблоны 40, 54, 104 области букв и цифр 51 области изображений 52 области с несколькими колонками 68 области таблиц 52 области цифр 51 область формы 52 обрабатываемые области 52 обработка вручную 31 из других приложений 31 основные этапы 20 пошаговая 31 с использованием рабочих процессов 88 этапы, обзор 20 обработка одним нажатием кнопки 25, 35, 38 обработка по расписанию 92 Обработка с помошью штрихкода 99 обработка фона 49 обработка форм с пропуском цветов 45 обрезка (Е) 44 объединение областей 53 объекты формы, редактирование 74 ограничение страниц для заданий 94 Озвучивание текста с помощью компонента RealSpeak 70 описание макета 39 описание макета документа 39 определение областей в Direct OCR 33 определение области для инструментов оптимизации изображений 43 оптимизация изображения в рабочих процессах 49 журнал 48 инструменты 43 основные этапы обработки 20

остановка рабочих процессов 89 открепление панелей 20 открытие конечное приложение 78 рабочих процессов с рабочего стола 89 открытие рабочих процессов с рабочего стола 89 открытие файлов изображений 33 отложенная обработка страниц 69 отметки 57, 59 отмеченные слова в редакторе 57 отображение PDF-файлов в Интернете 82 отображение и скрытие пометок 57 отображение позиции закрепления 20 отображение справки 20, 25 Отправка в Kindle 84 в буфер 76 страницы по почте 83 отправка по почте 83 отправка уведомлений по электронной почте о завершении задания 92 отслеживаемые папки 97 отслеживаемые почтовые ящики 98 отслеживание почтового ящика 98 отсутствующие в словаре слова 57 оттенки серого изображения 77 сканирование 37 оттенок/насыщенность (Е) 44 отчеты о ходе выполнения из рабочих процессов 97

П

Панели 20 панели 20 панели в виде вкладок 20 Панели инструментов 25 панели программы 20 панели, расположенные мозаикой 20 панель «Изображение страницы» 20 панель изображения 20 панель инструментов «Изображение» 20 панель инструментов «Создание формы» 72 Панель инструментов «Управление формами» 73 панель инструментов OmniPage 21 панель инструментов форматирования 20 параметры диалоговое окно «Параметры» 27 для Direct OCR 32

получение текста 32 типы областей 54 параметры проверки 58 параметры рабочих процессов 89 параметры сохранения 79 пароли для PDF-файлов 82 переустановка OmniPage 18 плавающие панели 20 поворот (Е) 44 повторное закрепление панелей 20 повторное создание эталонов 65 повторный экспорт 76 повторяющиеся задания 94, 98 повышение точности распознавания 36, 66, 103 Поддержка восточноазиатских языков 13 поддержка восточноазиатских языков 63 Подключение к Интернету для активации. 13 подчеркивание контуров букв 44 подчеркнутый текст 67 поиск отсутствующие в словаре слова 58 сомнительные слова 58 поиск в выводе PDF 81 поиск и устранение неполадок 102 полужирный текст 67 получение со сканеров 36 из PDF-файлов 34 из файлов PDF 33 из файлов изображений 33 с цифровой камеры 34 через простой загрузчик 35 получение из PDF-файла 34 получение из файла PDF 33 получение с цифровой камеры 34, 45 получить и преобразовать 35 пользовательские виды 25 пользовательские словари 58, 61 пользовательский макет 40 пометка текста 70 Помощник eDiscovery для PDF с возможностью поиска 82 помощник подшивки 101 помощник рабочего процесса 31, 89 порядок обхода 69 порядок элементов страницы 69 поточная страница 78 пошаговая обработка 20 предварительная обработка изображений 41

предложения при проверке 58 предотвращение закрепления панели с помощью клавиши CTRL 20 предупреждения из заданий 97 Преобразование документов 39 преобразование из формата PDF 82, 83 преобразование файлов изображений 89 приостановка рабочих процессов 89 проблемы с производительностью при распознавании 105 проблемы с факсами 105 проверка в рабочем процессе 88 параметры 58 проверка выбора языка 63 проверка выбранных языков 63 проверка результатов распознавания 60 проводник Windows 36, 89 проводник, загрузка файлов 36, 89 пропуск изображений при экспорте 78 пропуск интерактивных этапов задания 94 пропуск фона 49 пропуск цветов для форм 45 пропускаемые области 52 пропускаемый цвет (Е) 45 прослушивание озвучиваемого текста 71 просмотр вертикального текста 50 просмотр входных или выходных файлов 97 просмотр ход выполнения рабочего процесса 97 Просмотра рабочих процессов 97 Простой загрузчик 22 простой загрузчик 20, 35 простой загрузчик в быстром виде 25, 35 простой текст в редакторе 57, 78 профессиональные словари 58, 62 пункты меню «Получить текст» 32

Ρ

работа с Direct OCR 32 работа с областями 52 рабочие процессы выполнение 88 для извлечения данных формы 74 завершение 91 запуск со сканера 38 значок на панели задач 89 приостановка и остановка 89 просмотр состояния 97 создание 91

состав 87 участие пользователя 88 шаги и параметры 89 этапы оптимизации изображения 49 рабочий стол 20 рабочий стол программы OmniPage 20 разгруппировка элементов 68 разметка областей в рабочем процессе 88 разметка областей вручную 49 Разновидности PDF 81 разрешение 77, 104 разрешение (Е) 45 рамки 68, 79, 104 рамки для распознаваемого текста 104 распознавание Direct OCR 31 диспетчер пакетной обработки 92 низкая производительность при 105 параметры для Direct OCR 31 проблемы с факсами 105 проверка результатов распознавания 60 результаты проверки 58 сохранение результатов 77 точность 37, 65, 103 **ускорение** 105 языки 62, 104 распознавание изображения 42 Распознавание текста на азиатских языках 63 распознавание факсов 105 расширенные параметры сохранения 79 Регистрация 17 редактирование атрибуты абзацев 67 атрибуты символов 67 в режиме True Page 68 вертикальный текст 50 вывод в PDF 81 изображения 67 объекты формы 74 пользовательские словари 61 путем мгновенных изменений 69 распознанный текст 67 таблицы 54, 68 файлы эталонов 66 редактирование в режиме True Page 68 редактирование текста 70 редактор

уровни форматирования 57 рекомендации по устранению неполадок 102

С

сброс видов 22 свойства областей 51 связывание функций OmniPage с кнопками сканера 38 сжатие MRC 82 символы, сомнительные 57 синхронизовать панели (Е) 44 системные ошибки и проблемы производительности при распознавании 105 сканеры 105 драйверы 15 дуплексные 37 настройка 14 сканирование 37 в рабочих процессах 101 изображения 37 мастер 14 получение со 37 с помощью рабочих процессов 38 словари 58 создание данные эталона 66 новые рабочие процессы из существующих 91 рабочие процессы 91 создание пользовательского словаря 61 создание эталонов вручную 65 сомнительные слова 57 сообщения об ошибках из заданий 95, 97 состав рабочих процессов 87 состояние заданий 95, 97 состояние рабочего процесса 20, 25, 97 сохранение в почтовые сообщения 83 в файл 77 в файлах различных типов 80 документы 76 и открытие 78 исходные изображения 77 как документ OmniPage 76 параметры 79 пользовательские словари 61 результаты распознавания 77 текст 77 файлы PDF 81 файлы эталонов 66

шаблоны областей 55 сохранение и применение шаблонов оптимизации изображения 48 сохранение изменений в областях 69 сохранение исходных изображений 77 сохранение макета 58 сохранение стилей абзацев 78 сохранение текста 77 специализированные словари 62 сравнение распознанных слов с исходными 60 ссылки на веб-страницы 68 стандартная панель инструментов 21 страницы выравнивание 45 многостраничные файлы изображений 77 навигация 21 отправка в буфер обмена 76 отправка по почте 83 страницы с несколькими колонками 40 страницы электронных таблиц 40 страницы, отсканированные с перекосом 45 страницы, содержащие одну колонку с таблицами 40 существующий рабочий процесс как источник для нового рабочего процесса 91

Т

таблица символов 60 таблицы в текстовом редакторе 68 изменение разделителей 54 на страницах, содержащих одну колонку 40 области 52. 54 редактирование 68 строки в 54 удаление разделителей 54 текст, выделенный курсивом 67 текстовый редактор 20, 57, 67 тестирование программы OmniPage 103 техническая информация 102 типы областей 51 титульные страницы для обработки штрихкодов 99 тон/насыщенность (Е) 44 точность влияние на сканирование 36 влияние удаления дефектов 44 влияние эталонов 65 влияние яркости 37 повышение 36, 65, 103

Требования для поддержки азиатских языков 13 Требования к операционной системе 13 требования к памяти 12, 103 требования к системе 12

У

уведомление о завершении задания 92 увеличение памяти 103 увеличение рабочей области 23 удаление задания 96 пользовательские словари 61 файлы эталонов 66 удаление дефектов (E) 44 удаление краев изображений 44 удаление программы 18 удаление точек с изображений 44 удаление шаблонов областей 55 удаление шума с изображений 44 удаление этапов рабочего процесса 91 указатель (Е) 43 уменьшение области изображения 44 уменьшение разрешения изображения 45 управление заданиями 95, 97 управление рабочими областями 23 упрощенный интерфейс пользователя 24 уровень форматирования «Простой текст» 78 уровень форматирования «Форматированный текст» 78 уровни форматирования 58, 78 простой текст 57 Фактическая страница 58 Форматированный текст 58 установка OmniPage 13 сканеры 14 устройство автоматической подачи 33, 37 участие пользователя в рабочих процессах 88

Φ

файлы в качестве источника изображений 33 назначение экспорта 77 параметры разделения 77 сохраняемые при удалении 18 типы для экспорта 78 файлы OPD встраивание шаблонов 54 встраивание элементов 19 извлечение элементов 19 файлы PDF, преобразованные в документы Word 82 файлы изображений образцы 103 получение 33 порядок обхода 33 преобразование 89 файлы образцов изображений 103 финансовые словари 62 фон для разметки областей 49 Форматированный текст 58 фрагменты английского текста в азиатском тексте 63 функция преобразования текста в речь 71

Ц

цветные изображения 77 маркеры 59 сканирование 37

Ч

черно-белые изображения 77 сканирование 37

Ш

шаблон, формы 74 шаблоны в документах ОРD 54 шаблоны областей 40, 54, 104 шаги для рабочих процессов 89 шифрование PDF-файлов 82

Э

экспорт в Kindle 84 в PDF 81 в буфер 76 в поточной странице 78 в почтовые сообщения 83 в режиме True Page 79 в файл 77 изображения 78 повторный 76 экспорт в режиме True Page 79 эскизы 20, 21 эталон 65 автоматически (IntelliTrain) 65 вручную 65 файлы эталонов 66

Ю

юридические документы 40 юридические словари 58, 62

Я

языки 62, 104 японский 63 яркость 37, 104 яркость / контрастность (E) 44 яркость распознаваемого (E) 44

ЛИЦЕНЗИИ И ИНФОРМАЦИЯ СТОРОННИХ ПРОИЗВОДИТЕЛЕЙ

Компоненты данного программного обеспечения, связанные с верификацией и переносом слов, а также проверкой правописания, частично основаны на технологии Proximity Linguistic. Система расстановки переносов Proximity. © 1988. Все права защищены. Franklin Electronic Publishers, Inc.

Словарные базы данных для английского языка (США) Proximity/Merriam-Webster. © Merriam-Webster Inc., 1982, 1983, 1987, 1988. © Franklin Electronic Publishers, Inc., 1982, 1983, 1987, 1988. Для проверки слов используются словарные базы данных Proximity/Merriam-Webster, насчитывающие 116 000, 80 821, 92 641, 106 713, 118 533, 91 928, 103 792, 130 690 и 140 713 слов. Словарные базы данных для английского языка (Великобритания) Proximity/Collins. © William Collins Sons & amp; Co., Ltd., 1985. Юридические и медицинские словари. © Merriam-Webster Inc., 1982. © Franklin Electronic Publishers, Inc., 1982, 1985. Для проверки слов используются словарные базы данных Proximity/Collins, насчитывающие 80 307, 90 406, 105 785 и 115 784 слов. Словарные базы данных Proximity/Collins для испанского, итальянского, немецкого, португальского (Бразилия и Португалия) и французского языков. © William Collins Sons & amp; Co., Ltd., 1984, 1985, 1986, 1988. © Franklin Electronic Publishers, Inc., 1984, 1985, 1986, 1988. Для проверки слов используются словарные базы данных Proximity/Collins, насчитывающие 136 771, 150 893, 178 839, 207 119, 212 565 и 194 393 слов. Словарная база данных Proximity/Van Dale для нидерландского языка. © Van Dale Lexicografie bv., 1987. © Franklin Electronic Publishers, Inc., 1987 Для проверки слов используется словарная база данных Proximity/Van Dale, насчитывающая 119 614 слов. Словарная база данных Proximity/Munksgaard для датского языка. © Munksgaard International Publishers Ltd., 1988. © Franklin Electronic Publishers, Inc., 1988. Для проверки слов используется словарная база данных Proximity/Munksgaard, насчитывающая 113 000 слов. Словарные базы данных Proximity/IDE для норвежского и шведского языков. © IDE a.s., 1988. © Franklin Electronic Publishers, Inc., 1988. Для проверки слов используются словарные базы данных Proximity/IDE, насчитывающие 126 123 и 150 000 слов.

Словари INSO/Vantage Research: система проверки правописания International CorrectSpellTM © Lernout & Marpi, 1993.

База данных проверки правописания для словенского языка. © Ambeis d.o.o., 2002.

Словарь для эсперанто основан на данных Туна Виткама (Toon Witkam) и Стефана Макгилла (Stefan MacGill).

Функции распознавания текста на азиатских языках совместно разработаны компаниями Beijing Wintone Information Technology Corporation Ltd и Nuance Communications, Inc. Все права защищены.

Проект International Components for Unicode (ICU). © International Business Machines Corporation и др., 1995—2009.

Данное программное обеспечение частично основано на разработках компаний Independent JPEG Group и Colosseum Builders, Inc.

Программное обеспечение Independent JPEG Group. © Томас Дж. Лэйн (Thomas G. Lane), 1991—1995.

Части этого программного обеспечения. © FreeType Project <www.freetype.org>, 2006. Все права защищены. FreeType 2.3.1, Turner, Wilhelm, Lemberg.

Zlib. © Жан-Лу Гайи (Jean-loup Gailly) и Марк Адлер (Mark Adler), 1995-1998.

Данное программное обеспечение разработано с использованием программы Kakadu.

Элементы управления в диалоговом окне «Параметры экспорта» от Аллана Нильсена (Allan Nielsen), элемент управления Supergrid. © 1999.

Данный продукт включает компоненты, разработанные в рамках проекта OpenSSL http://www.openssl.org, включая программное обеспечение, написанное Эриком Янгом (Eric Young) и Тимом Хадсоном (Tim Hudson).

Часть данного программного обеспечения основана на алгоритме создания «дайджестов» сообщений MD5 компании RSA Data Security, Inc.

Шифрование и дешифрование по алгоритму AES для PDF-файлов. © Д-р Брайан Глэдман (Dr Brian Gladman), 2001, Вустер, Великобритания.

Kindle 2. © Amazon, 1999-2011.

Компоненты для обработки азиатских шрифтов. © Adobe Systems Incorporated, 2009. Все права защищены.

Некоторые элементы интеграции и другие компоненты. © Microsoft Corp., 2009. Все права защищены.

Создание PDF-файлов. © Zeon Corporation, 1993—2011. Все права защищены.

RealSpeak™ Solo. © Nuance Communications, Inc., 2002—2011. Все права защищены.

© Nuance Communications, Inc., 2011.

Все права защищены. Документ может быть изменен без предварительного уведомления.