

PERCEIVE[®]

personal

WINDOWS

USER MANUAL

Notice

Information in this manual is subject to change without notice. The software described in this manual is furnished under a license agreement or nondisclosure agreement. The software may be used or copied only in accordance with the terms of agreement. The purchaser may make one copy of the software for backup purposes.

Even though OCRON has tested the software and reviewed the documentation, **OCRON makes no warranty or representation, either expressed or implied, with respect to this software, its quality, performance, merchantability, or fitness for a particular purpose. As a result, this software is sold "As Is," and you, the purchaser, are assuming the entire risk as to its quality and performance.**

In no event will OCRON be liable for direct, indirect, special, incidental or consequential damages resulting from any defects or misuse in the software or its documentation.

No part of this manual may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying or recording, for any purpose other than the purchaser's personal use without the prior written permission of OCRON, Inc.

©Copyright OCRON, Inc., 1989-1993. All rights reserved.

Perceive Personal is a trademark of OCRON, Inc.

Microsoft, Microsoft Windows and MS-DOS are registered trademarks of Microsoft Corporation.

IBM is a registered trademark of International Business Machines Corporation.

Other products mentioned in this manual are trademarks of their respective holders.

Software Product License Agreement

IMPORTANT: PLEASE READ THIS DOCUMENT CAREFULLY BEFORE OPENING THE MEDIA ENVELOPE. BY OPENING THE MEDIA ENVELOPE, YOU ARE AGREEING TO BECOME BOUND BY THE TERMS OF THIS AGREEMENT, WHICH INCLUDES THE SOFTWARE LICENSE AND DISCLAIMER OF WARRANTY. IF YOU DO NOT AGREE TO THE TERMS OF THIS AGREEMENT, DO NOT OPEN THE MEDIA ENVELOPE, PROMPTLY RETURN THE UNOPENED MEDIA ENVELOPE TO THE PLACE WHERE YOU OBTAINED IT FOR A FULL REFUND.

The enclosed computer program ("OCRON Software") is licensed, not sold, to you for use only under the terms of this license, and OCRON reserves any rights not expressly granted to you. You own the disk on which the OCRON Software is originally or subsequently recorded or fixed, but OCRON retains ownership of all copies of the OCRON Software itself.

1. License. This license allows you to:

(a) Use the OCRON Software only on a single host computer. You must obtain a supplementary license from OCRON before using the OCRON Software in connection with systems, multiple central processing units, computer networks, or emulations on a mainframe or minicomputer.

(b) Make one copy of the OCRON Software in machine-readable form solely for backup purposes. The OCRON Software is protected by United States copyright law. You must reproduce on each copy the OCRON copyright notice and any other proprietary legends that were on the original copy.

(c) Transfer the OCRON Software and all rights under this license to another party together with a copy of this Agreement provided you give OCRON written notice of the transfer and the other party reads and agrees to accept the terms and conditions of this agreement.

2. Restrictions. You may not distribute copies of the OCRON Software to others or electronically transfer the OCRON Software from one computer to another over a network. The OCRON Software contains trade secrets and in order to protect them you may not decompile, reverse engineer, disassemble, or otherwise reduce the OCRON Software to a human-perceivable form. YOU MAY NOT MODIFY, ADAPT, TRANSLATE, RENT, LEASE, LOAN, RESELL FOR PROFIT, DISTRIBUTE, NETWORK, OR CREATE DERIVATIVE WORKS BASED UPON THE OCRON SOFTWARE OR ANY PART THEREOF.

3. Termination. This License is effective until terminated. This License will terminate automatically without notice from OCRON if you fail to comply with any provision of this License. Upon termination you must destroy the OCRON Software and all copies thereof. You may terminate this License at anytime by destroying the OCRON Software and all copies thereof.

4. Export Law Assurances. You agree and certify that neither the OCRON Software nor any direct product thereof is being or will be shipped, transferred, or reexported, directly or indirectly, into any country prohibited by the United States Export Administration Act and the regulations thereunder or will be used for any purpose prohibited by the same.

5. Government End Users. If this OCRON Software is acquired by or on behalf of a unit or agency of the United States Government this provision applies. This OCRON Software: (a) was developed at private expense, and no part of it was developed with government funds, (b) is a trade secret of OCRON for all purposes of the Freedom of Information Act, (c) is "commercial computer software" subject to limited utilization as provided in the contract between the vendor and the governmental entity, and (d) in all respects is proprietary data belonging solely to OCRON, Inc.

For units of the Department of Defense (DoD), this OCRON Software is sold only with "Restricted Rights" as that term is defined in the DoD Supplement to the Federal Acquisition Regulations, 52.227-7013 (b) (3) (ii) and:

Use, duplication, or disclosure is subject to restrictions as set forth in subdivision (b) (3) (ii) of the Rights in Technical Data and Computer Software clause at 52.227-7013. Manufacturer: OCRON, Inc., 3350 Scott Blvd. Bldg. 36, Santa Clara, California 95054.

If this OCRON Software was acquired under a GSA Schedule, the Government has agreed to refrain from changing or removing any insignia or lettering from the Software or the documentation that is provided or from producing copies of manuals or disks (except one copy for backup purposes) and:

(1) Title to and ownership of this OCRON Software and documentation and any reproductions thereof shall remain with OCRON, Inc.

(2) Use of this OCRON Software and documentation shall be limited to the facility for which it is acquired; and

(3) If the use of the OCRON Software is discontinued at the installation specified in the purchase/delivery order and the Government desires to use it at another location, it may do so by giving prior written notice to OCRON, Inc., specifying the type of computer and new location site. Governmental personnel using this OCRON Software, other than under a DoD contract or GSA Schedule, are hereby on notice that use of this OCRON Software is subject to restrictions that are the same as or similar to those specified above.

6. Disclaimer of Warranty. THE OCRON SOFTWARE IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION ANY WARRANTY WITH RESPECT TO ITS MERCHANTABILITY OR ITS FITNESS FOR ANY PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE OCRON SOFTWARE IS WITH YOU. SHOULD THE OCRON SOFTWARE PROVE DEFECTIVE, YOU (AND NOT OCRON OR AN OCRON AUTHORIZED REPRESENTATIVE) ASSUME THE ENTIRE COST OF ALL NECESSARY SERVICING, REPAIR, OR CORRECTION.

OCRON DOES NOT WARRANT THAT THE FUNCTIONS CONTAINED IN THE OCRON SOFTWARE WILL MEET YOUR REQUIREMENTS OR THAT THE OPERATION OF THE OCRON SOFTWARE WILL BE UNINTERRUPTED OR ERROR FREE OR THAT DEFECTS IN THE OCRON SOFTWARE WILL BE CORRECTED.

SOME STATES DO NOT ALLOW THE EXCLUSION OF IMPLIED WARRANTIES, SO THE ABOVE EXCLUSION MAY NOT APPLY TO YOU. THIS WARRANTY GIVES YOU SPECIFIC LEGAL RIGHTS AND YOU MAY ALSO HAVE OTHER RIGHTS, WHICH VARY FROM STATE TO STATE.

7. Limitation of Remedies and Damages. IN NO EVENT WILL OCRON BE LIABLE (i) TO YOU FOR ANY INCIDENTAL, CONSEQUENTIAL, OR INDIRECT DAMAGES (INCLUDING DAMAGES FOR LOSS OF BUSINESS PROFITS, BUSINESS INTERRUPTION, LOSS OF BUSINESS INFORMATION, AND THE LIKE) ARISING OUT OF THE USE OF OR INABILITY TO USE ANY OCRON SOFTWARE EVEN IF OCRON OR ANY AUTHORIZED OCRON REPRESENTATIVE HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES, OR (ii) FOR ANY CLAIM BY ANY OTHER PARTY.

BECAUSE SOME STATES DO NOT ALLOW THE EXCLUSION OR LIMITATION OF LIABILITY FOR CONSEQUENTIAL OR INCIDENTAL DAMAGES, THE ABOVE LIMITATIONS MAY NOT APPLY TO YOU.

OCRON's liability to you for actual damages for any cause whatsoever, and regardless of the form of the action, will be limited to the greater of \$50 or the money paid for the OCRON Software that caused the damages or that is the subject matter of, or is directly related to, the cause of action.

8. General. This Agreement will be construed under the laws of the state of California, except for that body of law dealing with conflicts of law. If any provision of this Agreement shall be held by a court of competent jurisdiction to be contrary to law, that provision will be enforced to the maximum extent permissible, and the remaining provisions of this Agreement will remain in full force and effect.

Limited Warranty and Limitations of Warranty

A. Limited Warranty on Media. OCRON warrants the disks on which the OCRON Software is recorded to be free from defects in materials and faulty workmanship under normal use for a period of ninety (90) days from the date of delivery as evidenced by a copy of the receipt. OCRON will, at its option, replace or refund the purchase price of the disk at no charge to you, provided you return the faulty disk with proof of purchase to OCRON or an authorized OCRON dealer. OCRON shall have no responsibility to replace or refund the purchase price of a disk damaged by accident, abuse, or misapplication. ANY IMPLIED WARRANTIES ON THE DISKS, INCLUDING IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, ARE LIMITED IN DURATION TO 90 DAYS FROM THE DATE OF DELIVERY.

B. Disclaimer of Warranty and Limited Warranty. THE OCRON SOFTWARE IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND AND OCRON EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES, INCLUDING BUT NOT LIMITED TO THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. OCRON DOES NOT WARRANT, GUARANTEE, OR MAKE ANY REPRESENTATIONS REGARDING THE USE, OR THE RESULTS OF THE USE, OF THE OCRON SOFTWARE IN TERMS OF CORRECTNESS, ACCURACY, RELIABILITY, CURRENTNESS, OR OTHERWISE. THE ENTIRE RISK AS TO THE RESULTS AND PERFORMANCE OF THE OCRON SOFTWARE IS ASSUMED BY YOU. OCRON DOES NOT WARRANT THAT THE OCRON SOFTWARE WILL WORK CORRECTLY IN A MULTIUSER ENVIRONMENT. IF THE OCRON SOFTWARE IS DEFECTIVE, YOU, AND NOT OCRON OR ITS DEALERS, DISTRIBUTORS, AGENTS, OR EMPLOYEES, ASSUME THE ENTIRE COST OF ALL NECESSARY SERVICING, REPAIR, OR CORRECTION. SOME STATES DO NOT ALLOW THE EXCLUSION OF IMPLIED WARRANTIES, SO THE ABOVE EXCLUSION MAY NOT APPLY TO YOU.

C. Limitations of Warranty and Liability. THE LIMITED WARRANTIES PROVIDED IN PARAGRAPHS A AND B ABOVE ARE THE ONLY WARRANTIES OF ANY KIND THAT ARE MADE BY OCRON ON THIS OCRON PRODUCT. NO ORAL OR WRITTEN INFORMATION OR ADVICE GIVEN BY OCRON OR ITS DEALERS, DISTRIBUTORS, AGENTS, OR EMPLOYEES SHALL CREATE A WARRANTY OR IN ANYWAY INCREASE THE SCOPE OF THIS WARRANTY, AND YOU MAY NOT RELY ON ANY SUCH INFORMATION OR ADVICE. THIS WARRANTY GIVES YOU SPECIFIC LEGAL RIGHTS. YOU MAY HAVE OTHER RIGHTS, WHICH VARY FROM STATE TO STATE.

Table of Contents

Chapter 1	Introduction	1-1
	A Simple but Powerful Tool	1-1
	Features	1-2
	Reading Text	1-3
	An Innovative On-line Text Merge	1-4
Chapter 2	Getting Started	2-1
	Installation	2-1
	What You Need	2-1
	What Comes with Perceive Personal	2-2
	Installing Your Scanner	2-2
	Installing Perceive Personal	2-2
	Readme File	2-3
	Starting Perceive	2-3
	Selecting Scanner	2-3
	Using Help	2-4
	Closing Perceive	2-4
Chapter 3	Tutorials	3-1
	Scanner Settings	3-1
	Exercise 1: Scanning Multiple Column Documents	3-3
	Exercise 2: Opening a Previously Scanned Image File	3-9
	Exercise 3: Scanning Across Single Column Documents	3-11
	Exercise 4: Scanning Down Single Column Documents	3-13
Chapter 4	Using Your Scanner	4-1
	The Scanner Controls	4-1
	Scanning Tips	4-3
	Judging Image Quality	4-5
	Adjusting Contrast	4-5
	Touching Characters	4-6
	Broken Characters	4-6
	Filled-in or Blurred Characters	4-7
	Other Factors	4-7
	Commonly Misread Characters	4-7
	Point Sizes	4-8
	Scanning Different Document Types	4-8

Single Column	4-8
Multiple Columns	4-9
Tables	4-9
Text Mixed with Graphics	4-10
Legal	4-10
Dot Matrix	4-10

Chapter 5

Command Summary	5-1
Toolbar	5-1
Scanning Icons	5-1
Zone Icons	5-2
Zoom Icons	5-4
OCR Icon	5-4
Append/Merge Icons	5-4
File Menu	5-6
Edit Menu	5-8
Image Menu	5-9
Settings Menu	5-9
Window Menu	5-11
Help Menu	5-12
Keyboard Shortcuts	5-12

Appendix A

Troubleshooting	A-1
----------------------------------	------------

Glossary

.	G-1
-----------	------------

Index

.	I-1
-----------	------------

Chapter 1 Introduction

Welcome to Perceive Personal™! Perceive Personal is a new-generation optical character recognition (OCR) tool that makes automated text input into your personal computer a reality. (Perceive Personal will be referred to as “Perceive” throughout this manual.)

Perceive uses *omnifont* technology, which means you can automatically read text composed of commonly used typefaces from typesetters, laser printers, typewriters and other sources. With Perceive and your hand-held scanner, you can read magazine or newspaper articles and distribute them through electronic mail or build a personal information database. You can even scan a full page of text (using two or more scanner passes) and Perceive will automatically merge the text for you.

After the text is recognized, it can be edited from the keyboard, then saved in a file format that can be used with word processors, desktop publishers, databases and any other text application.

A Simple but Powerful Tool

Perceive is a Microsoft® Windows-based software package, which means that all operations, control and setup is accomplished through easy-to-use icons and “pull down” menus. Text is input with just few “clicks” and “drags” of the mouse.

Perceive offers much more than simple text recognition. With Perceive, you can edit text, recognize text from stored image files, and merge vertical or horizontal strips of recognized text — all without leaving the application.

Perceive can operate on a 286 or higher PC/AT without a special adapter board or large additions of memory. And, because the program is fully compatible with the Microsoft Windows environment, it may be run side-by-side with other applications — a perfect complement to word processing, desktop publishing and database management applications.

Features

- Interacts with supported hand-held scanners to scan images. Text can be scanned from the top down, left to right or right to left.
- Easy-to-use toolbar icons.
- Automatic page segmentation for multi and single column text.
- Fast and accurate recognition of English and European documents.
- Innovative text merge lets you scan any size page using as many passes with your hand-held scanner as needed. Merges the left and right sides, or the top and bottom, and removes any overlap.
- Recognized text can be saved in a variety of formats so it can be used with word processors, desktop publishers, databases, spreadsheets, and any other text application.
- Scanned images can be saved in TIFF, Compressed TIFF, or PCX format.
- Extensive on-line help.

Reading Text

Converting printed text to computer-readable form is a two-step process. First, the document is scanned and entered into your computer as an electronic image. Second, the image is "recognized" by the OCR software to identify and convert the text characters so they can be edited on the computer. In the first step, an optical scanner creates an electronic image of the scanned text. On a microscopic level, the electronic image is composed of small dots which, in their entirety, comprise a "picture" of the text or graphics on the page. This electronic image cannot be used with text or numbered-based applications. These applications require characters from a computerized set of coded characters such as "ASCII" or "ANSI." This is where OCR offers a real benefit to the user. It converts the image into coded characters — eliminating the need for manual typing.

While recognition speed is a function of the image quality, computing power and other factors, in most cases an entire page of text can be read in one to three minutes. This compares to eight to fifteen minutes per page for a fast typist to type the text.

An Innovative On-line Text Merge

With Perceive and your hand-held scanner, you are not limited to scanning and reading just columns of text or numbers. Perceive's innovative on-line text merge offers you unparalleled ease-of-use when scanning a full page of text.

Perceive lets you assemble text two ways:

- Append or merge below the previously scanned text.
- Merge the left and right halves of the page.

When scanning multiple column documents, such as magazine articles, you scan and recognize the first column of text, then scan and recognize the second column. Perceive will append the second column below the first column. You can continue scanning, recognizing and appending text until you are finished with your document. If the scanned image is of poor quality, you can even rescan the last portion and continue appending to the previous column of text — all using the on-line text merge in Perceive.

When reading single column documents, such as legal contracts, Perceive offers you the flexibility of scanning the pages horizontally or vertically and assembling the pages into a single file. Vertical scanning can be used to scan pages from the middle of a thick book where the binding makes horizontal scanning difficult. You can scan the left half of the page then the right half. The Perceive's on-line text merge capability will automatically merge the right and left halves of the page.

With unlimited horizontal and vertical scanning and your hand-held scanner, Perceive brings you the same power that was previously only available with full-page scanners.

Chapter 2 Getting Started

The intent of this chapter is to familiarize you with Perceive and get you off to a quick start.

In this chapter you will learn how to:

- Install Perceive on your hard disk
- Start Perceive
- Configure Perceive to work with your scanner
- Use on-line help
- Exit from Perceive

Installation

This section describes how to install Perceive on your computer. Installation of Perceive is very simple and will only take a few minutes.

What You Need

- IBM PC/AT or compatible (286 or higher) with:
 - 1.44 Mb 3.5 inch floppy drive
 - Hard disk (minimum 5 Mb free space)
 - Memory: 4 Mb RAM
8 Mb recommended
 - DOS 3.3 or later
 - Windows-compatible mouse and graphics adapter
- Microsoft Windows version 3.1 or higher
- Any supported scanner

Perceive supports all TWAIN-compliant scanners and most other popular handheld scanners.

What Comes with Perceive Personal

Your Perceive package should contain the following:

- Perceive Personal User Manual
- Perceive Personal Program Diskette

If anything is missing or damaged, contact your dealer immediately for replacement.

Installing Your Scanner

Follow your scanner manufacturer's instructions to install your scanner and interface card. Test the scanner using the manufacturer's software before using it with Perceive.

Installing Perceive Personal

Before installing Perceive, make sure you have installed Microsoft Windows on your computer. Perceive installs and operates from Microsoft Windows. During installation a new Windows program group will be created for Perceive.

As a safeguard against damage, make a backup copy of the Perceive diskette before installing it on your computer. Put the original in a safe place and use the backup copy when installing Perceive.

To install Perceive Personal on your hard disk, perform the following steps:

1. Start Windows in the normal manner.
2. Insert the Perceive Personal Program diskette into drive A and close the door.
3. Select **Run** from the Program Manager File menu. The Run dialog box will be displayed.
4. Type **A:\INSTALL** in the Command Line box and click **OK**.
5. Follow the directions on the screen to install Perceive.

When installation is complete, a new Perceive Personal program group is created and the Perceive Personal and Readme icons are installed in the group.

Readme File

The Readme file provides supplemental information for configuring your scanner. It also provides any new information that was not available when this manual was printed.

To open the Readme file:

1. Double-click the **Perceive Personal** program group icon to open the group.
2. Double-click the **Readme** icon to read the file.

Starting Perceive

1. Double-click the **Perceive Personal** group icon to open the group.
2. Double-click the **Perceive Personal** icon to start Perceive.

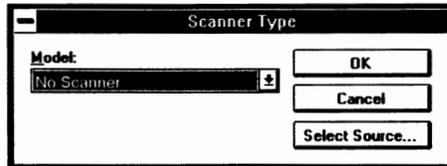


Perceive
Personal

Selecting Scanner

Perceive must be configured for your scanner before you can use it to scan documents.

1. Select **Scanner Type** from the Settings menu. The Scanner Type dialog box appears.



2. Click the scroll bar to display the list of scanners. Select your scanner from the list.

If you are using a TWAIN-compliant scanner, choose that option, then click **Select Source**. The TWAIN-compliant scanners that are installed on your computer are listed in the dialog box. Select the scanner you want to use with Perceive Personal and click **Select**.

3. Click **OK** to accept the scanner and exit from the dialog box.

If you are using a scanner that is not TWAIN-compliant:

1. Choose **Scanner Settings** from the Settings menu. The options that are available for the currently selected scanner are listed in the Scanner Setup dialog box.
 - A. If the I/O, IRQ, and DMA configuration options are listed, set them to match your scanner configuration.
 - B. If the resolution, mode and contrast (brightness or intensity) options are listed, set them for the type of document you are scanning. They should match any settings that are set by dials or switches on the scanner. See Chapter 4 for additional information.
2. Click **OK** to accept the settings and exit from the dialog box.

These settings will remain in effect until changed.

Using Help

To get help while using Perceive Personal, choose **Contents** from the Help menu and select a topic from the list. To check the Perceive Personal version number, choose **About**.

Closing Perceive

After you are finished using Perceive, choose **Exit** from the File menu to close Perceive.

Chapter 3 Tutorials

The exercises in this chapter will familiarize you with the basic features of Perceive.

- Exercise 1** You will learn how to scan and recognize the text from a page where the text is formatted into two columns. You will also learn how to save an image so it can be recognized at a later time.
- Exercise 2** You will learn how to open a previously scanned image document and recognize the text. This exercise uses the image file that you save in Exercise 1.
- Exercise 3** You will learn how to scan across a standard single column page and merge the top and bottom halves.
- Exercise 4** You will learn how to scan down a standard single column page and merge the left and right halves.

The sample documents on the next page will be used for the exercises. Carefully remove the page from the manual.

Scanner Settings

The scanner image tone, resolution and contrast (intensity or brightness) *should be* set as follows for these exercises:

Image Tone	Line Art
Resolution	300 dpi
Contrast (intensity or brightness)	Normal

The method used to select the scanner settings will vary depending on the type of scanner that you are using.

- If your scanner has dials or switches, set them for the desired settings, then select the same settings in Perceive.
- If your scanner is TWAIN-compliant, the Scanner Control dialog box opens when you click the Scan icon. The other options that are available in the dialog box will vary depending on the scanner you are using. Refer to your scanner documentation for a description of the options.
- If your scanner is *not* TWAIN-compliant, choose **Scanner Settings** from the Settings menu and select the scanner settings before clicking the Scan icon. These settings will remain in effect until changed.

OCR is a computerized process that converts optical images of text or characters into computer recognizable data. The process works in the following way: A page that has text data to be processed is scanned by an optical scanner. The image is then processed by an OCR program and a text file is produced. The processed text can then be printed, transmitted, or manipulated by a word processor.

Presently, two algorithms dominate OCR programming: matrix matching and feature extraction. Each of these algorithms has benefits and drawbacks.

Matrix Matching

Matrix matching converts each known character into a pattern within an individual matrix much like the scanner does with the entire document. The program then compares this matrix pattern to an existing index of known characters within a defined tolerance level. If no match is found, the program drops this pattern and continues with the next matrix pattern.

Matrix matching allows the program to be “trained” by providing the capability of entering user

defined font types into the index of known characters. It can, therefore, accommodate itself, through training, to an almost limitless number of font types. Although the matrix matching method provides great flexibility, it requires very large disk storage capacity and lots of main memory. It tends to be inaccurate and is slow if documents have variations in character sizes and font types.

Feature Extraction

Also called, “feature recognition,” “pattern recognition” and “topological analysis.” This method defines each character by the presence or absence of certain key features. These features include height, width, density, loops, straight line segments, stems and other characteristics. Feature extraction requires far less data to describe and identify a particular font character. Each character is scanned for these features and then compared to an existing index of known font types.

Although OCR has been around since the 1970s, it has only recently become practical and widely affordable for commercial and individual use. In large measure, this has been due to the increased performance of personal computers and the affordability of optical scanners.

Of even greater importance has been the emergence of omnifont character recognition software which can automatically read text composed of commonly used typefaces from typesetters, laser printers, typewriters and other sources. With the omnifont OCR software, you can greatly enhance office efficiency by scanning documents directly into your personal computer. Your OCR software can be used to scan magazine articles, newspaper clippings, legal contracts and many other types of documents directly into your personal computer.

The performance of the OCR software highly depends on the power of your computer, the amount of available memory and the type of document you are scanning. In most cases, an entire page of text can be read in one to four minutes. This compares to eight to ten minutes per page with a fast typist.

SAMPLE B

Exercise 1: Scanning Multiple Column Documents

In this exercise you will learn how to scan and recognize the text on a two-column page. You will manually select each column so they are recognized separately, then append the second column below the first column in the text document. You will also learn how to save an image so it can be recognized later.

Sample A will be used in this exercise.



Scan Topdown
Icon

1. Click the **Scan Topdown** icon.

If your scanner is TWAIN-compliant, the Scanner Control dialog box opens so you can select the scanner settings. Click the **Scan** button in the dialog box to initiate scanning.

Perceive will turn on the scanner's scan window light. This indicates that you can begin scanning. Refer to Appendix A if the light does not turn on.

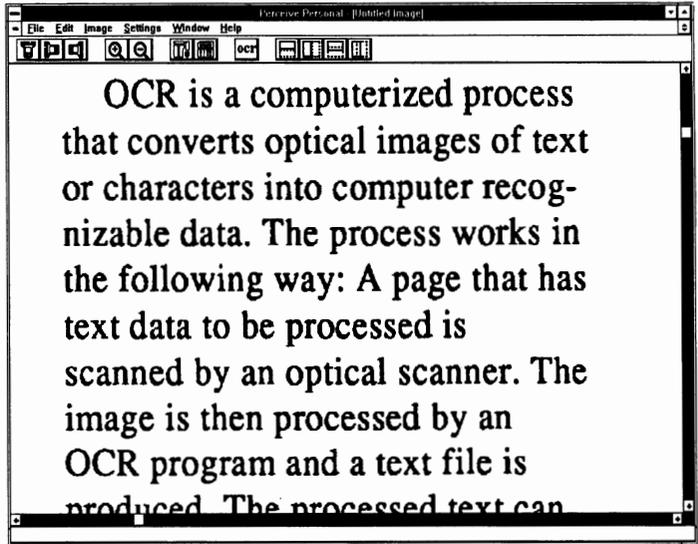
2. Place the scanner over the first column on Sample A, align the first line of text in the scan window, then move the scanner above the text. Place a book next to the scanner so you can use it as a guide when scanning. Press the Scan button and slowly move the scanner down the column.

The window that appears while scanning will vary depending on the scanner you are using. Some scanners require that you click **Done** in the dialog box when finished scanning.



Zoom In Icon

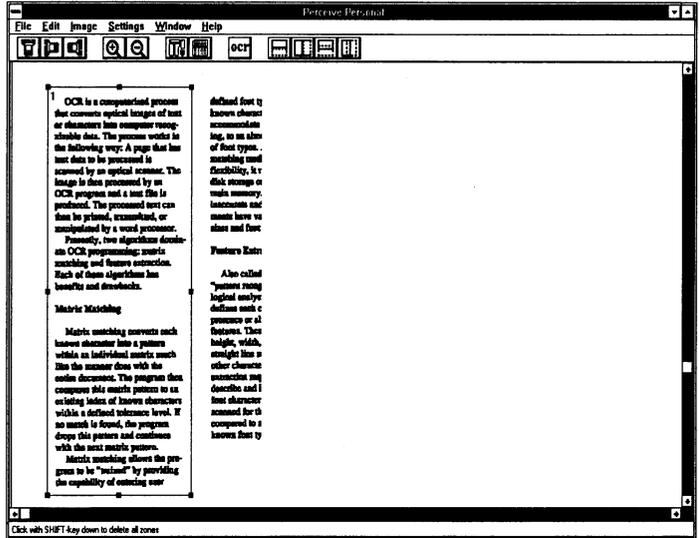
3. Click the **Zoom In** icon to magnify the image so you can check the quality of the characters. The characters should be crisp and clear. If the quality is poor, repeat steps 1 and 2 to scan the image again. The following is an example of a magnified image. Click the **Zoom In** icon as many times as required to make the text large enough to check the quality.



Zoom Out Icon

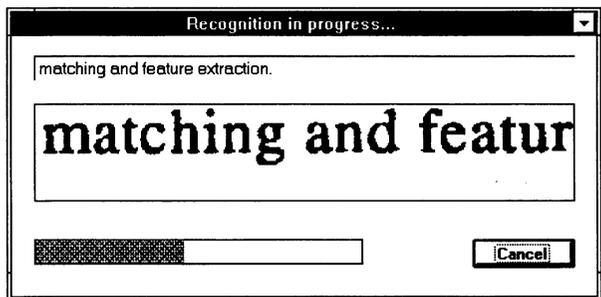
- Click the **Zoom Out** icon to return the image to the original size.
4. Choose **Save As** from the File menu to save the image so it can be used in Exercise 2. Under normal circumstances, you would only save the image if you want to recognize the text at a later time. The Save As dialog box appears.
 - A. Select **PCX** for the file format.
 - B. Type **IMAGE** as the file name. The .PCX extension is added automatically.
 - C. Select the directory where you want the file located. The Perceive directory is the default. Save all of the exercise files to the same directory so they are easy to locate and delete when finished.
 - D. Click **OK** to save the image.

- Place the pointer at the top of the first column and drag the mouse to draw the zone box around the first column of text.



OCR Icon

- Click the **OCR** icon to recognize the text. The Recognition in progress... window appears.



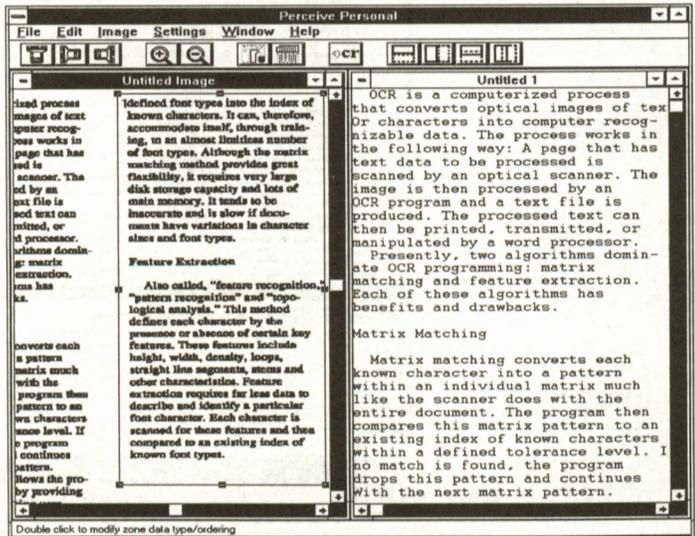
The recognized text is shown at the top of the window with portions of the scanned image below. The text appears in an untitled document when recognition is complete.

The Cancel button cancels recognition before it is complete.

☆ *The image document is minimized to an icon. You can activate the image document if you need to verify any of the text.*

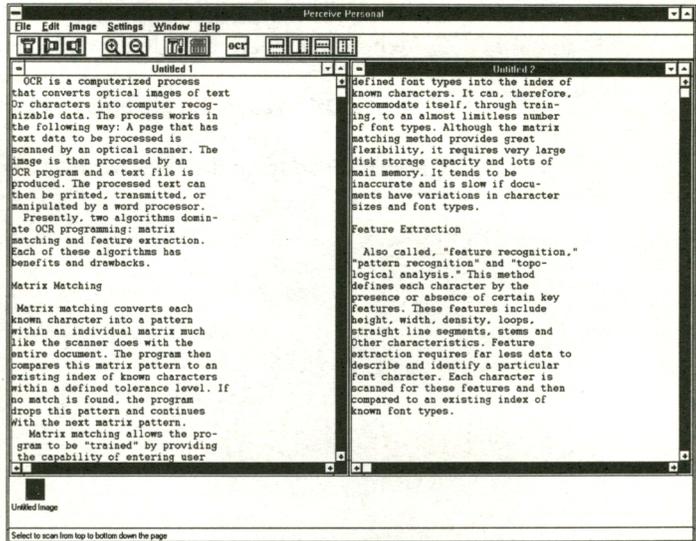
7. Scan the second column as in steps 1 through 5. The previous image is discarded and the new image appears in the image document.

The image and text document are both open.



☆ *The Window menu Cascade and Tile commands can be used to change how the two documents are displayed.*

8. Click the **OCR** icon to recognize the text. When recognition is complete, the text appears in a new untitled document. The two untitled documents appear as shown in the following example.



Append Icon

9. Click the **Append** icon. This appends the Untitled 2 document text below the previous text in the Untitled 1 document.

The text can now be edited and saved.

- The keyboard can be used to add or delete text, much as you would in a word processor.
 - The Edit menu lets you clear, move and copy selected portions of text. It can also be used to locate and replace unrecognized characters. Unrecognized characters are entered as ~ signs, unless a different character is selected in Preferences.
10. Choose **Save** from the File menu. The Save As dialog box appears.
 - A. Select the file format that you want to use from the list of file types on the dialog box.

- B. Type **EXAMPLE** as the file name and select the directory where you want the file located.
 - C. Click **OK** to save the file and exit from the dialog box.
11. Choose **Close** from the File menu to close the text document.

Exercise 2: Opening a Previously Scanned Image File

In this exercise you will learn how to open a previously saved image file so you can recognize the text. This feature can be used to recognize the text on scanned images at a more convenient time or to recognize the text on fax files.

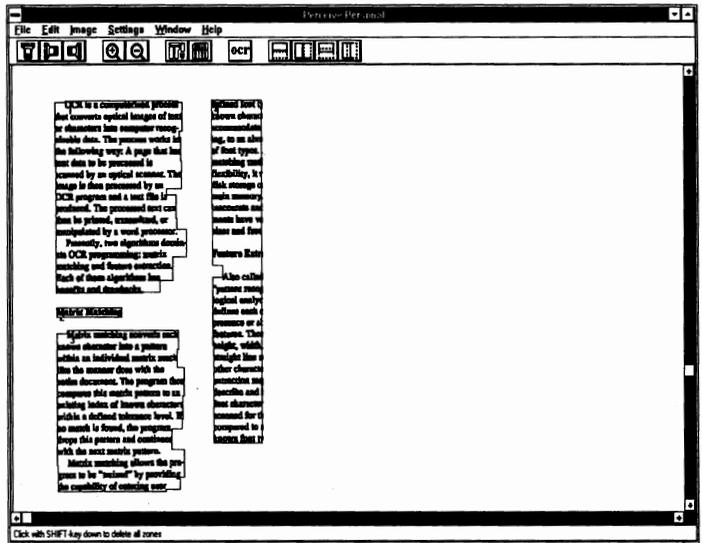
You will also learn how to automatically select the recognition zones.

The image you saved in Exercise 1 will be used in this exercise.

1. Choose **Open** from the File menu. The Open dialog box appears so you can select the image file that you want to recognize.
 - A. Select the directory where the **IMAGE.PCX** file from Exercise 1 is located.
 - B. Select **PCX** from the list of file types. A list of PCX files appears in the Files list.
 - C. Select the **IMAGE.PCX** file. The selected file will appear in the Filename box.
 - D. Click **OK** to open the image.
2. Click the **Multiple Columns** icon to automatically draw zone boxes around the text. A number appears in the upper left corner of the zones. This is the order in which the zones will be recognized.



Multiple Columns
Icon



Click any zones that you don't want to recognize and press the **Del** key to remove the zone box. In the above example you would delete zone 4.



OCR Icon

3. Click the **OCR** icon. When recognition is complete, the text appears in an untitled document.
4. Choose **Close** from the File menu. There is no need to save the text to a file.

Exercise 3: Scanning Across Single Column Documents

In this exercise you will scan across the text in Sample B in two passes and merge the two halves. This is the easiest way to scan single column documents.



Scan to Right
Icon

1. Click the **Scan to Right** icon. Perceive will turn on the scanner's scan window.

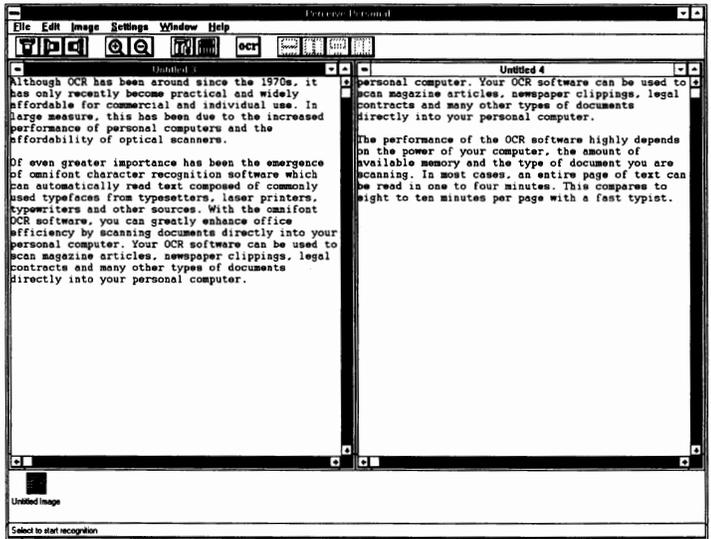
If the scanner is TWAIN-compliant, the Scanner Control dialog box opens so you can select the scanner settings.

2. Place the scanner so you can scan from left to right across the first two paragraphs. Align the first character of each line in the scan window then move the scanner above the text. Place a book next to the scanner so you can use it as a guide when scanning. Press the Scan button and slowly move the scanner across the paragraphs.
3. Perceive will automatically rotate the image clockwise so the text can be recognized from left to right.
4. Check the quality of the scanned image. If the quality is poor, repeat steps 1 through 3 to scan the image again.
5. Click the **OCR** icon to recognize the text. It is not necessary to draw zones around the text since any overlap will be deleted when you merge the two halves.
6. After the text appears in the untitled document, scan the third paragraph as in steps 1 through 5. The text appears in a new untitled document.



OCR Icon

Notice that the last four lines of the first document are the same as the first four lines in the current document. These overlapped lines will be merged when the two documents are merged. Your overlap may be different than the example. There should be a minimum overlap of two or three lines.



Merge Icon

7. Click the **Merge** icon to merge the current text below the previous text and automatically delete any overlap.
8. Choose **Close** from the File menu. There is no need to save the text to a file.

Exercise 4: Scanning Down Single Column Documents

In this exercise you will scan down Sample B in two passes, then merge the two halves together. It may take a little practice with the scanner to successfully use this method.



Scan Topdown
Icon

1. Click the **Scan Topdown** icon. Perceive will turn on the scanner's scan window light.

If the scanner is TWAIN-compliant, the Scanner Control dialog box opens so you can select the scanner settings.

2. Place the scanner over the left half of Sample B, align the first line of text in the scan window then move the scanner above the text. Place a book next to the scanner so you can use it as a guide when scanning. Press the Scan button and slowly move the scanner down the page.
3. Check the quality of the scanned image. If the quality is poor, repeat steps 1 and 2 to scan the image again.
4. Click the **OCR** icon to recognize the text.
5. When the text appears in an untitled document, repeat steps 1 through 4 to scan the right half of the page. There should be about 1/2 inch overlap between the first scan and the second.

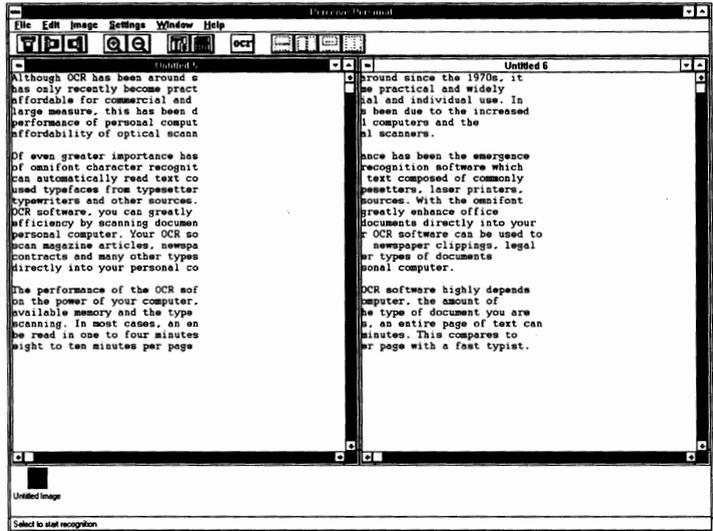


OCR Icon

Notice the overlap of text between the previous and current text in the example on the following page. This overlap and any misread characters will be deleted when you merge the text.

6. Check to be sure there is at least a 5 character overlap between the current and previous halves of the page.

If there is less than a 5 character overlap, repeat steps 1 through 4 to scan the right half again.



Merge to Right
Icon

- Click the **Merge to Right** icon and the two halves will merge into the first untitled document.

If the text is not merged correctly, you can choose **Undo** from the Edit menu and rescan the right half.

- Choose **Close** from the File menu. There is no need to save the text to a file.

CONGRATULATIONS! you have learned the basics for using Perceive and are now ready to start scanning your documents. Since the accuracy of text recognition is dependent on the quality of the scanned image, be sure to review the tips in Chapter 4.

Chapter 4 Using Your Scanner

Perceive recognizes the text from the scanned image. The accuracy of Perceive can only be as good as the quality of the scanned image. Therefore it is very important that you take care when scanning the page.

The Scanner Controls

Most handheld scanners have dials and switches that are used to select the scanner settings. Perceive should be set to the same settings as the scanner. If your scanner doesn't have dials or switches, the scanner settings are selected in Perceive.

The following scanner controls will be used when scanning into Perceive. Your scanner manual provides detailed information for using the scanner buttons and controls.

Scanning Window

The scanning window lets you view text before scanning so it can be properly aligned. After aligning the first line of text in the scanning window, move the scanner above the text before starting to scan.

The light in the scanning window indicates that scanning has been initiated from Perceive. The light will turn off when you finish scanning.

Scan Button

Depending on the scanner you are using:

- Press and release the Scan button to start scanning.
- Press and hold the Scan button while scanning.

**Contrast
Dial**

The contrast dial adjusts the amount of light that the scanner uses to scan the page. It can be used to improve OCR accuracy when scanning light or dark text. See "Adjusting Contrast" for additional information.

**Image Tone
Control**

The Line Art setting should be used for most documents. This will produce a high-contrast black and white image that can be recognized.

Use the Grayscale setting to improve recognition accuracy when scanning pages with shaded or colored documents, newspapers, light or thin characters, and glossy paper.

Set the Contrast to normal or medium when using the Grayscale setting

**Resolution
Control**

In general, the higher the resolution, the better the quality of the scanned image.

For optimum accuracy, use 300 dpi for 12 point or larger type and 400 dpi for 10 point or smaller type.

Scanning Tips

It is very important that you scan in a straight line at a SLOW and SMOOTH speed.

Take a few minutes to practice using the scanner before using it with Perceive.

1. Place the page on a flat surface where you have enough room to scan the entire page. Position the page so you can comfortably move the scanner over it.

If you are scanning a magazine that is bound using center staples, fold the magazine to avoid snagging the scanner rollers on the staples.

2. Place the scanner over the text you want to recognize. You can scan the text from top to bottom or from left to right. You will scan the page in two or three passes. There should be some overlap between passes. Perceive will assemble the text after it is scanned.
3. Position the scanner so the first line of text is visible through the scanning window.
4. Place the spine of a book next to the scanner so it can be used as a guide when moving the scanner.
5. Move the scanner above the first line of text so it is not visible in the scanning window.
6. Move the scanner at a SLOW, SMOOTH speed. A good speed is $\frac{1}{2}$ inch per second. Try not to jerk or start and stop when moving the scanner.

If you vary the speed or jerk when scanning, the characters will be stretched or compressed on the scanned image and cannot be recognized accurately.

OCR is a
that converts
or characters
nizable data.
the following

Normal

OCR is a
that converts
or characters
nizable data.
the following
text data to b

Degraded

If you do not scan in a straight line, the text will be skewed at an angle. It is almost impossible to hand scan without some skew and Perceive will compensate for a normal amount of skew. Perceive's Auto Deskewing feature automatically straightens skewed text, but if there is an excessive amount of skew the characters may be distorted. The Auto Deskewing feature can be turned on and off in Preferences.

OCR is a computer
that converts optical im
or characters into comp
nizable data. The proce
the following way: A p
text data to be processe
scanned by an optical s
image is then processe
OCR program and a tes

Normal

OCR is a computer
that converts optical im
or characters into comp
nizable data. The proce
the following way: A p
text data to be processe
scanned by an optical
image is then processe

Excessively Skewed

Judging Image Quality

When you look at the page, you can usually tell if the text is clear and easy to read, or if it is degraded and more difficult to read. Subtle differences may be more difficult to see. It goes without saying that seriously degraded text cannot be recognized accurately.

The scanner and OCR software are not nearly as sophisticated as the human eye, so the best way to tell the quality of the text is to look at the scanned image in Perceive. You can use the Zoom-in icon to magnify the image so you can see the characters in detail. Characters that are completely formed and separated from other characters as illustrated are the easiest to recognize.

OCR is a computerized

☆ *The image reproduced in your computer will not be as clear as the original document. This degradation, combined with the effects of multi-generation photocopies and image "noise" such as stray markings and smudges, can make the job of character recognition more difficult.*

Adjusting Contrast

As you develop experience in using Perceive, you will find that accuracy can be improved by adjusting the contrast setting, which will accentuate or de-emphasize certain text attributes, depending on the document scanned.

You will need to experiment a little to determine the best setting to use. The following are a few guidelines:

Light or Thin Characters Dark setting

Dark or Thick Characters Light setting

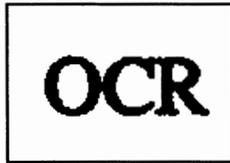
Glossy Paper Dark setting

Touching Characters

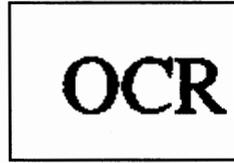
When characters touch, they are difficult to isolate and may be misread.

To improve the quality of these characters, you can try using a lighter contrast setting. This will make the characters lighter so they are easier to separate, but may make the text too light to recognize accurately.

The following are examples of how the scanner contrast setting can improve the quality of some characters.



Normal Contrast



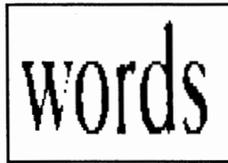
Lighter Contrast

Broken Characters

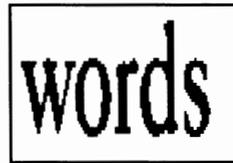
The characters may be broken when scanning light text, glossy pages such as magazines, or thin-stroke fonts.

Change the scanner contrast setting darker or scan in Grayscale mode to see if recognition improves.

The following are examples of how the scanner contrast setting can improve the quality of broken characters.



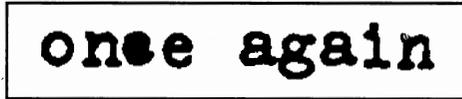
Normal Contrast



Darker Contrast

Filled-in or Blurred Characters

The features of the characters cannot be distinguished when they are thick or blurred. The centers fill and they lose their individual characteristics. Typewriter keys that get dirty will fill-in characters.



onee again

To improve the quality of these characters, you can try using a lighter contrast setting. This will make the characters lighter so they are easier to separate, but may make the text too light to recognize accurately.

The following are examples of how the scanner contrast setting can improve the quality of blurred characters.



Kinetic

Normal Contrast



Kinetic

Lighter Contrast

Other Factors

Commonly Misread Characters

Characters that are very similar may be misread. For example, the numbers 1 and 0 and the letters l and O look almost identical in some fonts.

Small characters such as the period, comma, apostrophe and hyphen are also easily misread. This is a common problem when recognizing small point sizes, because these characters get very small.

When editing the text, use the Edit menu Find or Replace command to locate and correct any misread characters.

Point Sizes

Perceive can recognize normal, bold and italic text in the following point sizes:

200 dpi	12 to 48 points
300 dpi	8 to 36 points
400 dpi	6 to 30 points

This is the normal size range for most documents. Anything smaller than 6 points is “fine-print” that is difficult to read without a magnifying glass. If you need to recognize text that is smaller than 6 points, try making an enlarged photocopy to scan. Text larger than 36 points can be entered from the keyboard after the text is recognized.

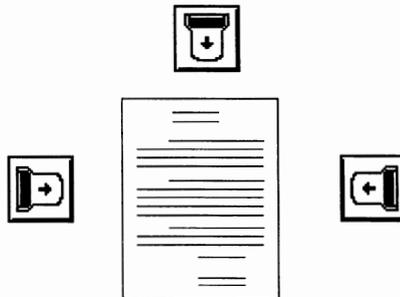
Scanning Different Document Types

The following are examples of the most common document types and the options to select when scanning them.

Single Column

Text that reads across the full width of the page is considered a single column. Most letters and reports are printed in this manner.

This type of page can be scanned either down or across. Scanning across is the easiest because you scan across the paragraphs. When you scan down the page in two passes, you can merge the two halves together.



Multiple Columns

Paragraphs of text that are formatted into columns on the page are considered multiple columns. Newspaper and magazine articles are formatted in this manner.

Scan down each of the columns.



Tables

It is easiest to scan across tables. If you scan down the table, use the Append to the Right icon to append the second pass.



Perceive will read across the table and insert spaces to separate the columns. Select all of the columns in the same zone. **DO NOT** select each column of the table separately.

Save the text in ANSI, ANSI spreadsheet, ASCII or ASCII spreadsheet file format to be sure a carriage return is entered at the end of each line. The ANSI and ASCII spreadsheet formats replace the spaces between columns with tabs so the files can be used with spreadsheet and database applications.

Text Mixed with Graphics

Perceive will automatically skip over graphics on the page and recognize only the text. However, it does take longer to search through the graphic looking for text. To speed recognition, use zones to select only the text portion of the page.



Legal

Legal documents are frequently printed on pleading paper that has line numbers printed down the side. Select the zones so the line numbers are not recognized.

Dot Matrix

There are two types of dot matrix printing: near-letter-quality (NLQ) and draft-quality. Both types can be recognized.

- The NLQ characters look like typewriter characters and can be recognized in standard mode.
- Draft-quality characters are printed in a dot pattern. You will need to select the Draft-mode Dot Matrix option in the Preferences dialog box to read draft-quality dot matrix. Be sure to deselect Draft-mode Dot Matrix when finished.

Since these printers use cloth ribbons, the print is usually very light, which makes the text more difficult to recognize.

To compensate for the dot pattern and light printing, set the scanner contrast for dark. Change it back to medium when finished scanning dot matrix documents. Another option would be to use a photocopier to darken the text.

Chapter 5 Command Summary

This chapter provides a brief description of each of the toolbar icons and the menu commands. The menus are listed in the order in which they appear in the menu bar.

Toolbar

The toolbar contains the icons that are used to scan, recognize the text, select the recognition zones, and merge or append the current text to the previous document.

Scanning Icons

The following icons determine the scanning direction and initiate scanning. There is a corresponding Image menu command for each icon.

If your scanner is TWAIN-compliant, the Scanner Control dialog box appears when you click any of the scanning icons. Select the resolution, contrast (intensity or brightness), and image tone that you want to use. If your scanner has switches or dials, the settings that are selected in the dialog box should match the settings on your scanner. The other options that are available in the dialog box will vary depending on the scanner you are using. Refer to your scanner documentation for a description of the options.



Scan Topdown Icon

Click this icon to scan from the top to the bottom of the page. To scan, press the Scan button and slowly move the scanner down the page.



Scan to Right Icon

Click this icon to scan from left to right across the page. To scan, press the Scan button and slowly move the scanner across the page. Perceive will automatically rotate the image clockwise so the text can be recognized.



Scan to Left Icon

Click this icon to scan from right to left across the page. To scan, press the Scan button and slowly move the scanner across the page. Perceive will automatically rotate the image counter clockwise so the text can be recognized.

Zone Icons

The Multiple Columns or Single Column icons select the zones automatically. The zones can also be selected manually.

The purpose for selecting zones is to:

- recognize only portions of the text in the image
- select columns of text
- change the recognition order of the zones

☆ *If no zones are selected, the text will be recognized from left to right across the entire image.*



Multiple Columns Icon

Automatically discerns the text areas and draws zone boxes around the text. Multiple columns, such as magazine articles, will be segmented into separate boxes. This icon performs the same function as the corresponding Image menu command.



Single Column Icon

Automatically segments the full width of the text into zones. Use for tables and similar documents where you don't want the text segmented into columns. This icon performs the same function as the corresponding Image menu command.

Selecting Zones Manually

Drag the mouse to draw a zone box around the text that you want to recognize. Repeat to draw additional boxes. The number in the corner of the zone indicates the recognition order.

To change the size of a manually drawn zone box:

Click inside the zone. Eight handles appear around the box. Drag the handles to resize the box. Click inside the zone to remove the handles.

To move a manually drawn zone:

Place the pointer inside the zone and drag to the new location.

Deleting Zones

To delete all zones:

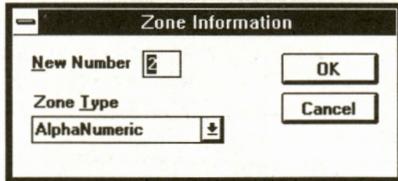
Hold the Shift key and click anywhere outside the boxes.

To delete only one zone:

Place the pointer inside the zone that you want to delete, then press the **Del** key or hold the Shift key and click.

Changing the Order

Double-click inside the region that you want to change. The Zone Information dialog box appears.



To change the recognition order:

Type the new number in the dialog box and click **OK**. The regions are automatically renumbered to reflect the change.

Zoom Icons

The zoom icons enlarge and reduce the size of the scanned image. There is a corresponding Image menu command for each icon.



Zoom In Icon

Magnifies the image one level each time the icon is clicked, until the highest level is reached. Use to get a closer view when selecting text and to examine the quality of the scanned image.



Zoom Out Icon

Reduces the image one level each time the icon is clicked, until the lowest level is reached.

OCR Icon



Click the OCR icon to start recognition of the text on the image. If needed, select the zones before clicking the OCR icon. If no zones are selected, the full width of the entire image will be recognized. The OCR icon performs the same function as the Image menu Recognize command.

Append/Merge Icons

The following icons determine how the current text is combined with the previous text. There is a corresponding Edit menu command for each icon.



Append Icon

Appends the current text to the end of previous text, without deleting any overlap. Also used for the first scan on a new page.



Append to Right Icon

Appends the current text to the right of previous text, without deleting any overlap. Use for tables and similar type documents.



Merge Icon

Merges the current and previous text, deleting any overlap. There should be a two or three line minimum overlap between the current and previous text. Use when scanning across a page in two or more passes.



Merge to Right Icon

Merges the current and previous text, deleting any overlap. There should be a minimum overlap of 5 characters between the current and previous text. Use when scanning down a page in two or more passes.

File Menu

File	
Open...	Ctrl+O
Close	
Save	Ctrl+S
Save As...	
Exit	Alt+F4

The File menu is used to open, save and close image and text files and to close the Perceive application.

Open...

Opens an image file so the text can be recognized. The file format and name are selected from the dialog box.

Close

Closes the active text or image document.

Save

Use to save new text files and to save changes to an existing file. The Save As dialog box will be displayed the first time a document is saved. Select the file format before entering the file name. Change the directory if you want the file located in a different directory.

Save As...

Use to assign a new file name or to change the file format. Also used to save images. The Save As dialog box appears. Select the file format before entering the file name. Change the directory if you want the file located in a different directory.

Text File Formats

Select a file format that can be used by the application where you plan to use the text. The ANSI formats can be used with other Windows applications. The ASCII formats can be used with DOS or Windows applications.

ANSI Inserts a carriage return at the end of each line. Use for tables and lists that you do not want wrapped into paragraphs.

ANSI Paragraph Formats text into paragraphs.

ANSI Spreadsheet	Replaces the spaces between columns with a tab and inserts a carriage return at the end of each line. Use for tables and lists that you want to open in spreadsheet or database applications.
ASCII	Inserts a carriage return at the end of each line. Use for tables and lists that you do not want wrapped into paragraphs.
ASCII Paragraph	Formats text into paragraphs.
ASCII Spreadsheet	Replaces the spaces between columns with a tab and inserts a carriage return at the end of each line. Use for tables and lists that you want to open in spreadsheet or database applications.

Image File Formats

Select a file format that can be used by the application where you plan to use the image. Any of the three file formats can be used to save files that you want to open in Perceive.

TIFF Group 1	TIFF files require a large amount of hard disk space.
Compressed TIFF	Uses about one-third less disk space than TIFF format. It will take longer to save and open compressed TIFF files.
PCX	This is the PC Paintbrush format. It requires less disk space than TIFF files.

Exit

Closes the Perceive application.

Edit Menu

Edit	
Undo	Ctrl+Z
Cut	Ctrl+X
Copy	Ctrl+C
Paste	Ctrl+V
Clear	
Append	Ctrl+A
Append to right	Ctrl+B
Merge	Ctrl+M
Merge to right	Ctrl+N
Find...	Ctrl+F
Replace...	Ctrl+R

The Edit menu is used to edit text. The Append, Append to Right, Merge, and Merge to Right commands perform the same functions as the Toolbar append and merge icons.

Undo

Cancels the append or merge so you can rescan the current text.

Cut

Removes selected text and places it in the Windows clipboard.

Copy

Places a copy of the selected text in the Windows clipboard.

Paste

Inserts the text from the clipboard at the cursor position.

Clear

Deletes the selected text from the document.

Find...

Use to search for a character, word or phrase. You can then replace it from the keyboard.

Replace...

Use to search and replace a character, word or phrase.

Image Menu

Image	
Scan <u>T</u> opdown	F2
Scan to <u>R</u> ight	F3
Scan to <u>L</u> eft	F4
<u>R</u> ecognize	F9
Zoom <u>I</u> n	F5
Zoom <u>O</u> t	F6
<u>M</u> ultiple columns	F7
<u>S</u> ingle column	F8

The Image menu commands perform the same functions as the corresponding Toolbar icons. Refer to the “Toolbar” section for details.

☆ *The Recognize command performs the same function as the OCR icon.*

Settings Menu

Settings	
Scanner <u>T</u> ype...	
Scanner <u>S</u> ettings...	
<u>P</u> references...	

The Settings menu is used to select your scanner, change the scanners settings, and to set preferences.

Scanner Type...

Opens the Scanner Type dialog box so you can set Perceive to work with your scanner. Select your scanner and click **OK**.

If you are using a TWAIN-compliant scanner, choose that option, then click **Select Source**. The TWAIN-compliant scanners that are installed on your computer are listed in the dialog box. Select the scanner you want to use and click **Select**.

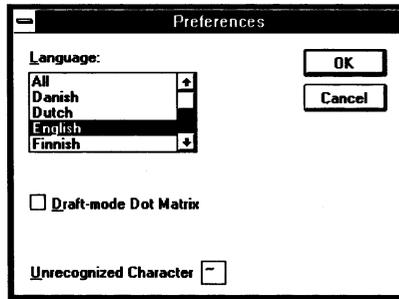
Scanner Settings...

Opens the Scanner Settings dialog box so you can set the scanning options for scanners that are not TWAIN-compliant. The options that are available will vary depending on the scanner you are using. If the I/O, IRQ, and DMA options appear, set them to match how your scanner is configured. Set the resolution, contrast (intensity or brightness), and image tone mode that you want to use. If your scanner has switches or dials, the settings that are selected in the dialog box should match the settings on your scanner.

The scanning size controls the area that will be scanned. The height should be set for the length of the page you are scanning. The width should normally be set for the width of the scanner. Scanning will stop automatically when the scan height is reached.

Preferences...

The Preferences dialog box lets you customize Perceive for specific types of documents.

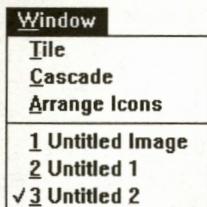


Language: Select the language to be recognized. The current selection is highlighted in the list. Select **All** to recognize documents that are printed in two or more languages.

Draft-mode Dot Matrix: Select this option to recognize 9-pin draft-quality dot matrix print. Perceive can only recognize draft dot matrix documents when this option is selected. Deselect the Draft-mode Dot Matrix option to recognize other types of documents.

Unrecognized Character: If you wish, you can change the character that replaces characters that Perceive cannot recognize. The default is ~. Recommended characters are: ~, #, @, *, and ?.

Window Menu



The Window menu commands let you change how the open documents are displayed and switch between documents.

Tile

Reduces the size of the open documents so they all fit in the window at the same time. After resizing a window, choose **Tile** again to fit the documents to the new size.

Cascade

Cascades the open documents. To switch between documents, click the title bar.

Arrange Icons

Arranges the minimized icons at the bottom of the window if they have been moved to different locations.

Open Documents

The open documents are listed at the bottom of the menu. Choose the document you want to be active.

Help Menu

Help
Contents F1
About...

The Help menu is used to access on-line help and to get information about Perceive.

Contents

The Contents command lets you get help while using Perceive. Select the topic you want help with from the list.

About...

Provides the Perceive version number.

Keyboard Shortcuts

The following keys perform the same function as clicking a Toolbar icon.

Scan Topdown	F2
Scan to Right	F3
Scan to Left	F4
OCR	F9
Zoom In	F5
Zoom Out	F6
Multiple Columns	F7
Single Column	F8
Append	Ctrl+A
Append to right	Ctrl+B
Merge	Ctrl+M
Merge to right	Ctrl+N

Appendix A Troubleshooting

This appendix provides information on common problems and solutions. The most common problems are communicating with the scanner or Microsoft Windows related problems.

Before you call for assistance, you may want to remove all of the memory-resident programs and unwanted drivers from your AUTOEXEC.BAT and CONFIG.SYS files. You may also want to remove other I/O cards such as a network or a fax card from your system and try to solve the problem on your own.

The light in the scanning window does not turn on when trying to scan

Possible causes:

1. The interface card is not firmly seated in the expansion slot of your system.
2. There is an I/O port address conflict with another device on the system.
3. Your scanner is not firmly connected to the interface card.
4. The connector or the cable on your scanner may be damaged.

Scanner not ready

This message appears when Perceive cannot communicate with your scanner.

Possible causes:

1. You may have selected the wrong scanner from the Scanner Type dialog box. Choose the **Scanner Type** command on the Settings menu and make sure your scanner is selected.
2. The scanner may not be powered-on or connected properly. Check to be sure the scanner cable is connected to the scanner interface card. It may have been accidentally connected to another card or to the serial/parallel port.

3. You may be using a scanner model that is not supported.

Make sure that your scanner works with the image scanning software that came with your scanner before calling for assistance.

Scanning or recognition is slow

Possible causes:

1. You may not have enough memory on your system. You may want to add extra memory on your system. Also, make sure that you have a minimum of 5 Mb free space on your hard disk.
2. You may be using a 6-8 MHz system or the speed of your hard disk is very slow.
3. Some I/O cards such as network cards may be conflicting with the scanner interface card. You may want to remove the network card from your system and try again.
4. Insure that the quality of the scanned image from the scanner reflects the original document. Click the Zoom In icon to get a closer look at the image quality. If the characters have a pattern such as bar code, you may have a problem with your scanner.
5. The version of Microsoft Windows that you are using may be corrupted or damaged.

Re-install Windows and try again. Or, get a new copy of Windows from your dealer.

6. Your document has too many touching characters. Change the scanner contrast dial to a lighter setting to improve the recognition speed.

There is no image after scanning

Possible causes:

1. The contrast is set too light or dark. Try setting the contrast to medium and rescan your document.

2. There is a DMA channel conflict on the system.
3. Test your scanner with the image scanning software that came with your scanner to be sure it is operating properly.

Perceive makes too many errors

Review the suggestions for improving accuracy in Chapter 4. There are, however, certain documents, such as a fourth generation photocopy, that are too poor quality to be recognized accurately.

Possible causes:

1. The characters are distorted. Rescan the page using a book to guide the scanner in a straight line and move the scanner at a smooth, slow speed.

2. Your document has too many broken characters.

To solve this problem, you can:

- A. Change the contrast dial on your scanner to a darker setting to improve recognition.
 - B. Make a photocopy to darken the characters before scanning.
3. Your document has too many filled-in or touching characters. To solve this problem, you can try using a lighter contrast setting on your scanner.
 4. Your document has smudges and other marks on it.
 5. The spacing between two lines in your document is very tight and there is no clear separation between them. Or, the text in your document is 6 point size or below.
 6. Documents with bold or italic characters below 10 point size will result in poor recognition accuracy.

Perceive does not recognize anything from the image

Possible causes:

1. The scanned image is totally black or has a black background. To check the image quality, click the Zoom In icon to get a closer look at the image.
2. Test your scanner with the image scanning software that came with your scanner to be sure it is operating properly.

Glossary

ANSI character set. American National Standards Institute 8-bit coded characters.

ASCII character set. American Standard Code for Information Interchange 8-bit coded characters.

Bit Map. A collection of bits, or dots, that make up the scanned image.

Character. A letter, number, punctuation or symbol.

DPI. Dots Per Inch. Used to measure resolution.

Font. All of the characters in one type face. Within a font family there can be normal, bold, italic and other variations of the same typeface.

Ligatures. Two or more characters linked together, as in cursive writing.

Line Art. Black and white graphic without shading or gray tones.

NLQ. Near Letter Quality. A dot matrix printing mode where the characters resemble typewritten, or letter-quality, characters.

Noise. Unidentifiable marks on image that are picked up when the page is scanned.

OCR. Optical Character Recognition. The process by which the text in an image is recognized and converted to coded characters.

Omnifont. An OCR process that is pre-programmed to read many different fonts and point sizes.

PCX. PC Paintbrush File Format. Can be used with most graphic applications.

Point Size. A measurement of type sizes. Each point measures .0138 inch. There are approximately 72 points in an inch.

Resolution. The density of dots on a page or screen, measured in dots per inch. The greater the resolution, the clearer the image.

Scanner. An input device used to convert the text and graphics on a printed page into an electronic file.

TIFF. Tagged Image File Format. An industry standard for storing halftone, line art and gray-scale images.

TWAIN. An application programming interface (API) that provides a standard method for acquiring images in a Windows application.

Window. The area of a screen that contains image or text information within the Microsoft Windows environment.

Index

A

ANSI 1-3, G-1
Append icon 3-7, 5-5
Append to Right icon 5-5
ASCII 1-3, G-1
Automatic page segmentation 1-2, 3-9, 5-2

C

Cancel recognition 3-6
Changing zone recognition order 5-3
Closing Perceive 2-4
Closing text document 3-8
Colored documents 4-2
Contrast setting 1-2, 4-2, 4-5, 4-10

D

Dark characters 4-5
Databases 4-9, 5-7
Degraded text 4-5
Deleting zones 3-10, 5-3
Dot matrix documents 4-10, 5-11

E

Edit menu 3-7, 5-8
 Append 5-8
 Append to right 5-8
 Clear 5-8
 Copy 5-8
 Cut 5-8
 Find 4-8, 5-8
 Merge 5-8
 Merge to right 5-8
 Paste 5-8
 Replace 4-8, 5-8
 Undo 3-14, 5-8
Editing text 3-7, 5-8

F

Features 1-2
File formats 4-9, 5-6
File menu
 Close 3-8, 5-6
 Exit 2-4, 5-7
 Open 3-9, 5-6
 Save 3-7, 5-6
 Save As 3-4, 5-6
Filled-in characters 4-6

H

Hardware requirements 2-1
Help menu 5-12
 About 2-4, 5-12
 Contents 2-4, 5-12

I

Image document 3-6
Image file formats 1-3, 5-7
Image menu 5-9
Image quality 4-3, 4-5
Installation 2-1

L

Language selection 5-10
Legal documents 4-10
Light characters 4-5
Line art mode 4-2

M

Memory requirement 2-1
Merge icon 3-12, 5-5
Merge to Right icon 3-14, 5-5
Microsoft Windows 1-1, 2-2, A-2
Misread characters 4-7
Multiple column documents 1-4, 3-3, 4-9
Multiple Columns icon 3-9, 5-2

O

OCR icon 3-5, 3-13, 5-5
Omnifont 1-1, G-1
Opening image files 3-9, 5-6

P

Package contents 2-2
Point sizes 4-8, G-1
Preferences dialog box 5-10

R

Readme file 2-3
Recognition in progress window 3-5
Resolution setting 4-2, 4-8

S

Saving image 3-4
Saving text 3-7, 5-6
Scan to Left icon 5-2
Scan to Right icon 3-11, 5-2
Scan Topdown icon 3-3, 3-13, 5-1
Scanner 2-2, 4-1
Scanner settings 4-1

- non TWAIN-compliant scanners 2-4, 3-2, 5-9
- scanner dials and switches 3-2, 4-1
- TWAIN-compliant scanners 3-2, 3-3, 5-1

Scanner Type dialog box 2-3
Scanning bound books 1-4, 4-3
Scanning glossy paper 4-2, 4-5
Scanning magazines 1-4, 4-3, 4-9, 4-11
Scanning newspapers 4-2, 4-9
Scanning tips 4-1, 4-3
Selecting zones 5-2

- automatically 3-9, 5-2
- manually 3-5, 5-3

Settings menu 5-9

- Preferences 3-3, 4-4, 4-10, 5-10
- Scanner Settings 2-4, 5-9
- Scanner Type 2-3, 5-9

Single column documents 1-4, 3-11, 3-13, 4-8
Single Column icon 5-2
Software requirements 2-1
Speed 1-3

Spreadsheets 4-9, 5-7
Starting Perceive 2-3
Stretched or compressed text 4-3

T

Tables 4-9
Text file formats 1-3, 4-9, 5-6
Text with graphics pages 4-10
Toolbar 5-1

- Keyboard shortcuts 5-12

Troubleshooting A-1
Tutorials 3-1

U

Undo merge 3-14, 5-8
Unrecognized character 3-7, 5-11
Untitled image document 3-6
Untitled text document 3-5, 3-6
Using your scanner 4-1

W

Window menu 5-11

- Arrange Icons 5-11
- Cascade 3-6, 5-11
- Tile 3-6, 5-11

Z

Zone Information dialog box 5-3
Zoom In icon 3-3, 4-5, 5-4
Zoom Out icon 3-4, 5-4



OCRON, INC.

OCRON INC. TM
PERCEIVE PERSONAL
S/N 0348964

81-100-01940 REV. 2.1