

COPYRIGHT INFORMATION

Copyright © 1998 by Scansoft, Inc., a Xerox Company. All rights reserved. No part of this publication may be transmitted, transcribed, reproduced, stored in any retrieval system or translated into any language or computer language in any form or by any means, mechanical, electronic, magnetic, optical, chemical, manual, or otherwise, without the prior written consent of Scansoft, Inc., 9 Centennial Drive, Peabody, Massachusetts 01960. Printed in the United States of America.

The software described in this book is furnished under license and may be used or copied only in accordance with the terms of such license.

IMPORTANT NOTICE

Scansoft, Inc. provides this publication “as is” without warranty of any kind, either express or implied, including but not limited to the implied warranties of merchantability or fitness for a particular purpose. Some states or jurisdictions do not allow disclaimer of express or implied warranties in certain transactions; therefore, this statement may not apply to you. Scansoft reserves the right to revise this publication and to make changes from time to time in the content hereof without obligation of Scansoft to notify any person of such revision or changes.

TRADEMARKS AND CREDITS

TextBridge is a registered trademark, and *Smart Zones* and *Instant Access OCR* are trademarks, of Scansoft, Inc., a Xerox Company. *Xerox*, *The Document Company*, and the *Stylized X* are trademarks of Xerox Corp.

Excel, *Word*, and *Windows* are trademarks of Microsoft Corp.

WordPerfect is a registered trademark of WordPerfect Corp.

Other terms used in this manual are the trademarks of their respective holders.

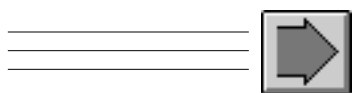
Portions of this product copyright © 1990–1997, Pixel Translations, Inc.

Portions of this product copyright © 1994–1997, Mastersoft Corp.

Designed, written, and illustrated by Lois West

© SCANSOFT, INC.
9 Centennial Drive
Peabody, Massachusetts 01960

TextBridge Classic User's Guide
Part Number 00–09066–00
April 1998



CONTENTS

PREFACE

About This User's Guide	v
Organization of this user's guide	vi
Documentation conventions	vii
Related Documentation	vii
Technical Support	viii

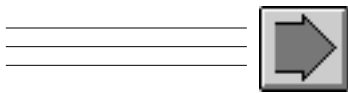
1 INTRODUCTION TO TEXTBRIDGE CLASSIC

Features and Benefits	1-2
Productivity features	1-2
Other TextBridge features	1-3
Characteristics of Documents TextBridge can recognize ..	1-4
System Requirements	1-6
Installing TextBridge, Scanner Drivers, and Languages	1-7
Setting Up TextBridge Instant Access	1-8
Uninstalling TextBridge	1-9
Input Image File Formats Supported	1-10
Output Text File Formats Supported	1-11
Where to Go From Here	1-12

2 OCR AND TEXTBRIDGE CLASSIC

What is TextBridge OCR?	2-2
Page types	2-2
Page sources	2-4
Running TextBridge	2-5
Standalone Application	2-5
Instant Access	2-5
TextBridge Functionality	2-6
Before You Start to OCR	2-6
Using TextBridge to OCR	2-7

Automatic Processing	2-7
Manual Processing	2-11
Selecting Page Type and Source	2-12
Previewing the Page	2-13
Zoning the Page	2-15
Recognizing the Page	2-17
Saving the Document	2-17
Improving Page Recognition with Settings	2-18
Page Type Settings	2-18
Scanner Settings	2-21
Processing Settings	2-22
Where to Go From Here	2-23
 3 LEARNING TO USE TEXTBRIDGE CLASSIC	
Starting TextBridge	3-2
Using the Help System	3-4
Using the Sample Documents	3-6
Session 1: Processing a Simple Document Using Auto	
Processing	3-11
Session 2: Using Instant Access OCR	3-17
Session 3: Processing a Complex Document Using Manual	
Processing	3-23
Where to Go From Here	3-29



PREFACE

ScanSoft, Inc., a Xerox Company, welcomes you to **TextBridge® Classic 2.0** for Windows 95™ and Windows NT. (Hereinafter TextBridge Classic 2.0 will be referred to as “TextBridge.”)

Before going on to find out more about TextBridge, please read this preface because it describes these important items:

- u About this user’s guide
- u Related documentation
- u Technical support

ABOUT THIS USER’S GUIDE

This user’s guide includes introductory information designed primarily for non-technical users as well as information designed for more technical users. It assumes that you are familiar with the management and operation of your computer and Windows.

The documentation that comes with TextBridge should provide all the information you need to operate TextBridge. TextBridge documentation includes this user’s guide, a Help system, and Release Notes. ScanSoft invites your comments about the information provided in the documentation. Please make sure to register your software and provide any comments to ScanSoft.

Organization of this user's guide

This user's guide is designed as a reference tool to provide basic information about TextBridge. It is organized as follows:

- u Chapter 1, "Introduction to TextBridge," discusses TextBridge's features. It also describes: documents TextBridge can recognize, what comes with TextBridge, system requirements, installation, setting up Instant Access, uninstalling TextBridge, and input and output file formats.
- u Chapter 2, "OCR and TextBridge," provides an explanation of the concepts of document recognition and OCR and the basic functionality of TextBridge.
- u Chapter 3, "Learning to Use TextBridge," walks you through several practice sessions designed to provide a firm basis on which to learn and use the important features of TextBridge.

This user's guide also provides a comprehensive index for you to quickly locate the information you need.

Documentation conventions

As described in Table P-1, TextBridge documentation uses certain graphical elements and formatting to emphasize information and give more meaning to text.

Table P-1. *Documentation Conventions*

bold	Introduces a new term or the first use of an important term in a chapter. Sometimes used to denote strong in-line emphasis.
<i>italic</i>	Denotes titles of other user's guides or books and generic representations of file name entries in examples; for example, <i>filename</i>

<code>monospace</code>	Denotes text that appears on the computer screen such as examples, menu text, and messages plus actual file names.
“ ” (quotes)	Denotes titles of chapters and sections in this user's guide.
+	Introduces tips that provide useful information about a procedural step or system function.
Note	Introduces information of note about the current subject.

RELATED DOCUMENTATION

TextBridge provides a comprehensive set of printed and online documentation designed to assist you in learning and operating the product. The documentation provided with TextBridge covers all aspects of installation and operation.

In addition to this *TextBridge Classic User's Guide*, refer to the following documentation for more information:

- u **Online Release Notes**—After you install TextBridge, **read the online Release Notes first**. These provide the most up-to-date information about TextBridge. Release Notes automatically appear in the TextBridge Classic folder and TextBridge program group in the Start menu. Simply point to Release Notes in the TextBridge Classic folder to open the Release Notes so that you can read them.
- u **Help**—An extensive online Help system comes with TextBridge. The Help provides you with information about the software in general; the menus, commands, and tools; step-by-step procedures; and a glossary.

- u **TextBridge online user's guide**—This includes an electronic version of this *TextBridge Classic User's Guide* in Adobe Acrobat format (.pdf). The documentation resides on the compact disk in the directory TB_Doc as well as in the TextBridge Classic folder and TextBridge program group in the Start menu. Please refer to the readme in that directory for information about using the online documentation. If the online user's guide is not available in either of these places, you may be able to get it from the TextBridge Classic Web site.u **TextBridge Classic Web sites**—General information about **TextBridge Classic** is available at www.xerox.com/scansoft/tbclassic. Commonly asked questions with their answers are available at www.xerox.com/scansoft/tbclass/ic/tbclassicqanda.htm.

Note You may need to refer to additional publications, such as the manufacturer's documentation for your scanner.

TECHNICAL SUPPORT

If you should experience problems with TextBridge that you cannot resolve with the documentation and software, contact the manufacturer of the scanner or multi-function device with which you received TextBridge Classic.

This information will assist Technical Support in solving the problem:

- u Your software version number
(This is on the back of the CD-ROM case and in the Help menu under About TextBridge.)
- u Your scanner make and model
- u A description of the steps that led up to the problem
- u If TextBridge generated an error message, a verbatim description of the error message or its number

Internet and electronic mail addresses

You can get information about TextBridge on the Internet at the addresses in the following list:

- u TextBridge site: www.textbridge.com

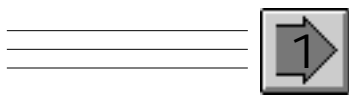
The TextBridge Web site provides a link to Technical Support with Frequently Asked Questions and technical information bulletins.

E-mail in the United States, Canada, or the Pacific Rim:

- u Upgrade information: textbridge_sales@xis.xerox.com

E-mail from European countries and the Middle East:

- u Upgrade information: xisuk@xis.xerox.com



INTRODUCTION TO TEXTBRIDGE CLASSIC

Welcome to ScanSoft's **TextBridge™ Classic 2.0**, optical character recognition (OCR) software for Microsoft Windows™ 95 and Windows NT. (Hereinafter TextBridge Classic 2.0 will be referred to as "TextBridge.")

This chapter provides an introduction to TextBridge including:

- u Features and benefits
- u What comes with TextBridge
- u System requirements
- u Installing TextBridge
- u Setting up TextBridge Instant Access
- u Uninstalling TextBridge
- u Input image file formats supported
- u Output text file formats supported

OCR is a technology that enables you to change the paper documents you use every day into fully-editable text on your computer.

You can use TextBridge to convert **printed documents** from fax machines, photocopiers, and dot matrix and laser printers to **electronic documents** for your word processor or text application as well as documents for some database and spreadsheet software. TextBridge OCR can also recognize **page image files** from scanners as well as fax machines and other sources.

FEATURES AND BENEFITS

Using Xerox's latest document recognition technology, DocuRT™, TextBridge OCR produces a fully-editable electronic document

TextBridge offers many productivity features. Whether you need to capture a simple one-page letter, a magazine article, a spreadsheet, or a long transcript, TextBridge can save you valuable time and effort.

Productivity features

TextBridge offers these major features:

- u **Improved OCR accuracy.** Dramatically saves time and eliminates retyping.
- u **Instant Access™.** You can start TextBridge within most Windows text programs. After recognizing and converting the page image to text, TextBridge then automatically pastes recognized data (text and pictures) directly into the text program's open document.

- u **Page type templates.** TextBridge provides many predesigned page type templates to make processing more efficient. Templates automatically provide appropriate settings for the type of page you want to process. For example, there is a magazine page type and a letter page type that automatically activate settings for improved results. Page types incorporate three page settings: page size, page layout, and print quality. You do not have to go through a complicated process of determining and specifying settings for common types of pages.
- u **Automatic zoning.** TextBridge automatically zones your page into text zones and reads them in the correct order.
- u **Zone editing.** You can manually create and edit zones to refine the zoning. Use zone editing to increase the accuracy and efficiency of page processing by reshaping zones.
- u **ToolTips and What's This? Help.** Instant context-sensitive information about commands, dialog boxes, and buttons on the interface.

Other TextBridge features

In addition to the features listed in the previous section, TextBridge provides these other productivity features:

- u **Windows 95** certification
- u **MMX** support
- u **Image processing.** TextBridge accepts a wide range of images from a variety of sources for processing. Specifically, the program imports and recognizes online document images in Alacrité ALA, AWD, BMP, Delrina, PCX, DCX, TIFF, and XIF formats originating from fax modems and other sources. For more information, see the "Supported Input Image File Formats" section in this chapter.

- u **Output text formats including HTML.** TextBridge supports a number of output text formats, including word processor, desktop publishing, portable document, spreadsheet, HTML, and database formats. Now you can process your text for publication on the Web. For more information, see the “Supported Output Text File Formats” section in this chapter.
- u **Preview with manual zoning.** TextBridge provides a set of tools for previewing page images before processing them. You can view a page before continuing with processing. You can manually define areas of page images as zones to be processed and capture only the text you want. You can also edit the automatic zoning by adjusting the text zones.
- u **Two-sided document processing.** If your scanner has a sheet feeder, you can scan the fronts (odd sides) of the pages first, then flip the stack and scan the reverse (even) sides. When scanning and recognition are complete, TextBridge automatically collates the text.

With these features, you can import virtually any paper document or document image file to your computer. TextBridge attains the highest degree of OCR accuracy and provides the output in fully editable form in your favorite text program.

Characteristics of documents TextBridge can recognize

TextBridge includes a number of advances developed by the Xerox Desktop Document Systems (DDS) division and by the Palo Alto Research Center (PARC), where modern computer interfaces were invented.

Consequently, TextBridge provides OCR on a wide range of documents. TextBridge can recognize:

- u Documents printed on typewriters, phototypesetters, and impact, ink-jet, dot-matrix, and laser printers
- u Photocopied, degraded, or dirty documents

- u Documents with single- or multiple-column layouts
- u Paper documents with black and white, grayscale, or color pictures including photos and line art
- u Page image files with black and white pictures
- u Online single- or multiple-page images from fax modems and other sources
- u Hard-copy faxes
- u Documents with point sizes ranging from 6-point to 72-point type in practically any typeface
- u TextBridge software with English, French, German, Italian, Brazilian Portugues, or Spanish user interface.
- u Documents composed in the languages in the following list:

Afrikaans	Albanian	Aymara
Baltic Rim	Basque	Breton
Bulgarian	Byelorussian	Catalan
Central European	Croatian	Czech
Cyrillic	Danish	Dutch
English	Estonian	Faroese
Finnish	Flemish	French
Friulian	Gaelic	Galician
German	Greek	Greenlandic
Hawaiian	Hungarian	Icelandic
Indonesian	Italian	Kurdish (Latin)
Latin	Latvian	Lithuanian

Lower Sorbian	Macedonian (Cyrillic)	Malaysian
Norwegian	Pidgin English	Polish
Portuguese	Romanian	Russian
Serbian	Serbo-Croatian	Slovak
Slovenian	Spanish	Swahili
Swedish	Tahitian	Turkish
Ukranian	Upper Sorbian	Welsh
Western European	West Frisian	Zulu

SYSTEM REQUIREMENTS

To install and run TextBridge, your Windows-compatible PC must be equipped with the following:

- u An Intel (or compatible) 80486 or Pentium™ microprocessor
- u VGA, SVGA, or multi-sync color monitor
- u A minimum of 16 megabytes of random access memory (RAM) for Windows 95 and Windows NT
- u Microsoft Windows™ 95 or Windows NT 4.0
- u A hard disk with a minimum of 20 megabytes (20 MB) of free space in which to install TextBridge. This enables installation of all TextBridge software and one language pack. Please allow one megabyte (1 MB) for each additional language pack you intend to install.

INSTALLING TEXTBRIDGE, SCANNER DRIVERS, AND LANGUAGES

To install TextBridge, follow the instructions provided with your scanner or multi-function device.

The information in this section provides additional explanations about installing TextBridge, including scanner drivers and language installation.

Note This scanner driver information may not be applicable to your scanner or multi-function device.

During installation, both ISIS and TWAIN drivers are installed. You can configure TWAIN sources during installation. You can configure ISIS drivers using the Select Scanner command in the File menu, which is available after TextBridge is installed. This command displays the Scanner Setup dialog box. This dialog box groups the scanners in three categories: All, TWAIN, and ISIS. TWAIN is the default.

TextBridge installation provides a Typical and Custom installation.

Typical installs the language of your interface.

Custom allows you to select language packs to add or remove.

For example, if you have the French version of TextBridge Classic, only French will be installed, unless you choose Custom. When you select Custom, the Language Selection dialog box appears, and you can select language packs to add or remove in this dialog box. If TextBridge is already installed and you want to add or remove languages, run the setup program and use the CD that contains the TextBridge program, then select custom installation.

SETTING UP TEXTBRIDGE INSTANT ACCESS

When you restart your PC, you can use the TextBridge Instant Access Control Panel dialog box to set up Instant Access (Figure 1–11). To set up TextBridge Instant Access from your other programs, use the following procedure:

1. **On the Windows task bar, click Start.**
2. **Point to Programs, then point to the TextBridge Classic folder.**
3. **Click Instant Access Control Panel.**

The TextBridge Instant Access Control Panel dialog box appears.

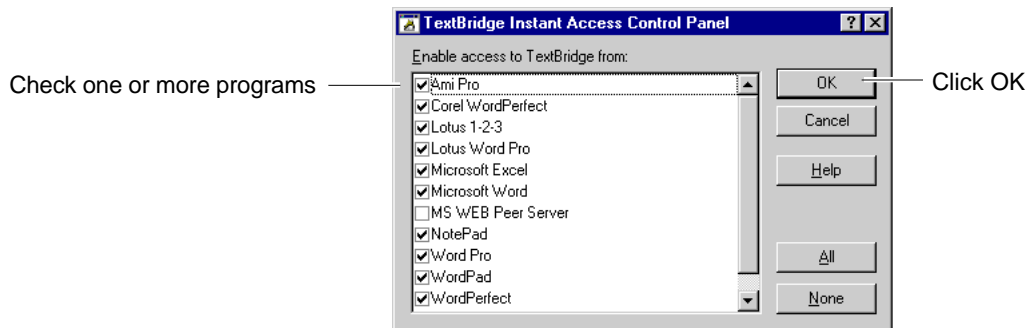


Figure 1–1. *TextBridge Instant Access Control Panel*

4. **Check one or more programs in the list.**
5. **Click OK.**

More information is available if you click the Help button in the dialog box.

TextBridge will now be available in the File menu of the program(s) you checked if they are installed on your PC.

UNINSTALLING TEXTBRIDGE

To restore your PC to the state it was in before you installed TextBridge, use the following procedure:

- 1. Close all active applications, including TextBridge.**
- 2. On the Windows task bar, click Start.**
- 3. Point to Programs, then point to the TextBridge Classic folder.**
- 4. Click TextBridge Uninstall.**

The TextBridge Uninstall dialog box appears.

- 5. Click Yes to continue the uninstall process.**

TextBridge automatically uninstalls.

Click No to exit the uninstall process.

- 6. The Uninstall Complete dialog box appears. Click OK to restart your computer.**

With the above steps completed, TextBridge is completely uninstalled from your PC.

INPUT IMAGE FILE FORMATS SUPPORTED

The source of page images for TextBridge can be your scanner or it can be image files. TextBridge can recognize the following types of image file formats:

Image File Format	File Name Extension
Delrina WinFax fax image files	.fxr, .fxd, fxm, .fxs
Extended image file	.xif
Microsoft fax image files	.awd
Multi-page PCX used in some fax programs Tag image file format (including Alacrity TIFF)	.ala, .dcx, .tif,
PCX	.pcx
Tag image file format (TIFF, including Alacrity)	.ala,.tif
Windows bitmap	.bmp

All the previous image files **must be black and white** with the exception of .xif, which can contain color or grayscale images. TextBridge can process images in resolutions from 72 to 900 dots per inch. However, you will not receive noticeably better OCR on images with resolutions higher than 400 dpi. In addition, you may encounter memory errors or at least slower processing time. It is recommended that you scan at 300 or 400 dpi.

Note	This list is subject to change. Refer to the online Release Notes for the latest information.
-------------	---

OUTPUT TEXT FILE FORMATS SUPPORTED

TextBridge can convert its recognized text to files for the following programs:

Program	File Name Extension
Ami Pro 2.0 and 3.0	.sam
ASCII Smart, Standard, and Stripped	.txt
dBase IV	.dbf
DCA/RFT	.rft
DisplayWrite 5	.rft
Excel 3.0, 4.0, and 5.0	.xls
Excel for the Macintosh 3.0 to 7.0	.xls
Excel 97	.xls
FrameMaker	.mif
HTML	.htm
HTML Editor	.htm
HTML Netscape	.htm
Interleaf	.wps
Lotus 1-2-3	.wk1
Lotus Word Pro	.lwp
MSWorks	.rtf
MultiMate Advantage	.doc
PostScript	.ps
Professional Write 2.0 and 2.2	.doc
Quatro Pro for Windows	.wb2
Rich Text Format	.rtf
RTF for the Macintosh	.rtf
Unicode text	.txt
Windows Write	.wri
Word for Windows 2.x	.rtf
Word 6.0 and 7.0	.rtf

Program	File Name Extension
Word 97	.rtf
WordPerfect 4.2 and 5.1	.wpf
Word Perfect 6.0, 6.1, 7.0 and 8.0	.wpd
WordStar	.doc

Note This list is subject to change. Refer to the Release Notes for the latest information.

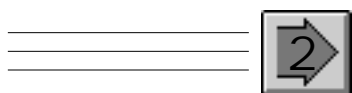
WHERE TO GO FROM HERE



To learn how TextBridge recognizes a document and how you prepare TextBridge to do this, read Chapter 2. This chapter explains the basic concepts and functions of the software.

To learn how you use TextBridge to process simple and complex documents, refer to Chapter 3. It also explains how to view, zone, and edit your document in your word processor.

The online Help system provides a complete reference to the user interface, including window areas, menus, commands, and tools as well as overview information and features, step-by-step procedures for using the software, tips, and a glossary.



OCR AND TEXTBRIDGE CLASSIC

This chapter provides information about the process of **page recognition**. Use this chapter to learn about **optical character recognition (OCR)** and other concepts that will help you use TextBridge effectively.

This chapter provides information about OCR and TextBridge including:

- u What is TextBridge OCR?
- u Running TextBridge
- u TextBridge functionality
- u Before you start to OCR
- u Using TextBridge to OCR
- u Automatic processing
- u Manual processing
- u Improving Page Recognition with Settings

Page recognition refers to the process during which a page is analyzed, and characters and words are saved as a text. **Optical character recognition** is the technology that converts documents that you can read into documents that your computer can read.

On the Scanner tab:

- u Choose the resolution value to reflect the actual resolution of your scanner. For most documents, use 300 dpi. For 8 point or smaller text, use 400 dpi for the best results.
- u Change the brightness based on whether your original page has light or dark text and pictures. For example, if the text and pictures on the original page are light, darken the brightness control.
- u Check the box to use the Automatic Document Feeder if your scanner has this feature and you are scanning multiple pages.

Processing Settings

You can view and change the settings for processing in the Processing tab of the Settings dialog box. Check the settings to be sure they are the best ones for processing the original page.

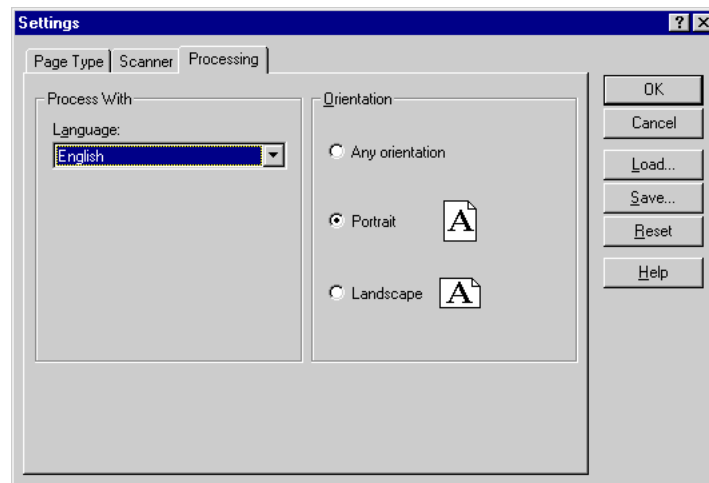


Figure 2-15. Settings dialog box and Processing tab

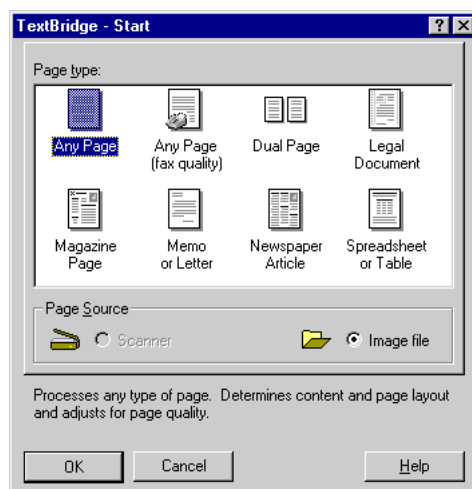


Figure 2-1. Page types in Start dialog box

Page types incorporate three settings: page size, print type, and page layout. The page types to choose from and their **default settings** are described in the following table:

Page Type	Page Size	Print Type	Page Layout
Any Page	Letter or A4	Any	Any
Any Page (Fax quality)	Letter or A4	Fax	Any
Dual Page	Tabloid Dual or A3 Dual	Any	Any
Legal Document	Legal	Good	Single column
Magazine Page	Letter or A4	Good	Multi-column
Memo or Letter	Letter or A4	Good	Single column
Newspaper	Legal	Newspaper	Multi-column

Page Type	Page Size	Print Type	Page Layout
Spreadsheet or Table	Letter or A4	Good	Table and one-column text

In addition to the settings connected with page types, there are several other settings that control page processing. You can modify these settings by using the Settings dialog box.

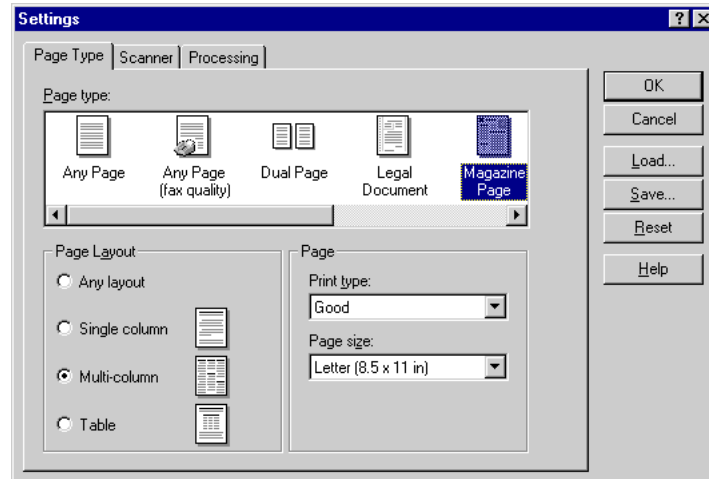


Figure 2-2. Page Type tab in Settings dialog box

Page sources

You can get pages to process from your scanner or from page images. Use your scanner as a source to input documents on paper to TextBridge, which then takes the scanned images, performs OCR, converts the recognized text to the text file format of your choice, and stores it on your PC. Alternatively, use TextBridge to recognize and convert page images stored in image files that come from fax modems or other sources. Refer to Figure 2-1, which illustrates where to select your page source.

RUNNING TEXTBRIDGE

You can run TextBridge as standalone application or invoke it from within an application with Instant Access.

Standalone Application

When you start TextBridge from the Start menu, it operates as standalone application and runs independently of any other application. TextBridge recognizes pages and saves them in the output format that you specify. You can then open the file in the application that uses the format you specified.

Instant Access

Instant Access gives you direct access to TextBridge from applications such as Microsoft Word. Programs with Instant Access have a TextBridge command in the File menu. Clicking TextBridge in the File menu starts TextBridge, which recognizes pages and pastes them directly into the open document in the program. You can also add the TextBridge Instant Access button to your application's toolbar using the Toolbars command in the View menu.

You operate TextBridge just as if you had started the standalone TextBridge software. The main differences between running TextBridge standalone compared to Instant Access are:

- u The options to process automatically or manually are available in the Start dialog box.
- u TextBridge automatically determines which format to use based on the application being used.

The Instant Access Control Panel, which is available from the Start menu, enables you to specify which programs have Instant Access to TextBridge. The programs in Figure 2-3 automatically have Instant Access.

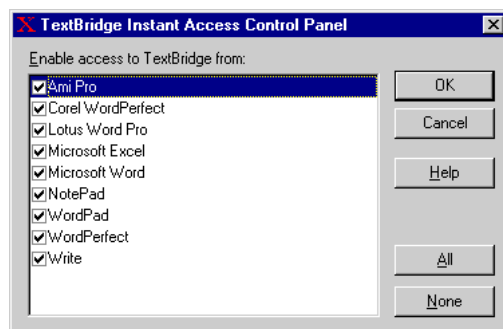


Figure 2-3. *Instant Access Control Panel*

TEXTBRIDGE FUNCTIONALITY

You can perform the activities in the following list with TextBridge:

- u Select a page type for the type of page to process.
- u Select page type settings for the entire document or change settings on a page-by-page basis.
- u Zone the page automatically or manually.
- u Select a portion of a page to process.
- u Save the recognition results in one of many file formats.

BEFORE YOU START TO OCR

The following checklist will take you through the most important questions to ask before you start to process a document.

1. Is my document coming from my scanner or an image file?
2. What type of page is the document?

3. Is this document a good candidate for OCR?
4. Are there any other settings I want to check and change?

The rest of this chapter provides information that helps you to answer these questions.

USING TEXTBRIDGE TO OCR

The next two sections provide information on automatic OCR processing as well as other more advanced processes that are options in manual processing. Refer to the Help system for the step-by-step procedures for these activities.

TextBridge provides an easy-to-use interface and a powerful set of built-in capabilities. You can use it in a number of ways to do OCR, depending on the complexity of the document to be recognized. You can use TextBridge to OCR in automatic mode or manual mode.

- u You can process all the pages automatically in the automatic mode or interact with the process in the manual mode.
- u You can preview and zone pages before OCR.

AUTOMATIC PROCESSING

When you use TextBridge's **automatic processing**, TextBridge processes pages automatically with very little interaction with you. In **automatic mode**, once you select the page type, TextBridge automatically recognizes your page(s). TextBridge only stops for you to save or add more pages.

The following steps describe the automatic process of using TextBridge for page recognition. Refer to the Help system for the step-by-step procedures for these activities.

- 1. Click the Auto Process button.**
- 2. Select a page type.**
- 3. Select the source of the page image, either your scanner or image file.**
- 4. Click OK.**
- 5. If you selected image file, select the image file(s) in the Open dialog box.**
- 6. TextBridge processes all the pages in your scanner or the selected image file(s).**
- 7. If scanning, click the More Pages button to add another page to the final document. (Optional)**
- 8. If scanning, click the No More button when there are no more pages to add.**
- 9. TextBridge recognizes the page(s).**
- 10. Save the text in a file format of your choice.**

During the automatic process, you will interact with these screens

Click the Auto Process
button

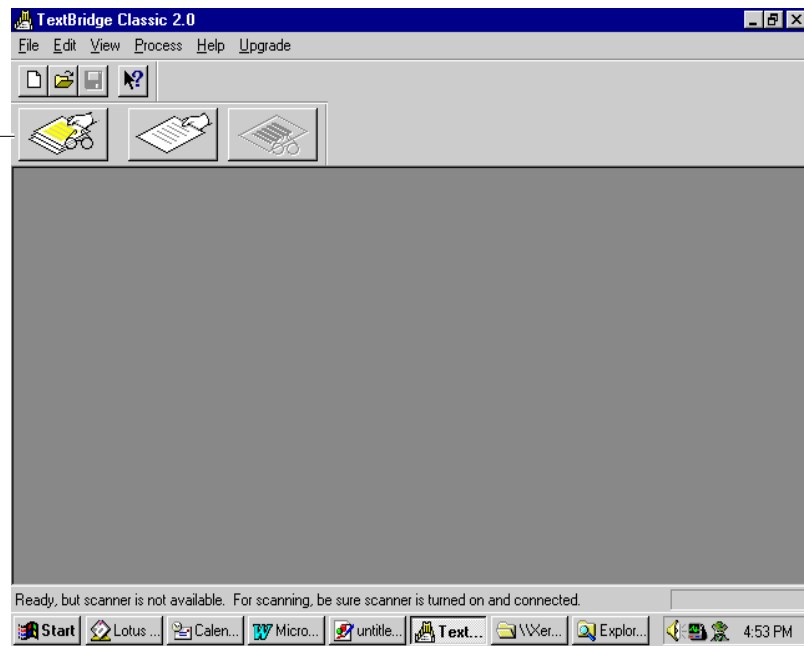


Figure 2-4. Click the Auto Process button in the TextBridge window

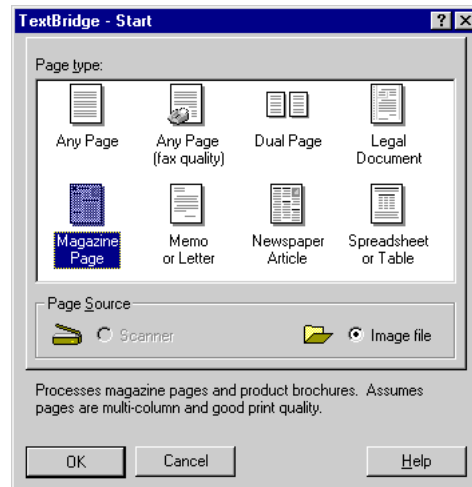


Figure 2-5. Start the automatic process using the Start dialog box

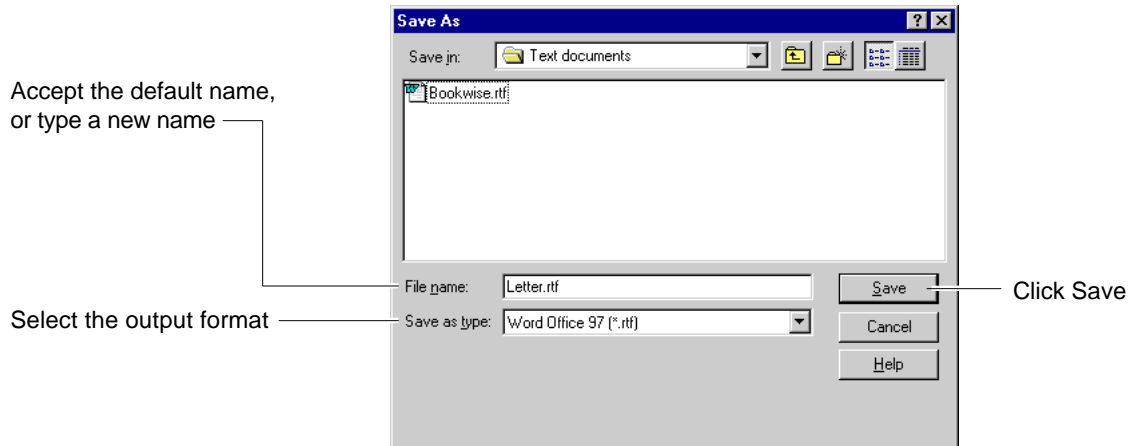


Figure 2-6. Save the document using the Save As dialog box

MANUAL PROCESSING

Page recognition is a complex process, and it can require your interaction with TextBridge to get the best output. There are a number of opportunities during the page recognition process that allow you to enhance the results for the particular document. During **manual mode**, you lead TextBridge through processing a document.

During **manual processing** you can stop the page recognition process to perform the activities in the following list:

- u Preview the page
- u Zone the page manually

The following steps describe the manual process of using TextBridge for page recognition. Refer to the Help system for the step-by-step procedures for these activities.

- 1. Click the Get Page button.**
- 2. Select a page type.**
- 3. Select the source of the page image, either your scanner or an image file.**
- 4. Click OK.**
- 5. If using an image file, select the file.**
- 6. TextBridge processes the first page image.**
- 7. Preview the page, including manual zoning, if desired.**
- 8. Click the Recognize Page button.**
- 9. TextBridge recognizes the page and zones it automatically if you have not zoned it manually.**

10. If scanning, add more pages to the document. (Optional)

11. Save the text in a file format of your choice.

Each of these activities is explained in the next sections.

Selecting Page Type and Source

When you start processing a new document, the TextBridge — Start dialog box appears, and you can perform the actions in the following list:

- u Indicate whether pages are from your scanner or an image file.
- u Select the Page type that best matches your original page(s).
- u View and change the settings for the page type you selected.



Figure 2-7. Start the manual process using the Start dialog box

Note You can use the optional Page toolbar to select the page type and source rather than the Start dialog box. From the Toolbars dialog box in the View menu, choose the Page toolbar.

Page toolbar

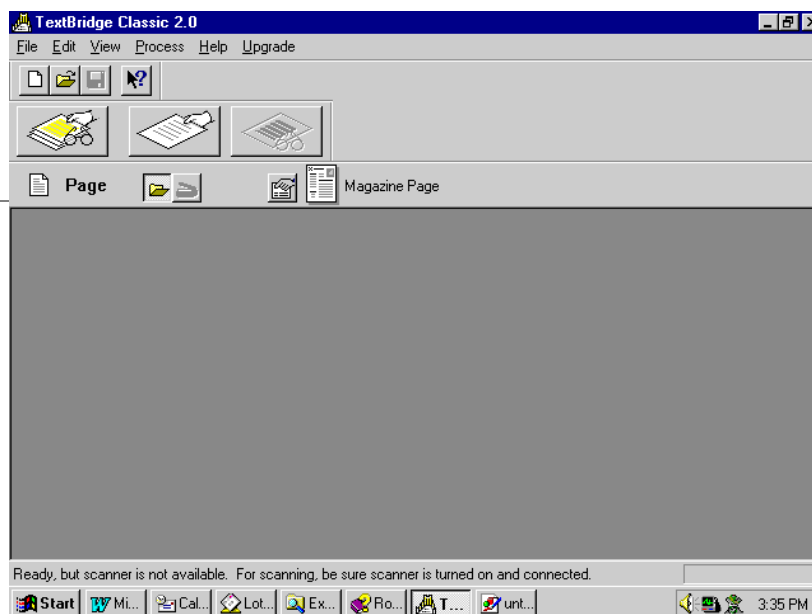


Figure 2-8. Page toolbar

Previewing the Page

After TextBridge gets a page of a document and before it begins page recognition, you preview the page. You commonly use Preview to check the contents, brightness, orientation, and quality of the page and delete unwanted pages from the document. After you check the page, zone it.

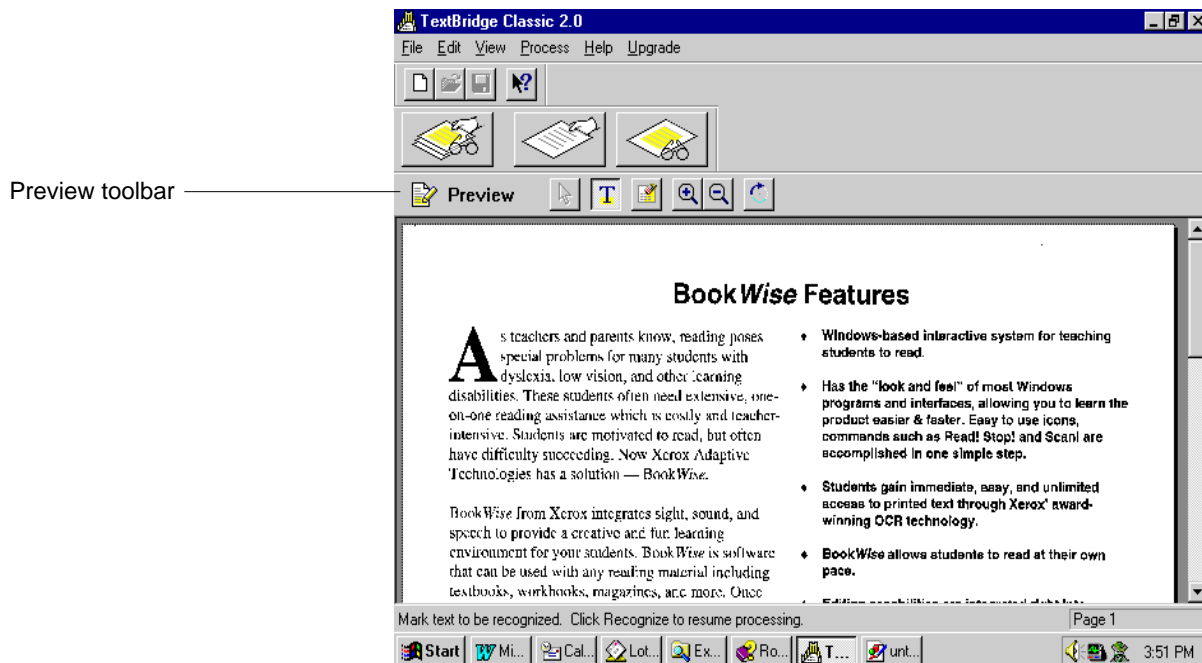


Figure 2-9. Preview the page using the Preview toolbar

Processing stops after TextBridge gets each page and displays the acquired image of the original page. At this point, you can perform one or more of the activities in the following list:

- u Check that this is the page you want.
- u Rotate the page to turn the page upright, if necessary.
- u Use the Zoom commands to magnify or reduce the page view to check the quality of the scanned page.
- u Delete the page from the document.
- u Adjust the settings for processing the page.
- u Cancel the process by creating a new file or opening another file.
- u Look at the properties of the page.
- u Continue processing the page.

Zoning the Page

You can use the Preview toolbar or View menu commands to examine and orient the acquired page.

During preview and before recognition can begin, the page must be zoned. TextBridge can automatically zone the page after you click the Recognize Page button, or you can zone the page manually. An acquired page is divided into one or more text zones. A **text zone** contains text that can be normal or reverse (white characters on a black background).

TextBridge assigns a color to distinguish text zones. You can change the assigned color in the Options dialog box. TextBridge also orders zones for output. TextBridge assigns a number to every text zone in the following order: headers, text including titles, headings, insets, and footers.

Only those parts of the page that are marked as text zones are recognized by TextBridge. If you want to recognize only part of a page, mark only that portion. Text zones are output in the order in which they were zoned. Pictures are not saved as part of the output.

When you stop to preview the page, you can manually zone the page. Use the zoning tools in the Preview toolbar called Text Marker and Erase Marker like highlighting markers to create and adjust zones.

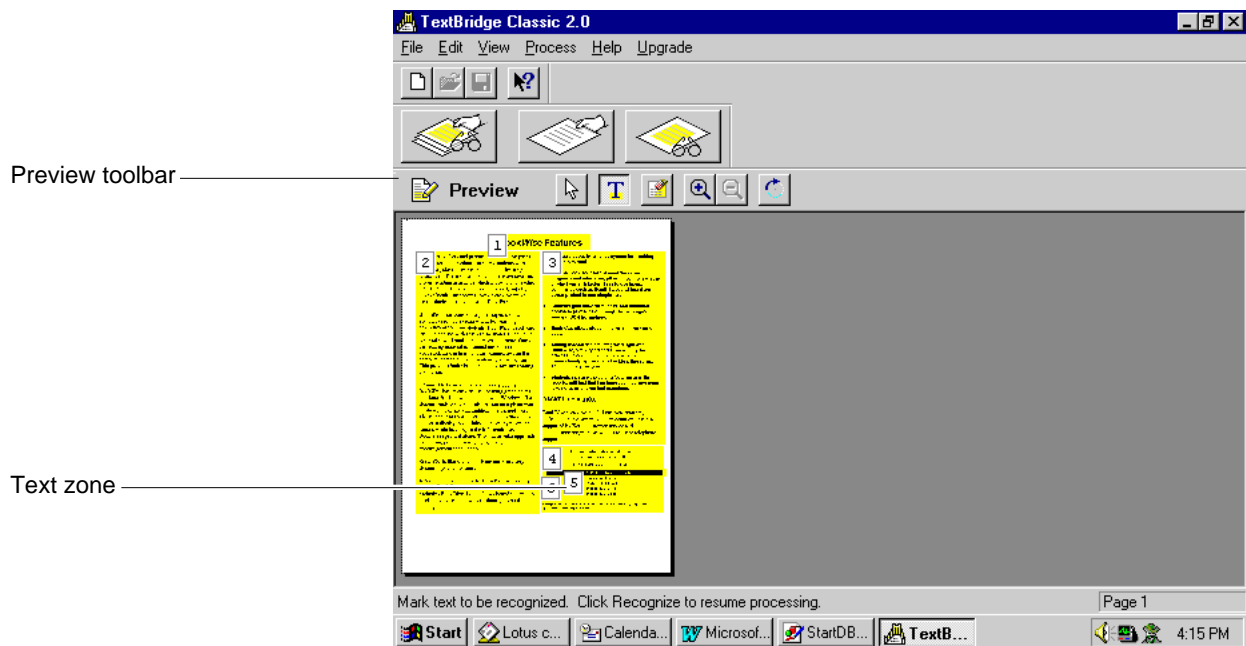


Figure 2-11. Zone the page using the Preview toolbar

You can perform these activities related to zones:

- u Use TextBridge automatic zoning.
- u Mark text zones.
- u Zone only part of a page.
- u View and adjust the text zoning by adjusting the size, merging, or splitting the zones.
- u Drag a selected zone to adjust its position.
- u Delete zones so that text is not included in the final document using the Clear command.

- u Enlarge or reduce the page to view the zones, using the Zoom In and Zoom Out buttons.
- u Use the zones you create on one page on all the following pages that you process in automatic mode in the current session.

You can also perform these less common activities related to zones:

- u Change the color used to highlight text zones.
- u Create polygonal zones with irregular shapes by using intersecting rectangular zones.

Use the Preview toolbar to quickly perform these activities. You can also use commands in the File, Edit and View menus for some activities related to zoning. After you complete the preview, tell TextBridge to recognize the page.

Recognizing the Page

TextBridge does all the work here. All you have to do is press the Recognize Page button, and TextBridge performs OCR on all the text zones on the page.

Saving the Document

After you finish previewing and TextBridge recognizes the document, you are ready to save it. If your page source is your scanner, once you save the document, you can not add any more pages to it. You can specify the location, name, and type of format of the output document. TextBridge converts the document to the format of your choice and saves it.

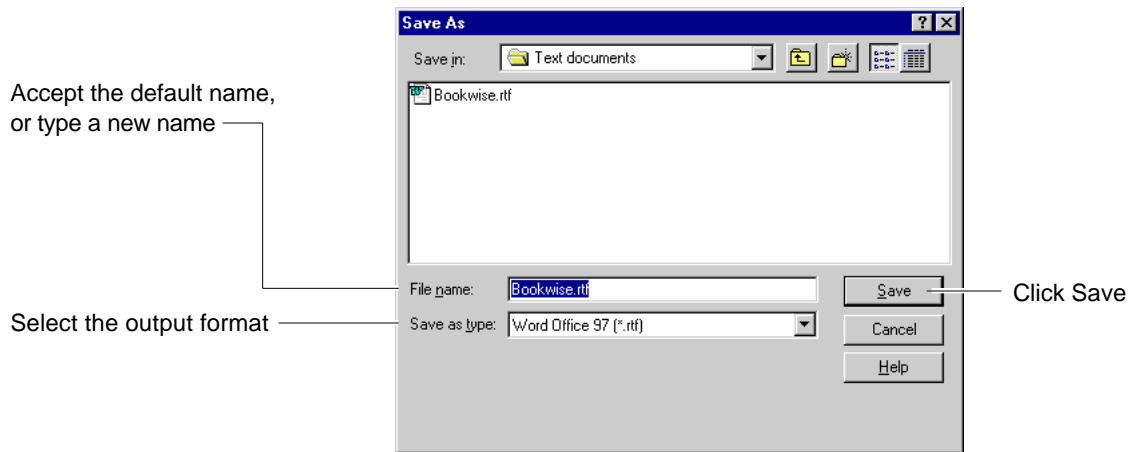


Figure 2-13. Saving the page using the Save As dialog box

After you save the document, the screen is blank and TextBridge is ready for you to begin a new job.

IMPROVING PAGE RECOGNITION WITH SETTINGS

There are a number of settings that you select in TextBridge at the beginning of the recognition process to help it recognize a document with more accuracy. Many of these options are related to the manual processes described in the previous section. Use the Settings dialog box to specify which options of the software you want to use.

Page Type Settings

Usually, you will want to use the settings automatically assigned to a page type. However, it is possible for you to change these settings.

You can view and change the settings for a page type in the Page Type tab of the Settings dialog box. Check the settings to be sure they are the best ones for processing the original page.

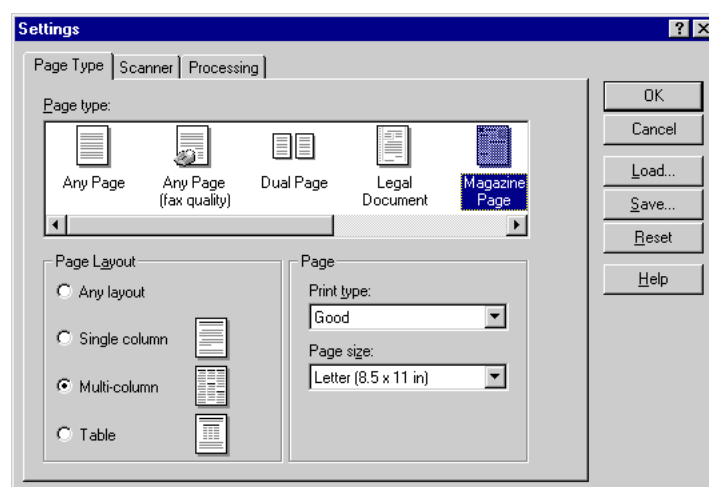


Figure 2-14. Page Type tab in the Settings dialog box

On the Page Type tab, you can choose the following page type settings for the specific page type you select:

- u Select the page layout of the original page:
 - Any layout
 - Single column with or without tables
 - Multi-column with or without tables
 - Table for pages with a table or spreadsheet and single-column text.

When you select Any layout, TextBridge automatically determines the page layout. Use Any layout when pages in your document have different layouts or when your pages do not fit the above layouts.

- u Set the print type of the document to be processed.

Any

Good

Fax

Dot matrix

Newspaper

When you select Any, TextBridge automatically determines the print type.

- u Set the page size to reflect the actual size of the original page:

Letter (8.5 X 11 in)

Legal (8.5 X 14 in)

A3 (297 X 420 mm)

A4 (210 X 297 mm)

Business Card (3.6 X 3.6 in)

Tabloid (11 X 17 in)

Letter Dual (8.5 X 11 in)

Legal Dual (8.5 X 14 in)

A4 Dual (210 X 297 mm)

Tabloid Dual (11 X 17 in)

A3 Dual (297 X 420 mm)

Note The available page sizes are based on your scanner's capabilities.

Scanner Settings

You can view and change the settings for your scanner in the Scanner tab of the Settings dialog box. For ISIS and TWAIN scanners that have been configured to allow TextBridge to control the scanner, check the scanner settings to be sure they are the best ones for processing the original page. If you are using a TWAIN driver with the TWAIN interface, TextBridge does not display the Scanner tab.

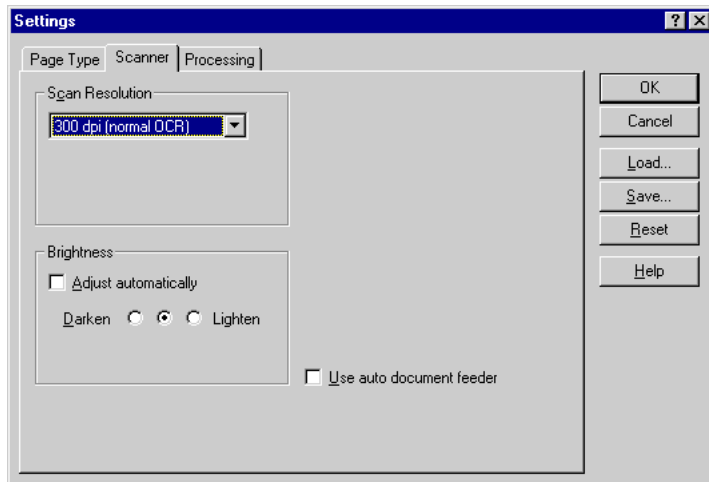


Figure 2-15. Scanner tab in the Settings dialog box

On the Scanner tab:

- u Choose the resolution value to reflect the actual resolution of your scanner. For most documents, use 300 dpi. For 8 point or smaller text, use 400 dpi for the best results.
- u Change the brightness based on whether your original page has light or dark text and pictures. For example, if the text and pictures on the original page are light, darken the brightness control.
- u Check the box to use the Automatic Document Feeder if your scanner has this feature and you are scanning multiple pages.

Processing Settings

You can view and change the settings for processing in the Processing tab of the Settings dialog box. Check the settings to be sure they are the best ones for processing the original page.

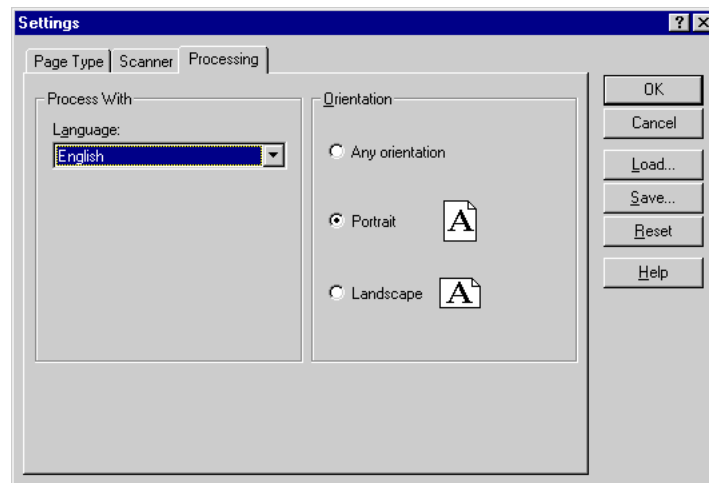


Figure 2-15. Settings dialog box and Processing tab

On the Processing tab:

- u Select the primary language of the document, which could be: English, German, French, Italian, Spanish, and other languages that you have.
- u Set the page orientation for the way text and images are printed on the original page:

Any orientation

Portrait

Landscape

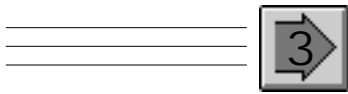
If you select Any orientation, TextBridge automatically determines the page orientation. This is the best selection for documents that contain both portrait and landscape orientations.

WHERE TO GO FROM HERE



To learn how you use TextBridge to process simple and complex documents, refer to Chapter 3. It also explains how to start TextBridge and use the Help system and sample documents, use Instant Access, plus view and zone your document in TextBridge.

The online Help system provides a complete reference to the user interface, including window areas, menus, commands, and tools as well as overview information and features, step-by-step procedures for using the software, tips, and a glossary.



LEARNING TO USE TEXTBRIDGE CLASSIC

The previous chapters have introduced you to TextBridge and document recognition. This chapter provides step-by-step instructions to teach you how to use the most important capabilities of TextBridge.

The learning sessions build on each other and assume that you understand the procedures explained in the previous sessions. It's best to do them in order or skim through prior sessions to familiarize yourself with the steps. Each learning session begins with introductory information including a list of what you will learn followed by step-by-step procedures and explanations.

The topics presented in this chapter are in the following list:

Starting TextBridge

Using the Help system

Using the sample documents Letter

Processing a simple document using auto processing Letter

Using Instant Access OCR Letter/Markplan

Processing a complex document using manual processing

STARTING TEXTBRIDGE

There are two ways to start TextBridge. You can start TextBridge as a standalone application or as Instant Access from any Windows-based text application.

In this section you will learn to start TextBridge as a standalone application.

To start TextBridge:

- 1. On the Windows task bar, click Start.**
- 2. Point to Programs, then point to the TextBridge Classic folder.**
- 3. Click the TextBridge Classic icon.**

The TextBridge Classic main window appears.

Note For these learning sessions you will be selecting the page type and page source from the Start dialog box. The instructions for these sessions assume that your screen looks like the following figure.

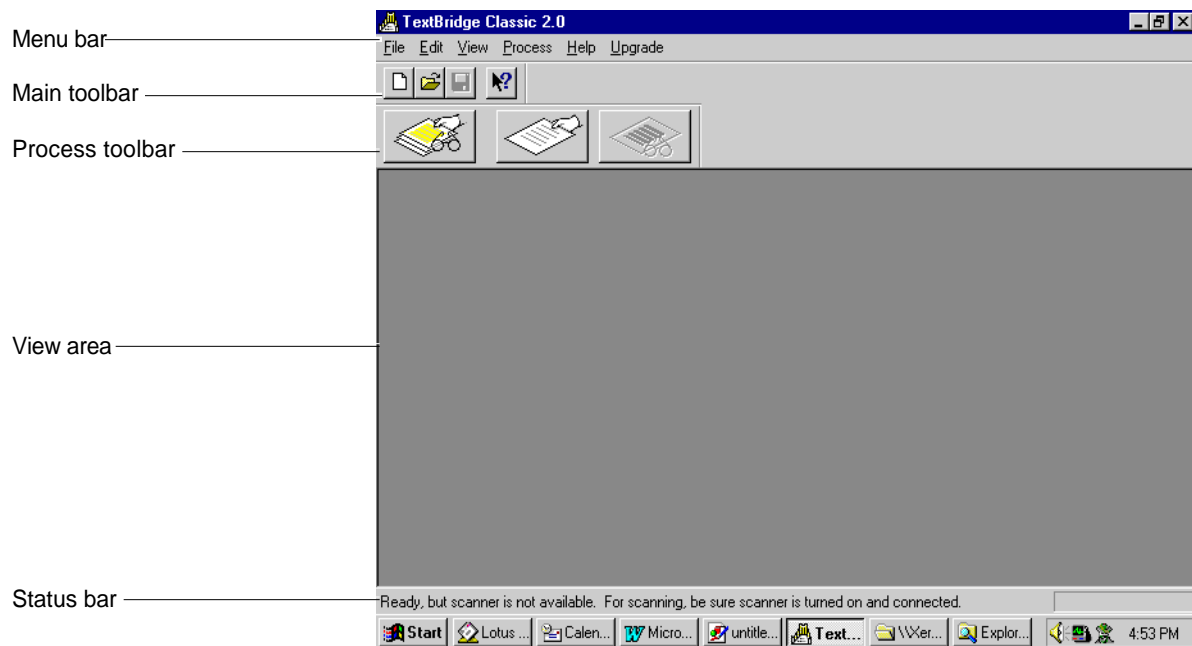


Figure 3-1. TextBridge Classic main window

If your screen does not look like this, in the View menu, select Toolbars. In the Toolbars dialog box, select Start dialog box in the Select Page Type and Source From box.

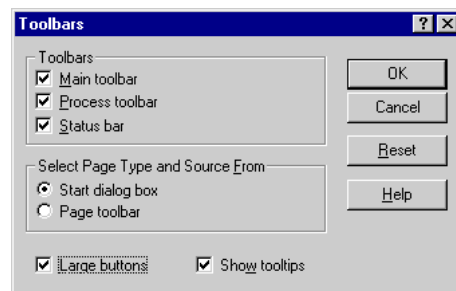


Figure 3-2. Toolbars dialog box

USING THE HELP SYSTEM



TextBridge is designed to be easy to learn and use. The online Help system provides general information about the program, step-by-step procedures for using the program, a glossary, and a complete reference to the user interface, including window areas, menus, commands, and tools.

- + In this section, you will learn to:

Get information from the Help system, including What's This? Help.

Use Help Topics window.

What you want to know about	How to get Help
Item in a dialog box or menu	Click the ? button then click the item <i>or</i> Right mouse click the item then click What's This? in the shortcut menu. <i>or</i> Select the item than press F1 or Shift+F1.
Entire dialog box	Click the Help button in the dialog box.
How to do something	Click TextBridge Help in the Help menu, then click Step-by-Step Procedures in the Contents tab. <i>or</i> Use the Index .
General information	Click TextBridge Help in the Help menu then click About TextBridge Classic in the Contents tab. <i>or</i> Use the Index .

What you want to know about

How to get Help

Meaning of a word

Click **Help** in the menu bar and in the **Contents** tab, click **Glossary**.

or

Use the Index.

A concept not listed in the Contents or Index

Click **TextBridge Help** in the Help menu then click **Find** and follow the directions.

You can get Help by using the main Help Topics window shown in the following figure.

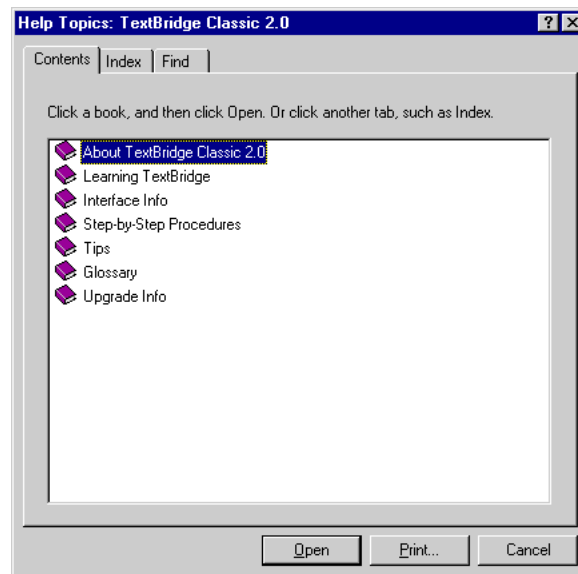


Figure 3-3. *Help Topics: TextBridge Classic Help window*

From the Help Topics window, you can get the help you want by performing one of the activities in the following list:

Select a topic from a book in the **Contents** tab.

Select a topic from the **Index** tab.

Search for information about a specific word or phrase using the **Find** tab.

Jump from one topic to a related topic.

USING THE SAMPLE DOCUMENTS



In this section, you will learn about the sample documents and how to open a sample document.

Use the sample documents provided on the installation CD with the learning sessions in this chapter. You can find the five sample documents in the installation folder in the following location:

`C:\Program Files\TextBridge Classic 2.0\Images\Samples`

This is the default location for these files; however, you may have installed them in another location. The sample documents are stored in TIFF format and are named:

Letter BookWise Markplan Plexis

The sample documents provide a cross-section of the page types that TextBridge can process:

Memo or Letter

Magazine Page

Any Page

Spreadsheet or Table

The sample documents are designed to provide you with documents on which to learn and to highlight the capabilities of the application. In each of the learning sessions, you are asked to use a specific sample document.

- + In this session, you will learn to open a sample document. For this session, use `Letter.tif`.

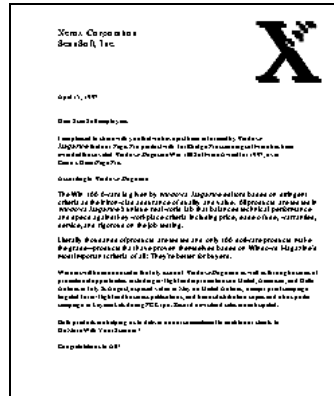


Figure 3-4. Letter sample document

To find and open a sample document:

1. **Click the Auto Process button.**



The Start dialog box appears.

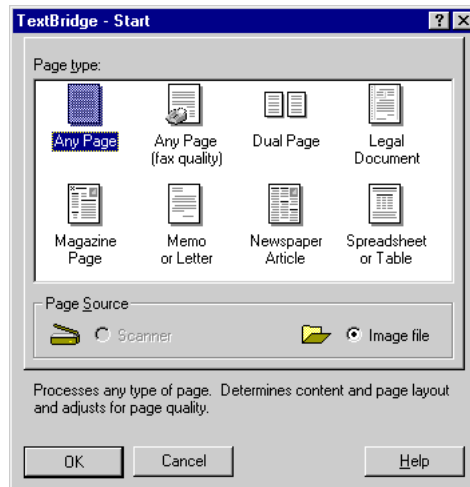


Figure 3-5. Start dialog box with *Any Page* and *Image file* selected

2. In the Start dialog box:

- Click *Any Page* in the Page type box.
- Click *Image file* in the Page source box.
- Click *OK*.

The Open dialog box appears. The default folder *Samples* is open. The sample TIFF files are listed in the Open dialog box.

Select an image file

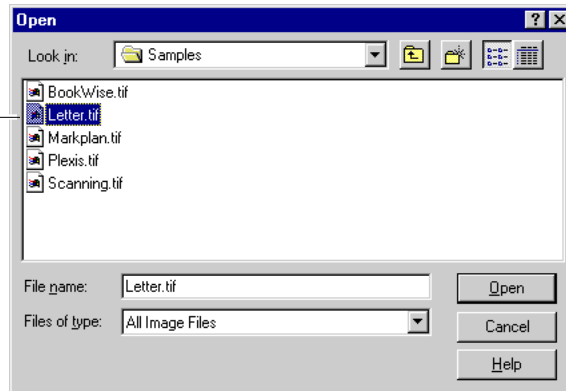


Figure 3-6. Open dialog box with *Letter.tif* selected

If *Samples* is not the open folder, access the sample documents folder in the following location from the Look In: box in the Open dialog box:

C:\Program Files\TextBridge Classic 2.0\Images\Samples

This is the default location unless you installed TextBridge in another place.

- 3. The Open dialog box, double-click a file name to open it. In this case, double-click *Letter.tif*.**

TextBridge gets the page as shown in the following figure.

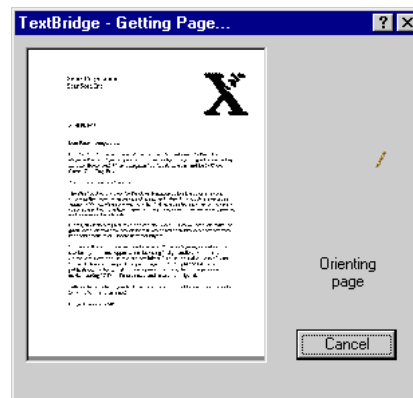


Figure 3-7. TextBridge - Getting Page dialog box

TextBridge automatically zones the page and identifies text as shown in the Zoning dialog box.

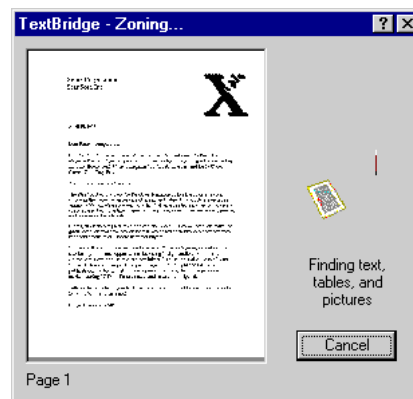


Figure 3-8. TextBridge - Zoning dialog box

TextBridge automatically recognizes the characters as shown in the Recognizing dialog box.

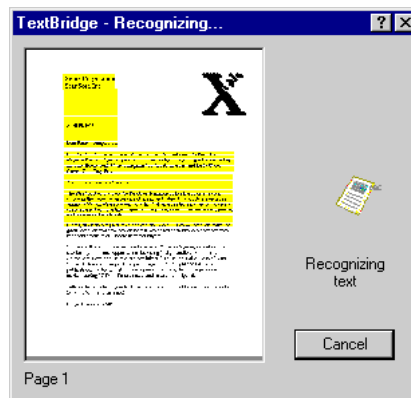


Figure 3-9. *TextBridge - Recognizing dialog box*

For this lesson, you just want to get back to where you started without saving the document.

4. Click Cancel in the Save As dialog box.

The current page will not be saved, and you return to the original TextBridge screen.

Now you know how to find and open a sample document .tif file.

Proceed to the learning sessions to work with TextBridge, and familiarize yourself with using its capabilities.

SESSION 1: PROCESSING A SIMPLE DOCUMENT USING AUTO PROCESSING



TextBridge provides a range of features. However, TextBridge is also designed to be very easy to use. For many documents, you can use default settings and automatically process a document.

- + For this learning session, use the sample document named *Letter*. This document has a single column of text and a logo.

In this session you'll learn to:

Use Auto Process

Use the Start dialog box

Select the Memo or Letter page type

Open an image file

Save a document after recognition

When you select **Memo or Letter** as the page type, it automatically specifies the following settings:

Single column page layout

Good print type

Letter size

TextBridge also uses the following default settings:

Scanner resolution 300 dpi

Scanner brightness at normal

Default language

Portrait orientation

Save in the Text Documents folder of the TextBridge folder

Save as standard format, which is .rtf

- + If you are using a TWAIN driver, check that the resolution is 300 dpi and brightness is normal.

Refer to Chapter 2 and Help for more information about these settings.

To process a simple document, use the following procedure:

1. **Start TextBridge.**

TextBridge appears.

2. **On the Process toolbar, click the Auto Process button.**



The Start dialog box appears.

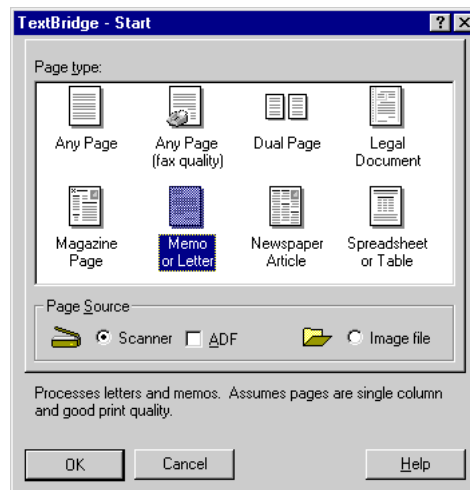


Figure 3-10. Start dialog box with Memo or Letter and Image file selected

3. In the Start dialog box:

- Click Memo or Letter in the Page type box.
- Select Image file in the Page source box.
- Click OK.

The Open dialog box appears.

Select an image file

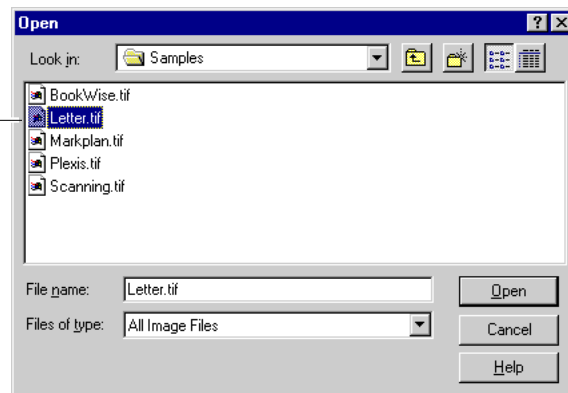


Figure 3-11. Open dialog box with Letter.tif selected

4. In the Open dialog box, double-click the sample document, Letter.tif.

TextBridge reads the image file, and automatically performs OCR on it, as indicated by the feedback display in the view area of the main window. Continue and save the document.

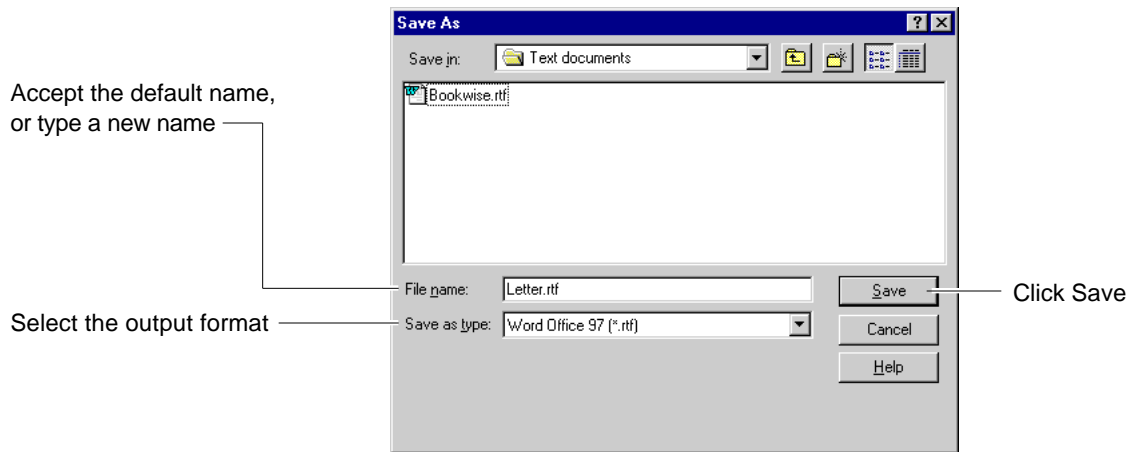


Figure 3-12. *Save As dialog box*

5. In the Save As dialog box, complete the following steps:

- In the Save in list, select the folder in which to save the text file.
- In the File name box, type a file name.
- In the Save as type list, select the output format for your word processor or other text application.
- Click the Save button.

TextBridge saves the document.

The status bar at the bottom of the screen confirms that you have saved the document. The main window remains open. You can start to process another document if you press the Auto Process button or the Get Page button.

- + Be sure to notice where the document is saved so that you can find it easily. The save location originally defaults to Text Documents in the TextBridge Classic folder in Program Files. You can check or change the default in the Settings dialog box Text Document tab.

7. Open the file in your word processor or other text application.

Unless you specified otherwise, open the file in the Text Documents folder of the TextBridge Classic folder in the Program Files folder. You can use the shortcut to this location.

- + The document is saved with an .rtf extension if your text application is Word. In the Open dialog box of your word processor, check that you can see files of this type listed.

Compare the recognized document in your word processor with the picture of the sample document, Letter.tif.

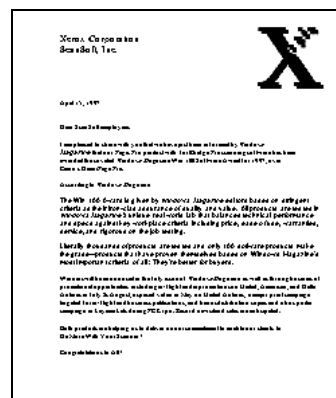
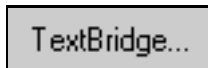


Figure 3-13. Letter sample document

The difference is that now you have formatted, fully editable text, just as if you had typed it in yourself. At this point, you could spell check the document and make any other changes in your word processor.

- + The large X does not appear in the final document because TextBridge interprets it as a picture, and pictures are not retained in TextBridge Classic.

SESSION 2: USING INSTANT ACCESS OCR



You can use TextBridge **Instant Access OCR** to run TextBridge from within another application, such as a word processor. To use Instant Access, simply start TextBridge from within an application, such as Word or WordPerfect. During Instant Access OCR, TextBridge processes a document then pastes it into the open document in your text application.

- + For this learning session, use the sample document named *Markplan*. This document has a single column of text, a title, headings, and bullet lists. The procedure is similar to processing a simple document.

In this session you'll learn to:

Use TextBridge Instant Access in your word processor.

Select the Any Page page type.

When you select **Any Page** as the page type, it automatically specifies the following settings:

Any page layout

Any print type

Letter size

TextBridge also uses the following default settings:

Scanner resolution 300 dpi

Scanner brightness at normal

Default language

Portrait orientation

Auto process all pages

One file for all pages

Refer to Chapter 2 and Help for more information about these settings.

- + If you are using a TWAIN driver, check that the resolution is 300 dpi and brightness is normal.
- + If TextBridge is still running from the previous learning session, exit from TextBridge. This will let you run TextBridge from your word processor. You can not have more than one copy of TextBridge running at the same time.

Before you run TextBridge as Instant Access, you may need to use the Instant Access Control Panel to choose which applications have Instant Access to TextBridge. TextBridge automatically provides Instant Access for the applications listed in the control panel, as shown by the check mark.

If you want to examine the status of Instant Access, click Start on the Windows taskbar, then Programs, then TextBridge Classic, and Instant Access Control Panel. You can also access the Instant Access Control Panel from the main TextBridge application in the File menu by clicking Instant Access Control Panel. Help provides additional information about Instant Access.

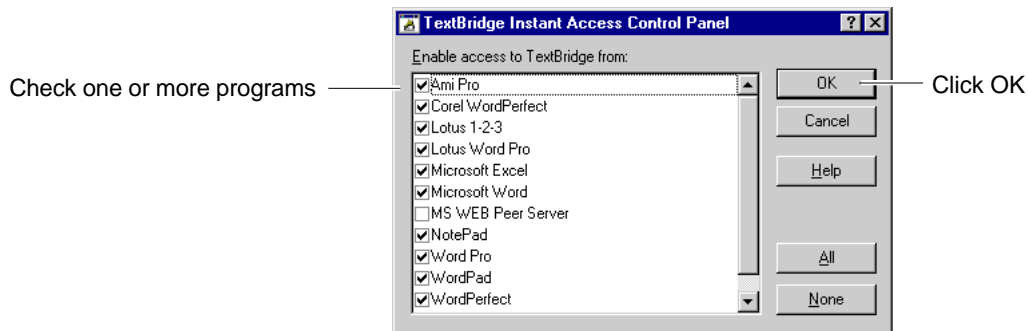


Figure 3-14. *TextBridge Instant Access Control Panel*

The Enable access to TextBridge list shows the text applications from which TextBridge can be invoked. The list includes applications commonly used with TextBridge and applications that are currently running. If your application does not appear in this list, close the TextBridge Instant Access Control Panel, start your application, and reopen the TextBridge Instant Access Control Panel. Your application should now appear in the list.

Click on applications in the list to check or uncheck them. Click All to check all items in the list. Click None to uncheck all items in the list. Instant Access to TextBridge will be available from all checked applications.

Click OK to close the Instant Access Control Panel and save any changes you specified.

To use Instant Access OCR from your word processor, use the following procedure:

1. **Start your word processor, and open a new document.**
2. **In the File menu, click the TextBridge... command.**

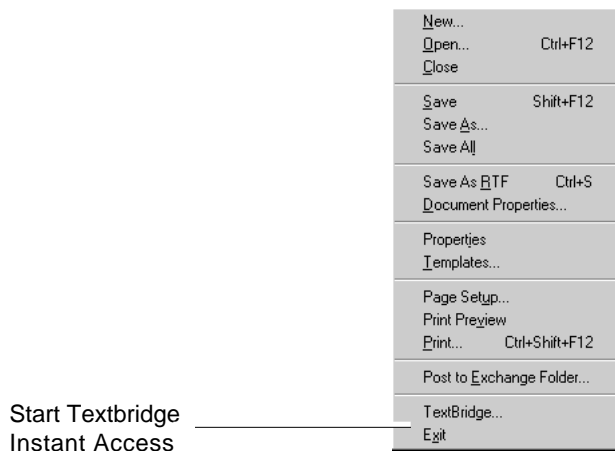


Figure 3-15. TextBridge... command in File menu

The Start dialog box appears. Notice that the Start dialog box is slightly different than the Start dialog box in the standalone version of TextBridge. Processing has been added.

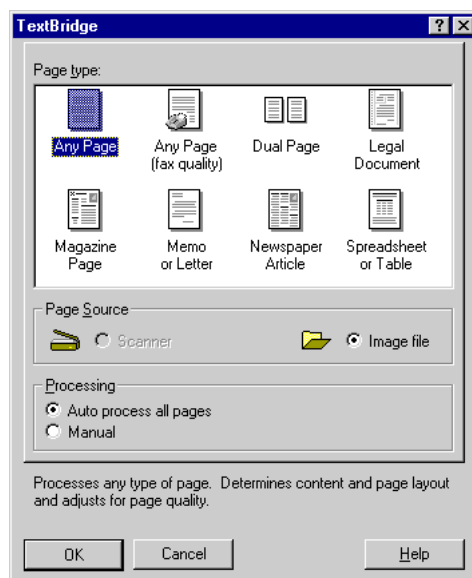


Figure 3-16. Start dialog box for Instant Access

3. In the Start dialog box:

- In the Page type box, click Any Page.
- In the Page Source box, select Image file.
- In the Processing box, select Auto process all pages.
- Click OK.

The Open dialog box appears.

Select an image file

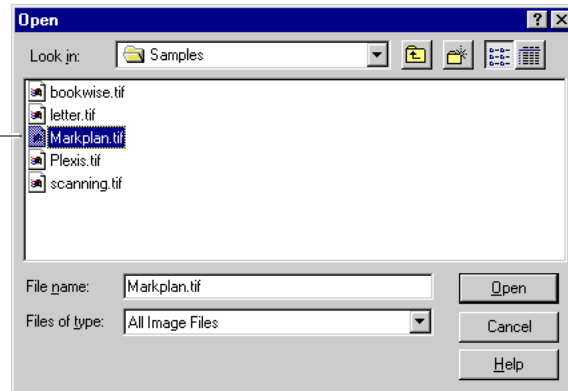


Figure 3-17. Open dialog box with *Markplan.tif* selected

- 4. In the Open dialog box, double-click the sample document, *Markplan*.**

TextBridge reads the image file, and automatically performs OCR on it, as indicated by the feedback display in the view area of the main window. After acquiring and recognizing the page, TextBridge pastes the document into the open document in your word processor.

If TextBridge can not do this, you will get a message that tells you, “TextBridge is unable to paste text automatically. Select Paste from the Edit menu to manually copy the recognized text from the clipboard.” You may have to open a new document before you can use the Paste command. After you follow these instructions, the processed document is pasted in the open document of your word processor.

Compare the recognized document in your word processor with the reproduction of the sample document, *Markplan.tif*.

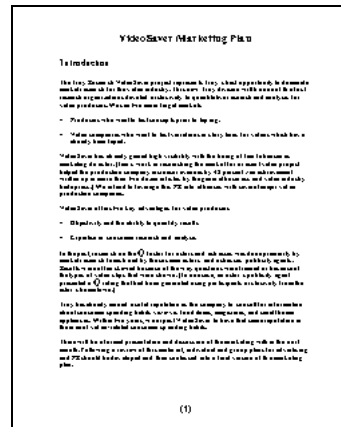


Figure 3-18. Markplan sample document

- + If this document continues to a second page, delete any additional spacing that was inserted into the document.

You can save the document or make any changes you'd like to the document just as if you'd typed it yourself. For example, you can spell check it and save it with your changes.

SESSION 3: PROCESSING A COMPLEX DOCUMENT USING MANUAL PROCESSING



For more complex documents such as magazine articles, you often can use TextBridge in automatic mode. However, simply using a few additional steps in manual mode can sometimes produce a more accurate result in less time.

- + For this learning session, use the sample document named BookWise. This document has multiple columns, a dropped capital letter, headings, paragraphs, bullet lists, and reversed video text.

In this session you'll learn to:

Use manual processing with the Get Page button.

Select Magazine Page type.

Zone a page.

Use the Zoom button.

Save a page.

Edit the document in your word processor.

Refer to Chapter 2 and Help to learn more about zoning.

When you select **Magazine Page** as the page type, it automatically specifies the following settings:

Multi-column page layout

Good print type

Letter size

TextBridge also uses the following default settings:

Scanner resolution 300 dpi

Scanner brightness at normal

Default language

Portrait orientation

Save as standard format

Save in Text Documents folder

- + If you are using a TWAIN driver, check that the resolution is 300 dpi and brightness is normal.

Run the standalone version of TextBridge from the Start button for this learning session.

1. **Start the TextBridge standalone version.**
2. **Click the Get Page button.**



The Start dialog box appears.

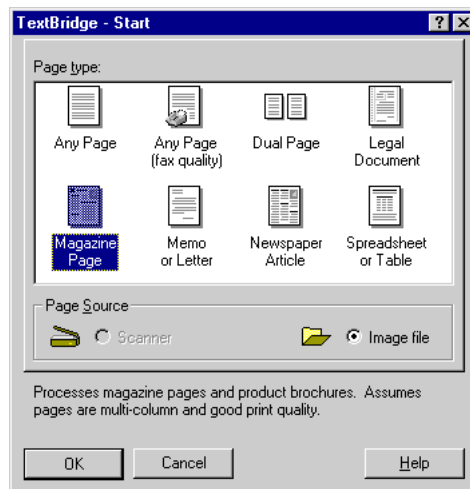


Figure 3-19. Start dialog box with Magazine Page and Image file selected

3. **In the Start dialog box:**
 - Click Magazine Page in the Page type box.

The settings are automatically set to multi-column page layout, good print type, letter size.

- Click Image file under Page source.

The application knows where to get the page.

- **Click OK.**

The Open dialog box appears.

4. Double click BookWise.tif.

TextBridge gets the page, displays it, and displays the Preview toolbar so that you can preview it.

The page you see should be a two-column magazine article beginning with a drop cap.

- + If this is not the correct page, in the File menu, click New. Click OK to close the current document. You can begin again by selecting Get Page.

5. Manually zone the page.

Create six text zones in the following order. The numbers on the text zones reflect the order in which these zones will appear in the final document.

- Mark the title as zone one.
- Mark the left column of text as zone two.
- Mark the right column of text as zone three.
- Mark “For more information” to “Ext. 1” as zone four.

- Mark the “Xerox Adaptive Products” reverse video text as zone five.
- Mark the remaining text as zone six.

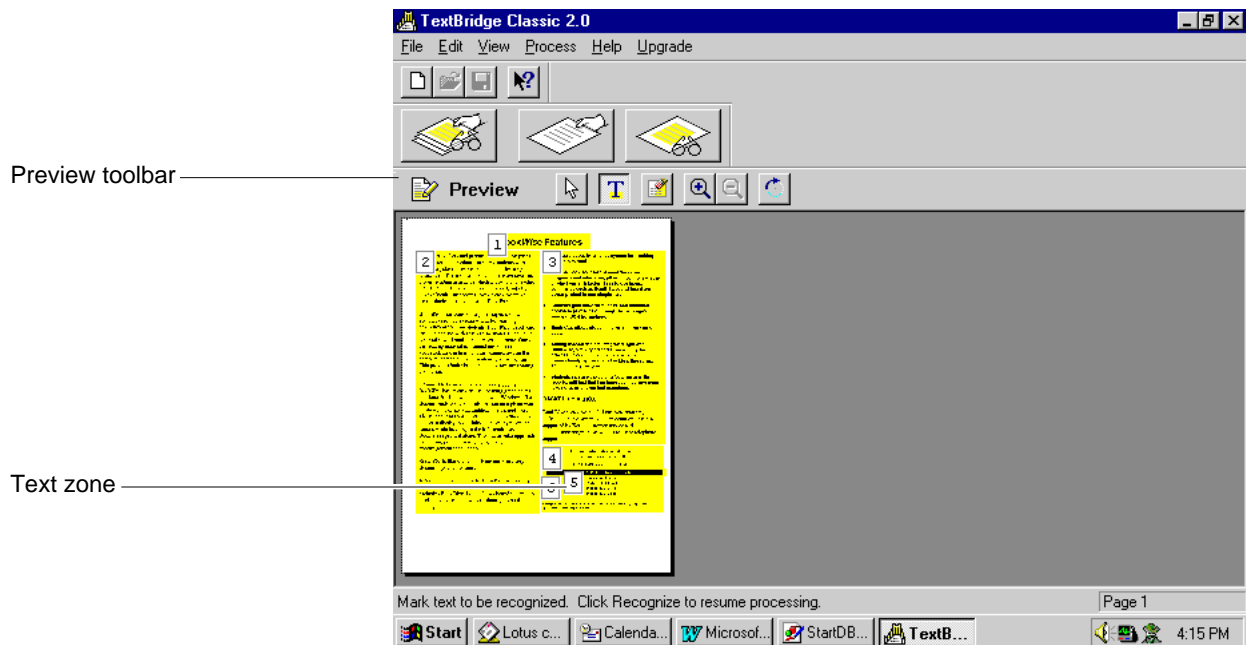


Figure 3-20. Zoned magazine page

- Click the Zoom In and Zoom Out buttons to enlarge and reduce the page to examine the zones, if necessary.

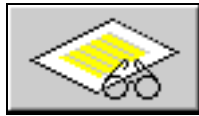


TextBridge magnifies the page.

- + Reverse video text must be in a separate text zone that includes no regular text. Also, the zoom tools are useful in determining that the text has been scanned at the appropriate scanner settings. If the result is too light or too dark, adjust the scanner settings, and rescan the document.

When the zones are accurate, continue with the next step, which is page recognition.

6. Click the Recognize Page button.



TextBridge performs OCR and recognizes the page.

The Save As dialog box appears.

7. Save the page as Magazine Page.rtf.

TextBridge provides a suggestion for the file name and uses the type of file you selected last. If you prefer another name, enter the new name in the File name box. Make any other changes, then click the Save button. TextBridge formats the document and saves the file.

8. Open Magazine Page.rtf in your word processor.

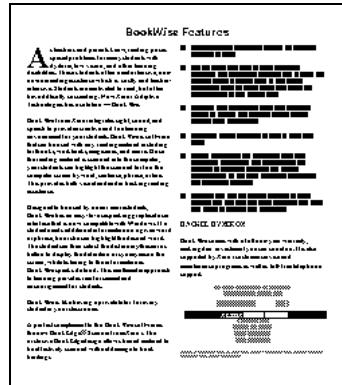


Figure 3-21. BookWise sample document

The page contains the text from the original page but does not include the original layout. The document is a fully editable version of BookWise in your word processor.

9. Edit the document in your word processor.

You can make changes to the text and layout in your word processor and spell check the page.

WHERE TO GO FROM HERE



The learning sessions in this chapter were designed to give you a solid basis on which to use TextBridge for your own documents.

For complete information about TextBridge, please refer to the online Help built into the application.