



Sleuth1

Optical Character Recognition for RISC OS

ProAction

3QD
DEVELOPMENTS LTD

APDL

Sleuth1

© 3QD Developments Ltd 2026
All rights reserved.

No part of this product may be reproduced in whole or part by any means without written permission of the publisher. Unauthorised hiring, renting, lending, public performance or broadcasting of this product or its parts is prohibited.

While every care is taken, the publisher cannot be held responsible for any errors in this product, or for loss of data or consequential effects from the use of this package.

Software and Documentation supplied by
APDL
a part of
3QD Developments Ltd
Single User version downloaded from the free of charge software section of
www.riscos.com

Please note that any free of charge downloads from RISCOS.com are supplied without any technical support package as per the licence agreement.

This document last edited on 29/09/26



If you have found this application and this manual useful please consider making a donation to help cover the costs of cataloguing, documenting, updating and hosting the APDL applications. Scanning the QR code will take you to our donations page.

Sleuth OCR software for RISC OS

Introduction

OCR is an acronym for Optical Character Recognition. This is a process whereby text is extracted from a scanned image and converted into a plain text file. OCR has been available for some time on other machine platforms and accuracy has been steadily improving. Sleuth 2 can be used with the majority of scanners available for RISC OS computers.

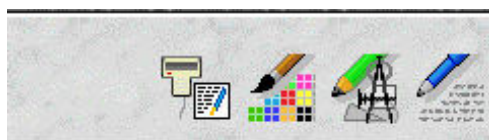
Installing the program

Sleuth is supplied in an compressed archive. Drag the entire contents of the archive, including the Sleuth program and the examples to a suitable directory on your hard drive.

Once Sleuth is installed on your hard drive it is ready to use.

Loading Sleuth

To load Sleuth, double-click on the application's icon and it will be installed on the icon bar. If you click Select over the icon bar icon an empty window will be opened. The application now requires a sprite to process.



Sleuth on the RISC OS iconbar

Loading a sprite

To load a sprite drag the sprite's icon onto the icon bar icon or into the open window. Only one sprite can be loaded at a time. If a further sprite is loaded it will replace the current one. A new sprite cannot be loaded if the package is already converting a sprite. If a sprite file contains more than one sprite only the first will be displayed. Most scanners, especially those using a TWAIN driver (see below), can produce monochrome or greyscale sprites suitable for use with Sleuth. Sleuth can automatically load sprites that have been compressed using !Squash.

Greyscale sprites

Sleuth will accept greyscale sprites and convert them into monochrome sprites. Using the Input and processing preferences dialogue box, the user can determine the levels of grey converted to white in this process. A Pre-sharpening facility is provided that will generally improve the quality of degraded text during conversion.

Icon bar menu

Click Menu over the icon on the icon bar and the following menu will be displayed:

Info

This option on both the sprite window and icon bar menu displays the usual dialogue box giving information about Sleuth. Please quote the version and serial number shown in this dialogue box in any correspondence to us.

Select Fonts

This allows you to tell Sleuth if the text you want to OCR is in a normal or a bold font.

Save Options

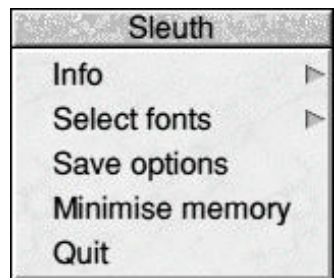
This will save as the default settings any changes you have made to the Sleuth options.

Minimise memory

Because scanned images take up so much memory this option has been provided to allow the memory used to store the sprite to be released. The sprite will be lost and the sprite window will be displayed empty. This option is greyed out when a sprite is being converted to text.

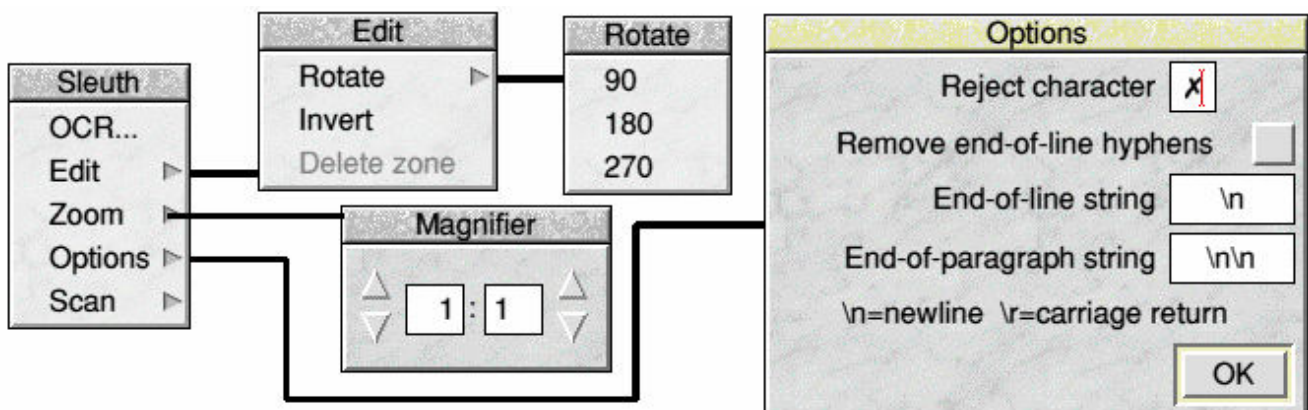
Quit

Choose this option to quit Sleuth. You will be warned if any converted text has not been saved and you will be given the option to save it before the application shuts down.



Sleuth Main menu

The Sleuth main menu is arranged as shown below



You will need to load a sprite into Sleuth before you can open the main menu. There are four example sprites included with Sleuth.

Manipulating the sprite

Before starting to convert a sprite you may need to alter it. The sprite must have black text on a white background and the text has to be upright. Sleuth allows the sprite to be zoomed, rotated or inverted. To rotate or invert the sprite choose the Edit option from the main menu and this submenu will be displayed:

Rotate offers a submenu from which you can choose to rotate the sprite by 90, 180 or 270 degrees. Choosing 90 degrees will rotate the sprite in an anti-clockwise direction; 270 degrees will rotate it in a clockwise direction.

Invert will swap the black and white elements of the sprite.

To zoom the sprite choose Zoom from the main menu and the zoom dialogue box will open.

Scanning using TWAIN

Sleuth supports TWAIN to allow direct scanning into the package. To use this facility you must have a copy of TWAIN, an appropriate scanner and Scanner Driver. TWAIN and scanner drivers can be downloaded from:

<https://www.davidpilling.com/wiki/index.php/TWAIN>

TWAIN supports the Select and Acquire options on the Scan submenu. Choosing Select opens a dialogue box that allows you to choose the scanner driver source that you require. Choosing Acquire lets you set up the scanner before scanning. Sleuth will only accept monochrome images, so make sure the BW option is selected. Full information about these options is provided with TWAIN.

The Save option allows you to save the scanned sprite.

Please note that !TWAIN must be seen before Sleuth can use these options.

Configuring Sleuth

Before beginning the OCR process it is important that Sleuth is configured correctly. The Options dialogue box is accessed by pressing Menu over the sprite to be scanned and going past the "Options" option on the menu.

The Reject character is the character used by Sleuth when, in its estimation, there is no equivalent character in the set of characters that it knows, or it is unsure what the correct character is. The Reject character should be set to an infrequently used character.

The Remove end-of-line hyphens option relates to the how the output file is saved. Any hyphens at the ends of lines will be removed if this option is selected.

End-of-line string can be set to any character. A carriage return is represented by \r escape sequence, a line feed by \n. These can be used separately or in combination e.g. \n\r. You could enter a space if you want to reformat the text in another package. These characters will be inserted at the end of each line of text converted by Sleuth.

End-of-paragraph string uses the same escape sequences as the End-of-line option. Sleuth will insert these characters at the end of each paragraph.

As these settings are read just before the converted text is saved it is possible to change them after the image is converted to text.

Clicking on OK will set the options for the current session.

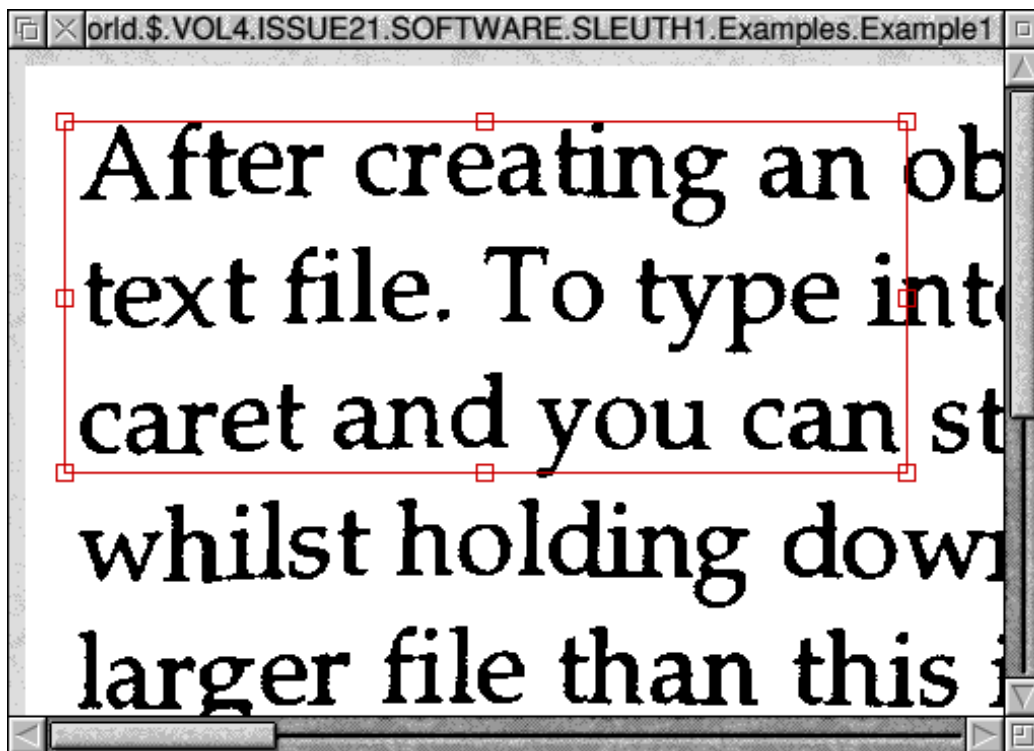
The OCR process

Sleuth cannot deal with complex page layouts automatically. A complex page layout might consist of text in columns and/or include graphics. Sleuth 1 can only OCR a block of text. Both Sleuth 2 and Sleuth 3 can cope with complex page layouts without assistance.

To OCR a simple block of text click Menu over the sprite window and choose OCR. A new window will open, this is the text window and this is where the converted text will appear. Once the OCR process has completed you can save the text by pressing Menu over this new window and clicking on Save.

Using Zones

It may be that you only want to OCR part of the sprite that you have loaded into Sleuth. In this case you can select the part you want with a zone. A zone is a user-defined area from which Sleuth will extract text. To create a zone drag with Select over the sprite window. As soon as you start dragging a zone rectangle will be drawn with eight handles and the pointer will be positioned in the bottom right-hand corner handle as shown below:



As you continue to drag the zone will be resized. The window will scroll automatically if necessary. Release Select to finish drawing the zone. The zone can be resized by dragging with Select over one of the handles and moved by dragging with Select inside the zone.

Sleuth1 only supports 1 zone, Sleuth 2 and Sleuth 3 both support multiple zones.

Manual editing

As text appears in the text window you have the option of clicking in the window with the mouse and manually editing the output to remove any mistakes. You can move the caret around with the cursor keys and remove incorrect letters with the delete key, just like you would in any word processor.

Getting the best results from Sleuth

There are many factors which affect the accuracy of the output from Sleuth. We recommend that the following suggestions are followed where possible to get the best results.

Use at least 300 or 400 dpi resolution scanning. Scan the original document where possible, not a photocopy. Scan a portion first and convert it to ensure that the scan is not too light or too dark. If you are using a hand-held scanner try to make the scan as straight as possible.

Here are some suggestions to help you get the best results:

- Review the document and plan the best method of scanning. Do not waste time scanning graphics that will be discarded in the OCR process.
- Scan a portion of the document and convert it to check that the contrast setting of the scanner is correct. Time spent getting this right will be saved later by having fewer errors in the output to correct.
- If you are using a hand-held scanner first run the scanner over the document without scanning to check there are no obstacles that will impede progress during scanning and that the scanner is running as straight as possible. Try to line up the body of the scanner with a straight line on the page or the left edge of the text being scanned.
- Scan the document carefully. If you are scanning to the end of a page you may find that a hand-held scanner will jump slightly as the body of the scanner rolls off the edge. This will almost certainly affect the output from Sleuth when it reaches this point and can have further detrimental effects if the scanner is tilted. Often in this case the image has smudges which degrade the text beyond recognition. One way to alleviate this scan the top half of the page and then turn the document round and scan the bottom half. The second half can be rotated in Sleuth before conversion.
- If using a flat-bed scanner make sure that the document sits flat on the scanner. If it is not flat smudges will probably appear on the scan.
- When scanning books keep the book flat and use the side of the scanner that has the least gap between the scannable area and the edge of the scanner. This will minimise problems with scanning to the inner margin.

Important points to remember

- Check that the image is not too bright or too dark and has a uniform contrast
- Check that the image is not overly skewed.
- Use a high resolution scan unless the text is large. 300-400 dpi should be sufficient.
- Sleuth has only been trained on a limited number of common fonts.
- There is a limitation on the size of text that can be converted.
- Sleuth has been trained on a specific set of characters, other characters will not be converted correctly.

Problem solving

These are possible problems that may occur when using Sleuth with explanations and possible solutions.

I'm getting output but it is all punctuation or apparently random characters.

The scan is either too light, too dark or contains a font that Sleuth has not been trained on. Re-scan the image to lighten or darken it. If there is a font that is not recognised by Sleuth in the image it will not be able to accurately convert it.

The output is good but the occasional word comes out badly.

It is likely that there is some defect in the scan where the conversion accuracy breaks down or the word is in a font unknown to Sleuth.

The output is good at the beginning of the page, but gradually deteriorates.

Check that the contrast of the image is consistent. Uniform contrast over the image is important to get good results. Inconsistent contrast is likely to occur if the document has not been printed well or is a photocopy. Try re-scanning the document in portions adjusting the contrast of the scanner to compensate for the change in contrast.

The output is good, but the order of the text is wrong.

The page layout is probably complex Sleuth will not be able to scan it. Select each section in turn with the zone tool and OCR that. Save the text for that section and then continue to the next section.

Scans have a black line or garbage down the right hand side

If your scans have a black line or garbage down the right-hand side of the scan you should obtain a version 2 scanner driver (source). The version number of the scanner driver can be found in !TWIN.Sources.!SourceName.!Info where !SourceName is the name of the scanner.

Sleuth 1 limitations

Although Sleuth has only been trained on the standard fonts it will recognise other similar fonts without further training. It will recognise the following characters :

!?"\$%&@'""%+=_{}#0{ }[]<>,-.::0123456789ABCDEFGHIJKLMNOPQRSTUVWXYZabc
defghijklmnopqrstuvwxyz'""fifl•—

Sleuth will convert characters between 9 and 24 point in size. This equates to approximately 1/8th to 1/3rd of an inch depending on the font. The actual limits are between 10 and 80 pixels measured from the top of an upper-case letter, like an 'A', to the bottom of a lower-case letter with a descender, like a 'g'. To convert 24 point text use a 200 dpi scan.

Sleuth will convert text at speeds between 80 to 250 words per minute depending on the machine used, the size and quality of the text and the resolution of the sprite.

Sleuth will automatically cope with slightly skewed scans (generally less than 2 degrees of skew, depending on line spacing) and lines of text that are slightly wavy.

RISC OS 2 Users

The templates as supplied are set-up for RISC OS 3. If you are using RISC OS 2 you can use an alternative set of templates that are in the !Sleuth.Resources.UK directory. To do this rename the Templates file to RO3Templs and then rename RO2Templs to Templates. Now when you run !Sleuth the RISC OS 2 templates will be used. Please note that the layout of the RISC OS 2 templates is different from the RISC OS 3 templates.

RISC OS 2 hard disc users should copy !Scrap into the root directory of their hard discs. !Scrap defines an area where temporary files can be stored.

Updates since original release

Sleuth Version 1.53

1. Sprite window not re-opened until scanning finished.
2. Degree of multitasking controlled via a message in the messages file.

Sleuth Version 1.52

1. TWAIN interface 'tweaked' to allow polling during scanning.

Sleuth Version 1.51

When using Sleuth with 400 d.p.i sprites on an ARM 2 machine to maximise the speed of conversion change to a low resolution screen mode (12 or 0) and remove all unnecessary applications from the icon bar.

Four examples scans are provided in the Examples directory.

Please note that the Info option has been removed from the main menu.

Sleuth can now import sprites that have been squashed using the !Squash application provided with RISC OS 3.1 and later.

Sleuth will now import mode 18, 25 and 29 sprites.

You will require a version 1 or later TWAIN scanner driver (source) to use TWAIN with this version of Sleuth.

Here are just a few of the other products you will be able to
download from RISCOS.com free of charge
(You can make a donation to help support future product releases but this is entirely voluntary)



This Documentation is © 3QD Developments Ltd 2026
Some of the products shown above might not be available immediately.
Not all products will run on all versions of RISC OS, check the product information page for details.

Other Products list last updated 29/2/26