

PRIME

EMC Business Partner Program for Technology Connect Partners

SELECT PARTNER

Partner Products

- PrimeOCRIA
- PrimeVerifyIA

EMC Products

- EMC Captiva ®

Solution highlights

- Converts images to PDF and other formats with very high voting OCR accuracy using PrimeOCR
- Allows users to verify/edit OCR output to 99.95% accuracy or higher with optional PrimeVerify module
- Supports high-volume, unattended production environments with built-in fault tolerance
- Will recognize text in English, Western European, Asian and Russian languages



PRIME OVERVIEW

Prime has always focused on the OCR needs of the largest organizations, specifically high accuracy, high volume, high reliability, and responsive support. This focus distinguishes it from most OCR companies, which concentrate their efforts on the larger but very different desktop/single user OCR market.

EMC's Captiva Capture was also focused on sophisticated, large imaging customers and is an excellent strategic match for Prime. Prime was one of the first companies to develop a Captiva Capture certified module in 1997 and has continued a strong investment in integrating new generations of PrimeOCR and Captiva Capture over the years.

Although conventional OCR technology has improved over the years, two major performance measures remain below the market's needs: accuracy and reliability. Accuracy on typical quality pages is still in the 98-99% range (which may sound good but represents 40-50 errors on a typical page, whereas <1 error is desired). And conventional OCR, an impressive but complex amalgamation of decades of artificial intelligence algorithm development, can still crash on any arbitrary image, leading to a process interruption and no output for that page.

Prime licenses the best underlying OCR technology in the world from the desktop vendors and via variety of techniques, including "voting" which it is best known for, has developed the most accurate machine print OCR available on the market, reducing error rates by 50-80% (or put another way, a factor of 2-5 times). Prime uses a modular internal architecture in which each different OCR technology runs independently to create both fault tolerance and speed (if multiple processors are licensed).

PRIMEOCRIA

PrimeOCRIA is a tightly integrated EMC certified Captiva Capture module, designed to provide all the interfaces required to utilize PrimeOCRIA with other EMC modules (see graphic).

It accepts as input TIFF, JPEG, PDF and other image types, in black and white, color or grayscale. These images can be enhanced via deskew, autorotation, despeckle, and other techniques.

Multiple levels of accuracy and speed tradeoffs can be licensed and configured. At the lowest level of accuracy PrimeOCRIA is faster than many conventional engines and at the highest levels of accuracy has much better accuracy (and is speed competitive). An "intelligent voting" setting is available that will analyze each page and if the quality is high, will process the page quickly, and if the quality is low, will apply as much

REDEFINE

SOLUTION OVERVIEW

The EMC² logo is displayed in white text on a blue background.

voting as possible for the best accuracy. This combines the best of speed and accuracy.

PrimeOCRIA can recognize machine print in any of the following languages: English (US or UK), Danish, Dutch, French, German, Italian, Norwegian, Portuguese, Spanish, Swedish, Russian, Japanese, Korean, and Chinese (Simple and Traditional).

Prime was one of the first OCR vendors to offer PDF output and it continues to lead the market in supporting PDF features such as PDF/A (1a and 1b) support, Accessible PDF (Section 508 compliant), zip compression (required by some government agencies), bookmark output, and retention of incoming PDF bookmarks, hyperlinks, and other PDF data. Other output formats include variations of ASCII, RTF, XML, HTML, PDA, and assorted other formats included many custom formats developed for customers. In addition, as a full Captiva Capture module, PrimeOCRIA outputs many of its data as IA values or IA files, strings, etc. so that they are available to IA process developers for processing flow control.

PRIMEVERIFYIA

PrimeVerifyIA is also a tightly integrated EMC certified Captiva Capture module. This is an optional module that takes OCR output from PrimeOCRIA and presents it to an operator. The operator reviews and if necessary corrects the text. In this way, output accuracy can be improved to 99.95% (or higher or lower as desired/configured by the customer). Output from PrimeVerify can then be reinserted into the IA process flow (as any of Prime's supported formats including PDF, XML, etc.)

This post OCR verification step is expensive since it uses manual labor. PrimeVerifyIA offers lower costs than competitive products because (1) the higher accuracy of PrimeOCR means that many fewer characters are presented to the user for review, and (2) there are many features of PrimeVerifyIA that are designed to make the user more efficient. For example, only "low confidence" characters are reviewed by the user. (The user can configure which of 9 different confidence levels constitutes low confidence in their application.)

ABOUT PRIME

Prime Recognition designs and develops advanced machine print character recognition software for the production OCR marketplace using "voting" technology. Prime Recognition markets worldwide, its customers include the Fortune 1000, service bureaus, system integrators, university/library systems, law firms, and the U.S. Federal Government which is Prime's largest segment and in which its largest installations are located. Prime Recognition was founded in 1994 and is headquartered in the Seattle area. Visit www.primerecognition.com for more information.

ABOUT EMC

EMC is a global leader in enabling businesses and service providers to transform their operations and deliver IT as a service. Fundamental to this transformation is cloud computing. Through innovative products and services, EMC accelerates the journey to cloud computing, helping IT departments to store, manage, protect, and analyze their most valuable asset—information—in a more agile, trusted, and cost-efficient way.

CONTACT US

To learn more about how EMC products, services, and solutions can help solve your business and IT challenges, [contact](#) your local representative or authorized reseller, visit www.emc.com, or explore and compare products in the [EMC Store](#).

EMC², EMC, the EMC logo, are registered trademarks or trademarks of EMC Corporation in the United States and other countries. VMware is a registered trademark or trademark of VMware, Inc., in the United States and other jurisdictions. © Copyright 2014 EMC Corporation. All rights reserved. Published in the USA. 02/15 Solution Overview, H12861.2.

EMC believes the information in this document is accurate as of its publication date. The information is subject to change without notice.

The EMC logo consists of the letters "EMC" in a bold, white, sans-serif font, with a small superscript "2" to the right of the "C". The logo is set against a solid blue rectangular background.