

2009

Cabinet NG, Inc



BATCH SCANNING AND DOCUMENT CAPTURE

Table of Contents

Introduction	3
Document Capture Synopsis	3
The Process	3
Scanning	4
Quality Assurance	4
Image Migration.....	5
Manual (Key from Image) vs. Automated Processing.....	5
Conclusion.....	6

Introduction

The need to quickly and efficiently organize, index, and file paper based documents in an Electronic Document Management System (EDMS) can be challenging. By implementing and properly configuring available software tools, this task can be managed effectively. Batch scanning is the commonly referred to process of turning paper into digital documents. The use of batch scanning in the proper manner can increase productivity and ensure success in the implementation of the EDMS. This paper will outline the batch scanning methodology to help in the understanding of the paper to digital image process.

Document Capture Synopsis

Whether we realize it or not, documents we deal with on a daily basis cost money to manage. The cost can be broken down into these primary functions: Document Creation, Document Capture, Document Workflow, Document Storage, and Document Retrieval. Effectively addressing these functions can save your business thousands of dollars in lost productivity, storage space, and efficiency.

Document Capture is an integral part of any solution to the paper problem we all face. So what is Document Capture? Simply put, it is turning paper into electronic format. So how do I capture my documents? The first thing that comes to mind is scanning the paper. For many of us, this is a logical first step because that is what we see daily, piles of paper covering every spare space we have available.

Document Capture usually consists of an input device (scanner), software to assist in sorting and identifying (indexing) the paper, and storage (disk space) to file the electronic documents for retrieval. It is also important to consider good document management software to manage the captured image.

Beyond simple document capture (imaging), many things can be done during the capture process to automate the process. Barcode recognition, database lookup, forms processing, and OCR are all advanced methods of capture that can boost productivity during the capture process. Finding a knowledgeable vendor to help analyze the capture process is critical to successfully implementing a solution.

So how do I get started? Review products that solve your problem. Make sure they are easy to use, easy to maintain, and provide the necessary tools to accomplish your goals. It is more important to find someone who understands the problem at hand and can effectively apply technology to solve the problem rather than simply buying hardware and software and hope it works.

The Process

Batch scanning typically falls into two basic categories; manual batch processing and automated batch processing. The two categories have different criteria and definition. Each has its place in the scanning realm and can complement each other. In fact, sometimes a combination of manual and automated processing is used to complete a scanning solution.

The “Document Imaging-Where to Begin” paper ([see: www.cabinetng.com/media/Document_Imaging-Where_to_Begin.pdf](http://www.cabinetng.com/media/Document_Imaging-Where_to_Begin.pdf)) defined the primary functions used in the scanning (capture) process as follows:

1. Document preparation
2. Scanning
3. Quality Assurance
4. Indexing or Classification
5. Migration to the document management solution

Preparation and Indexing were discussed in more detail in that paper. The discussion presented here will focus more on scanning, quality assurance, and migration to the EDMS.

Scanning

The scanning process involves the capture of documents. When considering scanning, several factors should be taken into account.

1. Get the right scanner hardware for the job. Based on the document criteria and processing requirements select the appropriate type of scanner. Without the correct hardware, the process may become frustrating and time consuming.
2. Select appropriate capture software. Review the documents to be scanned, necessary index values to be captured, and output requirements. Ensure the capture software meets the requirements.
3. Decide whether any of the indexing can be automated with zone OCR, database lookup, barcode recognition, etc.
4. Reliability, availability of parts and support, image quality, and feed capacities are all things to look for in a good scanner.

The scanner hardware itself becomes a key component and can make or break the success of imaging documents. Beyond scanning, the software that is coupled with the scanner for batch processing is the definitive factor which will determine success.

Quality Assurance

A key component of the batch processing software is the ability to provide a quality check on the documents and data which is captured. Keep in mind that once the document is scanned, the original likely will no longer be available. The electronic image must therefore be deemed of high quality as it will now become the ‘original’ document. Quality assurance is needed to review the documents that are scanned to make sure they meet the ‘original’ standard. An interesting point to make here is that in most cases, the scanned image is of better quality than the original due to the image processing capabilities of the capture software.

The features to look for in a good quality assurance module are:

1. The ability to append, insert, replace, rotate, flip, and delete pages.

2. The ability to provide a wide array of image processing capabilities such as deskew, border removal, cropping, page orientation, hole punch removal, etc.
3. Automatic image review, thumbnail viewing, page re-order, and image zooming.
4. Data correction and validation for indexed data.

These feature sets will provide the ability for the QA person to ensure accurate data and good image quality thereby insuring a good electronic 'original' document.

Image Migration

Once the scanning, QA, and indexing steps are completed, the images are typically ready for storage in the EDMS. The capture process should include a seamless mechanism for sending the images to the EDMS with the proper categorization as completed in the index step. Look for the ability of the batch processing software to define the output format of the document title and description, and ease of migration to the correct storage location in the EDMS system. Without proper output from the software, the documents may be misfiled, effectively making the electronic retrieval of the documents very time consuming. This would negate the primary purpose of an EDMS system.

Manual (Key from Image) vs. Automated Processing

The two primary types of document processing in batch mode are relevant depending on the documents being processed. As mentioned, many times a combination of manual and automated processing techniques are employed in a successful batch operation.

Manual processing typically is referred to as Key from Image processing. In this process, data indexes are defined and the data elements are keyed in by an operator who looks at the image on screen. This is the slowest and most time consuming method of batch image processing. In some cases, this may be the only method feasible. Another consideration is if it is faster to use this method, or spend additional time in the document preparation stage applying barcodes and configuring database lookups to assist in the index process.

When considering manual processing, look at the amount of data that will be required in the index step and the number of indices needed. It is best practice to keep the data input to a minimum so the input can be done as quickly as possible. Limiting the index values is critical to speed of document processing. There are functional tools available in most capture products to assist and speed up the manual entry process. Some of the functions to look for are template based document indexing, list box selection, field formatting or masking, data types, and OCR assisted indexing.

Automated processing can greatly increase the document throughput in the processing of images. The more that can be automated the better. Features of the software that enhance automated processing include zonal OCR, barcode recognition, and database lookup. By using some or all of these features, the indexing of documents in many cases can be fully automated.

Conclusion

The batch scanning and document capture process is very useful in processing many documents rapidly. It is important to review the document types, decide on the best scanning hardware and software to use for the capture process, and employ the most effective manual and automated data entry schemes. By implementing these techniques, a few hundred to thousands of pages of paper can be processed by a single scan station on a daily basis. It is critical to choose a knowledgeable consultant to guide you through the process of batch scanning. If these rules are followed, it will ensure a successful implementation. Visit www.cabinetng.com for more information about document management.