



MARKETFOCUS REPORT

Doculabs

Richard Medina and Bob Anders

Assessment of Kofax VirtualReScan



Table of Contents

Executive Summary	3
Maximizing Capture Throughput and Productivity.....	4
Optimizing Effective Throughput.....	5
Kofax VRS Approach to Effective Throughput	6
Company Overview.....	6
VRS Product Overview	6
Capabilities.....	7
Strategic Benefits	8
Image Quality	8
Input Range.....	11
Usability.....	11
Speed.....	12
Conclusion.....	13
Appendix: About Doculabs.....	14

EXECUTIVE SUMMARY

Organizations that need to quickly and accurately capture and digitize their business-critical documents must concern themselves with much more than scanner speed. Factors such as image quality, document input range, and usability all contribute to the overall efficiency of a scanning system. For example, poor image quality, inability to handle diverse document types, and hard-to-use interfaces all contribute to more time spent in scanner setup, adjustment, and rescanning – which increases the overall time and cost of a scanning operation.

One vendor offering organizations a compelling solution for reducing the operational costs and increasing the productivity of their document capture operations is **Kofax Image Products**. The company's VirtualReScan (VRS) solution provides an easy-to-use interface and tools for users to adjust critical scanning parameters such as brightness, skew, and contrast for multiple types of complex documents in a single batch. Then, at run-time, VRS automatically applies these adjustments and corrections to ensure that the best images are moved from the scan station into the application.

Kofax commissioned Doculabs, an independent research and consulting firm, to objectively test, analyze, and articulate the performance of Kofax's VRS technology and its ability to increase an organization's scanning productivity, reduce errors and rescanning, increase the usability of the scanning system for operators, and ultimately reduce overall scanning costs.

Doculabs found that Kofax's VRS provides a strong solution for making any document capture process more efficient, because it addresses image quality, document input range, and usability together in a single solution. Through the use of VRS, organizations can reduce their document preparation times, simplify their scanning process, and improve document image quality for OCR applications. The fact that VRS addresses all of these factors within a single product makes it a particularly strategic approach for maximizing the effective throughput of document capture operations.

MAXIMIZING CAPTURE THROUGHPUT AND PRODUCTIVITY

The goal of any capture system is to get document information into downstream systems as quickly and efficiently as possible, maximizing throughput. Thus, capture operation efficiency hinges on achieving acceptable output quality, in minimal time, with minimal cost, over a sustained duration. In addition, a capture system must work well with all the other components in the input process, including scanners, capture software, recognition engines, and downstream information systems such as imaging applications or line-of-business applications.

We call this “effective throughput.” There are four key factors that contribute to the overall performance and efficiency – the effective throughput – of a document capture system, including:

- ***Speed***
The speed of the scanner and the speed of the image processing stage of the scan process contributes to the overall throughput speed of a scanning process
- ***Image Quality***
High-quality images not only are easier to read by humans, but they lead to higher accuracy by automated recognition technologies such as OCR engines (which minimizes manual data entry and re-scanning)
- ***Input Range***
The ability for scanning systems to handle the documents you need to scan – which may include a wide variety of complex documents (multiple sizes, multiple colors, etc.) in single or multiple batches – in order to minimize document preparation, quality assurance, and re-scanning
- ***Usability***
If the scanning system is easy for administrators and operators to use, it will reduce time spent in configuring the system for optimal performance, and will minimize the training needed to use the system

Getting all of these factors to work together to maximize throughput is a difficult task, especially in high-volume scanning operations. For example, some scanning solutions provide good image quality, but configuring the scanner settings and parameters can be extremely complex and time-consuming. Similarly, easy-to-use systems may lack the input range or image quality needed in high-volume environments with diverse documents. Few technologies are designed to address all of the factors involved in maximizing effective throughput.

Optimizing Effective Throughput

To optimize the capture process, one must analyze it in stages, determine the dependencies and weak links, and then choose components and processes to optimize the overall input process. The input process consists of three main stages, pre-recognition, recognition, and post-recognition.

Pre-recognition consists of document design, processing, and delivery; document availability for input; document preparation; document feeding and handling; capture; image enhancement and correction; and possible rescan. The recognition phase consists primarily of recognition and indexing. The post-recognition phase includes error detection; error correction; re-keying; possible re-scanning and exception handling; and release to a downstream repository.

If you analyze the costs of each stage, you can divide the input process into relatively high cost stages and relatively low cost stages, where cost is a function both of the direct cost of a stage and the cost of the stages that depend on it. High cost stages generally include document preparation, all rescan and exception handling, and all post-recognition stages (except release). Low cost stages generally include capture, recognition and indexing, and release.

A quick look at the factors involved in effective throughput will show that there are more important factors than scanner speed. Speed is linear and predictable, as pages-per-minute is an easily calculable rate. However, other factors can have a much more dramatic impact on overall capture time. For example, higher OCR accuracy can reduce or eliminate re-scanning and manual data entry, which are extremely time-consuming. Similarly, a system that is easy to set up and configure for a range of document types will greatly reduce setup time and document preparation time. Clearly, maximizing efficiency in each of these areas – especially the high-cost areas – can pay big dividends in terms of overall effective throughput.

Kofax developed its VirtualReScan (VRS) solution with these factors in mind. By automatically improving image quality and ensuring that only high-quality images are sent into the application, the system minimizes errors and manual data entry, makes it easier for scanner operators to improve image quality on the fly, and reduces costly and time-consuming rescans and exception handling. Thus, unlike other document input or capture technologies, Kofax VRS is designed to address all of the factors involved in maximizing effective throughput within a single solution.

KOFAX VRS APPROACH TO EFFECTIVE THROUGHPUT

This section provides an overview of Kofax and highlights the way in which VirtualReScan contributes to the overall efficiency of a document capture operation.

Company Overview

Founded in 1985 and headquartered in Irvine, California, Kofax Image Products develops products that manage the conversion of paper documents into electronic information. Kofax is a division of DICOM Group plc, a publicly traded company (London Stock Exchange and Neuer Markt: DCM) that reported revenues of \$150.5 million in the financial year ending June 30, 2000. DICOM employed over 620 people at the time of this review.

VRS Product Overview

Kofax's scanning and image processing solution, VirtualReScan (VRS), uses several tools that enable users to analyze and determine the optimum settings for scanned documents (both color and black-and-white). At run-time, the product then automatically converts these documents into high-quality, black-and-white images for rapid transport into back-end systems.

VRS works by first converting documents to grayscale images, analyzing them, determining the proper settings for the document, and then converting the image to black and white. Thus, the system produces high-quality images without large file sizes and slower processing associated with grayscale images.

VRS includes both hardware and software components. For production scanning environments, VRS is assisted by Kofax Adrenaline scanner boards in order to drive documents at speeds up to 125 pages per minute. VRS-ready production scanners include high-volume units from Bell&Howell, Fujitsu, and Ricoh.

For both production and lower volume scanning environments, VRS software runs on Windows NT, 2000, 95, 98, and ME platforms. Not only can VRS software be used with VRS hardware and production scanners, but it can be used in stand-alone mode with many mid-volume to low-volume scanners, including models from Bell&Howell, Canon, Fujitsu, Kodak, Panasonic, Ricoh, and Hewlett-Packard. The software works with a range of SCSI controllers from Adaptec.

Capabilities

To set up the scanning system using VRS, scanner operators or administrators configure parameters for each document type. For example, users can quickly establish scanner settings for each document type, with defaults for parameters such as brightness, contrast, resolution, etc. When scanning is initiated, VRS inspects each document and automatically enhances the image to ensure that only high-quality images that meet user specifications are passed through the system.

VRS provides a number of capabilities to enhance image quality. For example, the system's automatic deskew capabilities are used to straighten each scanned image, which improves recognition accuracy as well as the visual alignment of the image. VRS can thin or thicken lines, eliminate noise, and reduce speckling in order to produce cleaner images. In addition, a black border cropping feature can either remove black border pixels from an image to reduce its height and width, or change these pixels to white in order to clean the image without changing its height or width.

VRS' interface is extremely intuitive, and enables less technical individuals to configure settings for optimal performance. For example, many settings are configured with simple slider bars. Users can fine-tune the settings right on the screen in real-time, with no need to re-configure the scanner itself, no need to re-scan the document in order to validate the modified settings, and, while in production, there is never a need to stop the scanning process if an image falls outside the user definable range. VRS allows scan operators to manually intercede and make adjustments on a particular image, which can be saved and used as new default settings for a given document type.

Another key capability of VRS is the input range of documents that it can handle. Scanner operators can place documents of any size, color, or orientation into the scanner and begin the process. VRS automatically detects paper jams, bent corners, or other equipment problems and alerts the scanner operator.

If scanner speed is important, VRS can be configured to scan at lower resolutions. However, VRS also accelerates scanning in other ways. For example, VRS allows users to create and save profiles for multiple document types, which saves time in document preparation. In addition, for new documents, VRS can display optimum settings for any scanned image, allowing users to create new default settings for a particular document type.

STRATEGIC BENEFITS

This section provides the strategic benefits that VRS provides organizations in terms of its ability to improve the overall efficiency of scanning operations. VRS achieves maximum productivity by focusing on the many factors that have an impact on overall system efficiency. Specifically, the VRS system provides benefits in the following areas:

- Image Quality
- Input Range
- Usability
- Speed

Image Quality

Kofax's VRS provides the following benefits for image quality:

- With its image enhancement capabilities, VRS produces sharp, crisp images that are easier for users to read, which helps minimize rescanning requirements.
- VRS can be optimized to produce images that are well-suited for viewing in a browser by suppressing noisy areas and shaded backgrounds, and by producing smaller file sizes that facilitate rapid display and retrieval
- VRS produces images that are better suited for automated recognition engines such as OCR and ICR software. The result is more accurate data capture and reduced re-keying or manual data entry. (The following section provides information on how much VRS can improve OCR accuracy.)

OCR Accuracy Improvement Testing

As part of this assessment, Doculabs evaluated the impact that VRS can have on image quality and the resulting accuracy of optical character recognition.

The test environment consisted of a Fujitsu 3097 DE scanner, industry standard capture application, and Caere OCR recognition software. The source documents were airbill forms, which were processed for OCR accuracy results. The scanning and OCR application was set up to capture data from three fields: a seven-character alpha numeric field, a seven-character numeric field, and an 11-character alpha numeric field that was printed on top of a shaded background.

We ran OCR accuracy tests on a batch of 500 documents without using VRS, and with VRS added to the configuration. Each had three fields per document, for a total of 1,500 evaluated fields. Each document had 25 characters per document, for a total of 12,500 evaluated characters. For the OCR testing under both scenarios, we never spent more than a few minutes to set brightness and contrast for each batch. We ran each batch multiple times, all with similar results.

The test results show that overall, VRS improved OCR character recognition accuracy by nearly 35%, and it improved OCR field accuracy by more than 36%.

In our testing, the shaded field (Field #3) proved particularly difficult for the OCR application. By eliminating this field and looking only at the results from Field #1 and Field #2, VRS provides even more dramatic improvement. In these cases, VRS improved OCR character recognition accuracy by more than 57% (achieving character accuracy of better than 92%), and it improved OCR field recognition accuracy by 54% (achieving field accuracy of nearly 70%).

The following tables show the results for OCR field accuracy and OCR character accuracy.

OCR Character Recognition Accuracy				
Results Without VRS	Field 1	Field 2	Field 3	Total
Total Characters	3,500	3,500	5,500	12,500
Character Errors	2318	2236	5471	10025
OCR Character Accuracy	33.8%	36.1%	0.5%	19.8%
Results With VRS				
Max Characters	3500	3500	5500	12500
Character Errors	271	257	5147	5675
OCR Character Accuracy	92.3%	92.7%	6.4%	54.6%
VRS Improvement Over Non-VRS Results	58.5%	56.5%	5.9%	34.8%

Table 1 – OCR Character Accuracy (VRS vs. Non-VRS)

OCR Field Recognition Accuracy				
Results Without VRS	Field 1	Field 2	Field 3	Total
Total Fields	500	500	500	1500
Fields with Errors	431	414	500	1346
OCR Field Accuracy	13.8%	17.2%	0.0%	10.3%
Results With VRS				
Total Fields	500	500	500	1500
Fields w/Errors	165	140	498	803
OCR Field Accuracy	67%	72%	0.004%	46.4%
VRS Improvement Over Non-VRS Results	53.2%	54.8%	0.004%	36.1%

Table 2 – OCR Field Accuracy (VRS vs. Non-VRS)

Input Range

Kofax's VRS provides the following benefits in terms of the input range it can handle:

- VRS can handle the full range of document sizes for the scanners it supports. With VRS, scanner operators can load documents of any size or color into the document feeder and begin scanning, reducing document preparation time.
- VRS improves productivity by minimizing the time required to obtain scanner settings for optimal image clarity for documents of mixed sizes, colors, contrasts, and brightness. VRS minimizes the need for operators to scan and re-scan a document in order to achieve acceptable image quality.
- VRS includes notification capabilities to alert operators of problems with the scanner, paper jams, folded corners, or other problems. This makes it easier for the operator to quickly fix the problem and continue scanning, minimizing downtime.

Usability

Kofax's VRS provides the following benefits in terms of usability:

- VRS makes scanning simple and easy. Operators do not need special knowledge of scanning, resolution, or image enhancement. They just press the scan button, and VRS does the rest. Thus, VRS reduces required operator interaction with the scanner, saving time and labor costs.
- VRS makes it easy for users to configure optimal parameters or ranges for factors such as contrast, brightness, and noise. Users can fine-tune the settings right on the screen in real-time, with no need to re-configure the scanner itself, and no need to re-scan the document in order to validate the modified settings. The system's interface includes simple controls such as slider bars to make the process straightforward and less time-consuming.
- VRS makes it easy for users to establish different parameter profiles for different document types. Moreover, VRS can display optimum settings for any scanned image, and allow users to quickly modify them and establish new default settings for a particular document type. This reduces the time required to establish proper settings, which results in higher OCR accuracy and fewer re-scans downstream.

Speed

Kofax's VRS provides the following benefits in terms of speed:

- If capture speed is important, VRS can be configured to scan documents at lower resolutions, while still creating the best possible image at that resolution. After analyzing the image, VRS can convert the image into a higher-resolution image for better on-screen viewing.
- VRS can also be configured to save images at different compressions, allowing operators to strike the proper balance between image quality and small file sizes for faster retrieval, faster viewing, and reduced storage requirements.
- Although this feature was not tested by Doculabs, Kofax's internal testing of its VRS scanner adaptor hardware and technology revealed a 25-50% scanner performance boost over rated speeds.

CONCLUSION

To make a document capture application as efficient and productive as possible, organizations must consider more than just scanner speed. Image quality, input range, and usability are also critical factors for an efficient operation. Maximizing all of these factors is the key to minimizing manual processing, improving recognition accuracy, minimizing errors and manual data entry, and reducing time-consuming re-scanning and exception handling.

In all of these areas, Kofax's VRS delivers. VRS provides a number of features that improve image quality, and offers the flexibility to configure image appearance and compression characteristics to meet the specific needs of the application. The system is extremely simple to set up and use, minimizing the time and effort required to establish settings for optimal image quality while allowing scanner operators to make quick modifications on the fly if required.

Thus, VRS is a straightforward way for organizations to improve their image quality across a wide range of documents types and sizes, ultimately improving their "effective throughput": the overall efficiency of their scanning operations. With the wide range of production scanners and mid- to low-volume scanners that VRS supports, the solution makes sense for a diverse range of document capture applications.

APPENDIX: ABOUT DOCULABS

Doculabs is an independent industry analyst firm specializing in e-Business technologies. We are guided by the principle that both end-users and vendors benefit from impartial feedback about product strengths and limitations to make both strategic and tactical business decisions. Founded in Chicago in 1993, Doculabs was one of the first industry analyst firms to ground its research, end-user and vendor advisory services in unbiased, reality-based product assessment results.

We use our benchmark methodology, combined with trend and market analysis, to help our clients make the right technology investment decisions and to help vendors attack new market opportunities. Doculabs' core research methodology builds a truly credible information bridge between end-user business needs and vendor product development.

Doculabs has a growing staff of analysts dedicated to research, advisory services, and market analyses. We specialize in emerging technology solutions that drive e-business, including technologies for e-content, e-commerce, e-fulfillment and delivery, and e-business infrastructure components.

Doculabs' analysts are regular speakers at industry conferences, and continue to author numerous articles in industry journals such as *InformationWeek*, *Network Computing*, *eWeek*, *Windows NT*, *KM World*, *KM Magazine*, *Group Computing*, *High-Volume Printer*, and *Transform*, to name a few.

Doculabs' service offerings include advisory and consulting services, assessment services, and publications (reports and annual subscriptions). Because Doculabs does not resell products, we remain completely objective in our research and our recommendations.

For more information about Doculabs, visit our Web site – <http://www.doculabs.com>.



1201 West Harrison Street, Third Floor
Chicago, IL 60607
312-433-7793
www.doculabs.com