## **About Microsoft Office Document Imaging**

With Microsoft Office Document Imaging, you can use a scanned document as easily as other Microsoft Office documents on your computer.

You can:

Scan both single and multi-page documents.

Perform optical character recognition (OCR).

Copy and export recognized text to Microsoft Word.

Read a scanned document as easily as a paper document.

Search for text within scanned documents as easily as you can in Word documents.

Reorganize scanned document pages as easily as rearranging papers in a folder.

Send scanned documents to others in e-mail.

## When should you use Microsoft Office Document Imaging?

Use Microsoft Office Document Imaging when you want to:

Access and work with the text or Tagged Image File Format (TIFF) images contained in scanned documents or faxes.

Extract the text from existing scanned documents or faxes for archiving.

Rearrange the page order of multi-page scanned documents.

Read scanned documents quickly and easily on the screen.

### **Scanning and imaging**

Microsoft Office Document Imaging actually has two components — a scanning component and an imaging component — listed separately in the Windows Start menu as Microsoft Office Document Scanning and Microsoft Office Document Imaging.

The first component controls the scanning of documents into your

computer by using any installed scanner. This is done by using scanning presets, which control your scanner by using settings that are optimized for specific purposes. For example, the **Black and white** scanning preset is ideal when scanning pages of text for OCR, while the **Color** scanning preset is best for scanning full-color pictures or artwork. Also, OCR is automatically performed by default on text documents immediately after scanning, and you can easily scan multiple pages into a single file.

The second component makes it easy to view scanned documents on the screen, rearrange multi-page documents, select and manipulate recognized text, and send documents to others in e-mail.

## **About scanning**

Microsoft Office Document Imaging combines and manages the separate tasks of scanning documents and performing optical character recognition (OCR) on scanned documents. In most cases, Office Document Imaging streamlines these tasks by controlling the scanner software and initiating the OCR process without further intervention.

### **Scan presets**

When you scan a new document, Office Document Imaging provides several presets, offering easily selectable settings for common scanning tasks. Each preset is optimized for different scanning needs. You can customize presets, or you can create your own presets for specific purposes. Built-in presets include the following:

**Black and white** Scans in monochrome at 300 dots per inch. This preset is designed for optimum OCR results when scanning black text on white paper, or when scanning line art. Scanning is faster than with other presets, and the resulting image files are small.

**Black and white from color page** Scans in grayscale at 300 dots per inch, but saves the resulting file in monochrome format. This preset is designed to provide maximum text resolution for OCR when scanning difficult originals containing colored backgrounds or colored text.

**Grayscale** Scans in grayscale at 200 dots per inch. This preset is useful when scanning pages containing continuous-tone, black and white images (such as photographs) and text, or colored text. This preset creates somewhat larger image files.

**Color** Scans in color at 150 dots per inch. This preset is designed for scanning full-color documents. Scanning is slower, small text may not be legible enough for OCR, and the resulting image files are large.

In addition to resolution and scan type, each preset controls the following settings:

Page size

Default folder for saving files

Language used for OCR

Save multi-page documents together or as separate pages

Automatically perform OCR after scanning

Automatically create file names

Automatically adjust pages scanned upside down or out of alignment

#### Scanner drivers

The first time you scan a new document, Office Document Imaging may check your scanner and ask you to scan a test document. This process might take a while, but it is recommended that you complete the operation so that Office Document Imaging can properly identify your computer hardware. If possible, Office Document Imaging determines the necessary settings, and if not, your scanner driver dialog box is displayed each time you scan. If you update drivers or install a new scanner, you can rerun the scanner testing procedure.

If possible, Office Document Imaging manages the scanning process for you. If you want to adjust settings yourself, you can choose to display your scanner driver dialog box each time you scan. To do so, click **Scan New Document** on the **File** menu, click the **Scanner** button, and then select the **Show scanner driver dialog before scanning** check box. If this option is not available, Office Document Imaging displays your scanner driver dialog box each time you scan.

#### **Automatic document feeders**

Automatic document feeders (ADFs) operate in different ways, depending on the type of scanner. Flatbed scanners sometimes have add-on feeder attachments, while multi-function scanners usually have integral feeders. The first time you scan, Office Document Imaging performs a test on your scanner and asks you to load pages into your ADF. If possible, Office Document Imaging detects the ADF for you.

#### **Notes**

If you are using Microsoft Windows Millennium, using a Windows Image Acquisition (WIA) scanner driver with an ADF is not recommended. For

best results, install the TWAIN scanner driver supplied by the scanner manufacturer.

If you want to check your settings or turn off ADF support, click **Scan New Document** on the **File** menu, click the **Scanner** button, and then select or clear the **Use automatic document feeder** check box. If this option is not available, Office Document Imaging displays your scanner driver dialog box each time you scan.

When scanning multi-page documents using an ADF, Office Document Imaging usually combines all the scanned pages into a single file on disk. If you prefer to save each page separately, click **Scan New Document** on the **File** menu. In the **Microsoft Office Document Scanning** dialog box, click **Preset options**, and then click either **Create new preset** or **Edit selected preset**. On the **Page** tab of the **Preset Options** dialog box, select the **Save each page as a separate document** check box.

TWAIN-emulation drivers will not appear in the **Scanner** list in the **Choose Scanner** dialog box.

## **About Tagged Image File Format (TIFF)**

TIFF is a commonly used format for various imaging applications, including those that scan and fax. Microsoft Office Document Imaging uses the TIFF format exclusively, utilizing the format's capability to contain text recognized by optical character recognition (OCR). When you scan new documents, they are saved in TIFF format (with a .tif extension), and any OCR text is stored in the TIFF file along with the image.

You can open and edit TIFF files created with Office Document Imaging by using many other graphics applications. When you do so, any OCR text that the file contains is lost. You will have to rerun OCR if you want to access the text in the TIFF file again in Office Document Imaging.

Technical details

# **Command line options**

You can use command line options when you start Microsoft Office Document Imaging from either the **Run** command (**Start** menu) or from a command prompt.

Command	Result
mspview <path:filename.tif></path:filename.tif>	Starts Office Document Imaging and opens the specified file.
mspview -r <path:filename.tif></path:filename.tif>	Starts Office Document Imaging and opens the specified file in Reading view.
mspview -o <path:filename.tif></path:filename.tif>	Starts Office Document Imaging, performs optical character recognition (OCR) on the specified file, and then saves the file with the new OCR text.
	This command has no effect if OCR has already been performed on the file.
	This command can be used for batch processing when you need to recognize text in multiple TIFF files.
	Notes
•	If the image is monochrome, it will be compressed in TIFF CCITT G4 format.
•	If the image is grayscale or color, it will be compressed in TIFF 6.0 JPEG format.
•	If the image is over 900 dots per inch (dpi), it will be saved at 900 dpi.
mspview -f <path:filename.tif></path:filename.tif>	Starts Office Document Imaging, performs OCR on the specified file, and then saves the file with the new OCR text.
	Use this command if you want to perform OCR on a file that has already had OCR performed on it.

This command can be used for batch processing when you need to recognize text in multiple Tagged Image File Format (TIFF) files.

#### **Notes**

- If the image is monochrome, it will be compressed in TIFF CCITT G4 format.
- If the image is grayscale or color, it will be compressed in TIFF 6.0 JPEG format.
- If the image is over 900 dpi, it will be saved at 900 dpi.

For example, the following command opens Mypic.tif in Reading view.

mspview -r "C:\My Documents\mypic.tif"

Return values

# **Keyboard shortcuts**

You can carry out commands in Microsoft Office Document Imaging using the keyboard instead of the mouse.

- Keys for editing
- Keys for file operations
- ▶ Keys for navigating and using Help
- ▶ Keys for navigating in the page pane
- ▶ Keys for navigating in the thumbnail pane
- ▶ Keys for navigating in Reading view

# **Work with scanning presets**

Presets are collections of settings used to initiate specific types of scans. When you create or edit a preset, you control the settings for that preset only. All other presets remain unaffected. This way, you can create presets for specific purposes. Microsoft Office Document Imaging comes with several built-in presets for common scanning tasks.

- ▶ Edit an existing preset
- Create a new preset
- Delete a preset

### Scan a document

- 1. Position the document on your scanner.
- 2. On the File menu, click Scan New Document.
- 3. In the **Microsoft Office Document Scanning** dialog box, select the options you want.
- For information about scanning options, click **Help**.
- To select additional scanning options, click **Preset options**, and then click either **Create new preset** or **Edit selected preset**.
- If you want to scan double-sided pages, select the **Original is double sided** check box. If you have an automatic document feeder (ADF) and you have more pages than will fit into the feeder, click the **Prompt for additional pages** check box.
- 4. Click Scan.

If you selected the **Original is double sided** or **Prompt for additional pages** check box, the **Microsoft Office Document Scanning** dialog box reappears as necessary.

5. To save the file, click **Done** after the last page is scanned.

Note By default, optical character recognition (OCR) is performed automatically when you scan documents. To turn off this feature, click Scan New Document on the File menu, and then click either Create new preset or Edit selected preset. On the Processing tab of the Preset Options dialog box, clear the Use OCR to recognize the text of the scanned image check box.

### About the scanner test

The first time you scan a document, Microsoft Office Document Imaging may check your scanner for compatibility. This is necessary if you want Office Document Imaging to control the scanner settings for you, and if your scanner is not listed in the database of scanners that is installed with Office Document Imaging.

During the test, Office Document Imaging attempts to determine the capabilities of your scanner, which will allow you to use the built-in scanner presets. If Office Document Imaging is unable to determine these settings, your scanner driver dialog box will appear each time you scan.

- Scanner drivers
- Bypassing the built-in scanner presets
- Updating the scanner driver database file
- Submitting scanner driver log files to Microsoft
- ▶ Rerunning the scanner test
- Automatic document feeders

# Scan a color document or photograph

When you scan and perform optical character recognition (OCR) on color documents or photographs, you can optimize scanner settings either for accuracy (when scanning for text recognition) or for maximum resolution (when scanning photographs).

- Scan a color document
- Scan a color photograph

# Scan a document of poor quality

You can use scanning presets to maximize the accuracy of optical character recognition (OCR) on documents of poor quality, whether they are in color or black and white.

- Scan a document for optimum OCR
- ▶ Increase the scanning resolution

## Scan a foreign-language document

By default, Microsoft Office Document Imaging performs optical character recognition (OCR) automatically when you scan a document, using your computer's normal language settings. Office Document Imaging can perform OCR on documents in any installed language dictionary.

- ▶ Scan and recognize text in a foreign-language document
- ▶ Recognize text in a previously scanned foreign-language document

# Scan multiple pages

You can scan double-sided documents, multi-page documents, or multiple single-page documents in one continuous operation. You can save the individual pages in a single Tagged Image File Format (TIFF) file, or you can save each page as a separate file.

- Scan multiple pages into a single file
- Scan multiple pages into separate files
- ▶ Tip

## **Align scanned images**

Use this procedure to correct the orientation of pages that were scanned upside down, sideways, or out of alignment.

- 1. On the **View** menu, click **Options**.
- 2. Select the **Auto rotate** and **Auto straighten** check boxes, and then click **OK**.
- 3. On the **File** menu, click **Recognize Text using OCR**. If you've already performed optical character recognition (OCR) on the current document, a confirmation message appears.

▶ Tip

### Rescan a document

Use Microsoft Office Document Imaging to rescan previously scanned documents if you want to try a different scanning preset for better optical character recognition (OCR).

- 1. Position the document on your scanner.
- 2. On the **File** menu, click **Scan New Document**, and then select the desired options in the **Microsoft Office Document Scanning** dialog box.

For more information about scanning options, click Help.

3. Select the **Original is double sided** or **Prompt for additional pages** check boxes, if appropriate.

For additional scanning options, click **Preset options**, and then click either **Create new preset** or **Edit selected preset**.

4. Click Scan.

If you selected the **Original is double sided** or **Prompt for additional pages** check boxes, the **Microsoft Office Document Scanning** dialog box reappears after each page is scanned.

5. After the last page is scanned, click **Done**.

## **About optical character recognition (OCR)**

Optical character recognition (OCR) translates images of text, such as scanned documents, into actual text characters. Also known as text recognition, OCR makes it possible to edit and reuse the text that is normally locked inside scanned images. OCR works using a form of artificial intelligence known as pattern recognition to identify individual text characters on a page, including punctuation marks, spaces, and ends of lines.

OCR can be performed in three scenarios:

**Automatic** OCR runs automatically each time you perform a new scan, unless you change the scanning presets.

**Manual** Run OCR manually for documents that were scanned using another program.

**Indexing** Indexing is a system service that helps you to quickly find files on your computer using text searches. When you perform OCR on Tagged Image File Format (TIFF) files, recognized text is available to the index, making it possible to find relevant TIFF files when you search. You can index any and all TIFF files on your computer.

- Setting OCR options
- Optimizing OCR accuracy

#### **About OCR international issues**

When optical character recognition (OCR) is performed on a scanned document, text is recognized using sophisticated pattern-recognition software that compares scanned text characters with a built-in dictionary of character shapes and sequences. The dictionary supplies all uppercase and lowercase letters, punctuation, and accent marks used in the selected language.

By default, Microsoft Office Document Imaging uses the dictionary for the same language that your computer's operating system uses. You can easily change the dictionary to scan documents in other languages.

**Note** Office Document Imaging does not provide OCR dictionaries for Hebrew or Arabic.

The OCR language dictionary is used in the following scenarios.

- Scanning new documents using Office Document Imaging
- Running OCR on previously scanned documents
- Adding foreign language text to the index for fast file searches

#### **About indexing text in TIFF files**

When Microsoft Office Document Imaging performs optical character recognition (OCR) on a scanned document, the text is stored within a Tagged Image File Format (TIFF) file when you save it.

This text is available not only when you open the document in Office Document Imaging, but also when you search for files using other Microsoft Office programs, or when you use Microsoft Windows search features.

#### The indexing service

Indexing is a special service that enables fast file searches on your computer. Text found in files on your computer is added to the index, which also stores a reference to the file where the text was found.

Without indexing, you can search for words only in files that have had OCR performed on them — either automatically at scan time, or manually from the **File** menu. With indexing, you can search for any TIFF file based on the words it contains.

The indexing service is implemented in different ways, depending on the operating system you are using.

- Windows 2000
- Windows 98 and Windows Millennium

If the indexing service is turned off, only the following text is available in TIFF files that have not had OCR performed on them: the file name itself and any file properties that might be available.

#### Indexing TIFF files without embedded OCR information

OCR is automatically performed on other TIFF files that exist on your computer, making the text available to the indexing service for file searches. In this case, the OCR text is stored only in the index, not within the TIFF files. This process takes several seconds per TIFF file encountered.

To turn off automatic indexing, click **Options** on the **View** menu, and then click **Indexing Service**. In the **Indexing Service** dialog box, clear the

# Use OCR to recognize the text in TIFF files when indexing check box. Indexing in other languages

If you want to index documents in languages other than your computer's default language, you can select a dictionary from the **OCR Language** list in the **Indexing Service** dialog box.

## Change the language for optical character recognition

The optical character recognition (OCR) feature of Microsoft Office Document Imaging can recognize text in any installed language.

- 1. On the View menu, click Options.
- 2. In the **OCR Language** box, select the language you want to use.

#### **About upgrading Optical Character Recognition**

## Special offer for Microsoft Office Document Imaging customers

As a Microsoft Office Document Imaging customer, you are entitled to upgrade to ScanSoft's OmniPage Pro OCR software at a special price. To learn more about OmniPage Pro OCR and how it can enhance your scanning capabilities, visit <u>ScanSoft's Information Page</u>.

**Note** The hyperlink in this topic goes to the Web. You can switch back to Help at any time.

#### Why upgrade?

When scanning documents with Microsoft Office Document Imaging, you can capture the words from a page, but you lose the layout and format of your original document. With OmniPage Pro OCR, your scanned documents retain their original formatting. Simply drag the document directly to your favorite Microsoft Office application, and you've got a fully formatted, editable file!

ScanSoft's OmniPage Pro achieves an accuracy rate of over 99%. It allows you to scan documents directly into Microsoft Office applications with just the click of a button, which makes the conversion of paper to editable text so fast and simple that you'll use it instead of retyping. This means you save time. No retyping, correcting, or reformatting — you simply scan and go!

#### **About ScanSoft's OmniPage Pro OCR**

Features and benefits:

ScanSoft's most powerful recognition engine has a proven accuracy rate of better than 99%.

True Page® retains color pictures, graphics, illustrations, and artwork, and accurately places them into the digital document.

ScanSoft's intuitive, user-friendly interface makes scanning simple. You

can:

Scan paper spreadsheets directly into Microsoft Excel.

Convert documents into Microsoft Word format.

Switch paper documents into HTML by using Microsoft FrontPage.

Add paper documents to your Microsoft PowerPoint presentations.

Powerful table-editing tools improve accuracy and give you more flexibility and control when working with cells. Word tables convert as true cell tables!

ScanSoft's AutoOCR Toolbar™ guides you through the entire optical character recognition process.

Voice Read-Back allows you to double-check for errors before exporting to Word, Excel, or another text-based application by reading OCR results aloud to you. Voice Read-Back uses the Microsoft Speech Application Programming Interface (SAPI), the most widely supported speech API in use today.

The benefit of using ScanSoft's OmniPage Pro? Less editing. Less proofing. Less work. More speed. They're results you really have to see to believe!

To order the upgrade, visit <u>ScanSoft's Information Page</u> or call 888-781-1189.

ScanSoft, Inc. 9 Centennial Drive Peabody, MA 01960 www.scansoft.com

#### Insert pages into a document

You can insert pages from a Tagged Image File Format (TIFF) file into any document that is currently open in Microsoft Office Document Imaging.

- 1. On the File menu, click Insert File.
- 2. Find the TIFF file (.tif) you want to insert, and then select it in the **Insert File** dialog box.
- 3. Click the **Insert** button, and then select the option you want.
- 4. Click OK.

## Insert pages using copy and paste

You can copy a page and paste it either into another Microsoft Office Document Imaging document or within the same document.

- Insert a page
- Insert multiple pages

## Move pages to a new document

You can move pages from an existing Microsoft Office Document Imaging document into a new document.

- ▶ Move a single page to a new document
- Move multiple pages to a new document

## Split a document

Use this procedure to divide multi-page documents into smaller or more meaningful pieces.

- Split a document by moving pages
- Split a document by cutting and pasting

## Copy and paste pages or text

- Copy and paste pages
- Copy and paste text
- ▶ Copy and paste images
- Copy and paste page contents
- ▶ Paste pages from another program

### Rearrange pages

You can change the order of pages in a multi-page Microsoft Office Document Imaging document, or move pages between documents.

- Rearrange pages by dragging
- Rearrange pages by cutting and pasting
- ▶ Copy or move pages between two open documents

## **Rotate pages**

You can change the orientation of scanned pages by rotating them 90 degrees to the left or right, or 180 degrees.

- Rotate a page
- Rotate multiple pages

## **Delete pages**

- Delete a page
- Delete multiple pages

#### **About Reading view**

Reading view is a special viewing mode that makes it easier to read scanned documents on the screen. When activated, Reading view presents the current document and hides most nonessential screen elements, including the Microsoft Windows taskbar.

To switch to Reading view, click **Reading View** on the **View** menu.

- ▶ The **Reading View** menu
- Navigating in Reading view
- Selecting text in Reading view
- Closing Reading view

**Note** To provide an unobstructed reading area, the mouse pointer is hidden until needed. When you stop moving the mouse, the pointer disappears after five seconds. To show the pointer again, move the mouse.

#### **Read a document**

For easier on-screen reading, switch to Reading view.

- 1. On the View menu, click Reading View.
- 2. Use the following keyboard shortcuts to read the document.

Press	То
SPACEBAR or PAGE DOWN	Move to the next section
BACKSPACE or PAGE UP	Move to the previous section
UP ARROW and DOWN ARROW	Scroll the page in that direction
LEFT ARROW	Move to the previous page
RIGHT ARROW	Move to the next page
HOME	Move to the top of the page
END	Move to the bottom of the page
CTRL+HOME	Move to the first page
CTRL+END	Move to the last page
ESC	Return to the page pane

#### About the page pane

Use the page pane to view a representation of each page in a Microsoft Office Document Imaging document at an easily readable size. When you click an image in the thumbnail pane, that page appears in the page pane.

You can view or hide the page pane by opening the **View** menu and then clicking **Page Pane**.



Thumbnail pane

Page pane

You can select and copy text in the page pane. Dragging the mouse in the page pane selects any recognized text within the selection. Only whole words are selected. If only part of a word lies within the selected text, the entire word is selected.

**Note** You can select text in the page pane only if optical character recognition (OCR) has been performed on a document.

#### Selecting text in the page pane

After OCR has been performed on a document, you can select text in the page pane by dragging, just as you can in other Microsoft Office programs.

When you drag the mouse to select text, only whole words are selected. If you drag across the middle of a word, the entire word is selected, even if it falls partially outside the selection rectangle.



The selected text is actually contained in a hidden layer, separate from the document image. You can select and copy text in this layer just as you can in a text document.

- Copying text or images
- Copying and cutting pages
- Rotating pages

## About the thumbnail pane

Use the thumbnail pane to select a page in a document for viewing in the page pane or in Reading view. When you click an image in the thumbnail pane, that page appears in the page pane.

You can show or hide the thumbnail pane by clicking **Thumbnail Pane** on the **View** menu.



- Thumbnail pane
- Page pane
- Sizing the thumbnail pane
- Sizing thumbnails
- Rearranging thumbnails

## Select a page thumbnail

In the thumbnail pane, you can select page images to view, copy, or delete.

- Select a thumbnail
- Select a series of thumbnails
- Select thumbnails not in a series

## Go to a specific page

You can quickly jump to any page in a multi-page scanned document.

- 1. On the **Edit** menu, click **Go To**.
- 2. Select or type the number of the page you want.
- 3. Click **OK**.

## View an open document

You can view any or all open documents in Microsoft Office Document Imaging.

- Switch to an open document
- View all open documents

### View page thumbnails

Use the thumbnail pane to navigate through multi-page documents.

- 1. On the **View** menu, click **Thumbnail Pane** if it is not already selected.
- 2. On the **View** menu, click **Thumbnails**, and then click the size you want. The default size is **Medium**.

▶ Tip

### Zoom and pan the page pane

You can zoom to change the size of images displayed in the page pane, or pan to change the on-screen position of documents that are displayed at a large zoom percentage.

- Zoom using the toolbar
- Zoom using the menu
- ▶ Pan using the mouse

#### **About finding text**

Use the Microsoft Office Document Imaging **Find** command on the **Edit** menu to locate specific words, word fragments, phrases, or punctuation marks in a scanned document. When text is found, it is highlighted in the page pane, which scrolls multi-page documents to the appropriate page. Each time you click **Next** or **Previous**, additional occurrences of the text are highlighted in the page pane.

Before you can find text in a scanned document, Office Document Imaging must perform optical character recognition (OCR) on that document.

If the **Use OCR to recognize the text of the scanned image** check box on the **Processing** tab of the **Preset Options** dialog box is selected, text recognition is automatically performed on any documents that you scan.

After text is recognized, it is stored in a Tagged Image File Format (TIFF) file with a .tif extension. If you then edit the TIFF file using another program, the text is lost and you must perform OCR again.

#### **Find text**

You can search a Microsoft Office Document Imaging document for specific text. Before you can find text, Document Imaging must perform optical character recognition (OCR) on the scanned document.

- Find text
- ▶ Find the next or previous occurrence of text
- Reuse a previous search

#### **Notes**

The **Find** command is case-insensitive, treating uppercase and lowercase characters the same.

Leading spaces, trailing spaces, and accent marks are ignored.

Common OCR errors are ignored. For example, the letter combination "rn" is often incorrectly recognized as "m," so the **Find** command will locate both. However, this can occasionally cause false matches.

## **Export text or pictures**

- ▶ Export text to a new Microsoft Word document
- ▶ Copy text to an existing document
- Copy pictures or artwork

#### Send a document in e-mail

- 1. Open your e-mail program.
- 2. Switch to the Microsoft Office Document Imaging document you want to send.
- 3. On the **File** menu, click **Send to Mail Recipient as Attachment**.

Your e-mail program creates a new message with the Office Document Imaging file attached.

#### Create a new document

You can create new documents in Microsoft Office Document Imaging by moving pages from other documents or by scanning new documents.

- Move pages into a new document
- Scan a new document

### Open a document

Microsoft Office Document Imaging can open most Tagged Image File Format (TIFF) files.

Do one of the following:

On the **File** menu, click **Open**.

On the **File** menu, click the name of a document that was previously opened.

Office Document Imaging keeps a list of the last four documents opened.

#### Save a document

You can save Tagged Image File Format (TIFF) documents created in Microsoft Office Document Imaging for use in other programs.

- Save a document
- Save a document with a new file name

**Note** When you save a TIFF file in Office Document Imaging, the optical character recognition (OCR) text is saved within the TIFF file. If you open and save the TIFF file using another program, the OCR text will be lost. You must run OCR again using Office Document Imaging if you want to access the text in the TIFF file.

#### **Print a document**

1. On the File menu, click Print.

For Help on a printing option, click the question mark ?, and then click the option.

2. Click OK.

**Note** In the **Print Options** list, the **Fit to page** option reduces images that are too large, or enlarges small images to fill the page.

### **Close a document**

On the **File** menu, click **Close**.

### **Troubleshoot scanning**

- ▶ The scanned text of my document is hard to read.
- My scanned photos or artwork are of poor quality.
- ▶ When I run optical character recognition (OCR), there are too many text-recognition errors.
- Scanning a document takes too long.
- My scanned files are too large.
- ▶ I can't select text in a scanned document.
- ▶ The automatic document feeder doesn't work.
- ▶ I can't scan multiple pages into the same file.

# **Advanced Scan Settings dialog box**

Option	Description
Resolution	Specifies the dots-per-inch resolution for the current preset. This sets the number of dots per inch scanned from your original document. Selecting a larger number will make finer detail visible, but will create larger files and may take longer to scan.
	Note Documents scanned at resolutions greater than 900 dots per inch will be converted to 900 dots per inch if Auto contrast, Save as monochrome, Auto rotate, or Auto straighten is selected.
Туре	Specifies the amount of information to record for each pixel to be scanned for the current preset. The options <b>24 bit color</b> and <b>8 bit gray</b> record more detailed information about an image than does <b>Monochrome</b> , but they create larger files and may take longer to scan.
Auto contrast, save as monochrome	Creates high-contrast images for optimum optical character recognition (OCR). After scanning by using the selected scan type, the resulting file is saved in 1-bit monochrome, resulting in a smaller file size. This option can be helpful when scanning for OCR those documents with background colors or colored text.

# **Choose Scanner dialog box**

Use this dialog box to select an available scanner.

Option	Description
Scanner	Lists the available scanners for your computer. Select the one you want to use.
Use automatic document feeder	Select this check box if your scanner has an automatic document feeder (ADF), and you want to use it.
Show scanner driver dialog before scanning	Select this check box to display the dialog box for the scanner software that came with your scanner. This option is useful if you want to control advanced features and settings of your scanner.

## Find dialog box

Use this dialog box to search for text in a scanned document.

Option	Description
Find	Type the text you want to find.
Next	Click to find the next occurrence of the text.
Previous	Click to find the previous occurrence of the text.
Cancel	Click to close the dialog box.

**Note** Before you can find text, optical character recognition (OCR) must first be performed on the document.

# Go To dialog box

In the Page box, select or type the number of the page you want, and then click OK.

### **Indexing Service dialog box**

Indexing is a special service that enables fast file searches on your computer. Text found in files on your computer is added to the index, which also stores a reference to the file in which the text was found.

**Note** You must have administrator permissions to change settings in this dialog box because indexing is a system service.

Option	Description
Use OCR to recognize the text in TIFF files when indexing	Select this check box to perform optical character recognition (OCR) on Tagged Image File Format (TIFF) files when the Windows Indexing Service compiles search indexes. This allows the Microsoft Windows search function to access text in TIFF files that do not already contain OCR information.
Auto rotate	Select this check box to detect the orientation of text on the page and perform OCR on images that are upside down or sideways.
OCR Language	Lists the available languages for OCR. Select the language you want to use.

### **Insert File dialog box**

Use this dialog box to insert the contents of a Tagged Image File Format (TIFF) file into the current document.

The message "Insert n page(s)" at the top of the dialog box indicates how many pages the selected file contains.

Option	Description
In front of page	The new pages are inserted before the page number you select in the box.
At the beginning of the document	The new pages are inserted before the first page.
At the end of the document	The new pages are inserted after the last page.

#### **Launching Scanner Driver dialog box**

This dialog box appears when Microsoft Office Document Imaging cannot set your scanner settings, or when you selected the **Show scanner driver dialog before scanning** check box in the **Choose Scanner** dialog box. In these cases, Office Document Imaging cannot control scanner settings from the currently selected scan preset, so you will need to do so manually.

When you click **OK**, you'll see the scanner driver dialog box for your currently selected scanner. Use this dialog box to set desired scan settings and begin scanning.

Option	Description
<b>Don't Show Me</b>	Select this check box if you want to hide the
This Dialog Again	<b>Launching Scanner Driver</b> dialog box in the future.

#### Note

If you are using Microsoft Windows Millennium Edition, then using a Windows Image Acquisition (WIA) scanner driver with an automatic document feeder (ADF) is not recommended. For best results, install the TWAIN scanner driver supplied by the scanner manufacturer.

### **Microsoft Office Document Scanning dialog box**

You can use the Microsoft Office Document Scanning feature to scan new documents, as well as to recognize text in existing documents.

Option	Description		
Select a preset for scanning	Scanning presets are optimized for different types of documents. The default preset is <b>Black and white</b> .		
Scan	Click to scan the document by using the selected preset.		
Preset	Click to choose one of the following options.		
options •	<b>Create new preset</b> Create additional custom scanning presets for specific tasks by using the <b>Preset Options</b> dialog box.		
•	<b>Edit selected preset</b> Modify settings in any existing preset by using the <b>Preset Options</b> dialog box.		
•	<b>Delete selected preset</b> Permanently remove a preset from the list.		
Scanner	Click to select the scanner you want to use. Opens the <b>Choose Scanner</b> dialog box.		
Original is double sided	If your scanner is equipped with an automatic document feeder (ADF), click this option to scan the front and back of multiple pages. Place the pages in the feeder so that the page fronts are scanned first.		
	The Microsoft Office Document Scanning dialog box will reappear, allowing you to turn the pages over and scan the page backs. Microsoft Office Document Scanning automatically collates the pages in the correct order after scanning is complete.		
Prompt for additional	Click this option to pause after each page is scanned, allowing you to place the next page in your scanner. Or, if you are using an ADF, click this option to pause when the feeder		

#### pages

is empty, allowing you to load more pages.

You can use this option with an ADF when you are scanning books or other documents that won't fit or are too long to load into the feeder. The **Microsoft Office Document Scanning** dialog box reappears after each page is scanned, or after the feeder in your ADF is empty.

# View file after scanning

Select this check box to open the newly scanned document and display it in the Microsoft Office Document Image Viewer.

#### **Options dialog box**

This dialog box controls various Microsoft Office Document Imaging settings.

The first three options control optical character recognition (OCR) settings when you are using the **Recognize Text Using OCR** command on the **File** menu.

Option	Description
Auto rotate	Select this check box to detect the orientation of text on the page and rotate page images that are upside down or sideways.
Auto straighten	Select this check box to detect the orientation of text on the page and correct any slightly misaligned pages.
OCR Language	Lists the available languages for OCR. Select the language you want to use.
Reset	Click to specify Office Document Imaging as the default application used to open Tagged Image File Format (TIFF) files.
Indexing Service	Click to open the <b>Indexing Service</b> dialog box.

**Note** Documents scanned at resolutions greater than 900 dots per inch (dpi) will be converted to 900 dpi if **Auto rotate** or **Auto straighten** is selected.

# **Preset Options dialog box - General tab**

Use this dialog box to specify settings for Microsoft Office Document Scanning presets.

Option	Description
Preset Name	Indicates the name of the current preset. When creating a preset, enter a descriptive name for your preset.
Scan Type	Specifies the type of scan (color, grayscale, or black and white) to be performed for the current preset.
Advanced	Click to open the <b>Advanced Scan Settings</b> dialog box to view or modify settings for the scan type.
Create Shortcut	Click to create and place on the desktop a shortcut icon that launches Microsoft Office Document Scanning and immediately starts a new scan by using the current preset.

#### **Preset Options dialog box - Output tab**

Use this dialog box to specify settings for saving and naming files created by Microsoft Office Document Scanning.

Option	Description
Save in folder	Specifies the current location for saving files. The default location is the My Documents folder.
Change	Click to change the current location for saving files.
Auto name based on first words on page	Select to create file names based on the first few words of text extracted using optical character recognition (OCR).
Auto number	Select to create sequentially numbered file names.
Date and time	Select to create file names containing the date and time of scanning.
Format	Specifies the format to be used for file names generated by using date and time.

Note If you want to use the Auto name based on first words on page option, you must select the Use OCR to recognize the text of the scanned image check box on the Processing tab.

### **Preset Options dialog box - Page tab**

Use this dialog box to specify settings for the size of pages being scanned, if applicable, and for automatic document feeders (ADF).

Option	Description
Page Size	Lists a variety of standard sizes for the page being scanned. When this option is available, you can select <b>Custom</b> to specify <b>Width</b> and <b>Height</b> . Available page sizes depend on the scanner model.
Width	Specifies the width of the page being scanned. Available only when the <b>Custom</b> page size is selected.
Height	Specifies the height of the page being scanned. Available only when the <b>Custom</b> page size is selected.
Save each page as a separate document	Select to save multi-page scans as separate documents. This is useful when scanning individual documents using an ADF.

#### **Preset Options dialog box - Processing tab**

Use this dialog box to specify settings that control optical character recognition (OCR) and scanning.

Option	Description
Use OCR to recognize the text of the scanned image	Microsoft Office Document Imaging runs OCR automatically after scanning is completed. If this option is turned off, the rest of the options on this tab are unavailable.
Auto rotate	Select this check box to detect the orientation of text on the page and rotate scanned pages that are upside down or sideways.
Auto straighten	Select this check box to detect the orientation of text on the page and correct pages that are slightly misaligned.
OCR Language	Lists the available languages for OCR. Select the language you want to use.

**Note** Using the **Auto rotate** and **Auto straighten** options may slow down the scanning process. If you are scanning at a higher resolution, using either of these options also reduces the resolution to 900 dots per inch.

# **Properties dialog box - General tab**

These properties pertain to the entire active document.

Option	Description
Туре	Indicates the file type, which is Tagged Image File Format (TIFF), the only supported file type.
Location	Indicates the location of the file on the disk.
Size	Shows the size of the file in bytes. Remains blank if the file has not been saved.
MS-DOS name	The name of the file as it appears as an "8.3" name in an MS-DOS window, where "8" indicates the eight-character file name and "3" indicates the three-character file type extension, in this case .tif.
Created	Time and date that the file was first saved.
Modified	Time and date that the file was last saved.
Accessed	Time and date that the file was last opened.
Attributes	MS-DOS file attributes include <b>Read-only</b> , <b>Archive</b> , <b>Hidden</b> , and <b>System</b> .

#### **Properties dialog box - Per Page tab**

These properties pertain only to a selected page in the current document. The thumbnail image at the top displays a small representation of a page in the current document.

Ontion	Deceription
Option	Description
Page scanned on	Time and date that the current page was originally scanned.
Page scanned by	Name of the user who scanned the current page.
Scanner	The model name of the scanner used to scan the current page.
Scanner software	The name and version number of the scanner software used to scan the current page.
lmage parameters	Lists technical statistics pertaining to the Tagged Image File Format (TIFF) image of the current page, including length and width, resolution in dots per inch, and type of image compression used.
Page statistics	Lists statistics pertaining to the results of optical character recognition (OCR) performed on the current page only, including numbers of characters, words, lines, and paragraphs. Also indicates <b>OCR quality</b> , a number between 0 and 1000 that reflects the relative accuracy of text recognition.

# **Properties dialog box - Statistics tab**

These properties pertain to the entire active document.

Option	Description
•	Time and date the document was scanned.
on	
Scanned by	Name of the user who scanned the document.
Scanner	The model name of the scanner used to scan the document.
OCR status	Indicates whether optical character recognition (OCR) was completed and successful.
	The name and version number of the scanner software used to scan the current page.
Last edit on	The time and date the file was last modified.
Last edit by	The name of the user who last modified the file.
Last saved by	The name of the user who last saved the file.
Statistics	Lists statistics pertaining to the results of OCR performed on the file, including the number of pages, paragraphs, lines, words, and characters. Also indicates <b>OCR quality</b> , a number between 0 and 1000 that reflects the relative accuracy of text recognition.

### Send Text to Word dialog box

Use this dialog box to send all or part of the recognized text in the current document to Microsoft Word.

Option	Description
<b>Current</b> selection	If you selected text in the page pane or in Reading view, select this option to send the selected text to Word.
Selected pages	Select this option to send all text on the selected pages to Word.
All pages	Select this option to send all text in the document to Word, regardless of the selection.