

# SOFTWARE PROGRAMS AND WEB SERVICES FOR OPTICAL CHARACTER RECOGNITION (OCR)

by Francis Chao

[fchao2@yahoo.com](mailto:fchao2@yahoo.com)



Web location for this  
presentation:

<http://aztcs.org>

Click on

“Meeting Notes”

# SUMMARY

OCR stands for "Optical Character Recognition". It means converting the graphical representation of text into text.

# TOPICS

- Definition of "OCR"
- "OCR" with "OneNote" of "Microsoft Office".
- "OCR" with Web-based Services
- "OCR" with software computer programs

# DEFINITION OF "OCR"

# DEFINITION OF "OCR"

- See

[http://en.wikipedia.org/  
wiki/Optical\\_character  
recognition](http://en.wikipedia.org/wiki/Optical_character_recognition)

"OCR" FROM INSIDE  
"ONENOTE" OF  
"MICROSOFT  
OFFICE.."

# "OCR" FROM INSIDE "ONENOTE" OF "MICROSOFT OFFICE.."

- How NOT to do "OCR" inside "OneNote":
  - Import the graphical file into the "OneNote" document.
  - Then do a "Send" to "Adobe PDF" or "Microsoft Word" to create the output document.



The screenshot shows the Microsoft OneNote application interface. At the top, the title bar reads "OneNote: one place for all of your notes - OneNote". The ribbon menu includes FILE, HOME, INSERT, DRAW, HISTORY, REVIEW, and VIEW. The notebook name is "Frank's Notebook", and the current section is "Quick Notes". A search bar is visible in the top right corner.

The main content area displays a list of notes. The first note is titled "1. Take notes anywhere on the page" and contains the text "Write your name here". A hand-drawn arrow points from this text to the "PRESENTATIONS FOR THE ONLINE APCUG SPEAKERS BUREAU" section below.

The "PRESENTATIONS FOR THE ONLINE APCUG SPEAKERS BUREAU" section is a bolded title with "APCUG SPEAKERS BUREAU" underlined in blue. Below the title, the text reads: "Francis Chao, who is an active member of [Windows Users Computer User Group](#) of California and [Tucson Computer Society](#) of Arizona, is prepared to present any of the following programs to your computer user group:".

The right-hand sidebar contains a "Search (Ctrl+E)" bar, an "Add Page" button, and a list of pages including "OneNote: one place for all of yo" and "OneNote Basics".



Info

New

Open

Print

Share

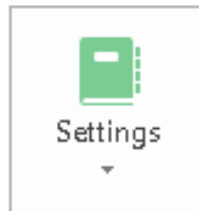
Export

Send

Account

Options

# Notebook Information



## Frank's Notebook

<https://onedrive.live.com/edit.aspx/Documents/Frank^4s%20Notebook?...>

 [Invite people to this notebook](#)

[View Notebook](#)



Info

New

Open

Print

Share

Export

Send



Info

New

Open

Print

Share

Export

Send

Account

Options

# Send

Current page: "OneNote: one place for all of your notes"



Email Page



Send as Attachment



Send as PDF



Send to Word



Send to Blog

Document1 - Word

FILE HOME INSERT DESIGN PAGE LAYOUT REFERENCES MAILINGS REVIEW VIEW **FORMAT** Frank Chao

Read Mode Print Layout Web Layout Outline Draft Show Zoom 100% New Window Arrange All Split Switch Windows Macros

1 2 3 4 5 6 7

## **PRESENTATIONS FOR THE ONLINE APCUG SPEAKERS BUREAU**

Francis Chao, who is an active member of [Windows Users Computer User Group](#) of California and [Tucson Computer Society](#) of Arizona, is prepared to present any of the following programs to your computer user group:

- **Using The HP "Stream 11" and Other Low End Laptops (B)**

The HP "Stream 11" laptop brings "Windows.." laptops down to the \$200 price point for the first time.

PAGE 1 OF 2 0 WORDS 75%

Document1 - Word

FILE HOME INSERT DESIGN PAGE LAYOUT REFERENCES MAILINGS REVIEW VIEW **FORMAT** Frank Chao

Read Mode Print Layout Web Layout Outline Draft Show Zoom 100% New Window Arrange All Split Switch Windows Macros

Views Zoom Window Macros

1 2 3 4 5 6 7

## **PRESENTATIONS FOR THE ONLINE APCUG SPEAKERS BUREAU**

Francis Chao, who is an active member of [Windows Users Computer User Group](#) of California and [Tucson Computer Society](#) of Arizona, is prepared to present any of the following programs to your computer user group:

- **Using The HP "Stream 11" and Other Low End Laptops (B)**

The HP "Stream 11" laptop brings "Windows.." laptops down to the \$200 price point for the first time.

PAGE 1 OF 2 0 WORDS 75%

# "OCR" FROM INSIDE "ONENOTE" OF "MICROSOFT OFFICE.."

- How NOT to do "OCR" inside "OneNote" (continued):
  - The "Send" in "OneNote" creates a "Microsoft Word" or "Adobe PDF" that has a graphic of the source text in it!

# "OCR" FROM INSIDE "ONENOTE" OF "MICROSOFT OFFICE.."

- How to do real "OCR" inside "OneNote":
  - Import the graphical file into the OneNote document.
  - Use the RIGHT mouse button to click on the imported text.



The screenshot shows the Microsoft OneNote application window. The title bar reads "OneNote: one place for all of your notes - OneNote". The ribbon menu includes FILE, HOME, INSERT, DRAW, HISTORY, REVIEW, and VIEW. The notebook name is "Frank's Notebook" and the current section is "Quick Notes". A search bar is visible in the top right corner. The main content area displays a list item:

1. Take notes anywhere on the page  
Write your name here

A hand-drawn arrow points from the text "Write your name here" to the right side of the page. Below the list item, there is a horizontal separator line. The main content area contains the following text:

**PRESENTATIONS FOR THE ONLINE  
APCUG SPEAKERS BUREAU**

Francis Chao, who is an active member of [Windows Users Computer User Group](#) of California and [Tucson Computer Society](#) of Arizona, is prepared to present any of the following programs to your computer user group:

The right-hand sidebar shows a "Search (Ctrl+E)" field, an "Add Page" button, and a list of pages including "OneNote: one place for all of yo" and "OneNote Basics".

ce for all of your notes - OneNote



Cut



Copy

Copy Text from This Page of the Printout

Copy Text from All the Pages of the Printout

Zoom to Document

View Original Printout



**Paste Options:**



Save As...



Link... (Ctrl+K)



Copy Link to Paragraph

Order

Move

Resize



Restore to Original Size



Alt Text...

Make Text in Image Searchable

Search



A

OneN

OneN

e p

RT

S RIRFALL



Cut



Copy

Copy Text from This Page of the Printout

Copy Text from All the Pages of the Printout

---

# "OCR" FROM INSIDE "ONENOTE" OF "MICROSOFT OFFICE.."

- How to do real OCR inside  
"OneNote" (continued):
  - Go to a blank "Microsoft  
Word" document and copy  
the text into it:

Document1 - Word

FILE HOME INSERT DESIGN PAGE LAYOUT REFERENCES MAILINGS REVIEW VIEW Frank Chao

Paste Calibri (Body) 11 B I U abc x<sub>2</sub> x<sup>2</sup> A Styles Editing

Clipboard Font Paragraph Styles

PRESENTATIONS FOR THE ONLINE

APCUG SPEAKERS BUREAU

Francis Chao, who is an active member of Windows Users

Computer User Group of California and Tucson Computer Society

of Arizona, is prepared to present any of the following programs

to your computer user group.

- Using The HP "Stream 11" and Other Low-End Laptops (B)

The HP "Stream 11" laptop brings "Windows..." laptops down

to the \$200 price point for the first time.

(Ctrl)

PAGE 1 OF 1 71 WORDS 100%

## PRESENTATIONS FOR THE ONLINE

### APCUG SPEAKERS BUREAU

Francis Chao, who is an active member of Windows Users Computer User Group of California and Tucson Computer Society of Arizona, is prepared to present any of the following programs to your computer user group.

- Using The HP "Stream 11" and Other Low-End Laptops (E)

The HP "Stream 11" laptop brings "Windows.." laptops down to the \$200 price point for the first time.

# "OCR" FROM INSIDE "ONENOTE" OF "MICROSOFT OFFICE.."

- Fonts, font colors, font effects, paragraph blocking, and graphics in the original document will be lost when you use "OneNote" of "Microsoft Office" to do "OCR".

# "OCR" WITH WEB-BASED SERVICES



# "OCR" WITH WEB-BASED SERVICES

- Size and type of input file is limited
- Varying levels of success
- Usually free

# "OCR" WITH WEB-BASED SERVICES (continued)

- The best Web-based service for doing "OCR" is <http://free-online-ocr.com/>

http://free-online-ocr.com/ Free Online OCR

g+1 415 Like Share 2,930 people like this. Sign Up to see what your friends like.

# Free Online OCR


Convert | [Features](#) | [FAQ](#) | [Support](#)

Select your image or PDF file

 Browse...


Select your output format

Word Document (DOC) ▼



Convert


## Free Online OCR



Convert scanned images into editable text.


Free Online OCR is a free service that allows you to easily convert scanned documents, faxes, screenshots and photos into editable and searchable text, such as DOC, TXT or PDF.

## Features



- Precise image to text conversion
- Keeps the layout and formatting
- Scanned PDF to DOC conversion
- Supports PDF, GIF, BMP, JPEG, TIFF or PNG

## What is OCR?



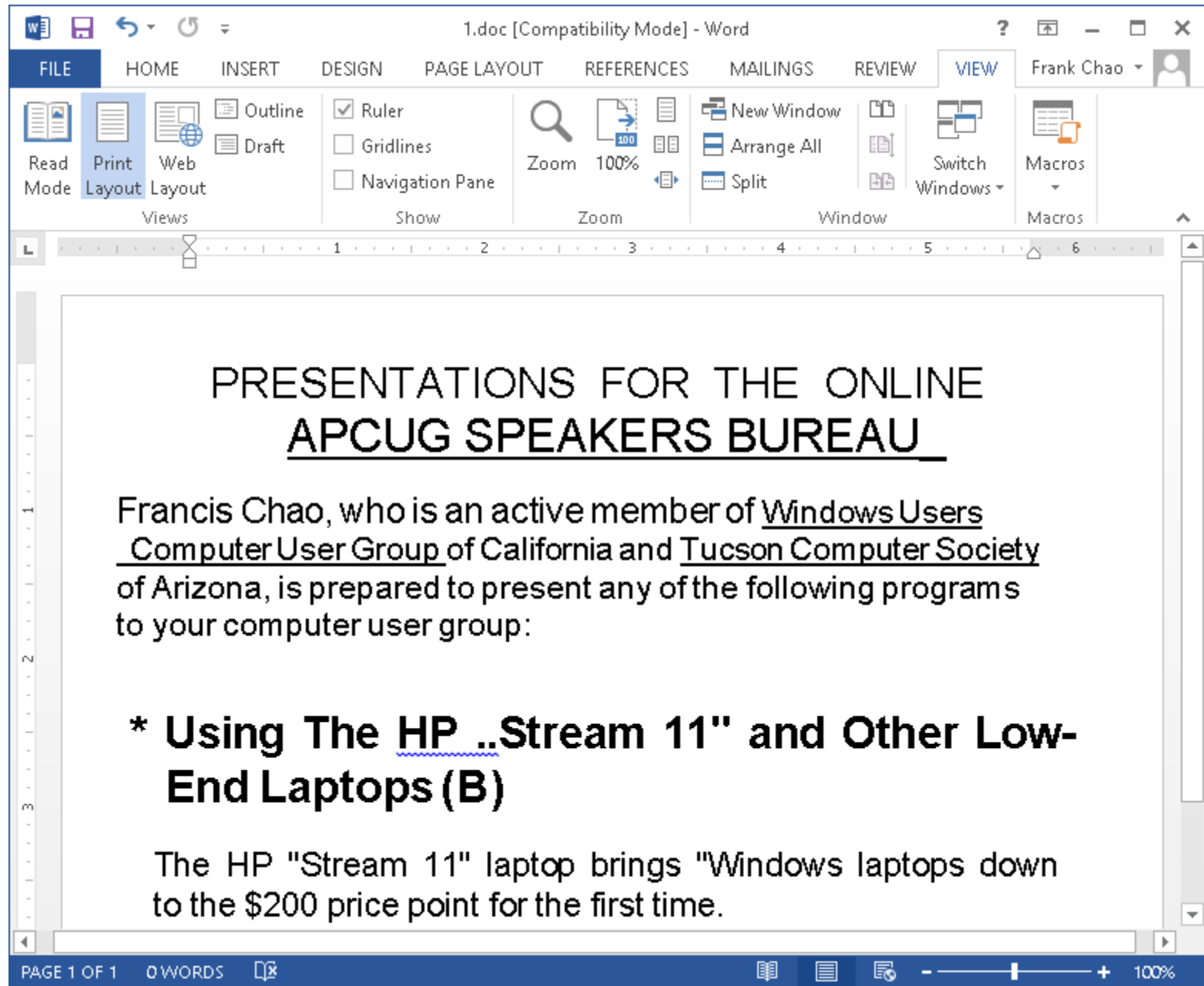
OCR (Optical Character Recognition) is a technology that extracts the text from an image or a scanned document so that it can be edited, formatted, searched, indexed, automatically translated or converted to speech.

OCR can be used to convert books and

# "OCR" WITH WEB-BASED SERVICES

(continued)

- Not as good as commercial "OCR" software programs:  
Some spacing inaccuracies between words  
and  
font colors were not detected  
and size of the input file is limited:



1.doc [Compatibility Mode] - Word

FILE HOME INSERT DESIGN PAGE LAYOUT REFERENCES MAILINGS REVIEW VIEW Frank Chao

Read Mode Print Layout Web Layout

Outline Ruler Gridlines Navigation Pane

Zoom 100%

New Window Arrange All Split

Switch Windows

Macros

1 2 3 4 5 6

## PRESENTATIONS FOR THE ONLINE APCUG SPEAKERS BUREAU

Francis Chao, who is an active member of Windows Users Computer User Group of California and Tucson Computer Society of Arizona, is prepared to present any of the following programs to your computer user group:

**\* Using The HP ..Stream 11" and Other Low-End Laptops (B)**

The HP "Stream 11" laptop brings "Windows laptops down to the \$200 price point for the first time.

PAGE 1 OF 1 0 WORDS 100%

# "OCR" WITH SOFTWARE COMPUTER PROGRAMS

# "OCR" WITH SOFTWARE APPLICATION PROGRAMS

- No limit to the size of the input file
- Usually more tolerant of the type of input file+
- Usually not free

# "OCR" WITH SOFTWARE APPLICATION PROGRAMS

- Nuance Omnipage
- ABBYY FineReader
- I.R.I.S.'s ReadIRIS



