



PrimeVerify

A High Productivity Post-OCR Error Correction Tool

Once an image has been OCR'd by PrimeOCR, the text output can be manually reviewed and corrected through PrimeVerify.

Key Design Features

Throughput. PrimeVerify was designed to speed manual OCR verification, often the dominant cost of the whole imaging system. Users see the recognized text and the original image together for easy "heads up" editing. Characters flagged as "suspicious" by PrimeOCR are highlighted in red. A red box is also drawn around the suspicious character on the image. Users can also quickly jump between highlighted characters using the TAB key.

Accuracy. Suspicious characters are highlighted based upon a user-defined confidence threshold. There are nine different thresholds to choose from, allowing the user to select the level of accuracy desired in the output. At the highest level of review the verified accuracy is typically over 99.995%.

LAN Based. PrimeVerify was designed as a stand alone application on a LAN. It runs independently of PrimeOCR, unlike conventional OCR products. You can have any number of copies of PrimeVerify that process the output of one or multiple PrimeOCR units.

PrimeVerify - Simultaneously edit PrimeOCR results and view the original image document in a "heads up" style split screen display.

The screenshot shows the PrimeVerify application window titled "PrimeVerify - c:\prdev\sample01.pro - page 1 of 1". The menu bar includes "File", "Edit", "Image", "Page", "Text Settings", "KeyTable", and "Options". The main text area displays OCR results for a document about "OmniPATH for Windows". Several callout boxes provide details:

- Text Selection:** "As you select text to edit, the same area on the image is selected for fast comparison." (points to a red box around the word "intuitive")
- Navigation:** "Use shortcut keys or the mouse to move between and select characters, words, or even full lines of text for editing." (points to the cursor on the word "intuitive")
- Font Size:** "Change the font size of recognized text for easier proofreading." (points to the word "so" in "so that")
- Image Display:** "Adjust the image display using Zoom In, Zoom Out, Invert, and Greyscale." (points to the right side of the window)
- Confidence Levels:** "Nine different text confidence levels are supplied by PrimeOCR. Suspect text is highlighted in red based upon the confidence threshold you select." (points to the word "intuitive" which is highlighted in red)
- Automatic I/O:** "PrimeVerify can perform automatic file I/O when you reach the end of the current file." (points to the end of the text)



"PrimeOCR's higher OCR accuracy means cleaner data and significantly less OCR verification for our customers."

Scott Blau, Datacap

PrimeVerify Features

PrimeVerify Functionality

The PrimeView application window presents a split screen view. The top section, the image display area, will show the original image used by PrimeOCR to generate the results file. The bottom section, the edit display area, will contain the recognized text in editable form.

The dual screen view allows users to quickly compare recognized text with the text as it appears on the original image. This, combined with the PrimeVerify proofreading features described below, makes for efficient and comprehensive manual error correction.

Make Image Display Adjustments

Once a PrimeOCR results file is opened, you can make changes to the way its corresponding image file appears in the image display area through menu options such as invert, grayscale, and zoom in/out.

Make Text Display Adjustments

The font size of text recognized by PrimeOCR and displayed in the edit display area can be changed for easier proofreading.

Highlight Low Confidence Characters

Suspect or Low Confidence characters in the edit display area, marked by PrimeOCR, are highlighted in red. This makes for easier identification of possible recognition errors that may need to be corrected. Users have nine different confidence level thresholds from which to choose. Selecting a lower threshold highlights more characters.

Text and Image Navigation

Use the TAB key to quickly move the cursor to the next low confidence character (displayed in red) or use keystrokes, menu options (i.e., find/replace) or mouse clicks to locate the characters you wish to edit. The corresponding image display is automatically synchronized to show the same text area highlighted on the image for heads up editing.

Edit Text

In addition to the mouse, a variety of keyboard shortcuts allow you to select and replace characters, words, or entire lines of text in the PrimeOCR results file. Powerful menu options give you the ability to convert character confidences as you tab through text, skip over user-defined characters, convert selected text to upper or lower case, and even replace selected text with other text selected in user-defined tables.

Saving and Exporting Text

You can store intermediate editing results for later completion or export corrected text into a variety of file formats. ASCII, Formatted ASCII, PDF, HTML, and RTF are some of the output file types supported and multiple output files can be exported from one input file.

Configure Options

PrimeVerify includes several processing options that can speed verification. For example, once you reach the end of the current file, PrimeVerify can be configured to automatically save or export the file and open a new file for editing. You can even automatically delete the current file after it has been exported or just move it to another directory.

View Image Statistics

PrimeOCR generates character confidence statistics on every image it processes. These statistics can be viewed in PrimeVerify and detail average image confidence, the number of characters marked as suspicious in each of PrimeOCR's nine confidence levels, and cumulative confidence totals (see below):

Image Statistics		
Confidence Level	Characters Marked	Cumulative %
1	1	0
2	0	0
3	0	0
4	0	0
5	1	0
6	6	1
7	0	1
8	8	1
9	1522	100
Total	1538	100
Average Confidence		8.97

OK

System Requirements

Software:

- Microsoft Win95/98
- Microsoft Windows NT 4.0 (Workstation/Server)
- Microsoft Windows 2000 (Professional/Server)

Hardware:

- Pentium based Intel compatible computer
- One CD drive to install software
- A hard disk with 5 MB of space for installation
- At least 32MB of free RAM