

SimpleIndex - DOCUMENT MANAGEMENT AUTOMATED



How many clicks *does* it take to scan your documents? Many small business scanning applications lack time-saving automation features like OCR and barcode recognition found in enterprise software. **SimpleIndex** is the most affordable fully-featured document capture solution available.

SimpleIndex provides easy 1-click scanning, unattended processing and even supports MS Office, PDF, audio and video files. It is the fastest, most affordable way to organize all of your digital files or index them for your document management system: it's a scanning program that not only scans documents, but reads the content. It does its work using a number of advanced techniques.

Optical Mark Recognition

The screenshot shows the SimpleIndex software interface. The main window displays a form for document processing. The form includes fields for Student's name (Annie May), Date (2006-04-25), Current Grade (F), and Teacher (Ms. Thompson). Below these fields, there are checkboxes for 'Help your grade improve by' with options: 'Completing homework assignments on time' (checked), 'Participating in class' (unchecked), 'Doing Extra Credit' (checked), and 'Other:'. On the right side, there are dropdown menus for 'Student's Name', 'Student's Grade', 'Teacher's Name', 'Help your grade improve by', 'Completing homework assignments on time', 'Participating in class', 'Doing extra credit', 'Date', and 'Date'. At the bottom right, there are buttons for 'Save All' and 'Save Index'. The status bar at the bottom indicates 'Source: IQR: 00119', 'File: 1 of 10', and 'Page: 1 of 1'.

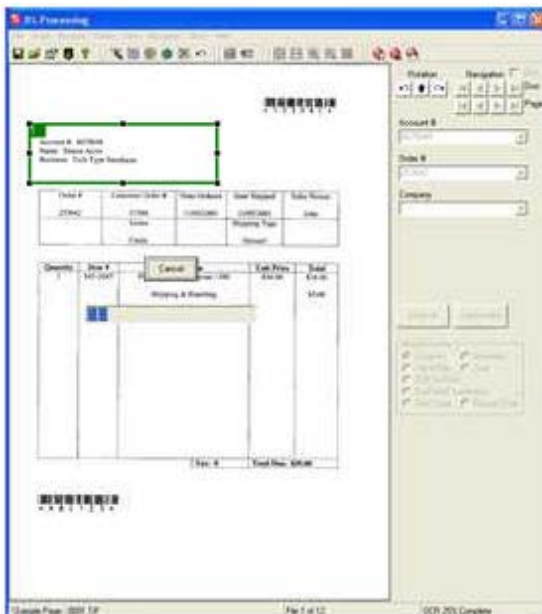
Optical Mark Recognition lets you define check box regions on scanned images. OMR is very fast and can be used for a variety of applications:

- Business reply mail
- Simple surveys
- Separate multi-page documents
- Document routing control
- Verify presence of signatures

To configure OMR, use an unfilled form to obtain baseline counts of how many black "pixels" are in the box. When processing, **SimpleIndex** compares the amount of black in each image to the baseline value to determine if the box is checked or not.

With OMR, it is very important that the check boxes appear in the same place on every scan, and that other text on the document does not move into your check box zone. For best results, use large boxes with plenty of white space around them.

Zone OCR



Zone OCR is used to read document indexes or tags from text on the page. Zone OCR is a great way to automate the data entry associated with scanning documents.

However, there are several limitations to TRADITIONAL zone OCR that must be overcome:

- Index information must be in the exact same place on every page
- Documents shift and skew during scanning, causing the zones to not line up
- If surrounding lines or text on the document are too close, they can encroach on the zone

Dynamic OCR

SimpleIndex overcomes these limitations by using Dynamic OCR technology to locate the desired text even when it moves around on the page. Our simplified version of Dynamic OCR works great for many types of documents at a fraction of the cost of other solutions.

- Index information can appear anywhere on any page
- Unwanted characters are automatically ignored
- Find unique patterns of letters and numbers using Template Matching (Social Security #, Date, etc.)
- Use Dictionary Matching to find a value from a list of possible values (Vendor Name, Document Type, etc.)

With **SimpleIndex** you can use large zones – even full page - that give a wide margin for error. Template and Dictionary matching are then used to extract the 7-digit Account Number, 6-digit Order Number and Company Name. **SimpleIndex** discards the surrounding text and keeps the correct value.

Another common example is finding a unique identifier, for example a social security number, that could appear anywhere on the page. Simply enter the template ###-##-#### and **SimpleIndex** will search the full OCR text until it finds a match. Since only one social security number is likely to appear on the page, a match on this pattern is almost certainly the required value.

With dictionary matching, you can give **SimpleIndex** a list of possible values and it will automatically search the zone or page for each possible value until it finds a match.

Many dynamic forms processing applications can be implemented using these simple algorithms. This makes **SimpleIndex** far more versatile than other zone OCR solutions that require the index value to be in the exact same location on every page. Yet **SimpleIndex** costs only a fraction of the price!

Barcode Recognition



This is the most efficient way to capture index data printed on documents. Some documents already have key information in barcode format on them. In many cases adding a barcode to a document is as simple as changing or adding a font. Adding barcodes to new documents is preferable as all the index data is on the document at the time it is created and in a format that can be read with near 100% accuracy.

As an alternative to placing barcodes on the individual documents, it is possible to print out a barcode cover page and place it on the file before it is scanned.

Barcode recognition can also be useful when you have documents with a variable number of pages that will all receive the same index values. If it is not possible to generate an indexed coversheet for these at the time they are created, a generic barcode coversheet can be used to separate the scanned images into multi-page files, one for each document. A second process can then be used to index these images one file at a time instead of one page at a time, greatly increasing throughput.

- Read barcodes printed on scanned paper documents
- Read barcodes embedded in PDF files
- Automatically rename files based on barcodes
- Export barcode data to CSV file or any database
- Separate multi-page documents with cover pages
- Recognize 2D formats like PDF417, DataMatrix, Aztec and QR Code
- Recognize 30 different 1D barcode formats
Code 39, Codabar, UPC, Code 128, EAN 13, 2 of 5, etc.
- Recognize postal barcodes like Planet, PostNet, Royal Post and Australian Post

Open Database Integration

Open database integration is a powerful feature of **SimpleIndex** and one that furthers its interoperability with custom programs. Instead of using a proprietary database, **SimpleIndex** allows you to map its index fields to cells in any database table. It can be configured to create new records, update existing ones or retrieve them for viewing. Using these three basic database functions, it is able to interface with and operate on any database.

The Autofill feature of **SimpleIndex** is an easy way to associate many index fields with one document without retyping data that already exists in another database. Autofill uses a database lookup to retrieve records that match a key value entered by the user. Blank index fields are then filled in automatically with the data from this lookup. The result is a document database with many different possible search fields, of which only one needed to be entered during scanning.

The key field may be typed by the user, or it may be read from the document automatically using barcode recognition or OCR. The lookup is performed either when the user changes this field or when the index values are saved. If the lookup finds multiple matching records, the user will be notified and the first set of values will be used by default.

When used with pre-index batches, key information can be read automatically from barcodes or OCR and matched to database records with a single click. Search on up to 99 index fields without a single keystroke!

SimpleIndex features "Match and Attach" mode, which combines the power of the Autofill feature with the ability to update existing records in your database. This lets you "attach" images to existing database records by "matching" them with values from **SimpleIndex**. With other systems you must create a temporary database and write a custom import to match the scanned images to existing records. When indexing with OCR or barcode recognition, the entire process can be accomplished with our 1-click interface.