OmniPage® Pro

User's Manual

Caere GmbH
Innere Wiener Strasse 5
81667 München, Germany

Caere UK Information Centre
Abbey House
4 Abbey Orchard Street
Westminster, London SW1P 2JJ

Centre d'informations Caere
72, rue Baratte-Cholet
94100 Saint-Maur, France

Please Note
In order to use this program, you should know how to work in the Microsoft Windows environment. Please refer to Windows documentation if you have questions about how to use menu commands, dialog boxes, scroll bars, edit boxes, and so on.

OmniPage Pro for Windows
Version 9

800-1288-030A

Welcome

Welcome to OmniPage Pro, and thank you for using our software! The following documentation has been provided to help you learn about OmniPage Pro.

This User's Manual

This manual introduces you to the basics of using OmniPage Pro. It includes installation and setup instructions, an introduction to OmniPage Pro, task-oriented instructions, ways to customize processing, settings guidelines, and technical information.

This manual is also available as an electronic PDF file. To open the file, click *Start* in the Windows taskbar and choose *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *OmniPage Pro Manual* after OmniPage Pro has been installed.

Online Help

OmniPage Pro's online Help contains detailed information on features, settings, and procedures. The online Help conforms to Windows 95 Help conventions and has been designed for quick and easy information retrieval. Please see "Getting Online Help" on page 14 for more information.

Readme File

The *Readme* file contains last-minute information about the software. Please read it before using OmniPage Pro. To open this text file, click *Start* in the Windows taskbar and choose *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *OmniPage Pro Readme* after OmniPage Pro has been installed.

Scanner Setup Notes

The *Scanner Setup Notes* contains information about supported scanners and related issues. To open this PDF file, click *Start* in the Windows taskbar and choose *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *Scanner Setup Notes* after OmniPage Pro has been installed.

Using This Manual

This manual is written with the assumption that you know how to work in the Microsoft Windows environment. Please refer to your Windows documentation if you have questions about how to use dialog boxes, menu commands, scroll bars, drag and drop functionality, shortcut menus, and so on.

The following conventions are used in this manual.

Convention	Purpose
Italicized text	• Emphasizes menu commands, dialog box options, labeled buttons, and file names
	For example:
	"Choose *Open...* in the File menu."
	• Emphasizes new terms the first time they are used
	• Emphasizes important words in a sentence
▶ Note symbol	Introduces a tip or an item of note
⚠ Warning symbol	Introduces important information

Installation and Setup

This chapter provides installation and setup information for OmniPage Pro and the Scan Manager.

For technical and troubleshooting information, please read Chapter 6, Technical Information.

For information on supported scanners and scanner setup, read the *Scanner Setup Notes*. To open this PDF file, click *Start* in the Windows taskbar and choose *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *Scanner Setup Notes* after OmniPage Pro has been installed.

This chapter contains the following topics:

- Minimum System Requirements
- Installing OmniPage Pro
- Starting and Closing OmniPage Pro
- Registering OmniPage Pro

Minimum System Requirements

You need the following setup, at minimum, to install and run OmniPage Pro:

- Computer with a 486 or higher processor
- Microsoft Windows 95, Windows 98, or Windows NT 4.0
- 16MB of memory (RAM)
- 45MB of free hard disk space to install application files, the Scan Manager, and one OCR language
  55MB to install above files and all OCR languages
- SVGA or VGA monitor with 256 colors
- Windows-compatible pointing device
- CD-ROM drive for installation
- A compatible scanner if you plan to scan documents
  Please see the *Scanner Setup Notes* for a list of tested scanners.

▶

Performance and speed will be enhanced if your computer's processor, memory, and available disk space exceed the minimum requirements.

Installing OmniPage Pro

OmniPage Pro's Setup program takes you through installation with onscreen instructions at every step.

Before installing OmniPage Pro:

- Make sure your scanner is connected, turned on, and compatible with your system.
- Close all other applications, especially anti-virus programs.
- Log into your computer with administrator privileges if you are installing on Windows NT.

⚠

If you own a previous version of OmniPage Pro, or if you are upgrading from OmniPage Limited Edition, it is strongly recommended that you uninstall that product first and then restart your computer.

To install OmniPage Pro:

1 Insert OmniPage Pro's CD-ROM in the CD-ROM drive.

   The Setup program should start automatically. If it does not start, locate your CD-ROM drive in Windows Explorer and double-click the Setup.exe program at the top-level of the CD-ROM.

2 Follow the instructions on each screen to install the software.

   During installation, you may be prompted to enter a serial number. You can find your serial number on the label of the CD-ROM envelope.

The Caere Scan Manager is installed during OmniPage Pro installation. You will be prompted to select your scanner manufacturer and model in the Scan Manager so that you can use your scanner with OmniPage Pro. Read the *Scanner Setup Notes* for the most detailed information about scanner support and setup. You can open this document after OmniPage Pro has been installed by clicking *Start* in the Windows taskbar and choosing *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *Scanner Setup Notes*.
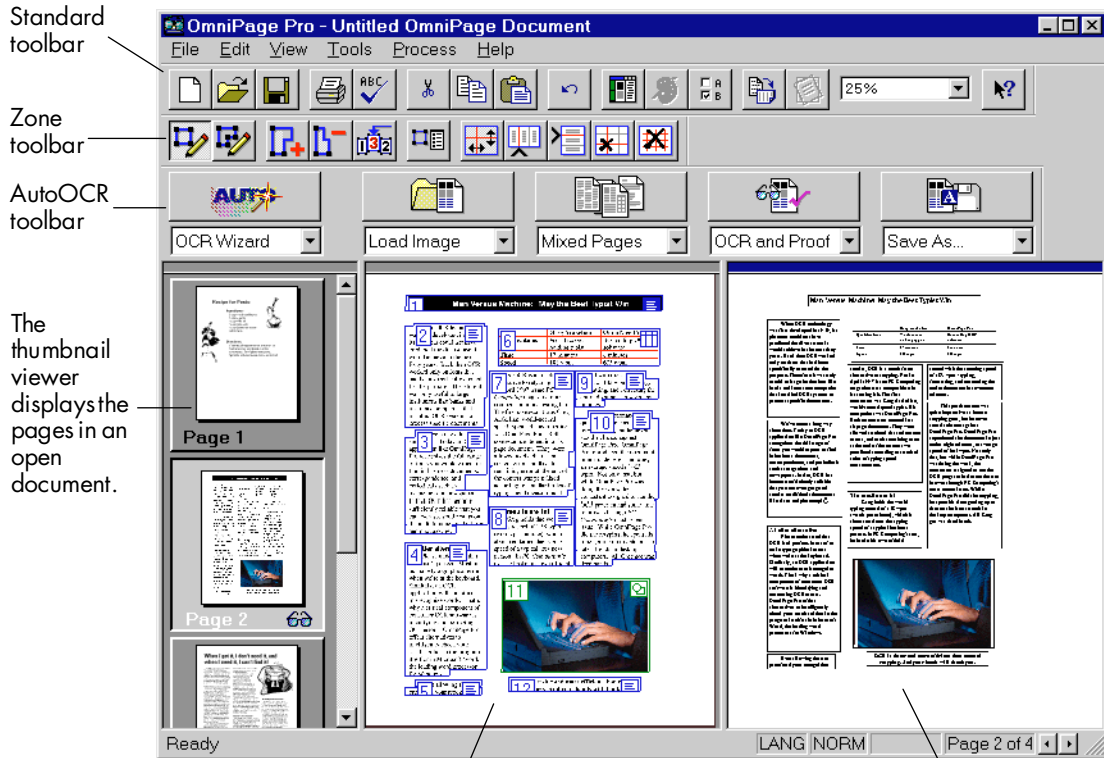
Starting and Closing OmniPage Pro

If you plan to scan, make sure your scanner is attached to your computer and turned on before you start OmniPage Pro.

To start OmniPage Pro, do one of the following:

- Click *Start* in the Windows taskbar and choose *Programs* ▸ *Caere Applications* ▸ *OmniPage Pro 9.0*.

  (Use the program group you selected during installation if it is different than *Caere Applications*.)

- Double-click the OmniPage Pro icon located in the folder where you installed OmniPage Pro.

OmniPage Pro's desktop appears when you open OmniPage Pro. See "The OmniPage Pro Desktop" on page 10 for an introduction to OmniPage Pro's user interface.

Standard toolbar

Zone toolbar

AutoOCR toolbar

The thumbnail viewer displays the pages in an open document.

The image viewer displays the current page's original image.

The text viewer displays the current page's recognized text and retained graphics.

Closing OmniPage Pro

Choose *Exit* in the file menu to close OmniPage Pro. You are prompted to save the current document if you have not saved it or have modified it since the last save.

Registering OmniPage Pro

Register your copy of OmniPage Pro with Caere Corporation to receive notification of special offers and the best prices on product upgrades.

▶

Some versions of OmniPage Pro will only launch 25 times if you do not register it.

If you purchased your product directly from Caere or if you were previously registered, you may not need to register again. Your version of OmniPage Pro will not display a *Register* menu if you do not need to register it.

To register OmniPage Pro:

1   Click the *Register* menu to open the Register dialog box.

2   Click *Register Now*.

3   Fill out the information requested on the screen and then click *Next*.

4   Follow the instructions on the screen.
     OmniPage Pro will decide on the best method of registration according to your country and computer system. It may try using modem, FTP, or HTTP connections to transmit your registration information directly. Or, it may prompt you to call a phone number or print out and mail in your registration information.

▶

After registration is complete, you will be given a registration number. Be sure to write that number down and keep it handy in case you need to use it for reinstallation. If you reinstall OmniPage Pro using your registration number on the same computer, you will not have to go through the entire registration process again to reregister it.

To reregister OmniPage Pro after reinstallation:

1   Click the *Register* menu to open the Register dialog box.

2   Click *Reregister*.

3   Type in your registration number and click *OK*.

Introduction to OmniPage Pro

You probably use your computer for most business correspondence and other written projects. The challenge is that certain sources of information cannot be immediately used on a computer.

For example, if you want to incorporate information from a magazine article into a document in your word processor, you somehow have to get the text from the article into your computer. Painstakingly retyping the article is not an appealing solution.

OmniPage Pro offers a smart solution to increase your work productivity. OmniPage Pro's *optical character recognition (OCR)* technology accurately and easily converts scanned paper documents and image files into editable text for use in your favorite computer applications. OmniPage Pro eliminates the need for manual retyping.

Please continue reading this chapter for information on these topics:

- What Is Optical Character Recognition (OCR)?
- The OmniPage Pro Desktop
- Getting Online Help
- Product Support

What Is Optical Character Recognition (OCR)?

Optical character recognition (*OCR*) is the process of turning an *image* into computer-editable text. An image is an electronic picture of text such as a scanned paper document or an electronic fax file. Images do not have editable text characters; they have many tiny dots (*pixels*) that together form a picture of text.

During OCR, OmniPage Pro analyzes an image and defines characters to produce editable text. After OCR, you can save the resulting text to a variety of word-processing, page layout, and spreadsheet applications.

OmniPage Pro's OCR Capabilities

In addition to text recognition, OmniPage Pro can retain the following elements of a document during OCR.

Graphics

Photos, logos, and drawings are examples of graphics.

Text formatting

Font types, font sizes, and font styles (such as **bold** or *italic*) are examples of text formatting.

Page formatting

Column structure, paragraph spacing, table formats, and placement of graphics are examples of page formatting.

The graphics, text formatting, and page formatting elements that OmniPage Pro retains are determined by the settings you select. See "Settings Guidelines" on page 51 for more information.

OmniPage Pro only recognizes machine-printed characters such as laser-printed or typewritten text. However, it can retain handwritten text, such as a signature, as a graphic.

Basic Steps of OmniPage Pro OCR

These are the basic steps of OmniPage Pro's OCR process.

1 Bring a document image into OmniPage Pro.

You can scan a paper document or load an image file. The resulting image appears in OmniPage Pro's image viewer. See "Bringing Document Images into OmniPage Pro" on page 20 for more information.

2 Create zones to identify areas you want to recognize as text or retain as graphics.

Zones are borders that enclose the areas of a document image that will get processed. You can create zones automatically, manually, or with a template. Any areas not enclosed by zones are ignored during OCR. See "Creating Zones for OCR" on page 22 for more information.

3 Perform OCR to convert text information into editable text characters.

During OCR, OmniPage Pro interprets text characters in an image. After OCR, you can check and correct errors in the text using the OCR Proofreader. See "Performing OCR on a Document" on page 23 for more information.

**4 Export the document to the desired location.**

You can save your document to a specified file format, place it on the Clipboard, or send it as a mail attachment. See "Exporting Documents" on page 34 for more information.

There are different ways to start the OCR process in OmniPage Pro. See "Ways to Process Documents" on page 18 for more information.

The OmniPage Pro Desktop

OmniPage Pro's desktop displays the pages of an open document in its thumbnail viewer, image viewer, and text viewer. You can use buttons in the Standard, AutoOCR, and Zone toolbars to perform various tasks on the document.
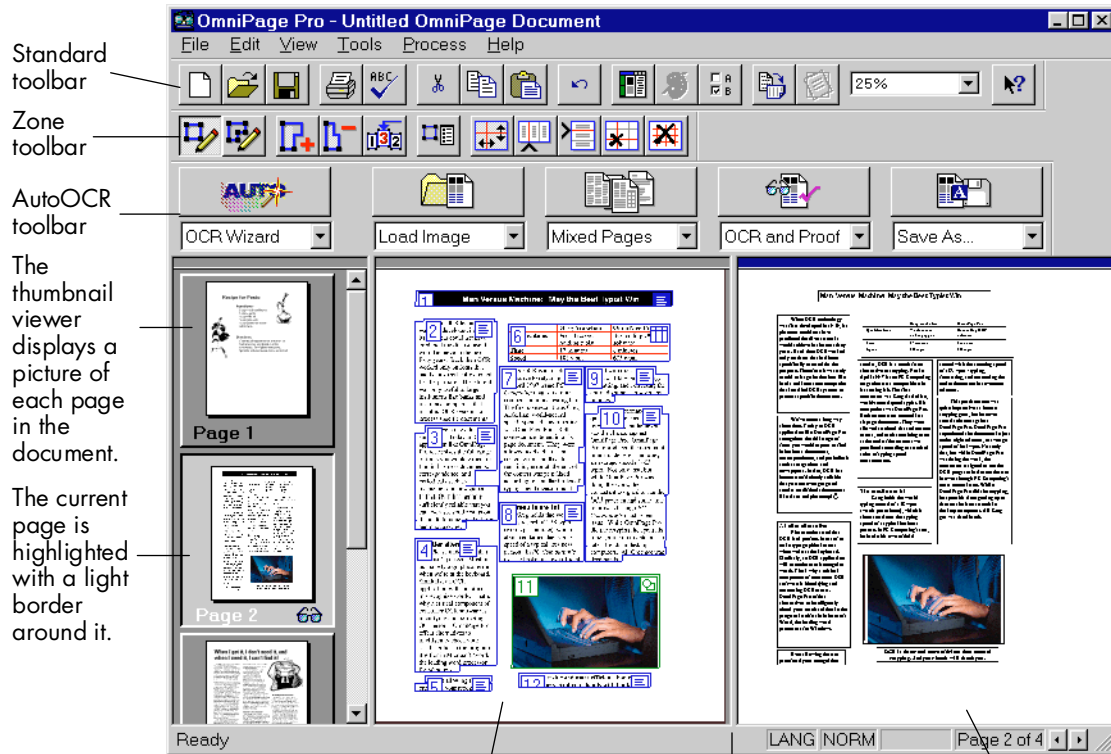
Standard toolbar

Zone toolbar

AutoOCR toolbar

The thumbnail viewer displays a picture of each page in the document.

The current page is highlighted with a light border around it.

The image viewer displays the current page's original image.

Drag this splitter to the left or right to resize a viewer.

The text viewer displays the current page's recognized text and retained graphics.

AutoOCR Toolbar

The AutoOCR® toolbar contains buttons that can activate each step of the OCR process.

Click the down arrow to display the commands in a button's drop-down list.
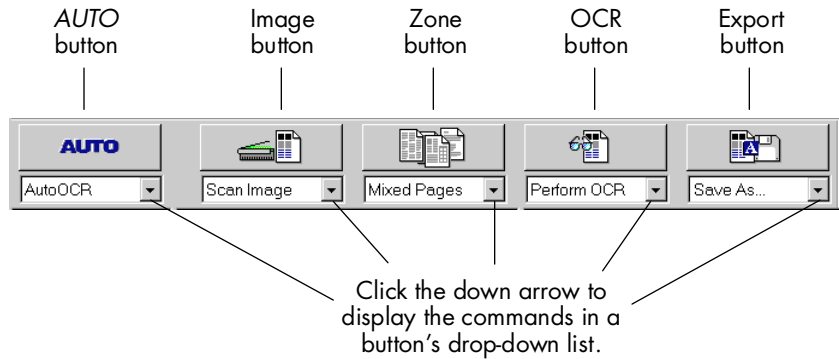
You can set different commands in the AutoOCR toolbar buttons for the operations you want to perform. Choose a command using each buttons's drop-down list.

- The AUTO button allows you to activate automatic processing or use the OCR Wizard.
- The Image button allows you to bring in images by scanning or loading pages.
- The Zone button allows you to automatically create zones on images based on their original page layouts or predefined templates.
- The OCR button allows you to perform OCR, train characters for OCR, or schedule OCR at a later time.
- The Export button allows you to save, copy, or send your recognized document as a mail attachment.

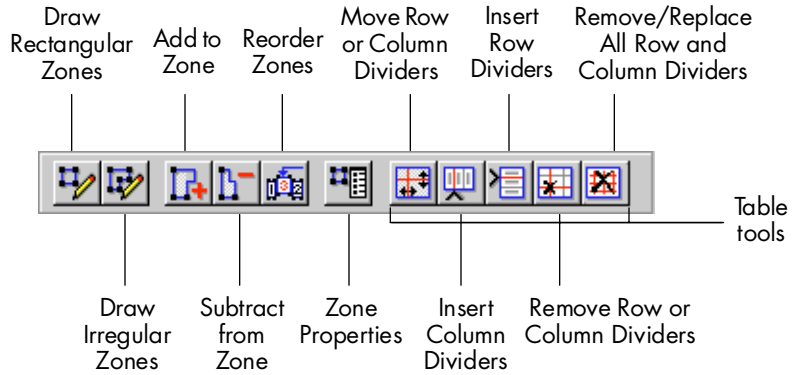Please see "Setting AutoOCR Toolbar Commands" on page 40 for more information on each toolbar button.

Standard Toolbar

The Standard toolbar contains buttons and a drop-down list for performing standard tasks.
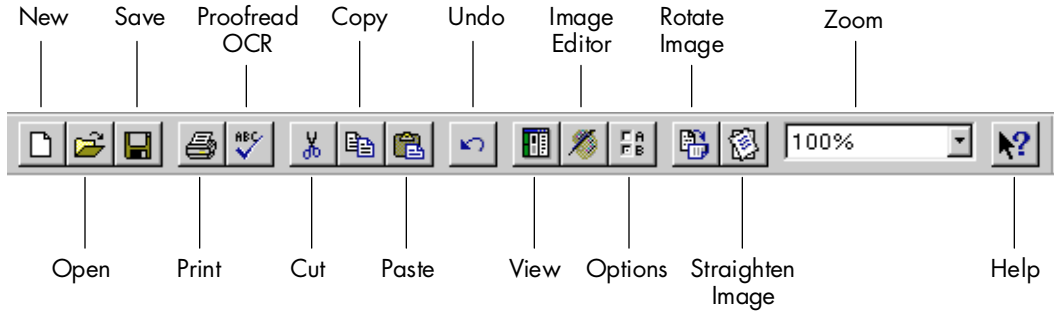
Zone Toolbar

The Zone toolbar contains buttons that allow you to draw and define zones on a page image.
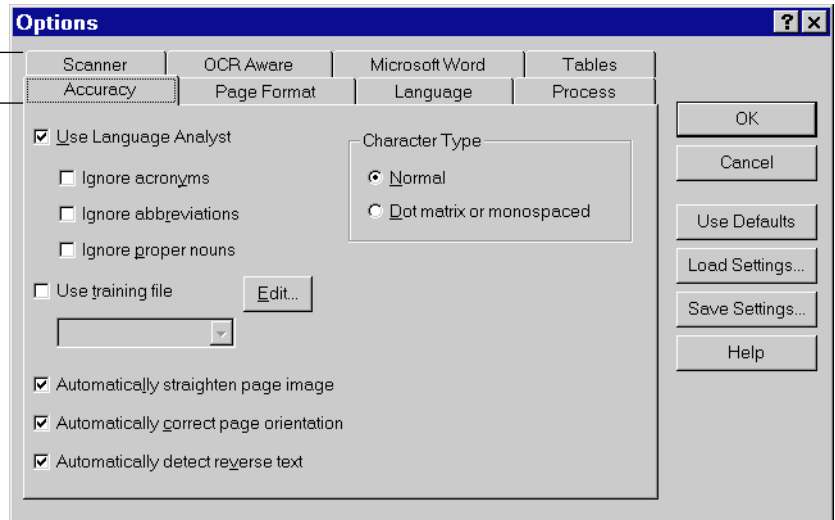

See "Customizing Zones" on page 63 for more information.

Options Dialog Box

You can select settings for OmniPage Pro in the Options dialog box. To open it, click the Options button or choose *Options...* in the Tools menu.

Click the tabs in the Options dialog box to view and select different settings.

Options				? X
Scanner	OCR Aware	Microsoft Word	Tables	
Accuracy	Page Format	Language	Process	

☑ Use Language Analyst

  ☐ Ignore acronyms

  ☐ Ignore abbreviations

  ☐ Ignore proper nouns

☐ Use training file   Edit...

Character Type

  ⦿ Normal

  ◯ Dot matrix or monospaced

☑ Automatically straighten page image

☑ Automatically correct page orientation

☑ Automatically detect reverse text

OK

Cancel

Use Defaults

Load Settings...

Save Settings...

Help

See Chapter 4, OmniPage Pro Settings, for more information on settings.

Getting Online Help

In addition to using this manual, you can use OmniPage Pro's online Help topics to learn about features, settings, and procedures. Online Help is available after you install OmniPage Pro.

OmniPage Pro's online Help follows the conventions of Microsoft Windows 95 Help. Choose *How to Use Help...* in OmniPage Pro's Help menu to get information on using Help.

Help Menu

One way to open OmniPage Pro's online Help is to choose commands in the Help menu.
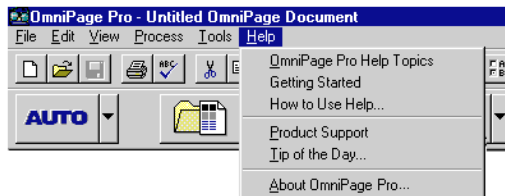
- Choose *OmniPage Pro Help Topics* to get contents and index listings for OmniPage Pro Help topics.
- Choose *Getting Started* to get introductory topics to OmniPage Pro.
- Choose *How to Use Help...* to get Microsoft Windows Help topics that explain how to use and customize Help.
- Choose *Product Support* to find out how to get product support services for OmniPage Pro.
- Choose *Tip of the Day* to get hints for using OmniPage Pro.
- Choose *About OmniPage Pro...* to get information about your version of OmniPage Pro.

Context-Sensitive Help

You can get on-the-spot information about a particular OmniPage Pro command, toolbar button, or dialog box option in the following ways:
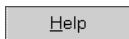
- Click the Help button in the Standard toolbar and then click any toolbar button, menu command, or area of the OmniPage Pro desktop to display a Help topic explaining that item.

- Click the question-mark button in the upper-right corner of a dialog box and then click an item in the dialog box to get a pop-up explanation for that item.

- Some dialog boxes have a *Help* button. Click *Help* to get information about that dialog box.

Product Support

For the fastest and easiest way to get help, please look for solutions in this manual or in the online Help. See "General Troubleshooting Solutions" on page 86 for troubleshooting tips.

If you need additional help, please use the following resources:

- **Caere's World Wide Web site**

  Go to Caere's World Wide Web site for common questions and answers, updates, patches, troubleshooting procedures, and product information. Caere's Web site address:

http://www.caere.com

- **OmniPage Pro Readme file**

  Read the *OmniPage Pro Readme* file for last-minute information about the software. This is available after installing OmniPage Pro. To open the file, click *Start* in the Windows taskbar and choose *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *OmniPage Pro Readme*.

- **Scanner Setup Notes**

  Read the *Scanner Setup Notes* document to learn about supported scanners and related issues. This document has been provided to you as an electronic document in PDF format. To open this document, click *Start* in the Windows taskbar and choose *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *Scanner Setup Notes*.

- **Caere Product Support document**

  Read the *Caere Product Support* document to get a list of support telephone numbers, including ones for international product support. This document has been provided to you as an electronic document in PDF format. To open this document, click *Start* in the Windows taskbar and choose *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *Product Support*.

◢▶

You must have Adobe Acrobat Reader 3.01 or greater installed if you want to read the *Caere Product Support* and *Scanner Setup Notes* PDF documents. To install the Reader, click *Start* in the Windows taskbar and choose *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *Acrobat Reader*.

Processing Documents

This chapter describes how to work with documents in OmniPage Pro, including each step of the OCR process.

There are different ways to accomplish the same tasks in OmniPage Pro. You can use toolbar buttons or menu commands to start procedures. OmniPage Pro can perform all OCR steps automatically, or you can start each step individually. You can even do different tasks at the same time.

Please continue reading this chapter for information on these topics:

- Ways to Process Documents
- Bringing Document Images into OmniPage Pro
- Creating Zones for OCR
- Performing OCR on a Document
- Proofreading OCR Results
- Using OCR in Other Applications
- Working with Documents
- Exporting Documents

For complete information on all OmniPage Pro commands, settings, and procedures, please use OmniPage Pro's online Help. See "Getting Online Help" on page 14 for more information.

Ways to Process Documents

Optical character recognition (OCR) is the process of turning an image into computer-editable text so you do not have to retype the text manually. The basic steps of OmniPage Pro's OCR process are explained on page 9. The following is a summary of those steps.

1. Bring a document image into OmniPage Pro.
   See page 20 for more information.

2. Create zones to identify areas you want to recognize as text or retain as graphics.
   See page 22 for more information.

3. Perform OCR to convert text information into editable text characters.
   See page 23 for more information.

4. Export the document to the desired location.
   See page 34 for more information.

Using the OCR Wizard

The OCR Wizard guides you through the entire OCR process by asking you questions about your document and selecting the appropriate settings for you.

To process your document using the OCR Wizard:

1. Set *OCR Wizard* as the command in the AUTO button's drop-down list.

2. Click AUTO or choose *OCR Wizard* in the Process menu.
   The first wizard screen appears.
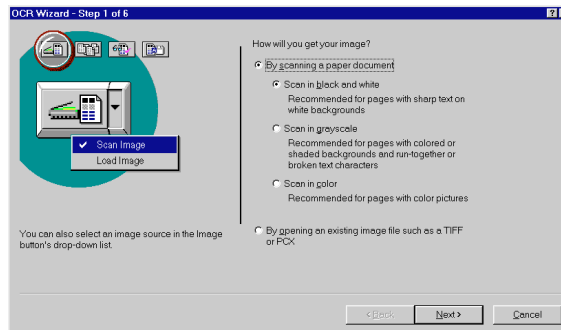
3. Answer the question in the first screen and click *Next*.

4. Continue answering questions in the screens that follow.

Automatic Processing

Use the AUTO button to process a new document from start to finish or to finish processing an open document.

To process your document automatically:

AUTO

1 Set *AutoOCR* as the command in the AUTO button's drop-down list.

2 Set the desired Image, Zone, OCR, and Export commands. See "Setting AutoOCR Toolbar Commands" on page 40 for more information.

3 Choose *Options...* in the Tools menu and check that settings are appropriate for your document. See "Settings Guidelines" on page 51 for more information.

4 Place your document in your scanner if you are scanning.

5 Click AUTO or choose *AutoOCR* in the Process menu. Each page of the document is processed and finished in order according to the selected commands. If page images in an open document already have zones, OmniPage Pro will skip zoning for those pages and continue with the selected OCR and export operations.

Performing Multiple Tasks at Once

OmniPage Pro takes advantage of your computer's ability to handle more than one process at a time. You can simultaneously scan, create zones, recognize, and edit documents. You do not have to wait for any process to complete before moving on to the next task.

For example, if you scan a multiple-page document, you can draw zones on an image as soon as the first page is scanned and you can edit recognized text as soon as it appears in the text viewer. These tasks can be done while other pages are being scanned and recognized.

Starting the OCR Process Outside OmniPage Pro

You can start the OCR process outside OmniPage Pro in a variety of ways. For example, you can use the *OCR Aware* feature to initiate OCR from another application and paste recognized text into an open document. See "Using OCR in Other Applications" on page 29 for more information.

Bringing Document Images into OmniPage Pro

You can bring document images into OmniPage Pro by scanning pages or loading image files.

Scanning Pages

You can scan paper documents to convert them to electronic images in OmniPage Pro. If a document is already open, scanned pages are inserted as new pages.

To scan in OmniPage Pro, you must install the Scan Manager and select your default scanner. See "Scan Manager is Needed with OmniPage Pro" on page 92 for more information.

To scan pages into OmniPage Pro:

1 Place your page in your scanner.

You can scan a stack of pages if you have an automatic document feeder (ADF).

2 Set *Scan Image* as the command in the Image button's drop-down list.

3 Choose *Options...* in the Tools menu and click the *Scanner* tab to make sure the appropriate settings are selected.

Select *Scan until empty* in the *Scanner* tab if you want to scan all pages in an ADF at once. Otherwise, you must click the Image button to scan each subsequent page.

4 Click the Image button or choose *Scan Image* in the Process menu.

Pages are scanned in order and combined into one working document.

Loading Image Files

You can load image files into OmniPage Pro. An image file is an electronic picture of text, such as a scanned paper document or an electronic fax, that is saved in an image file format such as PCX or TIFF. If a document is already open, loaded image files are inserted as new pages.
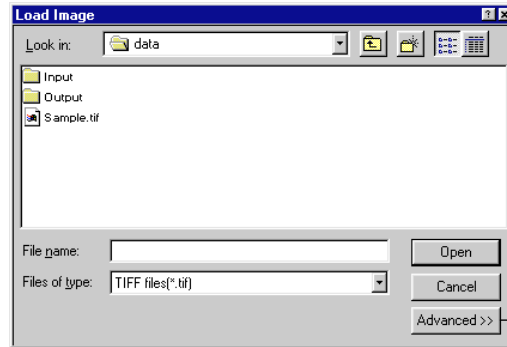
The following procedure is for loading image files only. To open an OmniPage Document (*.met), use the *Open...* command in the File menu.

To load image files into OmniPage Pro:

1   Set *Load Image* as the command in the Image button's drop-down list.

2   Click the Image button or choose *Load Image* in the Process menu.

    The Load Image dialog box appears.

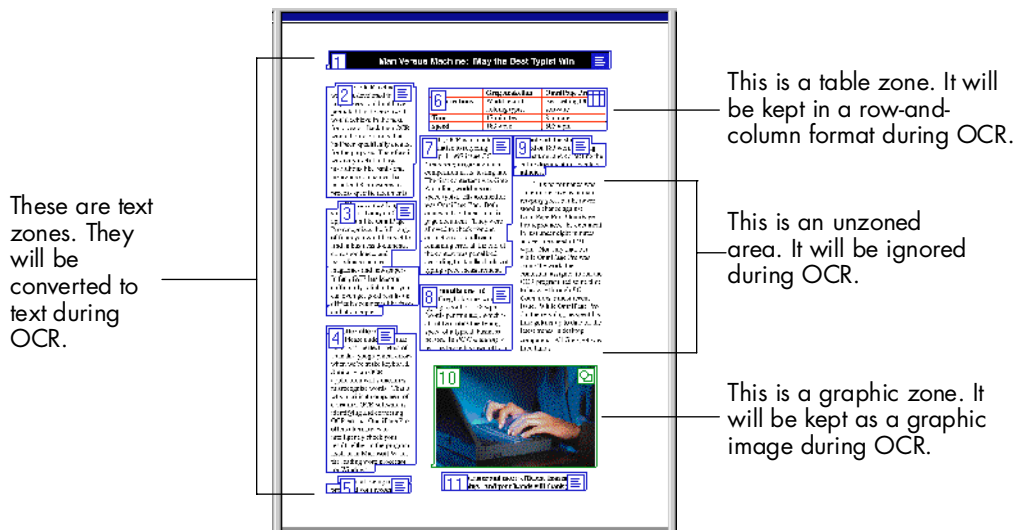Click *Advanced* if you want to select files from more than one folder.

3   Select the folder location and file type of the file you want to load.

    See "Supported File-Format Types" on page 89 for a complete list of supported file formats.

4   Select the files you want to load.

    You can Shift-click or Ctrl-click to select multiple files in the same folder.

5   Click *Advanced* if you want to select files from more than one folder.

    • Select a file and click *Add* to put it in the *Selected Files* list.

    • Click *Add All* to add all files from the current folder.

6   Click *Open* when you have selected all the files you want to load.

    Image files are loaded in the order selected and combined into one working document.

If you have electronic fax files that you want to convert to editable text, save the fax files in TIFF format and load them into OmniPage Pro using the *Load Image* command.

Creating Zones for OCR

Page images are displayed in OmniPage Pro's image viewer where *zones* are created before OCR. Zones are borders that identify areas of an image that will be recognized as text or retained as graphics. Any part of an image not enclosed by a zone is ignored during OCR.

These are text zones. They will be converted to text during OCR.

This is a table zone. It will be kept in a row-and-column format during OCR.

This is an unzoned area. It will be ignored during OCR.

This is a graphic zone. It will be kept as a graphic image during OCR.

The easiest way to create zones on a page is to let OmniPage Pro do it automatically for you. However, you may want to draw zones manually if you want to customize the way your page will be processed. For example, if you only want to process certain areas of a page, you would manually draw zones around the desired areas. For information on drawing zones manually, modifying zones, deleting unwanted zones, and using zone templates, please see "Customizing Zones" on page 63.

Creating Zones Automatically

OmniPage Pro can analyze a page and create zones automatically for you. It uses the selected setting in the Zone button to determine the text flow on a page and breaks it into ordered zones.

To create zones automatically:

1   Choose a setting in the Zone button's drop-down list that most closely matches the format of your document.

You can choose *Single-Column Pages*, *Multiple-Column Pages*, *Spreadsheet Pages*, *Mixed Pages*, or a template of your own. See "Zone Button Commands" on page 42 for more information on these settings.

2   Click the Zone button or choose *Auto Zones* in the Process menu.

OmniPage Pro automatically draws zones on the current page in the image viewer. Each zone has a number indicating its order and a picture indicating its zone type.

▶

Make sure zones are identified correctly before performing OCR. For example, if you want to retain an area as a graphic, that area should be identified as a *Graphic* zone type. See "Changing Zone Properties" on page 71 for more information.

Performing OCR on a Document

Performing OCR converts an image to editable text. This is also referred to as *recognizing text*.

▶

OmniPage Pro only recognizes machine-printed characters such as laser-printed or typewritten text. However, it can retain handwritten text, such as a signature, as a graphic.

To perform OCR:

1   Choose *Options...* in the Tools menu and click the *Page Format* tab.

2   Select an *Output Format* setting for your document.

OmniPage Pro uses this setting to determine the output formatting of a document during OCR.

3   Set *OCR and Proof* as the command in the OCR button's drop-down list.

Or, set *Perform OCR* as the command if you do not want the OCR Proofreader to begin automatically after OCR.

4   Click the OCR button.

The page is recognized according to the current zones and settings. If there are no zones on the page, zones are created according to the current command in the Zone button.

◇ ▶

To schedule a group of documents for OCR at a particular time, see "Scheduling OCR" on page 80.

Proofreading OCR Results

After performing OCR, recognized text appears in the text viewer where you can proofread the results. Proofreading starts automatically if you chose *OCR and Proof* as the OCR process command.

OmniPage Pro marks suspected errors in green and inserts a red "reject" character for any character it cannot recognize. To turn off these color markers, choose *Show Markers* in the View menu so that it is deselected.

To proofread OCR results and correct errors:

ABC ✓

1   Click the Proofread OCR button or choose *Proofread OCR...* in the Tools menu.

If a suspected error is detected, the OCR Proofreader dialog box displays the error and a picture of how it originally looked in the image.

This is what OmniPage Pro thought the word was.
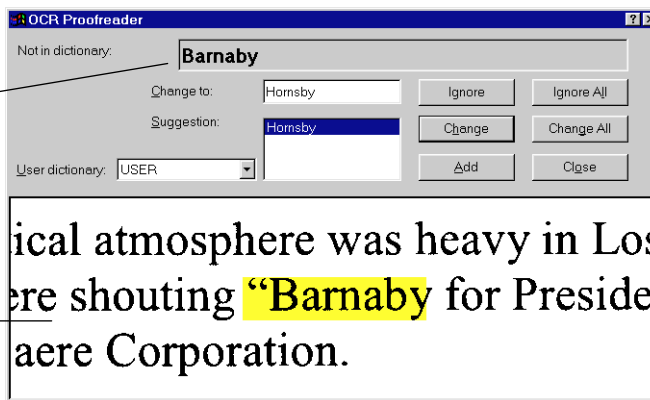
OCR Proofreader	? ☒

Not in dictionary: **Barnaby**

Change to: Hornsby | Ignore | Ignore All |
Suggestion: Hornsby | Change | Change All |
User dictionary: USER | Add | Close |

This window shows a picture of the original image. Click inside it to enlarge or reduce the picture. You can also drag a corner of the dialog box to see more areas of the image.

ical atmosphere was heavy in Los
ere shouting **"Barnaby** for Preside
aere Corporation.

2   Select one of these options for the word:

- Click *Ignore* to allow the word to remain as is.

- Click *Ignore All* to ignore all instances of the word in the current document.

- Click *Change* to replace the word with the word in the *Change to* edit box.

- Click *Change All* to replace all instances of the word with the word in the *Change to* edit box.

• Click *Add* to add the word to the current user dictionary.

After you choose an option for the word, the OCR Proofreader looks for the next possible error.

3   Click *Close* to stop proofreading OCR.

Color markers are removed from words that have been proofread.

Verifying Text

After performing OCR, you can compare recognized text against the original image to verify that the text was recognized correctly.

To verify text against its original image:

1   Double-click any word in the text viewer or select a word and choose *Verify Text* in the Tools menu.

The Verify Text window opens and shows a picture of the original word and its surrounding area.

This window shows a picture of the original image. Click inside it to enlarge or reduce the picture. You can also drag a corner of the window to resize it.

Close button

Verify Text

e Best **Typist** Win

2   Click inside the window to enlarge or reduce the picture.

The picture is enlarged on the first two clicks and reduced on the next two clicks.

3   Continue double-clicking words that you want to verify.

The display changes as you select new words.

4   Click the Close button to close the window.

Proofreading OCR Results in Microsoft Word

You can proofread OCR results directly in Microsoft Word 95 (version 7) or Word 97 if you have one of those versions installed on your computer.

To enable proofreading in Microsoft Word:

1   Select settings in the *Microsoft Word* tab of OmniPage Pro's Options dialog box.

See "Microsoft Word Settings" on page 50 for more information.

2   Make sure the `*.doc` file extension is associated with the version of Word you plan to use.

Refer to your Windows documentation for more information on associating file extensions with applications.

To proofread OCR results and correct errors in Microsoft Word:

1   Perform OCR on your document and then save it as the appropriate file type:

• Save as *Word for Windows 7.0, 95* if you are using that version.

• Save as *Word 97* if you are using that version.

2   Open the document in Microsoft Word.

⚠

The document must be opened on a system that has OmniPage Pro installed.

An OmniPage menu appears in Microsoft Word's menu bar as well as this corresponding toolbar:

Proofread OCR

Remove OCR Proofreader Support

Verify Text   Close Image Viewer
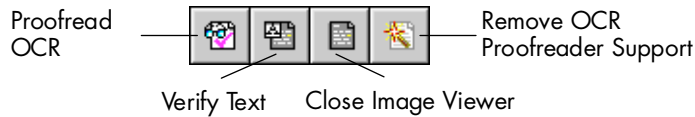
3   Choose *Proofread OCR...* in the OmniPage menu or click the Proofread OCR button.

If a suspected error is detected, the Verify Text window appears displaying the original image of the text.

Use these buttons to zoom in or out on the image.

original image

Verify Text - MEMO (1)

Please schedul

Conchita will
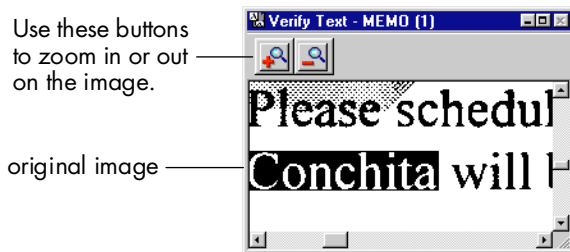
The OCR Proofreader dialog box also appears.
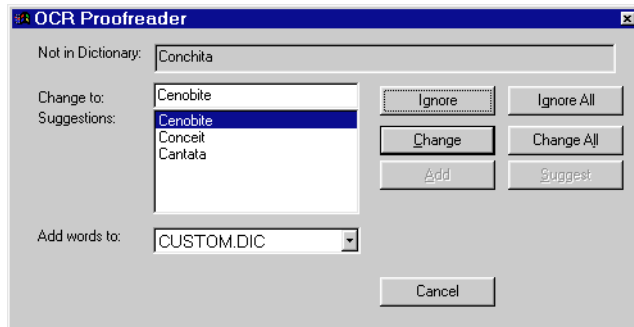
4   Select one of these options for the word:

  • Click *Ignore* to allow the word to remain as is.

  • Click *Ignore All* to ignore all instances of the word.

  • Click *Change* to replace the word with the word in the *Change to* edit box.

  • Click *Change All* to replace all instances of the word with the word in the *Change to* edit box.

  • Click *Add* to add the word to the current user dictionary.

  After you choose an option for the word, the OCR Proofreader looks for the next possible error.

5   Click *Close* to stop proofreading OCR.

  Color markers are removed from words that have been proofread.

⸻

To verify recognized text against its original image in Microsoft Word, you must process the document in OmniPage Pro and save it to the appropriate Word format. You cannot verify text against original images using the OCR Aware feature.

⸻

To verify text against its original image in Microsoft Word:

1   Follow steps 1 and 2 in the preceding instructions if your document is not already open in Microsoft Word.

2   Select a word that is a suspected error.

  Suspect words are marked in the color that was selected in the *Microsoft Word* tab of OmniPage Pro's Options dialog box.

> ◁▶
>
> You can only verify words that are marked as suspected errors. However, once the Verify Text window is open, you can use its scroll bars and zoom buttons to see any part of the original image.

3 Choose *Verify Text...* in the OmniPage menu.

The Verify Text window opens and shows a picture of the original word and its surrounding area.

Use these buttons to zoom in or out on the image.

4 Repeat steps 2 and 3 to continue proofreading other suspect words.

The display changes as you select new words.

5 Choose *Close Image Viewer* in the OmniPage menu to close the window when you are done.

◁▶

Removing OmniPage Pro Data from the Word Document

After proofreading OCR, you should remove OmniPage Pro data from your document to reduce its file size. You are automatically prompted to remove OmniPage data after all suspect words have been proofread. You can also choose *Remove OCR Proofreader Support* in the OmniPage menu. The OmniPage menu, toolbar, color markers, and image data will all be removed from the document.

Using OCR in Other Applications

You can use OmniPage Pro's *OCR Aware* feature to use OCR in other applications. For example, you can scan, recognize, and paste text directly into a document without ever leaving your word-processing application.

You can use OCR Aware with 32-bit applications that have been registered with OmniPage Pro. An application must be installed on your computer in order to use it with OCR Aware. See page 49 for more information on registering applications with OCR Aware.

For information on other ways to start OCR outside OmniPage Pro, please see the "Starting OCR Outside OmniPage Pro" online Help topic.

To use OCR Aware in an application:

1   Align your document in your scanner if you plan to scan.

2   Open the application in which you want to insert recognized text.
    The application must be registered to work with OCR Aware. You do not need to open OmniPage Pro itself.

3   Place the cursor at the location in your document where you want to insert recognized text.

4   Choose *Acquire Text Settings...* in the application's File menu if you want to check the current settings.

5   Choose *Acquire Text...* in the application's File menu when you are ready to start the OCR process.
    OCR processing occurs according to the selected settings. Recognized text appears at the cursor location in your application. If no document is open, text is copied to the Clipboard.

Text formatting, such as bold and italics, is retained if the application supports RTF information. Otherwise, only plain text will be pasted. Graphics are retained if the application supports bitmap images.

Working with Documents

OmniPage Pro's thumbnail, image, and text viewers allow you to look at and work with pages in the current document.

Thumbnail viewer

Image viewer

Drag this splitter to the left or right to resize a view.
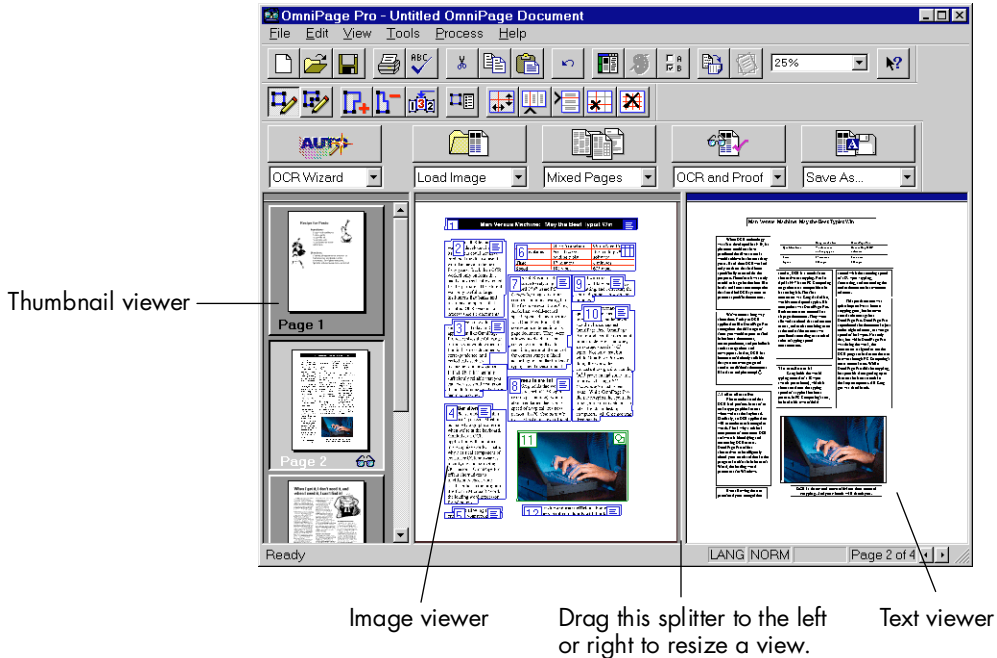
Text viewer

This section describes the following procedures:

- Resizing a Page View
- Changing Pages
- Reordering Pages
- Deleting Pages
- Printing a Document
- Closing a Document

Resizing a Page View

You can resize a page displayed in the image viewer or text viewer to enlarge or reduce the view.

To resize a page view:

1   Click in the viewer you want to enlarge or reduce to make it active.

2   Choose a size option in the Zoom drop-down list in the Standard toolbar.

Or, choose *Zoom* in the View menu and select a size option in the drop-down list.

The page resizes as specified.

▷

You can also click your right mouse button in the viewer you want to resize and select a size option in the shortcut menu. (If you are resizing the image viewer, click outside of a zone.)

Changing Pages

The thumbnail viewer, image viewer, and text viewer all display the same page of a document.

You can change pages in a document in the following ways:

• Click the thumbnail of the page you want to display.
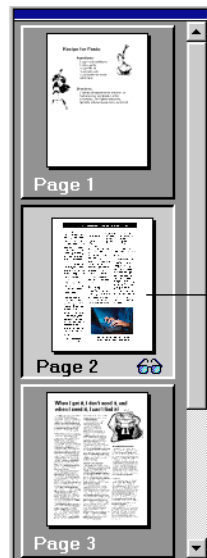
The thumbnail of the currently displayed page is highlighted with a light border around it.

- Click the Next Page or Previous Page buttons at the lower-right corner of the OmniPage Pro desktop.

◀ ▶

- Choose *Next Page*, *Previous Page*, or *Go to Page...* in the Edit menu.

Reordering Pages

You can reorder pages in a document by dragging their thumbnails to different positions in the thumbnail viewer.
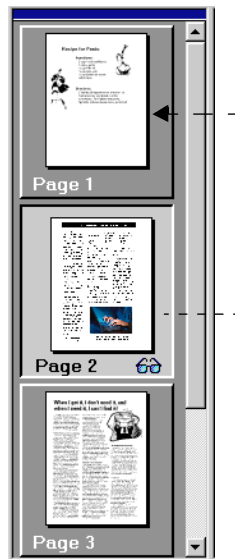
Click the thumbnail of the page you want to move and drag it above the desired page number.

▶

Hold down the Ctrl key while you click thumbnails if you want to select multiple thumbnails to move as a group.

Deleting Pages

If you delete a page from a document in OmniPage Pro, the thumbnail, original image, and recognized text for that page are all deleted.

To permanently delete pages:

- Choose *Delete Current Page* in the Edit menu to delete the currently displayed page.
- Select one or more thumbnails of pages you want to delete and press the Delete key.

▷

Undoing Changes

You can click the Undo button or choose *Undo* in the Edit menu to cancel the very last change you made in the text viewer. You can also choose *Undo* to cancel zone edits in the image viewer. However, page deletions cannot be undone.

Printing a Document

You can print the current document's original page images or recognized text.

To print a document:

1 Choose *Print...* in the File menu and choose one of the following in the submenu:

 • Choose *Image...* to print original page images.

 • Choose *Text...* to print recognized text.

2 Select the desired print settings in the Print dialog box.

3 Click *OK* to start the print job.

▷

As a shortcut, you can click either the text or image viewer to make it active and then click the Print button to print from that viewer.

Closing a Document

Choose *Close* in the File menu to close a document.

You are prompted to save your document if you have not saved it or have modified it since the last save. Save a document as an OmniPage Document (`*.met`) if you want to reopen it in OmniPage Pro again.

Exporting Documents

You can export a document to other applications by:

- Saving a Document
- Copying a Document to the Clipboard
- Sending a Document as a Mail Attachment

▷

After you export a document, a copy of the document remains open in OmniPage Pro. Save the document as an OmniPage Document (`*.met`) if you want to reopen it in OmniPage Pro again. OmniPage Documents retain all original images, zones, and recognized text.

Saving a Document

You can save recognized text and original images to disk in a variety of file types.

To save recognized text:

1   Choose *Save As...* in the File menu.

You can also click the Export button with *Save As* selected in the drop-down list.

The Save As dialog box appears.

The *Add to PageKeeper* setting only appears if you have PageKeeper installed on your computer. It puts a link to the saved document in PageKeeper's default folder.
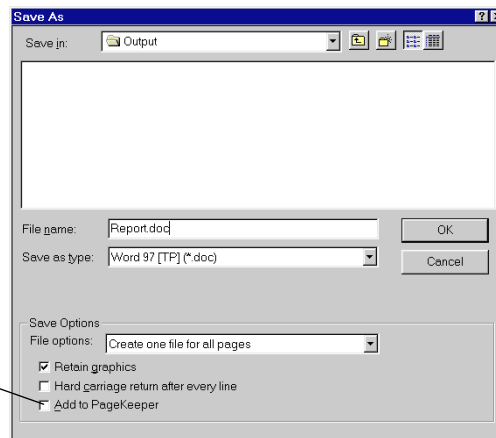
2   Select a folder location and file type for your document.

See "Supported File-Format Types" on page 89 for a complete list of supported file types.

3   Type in a file name and select save options.

4   Click *OK*.

The document is saved to disk as specified. Graphics and formatting are saved in the document only if the selected file type supports them.

To save original images:

1   Choose *Save Image...* in the File menu.

The Save Image dialog box appears.
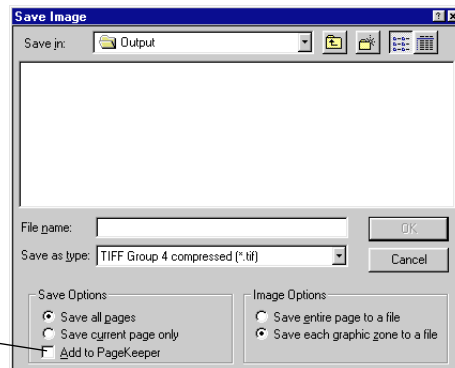
The *Add to PageKeeper* setting only appears if you have PageKeeper installed on your computer. It puts a link to the saved document in PageKeeper's default folder.

2   Select a folder location and file type for your document.

See "Supported File-Format Types" on page 89 for a complete list of supported file types.

3   Type in a file name and select *Save* and *Image* options.

4   Click *OK*.

The image is saved to disk as specified. (Zones and recognized text are not saved with the file.)

Saving a Document as You Work

Click the Save button in the Standard toolbar or choose *Save* in the File menu to save changes to the current document as you work.

> The Save As dialog box appears the first time you choose *Save* if a document has not been saved as an OmniPage Document or text-based file. See "Saving a Document" on page 34 for more information.

If a document has been saved as an OmniPage Document (*.met), all the changes you make in the open document are saved when you choose *Save*. If a document has been saved as a text-based file type, only the text changes are saved out to that file.

For example, suppose you save the current document as a text file called *Memo.txt*, but continue to make changes to the recognized text in OmniPage Pro. Whenever you choose *Save*, changes in the recognized text will be saved to the *Memo.txt* file.

Copying a Document to the Clipboard

You can copy every page of a recognized document to the Clipboard and then paste the text directly into another application.

To copy a document to the Clipboard:

1 Set *Copy to Clipboard* as the command in the Export button's drop-down list.

2 Click the Export button or choose *Copy to Clipboard* in the Process menu.
   The document is copied to the Clipboard.

> Text formatting, such as bold and italics, is retained when you paste into an application that supports RTF information. Otherwise, only plain text will be pasted. Graphics are retained if the application supports bitmap images.

Sending a Document as a Mail Attachment

You can send a recognized document as a file attached to a mail message if you have a MAPI-compliant mail application, such as Microsoft Outlook, installed.

To send a document as a mail attachment:

1   Choose *Send Mail...* in the File menu.

You can also click the Export button with *Send Mail* selected in the drop-down list.

The Send Mail dialog box appears.
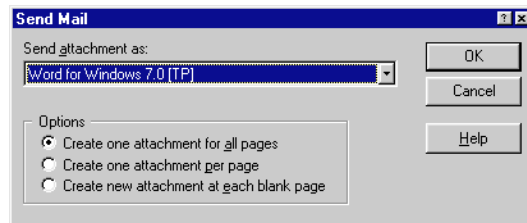
2   Specify a file type and attachment options for your document.

3   Click *OK*.

4   Log into your mail application if you are prompted to do so.
A new message appears ready for addressing.

5   Address your mail message as desired and click the Send button.
The document is sent as an attachment to the mail message.

OmniPage Pro Settings

This chapter describes the settings in the AutoOCR toolbar and Options dialog box. Please also look in OmniPage Pro's online help for more detailed information on settings.

The settings you select for processing documents can greatly affect OCR results. You may have to experiment with different settings to get the results you want. Settings guidelines are provided at the end of this chapter to get you started.

Please continue reading this chapter for information on these topics:

- Setting AutoOCR Toolbar Commands
- Selecting OmniPage Pro Settings
- Accuracy Settings
- Scanner Settings
- Page Format Settings
- Tables Settings
- OCR Aware Settings
- Process Settings
- Microsoft Word Settings
- Settings Guidelines

Setting AutoOCR Toolbar Commands

The AutoOCR toolbar buttons allow you to take a document through each step of the OCR process. Every toolbar button has different process commands that can be set for the operations you want to perform. OmniPage Pro can go through all steps automatically, or you can start each step individually.

AUTO	Image	Zone	OCR	Export
button	button	button	button	button
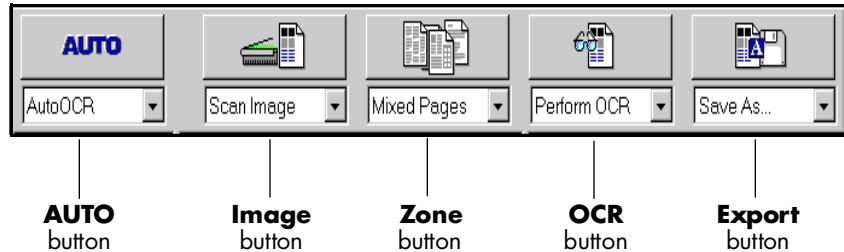
You can set AutoOCR toolbar commands in three locations:

- Click the down arrow below each AutoOCR toolbar button and select a process command in the drop-down list.
- Choose *Process Settings...* in the Process menu.
- Click the Options button and select process commands in the Options dialog box.

▶

The pictures in the AutoOCR toolbar buttons change as you set different process commands. The commands can be activated by clicking the AutoOCR toolbar buttons or choosing commands in the Process menu. A description of the selected process command is displayed below each AutoOCR toolbar button when *Large Buttons* is checked (default setting) in the *Toolbars dialog box. Choose Toolbars... in the* View menu to open the dialog box. Toolbars can be "torn off" and relocated anywhere on your desktop.

All AutoOCR toolbar commands are shown in their drop-down states on a separately enclosed OmniPage Pro 9 Reference Card.

AUTO Button Commands

Use the AUTO button to process a new document from start to finish or to finish processing an open document. The AUTO button's drop-down list contains *AutoOCR* and *OCR Wizard* commands.

AutoOCR

Select *AutoOCR* to finish processing a new or open document according to the selected process commands. See "Automatic Processing" on page 19 for more information.

OCR Wizard

For new documents, select *OCR Wizard* to have the OCR Wizard guide you through the entire OCR process. See "Using the OCR Wizard" on page 18 for information.

Image Button Commands

Use the Image button to bring a document image into OmniPage Pro's image viewer. The Image button's drop-down list contains the *Load Image* and *Scan Image* commands.

Load Image

Select *Load Image* to load existing image files such as TIFF, DCX, BMP, JPG, or PCX files.

Scan Image
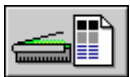
Select *Scan Image* to scan paper documents in your scanner. This command only appears in the drop-down list if you have installed the Caere Scan Manager and have selected your default scanner. Depending upon your scanner model and software drivers, you may have your scanner turned on.

Please see "Bringing Document Images into OmniPage Pro" on page 20 for more information.

Zone Button Commands

Use the Zone button to automatically create zones on document images. Zones are bordered areas that specify what will be recognized as text or retained as graphics on an image. The Zone button's drop-down list contains the *Single-Column Pages, Multiple-Column Pages, Spreadsheet Pages,* and *Mixed Pages* commands and the names of any zone templates you have created. See "Creating Zones for OCR" on page 22 for more information.

Single-Column Pages

Select *Single-Column Pages* to have OmniPage Pro automatically draw and order zones on single-column document images such as letters or memos.

Multiple-Column Pages

Select *Multiple-Column Pages* to have OmniPage Pro automatically draw and order zones on multiple-column document images such as magazine or newspaper articles.

Spreadsheet Pages

Select *Spreadsheet Pages* to have OmniPage Pro automatically draw and order zones on pages that have information arranged in rows and columns such as spreadsheets.

Mixed Pages

Select *Mixed Pages* if your document contains multiple pages with a variety of page layouts. OmniPage Pro will automatically draw and order zones on each page.

Zone Templates

Select a zone template to create zones on document images using that template. See "Creating Zone Templates" on page 73 for more information.

Zone Templates do not appear until you have saved a template. Once created, template names appear, proceeded by the word "Template:" in the drop-down list of the Zone button.

OCR Button Commands

Use the OCR button to perform the selected OCR operation on document images. The OCR button's drop-down list contains the *Perform OCR, OCR and Proof, Train OCR,* and *Defer OCR* commands.

Perform OCR

Select *Perform OCR* to recognize text on document images. During OCR, OmniPage Pro analyzes the image and identifies characters to produce editable text. See "Performing OCR on a Document" on page 23 for more information.

OCR and Proof

Select *OCR and Proof* to recognize text on document images and automatically start checking for errors after OCR. See "Proofreading OCR Results" on page 24 for more information.

Train OCR

Select *Train OCR* to teach OmniPage Pro how to recognize special characters. These pre-recognized characters are saved in a training file, which OmniPage Pro can use to compare with the characters in document images during OCR. See "Training OCR for Special Characters" on page 75 for more information.

Defer OCR

Select *Defer OCR* to delay text recognition during automatic processing. OmniPage Pro will process your document up to the point of OCR and then ask if you want to schedule the document to be finished later. See "Scheduling OCR" on page 80 for more information.

Export Button Commands

Use the Export button to export recognized text and retained graphics to other applications. The Export button's drop-down list contains the *Save As*, *Send Mail*, *Copy to Clipboard*, and *Defer Export* commands.

Save As

Select *Save As* to save a recognized document to disk in a specified file format. See "Saving a Document" on page 34 for more information.

Send Mail

Select *Send Mail* to send a recognized document as a file attached to a mail message if you have a MAPI-compliant mail application, such as Microsoft Exchange or Outlook, installed. See "Sending a Document as a Mail Attachment" on page 37 for more information.

Copy to Clipboard

Select *Copy to Clipboard* to place a copy of a recognized document on the Clipboard. See "Copying a Document to the Clipboard" on page 36 for more information.

Defer Export

Select *Defer Export* if you do not want to export your document right after automatic processing. OmniPage Pro will process your document up to the point of export and then stop.

Selecting OmniPage Pro Settings

Click the Options button or choose *Options...* in the Tools menu to open the Options dialog box. This is the central location for OmniPage Pro settings.

Click each tab to view and select different settings.
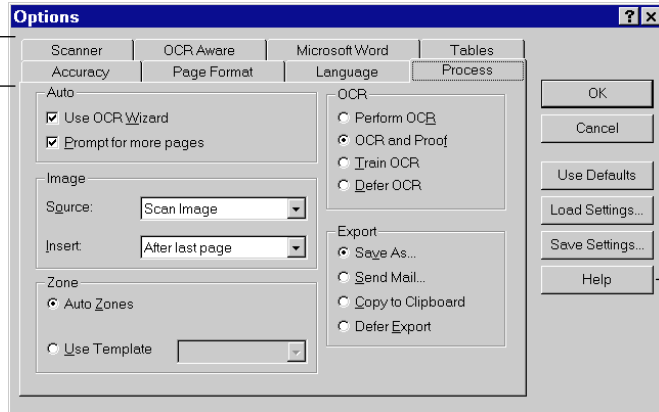
Click for a description of each setting.

▷

Default settings are shown in most examples that follow. However, documents require different settings depending on their input attributes and your output goals. To get the best results, learn how to identify document characteristics and make selections for them. You may have to experiment with different settings to get the results you want. Refer to the "Settings Guidelines" beginning on page 51 for more information.

Accuracy Settings

Click the *Accuracy* tab to select settings that affect OCR accuracy.

The *Language Analyst* evaluates and replaces unknown words with words most likely to be correct during OCR.

Training files help recognize special characters during OCR.

Usually, these setting should be selected for optimal accuracy. Deselect any that cause over correction.
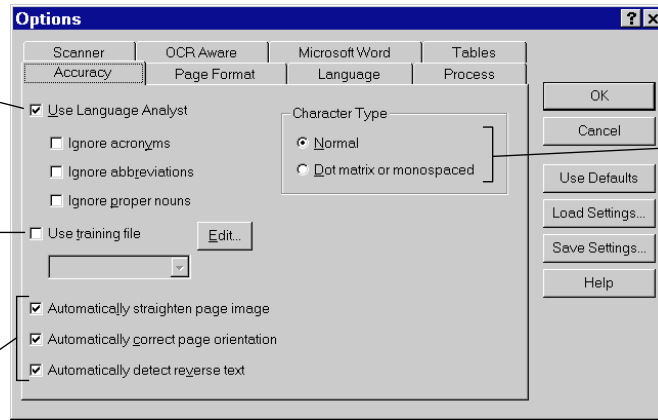
Select the type of characters that are in your document.

Scanner Settings

Click the *Scanner* tab to select settings for scanning pages.

The *Scanner* tab appears only if you have installed Scan Manager, and depending on your particular scanner, you might have to have your scanner connected and turned on for the *Scanner* tab to appear.

This is recommended for black and white pages.

This is recommended for pages with colored backgrounds, colored text, or pages containing grayscale graphics.

This is recommended for pages with color graphics that you want to save.
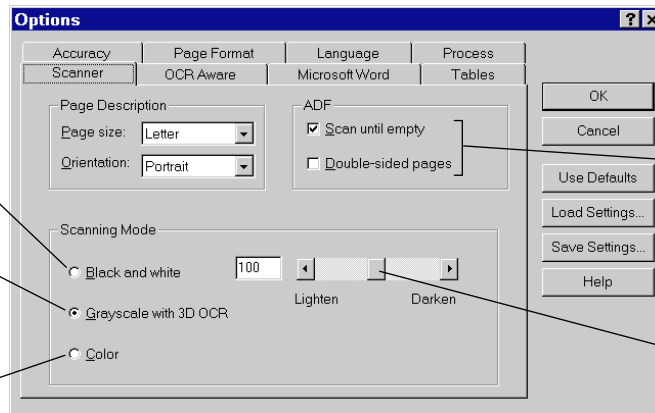
Use these settings if your scanner has an automatic document feeder.

Use the brightness slider to adjust for black and white, grayscale, or color scanning.

Page Format Settings

Click the *Page Format* tab to select settings that determine how the formatting of a page is handled during OCR.

Select a setting that best describes how your original page looks.

The page icons change to depict the general appearance of your page original.

Select a setting to determine what you want your page to look like after OCR.
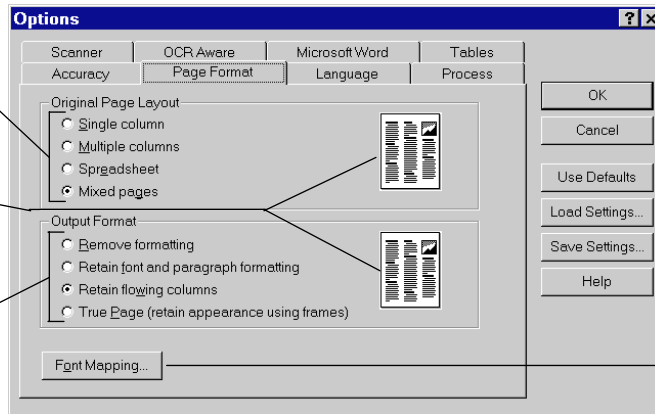
Click to select font options for recognized text

Tables Settings

Click the *Tables* tab to select table settings for your document.

Select to automatically detect tables that have grid lines between rows and columns.

If your target application is Microsoft Word or WordPerfect, you can select *Table objects...* to have tables saved with their grids. Otherwise, tables will be saved as tab-delimited text.
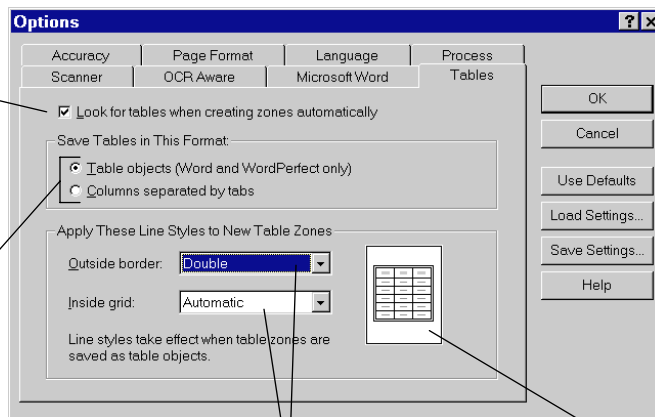
These drop-down menus determine how your table borders will look after export.

Changing the line styles using the drop-down menus will change the page icon to show the general appearance you can expect of the table grids and border after export.

Language Settings

Click the *Language* tab to select language settings for your document.

Select the language that appears most in your document.

Select additional languages for a multilanguage document. You must have installed those languages during installation.

This is the language that will be used in dialog boxes, windows, and menu commends.

This is the character used in place of unknown characters. You can enter your own choice.

OCR Aware Settings

Click the *OCR Aware* tab to select settings for the OCR Aware feature. OCR Aware allows you to initiate OCR from another application. See "Using OCR in Other Applications" on page 29 for more information.

OCR Aware allows you to start scanning and perform OCR from another application.

If your application is not listed, click *Browse...* to locate the application file (`*.exe`) and add it to the *Registered* list box.

An application must be registered to work with OCR Aware.

Click Register Office 97... to register Office 97 applications.

◈ ▶

Some applications may be pre-registered with OCR Aware during OmniPage Pro installation. These applications will display in the *Registered* list box.

To register an application with OCR Aware:

1 Launch the application you want to register and open a document in it.

This will ensure that the application name appears in the list box in step 5.

2 Choose *Options…* in OmniPage Pro's Tools menu.

3 Click the *OCR Aware* tab in the Options dialog box.

4 Make sure that *Enable OCR Aware* is selected.

5 Select the name of the application you want to register in the *Unregistered* list box.

6 Click *Add >>* to add the selected application to the *Registered* list box and then click *OK*.

OmniPage adds the *Acquire Text…* and *Acquire Text Settings…* commands to the File menus of registered applications.

Process Settings

Click the *Process* tab to set commands and settings for each step of OCR.

The OCR Wizard will guide you through the OCR process when you click the AUTO button on the AutoOCR toolbar.

Specifies where newly loaded or scanned images are to be added to an open document.
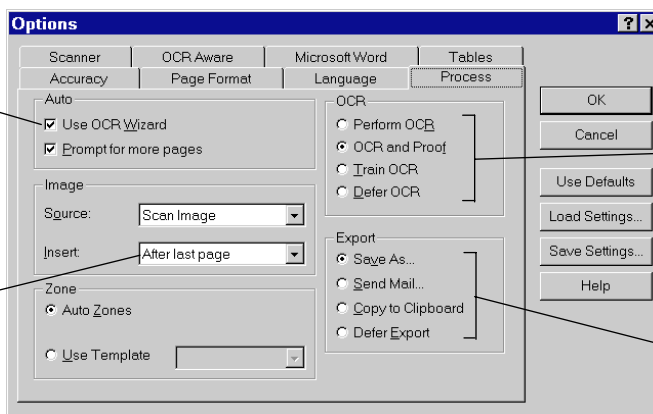
These specify the OCR steps that you want.

These specify how the recognized text is to be exported.

Microsoft Word Settings

Click the *Microsoft Word* tab to select settings for OCR proofreading directly in Microsoft Word. See "Proofreading OCR Results in Microsoft Word" on page 25 for more information.

Select this if you want to check for OCR errors in Microsoft Word.
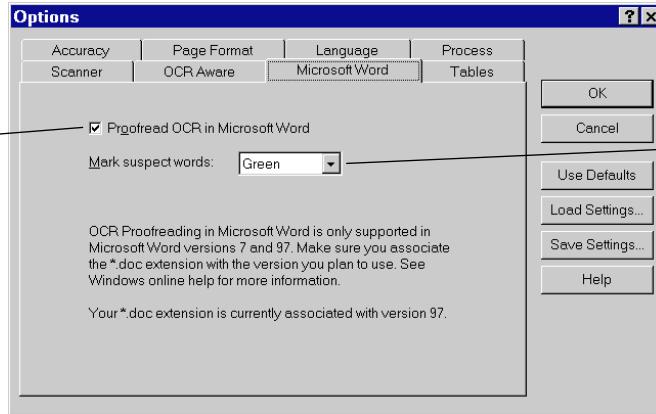
Select the color in which you want suspected errors to appear in Microsoft Word.

Proofreading OCR in Microsoft Word is only supported in Microsoft Word 95, Word 7.0, and Word 97. Make sure you associate the `*.doc` extension with the version you plan to use. Please refer to your Windows documentation for more information.

Settings Guidelines

The settings you select in OmniPage Pro can greatly affect OCR results. Make sure that settings are appropriate for your document *before* you begin processing. You may have to experiment with different settings to get the results you want.

Answer the following questions to get settings recommendations for your documents. Generally, if you indicate the characteristics of your documents to OmniPage Pro, you will receive better OCR results.

- **What type of document are you processing?**

  Magazine and newspaper pages, page 52
  Memos and letters, page 52
  Text and table, page 53
  Spreadsheets, page 53
  Legal documents, page 54
  Mixed formats or not sure, page 54

- **What is the quality of the original document?**

  Poor or not sure, page 55
  Good, page 55

- **How much original formatting do you want to keep?**

  Minimal, page 56
  Some, page 56
  As much as possible, page 57

- **Do you want to retain graphics in your document?**

  Yes, page 58
  No, page 58

- **How many languages are in your document?**

  One language, page 59
  More than one language, page 59

- **Are you processing a multipage document?**

  Yes, page 60
  No, page 60

What type of document are you processing?

Magazine and newspaper pages

Recommendations

- Select *Multiple columns* in the *Page Format* settings.

- Select the appropriate page size and orientation in the *Scanner* settings if you are scanning.

- Draw zones manually or modify automatically created zones if auto zoning does not successfully create zones around all page areas you want to process. See "Customizing Zones" on page 63, for more information. Keep associated sections of text, such as paragraphs, together in one zone. Omit unnecessary parts of the page such as separator lines between columns.

Memos and letters
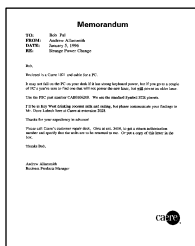
Recommendations

- Select *Single column* in the *Page Format* settings.

- Select the appropriate page size and orientation in the *Scanner* settings if you are scanning.

What type of document are you processing?

Text and table

Recommendations

- In *Page Format* settings, select *Single column* or *Multiple column* page layout depending on the number of columns in your document.

- Select the appropriate page size and orientation in the *Scanner* settings if you are scanning.

- If your table has no grid lines, draw a zone around the table, and set its properties to *Table,* and its content to *Numeric,* unless it has text headings, then select *Alphanumeric*. (Tables with grids are automatically detected).

- Choose the format you want to use to save the table by either selecting *Table objects* or *Columns separated by tabs* in the *Table* settings.

Spreadsheets

Recommendations

- Select *Spreadsheet* and *Retain flowing columns i*n the *Page Format* settings.

- Select the appropriate page size and orientation in the *Scanner* settings if you are scanning.

- Identify the zone content as *Numeric* if *only* numbers (no words or text headers) are in your document.

- Choose the format you want to use to save the table by either selecting *Table objects* or *Columns separated by tabs* in the *Table* settings.
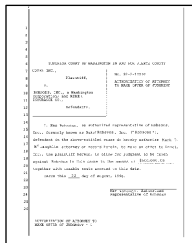
What type of document are you processing?

Legal documents

Recommendations

- Select *Single column* in the *Page Format* settings if the document has one, page-wide text column, even if document has pleading-line numbers.

- Select *Multiple columns* in the *Page Format* settings if text appears in two or more columns.

- Select the appropriate page size and orientation in the *Scanner* settings if you are scanning.

- Draw zones manually or modify automatically created zones to omit unnecessary parts of the page. For example, do not include line numbers in a zone if you plan to renumber lines in your word processor.

- Select *Hard carriage return after every line* in the Save As dialog box if you want to preserve line numbering.

Mixed formats or not sure
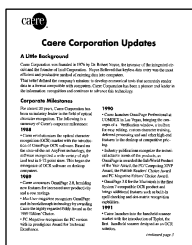
Recommendations

- Select *Mixed pages* in the *Page Format* settings.

- Select the appropriate page size and orientation in the *Scanner* settings if you are scanning.

- Draw zones manually or modify automatically created zones if auto zoning does not successfully create zones around all page areas you want to process. See "Customizing Zones" on page 63, for more information. Keep associated sections of text, such as paragraphs, together in one zone. Omit unnecessary parts of the page such as unwanted graphics.

What is the quality of the original document?

Poor or not sure

Degraded photocopies, colored or shaded backgrounds or text, run-together or broken text characters

thick, run-together text characters

thin, broken text characters

colored text or text on a colored background

Recommendations for scanning

- Select *Grayscale with 3D OCR* in the *Scanner* settings if you have a grayscale scanner and your page contains grayscale graphics, colored background, or colored text.

- For best accuracy, use the *Black and white* setting if your pages are black and white. By using the *Brightness* slider on the *Scanner* tab settings, lighten the setting for thick, run-together text characters or dark backgrounds. Darken the setting for thin, broken text characters.

- Try to scan original documents rather than photocopies.

Other recommendations

- Select *Use Language Analyst* in the *Accuracy* settings. OmniPage Pro will evaluate words and make logical replacements for hard-to-recognize characters.

- Draw zones manually to omit any smudges or scribbles on the page.

- Choose *Proofread OCR...* in the Tools menu to locate possible errors after OCR.

- Choose *Dot matrix or monospaced* in the *Accuracy* settings if you recognize the original font characteristics as such. Choose *Normal* if the text is not monospaced.

- Ask senders to select *Fine* or *Best* mode when they send faxes that you plan to recognize.

Good

Clear, well-formed, black text characters on a clean, white background
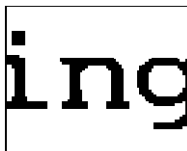
well-formed text characters

Recommendations

- Select *Normal* as the character type in the *Accuracy* settings for the fastest processing if you are scanning.

- Select *Use Language Analyst* in the *Accuracy* settings.

- For faster processing and more accurate results, in the *Language* settings, select only the language as the *Main language* that appears in your document.

- Choose *Proofread OCR...* in the Tools menu to locate possible errors after OCR.

How much original formatting do you want to keep?

Minimal	**Recommendations**
Keep one font and one font size only	• Select *Remove formatting* in the *Page Format* settings. • Click *Font Mapping...* in the *Page Format* settings and select the font and size you want mapped. • Select one of the text-only formats in the Save As dialog box if you want to be able to open the document in any text application.
Some Keep font characteristics and paragraph formatting	**Recommendations** • Select *Retain font and paragraph formatting* in the *Page Format* settings. • Click *Font Mapping...* in the *Page Format* settings and select the fonts you want mapped to various font types. • Save to the file format of your target word processing application for the best results. If you, or the eventual recipient of a file, do not have the exact application, saving to a RTF format will retain some text formatting, such as bold and italics.

How much original formatting do you want to keep?

As much as possible

Keep font characteristics, paragraph formatting, column formatting and graphic positioning

Recommendations

- Select *True Page* in the *Page Format* settings to retain the original appearance of a page using frames. The formatting will be precise but will be more difficult to edit.

- Select *Retain flowing columns* in the *Page Format* settings if your page contains multiple columns and you want text to flow between paragraphs and columns in your target application. The formatting may be less precise than *True Page* but will be easier to edit.

   Please note: The *Retain flowing columns* setting uses frames when necessary to maintain column formatting and graphic positioning. Although frames will appear in the text viewer, only required frames, such as frames around graphics, will be exported.

- Make sure all parts of the page are included within zones. Any part not enclosed within a zone is ignored during OCR and will not appear in the recognized document.

- Select *Retain graphics* in the Save As… dialog box.

- Save to the file format of your target word processing application for the best results.

   If you, or the eventual recipient of a file, do not have the exact application, saving to a RTF format will retain some text formatting, such as bold and italics.

Do you want to retain graphics in your document?

Yes

Keep graphics such as logos and photos during OCR processing

Recommendations for scanning

- Select *Color* in the *Scanner* settings if you are scanning pages with multiple-color graphics and you want to retain the graphics in color.
- Select *Grayscale with 3D OCR* in the *Scanner* settings if you are scanning with a grayscale scanner and you want to retain grayscale graphics.
- Select *Black and white* in the *Scanner* settings if you are scanning line-art drawings.

Other recommendations

- Manually draw zones around graphic areas if necessary.
- Make sure separate zones are drawn around graphic areas and text areas.
- Make sure graphic zones are identified as *Graphic* zone types. Select the zone and right-click with your mouse to determine its properties.
- Select *Retain graphics* in the Save As dialog box when you save a document to another file format.
- To save graphics separately from text after OCR, choose *Save Image…* in the File menu and select *Save each graphic zone to a file*.

No

Ignore graphics such as logos and photos during OCR processing

Recommendations

- Deselect *Retain graphics* in the Save As dialog box when you save a document to another file format.

How many languages are in your document?

One language	**Recommendations**
	• Select the document language as the *Main language* in the *Language* settings.
	If your document contains a language that is not installed in OmniPage Pro, you can add languages to OmniPage Pro by uninstalling and then reinstalling it.
	• For faster processing and more accurate results, select only the language that appears in your document in the *Language* settings.
More than one language	**Recommendations**
	• Select the main document language and any additional languages in the *Language* settings.
	If your document contains languages that are not installed in OmniPage Pro, you can add languages to OmniPage Pro by uninstalling and then reinstalling it.
	• For faster processing and more accurate results, select only the languages that appear in your document in the *Language* settings.

Are you processing a multipage document?

Yes	**Recommendations if you have an automatic document feeder (ADF)**
	• Select *Scan until empty* in the *Scanner* settings to scan a stack of pages at once. Otherwise, you must click the Image button to scan each subsequent page.
	• Select *Double-sided pages* to scan pages with print on both sides. You will be prompted to turn the stack over.
	• Insert blank (paper) pages to separate more than one job within a stack of pages. You can save pages between blank pages as separate files after OCR.
	Other recommendations
	• Set the desired process commands and click *AUTO* to automatically process each page of your document in order.
	• Create and use a zone template if all pages have similar zoning requirements. See "Creating Zone Templates" on page 73 for more information.
	• Choose *Schedule OCR...* in the Process menu to schedule processing for a specific time. Pick a time that you plan to be away from your computer.
	• After OCR, choose *Save As...* in the File menu. You can select an option to save the recognized document as a single file, one file per page, or a new file after each blank page.
No	**Recommendations**
	• Set the desired process commands and click *AUTO* to automatically process the page.
	• Click the Image button to add more pages to the document by scanning or loading images.

Customizing OCR

OmniPage Pro has many features that allow you to customize the way your documents are handled during OCR. This chapter describes how to use these features.

Please continue reading this chapter for information on these topics:

- Adjusting Page Images Before OCR
- Customizing Zones
- Specifying Fonts
- Training OCR for Special Characters
- Creating User Dictionaries
- Saving Settings Files
- Scheduling OCR

Adjusting Page Images Before OCR

You can rotate and straighten page images in OmniPage Pro's image viewer before zoning and OCR take place. This is recommended to improve OCR accuracy on pages that are not oriented correctly.

⚠️

If you need to rotate or straighten a page, be sure to do so *before* you create zones because all zones are deleted during these operations.

To rotate a page image:

1 Click on the page image to make the image viewer active.

2 Click the Rotate Image button to rotate the image 90-degrees (clockwise) at a time.

   Or, choose *Rotate* ▸ in the View menu and select 90, 180, or 270 degrees.

To straighten a page image:

1 Click on the page image to make the image viewer active.

2 Click the Straighten Image button.

   Or, choose *Straighten Image* in the View menu.

   OmniPage Pro straightens the page image up to a maximum of 10 degrees. OmniPage Pro will not straighten a page if it determines that it is unnecessary.

▷

It is recommended that you have OmniPage Pro automatically rotate or straighten pages if needed during OCR by simply selecting the *Automatically straighten page image* and *Automatically correct page orientation* options in the *Accuracy* tab of the Options dialog box.

Customizing Zones

Zones are borders created around areas of a page image to identify what will be recognized as text or retained as a graphic during OCR. Zones play a big part in determining OCR results.

You can create zones automatically, manually, or with a template. Topics in this section describe how you can customize zones including:

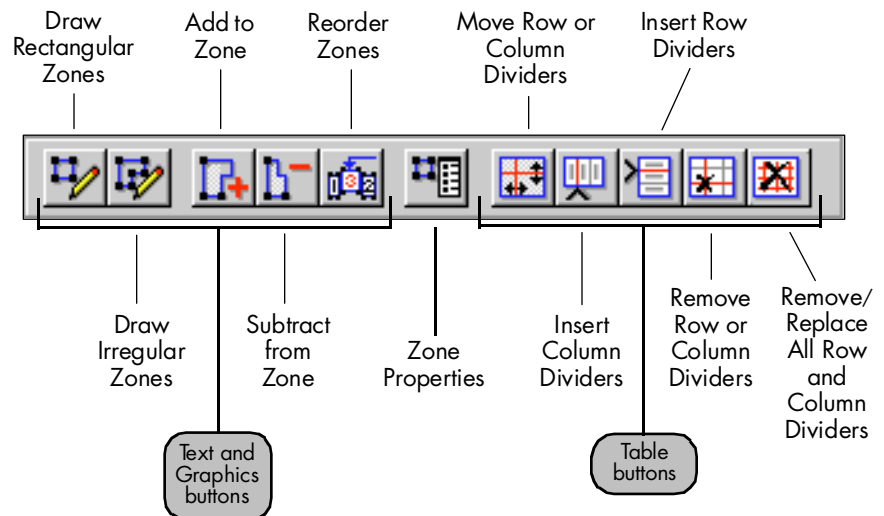- Drawing Zones Manually
- Modifying Text and Graphic Zones
- Modifying Table Zones
- Deleting Zones
- Changing Zone Properties
- Creating Zone Templates

For information on creating zones automatically, please see "Creating Zones for OCR" on page 22.

Zone toolbar

The Zone toolbar contains buttons for drawing and modifying zones. Toolbars can be "torn off" and relocated anywhere on your desktop.

Drawing Zones Manually

You can draw zones manually on an image using buttons in the Zone toolbar. Rectangular zones are the most common, but you can also draw irregular-shaped zones for graphics and text. Only rectangular (and square) zones are allowed for tables.

To draw rectangular zones:

1   Click the Draw Rectangular Zones button.

    The mouse pointer in the image viewer becomes a drawing tool.

2   Enclose an area of the image you want as a zone by holding down the mouse button and dragging the drawing tool to form a rectangular box.

    Try to keep areas of text, such as paragraphs or single columns, together in the same zone.

3   Release the mouse button when you are done.

    A number appears within the zone indicating its processing order.

You cannot draw overlapping zones. If you attempt to draw a zone over an existing zone, the borders of the new zone will wrap *around* the boundaries of the existing zone when you release the mouse button.

To draw irregular-shaped zones:

1   Click the Draw Irregular Zones button.

    The mouse pointer in the image viewer becomes a drawing tool.

2   Position the drawing tool where you want to start drawing the first side of the zone.

3   Click the mouse button once.

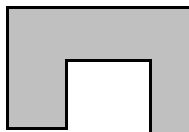4   Move the drawing tool to form the first side of your zone.

5   Click the mouse button when you have drawn the desired line length.

6   Draw a perpendicular line in either direction to form the next side of the zone.

7   Repeat steps 6 and 7 to finish drawing each side of your zone.

You will not be allowed to draw a line if it constitutes a restricted shape. The following zone shapes are restricted:

Indented along   Indented along
the bottom   the top
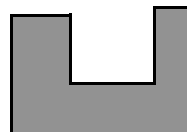
To draw a table zone:

1   Click the Zone Properties button and select *Table zone* as the zone type.
   See "Changing Zone Properties" on page 71 for more information.

2   Click the Draw Rectangular Zones button.
   The mouse pointer in the image viewer becomes a drawing tool.

3   Enclose an area of the image you want as a table zone by holding down the mouse button and dragging the drawing tool to form a rectangular or square box.

4   Release the mouse button when you are done.
   Row and column dividers appear in the table zone. You can adjust, add, or remove the dividers using other toolbar buttons.

5   Repeat steps 3 and 4 until you have finished drawing table zones around the desired areas of the page.

Modifying Text and Graphic Zones

You can modify zones by moving, resizing, reordering, extending, subtracting, connecting, or dividing them.

To move zones:

1   Deselect the buttons in the Zone toolbar.
   (If one of the first two drawing buttons is selected, you do not have to deselect it.)

2   Place the mouse pointer inside a zone.

3   Hold down the mouse button and drag the zone to the desired location.

To resize zones:

1 Deselect the buttons in the Zone toolbar.
(If one of the first two drawing buttons is selected, you do not have to deselect it.)

2 Select the zone you want to resize by clicking inside it.
The selected zone is shaded and handles appear on its border.

3 Place the mouse pointer over a handle so that it changes to a two-way arrow.

4 Hold down the mouse button and drag the handle in the direction that you want to enlarge or reduce the zone.

5 Release the mouse button when you are done.
The zone border changes to display the modified zone area.

To reorder zones:

1 Click the Reorder Zones button.
The numbers in the zones disappear.

2 Click within the zone you want recognized first.
The number *1* appears in the zone.

3 Click within the zone you want recognized next.
The number *2* appears in the zone.

4 Repeat step 3 until all the zones are appropriately ordered.
If you do not number all the zones, they are automatically numbered for you when you start OCR.

The numbered order of zones determines the order in which text will be placed on a recognized page. However, if you select *True Page* or *Retain flowing columns* as the Output Option for a page, the order of the text will be based on the order of the original page.

To extend an area of a zone:

1 Click the Add to Zone button.
The mouse pointer in the image viewer becomes a drawing tool with a plus sign.

2 Position the drawing tool at the point where you want to start extending the zone.
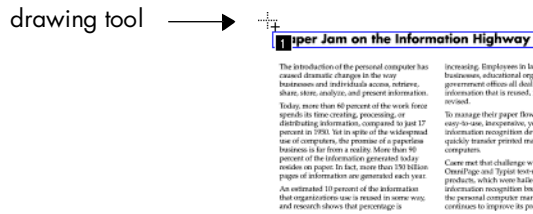
drawing tool ⟶

3 Hold down the mouse button and drag the drawing tool in the direction that you want to extend the zone.

4 Release the mouse button when you are finished extending the zone.

The zone border changes to display the modified zone area.

The left area of this zone has been extended downward.

To subtract an area of a zone:

1 Click the Subtract from Zone button.

The mouse pointer in the image viewer becomes a drawing tool with a minus sign.

2 Position the drawing tool at the point where you want to start subtracting from the zone.
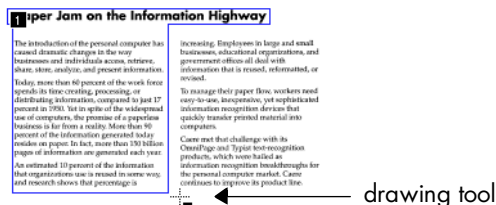
drawing tool

3 Hold down the mouse button and drag the drawing tool in the direction that you want to subtract from the zone.

4   Release the mouse button when you are finished subtracting from the zone.

The zone border changes to display the modified zone area.

Table zones are constrained to rectangular and square shapes. Attempting to modify the area of a table zone to an irregular shape is not allowed. Table zones, however, can be resized, and it is recommended that you resize the table zone as described on "To extend an area of a zone:" on page 66.

To connect two or more zones:
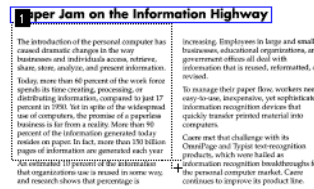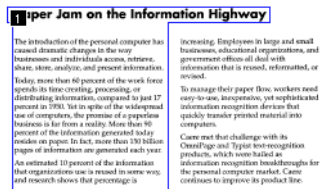
1   Click the Add to Zone button.

The mouse pointer in the image viewer becomes a drawing tool with a plus sign.

2   Hold the mouse button down and drag the drawing tool over the area where you want the zones to be connected.

3   Release the mouse button when you are done.

The zone border changes to display the modified zone area.

To divide a zone:

1   Click the Subtract from Zone button.

The mouse pointer in the image viewer becomes a drawing tool with a minus sign.

2   Hold the mouse button down and drag the area where you want to divide the zone.

3   Release the mouse button when you are done.

The zone border changes to display the modified zone area.

Modifying Table Zones

You can modify table zones by moving, resizing, reordering, extending, subtracting zones, and adding or removing table grids.

To move dividers in a table zone:

1   Click the Move Row or Column Dividers button.

2   Place the mouse pointer within the table zone in the image viewer.

The mouse pointer becomes a vertical- or horizontal-bar tool depending on which divider is being passed over.

3   Hold the mouse button down and drag the row or column divider you want to move. Ctrl-clicking a column divider will move only a column divider for single cell. Rows dividers, however, cannot be moved one cell at a time.

The selected divider moves. The cell divider, however, cannot be moved beyond its own cell.

4   Release the mouse button when you are done.

To insert column dividers in a table zone:

1   Click the Insert Column Dividers button.

2   Place the mouse pointer within the table zone where you want to insert a column divider.

The mouse pointer becomes an upward-facing caret (**^**) with a dimmed vertical bar.

3   Click the mouse button.

A new column divider is inserted in the table. Hold down the Ctrl key to insert a column divider only for a single cell; the cell that contains the mouse pointer.

To insert row dividers in a table zone:

1   Click the Insert Row Dividers button.

2   Place the mouse pointer within the table zone where you want to insert a column divider.

The mouse pointer becomes a right-facing caret **(>)** with a dimmed horizontal bar.

3   Click the mouse button.
A new row appears.

To remove a row or column divider from a table zone:

1 Click the Remove Row or Column Dividers button.

2 Place the mouse pointer within the table zone where you want to remove a row or column.
The mouse pointer becomes a small "x" with a dimmed bar.

3 Position the bar on the divider you want to remove and click the mouse button. Ctrl-clicking a column divider will remove only the column divider from a single cell. To remove a row divider, the whole row divider must be removed.
The selected divider disappears.

To remove all row or column dividers from a table zone:

1 Click the Remove/Replace All Row and Column Dividers button.

2 Place the mouse pointer within the table image zone.
The mouse pointer becomes a large "X".

3 Click the mouse button.
The internal row and column dividers disappear, and the mouse pointer changes to a grid. Clicking the mouse button again restores all the rows and column dividers to the originally drawn positions.

If you are dissatisfied with a change you have made to a table divider, you can cancel your last alteration with the Undo command, Ctrl-Z.

Additionally, you can insert a set of row and column dividers in a table zone by clicking the Remove/Replace All Row and Column Dividers button and then clicking in the table zone. This works only if your have previously used the Remove/Replace All Rows and Columns tool on the zone.

Deleting Zones

You can delete the current zones if you want to create new zones. You can also delete individual zones that you do not want to process during OCR. Any part of a page image not enclosed by a zone is ignored during OCR.

▶

To delete and replace the current zones automatically, click the Zone button in the AutoOCR toolbar. You will be prompted to replace the current zones.

To delete zones:

1 Select the zone you want to delete by clicking inside the zone.
  - Shift-click to select additional zones.
  - Choose *Select All* in the Edit menu to select all zones on the current page.

  Selected zones are shaded.

2 Press the Delete key or choose *Clear* in the Edit menu.
  The selected zones disappear.

Changing Zone Properties

You can set certain properties for zones to customize how each zone will be treated during OCR. The Zone Properties dialog box contains settings for *zone type* and *zone content*.
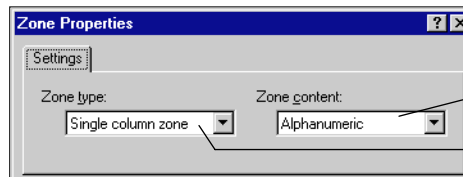
Zone content drop-down list

Zone type drop-down list

▶

When you change a zone type using the Zone Properties button, newly drawn zones *and* any previously selected zones will change zone type.

Zone Type

Every zone on a page has a zone-type setting. You can select the following zone types:

- *Single-column text zone* for text zones that contain a single column
- *Multiple-column text zone* for text zones that contain multiple columns
- *Table zone* for text or numeric zones that contain data in rows and columns
- *Auto-detect zone* for the automatic detection of zone-content type (not usually recommended)
- *Graphic zone* for photos, drawings, and areas of text that you want to retain as a graphic
- *Reverse-text zone* for single columns of light text on dark background

Zone Content

All text zones on a page also have a zone-content setting. This specifies the characters OmniPage Pro looks for within a zone during OCR. You can select *Alphanumeric* or *Numeric* as the zone-content setting.

For example, if a particular zone only contains numbers and mathematical signs, you can specify the contents of that zone to be *Numeric*. OmniPage Pro will only look for numeric characters in that zone during recognition.

▶

OmniPage Pro assigns a zone type and Alphanumeric contents to each zone when it creates zones automatically. This is true for all zones recognized automatically, except graphic-content zones. You do not need to change the zone properties unless you want to modify the way zones will be treated during OCR.

To change the properties of a zone:

1   Select the zone you want to modify by clicking it.
    You can Shift-click to select multiple zones. Selected zones are shaded.

2   Click the Zone Properties button to open the Zone Properties dialog box.

The settings in this dialog box will be blank if multiple zones with different settings are selected.
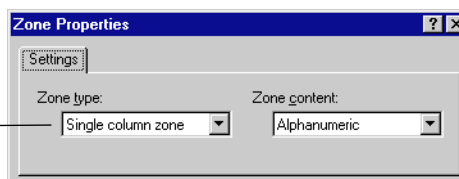
Close button

3   Select a zone type for the selected zones.

If you change an irregular-shaped zone to a Table type zone, OmniPage Pro substitutes the largest rectangle that fully encloses the irregular area.

4   Select a zone content for the selected zones.

You can select a zone-content setting for any zone type except Graphic.

5   Click the Close button when you are done.

You can also change a zone's type and content settings individually by clicking your right-mouse button over the zone and choosing a setting in the shortcut menu that appears.

Creating Zone Templates

A zone template has attributes including size, shape, position, order, type, and content. Zone templates are useful to create zones on an image if you process a lot of documents that have the same layout and content.

To create a zone template:

1   Load a page image and create the desired zones.

2   Choose *Save Zone Template*... in the Tools menu.
     The New Template dialog box appears.

3   Type a name for your file in the *File name* text box.

4   Click *OK*.
     The zone template file is saved in the data folder in your installation folder. Select it in the Zone button drop-down list.

To create zones with a template:

1   Select the zone template that you want to use in the Zone button drop-down list.

2   Click the Zone button or choose *Template* in the Process menu.
     OmniPage Pro creates your predefined zones on the page image using the zone template.

Specifying Fonts

You can retain the font characteristics in your document during OCR if you select an Output Format option other than *Remove formatting* in the *Page Format* tab of the Options dialog box.

OmniPage Pro automatically *maps* detected font types to specified fonts. To map fonts, OmniPage Pro analyzes text and categorizes it as one of these font types:

- Proportional Serif

    Character spacing varies depending on the character; short lines finish off the letter strokes. The body text in this manual is an example of this font type.

- Proportional Sans-Serif

    Character spacing varies depending on the character; letter strokes do not have finishing lines. The headings in this manual are an example of this font type.

- Monospaced Serif

    Character spacing is the same for each character; short lines finish off the letter strokes. `Courier` is an example of this font type.

- Monospaced Sans-Serif

    Character spacing is the same for each character; letter strokes do not have finishing lines. `Letter Gothic` is an example of this font.

To customize the font mapping for font types:

1   Choose *Options...* in the Tools menu to open the Options dialog box.

2   Click the *Page Format* tab.

3   Click *Font Mapping...* to open the Font Mapping dialog box.

The selected fonts are applied to text when their corresponding font types are detected during OCR.
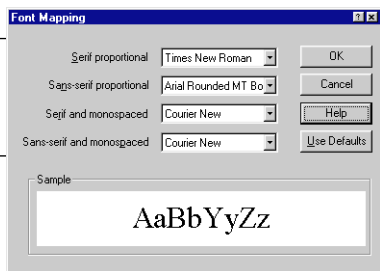
4   Select the font you want mapped to each font type.

    The fonts available in the drop-down lists depend on the True Type fonts installed on your system.

5   Click *OK* when you are done.

Training OCR for Special Characters

A *training file* is a set of pre-recognized text characters that OmniPage Pro compares with characters on a page image during OCR. You can create a training file for special characters that might normally be difficult to recognize such as the copyright symbol © or the registered trademark symbol ®.

To create a training file:

1 Open the image file or scan the page that includes characters you want to train.

2 Create zones around the text that you want to train.

3 Set *Train OCR* as the command in the OCR button's drop-down list.

4 Click the OCR button or choose *Train OCR* in the Process menu. OmniPage Pro analyzes the document and then opens the Train Characters dialog box.

Original character image

OmniPage Pro's
interpretation of the image
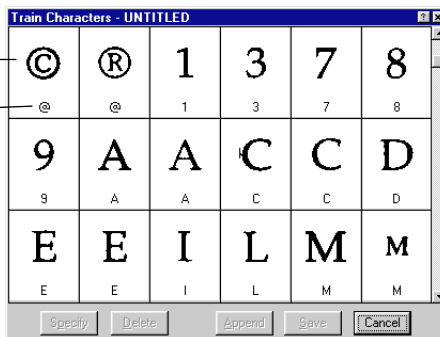
5 Double-click a character you want to train. Or select it and click *Specify*.

Most characters do not need to be trained. Look for uncommon characters such as the copyright symbol ©.

Do not train OmniPage Pro for regular characters because it may interfere with recognition.

The Specify Character dialog box shows how the selected character appeared in the original page image.

Click the character you want to associate with the selected character

The associated character appears here

The original image of the selected character
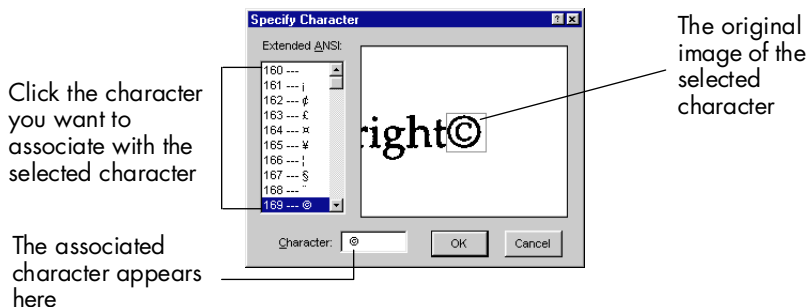
6   Specify how you want OmniPage Pro to interpret the character during OCR by entering a character in the *Character* edit box.

7   Click *OK* to return to the Train Characters dialog box.

8   Repeat steps 5–7 to continue specifying characters.

9   Click *Save* to save the specified characters to a training file.
    Or, click *Append* to add the specified characters to another training file.
    After saving or appending to a file, you are asked if you want to make this the current training file. Click *Yes* to recognize the current page using the training file you just created. Click *No* to return to the image without recognizing it.

◆▶

Training files are saved in the data folder in your installation folder. You can select them in the *Accuracy* tab of the Options dialog box.

To edit a training file:

1   Choose *Edit Training File...* in the Tools menu.
    A dialog box appears listing all your training files.
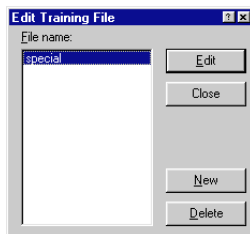
2   Double-click the training file you want to edit. Or, select it and click *Edit.*

The Train Character dialog box displays characters in the selected file.

Original image ———

Associated characters ———
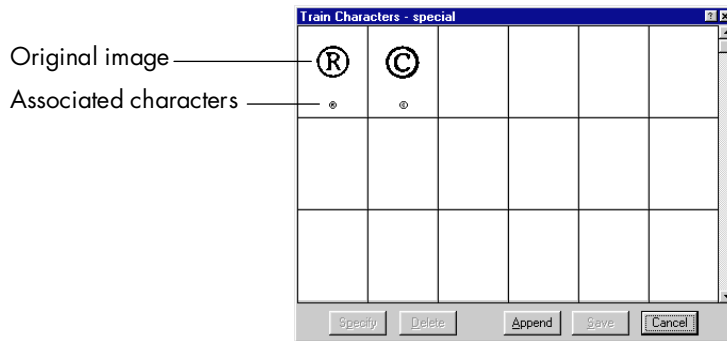
3   Edit the characters as desired.

   • Double-click a character that you want to edit.

   • Click a character that you want to remove and click *Delete*.

4   Do one of the following after editing the training file:

   • Click *Save* to save changes in the training file.

   • Click *Append* to add all trained characters to another training file.

   • Click *Cancel* to exit without saving the edits to the training file.

Creating User Dictionaries

Two dictionaries are used when you perform OCR and check for errors: the dictionary for the language you are using, and a user dictionary where you can add special words manually. You can create multiple user dictionaries, but you can only use one at a time. You can select a user dictionary in the *Language* tab of the Options dialog box.

To customize a user dictionary:

1   Choose *Edit User Dictionary...* in the Tools menu.
     A dialog box lists all user dictionary files.

This is Microsoft Word's user dictionary. You can use it with OmniPage Pro.

This is OmniPage Pro's default user dictionary.
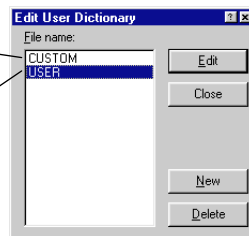
2   Do one of the following:

- Select a file and click *Edit* to edit an existing user dictionary.

- Click *New* to create a new user dictionary. Enter a name in the dialog box that appears and click *OK*.

  The User Dictionary dialog box appears.

Words in the user dictionary appear in this list box.
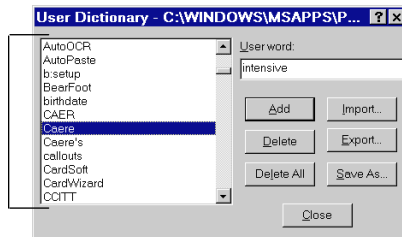
3   Add or delete words as desired:

- Type a word in the *User word* edit box and click *Add* to add it.

- Select a word in the list box and click *Delete* to delete it. Click *Delete All* to remove all words from the dictionary.

- Click *Import...* to add words from a text file.

4   Click *Close* when you are finished editing the user dictionary. OmniPage Pro's user dictionaries are saved in the data folder in your installation folder.

Saving Settings Files

You can save OmniPage Pro settings to a file. A settings file is useful for quickly loading particular settings that you need for certain documents.

The settings you select in OmniPage Pro can greatly affect OCR results. For help in selecting settings for different kinds of documents, see "Settings Guidelines" on page 51.

To save settings to a file:

1   Choose *Options...* in the Tools menu.

2   Select the desired settings in the Options dialog box.

3   Click *Save Settings...* to open the Save Settings dialog box.
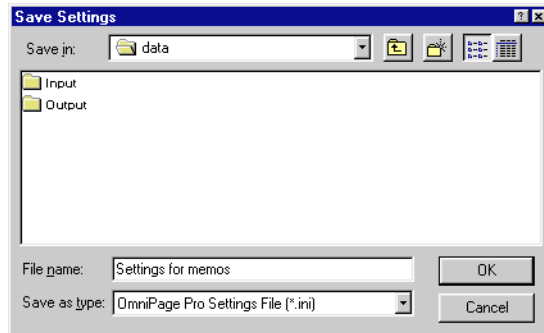
4   Select a folder location for the settings file.

5   Type in a file name for the settings file and click *OK*.
     All the current settings in the Options dialog box are saved into
     a settings file with an .ini extension.

6   Click *OK* to close the Options dialog box.

To load a settings file:

1   Choose *Options...* in the Tools menu to open the Options dialog
     box.

2   Click *Load Settings...* to open the Load Settings dialog box.
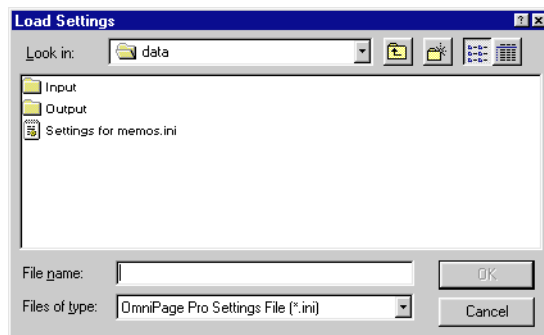
3   Select the folder location of the settings file you want to load.

4   Select the name of the settings file you want to load and click
     OK.
     The settings change according to the selected file.

5   Click *OK* to close the Options dialog box.

Scheduling OCR

You can schedule OCR to take place on one or more OmniPage documents, supported image files, and pages in your scanner. This processing can take place while you are away from your computer as long as OmniPage Pro is still running. Scheduled documents are opened at the specified time, unfinished pages are recognized, and the documents are saved in a preselected format and location.

⚠

Scheduled documents are deleted from the processing queue if you close OmniPage Pro. Therefore, you should keep OmniPage Pro running until the documents are processed.

Topics in this section include:

- Scheduling Individual Documents
- Scheduling Documents from an Input Folder
- Modifying Output Options for Documents

Scheduling Individual Documents

You can schedule individual documents from different folders. Scheduled documents are recognized at the specified time and then saved in the designated output folder.

To schedule individual documents:

1   Choose *Schedule OCR...* in the Process menu.
    The Schedule OCR dialog box appears.

All scheduled documents are displayed in this processing queue.

Click this to modify default output options.

OmniPage Pro starts processing scheduled documents, in order, at the specified time.

Click *Add...* to add documents to the processing queue.

Click *Remove* to remove a selected document from the processing queue.

2   Click *Add...* to open the Add Jobs dialog box.
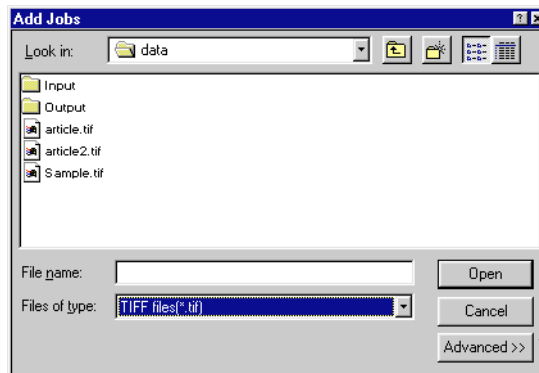
Click *Advanced* to select documents from more than one folder.

3   Locate and select the files you want to add to the schedule.
     You can select OmniPage Documents and supported image
     files.

4   Click *Open* after selecting the desired files.
     The Schedule OCR dialog box displays the newly added files.

5   Select the time that you want OmniPage Pro to process the
     scheduled documents.
     Select *Finish now* if you want OmniPage Pro to process all
     scheduled documents as soon as you close the dialog box.

6   Click *OK* in the Schedule OCR dialog box to save your settings
     as specified.
     All scheduled files are processed, in order, at the scheduled
     time.

Scheduling Documents from an Input Folder

You can set up OmniPage Pro to automatically schedule documents
from a specified input folder. Scheduled documents are recognized at
the specified time and then saved in the designated output folder.

To schedule documents from an input folder:

1 Choose *Schedule OCR...* in the Process menu.
The Schedule OCR dialog box appears.

All scheduled documents are displayed in this processing queue.

Click this to modify default output options.

OmniPage Pro starts processing documents in the queue at the specified time.
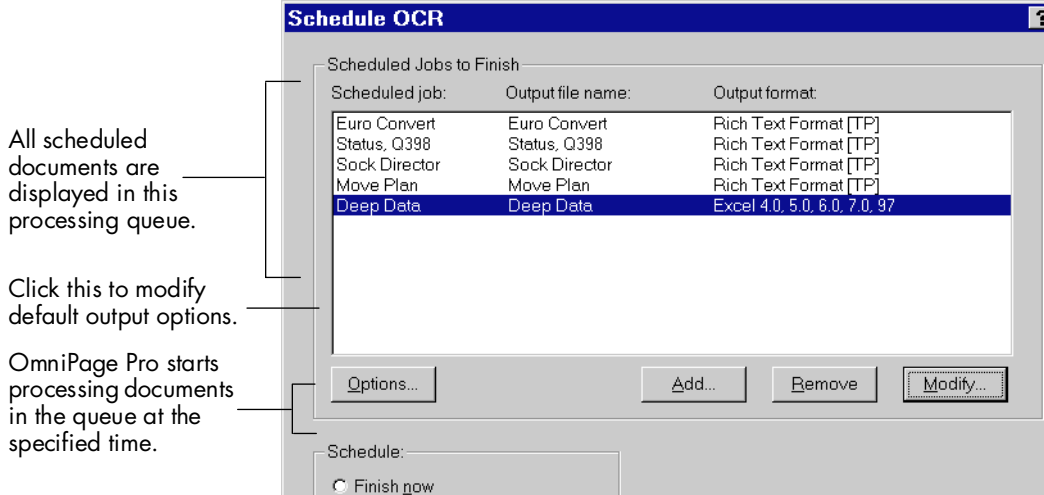
2 Click the *Options...* button to open the Schedule OCR Options dialog box.

The selected output options are used for all newly scheduled documents.

Select this to schedule documents in your scanner's automatic-document feeder (ADF).

Select this to automatically schedule documents in the specified folder.
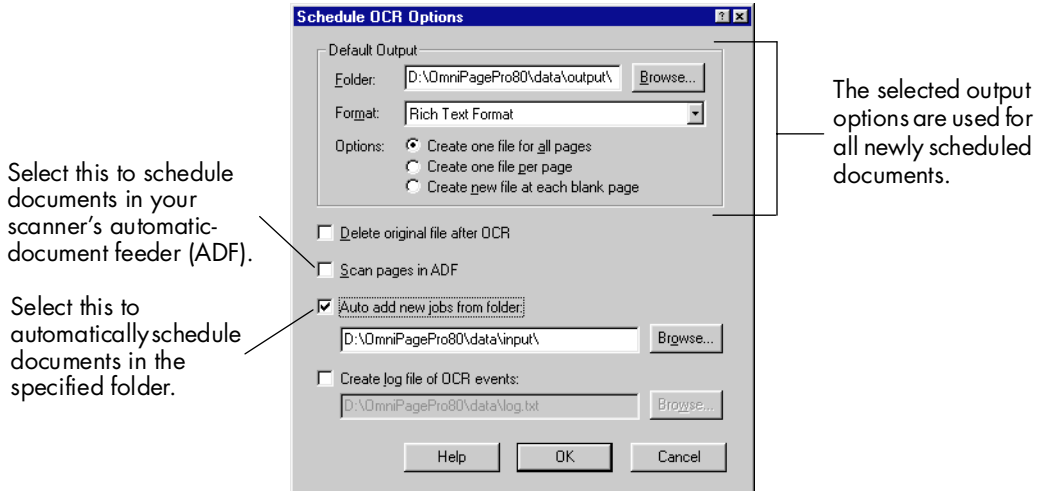
3 Select *Auto add new jobs from folder* and select the desired input folder.

▶

If you use the auto-add feature to schedule documents and you do not select *Delete original file after OCR* , original files will be moved from the input folder to the output folder after processing.

4   Click *OK* in the Schedule OCR Options dialog box to accept the selected settings.

The Schedule OCR dialog box reappears and adds documents from the input folder to the processing queue.

5   Select the time that you want OmniPage Pro to process scheduled documents.

6   Click *OK* in the Schedule OCR dialog box to save the settings and close the dialog box.

Processing begins at the specified time. Right before processing begins, OmniPage Pro checks the input folder again and adds any new documents to the processing queue.

⚠

After scheduled jobs are processed, the *Auto add new jobs from folder* option will be deselected.

Modifying Output Options for Documents

All newly scheduled documents have the same default output folder and file format assigned to them. The default output file name uses the original file name and the extension of the output file format. You can modify all of these output options for any scheduled document.

▶

Click the *Options...* button in the Schedule OCR dialog box to change the default options used for all newly scheduled documents.

To modify the output options for an individual document:

1 Choose *Schedule OCR...* in the Process menu.
The Schedule OCR dialog box appears.

Select the document for which you want to modify output options.

Click this to modify default output options.

Click this to modify the output options for the selected document.
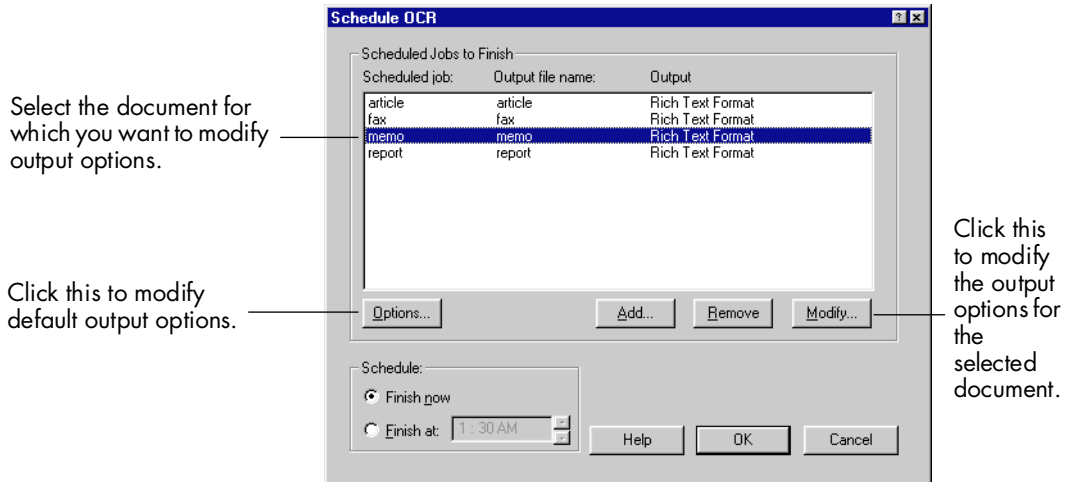
2 Select a scheduled file and click *Modify…* to open the Modify Scheduled Job dialog box.

Select output options for this particular document.

Select this if you want the original document deleted after processing.
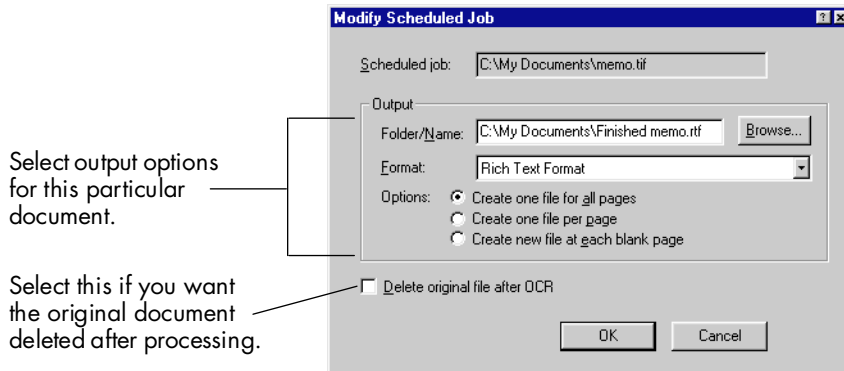
3 Select the desired options for the document.

4 Click *OK* to accept the selected options.
The Schedule OCR dialog box reappears.

5 Click *OK* to close the Schedule OCR dialog box.

Technical Information

This chapter provides troubleshooting and other technical information about using OmniPage Pro.

Please also read the online *Readme* file and the *Scanner Setup Notes*. The *Scanner Setup Notes* list all supported scanners and any connection or software-driver issues. The Readme file contains last-minute information relating to OmniPage Pro. To open these documents, click *Start* in the Windows taskbar and choose: *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *Scanner Setup Notes* or *Readme.*

Please continue reading this chapter for information on these topics:

- General Troubleshooting Solutions
- Supported File-Format Types
- Scanner Setup Issues
- OCR Problems
- Uninstalling the Software

General Troubleshooting Solutions

Although OmniPage Pro is designed to be easy to use, problems sometimes occur. Many of the onscreen error messages contain self-explanatory descriptions of what to do — check connections, close other applications to free up memory, and so on. Sometimes that is all the troubleshooting help you need.

Please see your Windows documentation for information on optimizing your system and application performance.

Topics in this section include:

- Solutions to Try First
- Testing OmniPage Pro
- Low Memory Problems
- Low Disk Space Problems

Solutions to Try First

Try these possible solutions if you experience problems using OmniPage Pro:

- Make sure that your system meets all requirements listed under "Minimum System Requirements" on page 2.

- Make sure that your scanner is plugged in and that all cable connections are secure.

- Turn off your computer and your scanner, turn your scanner back on, and then restart your computer. Make sure other applications are functioning properly.

- Use the software that came with your scanner to verify that the scanner works properly before using it with OmniPage Pro.

- Make sure you have the correct drivers for your scanner, printer, and video card. See the *Scanner Setup Notes* for more information by clicking *Start* in the Windows taskbar and choose *Programs>Caere Applications>Caere Documents >Scanner Setup Notes.*

- Run ScanDisk for Windows 95 or 98, or Check Disk for Windows NT to check your hard disk for errors. See Windows online help for more information.

- Defragment your hard disk. See Windows online help for more information.

- Uninstall and reinstall OmniPage Pro and the Scan Manager.

Testing OmniPage Pro

Restarting Windows 95 or 98 in *safe mode* or Windows NT in *VGA mode* allows you to test OmniPage Pro on a simplified system. This is recommended when you cannot resolve crashing problems or if OmniPage Pro has stopped running altogether. See Windows online help for more information.

⚠️

Your scanner will not run with OmniPage Pro in safe mode or VGA mode, so do not test scanner problems in this configuration.

To test OmniPage Pro in safe mode (Windows 95 or 98):

1   Restart your computer in safe mode by pressing F8 immediately after you see the "Starting Windows" message.

2   Launch OmniPage Pro and try performing OCR on an image. Use an existing image file such as the `Sample.tif` file.

    • If OmniPage Pro does not launch or run properly in safe mode, then there may be a problem with the installation. Uninstall and reinstall OmniPage Pro, and then run it in Windows safe mode.

    • If OmniPage Pro runs in safe mode, then a device driver on your system may be interfering with OmniPage Pro operation. Troubleshoot the problem by restarting Windows in Step-by-Step Confirmation mode. See Windows online Help for more information.

To test OmniPage Pro in VGA mode (Windows NT):

1   Restart your computer.

2   Select *Windows NT Workstation Version 4.00 [VGA mode]* and press Enter.

3   Press Ctrl+Alt+Delete and select *Task Manager*.

4   In the Task Manager dialog box, select all background applications and click End Process. See your Windows documentation for more information.

5   Launch OmniPage Pro and try performing OCR on an image. Use an existing image file such as the `Sample.tif` file.

Low Memory Problems

OmniPage Pro may run poorly under low-memory conditions. This may be indicated by various error messages or if OmniPage Pro works slowly and accesses the hard drive often. Try these solutions for low memory conditions:

- Restart your computer.
- Close other open applications to release memory.
- Close unnecessary OmniPage Pro windows.
- Defragment your hard disk to free up contiguous blocks of disk space. See Windows online help for instructions.
- Increase the amount of free hard disk space.
- Increase your computer's physical memory (RAM).
  More memory optimizes OCR performance. See "Minimum System Requirements" on page 2 for more information.

Low Disk Space Problems

Problems may occur if your system runs low on free disk space. Try these solutions for low disk space problems:

- Empty the Windows Recycle Bin.
- Close all open applications and delete the `*.tmp` files in the `Temp` folder. This folder is usually located in your Windows folder.
- Run ScanDisk for Windows 95 or 98, or Check Disk for Windows NT to check for errors that may be using disk space. See Windows online help for instructions.
- Back up unneeded files onto floppy disks or other media and delete them from your hard disk.
- Remove Windows applications that you do not use.
- Defragment your hard disk. See Windows online Help for instructions.
- Clear the cache for your web browser and limit its size.

Supported File-Format Types

OmniPage Pro can open these file-format types:

BMP, Bitmap (`*.bmp`)	OmniPage Document (`*.met`)[†]
DCX (`*.dcx`)	PCX (`*.pcx`)
JPEG (`*.jpg`)	TIFF uncompressed (`*.tif`)[‡]
TIFF Packbits (`*.tif`)	TIFF Group 3 or 4, compressed (`*.tif`)[‡]

[†]Caere Documents from version 8.0 can be opened if the original images were preserved as .tif or .jpg files.

[‡]TIFF files can be single- or multiple-page; line art, grayscale, or color. They can be up to 600 dpi, but 300 dpi is recommended for optimal OCR accuracy.

OmniPage Pro stores a 300 dpi line-art image or a 150 dpi grayscale or color image, depending on which is being viewed at the time.

Image files can be loaded at bit depths of 1, 8, or 24.

OmniPage Pro can save original images to these file-format types:

Bitmap (`*.bmp`)	TIFF Group 4, compressed (`*.tif`)
PCX (`*.pcx`)	TIFF Packbits (`*.tif`)
TIFF uncompressed (`*.tif`)	OmniPage document

Saving Image Files
OmniPage Pro saves each page of a multiple-page image separately.

If you select *Save all pages* in the Save Image dialog box, *Page#* is appended to file names to distinguish separately saved pages.

If you select *Save each graphic zone to a file*, then *Zone#* is appended to file names to distinguish separately saved graphic zones.

Images that are saved at low resolutions are not recommended for reloading for OCR.

OmniPage Pro can save recognized text to these file formats:

dBase III, III+, IV, 5.5 (*.dbf)	Microsoft PowerPoint (*.rtf)	Text only with line breaks (*.txt)
Excel 3.0, 4.0, 5.0, 6.0, 7.0, 97 (*.xls)	Microsoft Publisher 98 (*.rtf)	Ventura Publisher (MS Word) (*.doc)
FrameMaker 5.5.3 (*mif)	OmniPage Document (*.met)	Word for Windows 2.0, 6.0, and 7.0 (*.doc)
Freelance Graphics (*txt)	PageMaker 6.5.2 (MS Word) (*.doc)	Microsoft Word 95 and Word 97 (*.doc)
Harvard Graphics (*prn)	Quattro Pro for Windows 4.0, 8 (*.xls)	Wordpad (*.rtf)
HTML [†] (*.htm)	Rich Text Format (*.rtf)	WordPerfect for Windows 5.1, 5.2 (*.wp5), 6.0, 6.1, 95, 98 (*.wpd)
Lotus 1-2-3, 97 (*.wk1)	Text only	Word Pro 96, 97 (*.lwp)

[†]When saving to HTML, all graphics are saved as separate image files using JPEG format.

Scanner Setup Issues

This section contains information on setting up your scanner and solutions for scanning problems you may encounter.

▶

For more detailed scanner information, read the *Scanner Setup Notes* by clicking *Start* in the Windows taskbar and choose *Programs>Caere Applications>Caere Documents >Scanner Setup Notes.*

Topics in this section include:

- Scanner Drivers Supplied by the Manufacturer
- Scanner Drivers Supplied by Caere
- Problems Connecting OmniPage Pro to Your Scanner
- Missing Scan Image Command
- Scanner Message on Launch
- System Crash Occurs While Scanning

Scanner Drivers Supplied by the Manufacturer

Many scanners are shipped with one or more *scanner drivers*. This is software that allows your computer to communicate with your scanner. Some scanners do not require drivers and other scanners require more than one driver. Refer to your scanner documentation for information about installing any required scanner drivers.

Make sure that your scanner and scanner drivers are properly installed and configured before installing OmniPage Pro. Make sure that you have installed the appropriate scanner drivers supplied by the manufacturer.

▶

For HP IIp, IIc, IIcx, 3p, and 3c scanners, use the drivers that came with the scanners, or select a TWAIN driver in the Caere Scan Manager.

Scanner Drivers Supplied by Caere

OmniPage Pro is shipped with special scanner drivers that allow it to communicate with supported scanners. These scanner driver files are installed on your computer when you install the Caere Scan Manager. These drivers often work in conjunction with the drivers from your scanner manufacturer. To use your scanner with OmniPage Pro, you must select the appropriate scanner in the Caere Scan Manager.

Scan Manager is Needed with OmniPage Pro

To use your scanner with OmniPage Pro, you must install the Caere Scan Manager and select your scanner in it. The Scan Manager should have been installed during OmniPage Pro's installation.

To check if the Scan Manager is installed:

1   Click *Start* in the Windows taskbar and choose *Settings* ▸ *Control Panel* .

2   Look for the *Caere Scan Manager* icon.

The icon does not appear if the Scan Manager is not installed. Use the following procedure to install the Scan Manager if it has not been installed.

To install the Scan Manager:

1   Make sure your scanner is on before you start your computer.

2   Close OmniPage Pro if it is open.

3   Insert OmniPage Pro's CD-ROM.

4   Cancel the regular Setup program if it starts automatically.

5   Double-click the setup.exe program in the *Scanmgr* folder.

6   Select your scanner when you are prompted and follow the instructions on the screen.

Once your scanner is set up with OmniPage Pro, you can select scanner settings in OmniPage Pro's Options dialog box. See "Scanner Settings" on page 46 for more information.

Read the *Scanner Setup Notes* for the most detailed information about scanner support and setup. You can open this document after OmniPage Pro installation by clicking *Start* in the Windows taskbar and choosing *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *Scanner Setup Notes* .

Problems Connecting OmniPage Pro to Your Scanner

Try these solutions if you experience a problem between OmniPage Pro and your scanner or if you receive a scanner error message when you launch OmniPage Pro.

- Make sure the scanner is supported by OmniPage Pro with your version of Windows 95 or 98, or Windows NT.

  A list of tested scanners is provided in the *Scanner Setup Notes*. *Scanner Setup Notes* can be accessed by clicking *Start* in the Windows taskbar and choosing *Programs* ▸ *Caere Applications* ▸ *Caere Documents* ▸ *Scanner Setup Notes.* If your scanner is not listed, call your scanner manufacturer to find out if it is supported, or visit **www.caere.com**.

- Make sure the Caere Scan Manager is installed and that you have selected the correct scanner in the Scan Manager.

  See "Scan Manager is Needed with OmniPage Pro" on page 92.

- Make sure you have installed the appropriate scanner driver. See the *Scanner Setup Notes* for more information.

- Make sure your scanner is connected, turned on, compatible with your system, and runs with the software provided by the manufacturer *before* you use it with OmniPage Pro.

- Scanner drivers must be loaded at startup. Turn on your scanner first and then restart your computer.

- Make sure the scanner is not in use by another application.

- Uninstall and then reinstall the Caere Scan Manager. Refer to "Scan Manager is Needed with OmniPage Pro" on page 92

Missing *Scan Image* Command

The *Scan Image* command does not appear in the Image button's drop-down list in the following cases:

- You did not install the Caere Scan Manager or select an appropriate scanner. See "Scan Manager is Needed with OmniPage Pro" on page 92 for instructions.

- Your scanner is not connected to your computer or is not functioning properly. See "Scanner Setup Issues" on page 91.

- You use a Visioneer scanner or your scanner is set up to work with Visioneer's PaperPort software such as the HP ScanJet 5s. See the online *Scanner Setup Notes* for more information.

Scanner Message on Launch

The first time you launch OmniPage Pro after installing or changing your current scanner in the Caere Scan Manager, you may get this message: *This scanner's configuration is set using the system-level driver.* If it asks for no more information, click *OK* in the dialog box. You may also have the option to select the following:

- SCSI ID or scanner configuration information
  Consult your scanner documentation for the correct information.

- Page-size information
  Enter the largest size page that your scanner supports.

System Crash Occurs While Scanning

Try these solutions if a crash occurs during a scan:

- Turn your computer off. Power your scanner off and on again to return the scanner to its default state. Then restart your computer.

- Check your scanner setup. See "Scanner Setup Issues" on page 91 for more information.

- Check the *Scanner Settings* tab in the Caere Scan Manager if you are using a TWAIN scanner. (Click on your scanner icon in Scan Manager icon under Control Panel).

- Check with the scanner manufacturer to make sure you have the appropriate driver for your scanner.

- Resolve low memory problems. See "Low Memory Problems" on page 88 for more information.

- Resolve low disk space problems. See "Low Disk Space Problems" on page 88 for more information.

- Visit Caere Corporation's web site at **www.caere.com** for Scan Manager updates.

Scanner Not Listed in Supported Scanners List Box

Try these solutions if your scanner is not listed in the Scan Manager *Scanner* list:

- Check Caere Corporation's web site at **www.caere.com** for Scan Manager updates.
- Select *TWAIN Scanner* as your current scanner in the *Scanner* list.

Scanning Tips

OCR results will be poor if an image is not scanned properly. Remember the following tips when you scan:

- Scan documents at 300dpi.
- Take the color and quality of your document into account when scanning.
  High-quality documents return better recognition results than low-quality documents. Shaded, colored, or low-quality documents may result in poor recognition accuracy unless adjustments are made before scanning. See "What is the quality of the original document?" on page 55 for more information.
- Always try to scan an original document instead of a photocopy.
- If you are going to use FAX copies to OCR, ask your FAX sender to send them to you using their machine's Best or Fine mode.
- Make sure the page is properly aligned in the scanner.
  Select *Automatically straighten page image* in the *Accuracy* settings of the Options dialog box to automatically straighten a page image by up to 10 degrees if necessary.
- Check the glass, mirrors, and lenses on your scanner for dust, smudges, or scratches. Clean if necessary.
- Make sure the proper settings are selected in the *Scanner* tab of the Options dialog box before scanning.
  See "Scanner Settings" on page 46 for more information.

OCR Problems

This section contains information and solutions for possible OCR problems.

Topics in this section include:

- System Crash During OCR
- Text Does Not Get Recognized Properly
- Problems With Fax Recognition

System Crash During OCR

Try these solutions if a crash occurs during OCR or if processing takes a very long time:

- Resolve low memory problems. See "Low Memory Problems" on page 88 for more information.
- Resolve low disk space problems. See "Low Disk Space Problems" on page 88 for more information.
- Minimize all applications or click Alt+Tab to check for Windows error messages.
- Check the quality of the image you are recognizing. See "What is the quality of the original document?" on page 55 for more information.

  See "Scanning Tips" in the previous section for ways to improve the quality of scanned images.
- Break complex page images (lots of text and graphics or elaborate formatting) into smaller jobs. Draw zones manually or modify automatically created zones and perform OCR on one page area at a time. See "Customizing Zones" on page 63 for more information.
- Restart Windows 95 or 98 in safe mode or Windows NT in VGA mode and test OmniPage Pro by performing OCR on the included `Sample.tif`. See "Testing OmniPage Pro" on page 87.
- If you are performing multiple tasks at once, such as recognizing and printing, OCR may take longer.

Text Does Not Get Recognized Properly

Try these solutions if any part of the original document is not converted to text properly during OCR:

- Look at the original page image and make sure that all text areas are enclosed by text zones. If an area is not enclosed by a zone, it is ignored during OCR. See "Creating Zones for OCR" on page 22 for more information.

- Make sure text zones are identified correctly. Reidentify zone types and contents, if necessary, and perform OCR on the document again. See "Changing Zone Properties" on page 71 for more information.

- Adjust the Brightness slider in the *Scanner* settings of the Options dialog box. Lighten the setting for thick, run-together text characters or dark backgrounds. Darken the setting for thin, broken text characters.

- Make sure the correct main and secondary document languages are selected in the *Language* settings. Only languages included in the document should be selected. See "Language Settings" on page 48 for more information.

- Select *Use Language Analyst* in the *Accuracy*. The Language Analyst evaluates words and corrects likely errors during OCR. See "Accuracy Settings" on page 46 for more information.

- Train OmniPage Pro to recognize special characters that might normally be difficult to recognize, such as the copyright symbol © or the registered trademark symbol ®. Do not train OmniPage Pro for regular characters because it may interfere with recognition. See "Training OCR for Special Characters" on page 75 for more information.

- If you use *True Page* as the *Output Format* setting, recognized text gets put into frames (formatting boxes) in the text viewer. Some text may be hidden from view if a frame is too small. To view the text, place the cursor in the text frame and use the arrow keys on your keyboard to scroll to the top, bottom, left, or right of the frame.

- Check the glass, mirrors, and lenses on your scanner for dust, smudges, or scratches. Clean if necessary.

OmniPage Pro only recognizes machine printed-text characters such as typewritten or laser-printed text. However, it can retain handwritten text, such as a signature, as a graphic. See "Do you want to retain graphics in your document?" on page 58 for guidelines.

Problems With Fax Recognition

Try these solutions to improve OCR accuracy on fax images:

- Ask senders to select *Fine* or *Best* mode when they send you a fax. This produces a resolution of 200x200 dpi.

- Ask senders to transmit files directly to your computer via fax modem if you both have one. You can save fax images as image files and then load them into OmniPage Pro. See "Supported File-Format Types" on page 89 for more information.

- Ask senders to use clean, original documents if possible. Sans serif fonts (such as the one used for headings in this manual) are easier to recognize than serif fonts (such as the one used for body text in this manual).

Uninstalling the Software

Sometimes uninstalling and then reinstalling OmniPage Pro and the Caere Scan Manager will solve a problem.

OmniPage Pro's Uninstall program will *not remove* any files saved to the OmniPage installation folder or subdirectories, including the following files:

- Zone templates (`*.zon`)
- Training files (`*.trn`)
- User dictionaries (`*.ud`)
- Temp files (`*.tmp`)

To uninstall from Windows NT, you must be logged into your computer with administrator privileges.

To uninstall OmniPage Pro:

1 Close OmniPage Pro.

2 Click *Start* in the Windows taskbar and choose *Settings* ▸ *Control Panel* ▸ *Add/Remove Programs* .

3 Select OmniPage Pro and click *Add/Remove*.

4 Click *OK* to confirm that you want to remove OmniPage Pro.

5 Restart your computer.
Some icons and program files may remain on your system if they have been renamed, modified, or moved.

6 Restart your computer.

To uninstall the Caere Scan Manager:

1 Close OmniPage Pro.

2 Click *Start* in the Windows taskbar and choose *Settings* ▸ *Control Panel* ▸ *Add/Remove Programs* .

3 Select *Scan Manager* and click *Add/Remove*.

4 Click *OK* to confirm that you want to remove the Scan Manager.

5 Restart your computer.
Some icons and program files may remain on your system if they have been renamed, modified, or moved to different locations.