

OPTICAL CHARACTER RECOGNITION

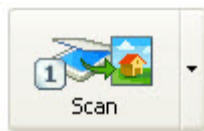
ABBYY FINEREADER

What is an OCR application?

Scanned pages you get from your scanner look very much like the original text, but they cannot be edited, say, in Microsoft Word or Excel. This is because they are essentially images, or "photos", of printed pages. To be able to edit such texts, you have to convert them into a text format. This is where an Optical Character Recognition (OCR) application comes in. It recognizes, or "reads," the scanned page and saves it in one of the available text formats or as an Excel file or even as a PowerPoint presentation.

How to Recognize a Document in Less than a Minute

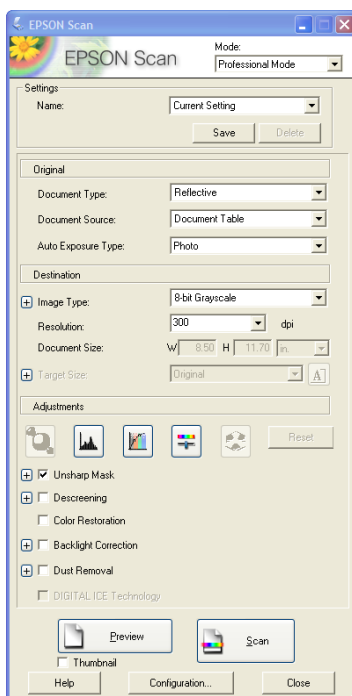
Turn on the scanner and place your document onto the scanner face down. Start ABBYY FineReader 6.0 Sprint from **Start/Programs/ABBYY FineReader 6.0 Sprint**.



1. Click the **1-Scan** button to start scanning. Adjust the scanning parameters to match the screenshot below. Once FineReader has scanned the document, the document image will be displayed in the **Image** window.

Tip: Images will be scanned and recognized faster if you select 300 dpi resolution and "gray" scan mode.

Setting scanning parameters



Recognition quality depends greatly on the scanned image quality. The image quality may be adjusted by setting the main scanning parameters: resolution, scan mode and brightness.

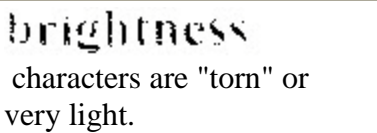
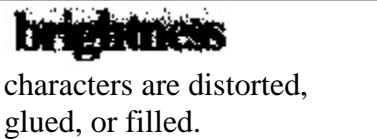
The main scanning parameters are:

- **Image type - 8-bit Grayscale**
- **Scan mode - gray.**
Scanning in grayscale mode will yield the best recognition results. If you scan your images in grayscale, the application tunes the brightness automatically.
- **Scan mode - black and white.**
The black and white scan mode enables the system to scan at a higher speed, but at the same time some character information is lost. This may have a negative effect on the recognition quality of the documents of medium and low print quality.

- **Scan mode - color.**
Select the color scan mode for scanning and retaining color documents. It scans and recognizes color documents with pictures, color text and background, and retains the color in electronic documents.
- **Resolution** - use 300 dpi resolution for regular texts (font size 10pts or greater) and 400-600 dpi resolution for texts set in smaller font sizes (9pts or less).
- **Brightness** - in most cases the medium brightness value (50%) will do.
- Some documents scanned in black and white mode may require some additional brightness tuning.

Tips on Brightness Tuning

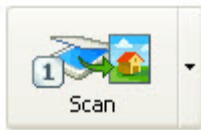
- The scanned image has to be legible.
 - an example of a good image (from the OCR point of view)
- If you see that the scanned image is far from perfect (characters are glued or torn), consult the table below to find how you can improve the image quality.

| Your image looks like this: | Possible remedy: |
|---|---|
|  <p>brightness characters are "torn" or very light.</p> | Try to decrease the brightness (it will make the image darker). Try to scan it in gray mode (the brightness autotuning is used in this case). |
|  <p>brightness characters are distorted, glued, or filled.</p> | Try to increase the brightness (it will make the image brighter). Try to scan it in gray mode (the brightness autotuning is used in this case). |

- **Brightness** - a scanning parameter reflecting the contrast between black and white image areas. Setting correct brightness increases the recognition quality.
- **Image type** - a scanning parameter determining whether an image must be scanned in black and white, gray or color palette.
- **Resolution** - a scanning parameter determining how many dpi to use during scanning. Resolution of 300 dpi is good for texts set in 10pt font size and larger, 400-600 dpi is good for texts of smaller font size (9pt and less).

Note: Scanning at 400-600 dpi resolution (instead of the default 300 dpi) or scanning in gray or color (instead of black & white) mode takes more time. For certain scanners, it may take up to four times longer to scan a document in 600 dpi than scanning in 300 dpi.

Tip: If you already have a scanned image of your document, simply open it in FineReader and go to step 2.



1. To open an existing image, click on the arrow to the right of the Scan button, select Open Image... and then select the required image file in the dialog that opens.



2. Click the **2-Read** button. The recognized text will soon be displayed in the **Text** window.

Tip: Before you start scanning, make sure that the right recognition language has been selected. Recognition languages can be selected from the **Language** menu.

3. Click the **3-Save** button to save the results.



The **Save Wizard** will help you save the recognition results in any of the available formats.

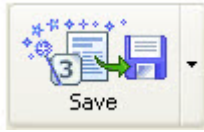
Saving and Exporting Recognition Results

Recognition results can be saved to a file, exported to an external application without saving them to disk, or copied to the Clipboard. Either all the pages or only the selected ones only may be saved.

FineReader Sprint can export recognition results to the following applications:

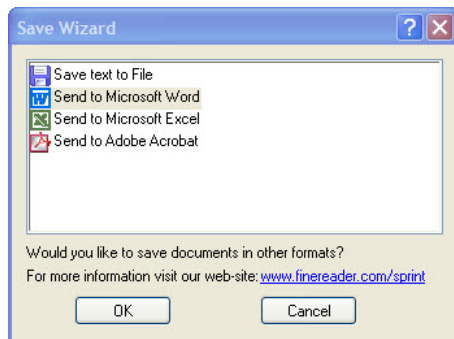
Microsoft Word 6.0, 7.0, 97 (8.0), 2000 (9.0), 2002 (10.0) and 2003 (11.0); Microsoft Excel 6.0, 7.0, 97 (8.0), 2000 (9.0), 2002 (10.0) and 2003 (11.0); Adobe Acrobat Reader.

Saving Recognized Text



The **3-Save** button saves recognition results in one of the available formats or exports them to an application of your choice. The appearance of the icon depends on the current saving mode. The **Save** button caption displays the name of the currently selected destination program.

To save the recognized text, click the arrow to the right of the **3-Save** button and select the required item from the menu that opens.



Once you have saved the recognized text or exported it to another application, the **3-Save** button icon will adopt the icon of the last saving or export mode. If you wish to save or export another portion of recognized text using the same mode, simply click the icon without browsing the button's menu.

Copying Recognized Text to Clipboard

1. Select a text fragment in the text editor.
2. Select **Copy** in the **Edit** menu or press **CTRL+C**. This will place the selected text into the Clipboard.
3. Switch to the desired application. Click somewhere in the document where you wish the copied text to be inserted and select **Paste** in the **Edit** menu or press **CTRL+V**.