

ПЕРЕВОД БУМАЖНОЙ ДОКУМЕНТАЦИИ В ЭЛЕКТРОННЫЙ ВИД

Опыт работы

Перевод бумажных архивов в электронную форму интенсивно идет во всем мире. Не осталось в стороне и Комсомольское-на-Амуре Авиационное производственное объединение (КнААПО). Четыре года назад для централизованной обработки документов был создан Участок перевода информации в электронную форму (Участок сканирования). Работая в тесной связи с подразделениями объединения (отделы Главного Конструктора, Главного Технолога, Главного Метролога, Планово-экономический отдел, Сектор стандартизации, цеха основного и вспомогательного производства, бухгалтерия, отдел кадров и многие другие), ему удалось перевести в электронный вид более 200 000 листов бумажных документов формата А4.

Подразделения, воспользовавшиеся нашими услугами, освобождаются от рутинного труда, связанного с большими затратами времени и требующего освоения программных продуктов, которые могут им в дальнейшем и не понадобиться. Процесс перевода в электронный вид регламентируется временным положением, действующим на предприятии.

Все необходимые сведения заносятся в Карту заказа, которую с помощью инженера по АСУП Участка сканирования заполняет заказчик при сдаче документов на обработку. В Карте заказа указываются параметры форматирования, имена файлов и каталогов, особенности обработки

и адреса хранения готовой электронной формы.

Сегодня мы предлагаем заказчикам следующие виды электронных документов:

- качественные копии бумажной технической документации, доступные для редактирования в графических редакторах (*.tif, *.jpg и др.);
- чертежи в виде рисунков AutoCAD (*.dwg);
- эскизы векторного качества для приложений Windows и баз данных в формате Метафайл Windows (*.wmf);
- книги Microsoft Excel (*.xls);
- web-страницы (*.html);
- документы, подготовленные для печати на типографском оборудовании в формате Acrobat (*.pdf) и т.д.

Кроме того, мы размещаем информацию в базах данных корпоративной сети предприятия. Таким образом, заказчики, сдав в обработку бумажные документы (зачастую ветхие и плохого качества), получают аналоги этих документов в электронном виде, доступные целой группе пользователей.

Стимулом к созданию участка стала потребность сопровождать изделия на экспорт документами в электронном виде. Первыми заказами, выполненными Участком сканирования, были Технологические карты и Руководства по эксплуатации изделий, послужившие основой для создания интерактивных электронных технических руководств.

На создание первых книг общим объемом 10 000 форматов А4 конструкторам и технологам было необ-



♦ ЛВС – локальные вычислительные сети
КВС – корпоративные вычислительные сети

ходимо (по их подсчетам) около полутора лет, но благодаря новой технологии эта работа была выполнена за три месяца.

Сегодня производительность участка составляет около 6 000 листов формата А4 в месяц. Количество обработанной документации зависит от качества оригинала и выбранной технологии. Разработаны технологические процессы разной сложности: например, перевод в электронный вид *Документ Microsoft Word* простых текстов без рамок форм имеет сложность 1, а перевод чертежей в электронный вид *Рисунок AutoCAD* — сложность 4. Благодаря систематическому обучению персонала поддерживается высокая квалификация специалистов, что позволяет создавать качественные электронные версии различных документов.

Поступающие на обработку документы регистрируются в специальной базе данных "Scandoc". В этой базе содержатся сведения о заказчиках и исполнителях работы, технологическом задании, сроках выполнения, адресах размещения электронных версий документов в корпоративной сети и т.д.

Технические средства

Первые работы по переводу в электронный вид были выполнены на сканерах Vidar TruScan Select с TruInfo — эти сканеры и сегодня продолжают исправно работать, позволяя выполнять сканирование на отдельных листах в черно-белом изображении шириной от 150 до 965 мм и длиной до 10 метров. Возможность подбора различных настроек монохромного сканирования в программе TruInfo, удобный интерфейс с укрупненным просмотром сканируе-

мого изображения, небольшой размер конечного файла в формате TIF CCIT Group4 заслуживают самых добрых слов. В то же время были и определенные трудности. Например, нарушение выравнивания камер сканирования требовало полной переустановки программы TruInfo с получением у фирмы-поставщика нового кода авторизации. Неудовлетворительный результат на этих сканерах дает и сканирование в режиме Grayscale.

Для сканирования книг, технических журналов и сшитых документов, формат которых не превышает А3, мы с успехом используем цветные планшетные сканеры Paragon 1200 A3 SP PRO.

Чтобы сканировать цветные документы большого формата и даже очень плохого качества, мы приобрели широкоформатный цветной сканер Contex с программным обеспечением RasterID, которое предлагает различные варианты настроек (это повышает качество сканирования). Хорошие отзывы получила автоматизация процессов обработки. Кроме того, программа поддерживает технологию ActiveX, что еще больше расширяет наши возможности.

Наш опыт

Для большинства документов хорошее качество изображений при небольших размерах файлов обеспечивает сканирование с разрешением 300 dpi. Сканирование калек, мятых и разорванных чертежей удобно проводить при помощи специальных прозрачных пакетов, приобретенных нами у компании Consistent Software.

Быстро и качественно сканируются документы, обладающие высоким контрастом, сочностью цвета, однородностью и постоянной плотностью. В этом случае оцифрованное изображение занимает незначительный объем дисковой памяти и практически не требует дальнейшей обработки. К сожалению, процент таких документов невелик (около 30%).

Для других документов получе-

НОВОСТИ

Новый информационный ресурс
www.rasterarts.ru

В сентябре 2003 года открылся новый информационный ресурс, посвященный программам серии Raster Arts: сайт www.rasterarts.ru. Он адресован специалистам, работающим со сканированными изображениями технического назначения (чертежами, планами, схемами, картографическими материалами и т.д.) или планирующим приобретение и внедрение программно-аппаратного комплекса для сканирования, обработки, редактирования, векторизации документов, преобразования "бумажного" архива в электронную форму.

На сайте представлена информация о программах Raster Arts, областях их применения, технологиях решения актуальных задач, связанных с использованием сканированной документации в машиностроении, промышленности, картографии и т.д.

Отвечив на несколько вопросов on-line анкеты, можно получить рекомендации по выбору программного и аппаратного обеспечения, позволяющего наилучшим образом использовать в работе бумажные документы из архива. Для более подробного ознакомления с возможностями программ предоставлена возможность загрузки учебных материалов и демонстрационных версий.

Реализованы регистрация программ и оперативная техническая поддержка в режиме on-line, даны ответы на часто задаваемые вопросы, размещена информация о последних изменениях и обновлениях текущих версий — с возможностью загрузки этих изменений и обновлений.

Раздел "Статьи" предлагает материалы, посвященные использованию программных продуктов Raster Arts. В разделах "Машиностроение", "Промышленность", "Картография", "Прокладка коммуникаций", "Электронный архив" приведены варианты применения программ Raster Arts в соответствующих областях проектирования. Информация построена по схеме "задача → вариант решения → результат".

Задать интересующий вопрос, проконсультироваться с другими пользователями можно, посетив конференцию (конференция раздела "Векторизация и импорт данных" ресурса www.autocad.ru).



▲ Окно базы данных "Scandoc", форма Книги, вкладка Регистрация

НОВОСТИ

WiseImage и RasterID включены в каталог решений, рекомендуемых компанией HP

Компания HP (Hewlett-Packard) включила разработанные компанией Consistent Software программные продукты WiseImage (Spotlight и RasterDesk) и RasterID в свой мировой on-line каталог рекомендуемых решений (HP Global Solutions Catalog — <http://www.hpgsc.com>).

HP — крупнейший производитель товаров в различных областях электронной техники. Производственные мощности компании размещены в Англии, Франции, Германии, Италии, Нидерландах, Ирландии и Испании. HP располагает десятками научно-исследовательских центров в Европе, а также собственным центром стратегических исследований (Бристоль, Великобритания). Обороты компании постоянно растут, она уже более десяти лет участвует в формировании индексов DownJones и NASDAQ. Общая численность сотрудников — более 100 000 человек.

Каталог рекомендуемых решений публикуется на сайте HP, который регулярно посещают не менее 650 000 потенциальных покупателей, практически весь персонал HP и дилеры компании по всему миру. Именно здесь пользователи оборудования от HP могут получить исчерпывающую информацию о программах, совместимых с этим оборудованием. Компания HP осуществляет специальные программы по продвижению данного Internet-портала и переадресации на него заинтересованных пользователей, посетивших другие сайты HP.

WiseImage (Spotlight/Raster Desk)

WiseImage (Spotlight/RasterDesk) — флагманский продукт компании Consistent Software. Идеальное решение для сканирования, обновления (актуализации), редактирования, векторизации и печати широкоформатных растровых изображений. Мощный интеллектуальный редактор и векторизатор цветного, полутонового и монохромного растра.

RasterID

Комплексное решение для сканирования, обработки, архивирования и печати широкоформатных чертежей. RasterID обеспечивает возможность быстрого ввода и регистрации сканированных чертежей в системе управления электронным архивом. Помимо реализованного на аппаратном уровне высококачественного скоростного сканирования, поддерживается интеллектуальное повышение качества растровых изображений, их индексация/архивирование и печать.

ние качественного растра требует времени. При обработке такого документа используются разработанные компанией Consistent Software программные продукты серии Raster Arts. С программами RasterDesk Pro и Spotlight Pro мы работаем начиная с третьей версии и считаем их наиболее эффективными для улучшения качества монохромных растровых изображений. В четвертой версии появилась возможность выбора растровых объектов, что позволяет редактировать растр, используя привычные методы работы с вектором. Длительное время мы, руководствуясь пожеланиями заказчиков, обрабатывали чертежи, сохраняя их либо в растровом, либо в векторном виде, однако считаем, что пора переходить к гибридным чертежам, так как работа с ними наиболее целесообразна для быстрого создания электронных архивов. Действительно, вряд ли имеет смысл перерисовывать в AutoCAD качественные растровые фрагменты!

В программе Spotlight весьма полезен режим пакетной обработки, позволяющий обрабатывать пакет растровых файлов без участия оператора. Правда, этот прием можно использовать только для изображений, одинаковых по качеству.

Для чертежей плохого качества быстрее получить векторную копию растрового оригинала средствами CAD-программ, чем проводить очистку сильно загрязненного растра. Программы Raster Arts позволяют

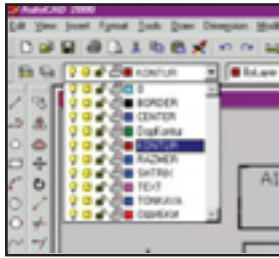
автоматически векторизовать растровые объекты, однако мы эту возможность используем редко, так как работаем для разных подразделений предприятия. Настройка параметров векторизации на разные виды часто меняющихся чертежей требует существенных затрат времени. Конечно, при больших объемах однотипных чертежей было бы целесообразно подобрать параметры векторизации, сохранить их в виде шаблонов и автоматизировать процесс перевода в электронный вид, но это требует участия специалистов подразделений, которые пока делают лишь первые шаги в освоении программ серии Raster Arts.

Обычно мы используем метод сколки: вставляем растр в рисунок AutoCAD в виде ссылки в масштабе 1:1, а затем средствами инструментов панели рисования получаем векторную копию растрового оригинала. Для работы с цветными документами (мы сканируем в цвете документы очень плохого качества) RasterDesk 5 и Spotlight 5 также предлагают набор инструментов, но их использование требует больших ресурсов памяти, которые есть не на всех рабочих местах, поэтому обработка таких документов замедляется. Цветные документы можно преобразовать в монохромные при помощи программы Color Image Processor и далее воспользоваться инструментами Spotlight, улучшающими битональные изображения.

Качественные тексты хорошо распознаются большинством OCR-



▲ Участок сканирования КНААПО



Слои
рисунка
AutoCAD

программ и требуют минимальной обработки.

Для текстовой документации плохого качества, которую трудно прочитать глазами, сканирование и распознавание нецелесообразны. Предпочтительнее ручной набор информации (затраты времени на сканирование, распознавание и редактирование могут оказаться больше, чем время ручного набора).

Информацию в электронном виде мы формируем так, чтобы при просмотре на экране компьютера ее расположение имело удобный для пользователя вид. Листы с книжным расположением информации должны иметь книжную ориентацию, а с альбомным расположением — альбомную. При выводе на печать соответствие бумажному оригиналу сохраняется путем настройки параметров принтера.

При наличии в бумажном оригинале технической документации символов, написанных или нарисованных от руки, мы определяем для них вариант замены из таблицы символов или используем качественно нарисованное электронное изображение данного символа. Для

рисования желательно выбирать векторный редактор и отслеживать, чтобы в электронной версии один и тот же символ не обозначался разными знаками. А для этого необходима единая библиотека символов предприятия. В бумажной документации, напечатанной в 70-80-е годы, символы чаще всего вписывались от руки — создавалась ситуация, когда один и тот же символ имел разное написание. Символы Windows имеют по несколько вариантов. При выборе того или иного из них наряду с внешним видом необходимо учитывать наличие библиотеки шрифтов, доступной на данном компьютере и в выбранной программе. Например, не все знаки, написанные в Microsoft Word (*.doc), будут отображаться в рисунке AutoCAD (*.dwg) или таблицах баз данных. Возникает следующая проблема: установку программного обеспечения и настройку компьютеров предприятия необходимо проводить так, чтобы одни и те же документы смотрелись на всех машинах одинаково и желательно в соответствии со стандартами на электронные документы.

Для CAD-программ необходимо использование (создание) типов линий по ГОСТ, единство текстовых и размерных стилей, систем слоев, выработка единых принципов формирования и сохранения чертежей (особенно при использовании гибридных технологий). Мы, например, для удобства работы в AutoCAD выбрали светло-серый цвет экрана.

Основной шрифт, принятый для использования на предприятии, — GOSTW. Но многие подразделения начали использовать AutoCAD раньше, чем были предприняты попытки стандартизации данного процесса. Следовательно, появился большой объем электронных чертежей с разнообразными настройками, удобными конкретным пользователям. Как результат, на сегодня существует большое количество текстовых и размерных стилей, типов линий и т.д.

Ряд проблем возникает при переводе в электронный вид текстовых документов, оформленных по ГОСТ. Текстовую часть документа нередко требуется поместить в рамки форм. Однако создать универсальные рамки, которые полностью печатались бы на любом принтере и при этом удовлетворяли требованиям стандартов — трудная задача.

Эту проблему можно решать по-разному. Создано много вариантов рамок в Word, Excel, в растровых и векторных редакторах. Но обидно тратить на это время да еще и подгонять результат под параметры принтера. К тому же нужно ли сохранять рамки для электронного вида? Можно, конечно, установить на все компьютеры специализированное программное обеспечение для работы с отсканированными формами, но насколько это целесообразно?..

Благодаря тесному сотрудничеству с компанией Consistent Software мы всегда в курсе ее новых разработок — к сожалению, не со всеми интересующими нас фирмами удалось достичь столь тесного контакта.

Из перспективных разработок Consistent Software нас заинтересовала программа PlanTracer — почти фантастическими возможностями создания чертежей в объеме из растровых рисунков. Хотелось бы попробовать ее на планировках наших зданий. Быть может, недалек и переход к моделям?

Надеемся, что с помощью таких партнеров, как Consistent Software, мы сможем более профессионально решать проблемы, возникающие в нашей работе.

Марина Авдеева,
инженер по АСУП,
Анатолий Чиркин,
инженер-программист
ИВЦ ОАСУП КнААПО
(Комсомольск-на-Амуре)



▲ Инженеры Марина Авдеева и Анатолий Чиркин