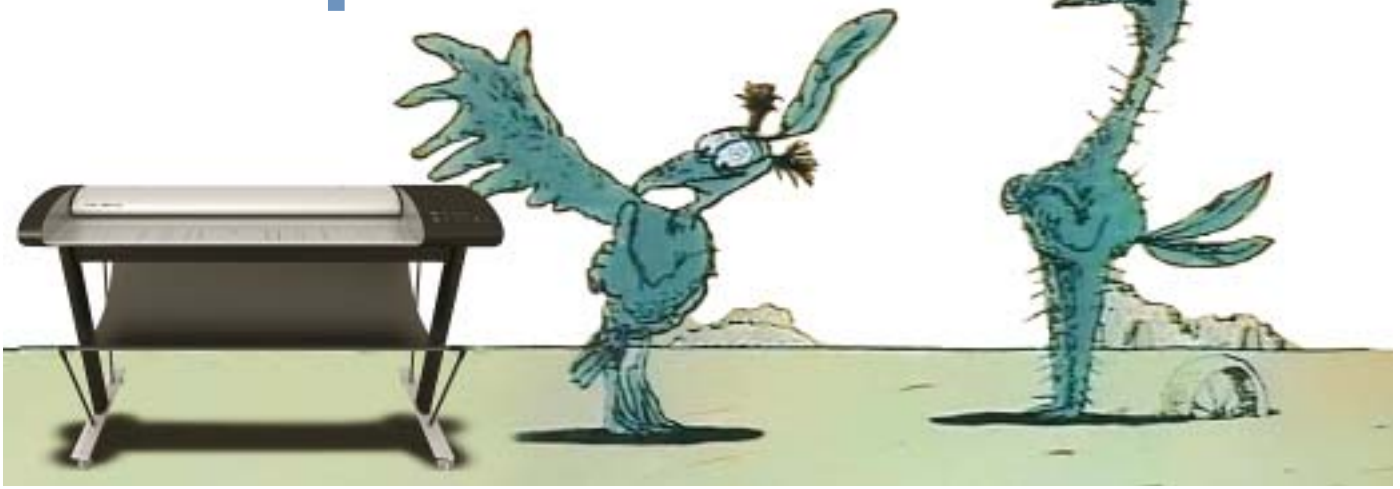


"Лучше день потерять, потом за пять минут долететь", или Как сократить время перевода бумажного архива в электронный



При сканировании большого количества документации, как правило, стоит задача максимально сократить затраченное время. Естественно, возникает вопрос, как ускорить сканирование и ввод отсканированных данных в электронный архив. В этой статье мы рассмотрим некоторые приемы, позволяющие оптимизировать данный процесс, на примере обработки чертежей и помещения их в структуру электронного архива с использованием программ RasterID, Spotlight и системы электронного архива TDMS. Поточковый ввод документов в электронный архив — набор рутинных операций, требующих от участников процесса четкости выполнения инструкций и усидчивости. Весь процесс наполнения электронного архива разобьем на два этапа: сканирование и регистрация документов в электронном архиве. Соответственно введем понятия "оператор сканера" и "архивариус".

Сканирование

Оператор сканера может настроить кнопки на панели сканера Contex и назначить через WIDESystem для каждой кнопки определенный режим запуска сканирования в RasterID. Например, установить для кнопки *Scan* функцию сканирования и сохранения в файл, а для кнопки *Color Copy* — функцию сканирования и ручной обработки в RasterID (рис. 1).

При сканировании документов разного формата в пакетном режиме удобно установить в параметрах сканирования RasterID опцию автоматического определения ширины — *Ширина Авто* (рис. 2). В этом случае сканер протягивает оригинал на установленную в программе величину, определяет размер, возвращает оригинал обратно и продолжает сканировать в определенных габаритах. Это позволяет ускорить процесс скани-



Рис. 1. Настройка кнопок сканера на режимы сканирования в RasterID



Рис. 2. Автоматическое определение ширины в параметрах сканирования

рования, уменьшить размер растрового изображения за счет исключения пустых полей на изображении и соответственно уменьшить время последующей обработки изображения.

Оптимизация процесса сканирования заключается не только в выборе самого быстрого режима, но и в исключении потери данных. Современное широкоформатное оборудование дает возможность компенсировать фон, подавлять растровый мусор. В потоковом режиме оператору сканера бывает трудно отследить правильность аппаратной обработки и исключить потери данных на сложных документах. Можно сделать предварительный просмотр для подбора параметров, но на это уходит время, и подобрать параметры можно только для отдельной части изображения. Более корректно провести чистку и подавление фона можно программными средствами в RasterID или в программах Raster Arts после сканирования.

В настройках RasterID есть возможность подключить командный файл при сканировании для обработки изображений, автоименования и сохранения (рис. 3). Использование командного файла позволяет автоматизировать выполнение отдельных команд для повышения качества. Такие операции, как *Автоматическое устранение перекоса*, *Обрезка изображения* и другие фильтры, установленные в командном файле, RasterID будет применять к отсканированному изображению автоматически (рис. 4).

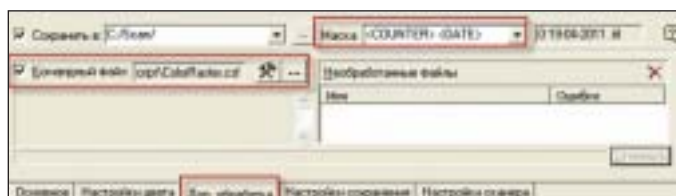


Рис. 3. Подключение командного файла при сканировании

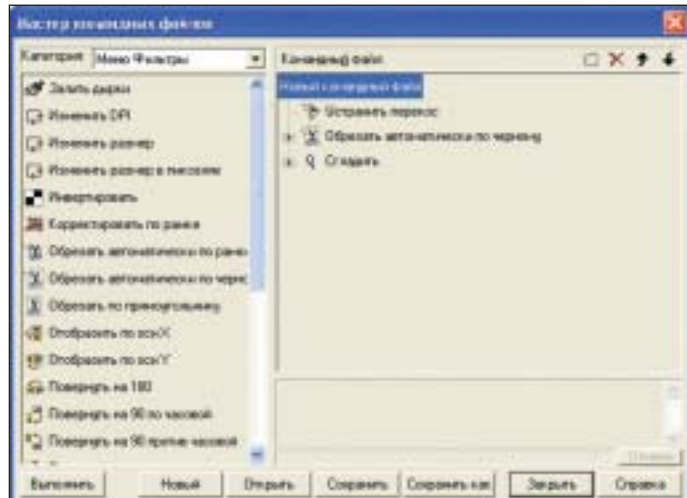


Рис. 4. Командный файл

После сканирования и автоматической обработки файл будет сохранен на диске с установленной маской имени, которая формируется из отдельных системных переменных: счетчика, даты, времени и т.д. (рис. 5)

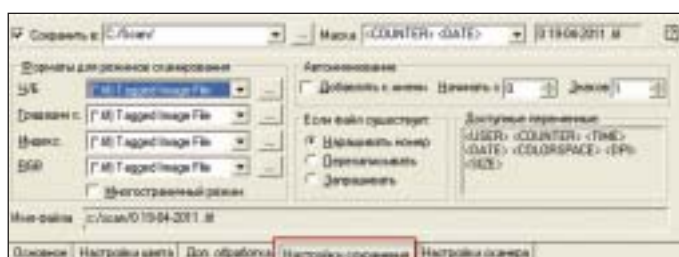


Рис. 5. Настройка сохранения

Обработка

Обработка отсканированных чертежей перед внесением их в электронный архив — достаточно трудоемкий и долгий процесс. Операции по устранению перекоса, обрезке и повороту изображения (при необходимости) должны быть выполнены при сканировании автоматически, чтобы не тратить на это время. Автоматическое удаление мусора в отдельных случаях может привести к потере данных с чертежа, когда вместе с растровым мусором удаляются и мелкие элементы (точки, запятые и т.д.). Перенос мусора на отдельный слой и его анализ — более корректный способ, но он требует достаточно много времени на поиск случайно попавших на слой с мусором полезных объектов и восстановления их на исходном слое. Оператор сканера (или специалист по обработке) может установить в команде *Автоматическое удаление мусора* наиболее крупные фракции, а очистку мелкого мусора при необходимости возьмет на себя пользователь при работе с растровым документом в дальнейшем. На этом этапе каждый может установить наиболее приемлемые для себя параметры удаляемого мусора.

Рассмотрим еще несколько приемов, которые позволят сократить время обработки изображений.

Интерфейс

Настройка удобного интерфейса в Spotlight позволяет значительно сэкономить время. Во-первых, это касается панелей инструментов. Пользователь может настроить необходимое ему рабочее пространство: отключить ненужные для его работы панели инструментов или создать свои панели, наполнив их необходимыми командами. Настроенный интерфейс может быть сохранен в отдельном профиле. Таких профилей под разные режимы работы пользователь может сделать сколько угодно. Переключение между профилями происходит простым выбором в списке. При этом мгновенно изменяется интерфейс под нужный режим работы (рис. 6).

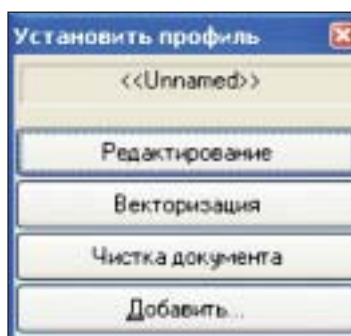


Рис. 6. Выбор профиля в Spotlight

Пользовательские кнопки

Воспользуйтесь функцией настройки интерфейса для создания пользовательских кнопок. Spotlight позволяет создать кнопку и назначить выполнение командного файла при ее нажатии (рис. 7). Это очень простой и удобный механизм, который значительно сокращает время работы. Нажав эту кнопку, вы сможете запустить целый набор команд с настройками под определенный тип растрового изображения.

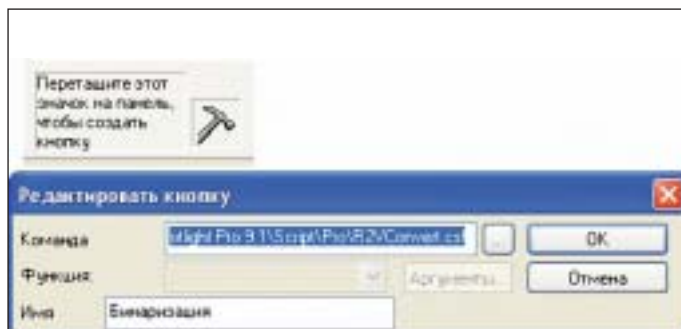


Рис. 7. Настройка пользовательских кнопок

Горячие клавиши

При обработке большого количества документов в ручном режиме приходится по многу раз вызывать различные команды. Spotlight позволяет настроить практически для любой команды сочетание клавиш для ее запуска (рис. 8). Подберите наиболее удобные вам сочетания клавиш для часто используемых команд и в процессе работы вы будете значительно экономить ваше время.

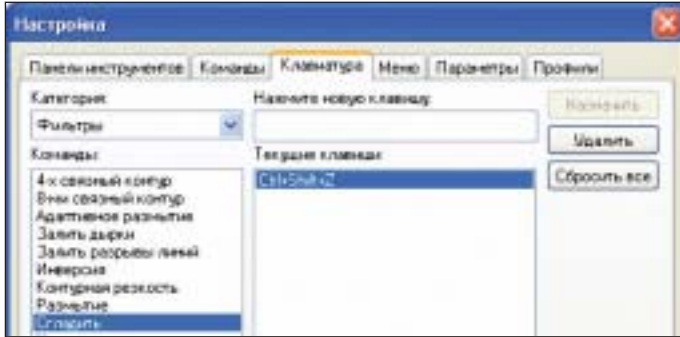


Рис. 8. Настройка горячих клавиш

Регистрация в архиве

Отсканированные и обработанные растровые изображения необходимо поместить в систему электронного архива. Рассмотрим этот процесс на примере системы электронного архива TDMS. Преимущества системы электронного архива по сравнению с другими методами хранения бесспорны. Это и неограниченные возможности поиска информации, высокий уровень безопасности, учет изменений, удобный интерфейс и много других полезных функций. Рассмотрим этапы занесения отсканированных документов в этот электронный архив и попробуем оптимизировать некоторые процессы.

Одним из основных элементов электронного архива является карточка документа, в которой содержится атрибутивная информация. Штамп отсканированного чертежа содержит всю информацию, необходимую для его регистрации в системе электронного архива. Одной из уникальных особенностей программы RasterID является функция автоматического поиска табличной формы (штампа чертежа), распознавания ее топологии и текста. Специально созданный интерфейс TDMS — RasterID позволяет автоматизировать процесс поиска штампа на отсканированном документе и распознавания текста, передачу отсканированного изображения в структуру электронного архива и автоматическое создание и заполнение карточки документа из полей штампа (рис. 9).

В RasterID в качестве приемника данных выбирается и настраивается TDMS. Поля карточки в TDMS связаны с полями шаблона штампа RasterID, используемого для поиска на отсканированном изображении. Кроме автоматического заполнения карточки в TDMS передается отсканированное изображение, которое привязывается к этой карточке.

Данный процесс может быть автоматизирован двумя способами: при сканировании и в пакетном режиме.

Сканирование чертежа с автоматической регистрацией в архиве

Этот способ позволяет оператору сканера, нажав всего одну кнопку на панели аппарата, автоматически зарегистрировать отсканированное изображение в архиве с заполнением карточки. В начале статьи мы рассматривали возможность назначения на управляющие кнопки сканера запуска режима сканирования в RasterID. Так вот, если в настройках сканирования RasterID задан *Командный файл*, то растровое изображение будет обработано по заданному алгоритму с автоматическим со-



Рис. 9. Интерфейс RasterID – TDMS

хранением обработанного файла на жестком диске. Теперь, если в этот командный файл включить команду *Распознать штамп* и заранее установить приемник данных TDMS, отсканированное и обработанное изображение будет помещено в электронный архив TDMS. Наиболее эффективно этот способ можно использовать при сканировании документации, созданной в электронном виде (чертежи с подписью). Как правило, такого рода документы имеют высокое качество изображения и не требуют затрат на восстановление и корректировку распознанного текста из штампа.

Пакетный режим ввода отсканированных документов в TDMS

Еще один инструмент автоматизации RasterID — *Мастер пакетных заданий*. Эта команда позволяет выбрать набор файлов или целиком папку с отсканированными изображениями на жестком диске и применить к ним командный файл. При этом RasterID в автоматическом режиме будет последовательно открывать каждый файл, выполнять команды обработки, регистрировать документ в системе электронного архива.

Размещение документов в структуре архива

Отсканированные документы после обработки попадают в раздел *Отсканированные непроверенные документы*. Архивариус (оператор электронного архива) может последовательно проверить правильность заполнения атрибутов карточки, руководствуясь растровым изображением штампа в карточке, внести необходимые изменения и установить статус *Проверен*. После проверки документ автоматически перемещается в раздел *Проверенные неструктурированные документы*. Распределение документа по структуре проекта может быть в дальнейшем проведено в ручном режиме или этот процесс может быть автоматизирован, если структура проекта заранее создана и настроены связи атрибутов карточки со структурой архива для автоматического размещения.

Автоматическое именование документов

Некоторые пользователи предпочитают именовать отсканированные изображения чертежей, используя информацию из документа. Если регистрировать отсканированный документ в электронном архиве, в этом нет особой необходимости, так как растровое изображение связано с карточкой, где содержится вся атрибутивная информация для поиска. Тем не менее, можно автоматизировать процесс именования файлов с использованием распознанной текстовой информации.

Анализируя настройки сканирования RasterID, мы рассматривали опцию сохранения файлов и автоматического именования их системными переменными: счетчиком, датой и т.д. Существует и второй способ сохранения. В настройках сканирования нужно задать папку, куда будут помещаться отсканированные файлы и выключить опцию *Сохранять в* (рис. 10). В приемнике данных перейти на закладку *Именование файлов* и в качестве имени файла вставить название поля из штампа документа. Можно комбинировать несколько полей и вставлять их, а также вставлять текст (рис. 11).

Чтобы проверить качество распознавания текста перед именованием документа, можно в параметрах приемника данных ус-

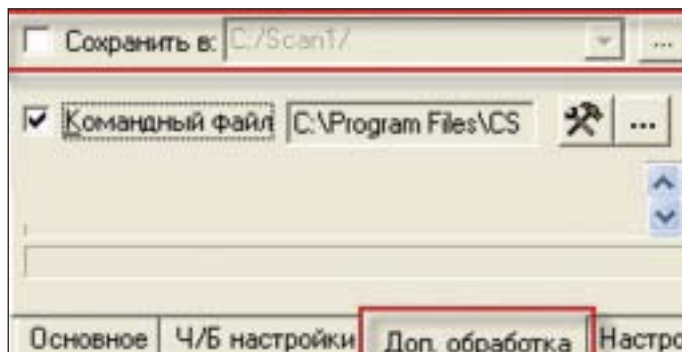


Рис. 10. Настройка пути сохранения

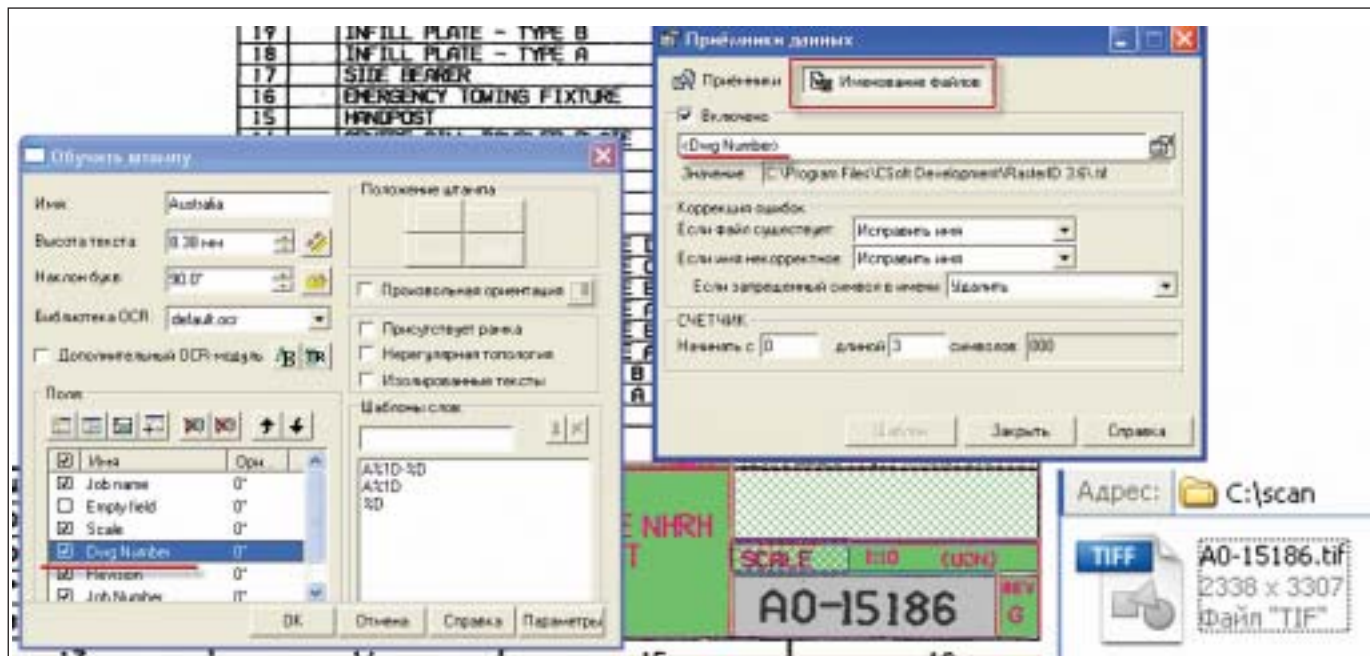


Рис. 11. Автоматическое именование документа с использованием распознанного текста

тановить параметр *Остановка для проверки* (рис. 12). При этом перед именованием, сохранением и размещением документа в электронном архиве RasterID будет отображать диалог проверки (рис. 13).

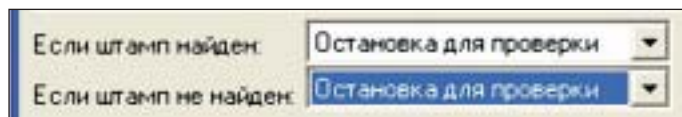


Рис. 12. Остановка для проверки штампа

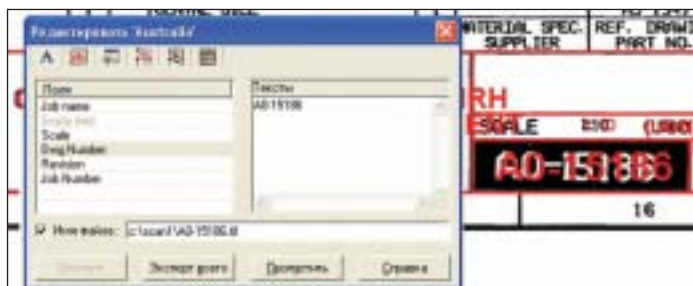


Рис. 13. Проверка распознавания текста в штампе

Выводы

Мы рассмотрели некоторые приемы, позволяющие сократить времени перевода бумажного архива в электронный вид. Скорость работы сканера во всем процессе — важный, но далеко не основной показатель. Гораздо больше времени занимают процессы обработки и регистрации изображения в электронном архиве. Если вы хотите оптимизировать эту работу, необходимо создать четкие инструкции для участников процесса, предварительно настроить оборудование и программное обеспечение, создать необходимые шаблоны, настроить оптимальный интерфейс. Для оценки времени перевода архива можно сделать подборку документов, в тестовом режиме завести их в систему электронного архива и по результатам выполнения такого пилотного проекта скорректировать параметры.

Илья Шустиков
ЗАО "CuSoft"
Тел.: (495) 913-2222
E-mail: Shustikov@csoft.ru