DIGITECH SYSTEMS

PaperVision® Capture Desktop

User Guide

PaperVision Capture Desktop Release 75

February 2012

Digitech Systems, Inc.
8400 E. Crescent Parkway, Suite 500
Greenwood Village, CO 80111
Phone: 303.493.6900 Fax: 303.493.6979
www.digitechsystems.com

Table of Contents

Table of Contents

Chapter 1 - Introduction

PaperVision Capture Desktop performs document scanning, indexing, barcoding, zonal OCR, image processing, and full-text OCR. The application provides an intuitive graphical interface that enables you to scan, index, export, and FTP documents; define barcode, image processing, and OCR zones; and, execute automated processing tasks such as image processing, barcoding, and zonal/full-text OCR. Various full-text OCR outputs are available including PDF, PaperFlow™, PaperVision® Enterprise, and text outputs. Batches can be exported in various formats that are compatible with external applications and web sites, such as PaperFlow, ImageSilo®, and Microsoft® SharePoint.

System Requirements

The following tables outline the minimum software requirements and recommended hardware requirements for the PaperVision Capture Desktop application. PaperVision Capture Desktop supports more than 300 ISIS-compatible scanners. Alternatively, PaperVision Capture Desktop also supports the use of TWAIN drivers provided by your scanner manufacturer.

Minimum Software Requirements	
Operating Systems	Windows XP SP3 or later (both 32- and 64-bit operating systems supported)
.NET Framework	Version 3.5 SP1 or later
Windows Installer	Version 3.1 or later

Recommended Hardware Requirements	
Processor	Current processor technology is recommended (typically, not older than four years).
RAM	2 GB
Hard Disk Space	1200 MB
Minimum Screen Resolution	1024 x 768

Maximum Image Sizes

This section outlines the approximate limits in image sizes that can be imported into PaperVision Capture Desktop and processed through full-text OCR, zonal OCR, and image processing tasks. The Thumbnails windows can handle substantially larger images. Additionally, images only stored in memory or simply ingested by PaperVision Capture Desktop (therefore not viewed in Thumbnails windows or processed through the Full-Text OCR, Zonal OCR, or Image Processing tasks), can also be significantly larger in size.

DISCLAIMER – PLEASE READ!

These dimensions are provided only as estimates to identify size limits in importing, viewing, and processing images in PaperVision Capture Desktop. Variations in technical environments may cause maximum image sizes to fluctuate across systems.

Maximum Image Sizes (in Pixels)	
Stored Images	10,000 x 10,000* * These dimensions can be greater in bitonal images.
Thumbnails	32,768 x 32,768
Image Processing	10,000 x 10,000* * These dimensions can be greater in bitonal images.
Full-Text OCR and Zonal OCR	32,000 (width) x 24,000 (height) *These dimensions will vary at different resolutions.

Terminology

The following PaperVision Capture Desktop terminology will be used throughout this user guide:

Batch

A batch is a collection of documents and their associated index name-value pairs.

Default Batch Path

The Default Batch Path is the local storage directory where PaperVision Capture Desktop stores all captured images.

Detail Sets

Detail sets expand on the capabilities of standard index fields because they define "many-to-one" relationships, which allow multiple sets of field data to reference a single document. In a many-to-one relationship, an index field contains a value that references another field or set of fields that contain unique values.

Document

A document is the equivalent of a file folder within a filing cabinet. A document holds all of the pages for a given set of index values.

Image

An image is a visual representation of a picture or graphic, such as an electronic file with the extensions .bmp, .jpg, or .tif.

Index

An index is a value that users apply to a document for reference and retrieval.

Job

A job is a defined process comprised of one or more manual or automated tasks (e.g., scanning, indexing, and exporting) through which batches are processed. Each batch is unique by name.

On the Fly Processing

Image processing, barcode reading, and zonal OCR can be executed while you scan documents. You can change the order in which these tasks are executed; however, image processing must be executed first or last.

Page

One or more images (files with extensions bmp, .jpg, and .tif,) comprise a single page within a document. For example, a page can include the originally captured image and a manipulated image after noise removal.

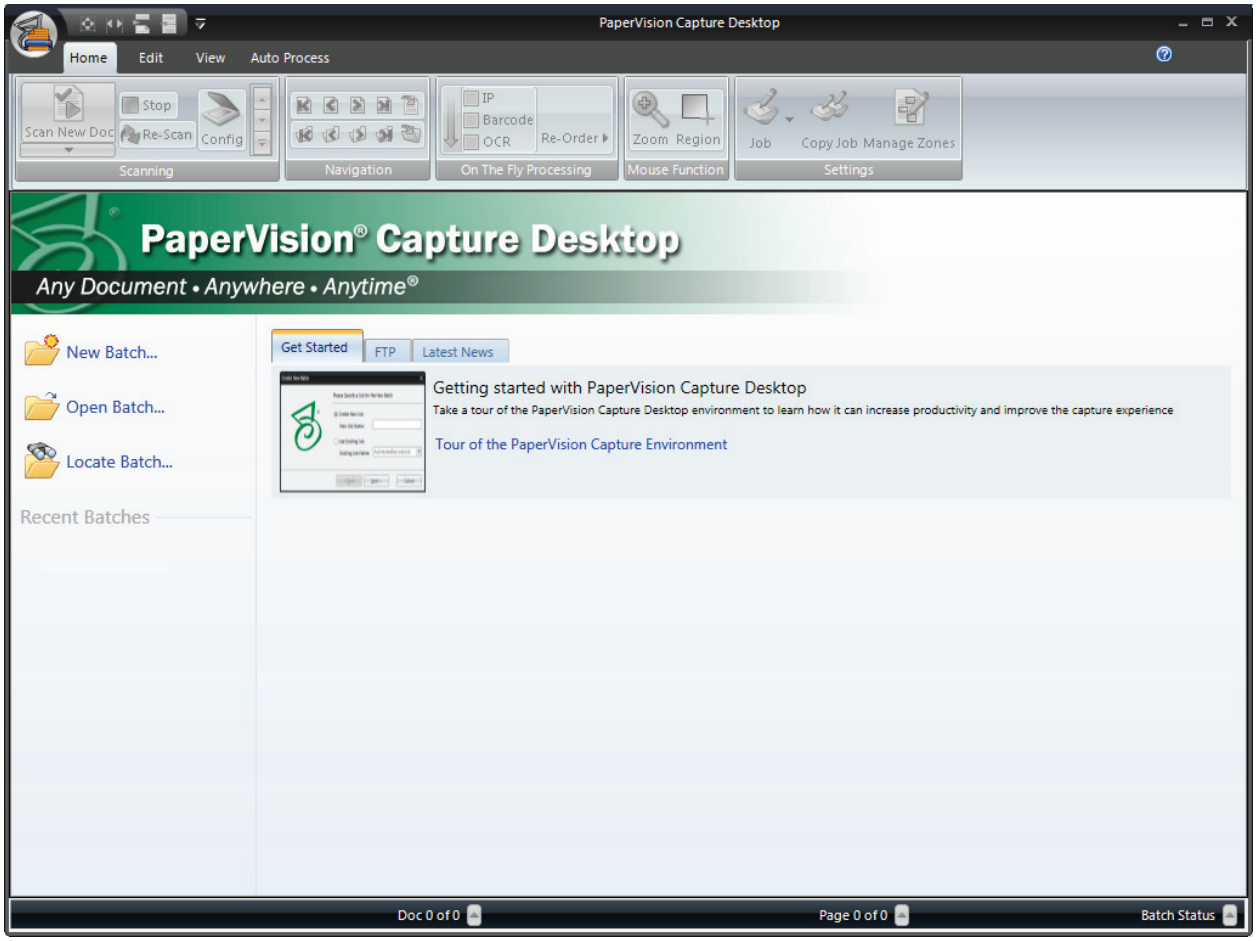
PaperVision Capture NOW!

PaperVision Capture NOW! is installed as a separate application in conjunction with PaperVision Capture Desktop. The PaperVision Capture NOW! application enables you to quickly scan and convert images to full-text PDF, image-only PDF, or multi-page TIFF files. You can execute operations in PaperVision Capture NOW! in the standard graphical user interface or via command line interface.

Starting the Application

To launch the PaperVision Capture Desktop application, double-click the PaperVision Capture Desktop icon located on your desktop. Alternatively, you can select **Start > Programs > Digitech Systems > PaperVision Capture Desktop**. The Start page will appear, displaying the batch operations and the Get Started, FTP, Guidance and Resources, and Latest News tabs.

> **Note**:
>
> The PaperVision Capture NOW! graphical user interface application (launched via the PaperVision Capture NOW! icon) and PaperVision Capture Desktop cannot be run simultaneously on the same machine.



Start Page

Get Started

Within the Get Started tab, you can insert links to your company's intranet, web site, training forums, etc. To launch a video (.exe file) of PaperVision Capture Desktop's features and operations, click the **Tour of the PaperVision Capture Environment** link. To view the video, Adobe Flash Player 9 or later must be installed on your workstation. For a free download of the latest Adobe Flash Player, visit **http://www.adobe.com**, and navigate to **Download > Adobe Flash Player**.

Creating a New Job and Batch

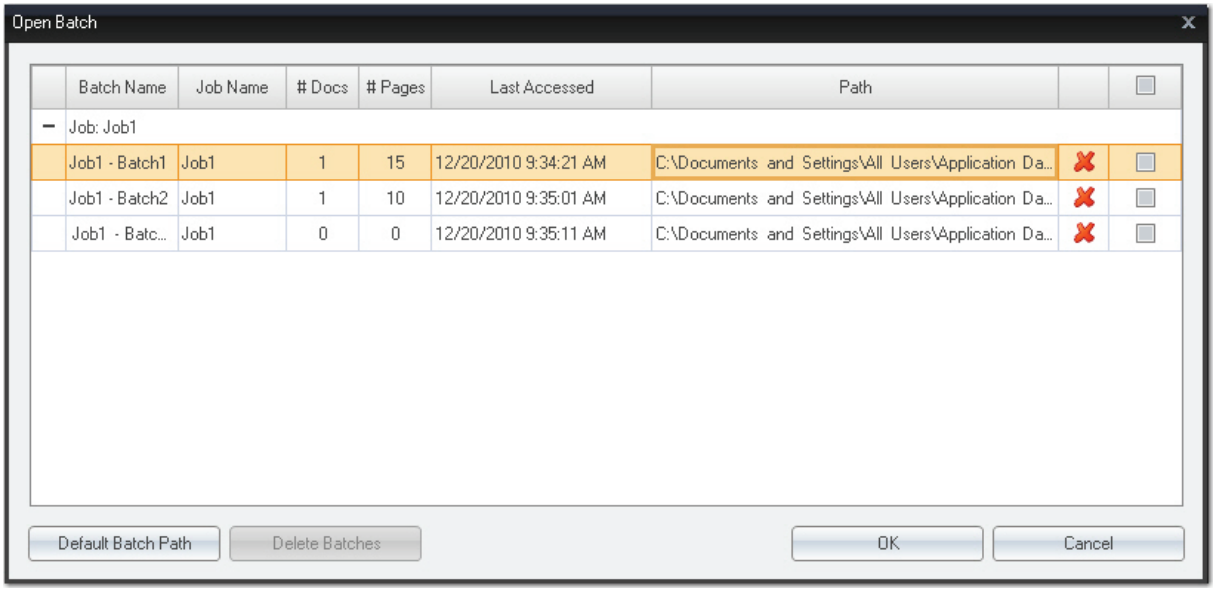
In the Start Page, the **New Batch**  option launches the New Batch Wizard. The wizard helps you create new jobs and batches to process your documents. See the **New Batch Wizard** section for more information.

Opening an Existing Batch

In the Start Page, you can open any existing batch.

To open an existing batch:

1. Click the **Open Batch**  icon**.** The **Open Batch** grid opens.



Open Batch

> **Note**:
> Alternatively, you can highlight the batch in the **Recent Batches** list to open it.

2. Select the batch.
    - Batches can be sorted by batch name, job name, number of docs/pages, directory path, and last opened date.
    - Click a column header to sort the list in ascending or descending order, and the sort order will be maintained for the duration of the user's session.
    - To reorder the columns, drag the column header to another location in the grid.
3. After selecting the batch, click **OK**. The **Home Page** appears, where you can execute scanning, indexing, and other primary operations. See the **Home Page** chapter for detailed information on all operations.

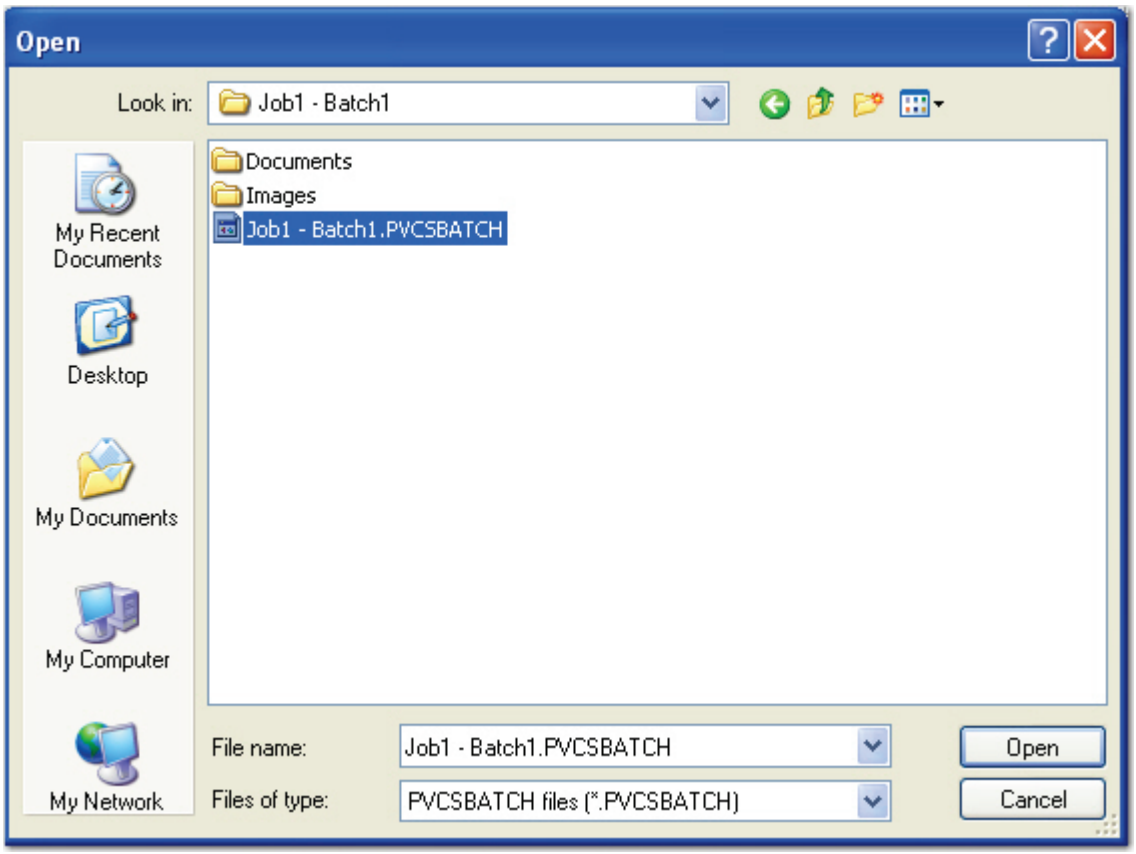
> **Note**:
>
> If another batch is currently open, you will be prompted to close the current batch in order to open the new batch.

Locating a Batch

In the Start Page, the Locate Batch operation enables you to find and open batches that are stored in locations such as your local machine, network, or shared drive.

To locate a batch:

1. Click the **Locate Batch**  icon. The Windows **Open** dialog appears.

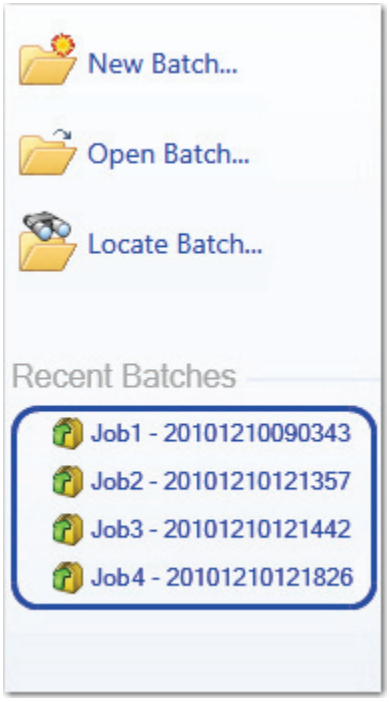


Open

2. Find the appropriate directory where the **.PVCSBATCH** file resides.
3. Highlight the **.PVCSBATCH** file, and then click **Open**. The located batch will appear in the **Recent Batches** list.

Recent Batches

From the PaperVision Capture Desktop Start Page, you can view the most recent batches that you have created.



Recent Batches

To open a batch you have recently worked on but not yet processed, highlight the batch in the **Recent Batches** list. The Home page will open, where you can resume scanning, indexing, etc. documents. See the **Home Page** chapter for specific information on scanning, indexing, and other document operations.

FTP

The FTP operation securely transfers your documents, images, and associated index values to an FTP site. Before processing FTP, ensure the required FTP settings have been configured in the PaperFlow or ImageSilo/PVE XML custom code generator wizard (**Auto Process page** > **Wizard** > **PaperFlow** or **ImageSilo/PVE XML**). You can execute FTP from the Start Page, PaperVision Capture Desktop Button, or from the Auto Process page. For detailed instructions on the FTP operation, see the section on **FTP** in the Auto Process chapter.

> **Note**:
>
> For more information on PaperFlow or ImageSilo/PVE XML export configuration, see the **Exports** chapter.

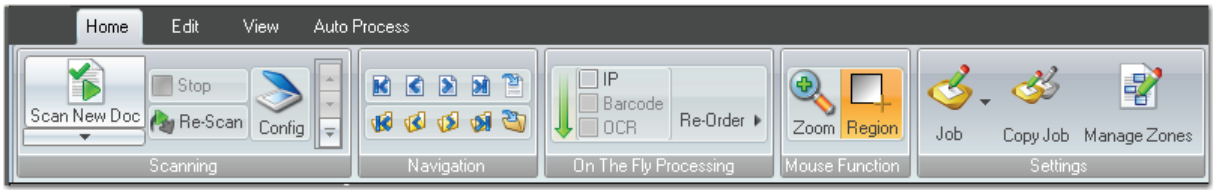
Latest News

The Latest News tab can hold your company's latest RSS feeds to provide updated information on your company's products and technologies. To subscribe to an RSS feed, enter the URL address, and then click the **Enable RSS Feed** button.

Navigating the Workspace

PaperVision Capture Desktop features an intuitive ribbon interface. Within each tab (page), a large toolbar displays graphical representations of related operations, such as the Scanning operations toolbar group in the Home page. To navigate to another page, simply click on the tab name.

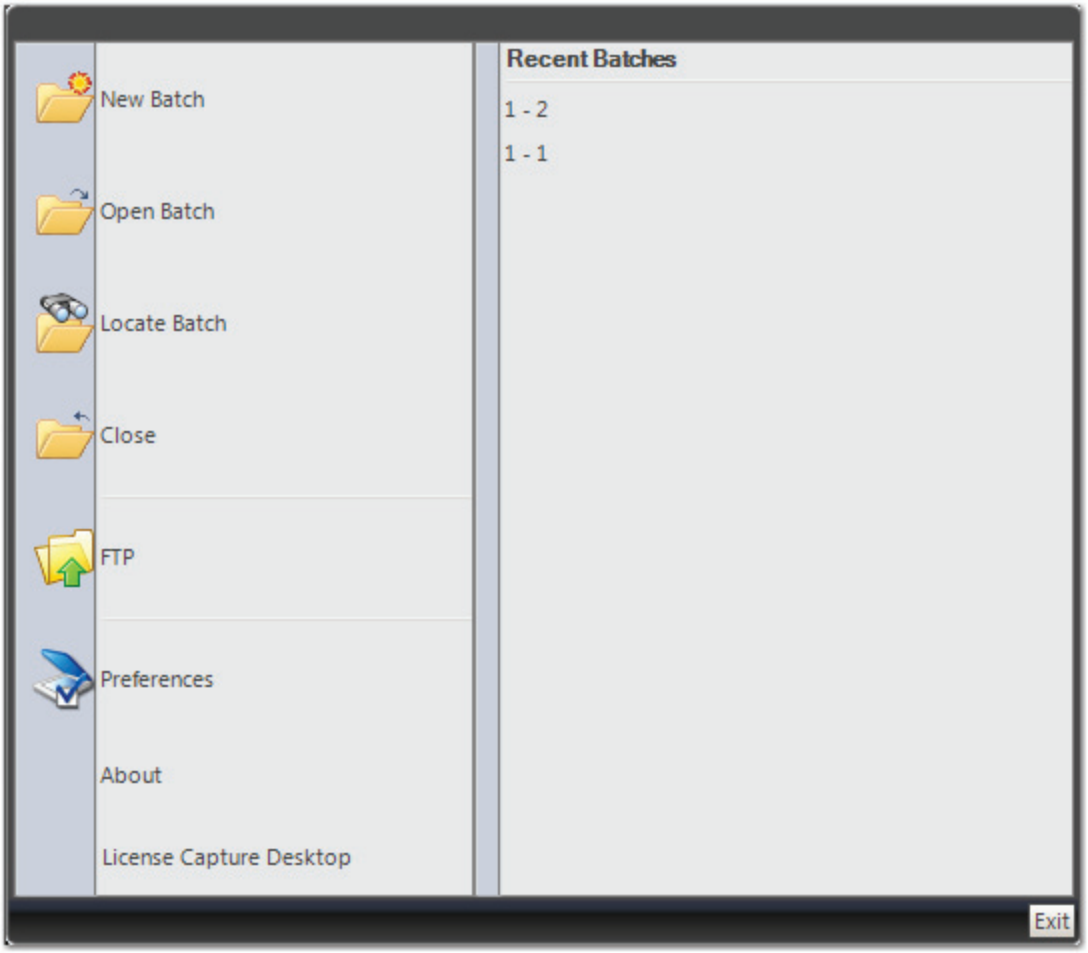
- The Home page contains the application's primary scanning, indexing, document/page navigation, zonal, on the fly processing, and mouse function operations.
- The Edit page contains the document and page manipulation operations.
- The View page contains display, scaling, and zoom operations.
- The Auto Process page contains zonal OCR, barcoding, image processing, full-text OCR, and export operations.



PaperVision Capture Desktop Ribbon

PaperVision Capture Desktop Button

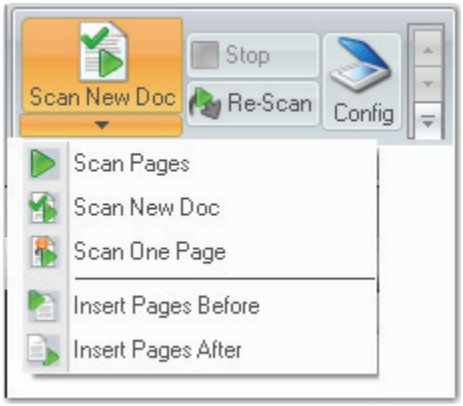
The PaperVision Capture Desktop Button is located in the upper left side of the application and contains all batch operations and user preferences, such as system, display, and import preferences. For detailed instructions on all operations, see the section on **PaperVision Capture Desktop Button Operations** in this chapter.



PaperVision Capture Desktop Button Operations

Toolbar Groups and Drop-Down Menus

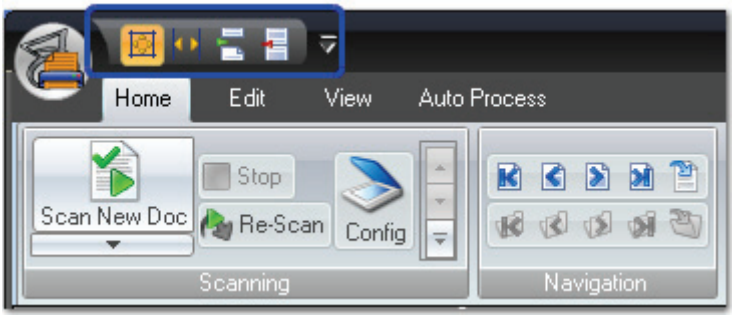
Within each page, related operations are grouped together in the main toolbar. Primary operations, such as the Scan New Document operation, contain drop-down menu options such as Scan Pages, Scan One Page, and Insert Pages Before/After. Once you select the Scan New Document, Scan Pages, or Scan One Page operation, your selection will be retained as the main operation in the Scanning toolbar group.



Scanning Toolbar Group and Drop-Down Menu

Quick Access Toolbar

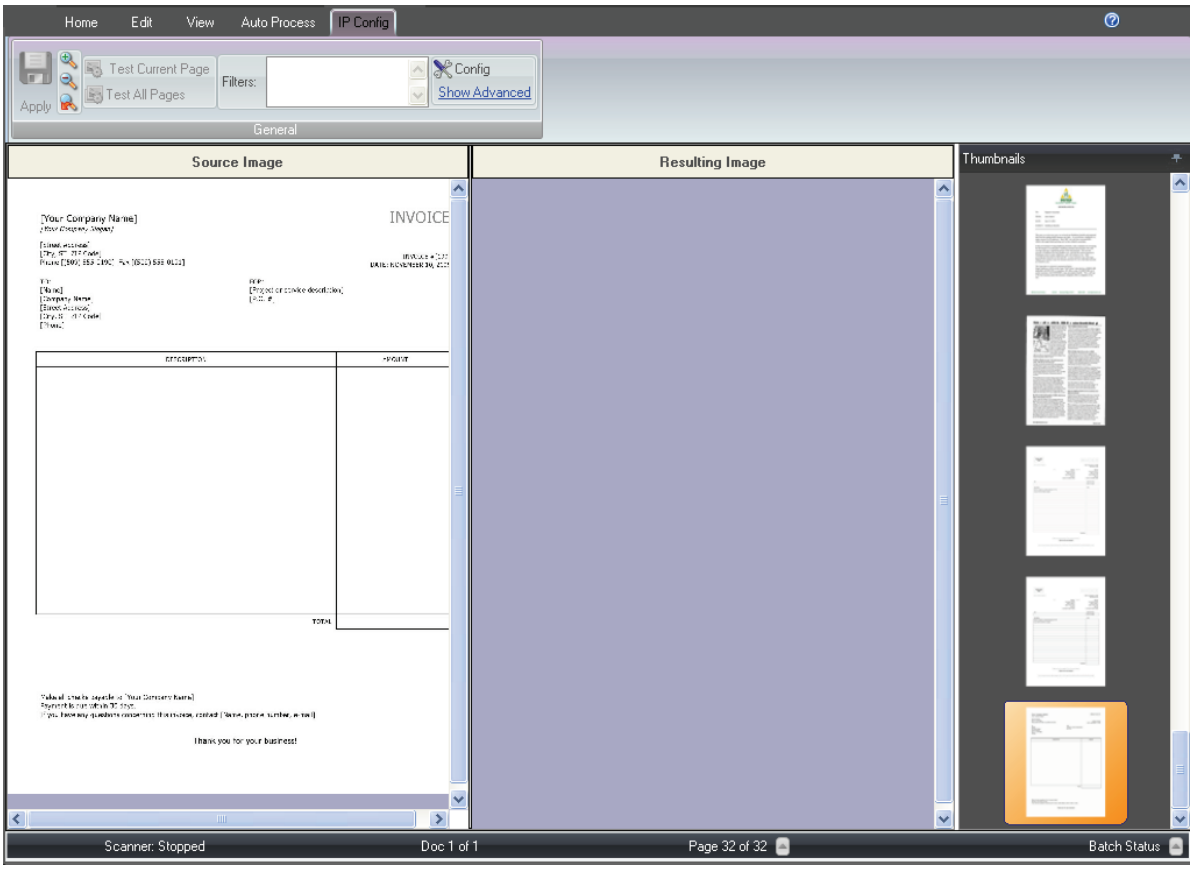
The Quick Access Toolbar contains a set of commands that are independent of the currently displayed page. You can customize this toolbar with commands that you commonly use in the application.



Quick Access Toolbar

Context Sensitive Pages

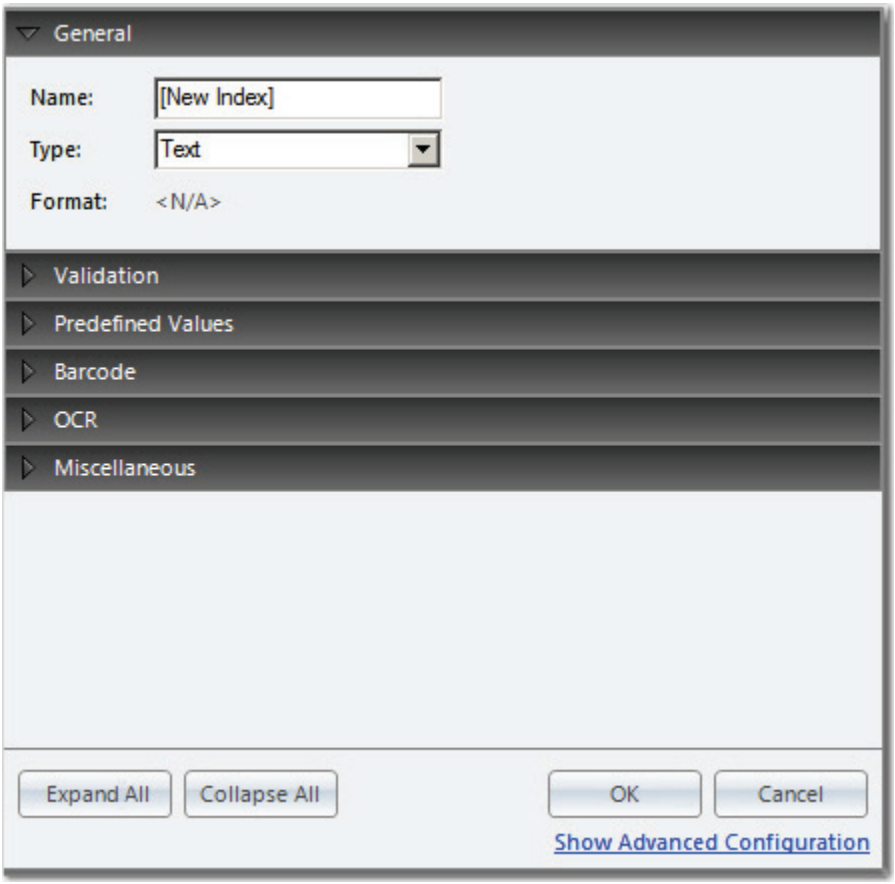
Context Sensitive pages appear when you select a certain type of operation in the Home and Auto Process pages. For example, when you select the **Config** icon in the IP toolbar group in the Auto Process page, the IP Config context sensitive page opens where you can configure, test, and apply IP filters.



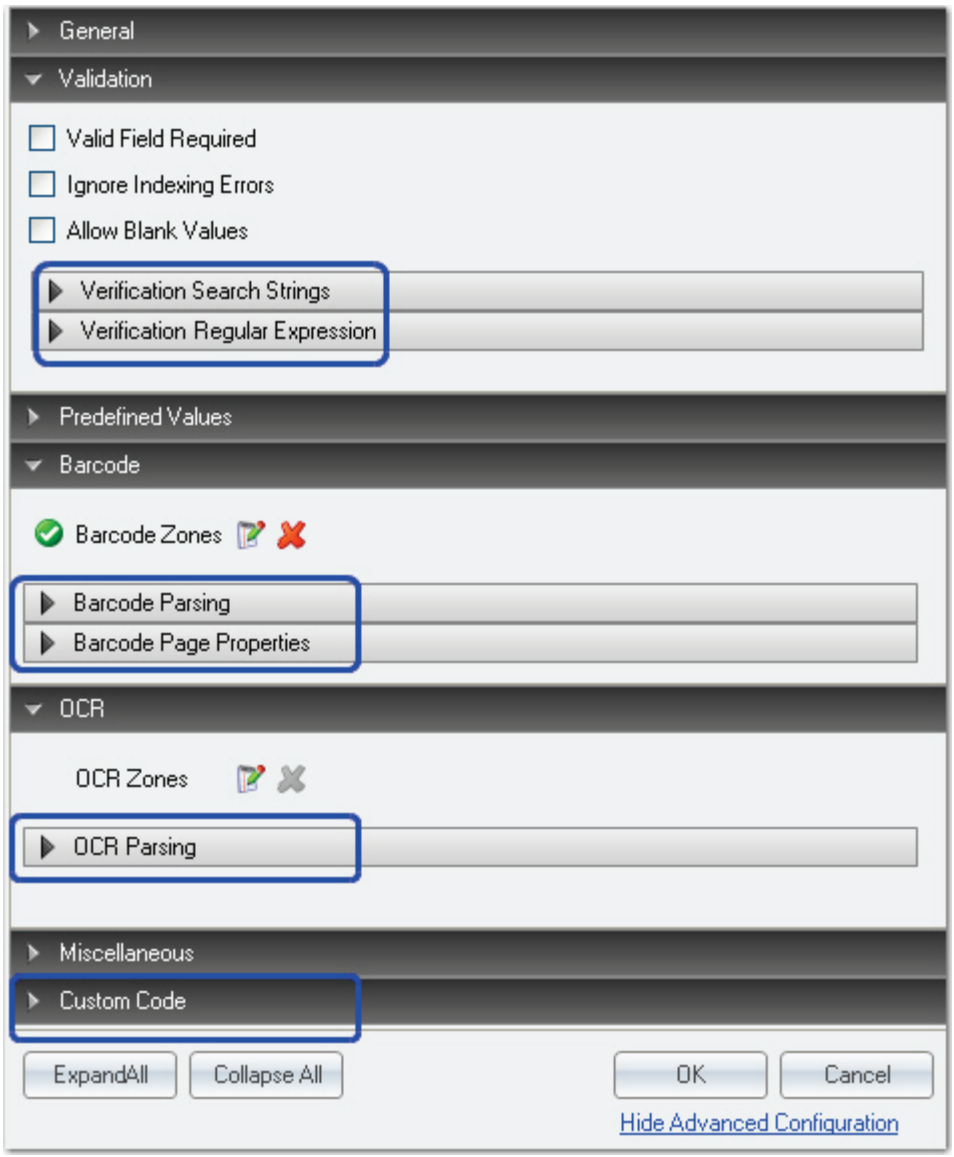
IP Config Context Sensitive Page

Expandable and Collapsible Properties

When you assign indexing properties (via the Indexing window in the Home page), expandable and collapsible panel menus will appear. Additionally, these panel menus will appear when you modify job settings in the Home Page. The Show/Hide Advanced Configuration link in the lower right corner enables you to toggle between basic and advanced settings for index and job properties. If you generally assign only basic indexing properties, such as index name, type, format, predefined values, etc., you may not need to view the advanced settings. Advanced index properties include index verification search strings and regular expressions; barcode and OCR parsing; and, custom code events. Advanced job properties include custom code configuration and Merge Like Documents settings.



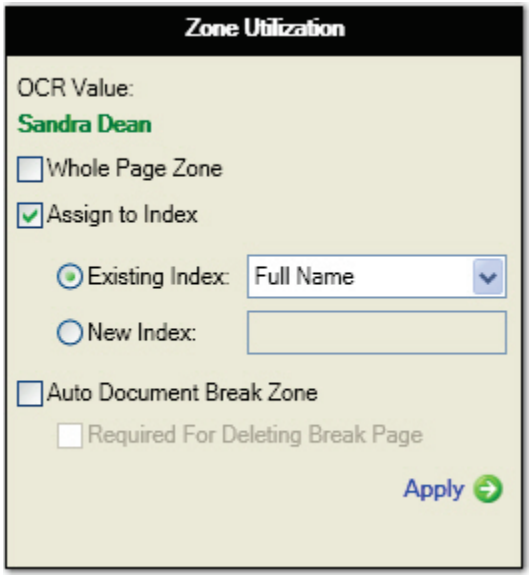
Basic Index Properties



Advanced Index Properties

Flyout Windows

When you apply zones (via the Manage Zones operation) or regions (via the Region mouse function) to images in the Home page, a flyout window immediately appears with barcode values, OCR values, and/or zone utilization options. Zone utilization options enable you to assign barcode and OCR zones to automatically break documents or assign to indexes. You can also apply zoom zones on images with the Region mouse function.



Zone Utilization Flyout Window

Status Bar

At the bottom of the application, the Status Bar provides the status of current operations in each screen (Scanner: Stopped; Automated Process: OCR; etc.). The Status Bar also shows the current document and page location; image information; and batch status.

Status Bar

To view the image information, such as its file name, dimensions, size, etc., click the arrow next to the **Page** information in the Status Bar.

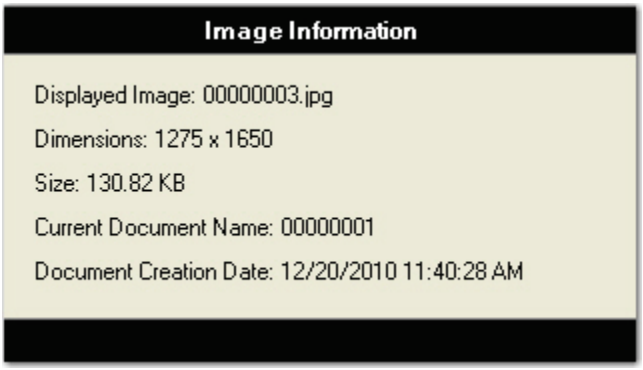
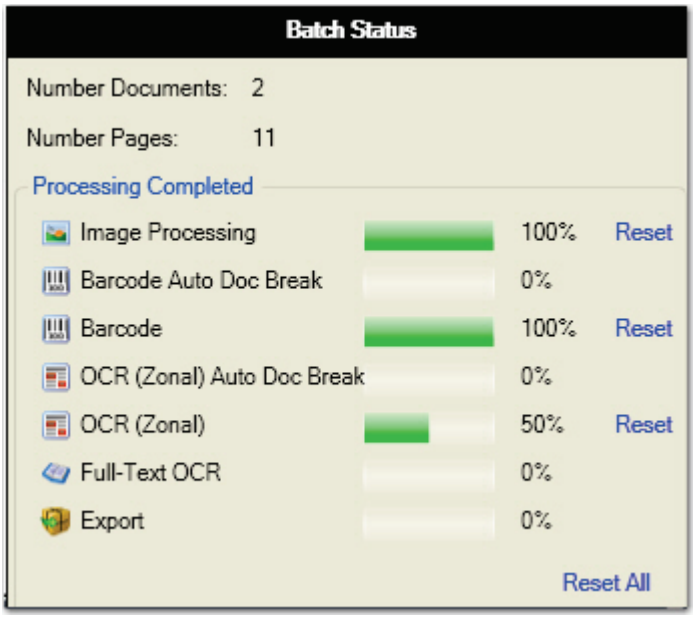


Image Information

To view the batch information, including the number of documents and pages comprising the batch; automated processing status, etc., click the arrow next to the **Batch Status** information in the Status Bar. To reset an automated process, click **Reset**; to reset all automated processes, click **Reset All**.

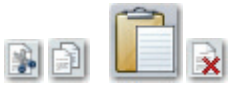


Batch Status Flyout Window

Right-Click Operations

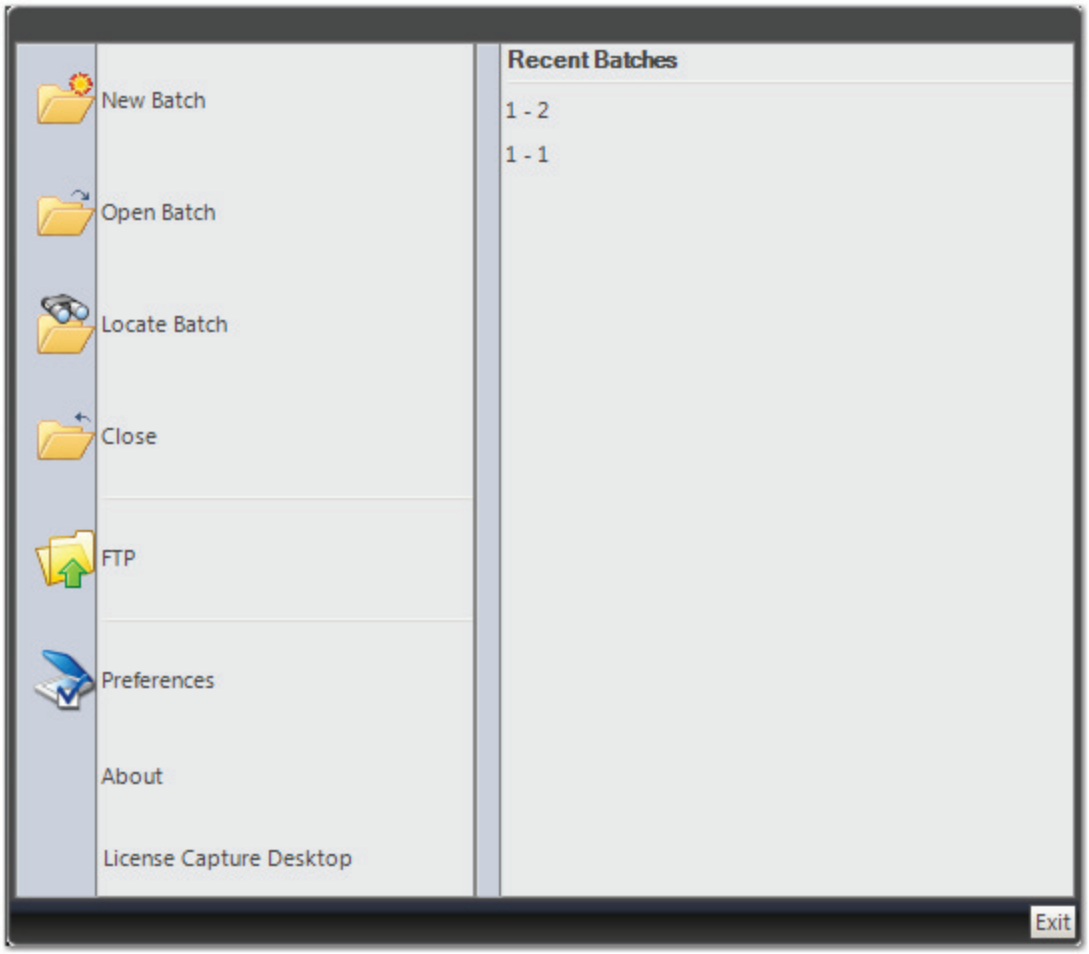
You can execute a variety of operations when you right-click on an image in the main viewing window in the Home, Edit, View, and Automated Processing Tasks pages. The following table displays each operation's default hot keys (if applicable), toolbar icon, and main page where the operation can be executed.

Right-Click Operations	
Reset Page 	Resets the page to its original view Page: Edit Hot Key: Ctrl+T
Rotate Image 	Rotates the image 90 degrees clockwise Page: Edit Hot Key: R
Rotate and Save Image 	Rotates and saves one or multiple images in one document or across multiple documents Page: Edit
Scale to Height, Width, Window 	Fits images to the height, width, or both height and width of the screen. Page: View
Copy Document 	Copies all pages and appends the new document after the selected document Page: Edit Hot Key: Y
Delete Document 	Deletes the current document and its associated images Page: Edit Hot Key: Ctrl+Delete
Insert Document Break 	Makes your selected page the first page of the new document Page: Edit Hot Key: Ctrl+Insert

Right-Click Operations (continued)	
Remove Document Break 	Removes a previously-inserted document break Page: Edit Hot Key: Ctrl+R
Cut, Copy, Paste, Delete Pages 	Cuts, copies, pastes, or deletes the current page Page: Edit Hot Keys: • Cut: Ctrl+X • Copy: Ctrl+C • Paste: Ctrl+V • Delete Page: Delete
Invert Page Polarity 	Reverses the black text on white background to white text on black background (or vice versa) Page: Edit
Invert and Save Page Polarity 	Page: Edit

PaperVision Capture Desktop Button Operations

Batch operations, such as opening/closing, importing, and creating new batches, are located within the PaperVision Capture Desktop Button in the upper left side of the application. User Preferences are also located under the PaperVision Capture Desktop Button. You can modify user preferences including display, hot keys, importing, indexing, system, and confirmation messages.



PaperVision Capture Desktop Button Operations

Creating a New Batch

From the PaperVision Capture Desktop Button, you can create a new batch. See the **New Batch Wizard** section for more information.

Opening an Existing Batch

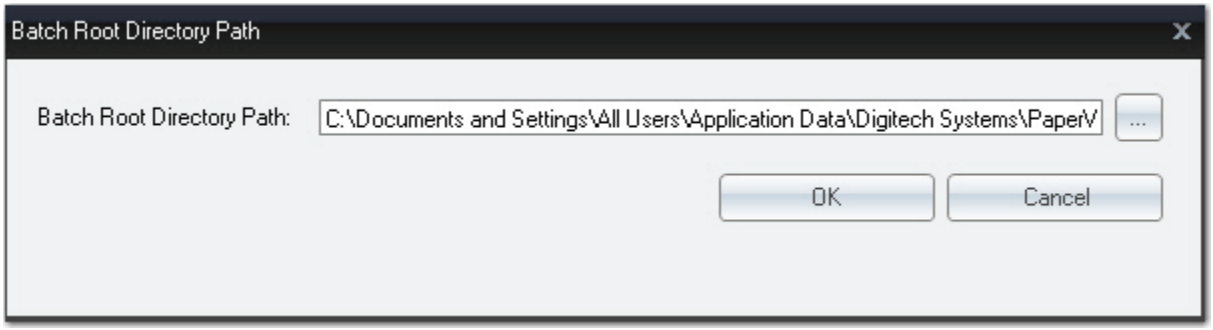
From the PaperVision Capture Desktop Button  , you can open a batch that was recently created. For detailed instructions on opening existing batches, see the previous section on **Opening Existing Batches**.

Changing the Default Batch Path

You can change the default batch path where your new batches are stored.

To change the default batch path:

1. In the **Open Batch** dialog box, click the **Default Batch Path** button. The **Batch Root Directory Path** dialog box appears.



Batch Root Directory Path

2. To change the default batch path, enter the location, or click the ellipsis button.
3. In the **Browse for Folde**r dialog, select the new directory, and then click **OK**.
4. To confirm the change, click **OK**.

Deleting Batches

To delete batches:

1. In the last column of the **Open Batch** dialog box, insert check marks for one or more batches.

> **Note**:
>
> To delete only one batch, highlight the batch, and then select the **Delete**  icon.

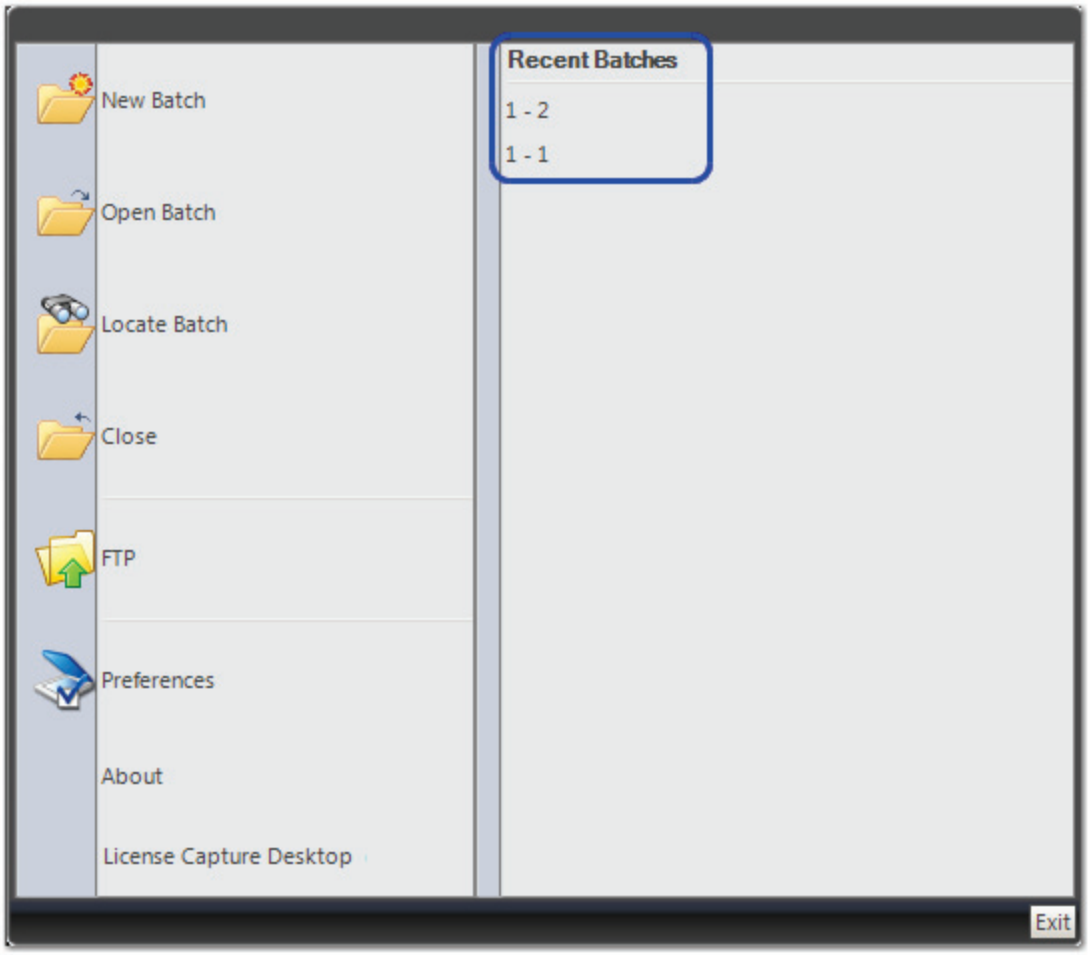
2. Click the **Delete Batches** button.
3. Click **Yes** to confirm the deletion.

Opening Recent Batches

From the PaperVision Capture Desktop Button, you can view the ten most recent batches that you have created. If applicable, you can also open a recent batch.

To view and open a recent batch:

1. Select the **PaperVision Capture Desktop Button**, and the **Recent Batches** list appears in the right column.



Recent Batches

2. To open a recent batch, select the batch from the **Recent Batches** list.
3. If another batch is currently open, you will be prompted to close the current batch and open the selected batch. If you select **Yes**, The **Home** page will appear. See the **Home Page** chapter for more information.

Locating a Batch

The Locate Batch operation in the **PaperVision Capture Desktop Button**  enables you to find and open batches that are stored in a location other than your local machine, such as a network or shared drive. See the section on **Locating a Batch** for more information.

Closing a Batch

You can close a current batch and work on it at a later time. The ten most recent batches that you close will remain in the Recent Batches list.

To close a batch:

1. Click the **PaperVision Capture Desktop Button**  , and then click the **Close Batch**  icon.
2. Click **Yes** to close the current batch. The main **Start** page appears once again.

FTP

The FTP operation securely transfers your documents, images, and associated index values to an FTP site. Before processing FTP, ensure the required FTP settings have been configured in the PaperFlow or ImageSilo/PVE XML custom code generator wizard (**Auto Process page** > **Wizard** > **PaperFlow** or **ImageSilo/PVE XML**). You can execute FTP from the Auto Process page, PaperVision Capture Desktop Button, or from the Start page. For complete instructions on processing FTP folders, see the previous section on **FTP**.

> **Note**:
>
> For more information on PaperFlow or ImageSilo/PVE XML export configuration, see the **Exports** chapter.

Preferences

Click the **PaperVision Capture Desktop Button**  , and then click the **Preferences**  icon to view all user preferences for the application's indexing and import operations, display and system settings, confirmation messages, etc. For more detailed information on all user preferences, see the **Preferences** section.

About

Click the **PaperVision Capture Desktop Button**  , and then select **About** to display the product name, current released version of the application, license information, and copyright information.

Licensing PaperVision Capture Desktop

If you are running PaperVision Capture Desktop in demonstration mode and would like to add a purchased license, click the PaperVision Capture Desktop Button  , and then select **License Capture Desktop**. The PaperVision Capture Desktop License Wizard will launch and help you add a purchased license. For more information on licensing, see the accompanying **PaperVision Capture Desktop Installation and Getting Started Guide**.

> **Note**:
>
> This option is not available if you have already activated a purchased license.

Exiting the Application

To close the application, select the **PaperVision Capture Desktop Button**  , and then click the **Exit** button. If you have any unsaved changes, you will be prompted to save those changes before the application closes. Alternatively, you can click the upper right "X" in the application.

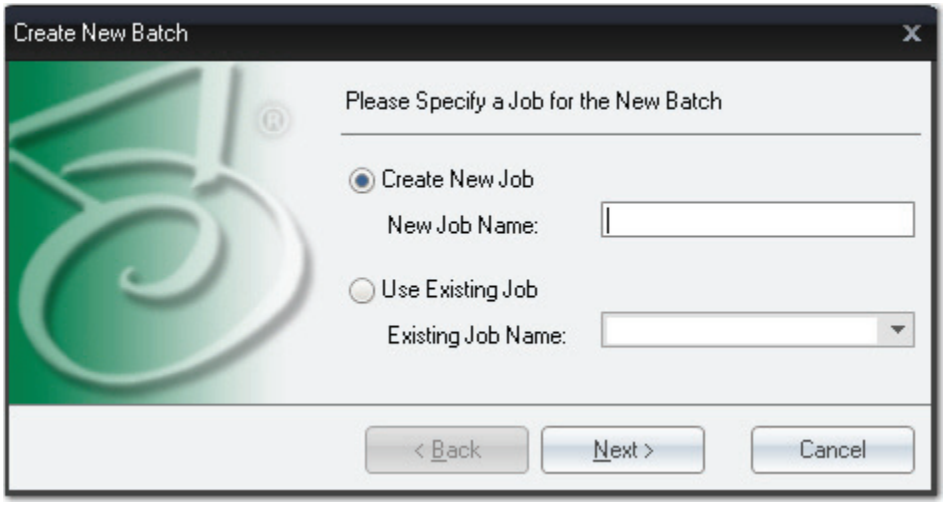
If the **Exit to System Tray** option in the System Preferences screen is enabled, you can also close PaperVision Capture Desktop by right-clicking the PaperVision Capture Desktop  or PaperVision Capture NOW!  icon in the System Tray. After you right-click the application icon, select **Shut Down PaperVision Capture Desktop** (or **Shut Down PaperVision Capture NOW!**).

New Batch Wizard

Once you have successfully installed and launched PaperVision Capture Desktop, you must first create a new job and batch from which to execute scanning, indexing, barcoding, exporting, etc., operations. As described in the first chapter, a job is a defined process comprised of one or more manual or automated tasks (e.g., scanning, indexing, and exporting documents) through which batches are processed. A batch is a collection of documents and their associated index name-value pairs and statistics that are moved as a logical unit of work through a job.

You can create new jobs and batches with the New Batch operation that is accessible in the Start page. Alternatively, you can create new jobs and batches from the PaperVision Capture Desktop Button located on the upper left side of the Welcome or Home pages. In the New Batch Wizard, you can also select an existing job from which to create a new batch of documents. Once you create a job, you can save the job in a local directory that can be used repeatedly by any user who logs into the computer to use PaperVision Capture Desktop.

> **Note**:
>
> Batch names must be unique.



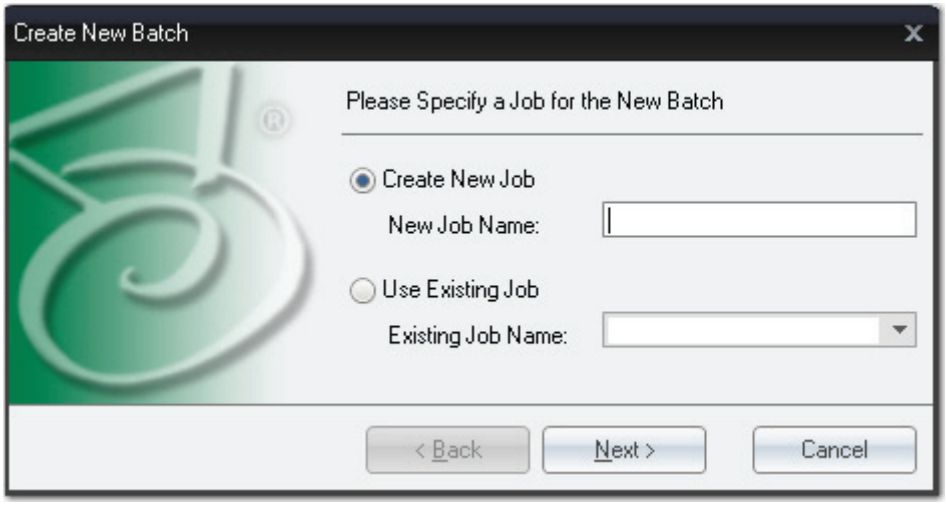
New Batch Wizard

Creating a New Batch

Creating new jobs and batches from the PaperVision Capture Desktop Button or the Start page involves a few simple steps outlined below.

To create a new job and batch:

1. Click the **PaperVision Capture Desktop Button**  , and then click the **New Batch**  icon. The **Create New Batch - Specify Job Name** dialog box opens.



Create New Batch - Specify Job Name

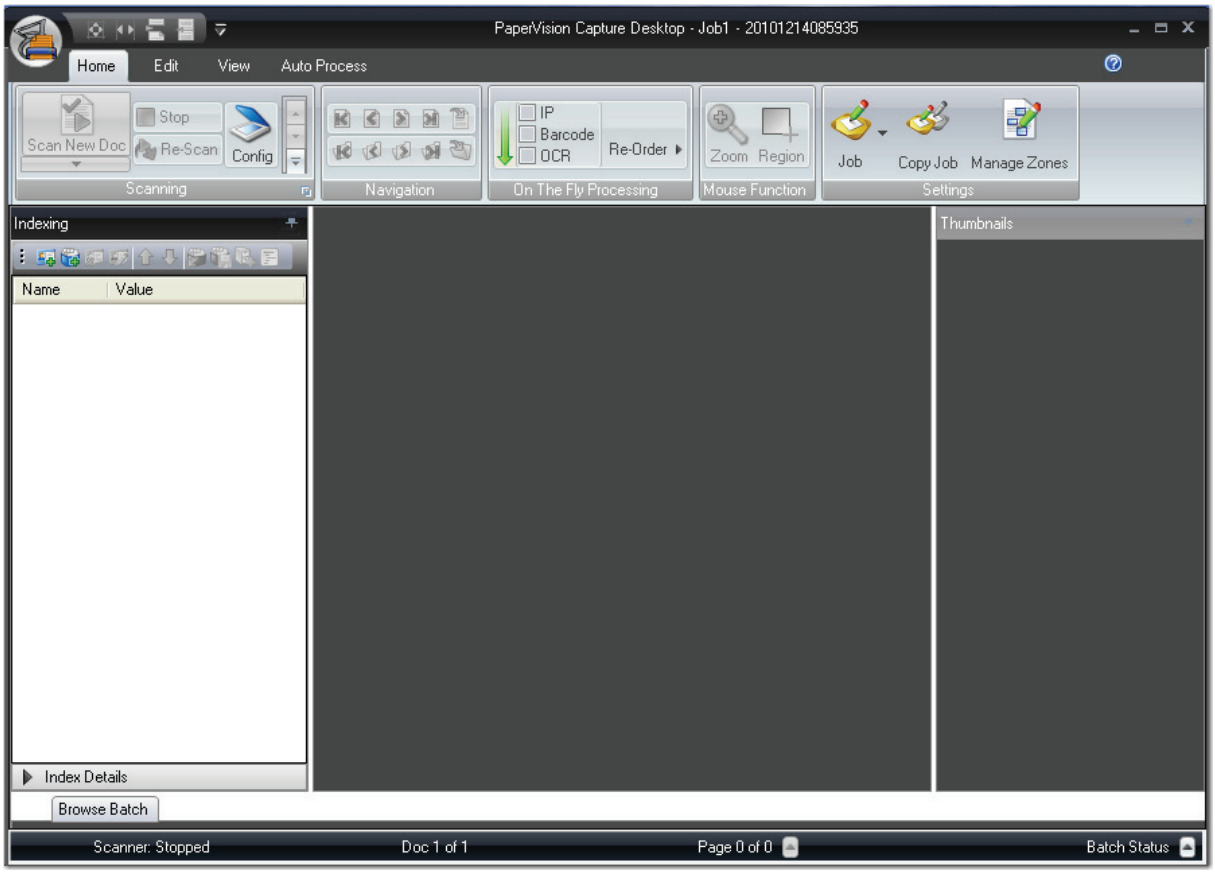
2. Select either the **Create New Job** or **Use Existing Job** option.
   - If you are creating a new job, enter its name in the **New Job Name** field.
   - If you are using a job that has already been created, select the job from the **Existing Job Name** drop-down list.
3. Click **Next**. The **Create New Batch - Specify Batch Name** dialog box appears.

Create New Batch - Specify Batch Name

4. Enter the **Batch Name**. By default, the batch name is formatted using the job name, current year, month, day, hour, minute, and second (<Job Name> - yyyymmddhhnnss).

> **Note**:
>
> The batch name must be unique.

5. Click **Finish**. The **Home** page appears, where you can configure your scanner; scan and index documents; enable on the fly processing; etc. See the **Home Page** chapter for more information.



Home Page

> **Note**:
>
> If you are running PaperVision Capture Desktop in demonstration mode, "Demonstration" will appear in the title bar of the application.

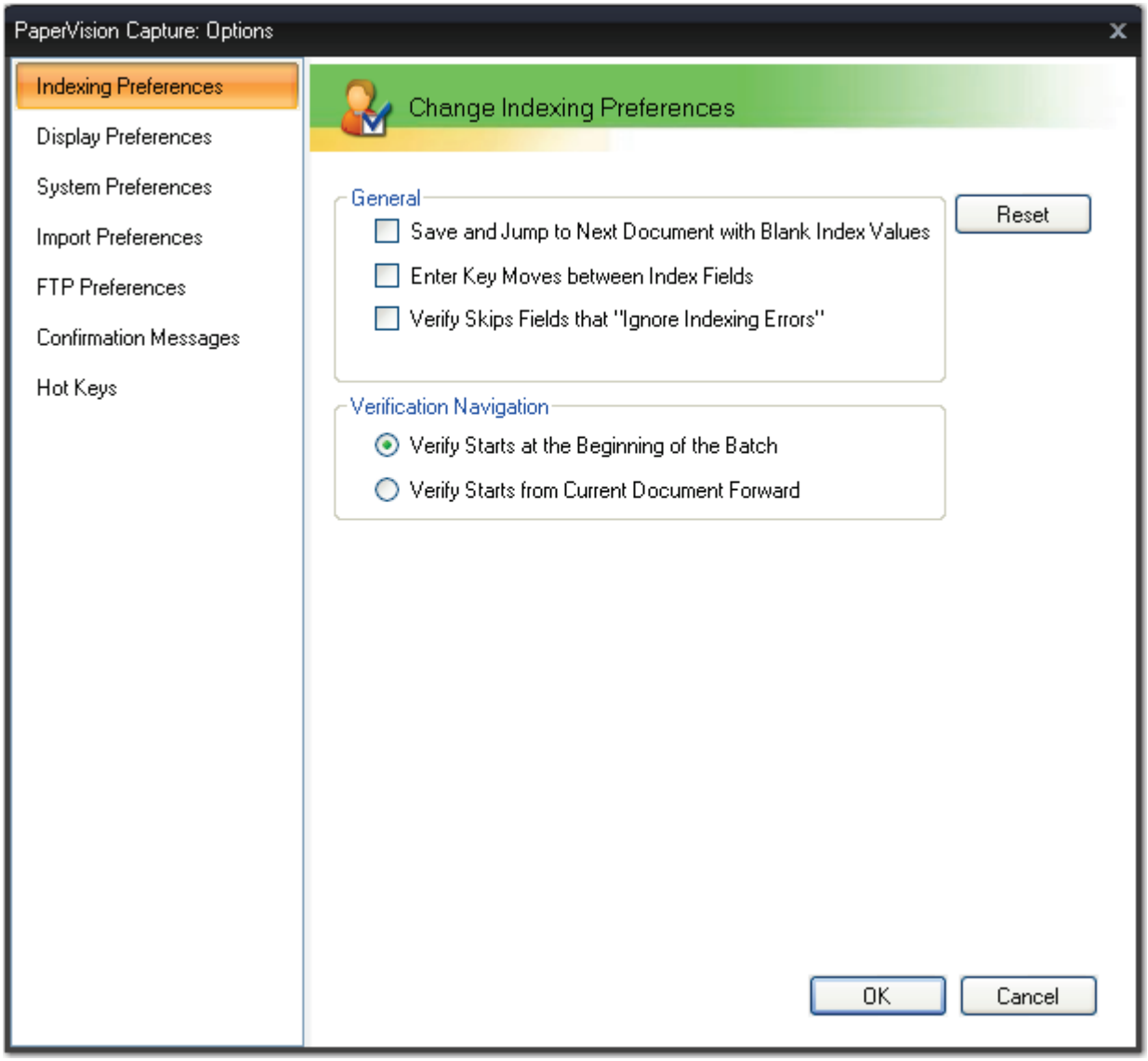
Preferences

You can configure indexing settings, display and system preferences, hot keys, and confirmation messages. To access these settings, select the **PaperVision Capture Desktop Button**  , and then select the **Preferences**  icon.

- **Indexing Preferences** allow you to customize hand-key indexing settings specific to individual workstations.
- **Display Preferences** define how single- and multiple-page documents display on-screen and how index verification is performed.
- **System Preferences** define document/page navigation and System Tray settings.
- **Import Preferences** allow you to select default PDF resolution and smoothing settings for import operations.
- **FTP Preferences** allow you to determine whether FTP error messages should appear and the subsequent action taken should an error occur.
- **Confirmation Messages** allow you to display or suppress confirmation messages that appear during specific operations.
- **Hot Keys** enable you to personalize hot key settings.

Indexing Preferences

To view the list of indexing preferences, select the **PaperVision Capture Desktop Button**  , and then select the **Preferences**  icon. Within the Indexing Preferences screen, you can select how to progress between documents and index fields. You can also determine where the index verification process begins in the batch.



Indexing Preferences

General

If you are tasked to enter index values, you can enable the following indexing settings:

- To save the current index values and proceed to the next document that contains blank index values, select **Save and Jump to Next Document with Blank Index Values.**
- To use the Enter key to move between index fields, select **Enter Key Moves Between Index Fields**.
- If you want the index verification process to skip fields that ignore an invalid index value, select **Verify Skips Fields that "Ignore Indexing Errors".**

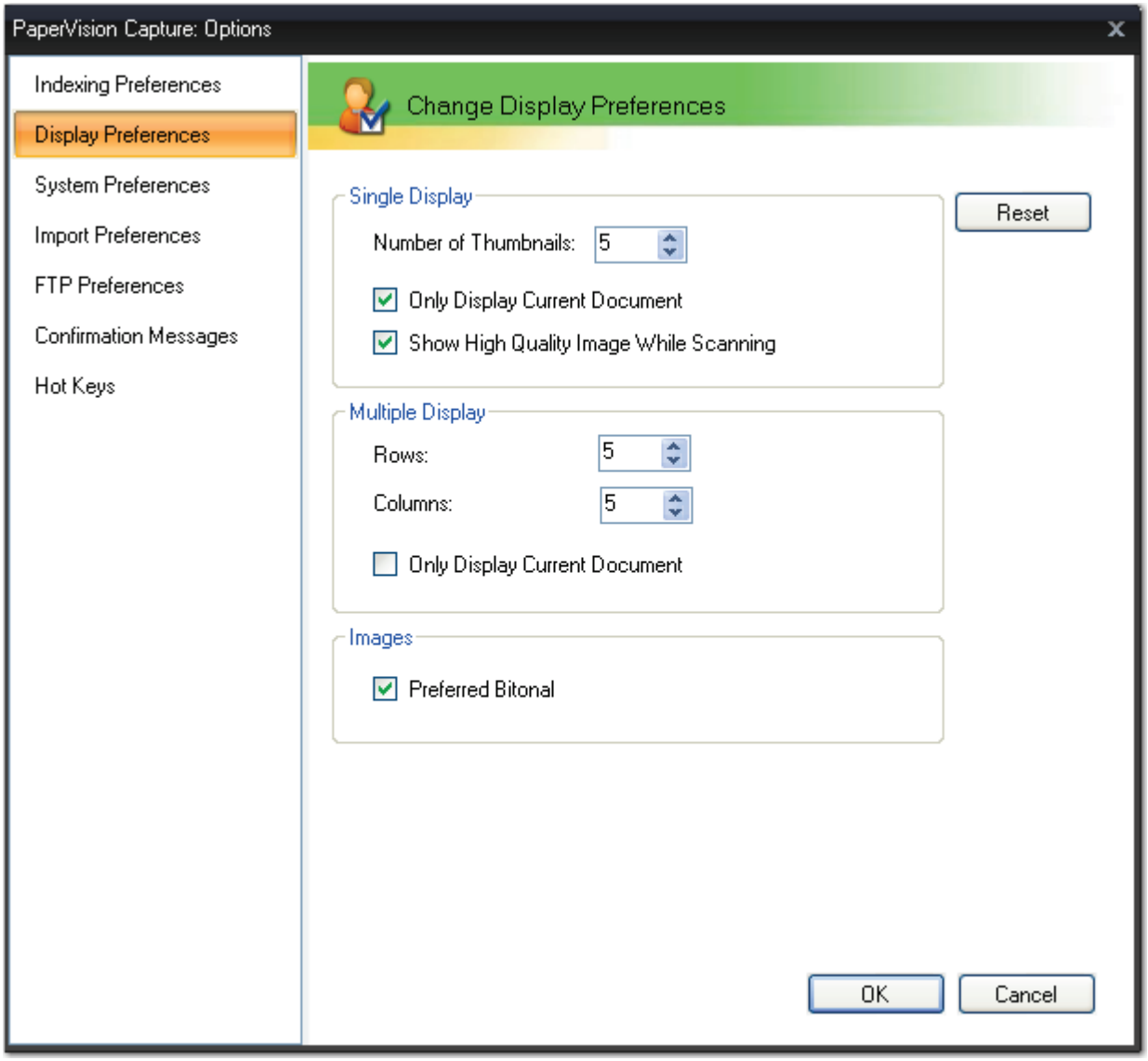
Verification Navigation

You can select the location in the batch where the index verification process begins.

- To start the index verification process at the beginning of a batch, select **Verify Starts at the Beginning of the Batch.**
- To start the index verification process from the current document forward, select **Verify Starts from Current Document Forward.**

Display Preferences

To view the list of display preferences, select the **PaperVision Capture Desktop Button**  , and then select the **Preferences**  icon. In the Display Preferences screen, you can manipulate the single- and multiple-page display options. Additionally, you can customize the image display quality for thumbnail views.



Display Preferences

Single Display

If you view images in single display (**View page > Single Display**), only one image appears at a time in the main viewing window. Optionally, the Thumbnails window allows you to preview a specified number of images.

- To preview document(s) as thumbnails, enter the number to display at a time in the **Number of Thumbnails** field.
- If you want to display only one document at a time, select **Only Display Current Document**. Leave this check box blank to view all of the documents in the batch.
- If you want to display a high quality image while scanning documents, select **Show High Quality Image While Scanning.** When this option is enabled, scanning speeds will decrease.

Multiple Display

You can view multiple images at a time in the main window (**View page > Multiple Display**) and specify how many images appear per row and column.

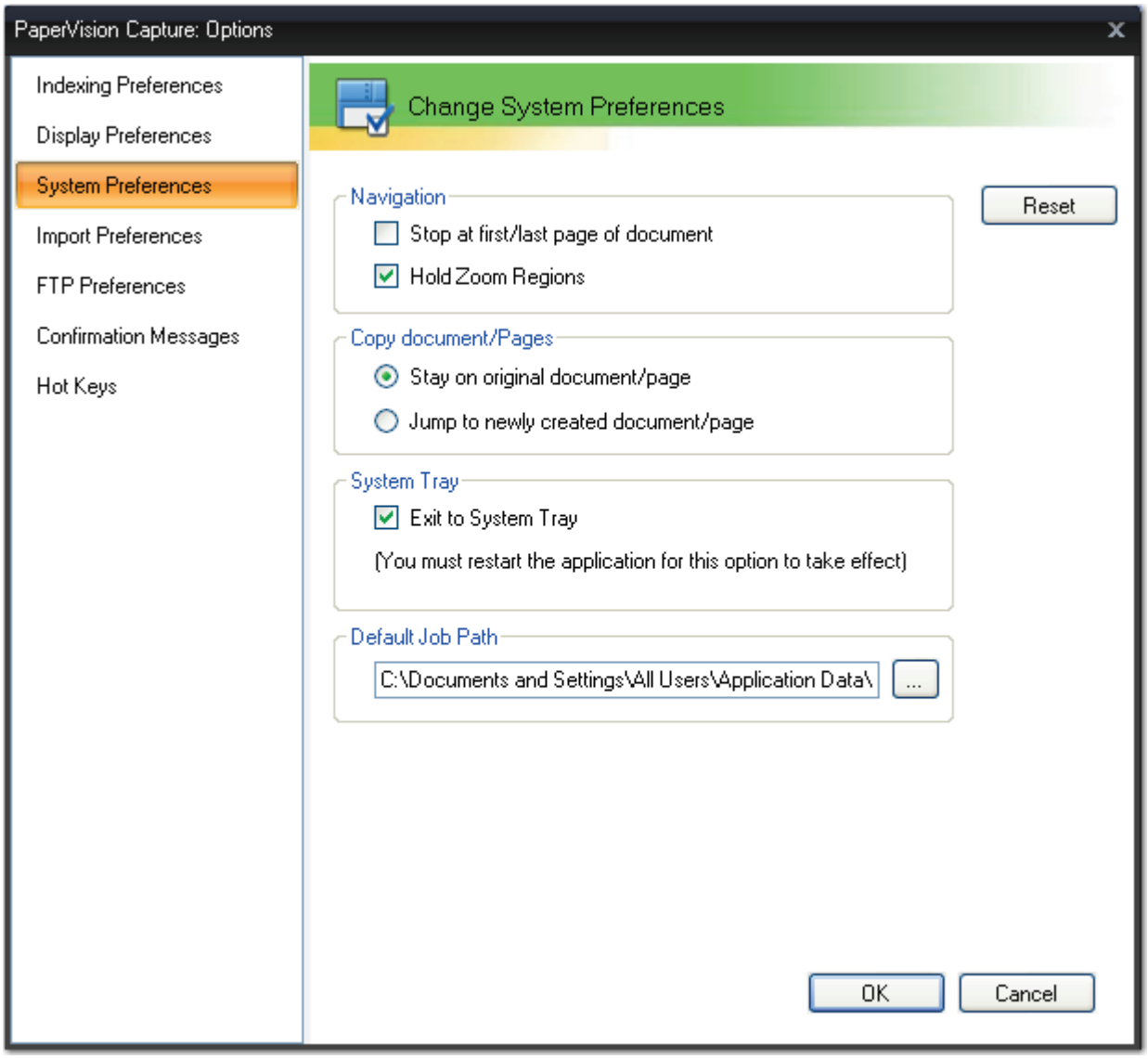
- For multiple-page documents, enter the number of rows/columns to display at a time in the **Rows** and **Columns** fields.
- Select **Only Display Current Document** to view only the current document in the batch. Leave this check box blank to view all of the documents in the batch.

Images

Select the **Preferred Bitonal** setting for dual-stream scanners.

System Preferences

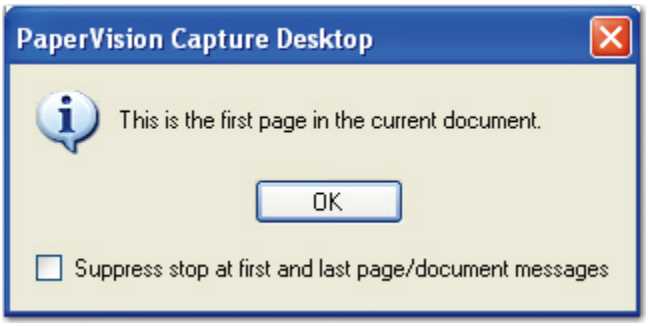
To view the list of system preferences, select the **PaperVision Capture Desktop Button**  , and then select the **Preferences**  icon. The System Preferences screen contains document/page navigation, System Tray, and Default Job Path settings.



System Preferences

Navigation

If you enable Stop at first/last page of document, an informational message appears if you attempt to navigate before the first page/document or beyond the last page/document. When this option is disabled and you navigate to the previous page of a document (when you are already on the first page), you are taken to the last page of the current document. Conversely, when you navigate to the next page of a document (when you are already on the last page), you are taken to the first page of the current document. The same results occur when you move between documents. Each time the message appears, you can choose to suppress the message in the future.



First/Last Page Message

Hold Zoom Regions

If you enable the **Hold Zoom Regions** setting, a zoom region that you set within an image will be retained as you navigate through the document.

Copy document/pages

This option assigns whether the original document/page (Stay on original document/page) or the copied document/page (Jump to newly created document/page) will open after you copy the existing document or page.

System Tray

When you exit PaperVision Capture Desktop or PaperVision Capture NOW!, both applications will automatically minimize to the System Tray by default (instead of completely shutting down). When this option is enabled, you can quickly resume running an application by right-clicking the PaperVision Capture Desktop or PaperVision Capture NOW! icon in the System Tray, and then selecting **Open PaperVision Capture Desktop** (or **Open PaperVision Capture NOW!**). Alternatively, you can close (shut down) an application by selecting the **Shut Down PaperVision Capture Desktop** (or **Shut Down PaperVision Capture NOW!**) option in the System Tray.

If you do not want the applications to minimize to the System Tray by default (and completely shut down upon exiting), disable the **Exit to System Tray** option.

> **Note**:
>
> If you disable or enable the **Exit to System Tray** option, you must restart (close and reopen) the application for the option to take effect.

Default Job Path

The Default Job Path stores PaperVision Capture Desktop job configurations. This path can be a local path or a network-accessible shared path such as **SERVER_NAME****SHARE**. Job configurations are named with the .PVCJOB extension and are stored by default in the local "Application Data(or, AppData)\Digitech Systems\Capture Desktop\Jobs" folder.

The following guidelines summarize how job path permissions and job modifications are handled in PaperVision Capture Desktop:

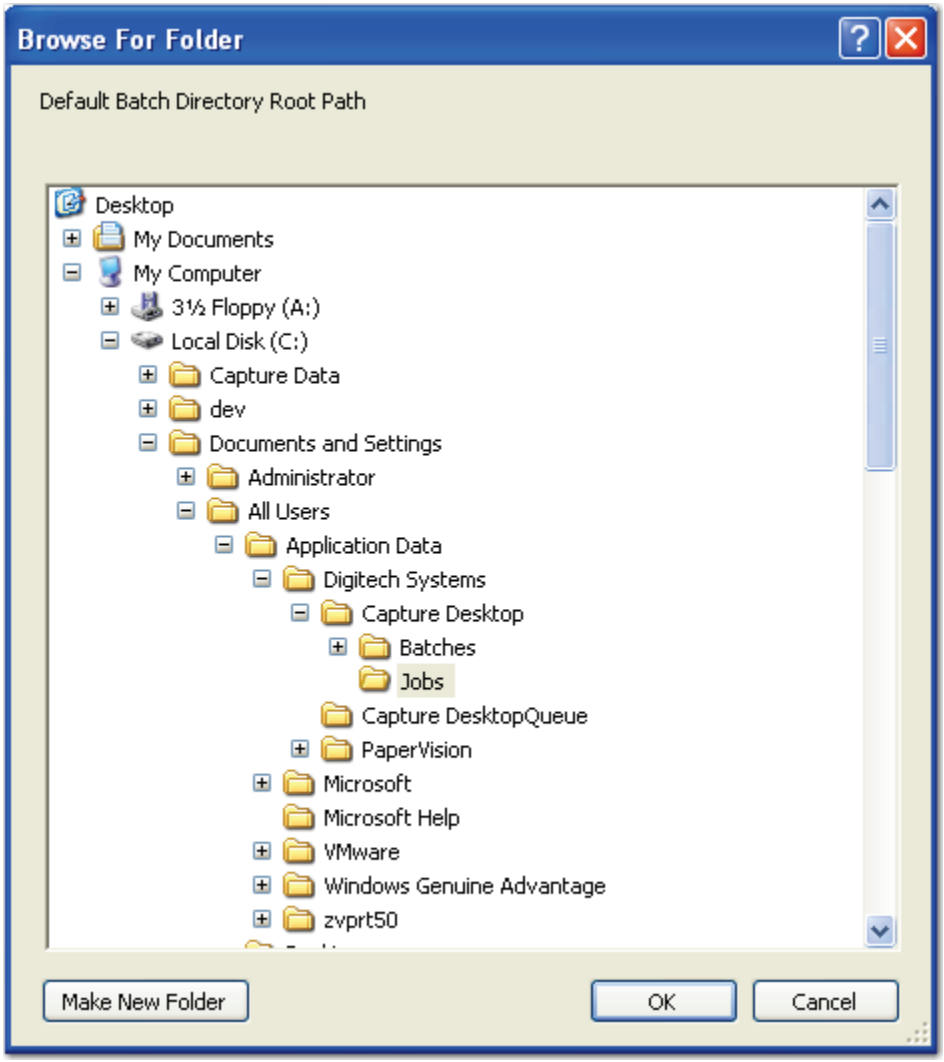
- The read/write permissions assigned in Windows determine who can view and modify PaperVision Capture Desktop jobs. For example, if jobs are saved to "c:\PaperVision Capture Desktop\Jobs" and you have read-only permissions to the Jobs folder, you cannot save any changes made to the job configuration. If you have not been granted access to the Jobs folder, then you will not see any jobs stored in that location.

  > **Note**:
  >
  > A user must be granted full read/write access to the Default Job Path in order to view jobs and make changes to job configurations.

- If multiple users make changes to the same job (concurrently from different workstations), only the most recent changes will be saved. In this scenario, any modifications made prior to the latest change may not be saved.
- If a job with one or more active batches is modified, active batches will not be impacted. For example, User A is currently working on a batch for Job 1. User B modified and saved Job 1's configuration while User A was still working on the batch. As a result, User A can continue to work on the batch, as it will not be impacted by the job modification.
- If you change the Default Job Path after you have created and saved jobs in the original location, only jobs from the updated location will be displayed when you create new batches (depending on permissions). For example, you changed the job path from "c:\PVCD\Jobs" to "c:\PVCD\Desktop" and created and saved jobs in the new location. When you create new batches, only jobs saved in the new location will be available for selection (from the "Use Existing Job" list in the New Batch Wizard).

To change the Default Job Path:

1. Enter the location of the **Default Job Path**, or click the ellipsis button to browse to the location. The **Browse for Folder** dialog box appears.

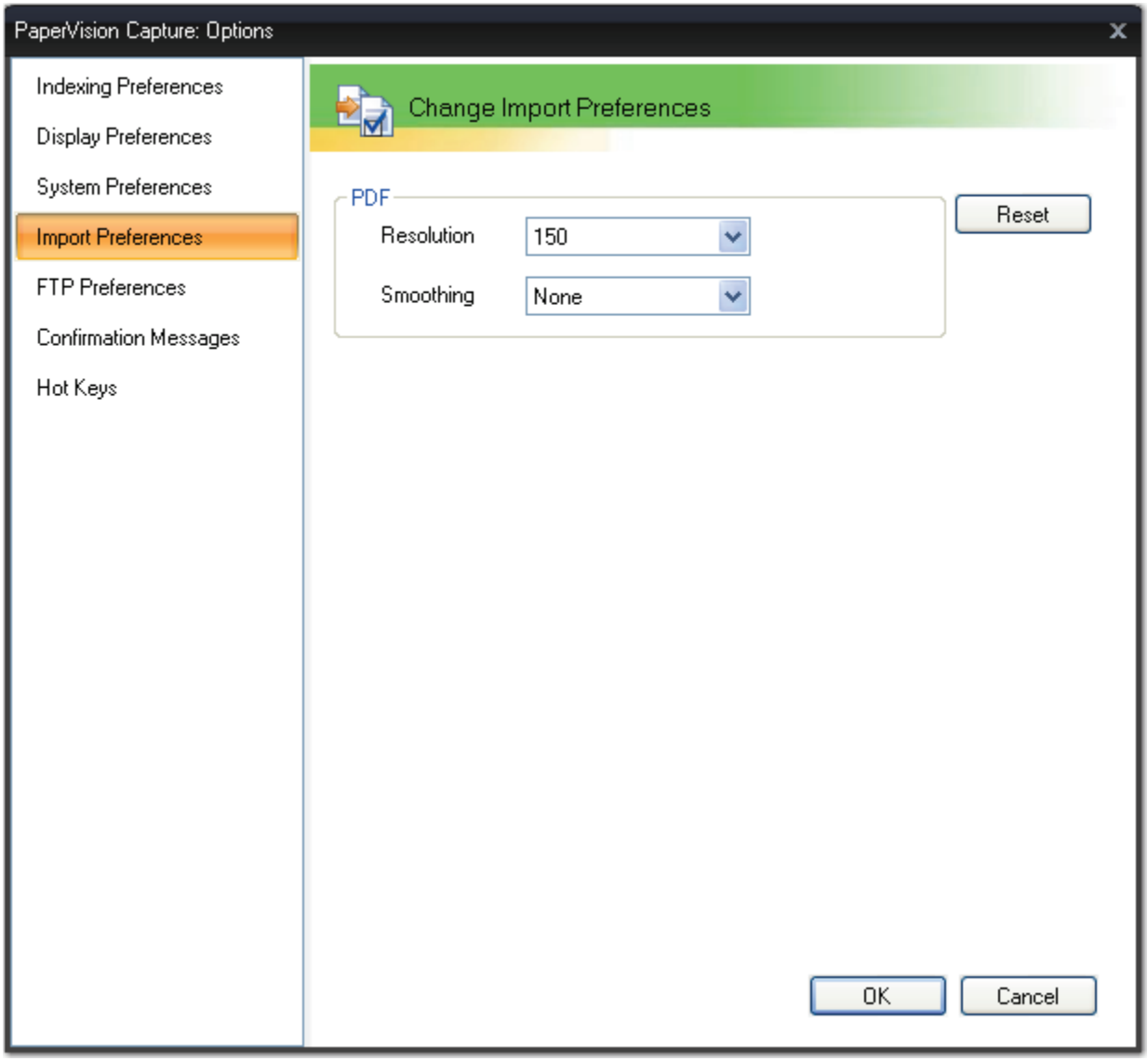


Browse for Folder

2. Highlight the new location, and then click **OK**.
3. Click **OK** in the **System Preferences** screen.

Import Preferences

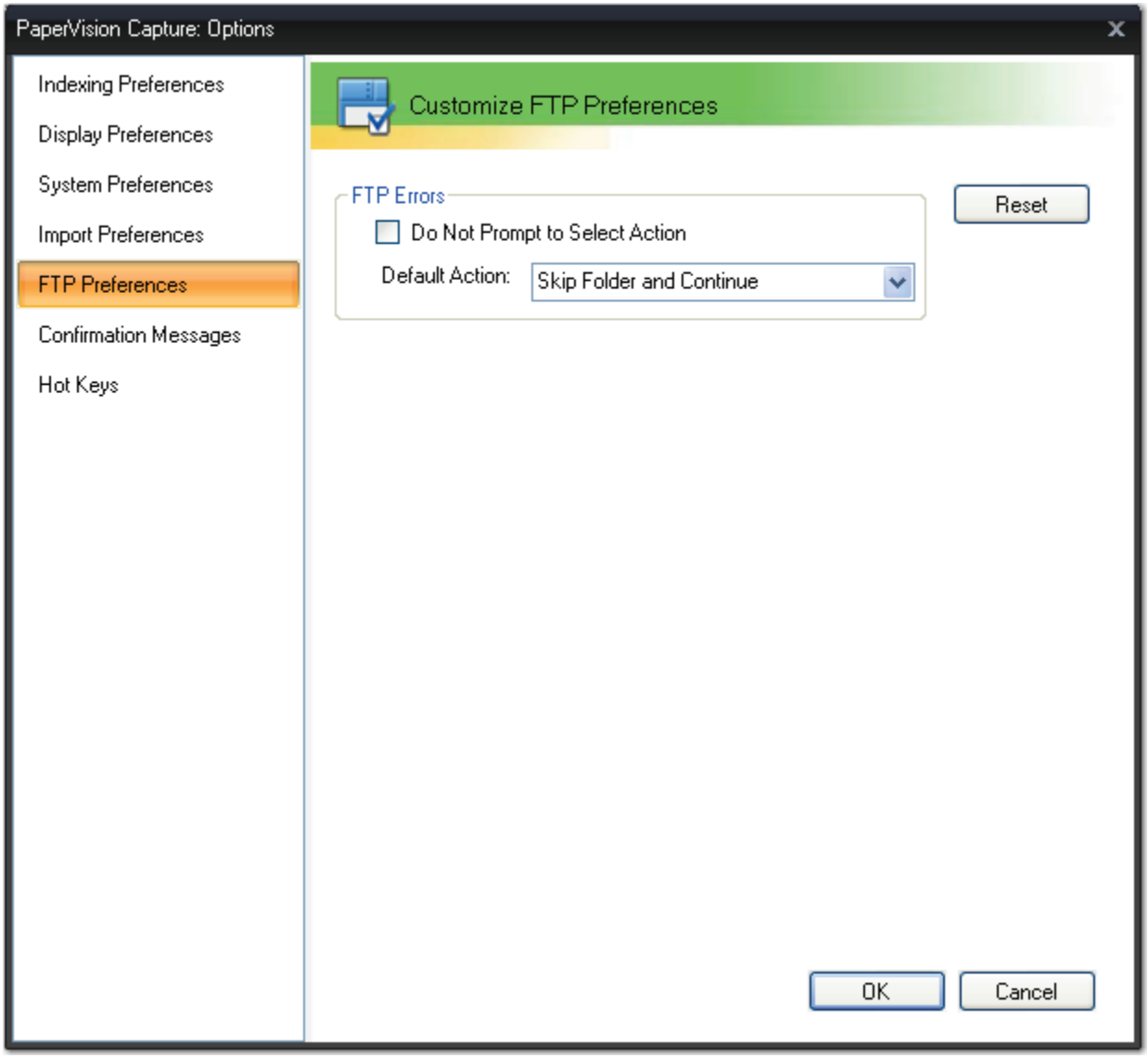
Import Preferences allow you to select default PDF resolution and smoothing settings for import operations. During the Import Batch and Import Images operations for PDF files, you can select the default resolution and smoothing settings. Higher resolution and smoothing settings may help improve the quality of PDF files during import operations. The default resolution is 150 dpi, and no smoothing is enabled by default.



Import Preferences

FTP Preferences

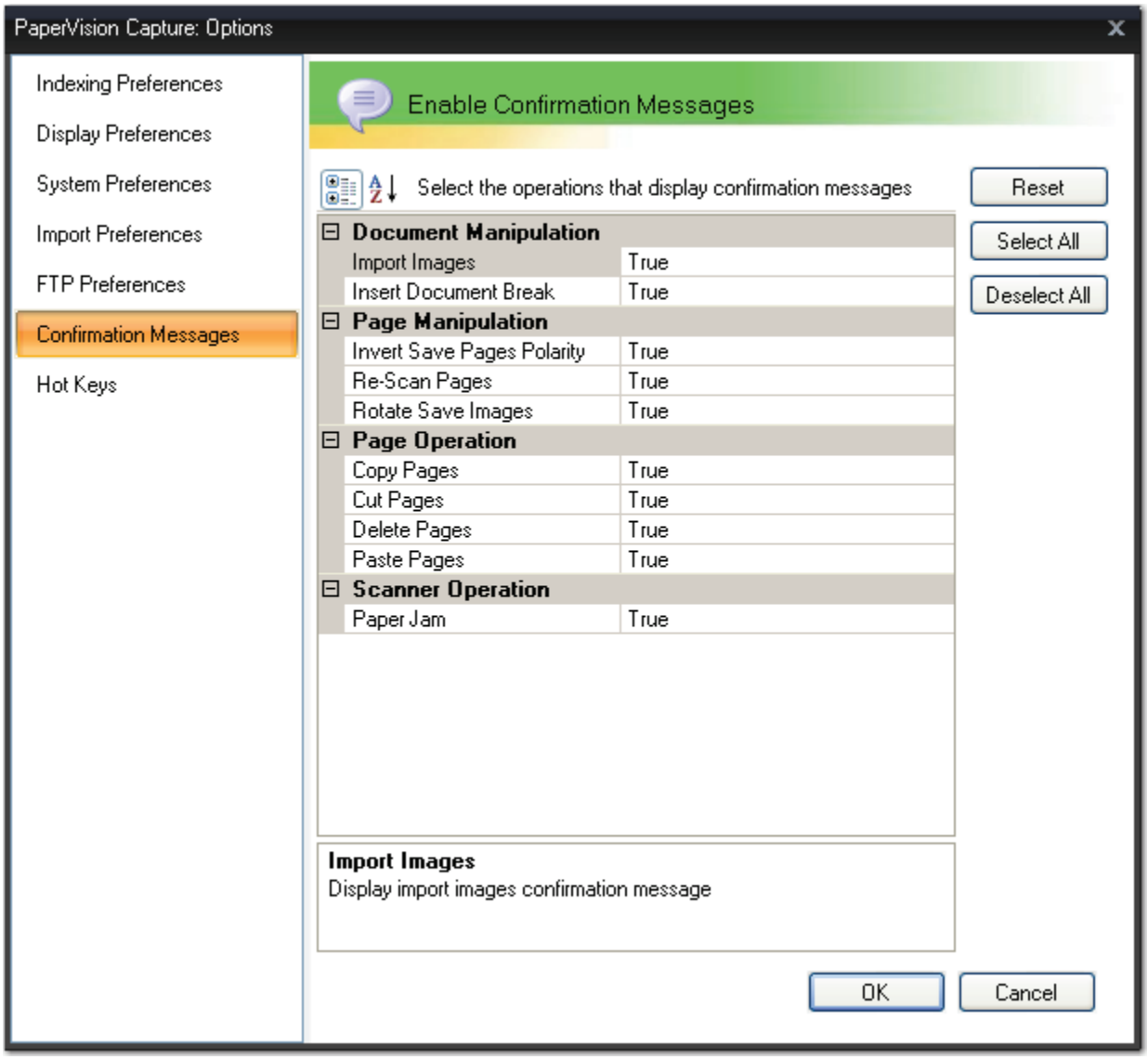
If you use the PaperFlow or ImageSilo/PVE XML exports to send data to ImageSilo and other FTP sites, you can determine whether FTP notifications should appear each time an error occurs during FTP transmission. To suppress FTP error messages and select a default action that will automatically occur should FTP fail to process, select **Do Not Prompt to Select Action**. Then, select the **Default Action** from the drop-down list.



FTP Preferences

Confirmation Messages

During specific operations in the application, such as cutting/pasting, copying/pasting, and deleting pages, you are presented with a confirmation message before the operation executes. To view the list of confirmation messages, select the **PaperVision Capture Desktop Button**  , and then select the **Preferences**  icon. In the Confirmation Messages screen, you can enable or suppress confirmation messages that appear during specific events. The following confirmation messages appear by default (except for scanner paper jam messages). However, when these messages appear, you can enable the options (e.g., Suppress future copy pages messages), so these messages do not appear each time you execute the operation.



Confirmation Messages

Confirmation messages for the following operations can be enabled or suppressed:

Document Manipulation

- Import Images
- Insert Document Break

Page Manipulation

- Invert Save Pages Polarity
- Re-Scan Pages
- Rotate Save Images

Page Operations

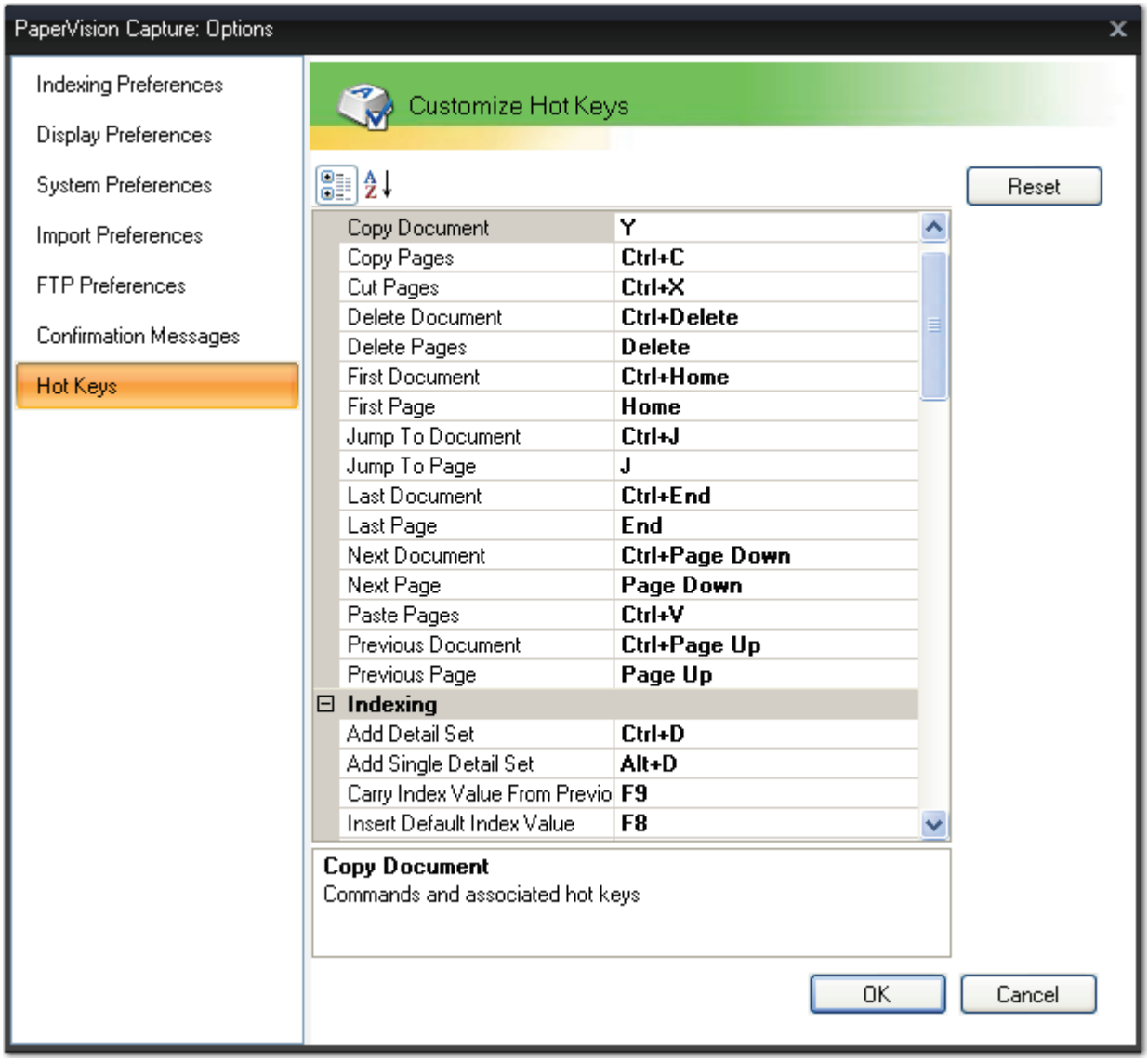
- Copy Pages
- Cut Pages
- Delete Pages
- Paste Pages

Scanner Operation

- Paper Jam

Hot Keys

To view the list of system hot keys, select the **PaperVision Capture Desktop Button** , and then select the **Preferences** icon. In the Hot Keys tab, you can customize the default hot key settings described below.



Hot Keys

The Customize Hot Keys screen includes the following sorting views:

- To view properties by category, click the **Categorized** icon.
- To view properties in alphabetical order, click the **Alphabetical** icon.

Batch Navigation

- Copy Document: Y
- Copy Pages: Ctrl+C
- Cut Pages: Ctrl+X
- Delete Document: Ctrl+Delete
- Delete Pages: Delete
- First Document: Ctrl+Home
- First Page: Home
- Jump to Document: Ctrl+J
- Jump to Page: J
- Last Document: Ctrl+End
- Last Page: End
- Next Document: Ctrl+Page Down
- Next Page: Page Down
- Paste Pages: Ctrl+V
- Previous Document: Ctrl+Page Up
- Previous Page: Page Up

Indexing

- Add Detail Set: Ctrl+D
- Carry Index Value From Previous Document: F9
- Insert Default Index Value: F8
- Verify Index Values: F7

Scanning

- Configure Scanner Settings: E
- Decrease Brightness: Shift+F2
- Decrease Contrast: Shift+F4
- Increase Brightness: F2
- Increase Contrast: F4
- Insert Pages After: Ctrl+A
- Insert Pages Before: Ctrl+B
- Mark New Document and Scan: M
- Re-Scan Page: A
- Scan One Page: O
- Scan Pages: S
- Stop Scanner: Escape

Note:

The Single Display and Multiple Display commands can be configured using the hot keys that you define. By default, these commands are undefined.

Standard

- Add Document: Shift+Ctrl+A
- Batch Information: Ctrl+I
- Clear Filters: Shift+Ctrl+F
- Document History: Ctrl+H
- Help Topics: F1
- Image Information: I
- Insert Document Break: Ctrl+Insert
- Merge Index Values: F5
- Merge Like Documents: Ctrl+F5
- Print Current Document: P
- Remove Document Break: Ctrl+R
- Re-Save Page: V
- Reset Image: Ctrl+T
- Rotate 90°: R
- Scale to Height: H
- Scale to Width: W
- Scale to Window: Ctrl+W
- Scroll Image Down: Ctrl+Down Arrow
- Scroll Image Left: Ctrl+Left Arrow
- Scroll Image Right: Ctrl+Right Arrow
- Scroll Image Up: Ctrl+Up Arrow
- Set Focus to Browse Batch: F6
- Set Focus to Indexing Window: Ctrl+Q
- Shuffle Document to Duplex: U
- Zoom In: Ctrl+Plus Sign
- Zoom Out: Ctrl+Minus Sign

Additional Standard Hot Keys

Although the following commands are undefined by default, you can configure these as well.

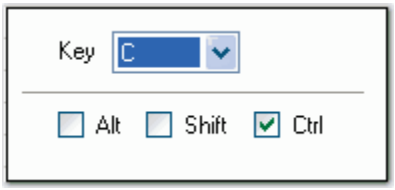
- About
- Auto Import
- Execute Custom Code
- Exit Application
- Help Manual
- Import Images
- Invert Page Polarity
- Invert and Save Page Polarity
- Move to Next Field
- Move to Previous Field
- Options
- Preferred Bitonal
- Rotate 270°
- Rotate and Save Images

Customizing Hot Keys

You can customize hot keys for any of the commands found in this screen.

To customize hot keys:

1. Expand the main category node (Batch Navigation, Indexing, Scanning, or Standard).
2. Click the row of the command to modify.
3. Open the drop-down menu in the right column, which opens the dialog box below.

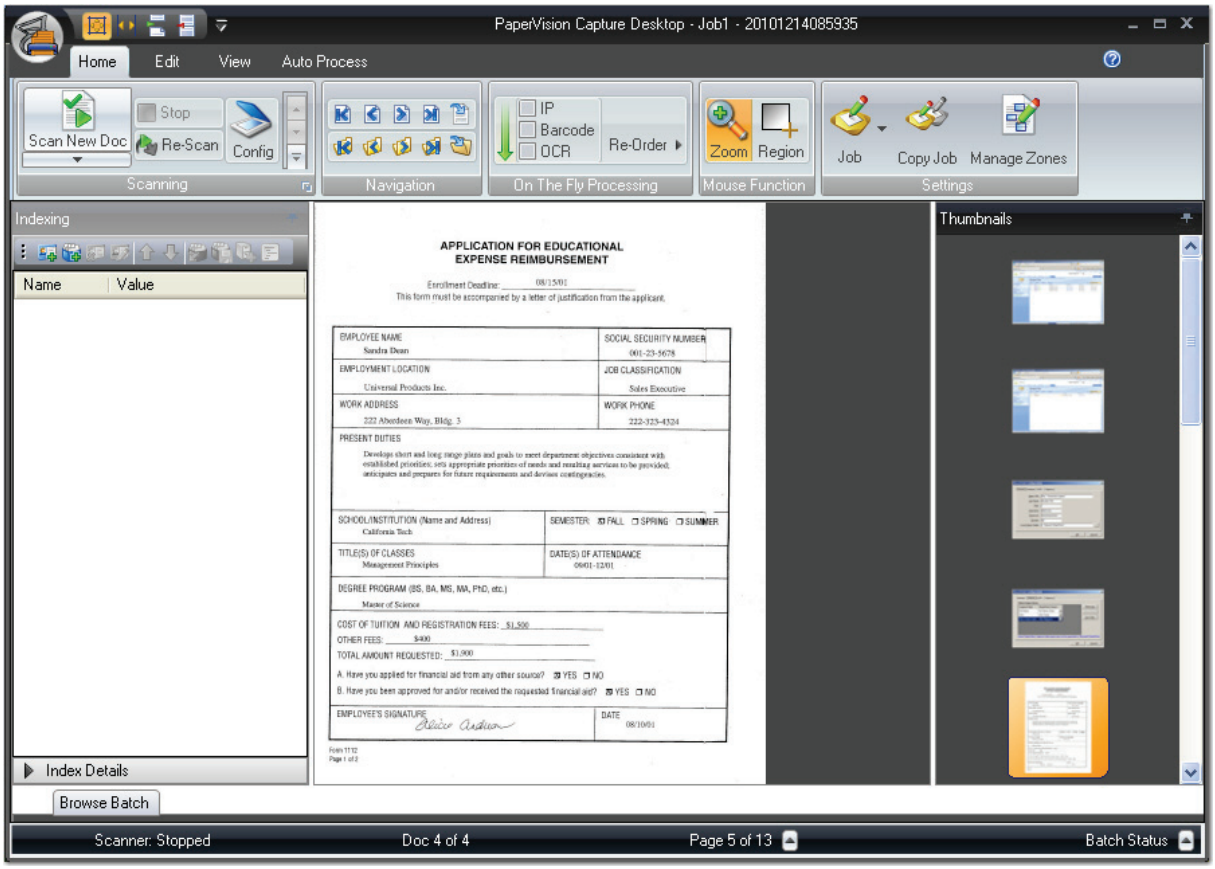


Customizing Hot Keys

4. Select from the **Key** drop-down list.
5. Select the additional keys (**Alt**, **Shift**, or **Ctrl**) that comprise the hot key setting (optional).
6. Click outside the command row to apply the changes.

Chapter 2 - Home Page

The Home Page contains the application's primary scanning, indexing, zonal setup, job settings, and navigation operations. You can also enable image processing, barcode, and zonal OCR processing on the fly while you scan documents in the Home page. General page and document navigation operations are also accessible in this page, such as First Page/Document, Next/Previous Page/Document, etc. If applicable, you can also apply zoom, barcode, and OCR zones directly on an image using the Region mouse function.



Home Page

Thumbnails Window

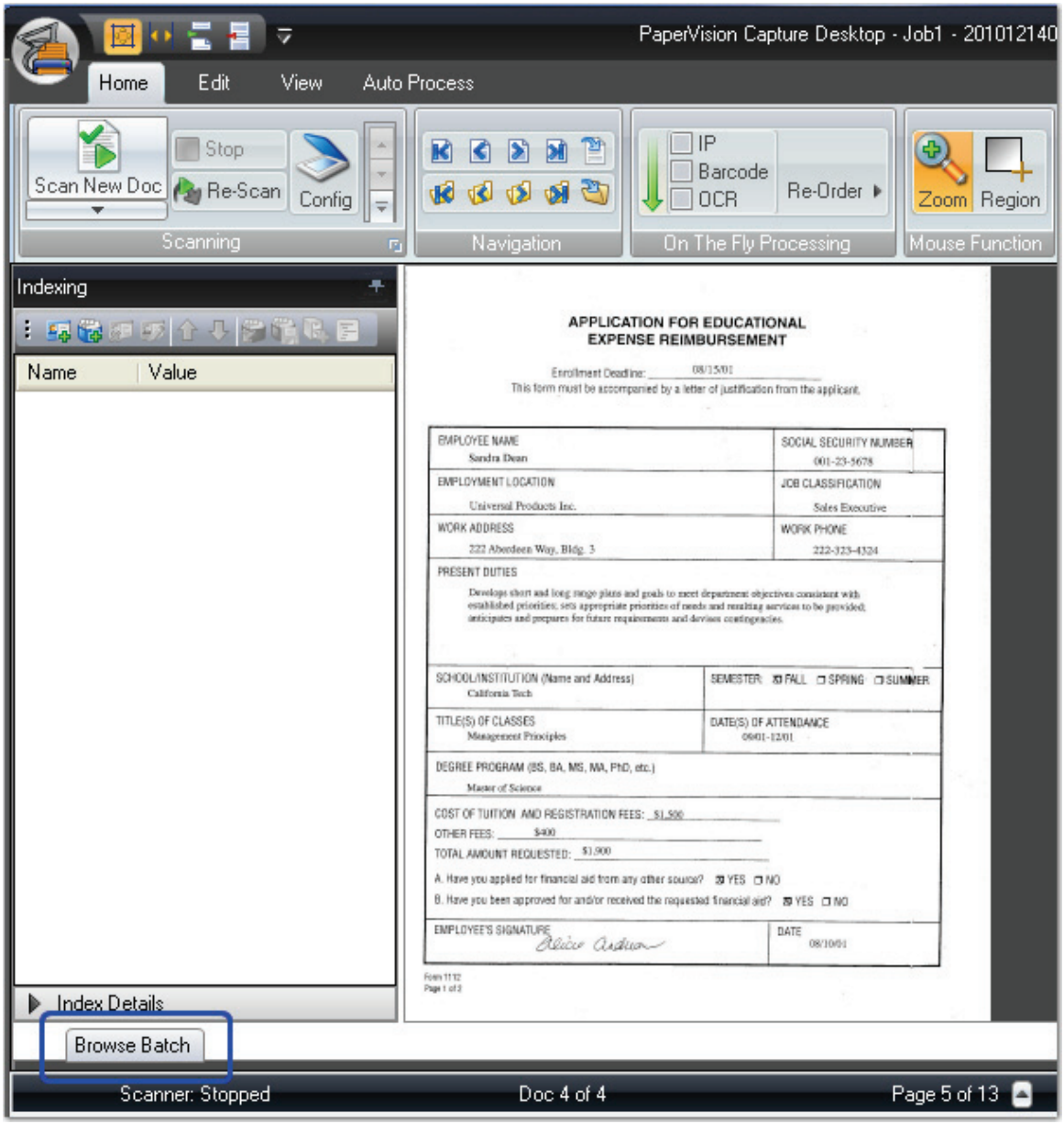
You can customize the Thumbnails window to display a certain number of pages by selecting the **PaperVision Capture Desktop Button**  **> Preferences > Display Preferences**. Thumbnails can help you navigate quickly through a document while previewing high-quality thumbnail images. In addition, you can perform basic operations in the Thumbnails window, such as cut, copy, and paste pages.

Browse Batch Window

The Browse Batch window, which is minimized by default, displays all documents, pages, and associated index values in the open batch. You can use the Browse Batch Filter to locate documents containing specific index values (or, blank index values), document numbers, and page counts. As you type characters in any of the top Filter fields, only those documents containing your specified criteria will appear in the Browse Batch window.

You can use the toolbar or right-click operations in the Browse Batch window to execute operations on the batch. First, you can delete one or more documents and associated images from the batch. You can also print and export one or more documents' metadata to an XML file, such as index values, batch name/order, and number of pages. The Browse Batch Filter can help you locate documents by document position number, number of pages comprising the document, or specific index values. Additionally, you can reorder and hide columns within the grid with the Configure Grid operation.

To open the Browse Batch window, click the **Browse Batch** tab located on the lower left side of the Home page:



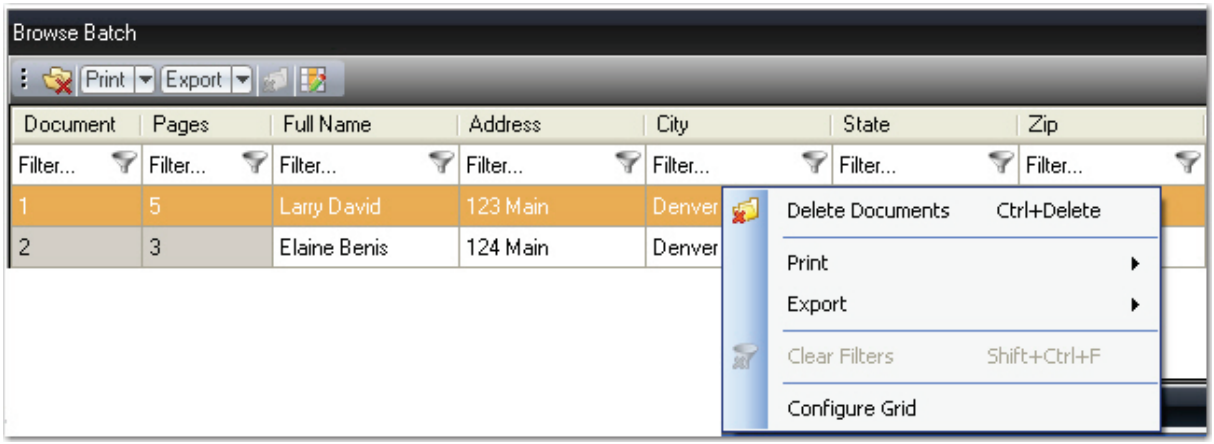
Browse Batch Tab

To pin the Browse Batch window in place, click the Auto-Hide pin located in the upper right side of the window:

Browse Batch Auto-Hide Pin

Browse Batch Window Operations

The following operations can be performed in the Browse Batch window using the toolbar or right-clicking in the window.



Browse Batch Window (and Right-Click Context Menu)

> **Note**:
>
> The Print and Export operations export document metadata; they do not print nor export associated images.

To delete metadata for one or more documents:

1. Highlight one or more rows.
2. Click the **Delete Documents** icon.
3. To proceed with the deletion, click **Yes**.

To print metadata for one or more documents:

1. Highlight one or more rows.
2. In the toolbar, select **Print > Selected Rows**.
3. Or, select **Print > All Rows**.
4. Select the printing parameters, and click **OK**.

To export metadata for one or more documents:

1. Highlight one or more rows.
2. In the toolbar, select **Export > Selected Rows.**
3. Or, select **Export > All Rows**.

4. In the **Save As** dialog box, browse to the appropriate directory where the XML file will be saved.
5. Enter the file name.
6. Click **Save**.

To clear all Browse Batch Filter criteria, right-click and select **Clear Filters**.

To reorder the columns:

1. In the toolbar, click the **Configure Grid** icon.

> **Note**:
>
> If no indexes have been defined, the Configure Grid operation will be disabled.

2. In the **Show/Hide Columns** dialog box, highlight the column to move, and then select **Move Up** or **Move Down**.

> **Note**:
>
> You cannot move the Document and Pages columns.

3. To restore the columns to their original order, click **Reset**.

To show or hide columns:

1. In the toolbar, click the **Configure Grid** icon.
2. In the **Show/Hide Columns** dialog box, select the columns to display in the Browse Batch grid.

> **Note**:
>
> You cannot hide the Document and Pages columns.

3. To restore the original columns that will be displayed in the grid, click **Reset**.

Scanning Operations

In the Home page, scanning operations enable you to scan new documents; scan pages to the end of the current document; scan single pages; configure scanner settings; and various other scanning tasks. Before you initially scan documents, you must first configure a scanner by selecting the **Configure Scanner**  icon in the toolbar. The Scan New Doc operation is the default operation in the Scanning toolbar group. However, if you select the Scan Pages or Scan One Page operation, that operation will appear as the default operation in the Scanning toolbar group.

The table below summarizes each operation and associated hot key:

Scanning Operations	
Scan Pages 	Scans from your specified page to the end of the current document Hot Key: S
Scan New Doc 	Creates a new document at the end of the batch and begins scanning pages in the new document Hot Key: M
Scan One Page 	Scans your specified page at the end of the current document (If your scanner is configured for duplex scanning, only the front image is scanned) Hot Key: O
Insert Pages Before/After 	Inserts pages before or after a specified page and applies page rotation (if defined by the administrator); does not apply barcode recognition or auto document breaks Hot Keys: Ctrl+B (Before); Ctrl+A (After)
Re-Scan Pages 	Re-scans your specified range of pages Hot Key: A
Stop Scanning 	Stops the scanning process Hot Key: Escape
Configure Scanner 	Allows you to configure your scanner's settings Hot Key: E See **Scanner Settings** for more information.

Scanning Operations (continued)	
Saved Scanner Settings	Drop-down menu that contains saved scanner settings and allows you to toggle among multiple scanner configurations on-the-fly Hot Key: Ctrl+F6

Configuring the Scanner

The Configure Scanner operation allows you to assign scanner settings. To add a scanner profile, click the **Configure Scanner**  icon. See the **Scanner Settings** section for details on each setting.

Scan New Doc

This operation creates a new document and begins scanning pages in the new document. To scan a new document, select the **Scan New Doc**  icon. The new document will be created at the end of the batch, and pages in the new document will be scanned.

Scan Pages

To scan pages from your specified page to the end of the current document, click the **Start Pages**  icon.

Scanning One Page

This operation scans your specified page at the end of the current document (if your scanner is configured for duplex scanning, only the front image is scanned). To scan one page, navigate to the page, and then select the **Scan One Page**  icon.

Stopping the Scanning Process

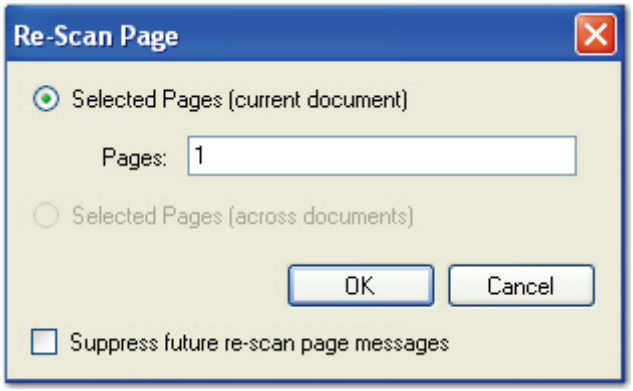
To stop the scanning process, click the **Stop Scanning**  icon.

Re-Scanning a Page

If a page did not scan properly, you can use the Re-Scan Page operation in the toolbar.

To re-scan a page:

1. Navigate to the page that needs to be re-scanned.
2. Click the **Re-Scan** icon. The **Re-Scan Page** dialog appears.



Re-Scan Page

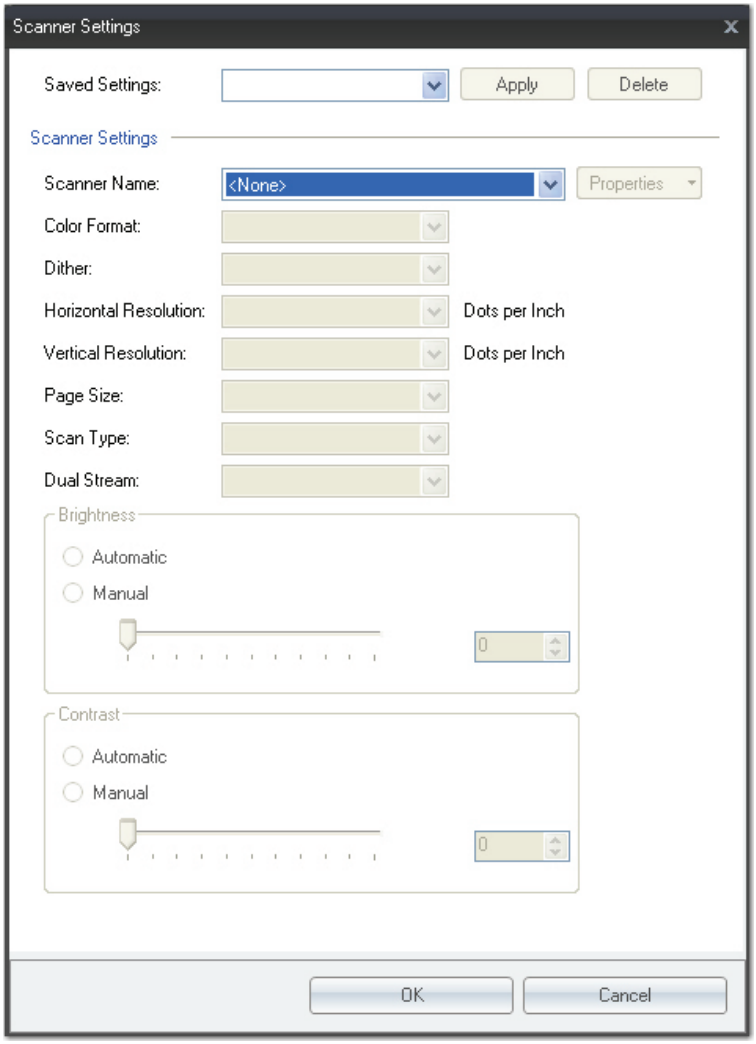
3. Enter the page numbers in the current document or across documents to be scanned again. You can place commas between page numbers and insert dashes for page ranges (e.g., 1, 2, 3-5).
4. Click **OK**.

> **Note**:
>
> Any image processing previously applied to the image will be removed when you rescan a page.

Scanner Settings

This drop-down menu displays scanner settings that were previously saved. To save a new scanner setting, enter the name directly in the **Saved Settings** field; then click **Apply**. To remove a setting, select it from the **Saved Settings** drop-down menu; then click **Delete.**

In PaperVision Capture Desktop, black and white images are saved in an industry standard Group IV TIFF file format, while color or grayscale images are saved in a standard JPG or BMP file format. To view scanner settings, open the Home page, and then select the **Config**  icon in the Scanning toolbar group.



Scanner Settings

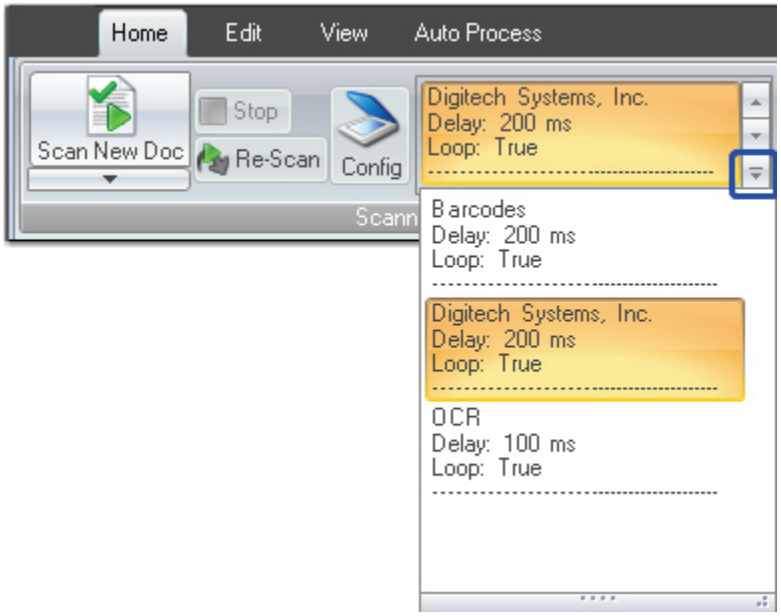
Note:

Depending on the type of scanner that is used, some scanner options may be disabled, and the number of options available in the drop-down menus may vary.

Saved Settings

This drop-down menu displays scanner settings that were previously saved. To save a new scanner setting, enter the name directly in the Saved Settings field; then click **Apply**. To remove a setting, select it from the Saved Settings drop-down menu; then click **Delete**. To use a scanner profile, select the scanner profile from the drop-down list.

In the Home page, you can view all scanner profiles that you created by selecting the down arrow next to the current scanner profile in the Scanning toolbar group:



Scanner Profiles

Scanner Name

Click the **Scanner Name** drop-down menu to select a scanner that has been installed and detected by PaperVision Capture Desktop. The Properties menu allows you to configure scanner and file import devices. Depending on the type of scanner, the menu options will display different settings.

The Properties drop-down menu contains the following options:

- More Settings may contain additional scanner settings that are available for configuration.
- About displays the driver's version, copyright, and other information specific to the scanner.
- Area Settings allow you to assign the scanning area.
- Extended Settings may contain additional scanner settings that are available for configuration.
- Calibrate allows you to calibrate the scanner driver.
- Configure allows you to configure the scanner driver settings.

Color Format

Also known as the mode, you can select from options such as black and white, color, etc.

Dither

Dithering converts and simulates unavailable colors. When dithering is turned on, the system combines two or more colors to approximate the unavailable color.

Horizontal Resolution

Select the horizontal dots-per-inch resolution setting to apply during the scanning process.

Vertical Resolution

Select the vertical dots-per-inch resolution setting to apply during the scanning process.

Page Size

This setting determines the default page size of the image as it is scanned.

Scan Type

This setting determines if scanning should be two-sided (duplex), one-sided (simplex), etc.

Dual Stream

Applicable to dual stream scanners, you can select how pages will be scanned (Front/Back, Front Only, Back Only). To disable dual stream scanning, select None.

Brightness

Brightness defines a pixel's lightness value from black (darkest) to white (brightest). Select the brightness level to apply during the scanning process and whether it should be applied manually or automatically. If applying the contrast manually, use the slider to increase or decrease the amount of contrast.

Contrast

Contrast is a measure of the rate of change of brightness in an image. A high-contrast image contains defined transitions from black to white. Select the contrast level to apply during the scanning process and whether it should be applied manually or automatically. If applying the contrast manually, use the slider to increase or decrease the amount of contrast.

Document and Page Navigation

The document and page navigation operations are accessible in the Home Page.

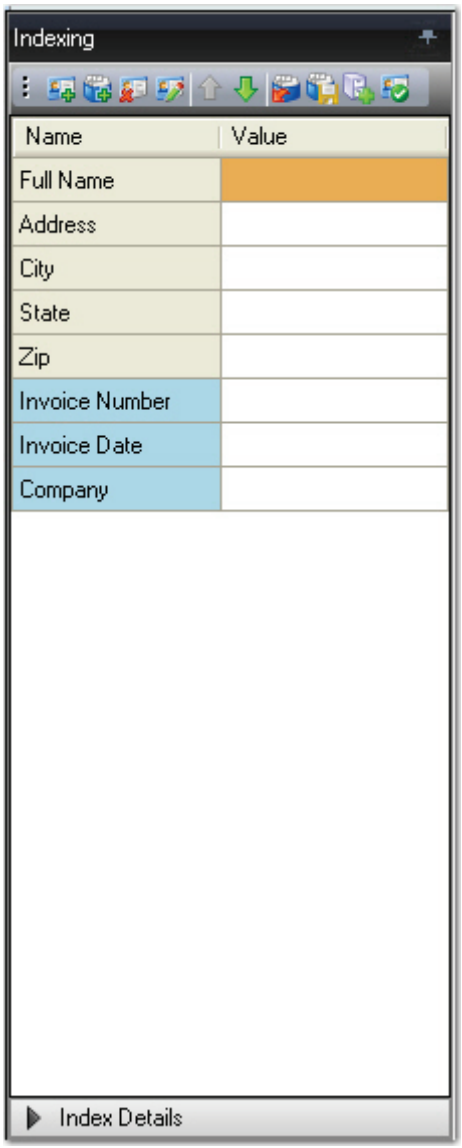
Toolbar icons and hot keys are outlined in the table below:

Document Navigation Operations	
First Document 	Displays the first page of the current document Hot Key: Ctrl+Home
Previous Document 	Displays the previous page of the current document Hot Key: Ctrl+Page Up
Next Document 	Displays the next page of the current document Hot Key: Ctrl+Page Down
Last Document 	Displays the last page of the current document Hot Key: Ctrl+End
Jump To Document 	Retrieves your specified document number Hot Key: Ctrl+J

Page Navigation Operations	
First Page 	Displays the first page of the current document Hot Key: Home
Previous Page 	Displays the previous page of the current document Hot Key: Page Up
Next Page 	Displays the next page of the current document Hot Key: Page Down
Last Page 	Displays the last page of the current document Hot Key: End
Jump To Page 	Retrieves your specified page number Hot Key: J

Indexing

In the Home page, the Indexing window, located on the left side of the Home page by default, provides both index configuration and hand-key indexing functionality. You can also configure indexes and hand-key index values while you are in the Edit, View, and Auto Process pages. In each of these pages, you can add, edit, and delete index names; enter and verify index values; and add and configure detail sets within the Indexing window. You can also reset (clear), save, and reorder index fields that you have added.



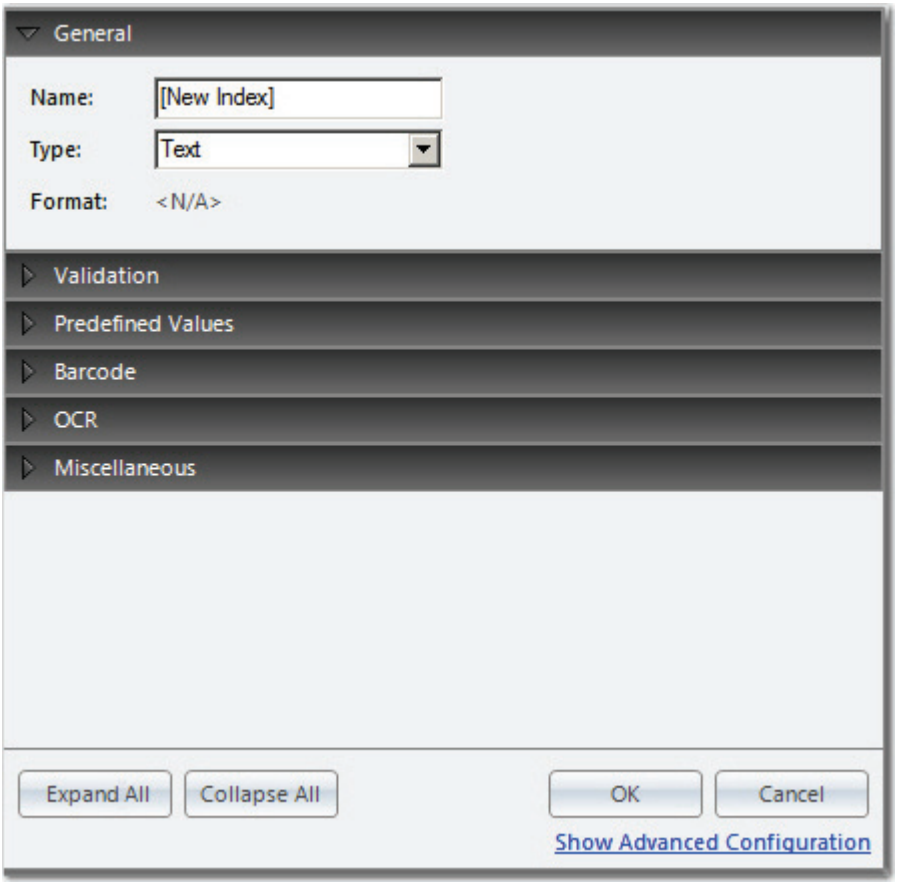
Indexing Window

Adding Indexes

You can add and configure indexes in the Indexing Properties window.

To add an index:

1. Click the **Add Index** icon. The **Index Properties (New Index)** window appears.



Indexing Properties (New Index)

2. Enter the name of the new index, and select the index type. Depending on the index type you select, you may be able to select the index format. See the **Index Types and Formats** section for more information.
3. Configure any remaining Validation, Predefined Values, Barcode, OCR, and/or Miscellaneous properties. See the **Basic and Advanced Index Properties** section for more information.

> **Tip**:
>
> Click the **Expand All** button to view all first-level properties found in each section. Click the **Collapse All** button to hide all properties in each section.

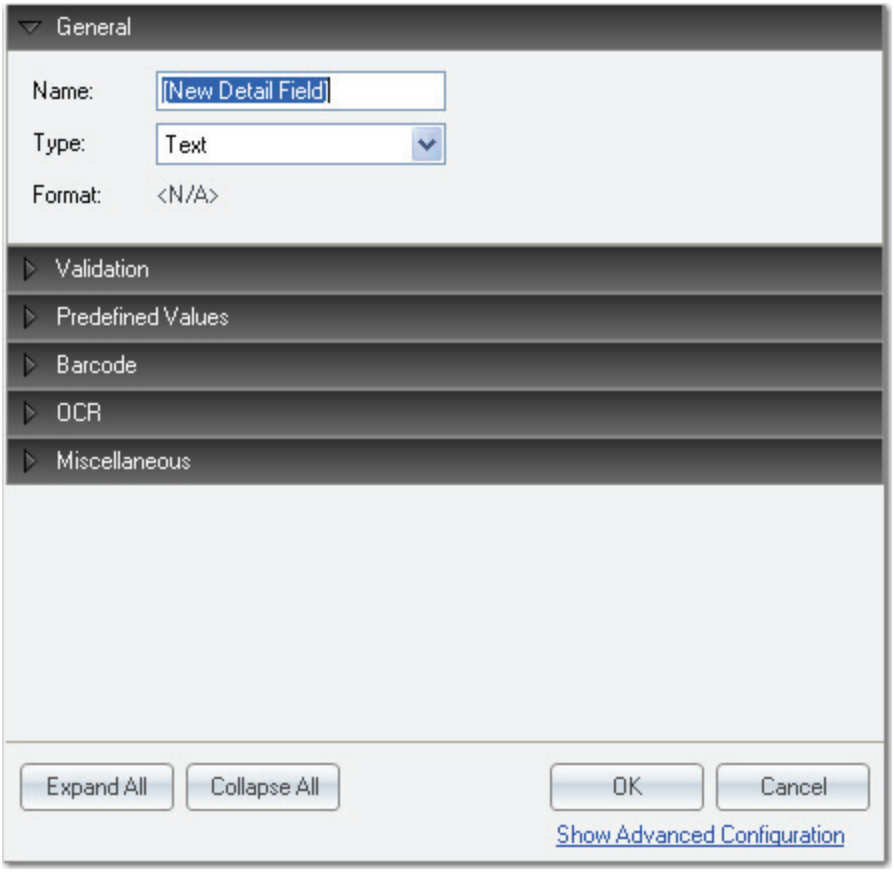
4. When finished configuring the index, click **OK**.

Adding Detail Fields

You can add individual detail fields that can (optionally) comprise a detail set. In PaperVision Capture Desktop, detail sets define a collection of indexes that allow multiple sets of field data to reference a single document. See the **Detail Sets** section for more information.

To add a detail field:

1. Click the **Add Detail Field** icon. The **Index Properties (New Detail Field)** window appears.



Index Properties (New Detail Field)

2. Enter the name of the new detail field, and select the index type. Depending on the index type you select, you may be able to select the index format. See the **Index Types and Formats** section for more information.
3. Configure any remaining Validation, Predefined Values, Barcode, OCR, and/or Miscellaneous properties. See the **Basic and Advanced Index Properties** section for more information.

> **Tip**:
>
> Click the **Expand All** button to view all first-level properties found in each section. Click the **Collapse All** button to hide all properties in each section.

4. When finished configuring the index, click **OK**.

Deleting Indexes and Detail Fields

You can delete indexes and detail fields from the Indexing window.

To delete an index or a detail field:

1. Highlight the index or detail field in the Indexing window.
2. Click the **Delete Index (Detail Field)** icon.
3. Click **OK** to confirm the deletion.

Editing Indexes and Detail Fields

You can change the properties of an existing index or detail field. Individual index fields and detail fields are differentiated by color in the Indexing window. See the **Basic and Advanced Index Properties** section for more information. Or, see the **Detail Sets** section for more information.

To change the properties of an index:

1. Highlight the index name in the Indexing window, and then click the **Configure Index** icon.
2. Edit the appropriate properties.
3. Click **OK** to save your changes.

To change the properties of a detail field:

1. Highlight the detail field name in the Indexing window, and then click the **Configure Detail Field** icon.
2. Edit the appropriate properties.
3. Click **OK** to save your changes.

Moving Indexes Up/Down

You can change the order in which indexes and detail fields appear in the Indexing window.

To move an index or detail field up or down the list:

1. Highlight the index or detail field.
2. Click the **Up** or **Down** arrow to move it to another position.

Resetting Index Values

After you have entered one or more index values, you can reset (clear) all index values that you have entered in the Indexing window.

To reset index values:

1. If applicable, navigate to the appropriate document.
2. Click the **Reset Values** icon. All previously-entered index values will be cleared.

Adding Detail Sets

You can add detail sets in the Indexing window. In PaperVision Capture Desktop, detail sets define a collection of indexes that allow multiple sets of field data to reference a single document. See the **Detail Sets** section for more information.

To add a detail set:

1. Click the **Add Detail Set** icon.
2. Enter the number of detail sets to add.
3. Click **OK**.

Saving Index Values

After you enter your index values in the **Value** column in the Indexing window, you can save them prior to exporting your batch. Unsaved index values display with an asterisk (*) next to the index field name. If you have not entered a required field (or entered an invalid index value) in the Indexing window upon saving the index value, a red exclamation icon will appear in the field.

> **Tip**:
>
> To receive a message that notifies you when all documents have been indexed, enable the general display preference option, **Save and Jump to Next Document with Blank Index Values.**

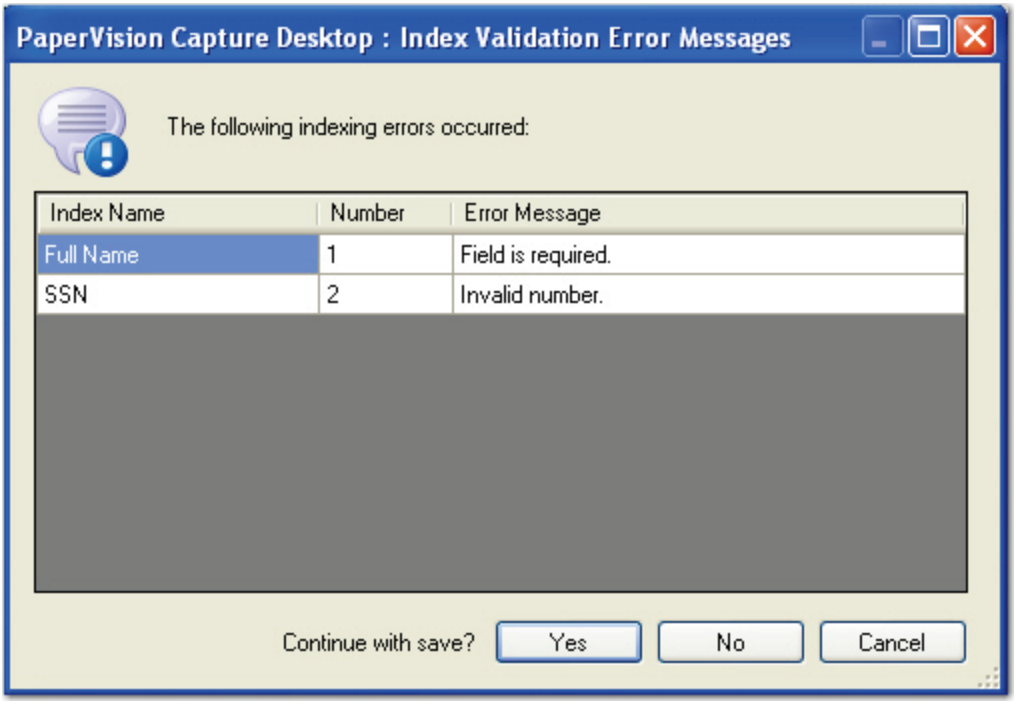
To save index values:

1. After you have finished entering the required index values in the **Values** column, click the **Save Indexes** icon.

> **Tip**:
>
> You can also press **Enter** to save index values in the Indexing window.

2. If you defined any index validation properties (see the **Basic Indexing Properties** section for more information), any indexing errors will appear in a popup message before you save them. Click **OK** to return to the Indexing window to correct the errors before proceeding.

> **Note**:
>
> If your Browse Batch window is viewable, its index values will be updated once you execute the Save Indexes command.



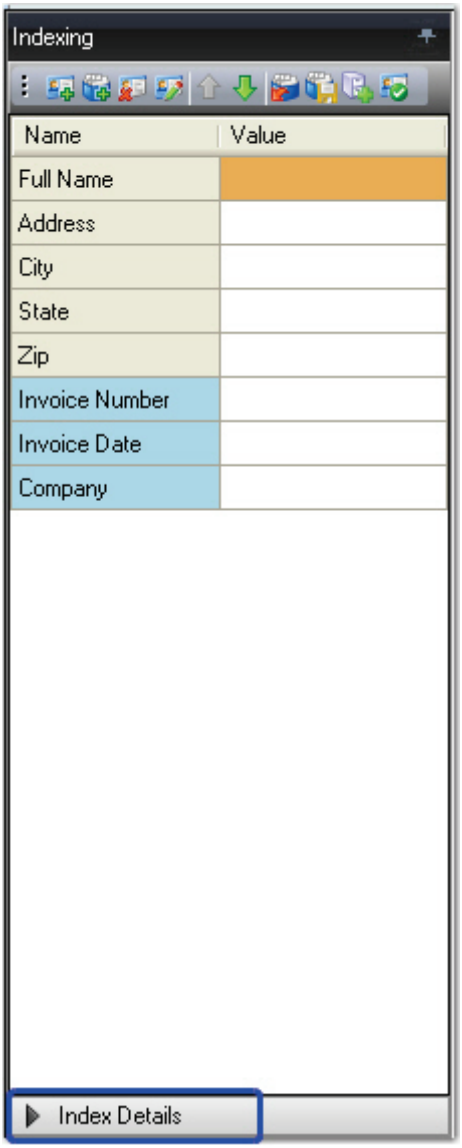
Index Validation Error Messages

Verifying Index Values

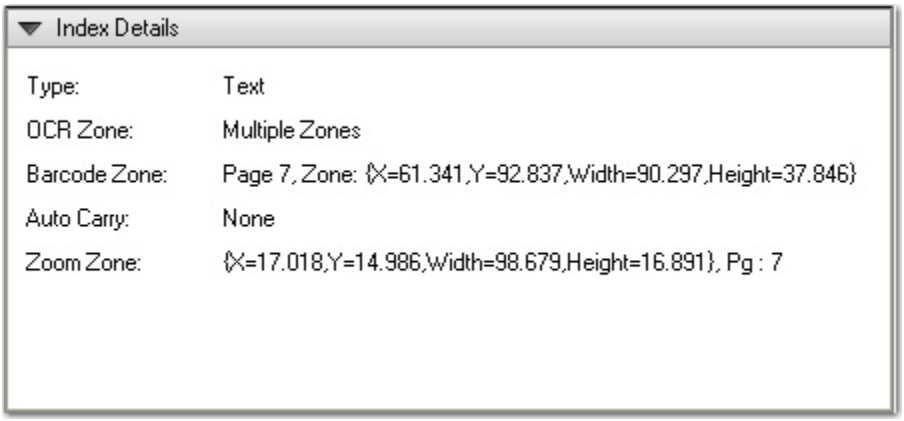
To ensure the accuracy of hand-key indexing, the Verify Index Values operation searches for blank values or other search string values that have been defined in the current batch. Depending on your general index verification settings in Indexing Preferences (**Verify Starts from Current Document Forward** or **Verify Starts at the Beginning of the Batch**), the index verification process starts with the appropriate document in the batch and will highlight the next document that contains a blank value or your defined search string value. To verify that blank index values or other defined search string values are not present, click the **Verify Index Values** icon. If no blank values or defined search string values exist, you will see a message that indicates all indexes have been verified.

Index Details

Although index details are hidden by default, you can click the Index Details arrow located at the bottom of the Indexing window. Index details reveal whether the selected index is assigned to a barcode, OCR, or zoom zone. The index type (date, currency, etc.) is shown and If auto-carry settings have been applied, they will appear in the index details.



Index Details (Hidden)



Index Details

Detail Sets

You can insert detail set values in the Indexing window. In PaperVision Capture Desktop, detail sets define a collection of indexes that allow multiple sets of field data to reference a single document.

For example, in an accounts payable job, index fields may be set up for check number, check date, payee, invoice number, and invoice date. If you set up all of these fields as index fields, a single document may be represented as follows:

Check Number	Check Date	Payee	Invoice Number	Invoice Date
12345	08/19/2008	ABC Corp	A0001	08/01/2008
12345	08/19/2008	ABC Corp	A0002	08/02/2008
12345	08/19/2008	ABC Corp	A0003	08/03/2008

The first three index fields (Check Number, Check Date, and Payee) will be duplicated per changing invoice number. Rather than duplicating the information in the first three fields, you can represent the first three fields as index fields and assign the remaining two fields, Invoice Number and Invoice Date, as detail sets.

Index Fields

Check Number	Check Date	Payee	Document ID *
12345	08/19/2008	ABC Corp	654

***** The system Document ID is generated behind the scenes, hidden from your view.

Detail Sets

Invoice Number	Invoice Date	Document ID *
A0001	08/01/2008	654
A0002	08/02/2008	654
A0003	08/03/2008	654

Basic and Advanced Indexing Properties

When you view properties for an index, you can toggle between basic and advanced settings to customize the properties required to index your documents. Basic settings include more commonly configured properties, and advanced settings offer more fine-grained properties to help you customize more complex indexing requirements. Basic settings may be sufficient in simpler jobs; in more advanced jobs, such as those involving barcode/OCR parsing and custom code events, advanced settings may need to be configured.

Basic properties include:

- Index Name, Type, and Format
- Validation
- Predefined Index Values
- Barcode, OCR, and Zoom Zone Configuration
- Index Masking
- Auto-Carry/Auto-Increment

Advanced properties include these additional properties:

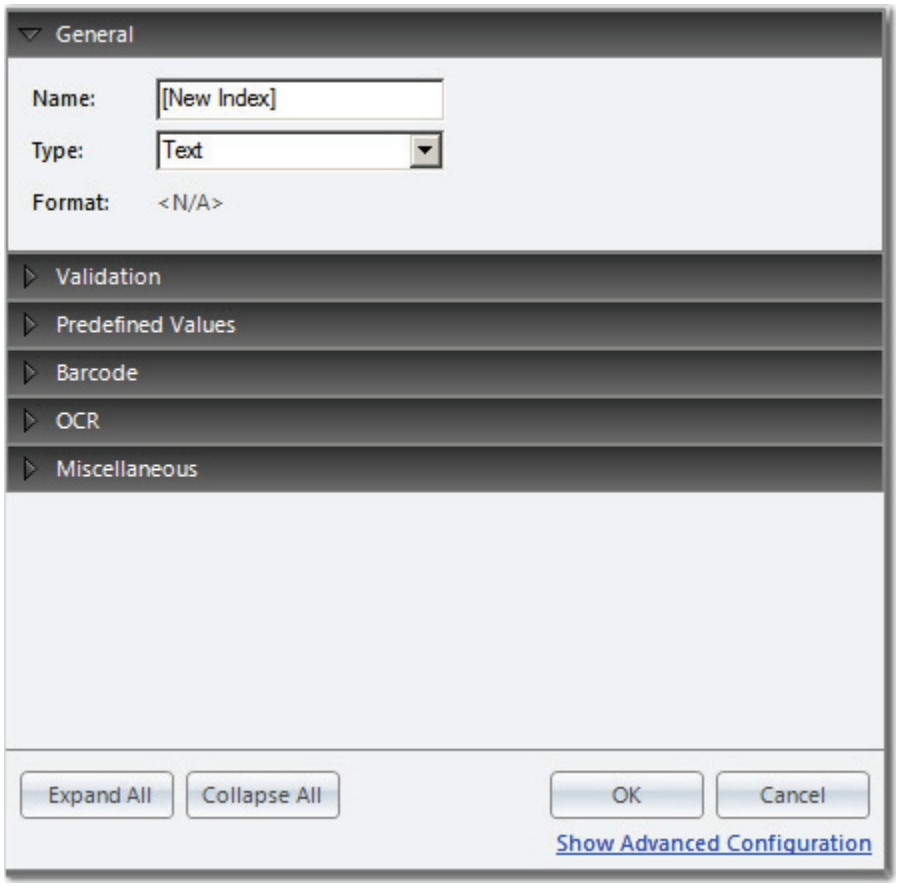
- Verification Search Strings
- Verification Regular Expressions
- Barcode Parsing
- OCR Parsing
- Custom Code Events (Index Populate and Index Validate)

Switching Between Basic/Advanced Index Properties

The following instructions describe how to toggle between basic and advanced indexing properties.

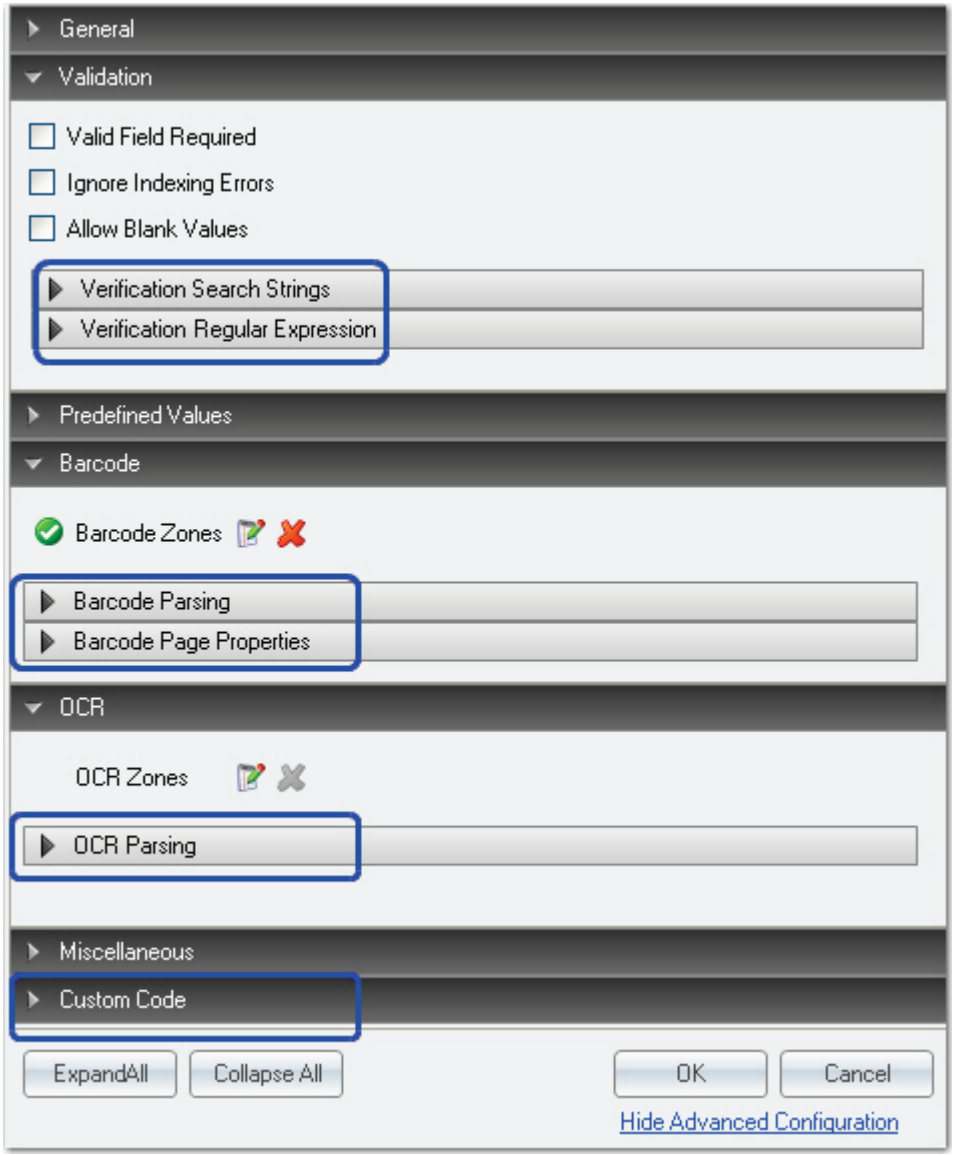
To toggle between Basic and Advanced indexing properties:

1. Click the **Add Index** icon (if you select an existing index to edit, click the **Configure Index** icon). By default, basic index properties appear. See the **Basic Indexing Properties** section for details on specific settings.



Basic Index Properties

2. Click the **Show Advanced Configuration** link in the lower right side. The advanced indexing properties appear.



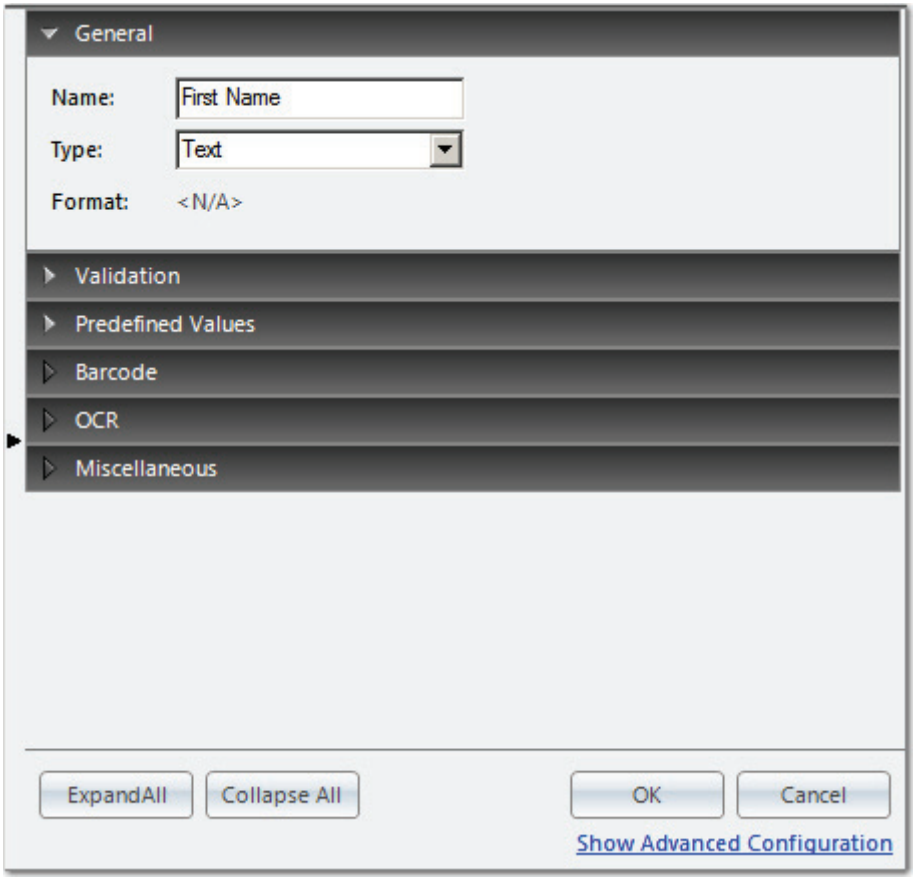
Advanced Index Properties

3. To expand one of the property groups, expand the section title bar (e.g., General, Validation, Predefined Values, etc.).
   - To expand all first-level properties under each section, click **Expand All**.
   - To close all sections, click **Collapse All**.
4. To return to basic indexing properties, click the **Hide Advanced Configuration** link.

When you add an index to the Indexing window, the expandable properties window opens and displays basic properties by default, including index validation, predefined values, barcode, OCR, and miscellaneous properties. Although the basic properties are displayed by default, you can click the **Show Advanced Configuration** link to show more complex index properties.

Miscellaneous properties include:

- Hot Key Default Value
- Zoom Zone
- Auto-Carry/Auto-Increment
- Auto-Carry Entire Index Value
- Auto-Carry Characters Following/Preceding Number
- Auto-Increment Number
- Overwrite Existing Values
- Carry Values to Copied Document
- Auto-Fill Cursor Location



Basic Index Properties

General Properties

Document index fields contain values that enable you to identify key elements of documents within a project during the capture process. For more information, see **Index Types and Formats.** The index field name, type (and format, if applicable) are assigned under the General panel menu.

Name
This editable field contains the name of the index value.

Type

The index type indicates whether it is a currency, date, number, text, etc. value.

Format

Depending on the index type you select, you can choose its format (e.g., Boolean can be True/False, Yes/No, On/Off, etc.).

Validation Properties

Validation properties allow you to determine whether index values are required to be a certain type; whether index values can remain blank; and, whether indexing errors can be skipped.

Valid Field Required

If this setting is **True**, you will be required to enter a valid index value for the field type, such as a date-formatted value for a date field. If this setting is **False**, the operator will be allowed to continue and keep the invalid value.

Ignore Indexing Errors

If this setting is **True**, incorrect operator input will be ignored and no prompt will appear for the operator. If this setting is **False**, the operator will be notified of an incorrect indexing entry.

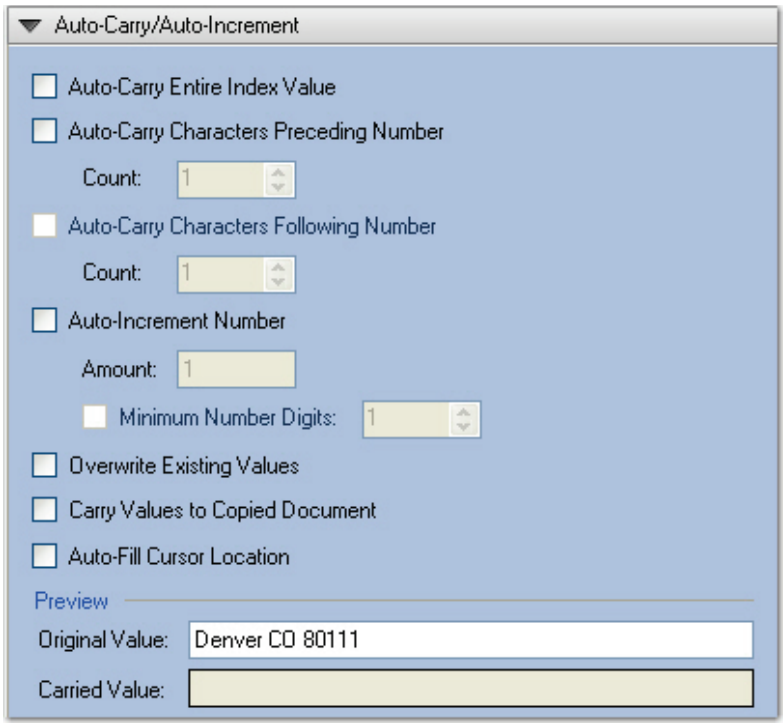
Allow Blank Values

This setting is helpful when indexing a match and merge project and a number of fields are left blank that will be matched at a later time.

- Select **True** if the index value can be omitted.
- Select **False** if it is required to enter an index value.

Auto-Carry/Auto-Increment

The Auto-Carry/Auto-Increment settings can greatly increase your productivity while hand-keying repetitive or incremental values or characters. Both tools operate during scanning (optional) and hand-keying. To configure these settings, click the ellipsis button in the **Auto-Carry/Auto-Increment** field.



Auto-Carry/Auto-Increment

> **Note**:
> Auto-Carry settings only apply when you save index values.

Auto-Carry Entire Index Value

This setting allows you to carry all characters from an index in one document to the corresponding index in the next document. You can then enable **Overwrite Existing Values** and/or **Carry Values to Copied Document.**

Auto-Carry Characters Preceding Number

This setting allows you to define the number of characters that precede a number. Your specified number of characters will carry from an index in one document to the corresponding index in the next document. For example, if you have an index that is always (or nearly always) the letters ABC followed by a number, you may not want to continuously re-enter ABC on each index value. You could set the number of characters to carry to 3. When you are keying the information, ABC would automatically get carried forward to the next document and you would only have to enter the numeric portion of the index.

Auto-Carry Characters Following Number

This setting allows you to define the number of characters that follow a number. Your specified number of characters will carry from an index in one document to the corresponding index in the next document. For example, if you have an index that is always (or nearly always) a number followed by the letters ABC, you may not want to continuously re-enter ABC on each index value. You could set the number of characters to carry to 3. When you is keying the information, ABC would automatically get carried forward to the next document and they would only have to enter the numeric portion of the index.

Auto-Increment Number

Auto-Increment takes Auto-Carry one step further. For example, if the numeric portion of the value was an incremental numeric value, you could set Auto-Carry to 3 and Auto-Increment to 1. This would increment the numeric value of any characters remaining after the first three characters by a value of one.

- The Auto-Increment Number can also be used without Auto-Carry if the value is completely numeric.
- The value entered in the Minimum Number Digits field allows you to pad the new value with zeros.
- The Preview section displays the original value and displays a preview of the carried value.

Overwrite Existing Values

By default, Auto-Carry and Auto-Increment do not fill in an index value if there is already information in the index. Selecting this check box will force Auto-Carry and Auto-Increment to update the index regardless of whether information previously existed.

Carry Values to Copied Document

By default, when documents are copied, no index values are carried through to the copies. This allows you to specify that the current index should also be copied, leaving the other indices blank.

Auto-Fill Cursor Location

If you enable this setting, operators are allowed to append to an existing index value. The setting places the cursor's focus at the end of the original index value so the original value is retained.

> **Note:**
>
> This determines whether data will be highlighted or the cursor will be placed at the end of the data when hand-keying an index that has the Auto-Carry or Auto-Fill option selected.

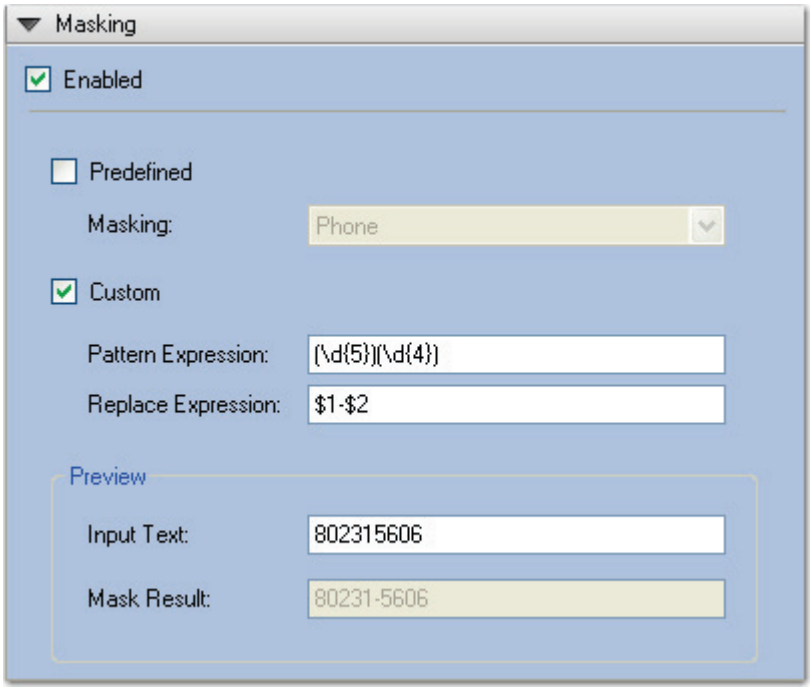
Masking

The Masking property allows you to predefine a specific format for index values entered during hand-key indexing. As operators enter index values, their entries will be formatted (masked) automatically. For example, you can predefine social security numbers to automatically insert dashes; as a result, operators only have to hand-key the 9-digit social security numbers and not the dashes.

> **Tip**:
>
> Configuring this property does not validate the operator's index value entries. Validation is performed as operators enter index values in the Indexing window.

To configure index masking:

1. In the Indexing Properties window, expand the **Masking** panel menu, and the **Masking** properties appear.



Regular Expression Mask - 5 + 4-Digit Zip Code

2. If you select a **Predefined Value**, select from the **Masking** drop-down list, and then proceed to step 6.
3. If you select a **Custom** mask, enter the **Pattern Expression**. The Pattern Expression is a regular expression that you define for the index mask. For example, for 5 + 4-digit zip codes such as 80231-5606, type the following:

   ```
   (\d{5})(\d{4})
   ```

4. If necessary, you can define a **Replace Expression** that will automatically format the operator's entry. To format an operator's 9-digit entry to appear as 80231-5606, type the following:

   ```
   $1-$2
   ```

> **Note:**
>
> If you do not define a Replace Expression, the operator's entry will not be formatted.

5. To preview how masking formats the number, enter a sample index value that an operator would hand-key in the **Input Text** field. The resulting masked index value appears in the **Mask Result** field.
6. Click **OK.**

> **Note:**
>
> Only the Text, Long Text, and Text (900) index types apply to the Masking property.

Date Regular Expression Mask

The following pattern expression formats either a one- or two-digit month and day followed by a two- or four-digit year:

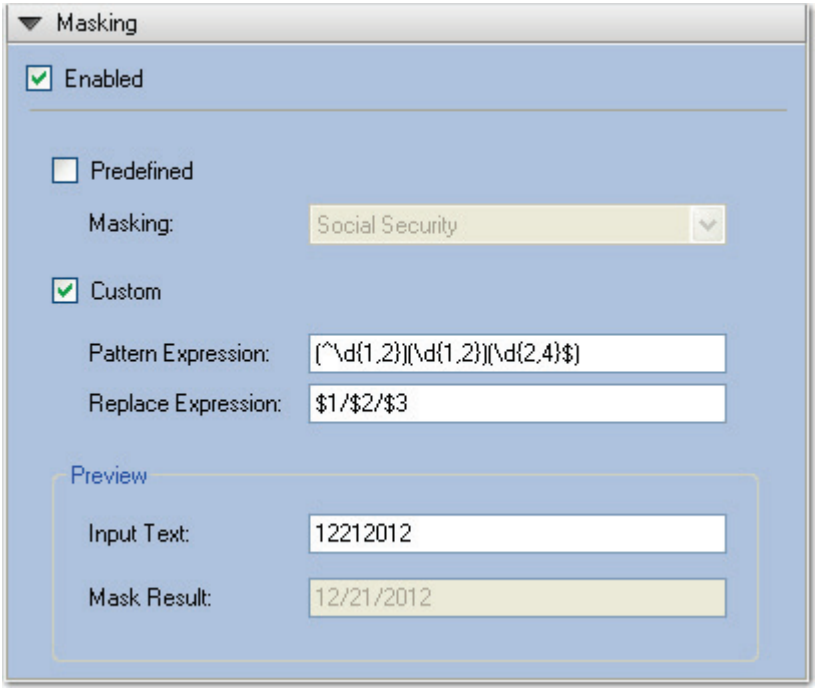
```
(^\d{1,2})(\d{1,2})(\d{2,4}$)
```

Enter the following replace expression to separate the month, day, and year with a dash:

```
$1-$2-$3
```

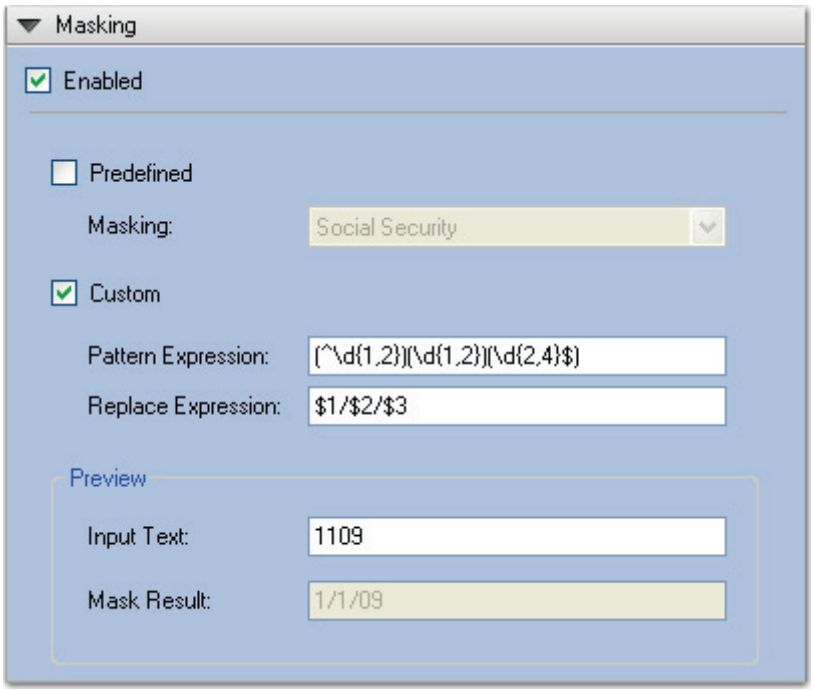
To separate the month, day, and year with a slash mark, enter:

```
$1/$2/$3
```



Two-Digit Month and Day with Four-Digit Year

The same pattern expression formats a one-digit month and day followed by a two-digit year:



One-Digit Month/Day and Two-Digit Year

Credit Card Regular Expression Mask

The following pattern expression formats a 16-digit credit card number:

(\d{4})(\d{4})(\d{4}$)(\d{4})

Enter the following replace expression to separate the digits with a dash:

$1-$2-$3-$4

Index Types and Formats

Document indices contain values that enable you to identify key elements of documents within a project during the capture process. Indices contain values that enable you to identify key elements of documents during the capture process.

The following types of index fields are supported:

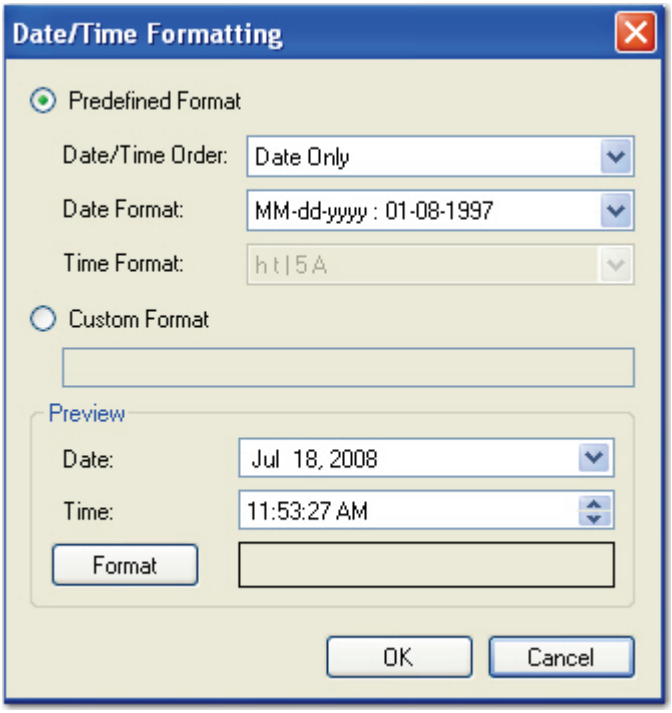
- Boolean stores Boolean values such as yes/no, on/off, and true/false.
- Currency stores currency (monetary) values.
- Date stores date/time values ranging from 12:00:00 midnight, January 1, 0001 through 11:59:59 P.M., December 31, 9999 A.D. This index type also supports searches on date ranges.
- Double Number represents a double-precision 64-bit number with values ranging from -1.79769E+308 to 1.79769E+308.
- Long Text stores textual data that exceeds 255 characters in length (up to approximately 64,000 characters in total).
- Number stores whole-number values between -2,147,483,648 and 2,147,483,647. This index type supports hyphens or dashes at the beginning of the number to indicate a negative value, but it does not support hyphens or dashes within the number, such as dashes within a social security number (555-55-5555). This index excludes these dashes from the number.
- Text stores textual data up to 255 characters in length. This type of index is the most common.
- Text(900) stores textual data up to 900 characters in length.

Date/Time Formatting

When you select a date index type, you can select from a predefined date/time format or you can customize a date/time format.

To define the date/time format:

1. In the **Indexing Properties** window, expand the **General** panel menu.
2. Select **Date** from the **Type** field.
3. Click the date link next to the Format property. The **Date/Time Formatting** dialog box opens.



Date/Time Formatting

4. Select either a **Predefined Format** (proceed to the next step) or a **Custom Format** (proceed to step 7)**.**
5. If you select a **Predefined Format**, select from the following **Date/Time Order** options:
   - Date Only
   - Time Only
   - Date/Time
   - Time/Date
6. Depending on your **Date/Time Order** selection, you can choose from the **Date/Time Format** drop-down menus.
7. If you select a **Custom Format**, enter the format in the blank field.

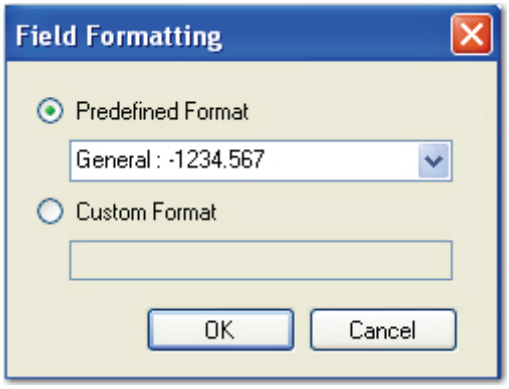
8. To preview a Predefined or Custom format, click the **Format** button in the **Preview** section.
9. If you need to preview a calendar, click the **Date** drop-down menu.
10. If you need to set the time, enter it in the **Time** field or use the up or down arrows to set the time.
11. Click **OK.**

Double Number Formatting

When you select a Double Number index type, you can select a predefined or custom format.

To define the double number format:

1. Click the ellipsis button in the right column of the **Index Format** field, which opens the **Field Formatting** dialog box.



Field Formatting

2. Select either a **Predefined Format** (proceed to the next step) or a **Custom Format** (proceed to the fourth step)**.**
3. If you select a **Predefined Format**, select from the following format types:
   - Currency
   - Fixed
   - General
   - Percent
   - Scientific
   - Standard
4. If you select a **Custom Format**, enter the format in the blank field.
5. Click **OK** when finished.

Predefined Values Properties

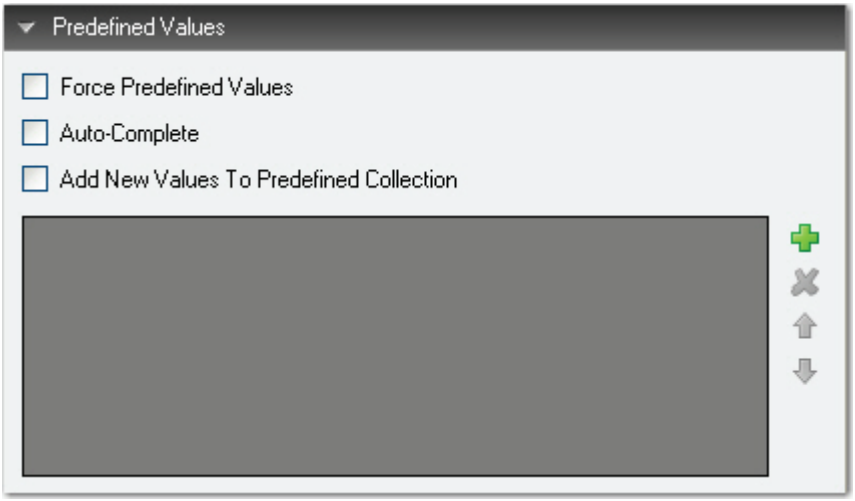
These settings allow you to predefine index field values that can be used repeatedly. You can predefine these values as you configure the index field or these entries can be added to the predefined values list. Your specified predefined values can be used with the Auto-Complete feature to finish information as you type.

Force Predefined Values

If this setting is enabled, you can only select from your predefined index values. If the entered data is not one of the predefined values, you will be alerted. If this setting is disabled, you can enter a value in the index field.

To assign predefined values to an index:

1. After you have added a new index to the Indexing window (or selected an existing index), expand the **Predefined Values** menu. The **Predefined Values** properties appear.



Predefined Values

2. For each value, click the **Add** icon.
3. Enter the value in the field.
4. To add all newly-entered values to the Predefined Values list, select **Add New Values to Predefined Collection**.
5. To automatically fill in an index field while you type select **Auto-Complete**.
6. If you only want to select from predefined index values, select **Force Predefined Values**.
    - If the entered data is not one of the predefined values, a notification will appear.
    - If you do not select this option, you will be allowed to enter a value in the index field.
7. When you are finished entering all values, click **OK**.
8. To move a value up or down the list, click the **Up** arrow or **Down** arrow.
9. When finished configuring the index, click **OK**.

To delete a value:

1. Highlight the value.
2. Click the **Delete**  icon.
3. Click **OK**.

Auto-Complete

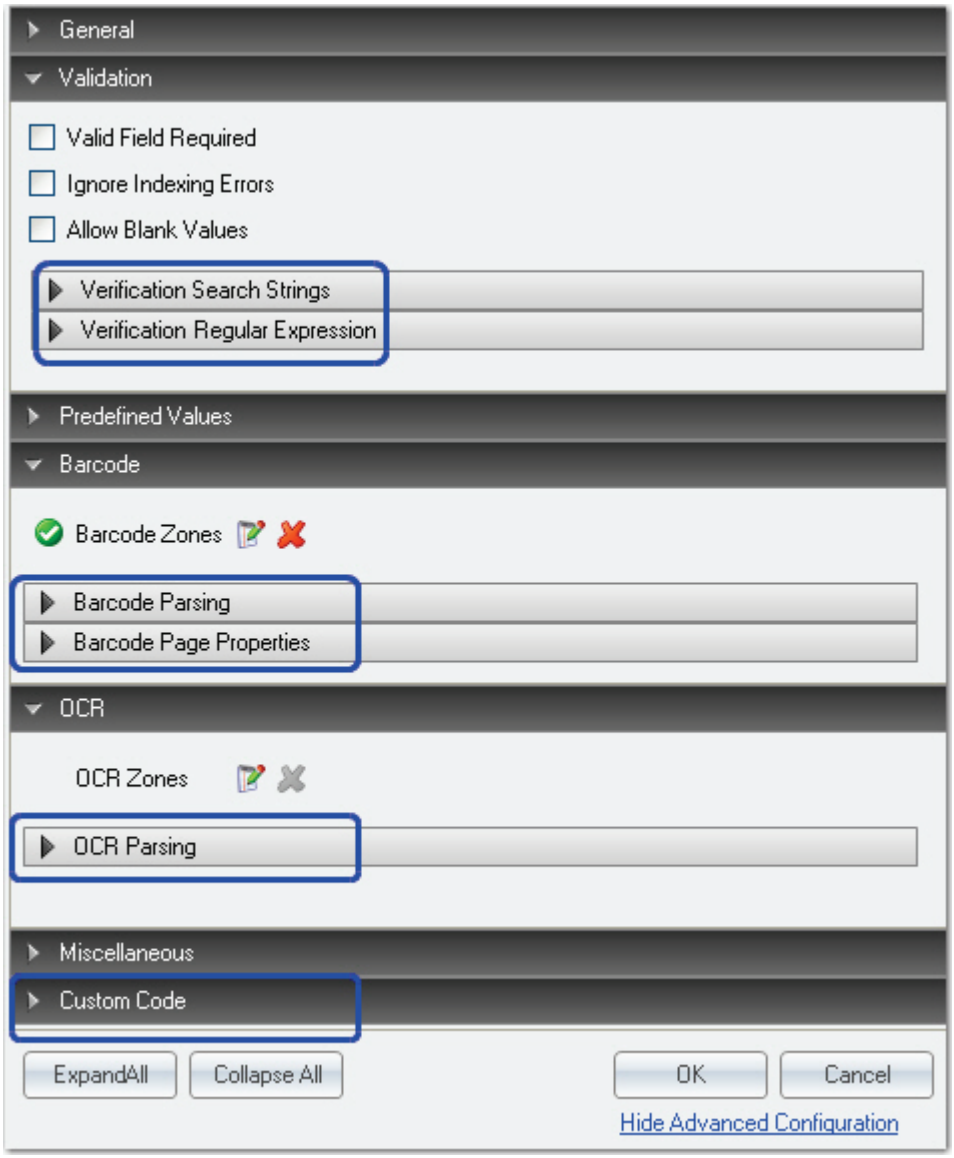
If this setting is enabled, the index field will automatically be completed as you type.

Add New Values to Predefined Collection

If this setting is enabled, all new entered values are added to the Predefined Values list.

Advanced Indexing Properties

Advanced indexing properties enable you to configure index verification search strings/regular expressions, barcode and OCR parsing, and custom code events. Additional (Miscellaneous) properties available for configuration include automated index validation/populating of custom code events. To view advanced properties for an index, Click the **Add Index** icon (if you select an existing index to edit, click the **Configure Index** icon). By default, basic index properties appear. See the **Basic Indexing Properties** section for details on specific settings. To view advanced properties, click **Show Advanced Configuration** in the bottom right of the Indexing Properties window.



Advanced Indexing Properties

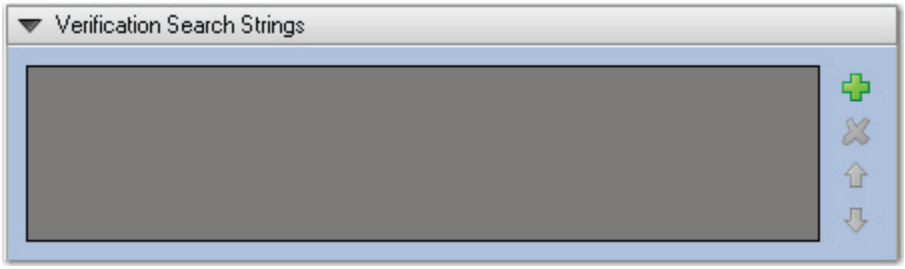
Verification Search Strings

The Verification Search Strings setting is used to validate index values when you save index values, tab to the next field, or export the batch. To ensure the accuracy of hand-key indexing, you can define multiple search strings that can be verified when you save your index values or when you execute the Verify Index Values operation. For example, you can assign individual characters or numbers to search for during the index verification process. By default, the verification process will highlight the first document in the batch that contains a blank value. However, you can exclude blank values from the index verification process by removing <Blank> from the list of search strings.

Depending on your index verification settings in the user **Display Preferences** (Verify Starts from Current Document Forward or Verify Starts at the Beginning of the Batch), the index verification process starts with the appropriate document in the batch and will highlight the next document that contains your defined search strings.

To assign verification search strings:

1. Click the **Add Index** icon (if you select an existing index to edit, click the **Configure Index** icon). By default, basic index properties appear. See the **Basic Indexing Properties** section for details on specific settings.
2. In the properties window, switch to the advanced properties view (click the **Show Advanced Configuration** link in the lower right).
3. Expand the **Validation** menu.
4. Expand the **Verification Search Strings** submenu.



Verification Search Strings

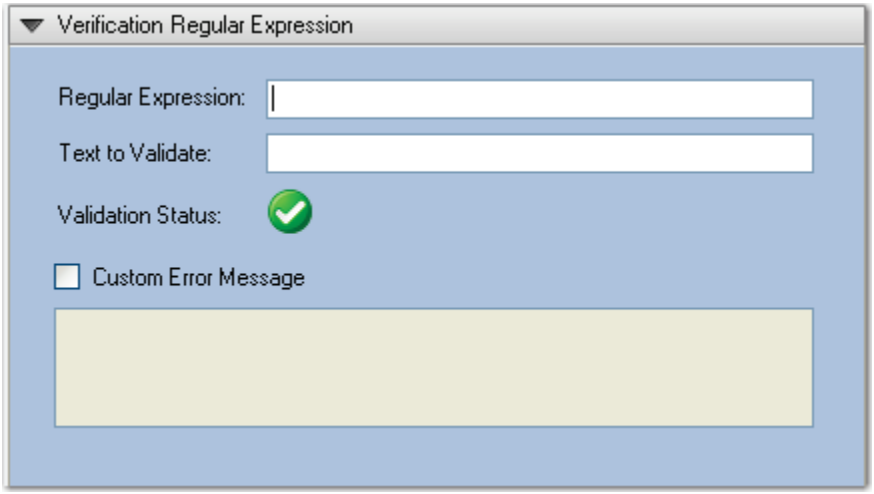
5. Click **Add**, and a blank line will appear.
6. Enter the search string.
7. Enter subsequent search strings, if necessary.
8. To remove a search string, highlight the string, and then click the **Remove** icon.
9. To move a search string up or down in the list, highlight the string, and then click the **Up** or **Down** arrow.
10. Click **OK**.

Verification Regular Expression

You can create a regular expression to validate data entry. A regular expression is a pattern of text that consists of ordinary characters (for example, letters A through Z) and special characters, known as metacharacters. The pattern describes one or more strings to match when searching a body of text. The regular expression serves as a template for matching a character pattern to the string being searched.

To assign verification regular expressions:

1. Click the **Add Index** icon (if you select an existing index to edit, click the **Configure Index** icon). By default, basic index properties appear. See the **Basic Indexing Properties** section for details on specific settings.
2. In the properties window, switch to the advanced properties view (click the **Show Advanced Configuration** link in the lower right).
3. Expand the **Validation** menu.
4. Expand the **Verification Regular Expression** submenu.



Verification Regular Expression

5. Enter the regular expression.
6. Enter the text to validate. Your entry will automatically be validated.
    - A successful validation displays with a green icon.
    - Invalid entries display with a red icon.
- Optionally, you can create a custom error message if users enter incorrect index values. To enable this feature, select the **Custom Error Message** option.
- Enter the custom error message text.
- Click **OK**.

Barcode Zone Properties

During index configuration, you can manage barcode zones that you have defined for the selected index.

To manage barcode zones during index configuration:

1. In the Index Properties window, expand the **Barcode** panel menu. If you configured a barcode zone for the index, you will be able to edit and/or delete the zone.

Barcode Zones

2. Click the **Manage** icon. The **Manage Zones** grid appears. See the **Manage Zones** section in the **Home Page** chapter for more information.
3. To delete the barcode Zone, click the **Delete** icon.

Barcode Parsing

During indexing configuration, you can configure a text delimiter or a regular expression to parse specific index fields from a barcode. You can then specify which field's index is parsed from the barcode (e.g., you can select the third field's index so only the last four digits of a social security number are parsed). Optionally, you can verify that an exact number of index fields results from the parse operation (e.g., three index fields indicative of a social security number in the format xxx-xx-xxxx).

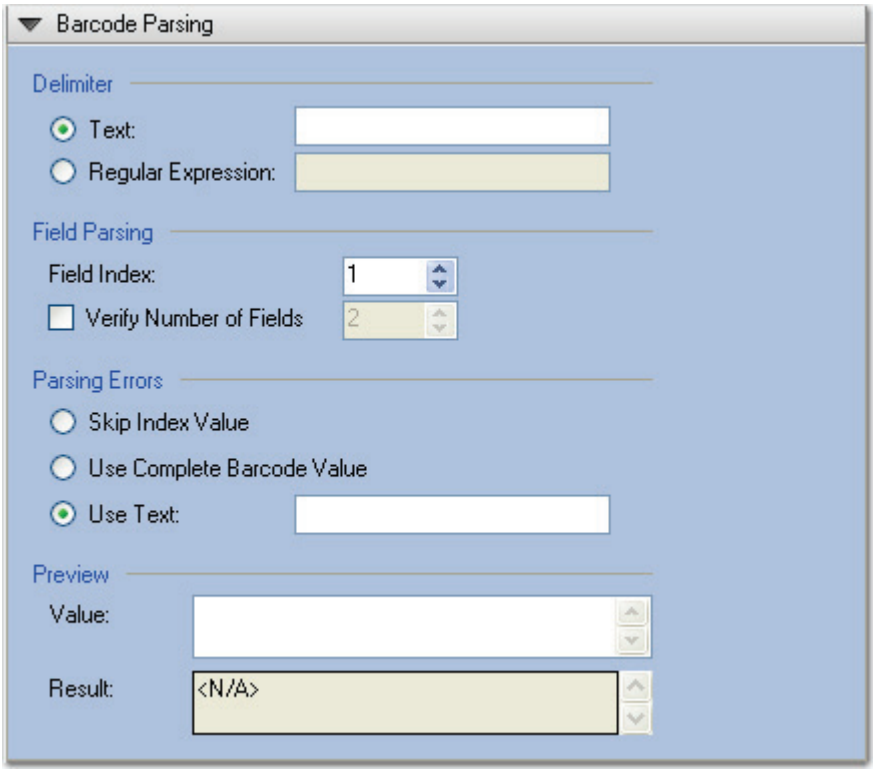
> **Note**:
>
> The **Verify Number of Fields** setting is intended to verify that an exact number of index fields (two or more) results from the parse operation.

If errors occur during barcode parsing, such as when the parsed number of index fields differs from your specified number of fields, you can select one of three subsequent actions. First, the entire index value can be skipped (therefore, no barcode parsing occurs). In the second option, the entire barcode value is used (therefore, no barcode parsing occurs). In the last option, you can specify the text used as the parsed value (e.g., you can enter "unknown value").

To configure barcode parsing:

1. Click the **Add Index**  icon (if you select an existing index to edit, click the **Configure Index**  icon). By default, basic index properties appear. See the **Basic Indexing Properties** section for details on specific settings.
2. In the properties window, switch to the advanced properties view (click the **Show Advanced Configuration** link in the lower right).
3. Expand the **Barcode** menu.

4. Expand the **Barcode Parsing** submenu. The **Barcode Parsing** properties appear.



Barcode Parsing

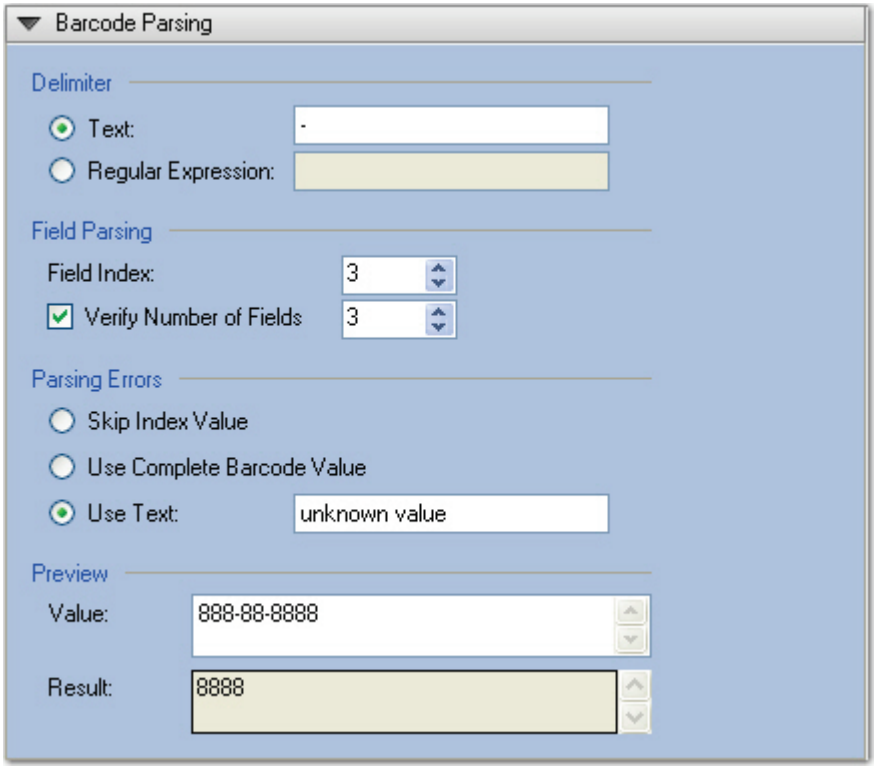
5. In the **Delimiter** section, select whether to use a text delimiter or regular expression to split the original index value into fields. If you enter an invalid text delimiter or regular expression, the error symbol will appear to the right of the field.

> **Note**:
>
> Additional information on regular expressions can be located at: **http://msdn.microsoft.com/library/default.asp?url=/library/en-us/script56/html/js56reconIntroductionToRegularExpressions.asp**

6. In the **Field Parsing** section, specify the field index position from which to parse data.
7. Optionally, you can verify that an exact number of index fields (two or more) results from the parse operation.

   For example, you can set the **Field Index** value to "3" to parse only the last four digits of a social security number that exists in the format xxx xx xxxx. You can then select the **Verify Number of Fields** option to verify that three index fields (indicative of a social security number) result from the parse operation.

8. In the **Parsing Errors** section, select the action that will be executed if parsing errors occur:
   - **Skip Index Value**: The entire index value is skipped, so no barcode parsing occurs.
   - **Use Complete Barcode Value**: The complete barcode value is used, so no barcode parsing occurs.
   - **Use Text**: Your specified text is used as the parsed value.
9. In the **Preview** section, you can enter a sample index value to ensure the text delimiter or regular expression parses the value correctly.



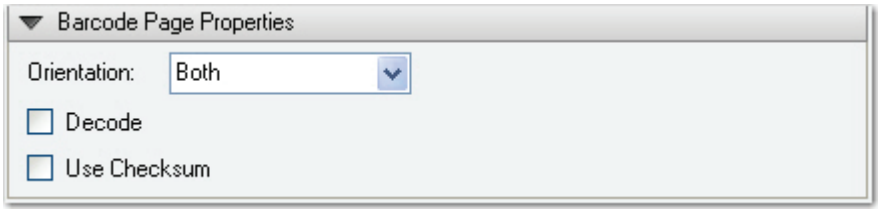
Barcode Parsing - Configured

Barcode Page Properties

Advanced barcode page properties include settings for barcode orientation detection, decoding special characters, and checksum error detection.

To configure advanced barcode page properties:

1. Click the **Add Index** icon (if you select an existing index to edit, click the **Configure Index** icon). By default, basic index properties appear. See the **Basic Indexing Properties** section for details on specific settings.
2. In the properties window, switch to the advanced properties view (click the **Show Advanced Configuration** link in the lower right).
3. Expand the **Barcode** menu.
4. Expand the **Barcode Page Properties** submenu. The **Barcode Page Properties** appear.



Barcode Page Properties

Orientation

PaperVision Capture Desktop detects horizontal and vertical barcodes with skew angles of no more than fifteen degrees from the horizontal and vertical axes, respectively. Horizontal barcode detection is slightly faster than vertical barcode detection. If you are unsure of the expected barcode orientation or if the documents might contain barcodes with different orientations, select **Both** from the drop-down menu.

Decode

Some barcode types, such as Code 128, do not represent their data as ASCII characters. Other barcode types, such as Code 3 of 9, use special characters to extend the basic character set to include the entire ASCII set. When this setting is enabled, barcode values are converted into human-readable ASCII strings. For example, if the barcode uses escape characters, as in "*%K123%M?*", and the Decode property is **True**, then "[123]" will be returned. If the Decode property is **False**, the raw barcode is returned.

> **Note:**
>
> You should enable this setting unless the barcode results should not be converted into ASCII strings. For example, this setting should be disabled if you are detecting Code 3 of 9 barcodes that represent dates using the slash mark ”r;/” character (e.g. 01/01/1999). If this setting is enabled, no results are returned because ”r;/0” and ”r;/1” are not valid ASCII character.

Use Checksum

A checksum is an error detection process where additional characters are appended to a barcode to ensure more accurate readings. Enable this setting if you want the checksum to be recognized during the scanning process

OCR Zone Properties

During index configuration, you can manage OCR zones that you have defined for the selected index.

To manage OCR zones during index configuration:

1. In the Index Properties window, expand the **OCR** panel menu. If you configured an OCR zone for the index, you will be able to edit and/or delete the zone.

OCR Zones

2. Click the **Manage** icon. The **Manage Zones** grid appears. See the **Manage Zones** section in the **Home Page** chapter for more information.
3. To delete the OCR Zone, click the **Delete** icon.

OCR Parsing

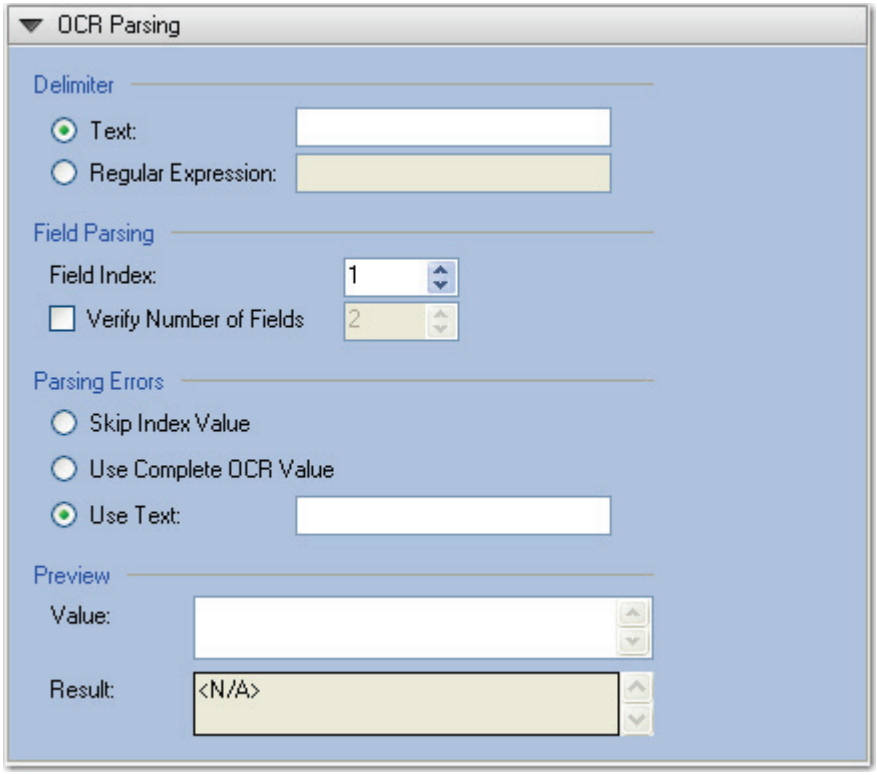
During indexing configuration, you can configure a text delimiter or a regular expression to parse specific index fields from OCR text. You can then specify which field's index is parsed (e.g., the fourth field's index from a credit card number). Optionally, you can verify that an exact number of index fields results from the parse operation (e.g., four index fields indicative of a complete credit card number).

> **Note**:
>
> The **Verify Number of Fields** setting is intended to verify that an exact number of index fields (two or more) results from the parse operation.

If errors occur during OCR parsing, such as when the parsed number of index fields differs from your specified number of fields, you can select one of three subsequent actions. First, the entire index value can be skipped (therefore, no OCR parsing occurs). In the second option, the entire OCR value is used (therefore, no OCR parsing occurs). In the last option, you can specify the text used as the parsed value (e.g., you can enter "unknown value").

To configure OCR parsing:

1. Click the **Add Index** icon (if you select an existing index to edit, click the **Configure Index** icon). By default, basic index properties appear. See the **Basic Indexing Properties** section for details on specific settings.
2. In the properties window, switch to the advanced properties view (click the **Show Advanced Configuration** link in the lower right).
3. Expand the **OCR** menu.
4. Expand the **OCR Parsing** submenu. The **OCR Parsing** properties appear.



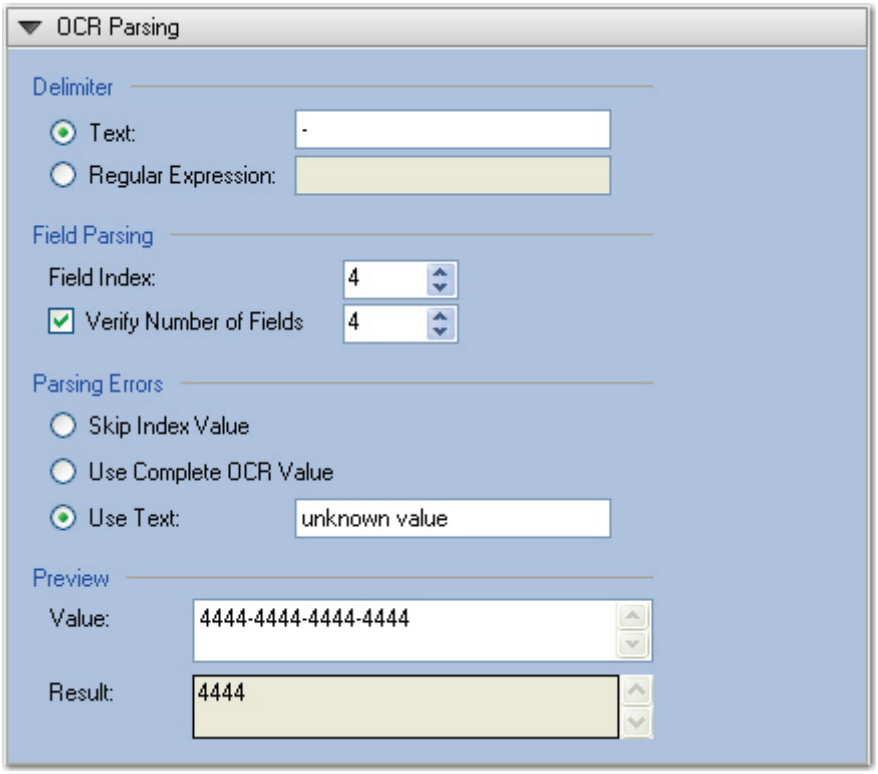
OCR Parsing

5. In the **Delimiter** section, select whether to use a text delimiter or regular expression to split the original value into fields. If you enter an invalid text delimiter or regular expression, the error symbol will appear to the right of the field.

> **Note**:
>
> Additional information on regular expressions can be located at: **http://msdn.microsoft.com/library/default.asp?url=/library/en-us/script56/html/js56reconIntroductionToRegularExpressions.asp**

6. In the **Field Parsing** section, specify the field index position from which to parse data.
7. Optionally, you can verify that an exact number of index fields (two or more) results from the parse operation.

8. For example, you can set the **Field Index** value to "4" to parse only the last four digits of a credit card number. You can then select the **Verify Number of Fields** option to verify that four index fields (indicative of a social security number) result from the parse operation.
9. In the **Parsing Errors** section, select the action that will be executed if parsing errors occur:
    - **Skip Index Value**: The entire index value is skipped, so no OCR parsing occurs.
    - **Use Complete OCR Value**: The complete OCR value is used, so no OCR parsing occurs.
    - **Use Text**: Your specified text is used as the parsed value.
10. In the **Preview** section, you can enter a sample index value to ensure the text delimiter or regular expression parses the value correctly.



OCR Parsing (Configured)

Miscellaneous Properties

The remaining properties can be defined during index configuration:

Hot Key Default Value

As operators are keying in index fields and press the assigned hot key, the specified default value will populate the index field.

Zoom Zone

If you configured a zoom zone for the index, you can edit and/or delete the zone.

To edit the zone, click the **Manage**  icon. The **Manage Zones** grid appears. See the Manage Zones section of the **Home Page** chapter for more information. To delete the zone, click the **Delete**  icon

Custom Code

During index configuration, you can configure custom code events that are triggered when index fields are populated or validated.

To configure advanced custom code properties:

1. Click the **Add Index** icon (if you select an existing index to edit, click the **Configure Index** icon). By default, basic index properties appear. See the **Basic Indexing Properties** section for details on specific settings.
2. In the properties window, switch to the advanced properties view (click the **Show Advanced Configuration** link in the lower right).
3. Expand the **Custom Code** menu, and the following properties can be configured:

Index Populating and Index Validating Custom Code Events

The advanced custom code event properties, Index Populating and Index Validating, allow you to select either Visual Basic or C# code to configure an action triggered immediately after an index field is populated (and you return to re-enter the index value) or validated by the system. The Index Validate event is triggered after you return to edit an index value, re-enter the index value, and then proceed to a subsequent index field (or save the edited index value).

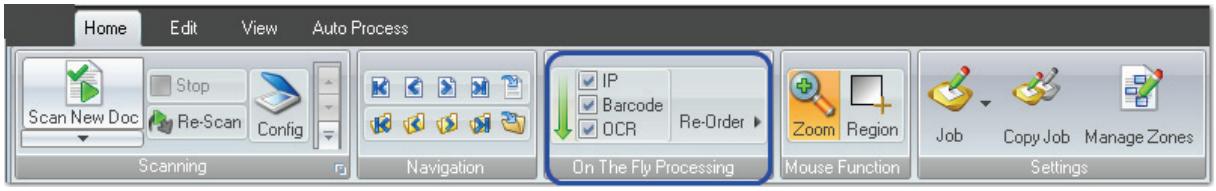
If you use either of these Custom Code Events to change an index value, the Indexing window will remain synchronized using the UIRefreshLevel property (e.g., "base.UIRefreshLevel =UIRefreshLevel.Index"). See the section on API Functions in the Custom Code Configuration section for a list of API functions and associated enumerations that can be used within Custom Code.

Custom Code Indexing Events

On the Fly Processing

You can enable on the fly image processing, barcoding, and zonal OCR processing while you scan documents in the Home page. Multiple processes can be enabled, and the order in which these processes execute can be modified. However, image processing must be designated as the first or last process. While you scan documents, for example, you could enable on the fly image processing to execute binary noise removal; you could then execute barcode detection to automatically break documents; lastly, you could execute zonal OCR processing to read social security numbers to populate an index.

Once you configure barcode zones, OCR zones, and IP zones/filters, you must enable the corresponding task in the On the Fly Processing toolbar group. On the Fly processing for barcode and OCR zones is configured through the Home page with the Region mouse function. On the Fly image processing is configured in the **Auto Process** page, where you can apply and configure one or more image processing filters.



On the Fly Processing - All Processes Enable

Configuring Barcoding On the Fly

To enable barcoding while scanning documents (on the fly processing), you must first define one or more barcode zones. You can add multiple barcode zones on each page, and you can apply unique properties to each zone. As you draw a barcode zone on an image, the barcode value will appear immediately in an adjacent flyout window. After you have configured all barcode zones, you can enable automated barcode processing in the Home page.

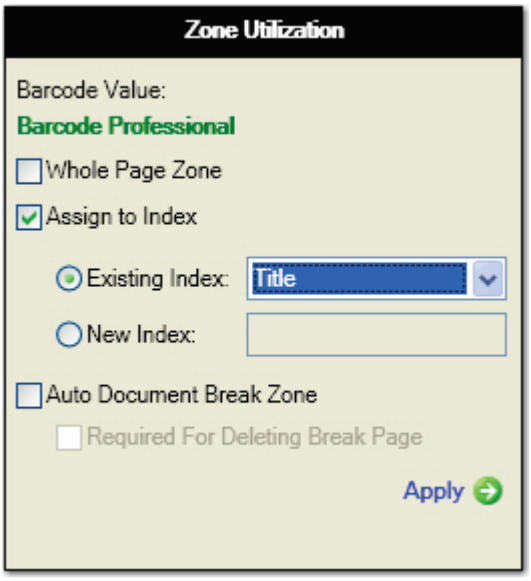
To configure barcoding on the fly:

1. In the Home page, click the **Region** 

icon in the **Mouse Function** toolbar group.
2. Drag the crosshair cursor around the barcode on the image. The Region toolbar will appear next to your zone, where you can select how to use the zone.

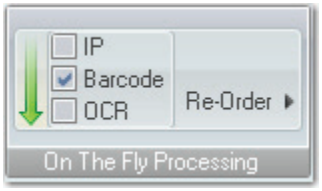
Region Toolbar

3. Adjust the borders of the barcode zone if necessary.
4. Click the **Add Barcode Zone**  icon. The zone's barcode value will populate the **Zone Utilization** flyout window.



Zone Utilization

5. If you want to assign the zone to an index, select one or more options:
   - **Whole Page Zone**: Assign the zone to the entire page.
   - **Assign to Index**: Select the existing index from the drop-down list, or assign to new index (enter the new index name).
   - **Auto Document Break Zone**: In addition to assigning the zone to the entire page or to an index, you can also assign the zone to automatically break documents. Optionally, you can require that the OCR zone be read successfully in order to delete the break page.
6. Click the **Apply** link in the **Zone Utilization** dialog box.
7. To enable barcoding on the fly, select the **Barcode** check box in the **On the Fly Processing** toolbar group.



On the Fly Processing (Barcode)

Configuring OCR On the Fly

To enable zonal OCR while scanning (on the fly processing), you must first define one or more OCR zones. You can add an infinite number of OCR zones on each page, and you can apply unique properties to each zone. As you draw a zone around text, you can ensure the text can be read successfully prior to processing batches.

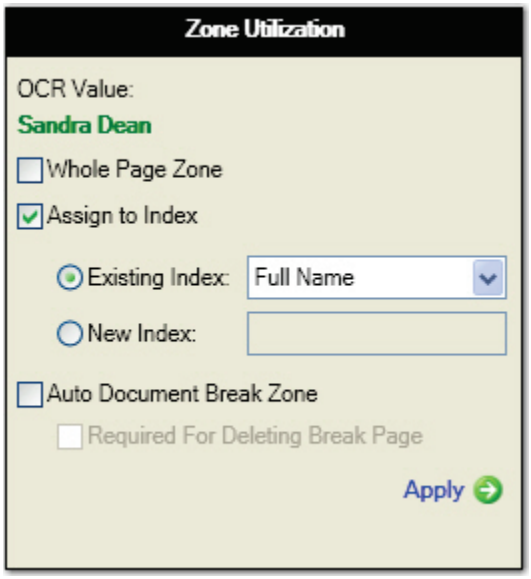
To configure OCR processing on the fly:

1. In the Home page, click the **Region** 

icon in the **Mouse Function** toolbar group.
2. Drag the crosshair cursor around the text on the image. The **Region** toolbar will appear next to your zone, where you can select how to use the zone.

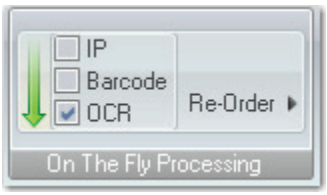
Region Toolbar

3. Adjust the borders of the OCR zone if necessary.
4. Click the **Add OCR Zone to Job**  icon. The zone's OCR value will populate the **Zone Utilization** flyout window.



Zone Utilization

5. In the **Zone Utilization** flyout window, select how to use the OCR zone:
   - **Assign to Index**: Select the existing index from the drop-down list, or assign to new index (enter the new index name).
   - **Auto Document Break Zone**: In addition to assigning the zone to the entire page or to an index, you can also assign the zone to automatically break documents. Optionally, you can require that the OCR zone be read successfully in order to delete the break page.
6. Click **Apply** in the **Zone Utilization** flyout window.
7. To enable zonal OCR processing on the fly, select the **OCR** check box in the **On the Fly Processing** toolbar group.



On the Fly Processing (OCR)

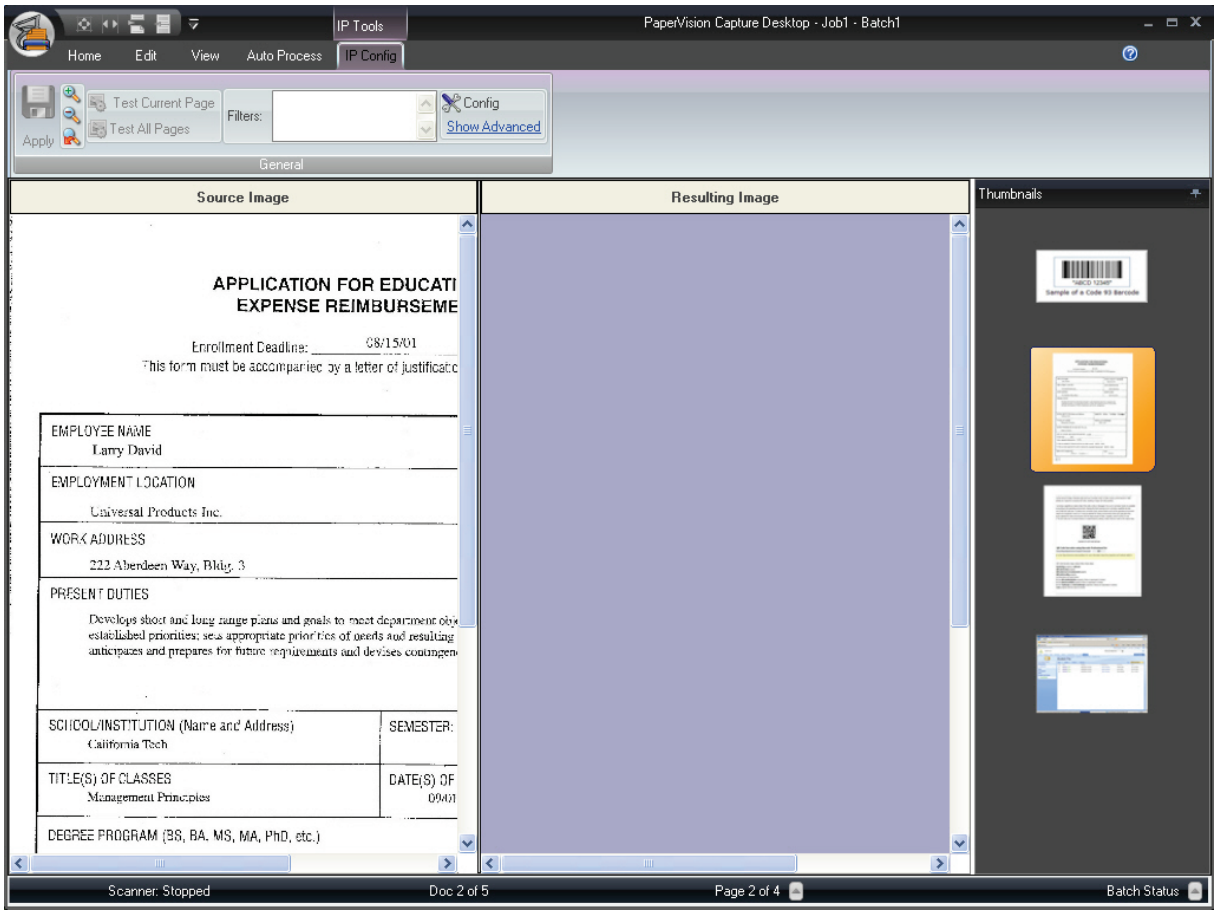
Configuring Image Processing On the Fly

To enable image processing while scanning (on the fly processing), you must first configure the appropriate image processing zones and filter(s) in the **Auto Process** page. You can apply IP filters on the entire image or specific zones. In both instances, you can test the current page or all pages to ensure the IP filters are acceptable.

> **IMPORTANT!**
>
> It is not recommended to apply the page deletion IP filter on the first page during on-the fly processing, as all pages will be deleted.

To configure image processing on the fly:

1. Open the **Auto Process** page, and then click the **Config**  icon in the **IP** toolbar group. The **IP Zone Configuration** context sensitive page appears.



IP Config Context Sensitive Page

2. In the IP Config page, click the **Config** icon. The **Image Processing Filters** dialog appears. Filters supported in zones are marked with an asterisk (*).

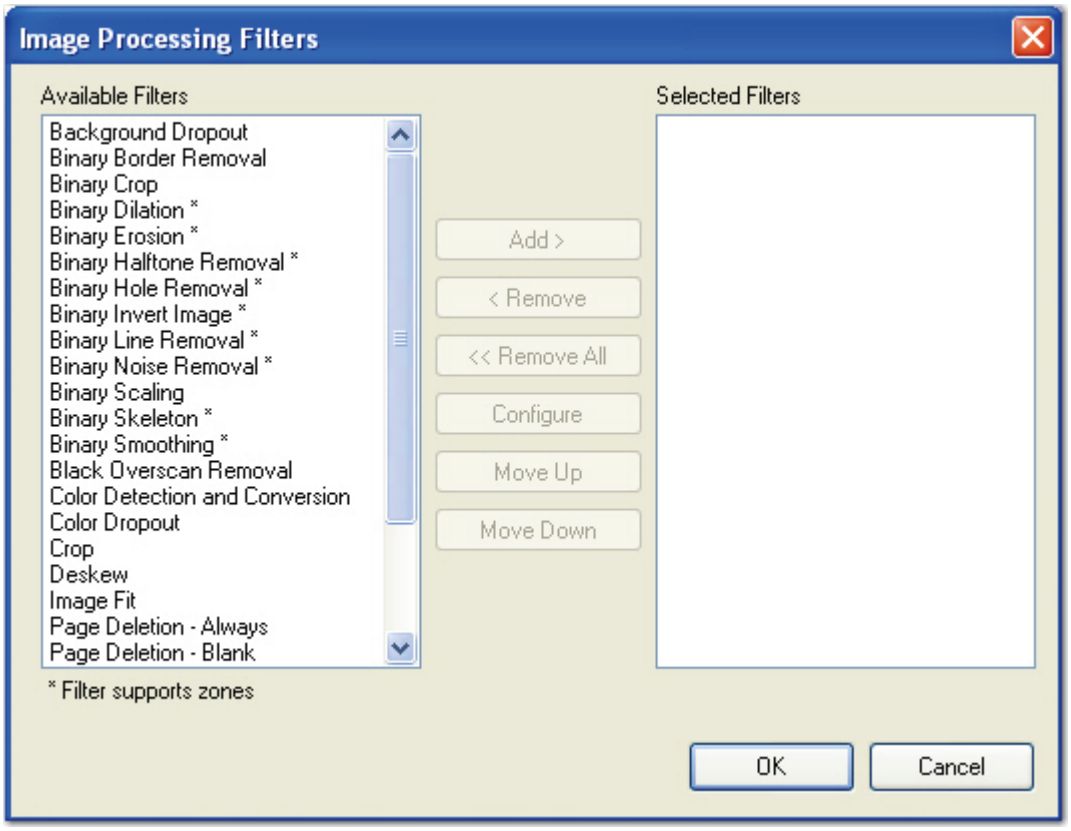


Image Processing Filters

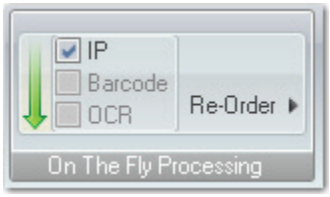
3. From the **Available Filters** list, highlight the filter(s), and then click **Add**.
4. To configure a selected filter, highlight the filter in the **Selected Filters** list, and then click **Configure**.
   - To remove a filter from the **Selected Filters** list, highlight the filter, and then select the **Remove** button.
   - To remove all filters from the **Selected Filters** list, click the **Remove All** button.
   - To move a filter up or down in the **Selected Filter** list, highlight the filter, and then select **Move Up** or **Move Down**.

> **Note**:
>
> See the **Image Processing Filters** section in Chapter 6 for descriptions of each filter.

5. Click **OK** after you have configured all filters.

> **Note**:
>
> For complete details on configuring IP Zones, see the **IP Configuration** section in Chapter 6.

6. When you are finished configuring all filters in the **IP Config** page, click the **Apply**  icon.
7. Open the **Home** page.
8. To enable IP on the fly, select the **IP** check box in the **On the Fly Processing** toolbar group.



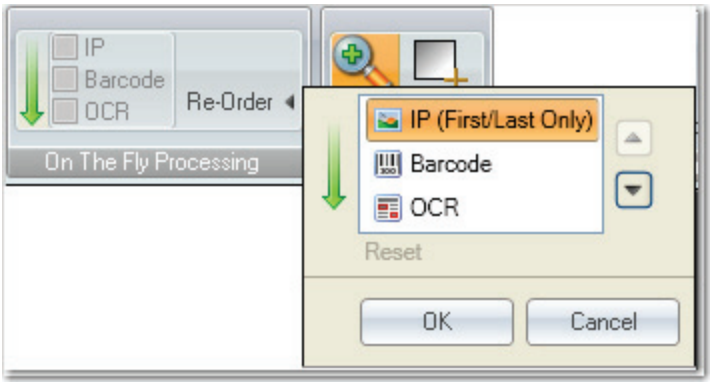
On the Fly Processing (IP)

Re-ordering Processes

In the Home page, you can customize the order in which tasks are executed during on the fly processing. However, image processing must be designated as the first or last process.

To re-order the processes:

1. In the Home page, select the **Re-order** button in the **On the Fly Processing** toolbar group. The **Reset** flyout window appears.



Reset

2. Select the process in the list, and then click the up or down arrow.
3. Click **OK**.

Mouse Functions

Mouse functions enable you to switch between the Zoom and Region (barcode, OCR, zoom zone, etc.) operations. By default, the Zoom mouse function is enabled. The Region mouse function allows you to apply a zone (you can define the type of zone and select its settings) on the image.

To zoom in on the image:

1. If the Zoom mouse function is not already enabled, click the **Zoom** 

icon.
2. Click on the area of the image to enlarge.

> **Note**:
>
> To reset the original view of the image, right-click on the image, and then select **Reset Image** from the context menu.

To apply a region on the image:

1. Select the **Region** 

icon.
2. Drag the crosshair cursor around the area on the image. A toolbar appears above the zone where you can execute one of several operations described below.

Region Toolbar

Deleting Regions

To delete the region you just applied, click the **Delete**  icon.

Assigning Barcode Values to Indexes

To assign a barcode value to a new or existing index, click the **Assign Barcode Value**  icon.

Assigning OCR Values to Indexes

To assign an OCR value to a new or existing index, click the **Assign OCR Value**  icon.

Assigning Barcode Zones to Jobs

To assign a barcode zone to a job, click the **Assign Barcode Zone to Job** icon. See the **Barcode Zone Configuration** section for more information.

Assigning OCR Zones to Jobs

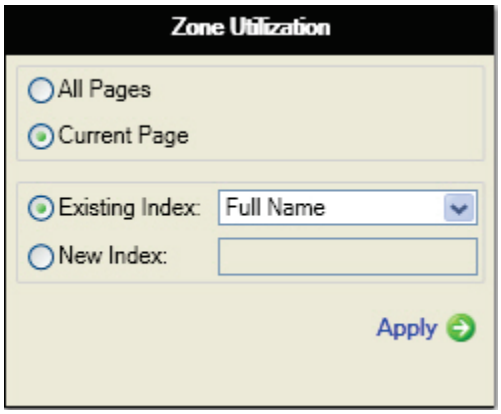
To assign an OCR zone to a job, click the **Assign OCR Zone to Job** icon. See the **OCR Zone Configuration** section for more information.

Assigning Zoom Zones

You can assign a zoom zone to the current page or all pages in the batch. You can then assign the zone to a new or existing index. Alternatively, you can apply the zoom zone to the entire page instead of a region.

To assign a zoom zone:

1. Click the **Add Zoom Zone to Job** icon. The **Zone Utilization** flyout window appears.



Zone Utilization

2. Select whether to apply the zoom zone to **All Pages** or the **Current Page**.
3. If you want to assign the zone to an index, select one option:
   - Assign to Existing Index (select the index from the drop-down list)
   - Assign to New Index (enter the new index name)
4. Click **Apply**.

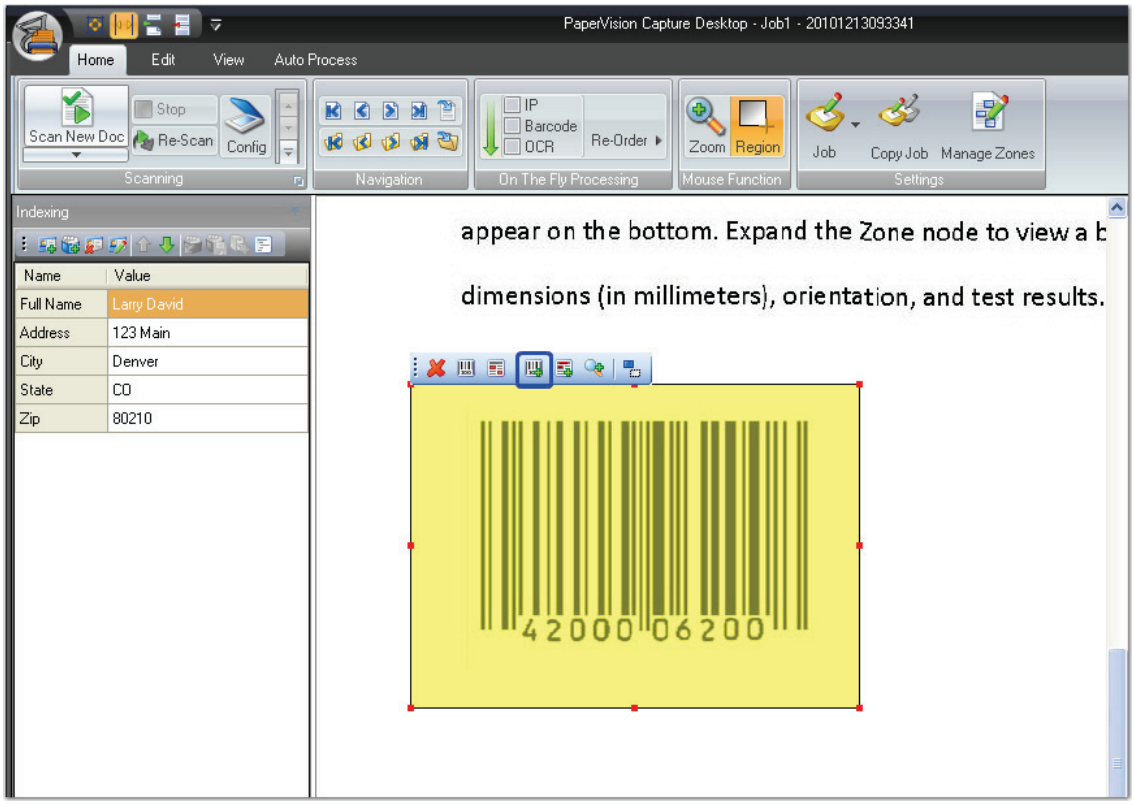
Barcode Zone Configuration

In PaperVision Capture Desktop, you can use barcodes to automatically populate index values and insert document breaks. You can also apply barcode zones manually so you can easily extract and index barcode data that may shift across pages and documents. PaperVision Capture Desktop recognizes one- and two-dimensional, black and white, and color barcodes.

> **Note**:
>
> To prevent unpredictable results, new line characters and tab characters will not be removed from 2D barcodes during barcode recognition.

In the Home page, you can configure barcode zones to automatically insert document breaks or populate index fields while scanning documents (on the fly). Alternatively, you can configure a barcode reading process that executes during post-processing (after scanning and indexing) when you process or export the batch from the Auto Process page. During barcode zone configuration, all barcodes types will be recognized automatically during the reading process (i.e., by default, all barcode types are selected when you view properties for a barcode zone). As you add a barcode zone in the Home page, you can view its value immediately in an adjacent flyout window, and you can define whether the zone automatically breaks documents, populates indexes, or both.

> **Note**:
>
> Use of the binary scaling image processing filter can improve the recognition rate of barcode detection. For more information, see the **Image Processing Filters** section.



Add Barcode Zone to Job

Adding Barcode Zones to Jobs

You can define barcode zones in the Home page with the Region mouse function which allows you to insert the zone. Barcode zones can be configured to read barcodes while you scan (on the fly processing) or when you process or export the batch (after scanning and indexing) in the Auto Process page. The following instructions describe how to assign barcode zones to jobs.

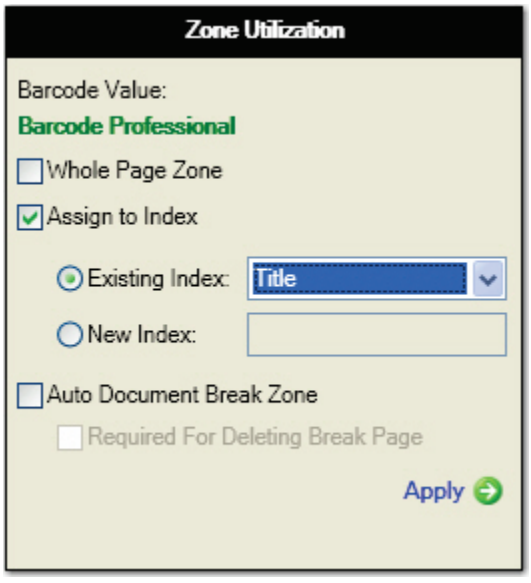
To add a barcode zone to a job:

1. In the Home page, click the **Region** 

icon in the **Mouse Function** toolbar group.
2. Drag the crosshair cursor around the barcode on the image. The **Region** toolbar will appear next to your zone, where you can select how to use the zone.

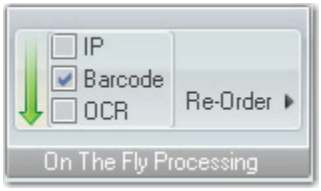
Region Toolbar

3. Adjust the borders of the barcode zone if necessary.
4. Click the **Add Barcode Zone to Job**  icon. The zone's barcode value will populate the **Zone Utilization** flyout window.



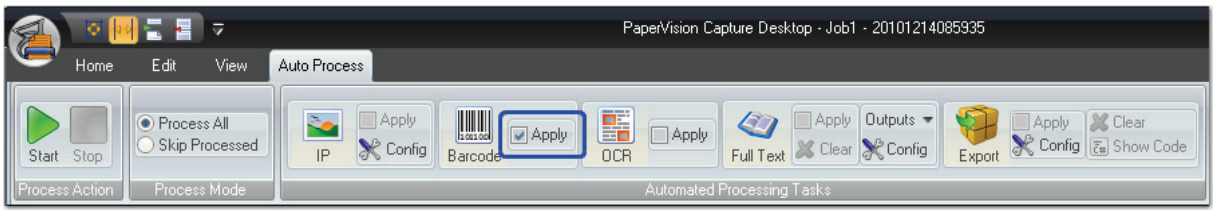
Zone Utilization

5. If you want to assign the zone to an index, select one or more options:
   - **Whole Page Zone**: Assign the zone to the entire page.
   - **Assign to Index**: Select the existing index from the drop-down list, or assign to new index (enter the new index name).
   - **Auto Document Break Zone**: In addition to assigning the zone to the entire page or to an index, you can also assign the zone to automatically break documents. Optionally, you can require that the OCR zone be read successfully in order to delete the break page.
6. Click **Apply** in the **Zone Utilization** flyout window to save the settings.
7. To execute the barcode reading process while scanning, select the **Barcode** check box in the **On the Fly Processing** toolbar group. If you want to apply the barcode reading process after scanning and indexing (when you process or export the batch from the Auto Process page), proceed to the next step.



Barcode On the Fly Processing

8. Open the **Auto Process** page, and then select the **Barcode** check box in the **Automated Processing Tasks** toolbar group. See **Chapter 5 - Auto Process Page** for more information.



Automated Processing Tasks - Barcode

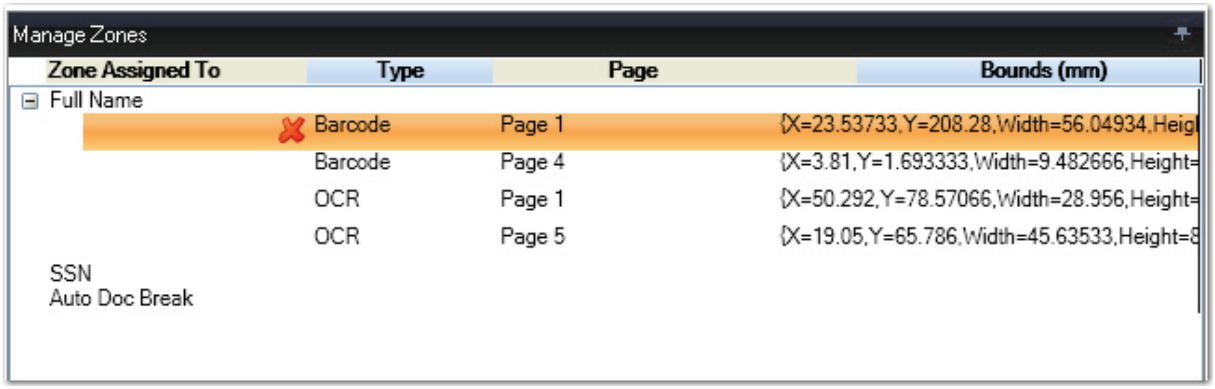
9. If you are ready to execute automated processing, click the **Start**  icon.

Editing Barcode Zones

You can view all barcode zones that you have applied by clicking the **Manage Zones** icon in the toolbar. You can view each barcode zone's X and Y coordinates, dimensions (in millimeters), orientation, and page location.

To edit a barcode zone:

1. In the Home page, click the **Manage Zones** icon. The **Manage Zones** window appears.

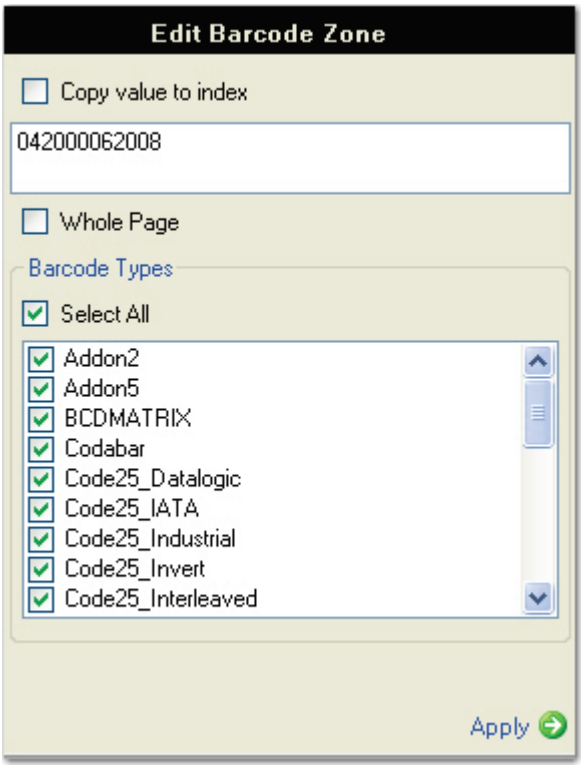


Manage Zones

2. Highlight the barcode zone. The **Manage Zone** toolbar appears.

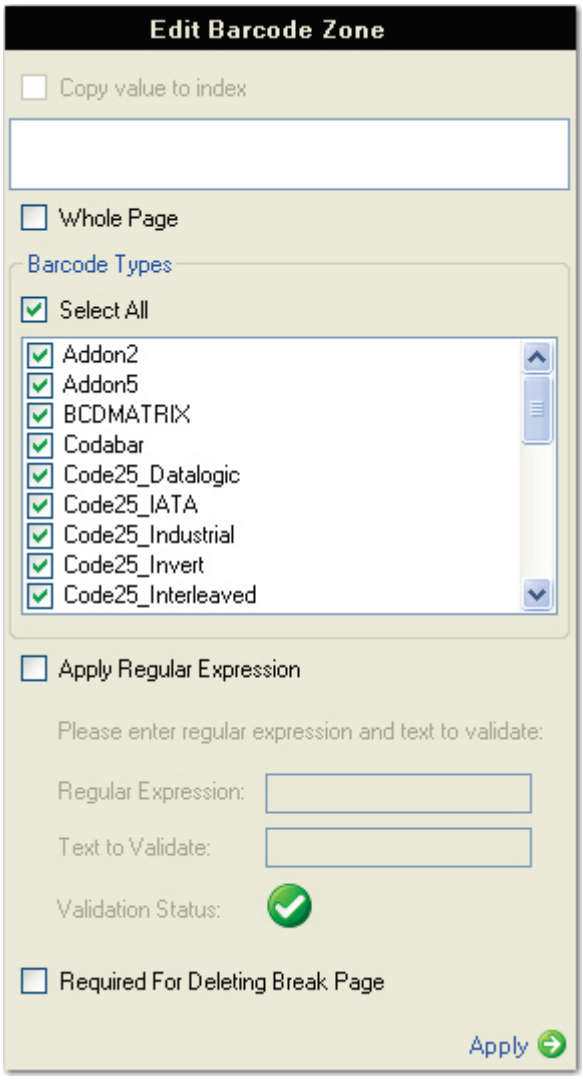
Manage Zone Toolbar

3. Click the **Edit Zone** icon. The **Edit Barcode Zone** flyout window appears. The barcode/OCR value appears in the first field. If you configured the zone as a barcode break page, see the next step.



Edit Barcode Zone

4. If you configured the zone as a barcode break page (in **Job Configuration**), you can assign a regular expression (Apply Regular Expression). In addition, you can delete the break page if the barcode zone is read successfully (Required For Deleting Break Page).



Edit Barcode Zone (for Auto Document Break Zone)

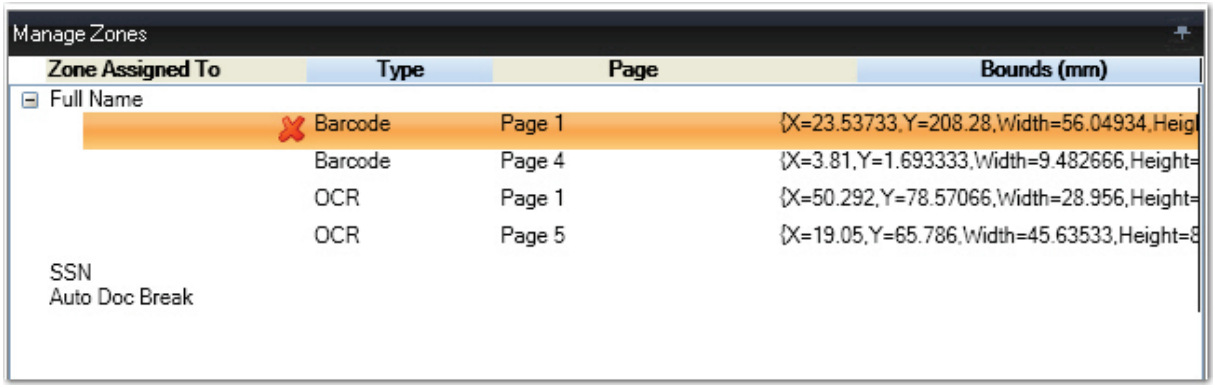
5. Adjust the barcode zone properties, if applicable.
6. To copy the value to the selected index, select **Copy value to index**.
7. To apply the barcode zone to the entire page, select **Whole Page.**
8. If you adjusted the borders of the zone and want to save the new dimensions, click the **Save** icon.
9. Click **Apply** to save your changes.

Deleting Barcode Zones

You can delete barcode zones that you have applied by clicking the **Manage Zones** icon in the toolbar.

To delete a barcode zone:

1. In the Home page, click the **Manage Zones** icon. The **Manage Zones** window appears.



Manage Zones

2. Highlight the barcode zone. The **Edit Zone** toolbar appears.

Edit Zone Toolbar

3. Click the **Delete** icon.
4. Click **Yes** to confirm the deletion.

> **Note**:
>
> Alternatively, you can select the **Delete** icon next to the barcode zone row in the grid.

Supported Barcode Types

To streamline the barcode reading process, all supported one- and two-dimensional barcode types (listed below) will be detected. If desired, you can exclude certain types from being read during the barcode reading process.

The following two-dimensional (2D) barcode types are supported:

- DataMatrix
- PDF417
- QR Code
- Royal Post
- Australian Post
- Intelligent Mail

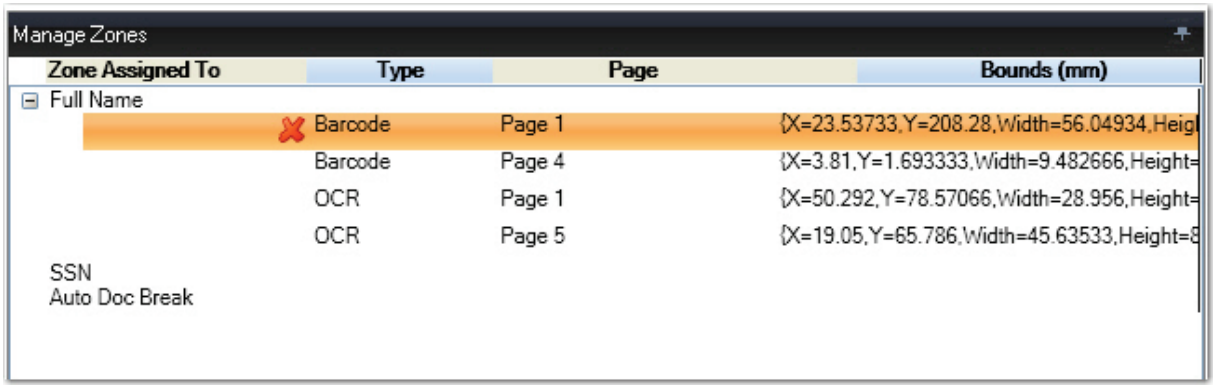
> **Note**:
>
> To prevent unpredictable results, new line characters and tab characters will not be removed from 2D barcodes during barcode recognition.

The following one-dimensional (1D) barcode types are supported:

- Addon 2
- Addon 5
- BCD Matrix
- Codabar
- Code25 Datalogic
- Code25 IATA
- Code25 Industrial
- Code25 Interleaved
- Code25 Invert
- Code25 Matrix
- Code 32
- Code 39
- Code 93
- EAN 13
- EAN 8
- Postnet
- Type 128
- UCC 128
- UPC-A
- UPC-E

To exclude certain barcode types from being detected:

1. In the Home page, click the **Manage Zones** icon. The **Manage Zones** window appears.

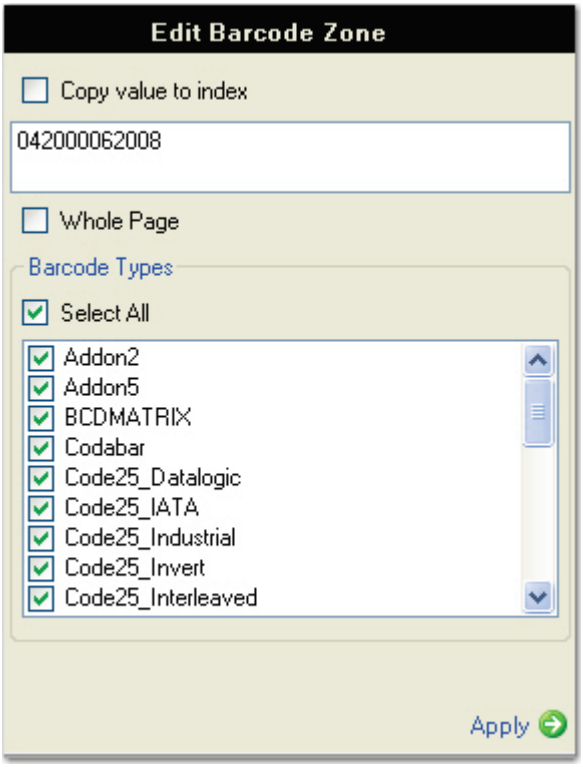


Manage Zones

2. Highlight the barcode zone. The **Edit Zone** toolbar appears.

Edit Zone Toolbar

3. Click the **Edit Zone** icon. The **Edit Barcode Zone** flyout window appears.



Edit Barcode Zone

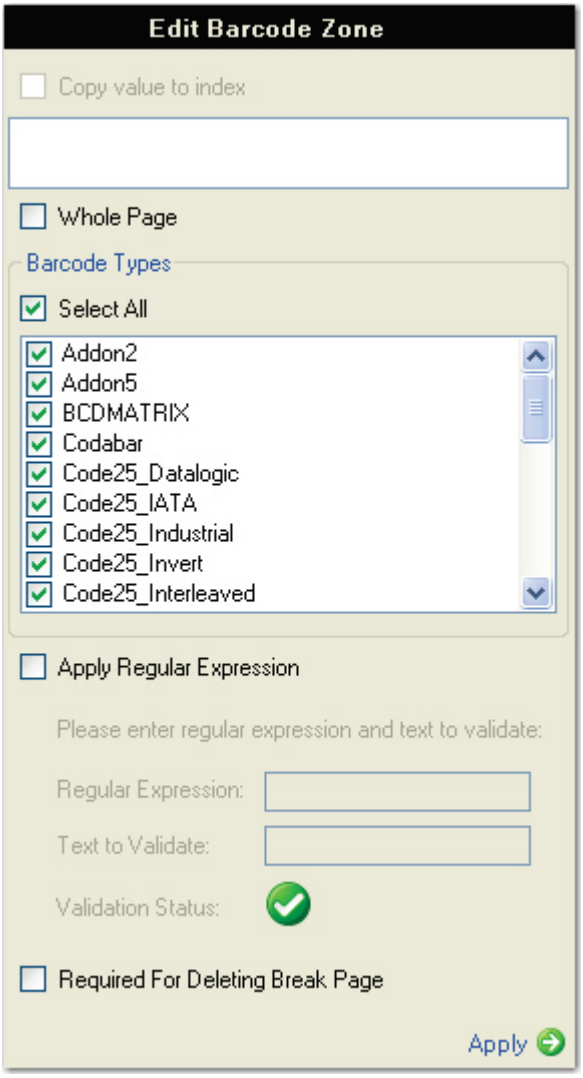
4. To exclude certain types from being detected, remove the check marks from the applicable check boxes in the **Barcode Types** section.
5. Click **Apply** to save your changes.

Regular Expression Verification (for Auto Document Breaks)

This field is applicable when you define Auto Document Breaks with barcodes. If you enter an exact value or regular expression into the **Regular Expression Verification** field, a document break is only inserted when the system reads barcodes matching your exact value or regular expression. If you leave this field blank, any barcode read by the system will cause a document break to be inserted. A regular expression is a pattern of text that consists of ordinary characters (for example, letters A through Z) and special characters, known as metacharacters. The pattern describes one or more strings to match when searching a body of text. The regular expression serves as a template for matching a character pattern to the string being searched.

To configure a regular expression:

1. In the **Edit Barcode Zone** flyout window, click **Apply Regular Expression** to enter a regular expression.



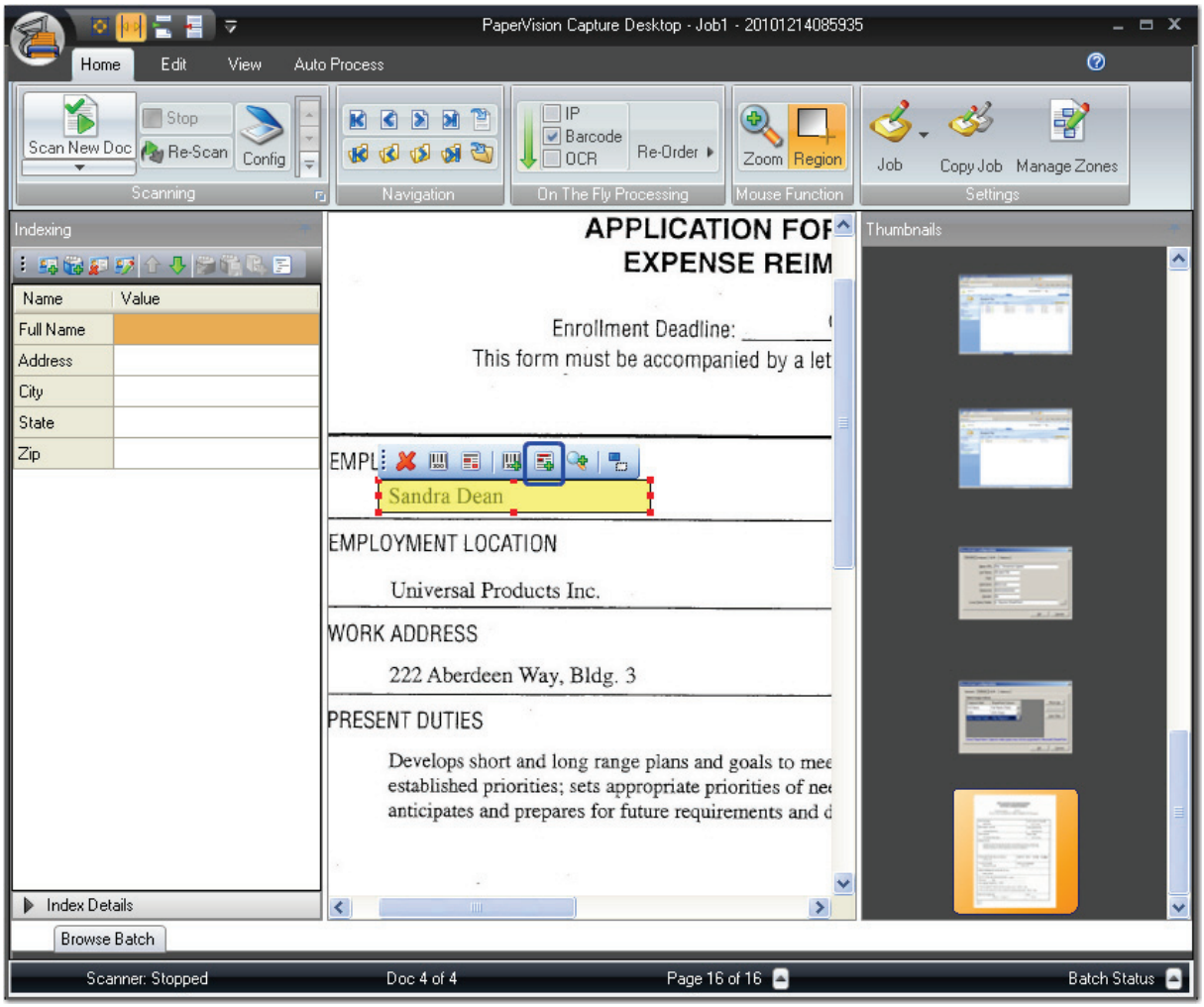
Regular Expression

2. In the **Regular Expression** field, enter the regular expression.
3. Enter the text to validate.
   - A successful validation displays with a green icon.
   - Invalid entries display with a red icon.

OCR Zone Configuration

Zonal OCR processing is performed by the Open Text® engine. PaperVision Capture Desktop recognizes text contained in OCR zones that you define in the Home page. You can define the OCR zones that will be recognized while you scan (on the fly processing) or when you export the batch in the Auto Process page. The Region mouse function equips your cursor to draw an OCR zone on the image.

> **Note**:
>
> New line characters will be removed during OCR processing.



Home Page - Adding an OCR Zone

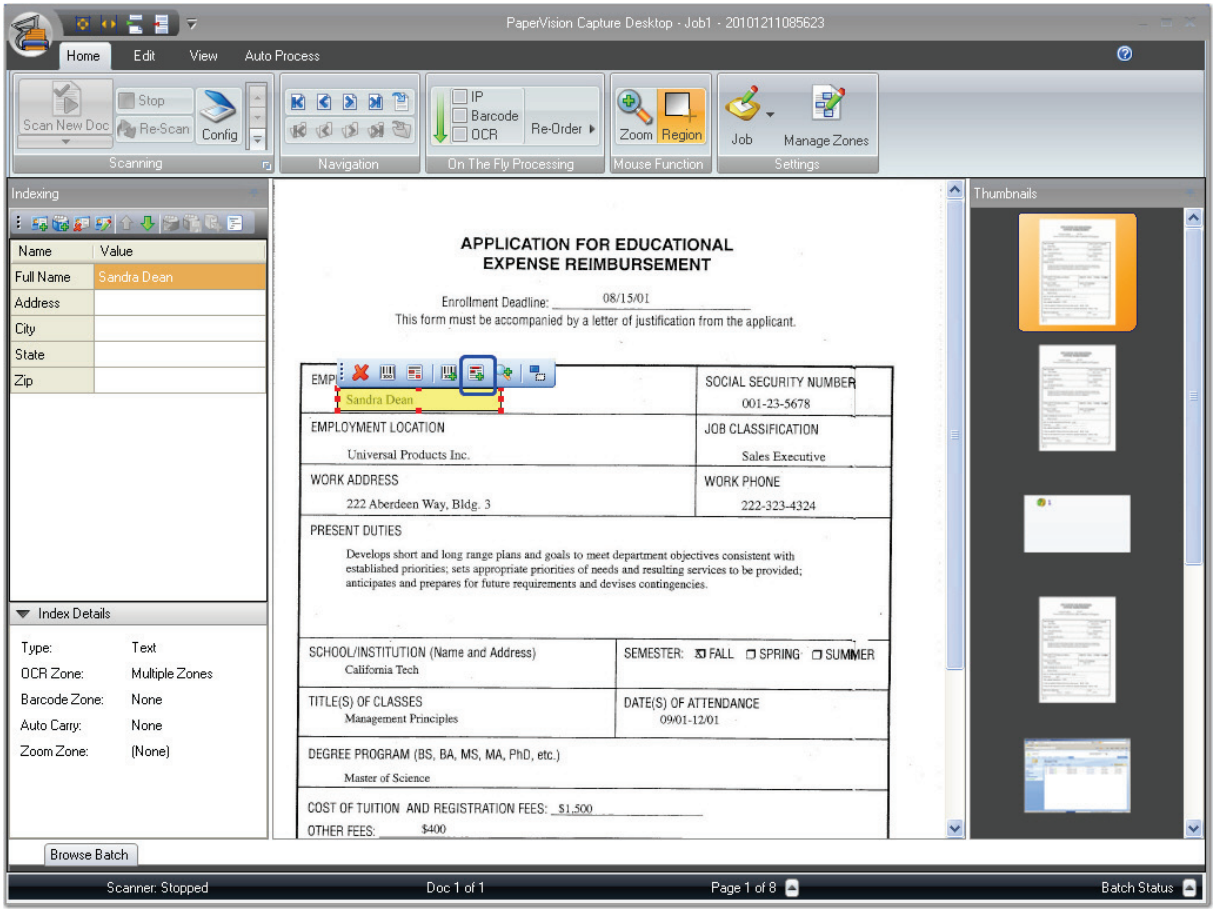
Adding OCR Zones to Jobs

You can define OCR zones in the Home page to extract text while you scan (on the fly processing) or when you execute automated processing (after scanning and indexing) in the Auto Process page. The following instructions describe how to assign OCR zones to jobs.

To add an OCR zone to a job:

1. In the Home page, click the **Region** 

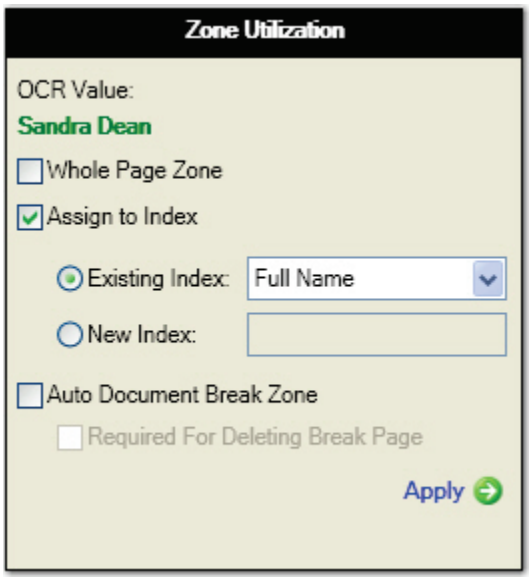
icon in the **Mouse Function** toolbar group.
2. Drag the crosshair cursor around the text on the image. The **Region** toolbar will appear next to your zone, where you can select how to use the zone.



OCR Zone Configuration - Add OCR Zone to Job

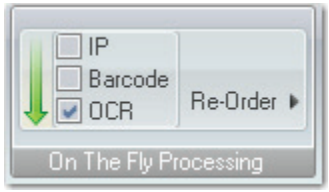
3. Adjust the borders of the OCR zone if necessary.

4. Click the **Add OCR Zone to Job** icon. The zone's OCR value will populate the **Zone Utilization** flyout window.



OCR Zone Utilization

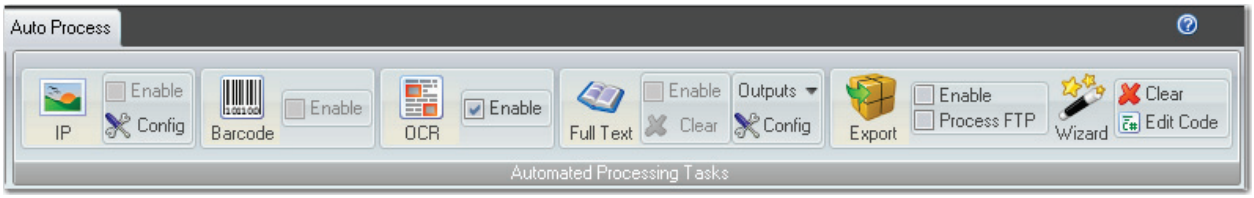
5. In the **Zone Utilization** flyout window, select how to use the OCR zone:
   - **Assign to Index**: Select the existing index from the drop-down list, or assign to new index (enter the new index name).
   - **Auto Document Break Zone**: In addition to assigning the zone to the entire page or to an index, you can also assign the zone to automatically break documents. Optionally, you can require that the OCR zone be read successfully in order to delete the break page.
6. Click **Apply** in the **Zone Utilization** flyout window.
7. If you want to extract text from this OCR zone while you scan documents (on the fly), select the **OCR** check box in the **On the Fly Processing** toolbar group in the Home page. You are now ready to scan documents.



OCR Applied On the Fly

8. Alternatively, if you only want OCR text to be extracted during automated processing (post-scanning and indexing), proceed to the **Auto Process** page.

9. In the Auto Process page, select the **Enable** check box next to **OCR**. You are new ready to execute zonal OCR during automated processing. For more information on executing Zonal OCR automated processing, see **Chapter 5 - Auto Process Page.**



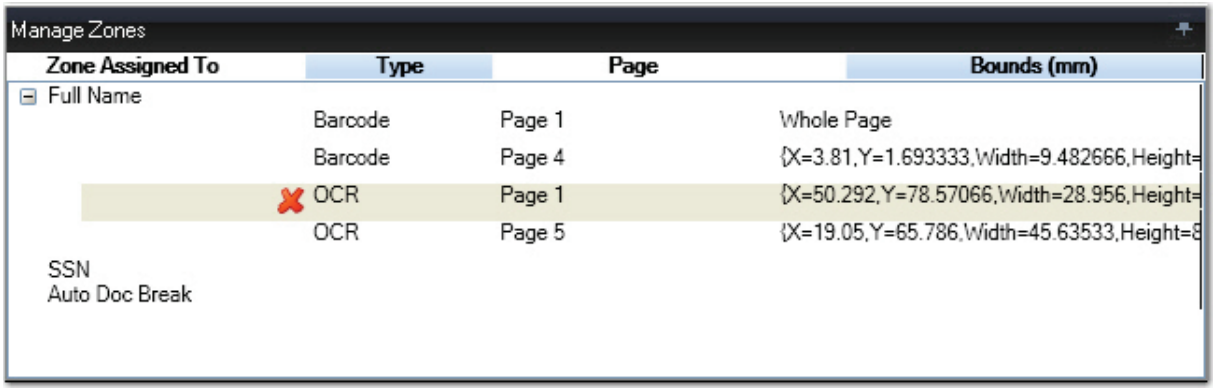
Zonal OCR - Enabled During Automated Processing

Editing OCR Zones

OCR Zones can be edited with the Manage Zones  operation in the Home page.

To edit an OCR zone:

1. In the Home page, click the **Manage Zones**  icon. The **Manage Zones** window appears.

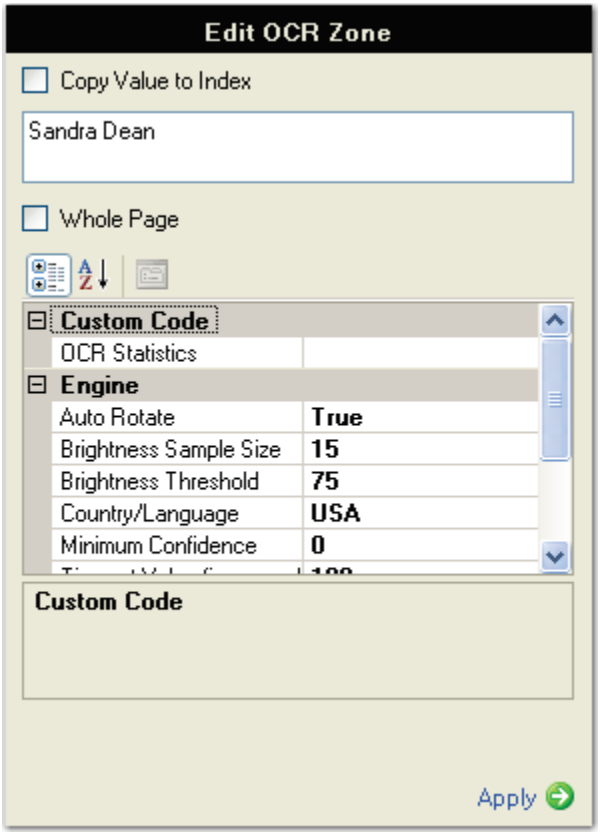


Manage Zones Window

2. Highlight the OCR zone. The **Edit Zone** toolbar appears.

Edit Zone Toolbar

3. Click **Edit Zone** icon. The **Edit OCR Zone** flyout window appears. The OCR value appears in the first field.



Edit OCR Zone

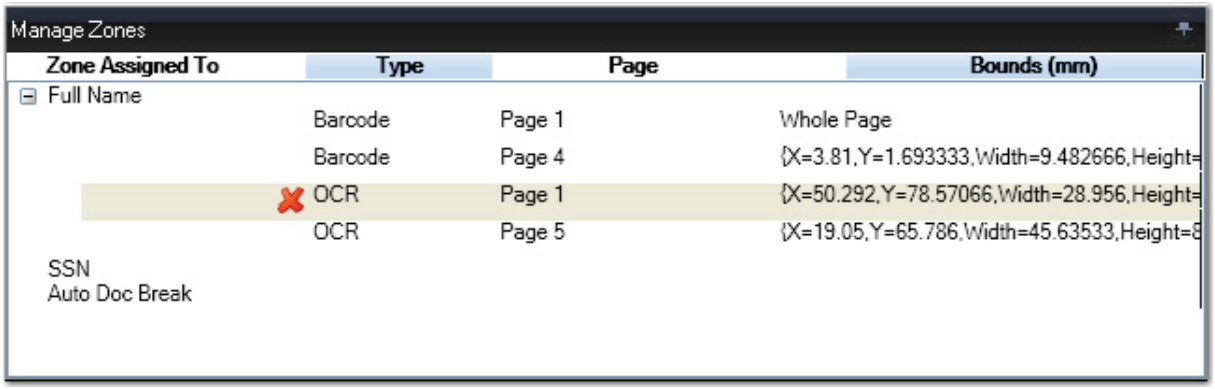
4. To copy the value to the selected index, select **Copy value to index**.
5. To apply the OCR zone to the entire page, select **Whole Page**.
6. Modify the properties, if applicable. See the section on **Zonal OCR Properties** for more information on each property.
7. Click **Apply** to save your changes.

Deleting OCR Zones

OCR Zones can be deleted with the Manage Zones operation in the Home page.

To delete an OCR zone:

1. In the Home page, click the **Manage Zones** icon. The **Manage Zones** window appears.



Manage Zones Window

2. Highlight the appropriate OCR zone in the grid. A toolbar appears above your selection.
3. Click the **Delete** icon.
4. Click **Yes** to confirm the deletion.

> **Note**:
>
> Alternatively, you can select the **Delete** icon next the OCR zone row in the grid.

Zonal OCR Properties

During zonal OCR configuration, custom code, engine-level, and zonal OCR properties are available for selection.

Auto Rotate

By default, this property is set to **True**, and the Open Text Zonal OCR engine will attempt to recognize text in all orientations (vertically and horizontally) within the zone. If you do not want the Open Text Zonal OCR engine to recognize text in all orientations (vertically only) within the zone, set this property to **False**.

Brightness Sample Size

This value (indicating both width and height) specifies the rectangle size used to calculate the brightness threshold. You can specify a value between 1 and 32, and the default value is 15.

> **Note**:
>
> Smaller brightness sample sizes may cause the OCR engine to recognize extraneous noise on the image.

Brightness Threshold

You can assign a brightness threshold value (between 0 and 255) for the image. The default value is 75.

Country/Language

When you select from the **Country/Language** property, your selection may reflect not only a country or language, but country groups (e.g., Western Europe), language groups (e.g., Latin), and character sets (e.g., OCR). Each country corresponds to one or more languages, and countries are automatically expanded into language sets (e.g., German corresponds to the German language; Switzerland corresponds to the German, French, Italian, and Rhaeto-Romantic languages).

Specific languages are also available for selection under the Country/Language property (e.g., English, German, Dutch, Italian, etc.). It is recommended to narrow your selection as much as possible since OCR recognition may become slower with a greater number of selected countries or languages. It is also recommended to select a country rather than a language or country group (e.g., Western Europe, South America, Scandinavia) since the recognition of certain types of addresses and money transfer forms may improve.

Language Groups

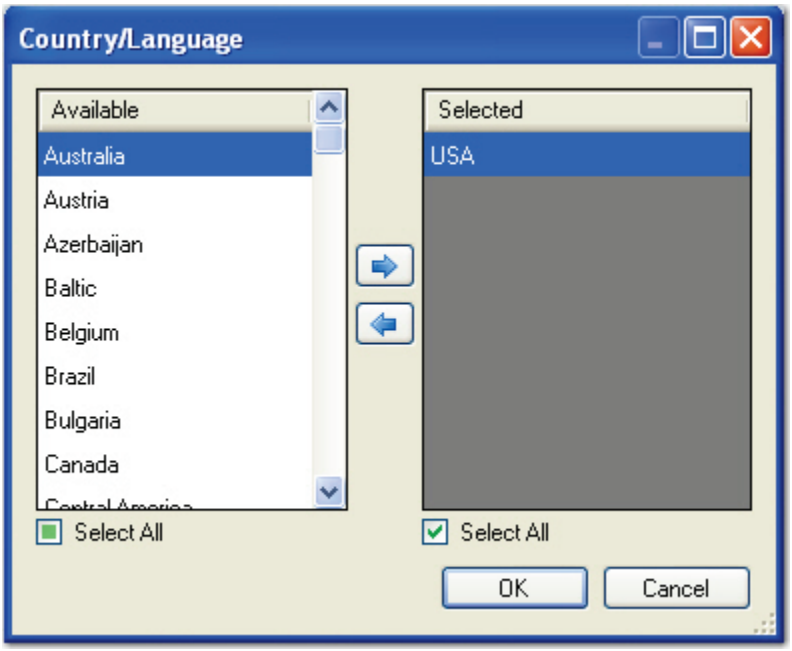
If you select a language group, it is recommended to select only one, since they encompass multiple languages, countries, and code pages.

1. Cyrillic (Codepage 1251)
2. Greek (Codepage 1253)
3. Latin (Codepage 1250, 1252, 1254, and 1257; i.e., Central Europe, Western Europe, Turkey, and Baltic)
4. Azerbaijanian

> **Note**:
>
> For language groups, recognition results are always represented by Unicode characters. The English character set (A-Z, a-z) is implicitly available with all country-language selections, even Greek or Cyrillic.

To select a country or language for full-text OCR output:

1. After selecting an output type, click the ellipsis button to the right of the **Country/Language** property. The **Country/Language** dialog box appears.



Country/Language

2. Highlight one or more selections from the **Available** list, and then click the right arrow.
3. To remove one or more selections from the **Selected** list, highlight the selections, and then click the left arrow.
4. When finished with your selections, click **OK**.

Supported Countries, Languages, Country Groups, Language Groups and Character Sets

If you narrow the search for specific languages or countries, the OCR engine will process more rapidly during OCR recognition. See **Appendix B** for a list of all supported countries, languages, country groups, and character sets.

Minimum Confidence

The confidence level reflects the reliability of the OCR recognition results. Values range from zero (the default setting), the lowest confidence level, to 255, the highest confidence level indicating the most reliable recognition results. Characters with lower confidence levels than your specified value will display as the rejection symbol, which is the tilde (~) character by default.

Timeout Value (sec)

This property allows you to define the maximum amount of time that the OCR engine processes a single image before it fails. By default, this property is set to 180 seconds (3 minutes). You can assign a timeout between one second and 3,600 seconds (1 hour).

> **Note**:
>
> Raising the timeout setting may increase the amount of time to process all images.

Reader Engine

Two internal OCR reader engines, RecoStar and AEGReader, are available for selection during zonal OCR configuration. Document content may cause one engine to generate more accurate recognition results, so the Voter option is selected by default. The Voter option automatically "votes" between both engines' recognition results, and generates results from the engine with the highest confidence level.

Rejection Symbol

This property represents rejected characters in output documents. A rejected character is not recognized by the active OCR recognition engine configuration. The default value is the Tilde character (~). Only a single character can be entered in this field.

> **Tip**:
>
> To prevent unrecognized characters from appearing in output documents, leave this field blank.

Syntax Mode

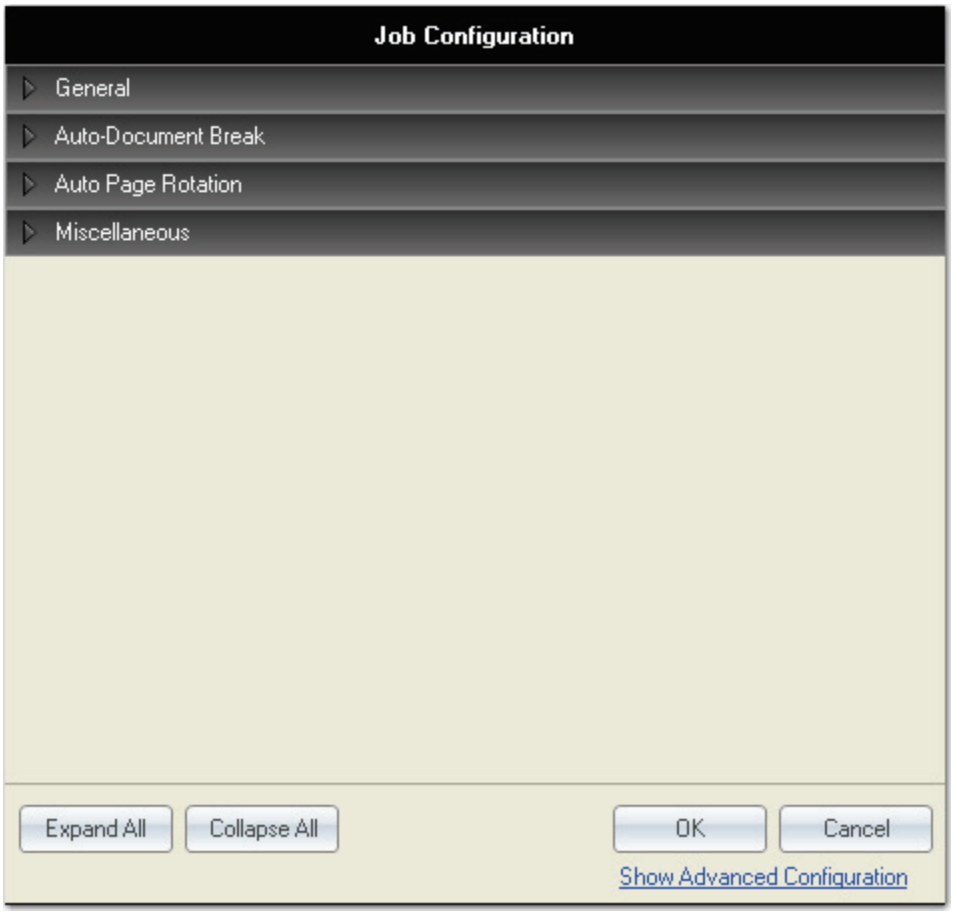
When you assign the syntax mode to alphanumerical, the default character set is alphanumeric. If a character is ambiguous, the OCR engine will attempt to process the character as a letter before a number. For example, the OCR engine will process a "G" before "6", "S" before "5", etc.

When you assign the syntax mode to numerical, the default character set is numeric. If a character is ambiguous, the OCR engine will attempt to process the character as a number before a letter. For example, the OCR engine will process a "6" before "G", "5" before "S", etc.

Job Configuration

In the Home page, job configuration settings contain properties for various tasks that are completed automatically in PaperVision Capture Desktop. During job configuration, you can assign automatic document break and page rotation settings. In addition, you can apply settings to limit the number of pages comprising each document, and you can automatically delete blank pages that are scanned in the application.

For example, document breaks can be inserted automatically based on page count, file size, barcode content, and OCR text. Additionally, you can configure custom code events that can automatically execute during batch processing. To view job settings, select the **Job**  icon in the **Settings** toolbar group.



Job Configuration (Basic Settings)

General

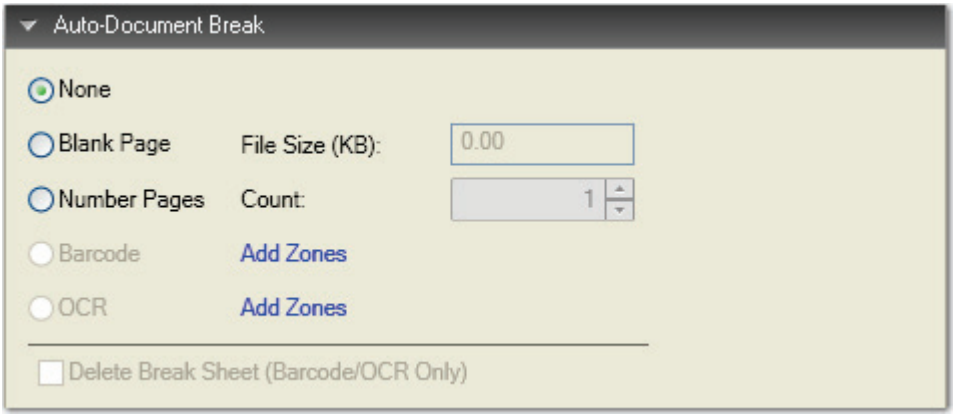
This read-only field displays the name of the job.

Auto Document Break

While scanning documents, you can determine where one document ends and the next document begins using the Auto Document Break properties. Although you can separate documents manually, you can select from options that are described below.

To assign the auto-document break settings:

1. In the Job Configuration window, expand the **Auto-Document Break** menu. The **Auto-Document Break** settings appear.



Auto Document Break

2. Select an option:
   - **None:** This is the default auto-document break type. When set to **None**, the system will expect you to manually separate new documents. No options are available for this setting.
   - **Blank Page**: To automatically insert document breaks based on the file size of the image, select **Blank Page**. Enter the size (in Kilobytes) of images to be considered blank. You can enter the file size whole numbers with up to two decimal places. Select **True** to leave the blank page in the batch, or select **False** to remove the blank page from the batch.

     Note:

     An error will appear if both the Auto Document Break and Minimum Page Size Detection properties are enabled.

   - **Number Pages:** To assign a fixed number of pages per document, enter the number of pages that PaperVision Capture Desktop will scan before starting a new document.

Barcode Zones

To select the **Barcode** mode, click the **Add Zone** link to define the barcode zone that will automatically break documents. See the **Barcode Zone Configuration** section for more information on specific settings.

OCR Zones

To select the **OCR** mode, click the **Add Zone** link to define the OCR zone that will automatically break documents. See the **OCR Zone Configuration** section for more information on specific settings.

Delete Break Sheet (Barcode/OCR Only)

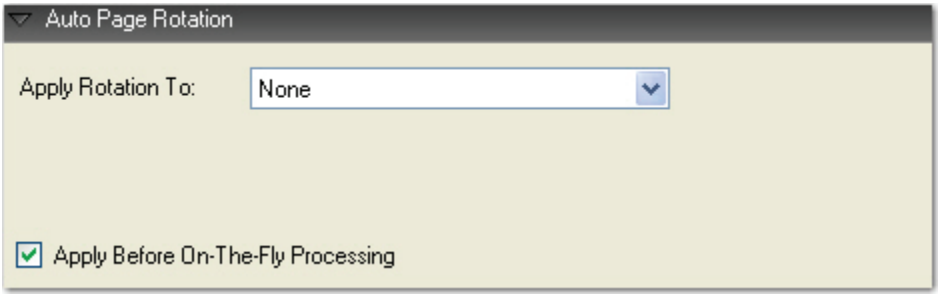
This property is applicable when you define Auto Document Breaks with barcode or OCR zones. When enabled, the break page will be deleted when all defined barcode or OCR zones are read successfully.

Auto Page Rotation

The Auto-Page Rotation setting allows you to configure how pages are rotated as images are scanned in the application.

To assign the page rotation settings:

1. In the Job Configuration window, expand the **Auto Page Rotation** menu. The **Auto Page Rotation** settings appear.



Auto Page Rotation

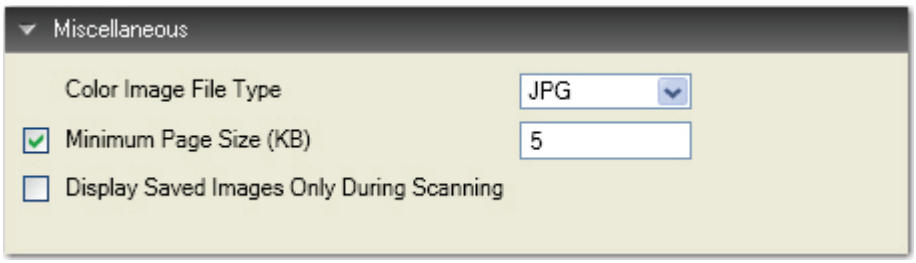
2. Select the page rotation setting from the **Apply Rotation To** drop down menu.
   - **None** disables the automatic page rotation feature.
   - **All Pages** automatically rotates all pages in a document by the specified rotation value as the documents are scanned.
   - **Even Pages** automatically rotates only the even numbered pages in a document by the specified rotation value as the documents are scanned.
   - **Odd Pages** automatically rotates only the odd numbered pages in a document by the specified rotation value as the documents are scanned.
   - **Even Pages/Odd Pages** automatically rotates the odd and even numbered pages in a document by the specified rotation values as the documents are scanned. Even pages and odd pages can be assigned different rotation values.
   - **First Page Only** automatically rotates the first page of a document by the specified rotation value as the documents are scanned.

- **All Pages Except First** automatically rotates all pages except the first page of a document by the specified rotation value as the documents are scanned.
- **First Page Only/All Page Except First** automatically rotates the first page of a document by the specified rotation value as the documents are scanned. The remaining pages can be assigned a different rotation value.

3. Select the rotation value from the **All Pages** (or the applicable selection will appear) drop-down list. You can select from 90°, 180°, or 270°.
4. If you are applying on the fly processing (for barcode, IP, or OCR zones) on images and you want the auto-rotation settings to be completed before you apply on the fly processing, select the option, **Apply Before On the Fly Processing**. Otherwise, images will be rotated after on the fly processing has been completed.

Miscellaneous

Miscellaneous job settings include Color Image File Type, Minimum Page Size, and Display Saved Images Only.



Miscellaneous Job Settings

Color Image File Type

You can specify the file type when storing scanned images that are not black and white. Click the **Color Image File Type** drop-down menu in the right column to make the selection. If you change this property after images have already been scanned into the batch, the file type will change for only those images subsequently scanned into the batch. For example, you change the Color Image File Type property from .bmp to .jpg after scanning ten out of twenty images in the batch. Images 1-10 will be .bmp file types; images 11-20 will be .jpg file types.

- BMP files are not compressed and can be large. These files contain pixels and can degrade when you increase resolution.
- JPG images are compressed, so they contain less data and smaller file sizes than other image types.

Minimum Page Size

Blank pages can be scanned accidentally or as the blank side of a duplex page. The Minimum Page Size setting allows you to delete blank pages as they are scanned. In the **Minimum Page Size** field, enter the minimum page size detection (in Kilobytes) to be deleted. You can enter the size in whole numbers with up to two decimal places.

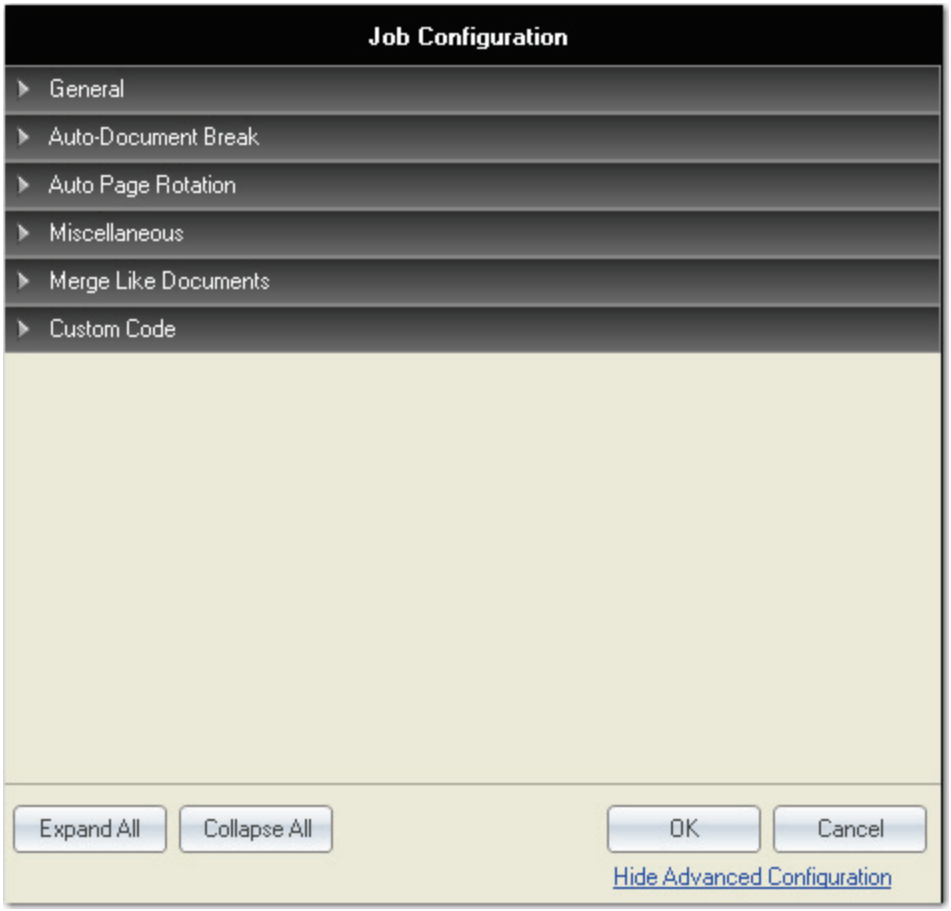
> **Note:**
>
> An error will appear if both the Auto Document Break and Minimum Page Size Detection properties are enabled.

Display Saved Images Only

When this property is enabled, PaperVision Capture Desktop only displays the images that are saved (in the manner that they are being saved). For example, if images are rotated as they are scanned, only the correct rotation orientation will display. If you enable this setting, and you have specified a minimum page size detection, blank pages will not display. If you do not enable this setting, all images will display, including blank images.

Advanced Job Configuration

Advanced job settings include properties to merge documents containing matching index values (Merge Like Documents operation). Custom code properties allow you to configure automated actions triggered after certain events (Add Page, Save Indexes, Match and Merge, and Custom Code Execution) are completed.



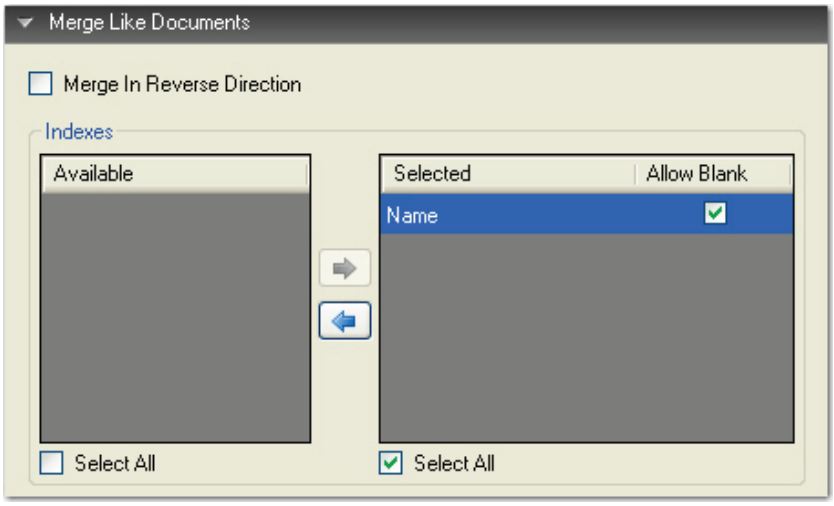
Advanced Job Configuration

Merge Like Documents

The Merge Like Documents job setting merges pages from multiple documents with the same index values into a single document. Documents that have not been indexed are not included in the merge process. The Merge Like Documents operation is performed on all documents in the batch.

To configure the Merge Like Documents setting:

1. Click **Show Advanced Configuration** on the bottom right of the **Job Configuration** flyout window.
2. Expand the **Merge Like Documents** menu. The **Merge Like Documents** settings appear.



Merge Like Documents

3. You can determine the page order of the merged document. Select **Merge in Reverse Direction** to place the last page at the beginning of the resulting document. If all pages should appear in the order in which they are merged, do not select this option.
4. All index values defined for the job appear in the **Available** list. Highlight the index values to be included in the Merge Like Documents operation, and click the right arrow. Your selected index values will appear in the **Selected** list.
    - Or, choose **Select All**, and then click the right arrow.
    - To remove a selected index value, highlight the index value in the **Selected** list, and then click the left arrow.
    - Or, choose **Select All** to remove all index values from the **Selected** list, and then click the left arrow.
5. By default, blank index values are not included in the merged document. If blank index values should be included in the merged document, select the **Allow Blank** check box for the appropriate index value. For example, if you select the **Allow Blank** check box for the **Invoice Number** index value, all documents must contain blank index values in order to be merged into one document. If at least one Invoice Number index value is defined and the remaining index values are blank (or vice versa), the documents will not be merged.

Custom Code Event Configuration

Custom code is automatically executed when certain events occur in PaperVision Capture Desktop. These events include the following:

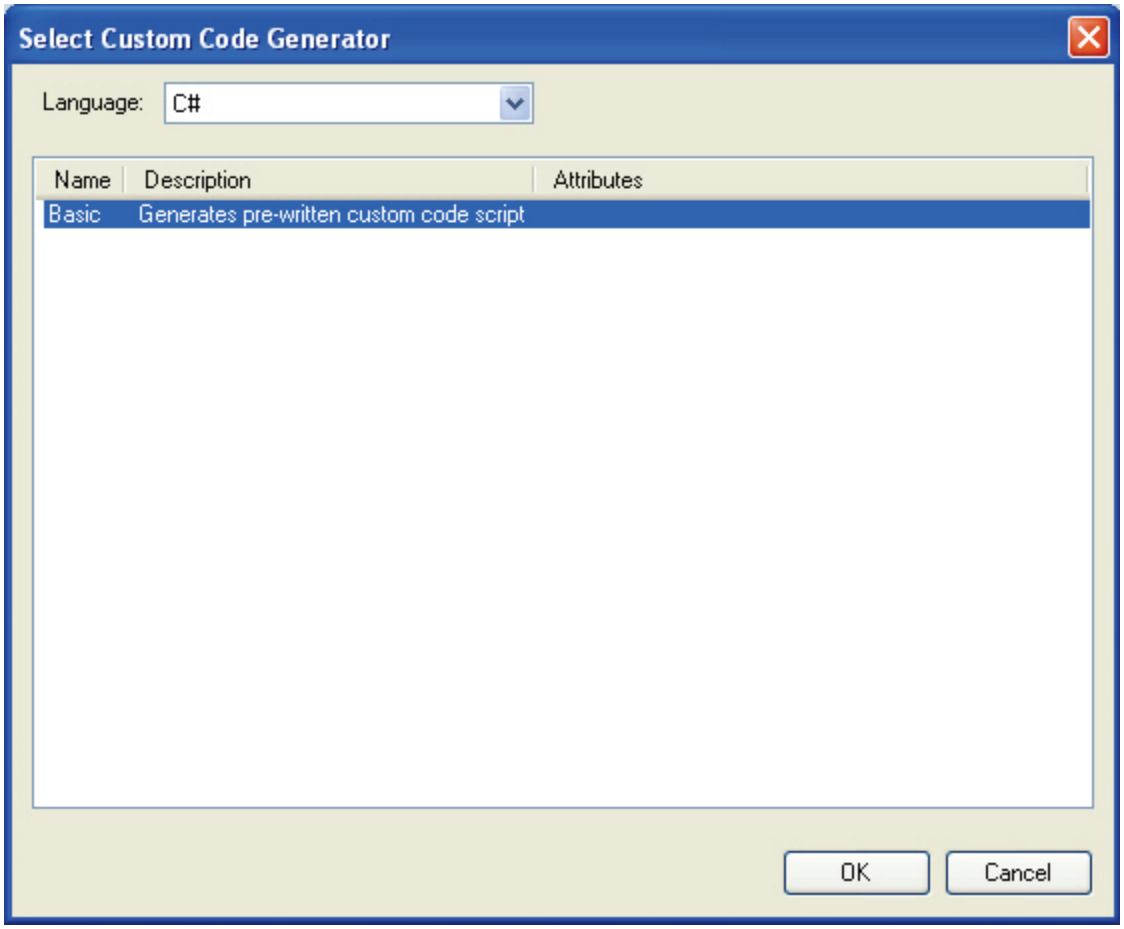
- Index Populating - when a specific index is populated with data
- Index Validating - when a specific index is to be validated
- Add Page - each time a page is added to the batch
- Barcode Detected - each time a barcode is successfully read
- Save Indexes - when index values are being saved to the batch
- Custom Code Execution - when a user clicks the "Custom Code" button
- Match and Merge - when a user clicks the "Match and Merge" button

To configure the code for the Index Populating and Index Validating events:

1. In the **Indexing** window, select the index.
2. Click the **Configure Index** icon to open the properties window.
3. Click **Show Advanced Configuration** to view advanced indexing properties.
4. Expand the **Custom Code** menu to view the Index Populating and Index Validating events.

Index Populating and Index Validating Events

5. Click the **Add** icon next to the event, and the **Select Custom Code Generator** dialog box opens.



Select Custom Code Generator

6. Select the **Basic** generator, and click **OK**. The **Script Editor** opens, where you can edit the script.
7. Click **OK** in the Script Editor to save the script.

The remaining events are configured in the Job Configuration window, located in the **Home page > Job (Settings toolbar group) > Show Advanced Configuration > Custom Code** menu. After you navigate to the appropriate screen, follow these steps to generate custom code:

To configure the code for the remaining events:

1. In the Job Configuration flyout window, click **Show Advanced Configuration** on the bottom right.
2. Expand the **Custom Code** menu. The **Custom Code** events appear.

Custom Code

Add Page

Add Page executes custom code just before images are appended to the batch, including rotation or barcode indexing. When the script is enabled for this option, it will be executed for all images scanned in the application or when you import a batch. This script is not executed if you execute the Import Images operation.

Barcode Detected

The Barcode Detected event executes custom code after a barcode's value, location, size, orientation, and type have been successfully read during scanning. When a script is enabled for this option, it will be executed every time a barcode is successfully read during scanning (multiple barcodes can be read per page). This event can also be used to apply a page-level custom tag. The script is not executed if a barcode cannot be successfully read.

Save Indexes

Save Indexes executes prior to the operator saving the index values in the Indexing window.

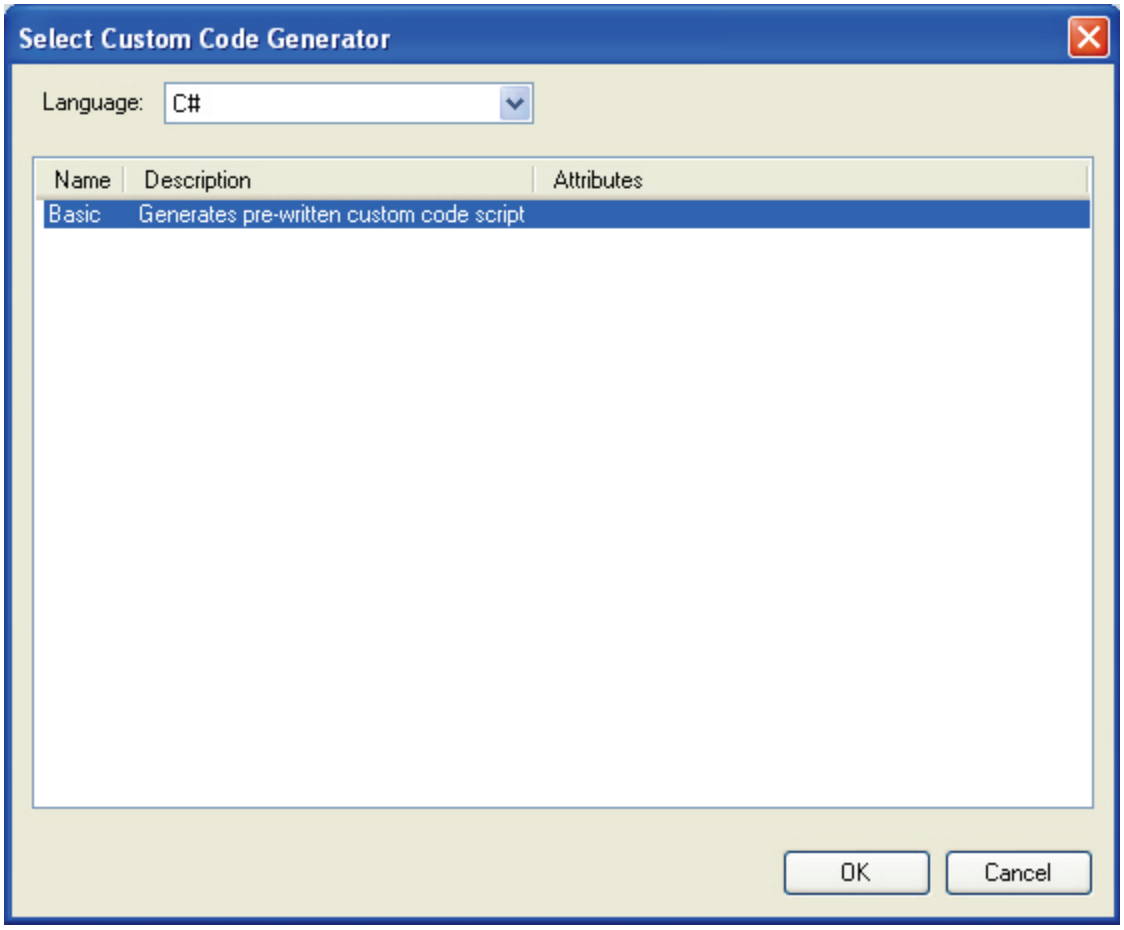
Custom Code Execution

Custom Code Execution executes when the user clicks the Custom Code button in the Edit page.

Match and Merge

Match and Merge executes when you click the **Match and Merge** button in the Edit page. See the Match and Merge Wizard section for specific configuration information.

3. To configure one of the custom code events, click the **Add**  icon. For the Add Page, Save Indexes, or Custom Code Execution events, the **Select Custom Code Generator** wizard opens, where you can configure a basic script generator.



Custom Code Generator Wizard

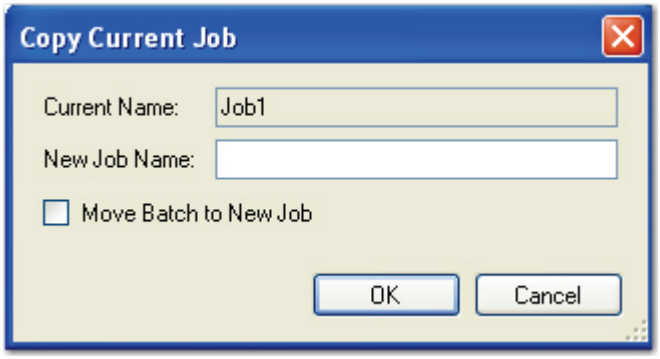
4. Select either the **C#** or **Visual Basic** programming language from the drop-down list.
5. Select the **Basic** generator, and click **OK**. The **Script Editor** opens, where you can edit the script.
6. Click **OK** in the Script Editor to save the script.

Copying a Job

The Copy Job operation makes a copy of the current job and allows you to move the open batch to the new job.

To copy a job:

1. In the Home page, click the **Copy Job**  icon in the **Settings** toolbar group. The **Copy Job** dialog box appears.



Copy Job

2. Enter the new job name.
3. To move the current batch to the new job, select the **Move Batch to New Job** check box.
4. Click **OK**.

Managing Zones

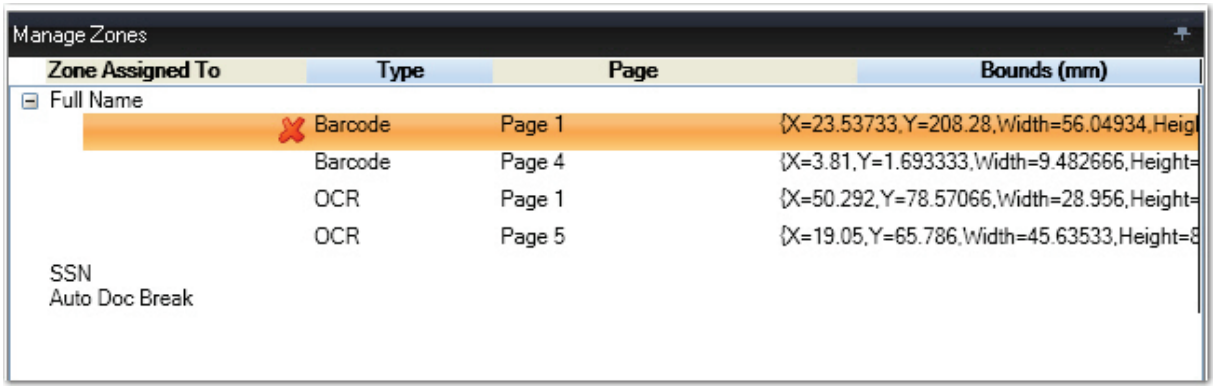
When you select the **Manage Zones** operation from the Home page, you can view, modify, and delete zones that you have assigned to indexes. The Manage Zones window displays all zoom, barcode, and OCR zones per index. You can also view the zone's page number, type (Barcode, OCR, and Zoom Zone), and parameters (X and Y coordinates; width and height).

- Index (indicates the index to which the zone is assigned)
- Type (indicates the type of zone, including barcode, OCR, or zoom zone)
- Page

For specific OCR and barcode zone configuration settings, see the **Barcode Zone Configuration** section or the **OCR Zone Configuration** section.

To edit existing zones:

1. In the Home page, click the **Manage Zones** icon. The **Manage Zones** grid opens, where you can view all zones that have been applied to indexes.

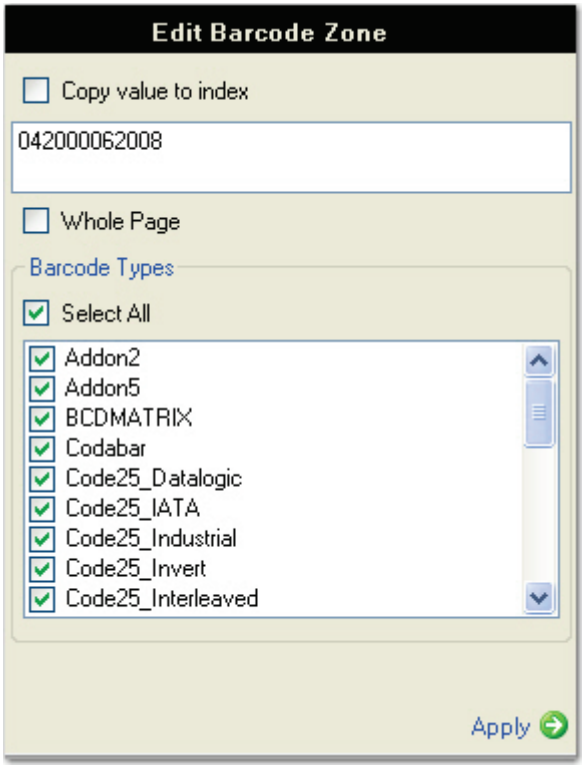


Manage Zones

2. To view a zone's properties, highlight the zone in the grid. The **Manage Zones** toolbar appears above the region on the image, where you can save, delete, or edit the zone's properties.

Manage Zones Toolbar

3. To edit the zone's properties, click the **Edit Zone** icon. The corresponding flyout window appears with a preview of the barcode/OCR value.



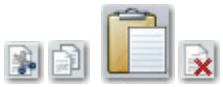
Edit Barcode Zone Flyout Window

4. To copy the zone's value to the selected index, select **Copy value to index.**
5. To apply the zone to the entire page, select the **Whole Page** option.
6. Edit the properties for the zone, and then click **Apply**.
7. If you adjusted the borders of the zone and want to save the new dimensions, click the **Save** icon.
8. To delete an existing zone, highlight the zone, and then select the **Delete** icon.

Right-Click Operations

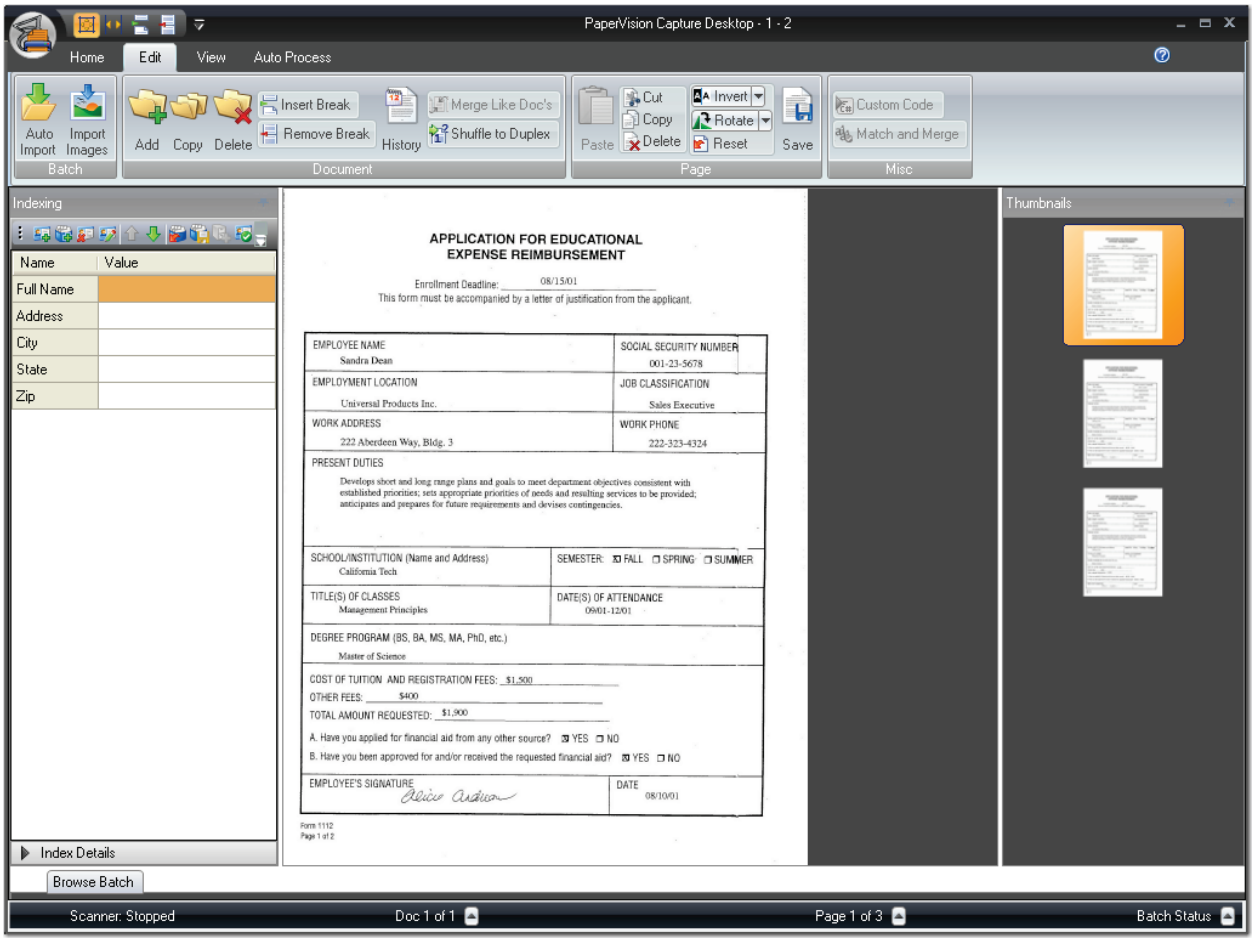
You can execute a variety of operations when you right-click on an image in the main viewing window in the Home, Edit, View, and Automated Processing Tasks pages. The following table displays each operation's default hot keys (if applicable), toolbar icon, and main page where the operation can be executed.

Right-Click Operations	
Reset Page 	Resets the page to its original view Page: Edit Hot Key: Ctrl+T
Rotate Image 	Rotates the image 90 degrees clockwise Page: Edit Hot Key: R
Rotate and Save Image 	Rotates and saves one or multiple images in one document or across multiple documents Page: Edit
Scale to Height, Width, Window 	Fits images to the height, width, or both height and width of the screen. Page: View
Copy Document 	Copies all pages and appends the new document after the selected document Page: Edit Hot Key: Y
Delete Document 	Deletes the current document and its associated images Page: Edit Hot Key: Ctrl+Delete
Insert Document Break 	Makes your selected page the first page of the new document Page: Edit Hot Key: Ctrl+Insert

Right-Click Operations (continued)	
Remove Document Break 	Removes a previously-inserted document break Page: Edit Hot Key: Ctrl+R
Cut, Copy, Paste, Delete Pages 	Cuts, copies, pastes, or deletes the current page Page: Edit Hot Keys: • Cut: Ctrl+X • Copy: Ctrl+C • Paste: Ctrl+V • Delete Page: Delete
Invert Page Polarity 	Reverses the black text on white background to white text on black background (or vice versa) Page: Edit
Invert and Save Page Polarity 	Page: Edit

Chapter 3 - Edit Page

The Edit page contains bath, document, and page manipulation operations such as the Cut, Copy, and Paste Pages/Documents. The Auto Import operation allows you to manually import directories containing images. Auto Import creates a new document at the end of the current batch and then inserts the images. In addition, you can insert and remove document breaks from this page; invert and rotate pages; merge documents containing matching index values; and, shuffle documents to duplex. In addition, you can save images after you rotate or invert them. You can also index documents (and all associated indexing operations) within the Edit page.



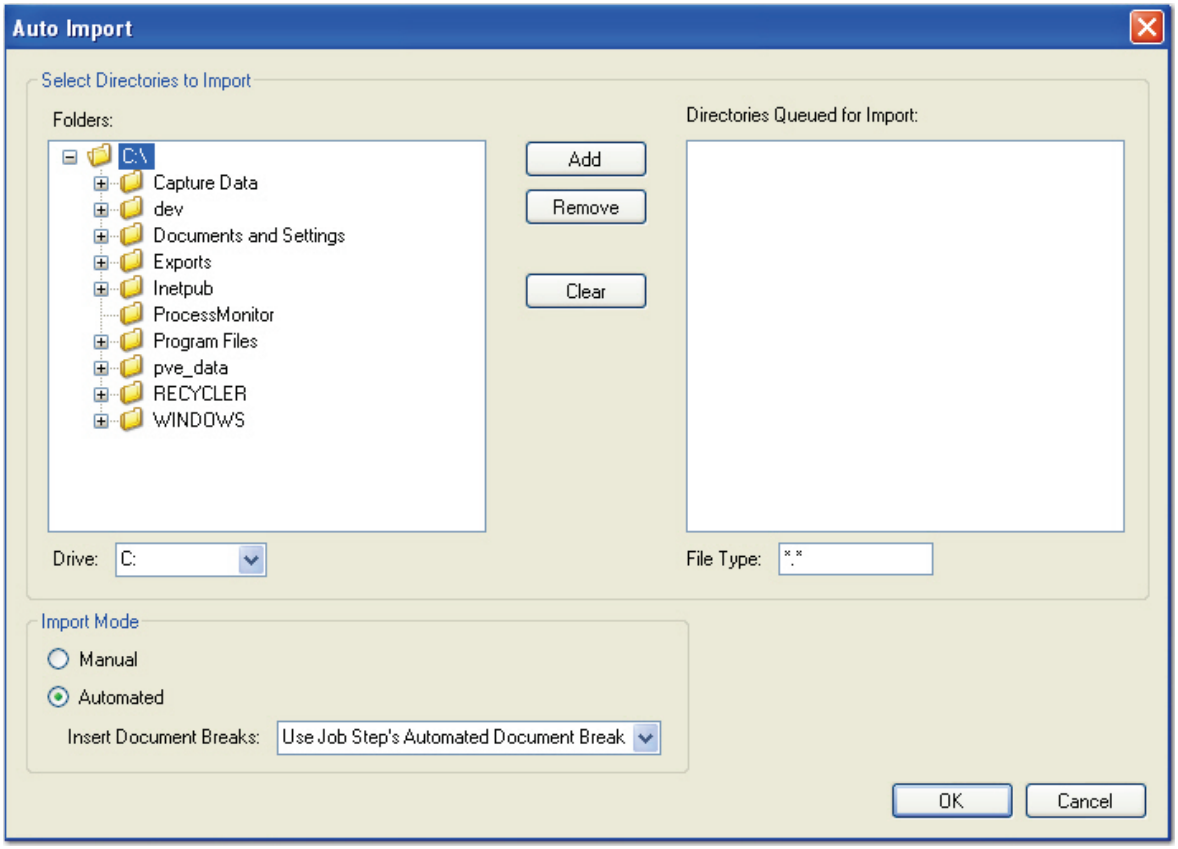
Edit Page

Auto Import

The Auto Import operation allows you to manually import directories containing images. This operation creates a new document at the end of the current batch and then inserts the images.

To execute the Auto Import operation:

1. Click the **Auto Import** icon. The **Auto Import** dialog box opens.



Auto Import

2. In the **Select Directories to Import** section, highlight the directory and click **Add.**
3. Repeat the second step for each directory.

> **Note:**
>
> A message appears if you attempt to import a directory that has already been added to the **Directories Queued for Import** list.

4. If applicable, select the drive from which to import the batch. By default, the contents of your local hard drive appear in the **Folders** list.

5. To remove one or more directories, highlight them in the **Directories Queued for Import** list, and then click **Remove.**

> **Note:**
>
> Use the **Ctrl** key when you select multiple directories or the **Shift** key when you select multiple, consecutive directories.

6. To remove all selected directories, click **Clear.**
7. To specify an image file type to import, enter it in the **File Type** field.
8. To enter multiple image file types, use the filter format with semicolons placed between file types:

   *.bmp; *.jpg; *.tif

9. In the **Import Mode** section, select **Manual** or **Automated.** If you select **Manual**, proceed to the next step. If you select **Automated**, proceed to step 11 on the next page.
10. If you select the **Manual** option, the **Manual Auto Import** screen opens where you can save or skip certain images and insert document breaks. Additionally, you can rotate the pages, invert page polarity, and reset the page view. The progress bar at the bottom of the screen helps you navigate through the images by displaying batch, image, document, and page information.

Manual Auto Import Operations

- To save the image, click the **Save Image** icon.
- To create a new document, click the **New Document** icon.
- To skip the image, open, click the **Skip Image** icon.
- To invert the page polarity, open **View > Invert,** click the **Invert** icon.
- To rotate the page view by 90 degrees clockwise, click the **Rotate** icon.
- To reset the image to its original view, click the **Reset** icon.
- To exit out of the screen, click the **Exit** icon.

The following hot keys can be used in the Manual Auto Import screen:

- Save Image (Ctrl+S)
- Create New Document and Save Image (Ctrl+Insert)
- Skip Image (Ctrl+I)
- Invert Page Polarity (Ctrl+N)
- Rotate Image (R)
- Reset Image (Spacebar or Ctrl+T)
- Exit (Ctrl+E)
- Help (F1)

11. If you select **Automated**, select from one of the following **Insert Document Break** options:
    - **Use Job Step's Auto-Document Break Setting** is assigned by your administrator. One of the three settings may have been assigned:
        a. **None:** This is the default auto-document break type for a newly created step. When set to **None**, the system expects you to manually separate new documents.
        b. **Number of Pages Per Document:** This is the total number of pages that PaperVision Capture will scan before starting a new document. Your administrator may have the system display a message that asks you for a fixed number of pages before breaking to a new document.
        c. **Barcode**: Your administrator pre-defined barcode zones to be read during the scanning process.
    - **Each Directory** automatically imports images into multiple documents, placing a document break between each directory. For example, if you select a parent directory that contains subdirectories, document breaks are automatically inserted between the subdirectories.
    - **Each Image** automatically imports multi-page image files and separates them into multiple documents, placing a document break between each image file.
12. Click **OK** to begin the import process.
13. If invalid images are found during the import process, a notification appears with each invalid image and prompts you whether to import the remaining images.
    - Click **Yes** to import the remaining images in the batch.
    - Click **No** to abort the import process from that image forward. The remaining images (located after the invalid image) will not be imported.

Importing Images

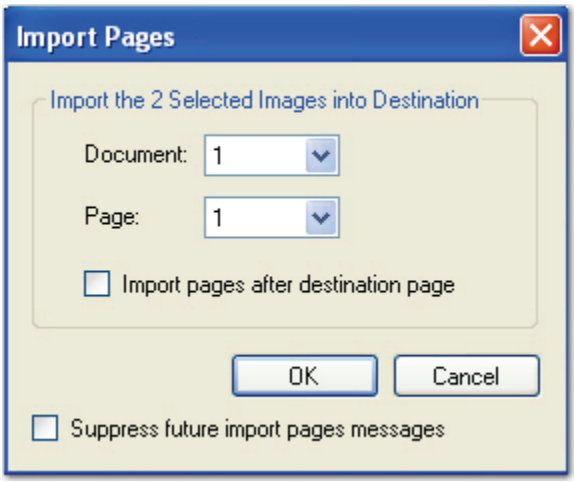
This operation imports images before or after your selected page in the current document. This command does not perform barcode recognition, apply page rotation, or insert auto document breaks.

To import images:

1. Navigate to the destination page in the document.
2. Click the **Import Images**  icon.
3. In the **Select Images to Import** dialog box, browse to the directory and select the image(s) to import.

> **Note**:
>
> Press the **Ctrl** key to select multiple images or the **Shift** key to select multiple, consecutive images.

4. Click **Open**, and the **Import Pages** dialog box appears.



Import Pages

5. Select the document number where the images will be inserted.
6. Select the page number where the images will be inserted.
7. If you want the images inserted after the destination page, select **Import pages after destination page**. Otherwise, images are inserted before the destination page.
8. Click **OK**.

Adding a Document

The Add Document command appends a blank document to the end of the batch. You can subsequently scan and import images into the new document and enter index values.

To add a document:

1. Open the appropriate document.
2. Click the **Add Document**  icon. The blank document appears at the end of the batch. You can then import or scan images into the new document (and enter index values if required).

Copying a Document

The Copy Document command copies all pages and appends the new document after the selected document.

To copy a document:

1. Open the appropriate document.
2. Click the **Copy**  icon.
3. Enter the number of copies to create, and then click **OK**.

Deleting Documents

The Delete Document command deletes the current document and its associated images.

To delete a document:

1. Open the appropriate document.
2. Click the **Delete**  icon.
3. To proceed with the deletion, click **Yes**.

If the Browse Batch window is viewable, you can delete one or more documents from the list.

To delete one or more documents from the Browse Batch window:

1. Select one or more documents in the list.
2. Right-click and select **Delete Documents**.
3. To proceed with the deletion, click **Yes**.

Inserting an Auto Document Break

While scanning documents, you can determine where one document ends and the next document begins by inserting an Auto Document Break. This operation makes your selected page the first page of the new document.

To insert an auto document break:

1. Navigate to the page where the document break will be placed.
2. Click the **Insert Document Break** icon.
3. If you want to save the index values, click **Yes**.
   - Selecting **No** will not save the index values, but will insert a document break.
   - Selecting **Cancel** will abort both operations.
4. If you want the marked page to be the first page of the new document, click **Yes**.

Removing a Document Break

If you inserted a document break and no longer want to keep it, you can remove the break and place the pages from the current document at the end of the previous document.

To remove a document break:

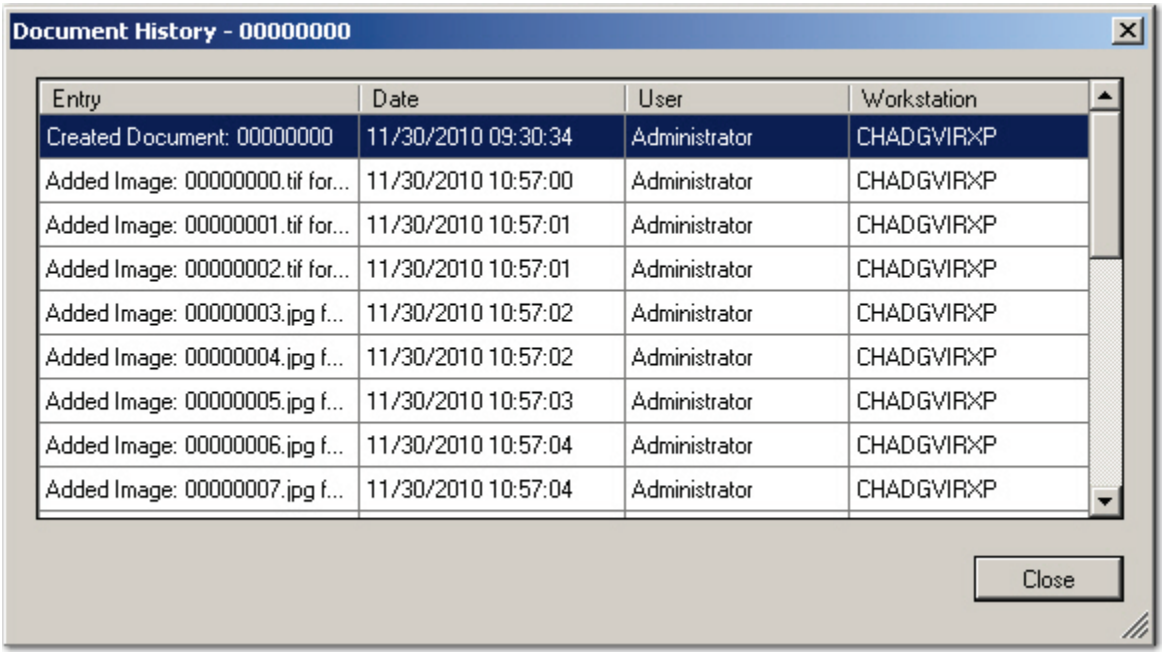
1. Navigate to the page where the document break was placed.
2. Click the **Remove Break** icon.

   Note:
   You cannot remove a document break for the first document in the batch.

3. Click **Yes** to remove the document break, and the current document will be merged with the previous document.
4. Or, select **No**, and the document break will not be removed.

Document History

The Document History operation lists all actions performed on the document, such as its creation date and name, image additions and deletions, and index additions and updates. To view a document's history, click the **Document History** icon.



Document History

Merging Like Documents

This operation merges selected documents containing the same index values. The Merge Like Documents job setting must be configured for this operation. See the **Job Configuration** section in Chapter 2 for more information.

To merge documents containing the same index values:

1. Enter the required index value(s) for all documents.
2. In the Edit page, click the **Merge Like Documents** icon. If the Browse Batch window is viewable, documents with matching index values will be combined into one document.

> **Note:**
>
> Documents must contain matching index values in order to be merged. For example, in a batch containing three documents, one document contains a blank index value; the other two documents have a matching index value of 80111. Only the two documents with 80111 index values will be merged.

3. If you are merging a large volume of documents, a progress message will appear during document analysis, prior to the document merge. During analysis, you can click the **Cancel** button to cancel the merge process and return to the job step screen.

> **Note:**
>
> Depending on the number of documents that comprise the batch, document analysis may take a few minutes to complete.

4. Click **OK** to the confirmation message.

Shuffling a Document to Duplex

This operation shuffles the first half of the pages, and they become the odd pages. The remaining half of the pages becomes the even pages, and these pages are placed in reverse order. The document must contain an even number of pages to shuffle.

> **WARNING!**
>
> This command is irreversible.

To shuffle a document to duplex:

1. In the Edit page, click the **Shuffle Document to Duplex** icon.
2. Click **Yes** to proceed with the shuffle.

Cutting and Pasting Pages

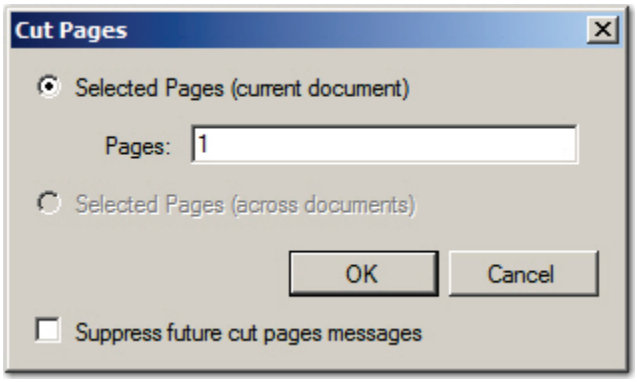
You can cut, copy, and paste one or multiple pages in the Single Display View  or Multiple Display View  within the Edit page. The cut, copy, and paste operations can be performed on consecutive or non-consecutive pages in one document or across multiple documents.

Tip:

You can also right-click on a page or within the multiple display view or the Thumbnails window to cut, copy, and paste pages.

To cut and paste pages:

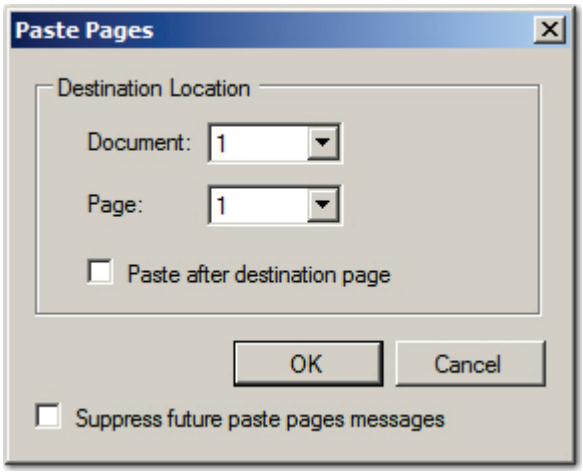
1. Navigate to the page in the document (or, select one or more pages in multiple display or Thumbnails).
2. Click the **Cut**  icon. The **Cut Pages** dialog box appears.



Cut Pages

3. If you selected pages within the same document, the first option, **Selected pages (current document)**, is selected. For this option, the selected pages are listed in the **Pages** field. If you selected pages across multiple documents, the second option, **Selected pages (across documents)**, would be selected. Ensure the appropriate option is selected, and then click **OK**.
4. In the Thumbnails window or multiple display view, selected pages appear with the **Cut**  icon in the upper right. Navigate to the page in the document where the pages will be pasted.

5. Click the **Paste**  icon. The **Paste Pages** dialog box appears.



Paste Pages

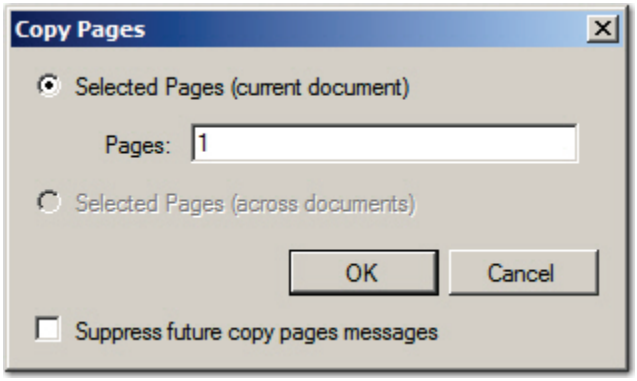
6. Your selected document and page number appear in the **Paste Pages** dialog box. To change your selection, enter the appropriate document and page number.
7. By default, the pages are pasted directly before your selected location. If you want them pasted after your destination page, select the option, **Paste after destination page**.

> **Tip**:
>
> To prevent the confirmation message from appearing each time you perform this operation, select the **Suppress future cut/paste pages messages** option. To show the message once again, select the **PaperVision Capture Desktop Button**  > **Preferences** > **Confirmation Messages**, and select the applicable check box.

Copying and Pasting Pages

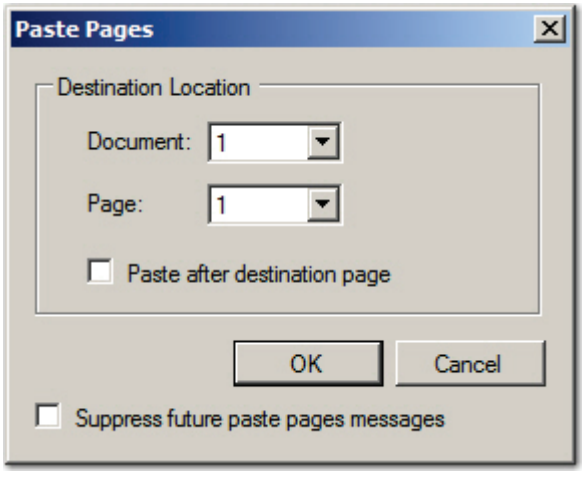
To copy and paste pages:

1. Navigate to the page in the document (or, select one or more pages in multiple display or Thumbnails).
2. Click the **Copy**  button. The **Copy Pages** dialog box appears.



Copy Pages

3. If you selected pages within the same document, the first option, **Selected pages (current document)**, is selected. For this option, the selected pages are listed in the **Pages** field. If you selected pages across multiple documents, the second option, **Selected pages (across documents)**, would be selected. Ensure the appropriate option is selected, and then click **OK**.
4. In the **Thumbnails** window or multiple display view, selected pages appear with the **Copy** icon in the upper right. Navigate to the page in the document where the pages will be pasted.
5. Click the **Paste**  icon. The **Paste Pages** dialog box appears.



Paste Pages

6. Your selected document and page number appear in the **Paste Pages** dialog box. To change your selection, enter the appropriate document and page number.
7. By default, the pages will be pasted directly before your selected location. If you want them pasted after your destination page, select the option, **Paste after destination page**.

> **Tip**:
>
> To prevent the confirmation message from appearing each time you perform this operation, select the **Suppress future copy/paste pages messages** option. To show the message once again, select the **PaperVision Capture Desktop Button** > **Preferences** > **Confirmation Messages**, and select the applicable check box.

Deleting a Single Page

To delete a page:

1. In the **Thumbnails** section, select the page to delete.
2. Click the **Delete** icon in the toolbar.
3. Click **Yes** to the confirmation message.

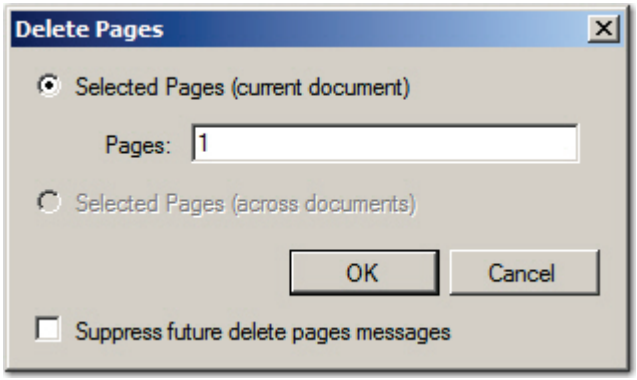
Deleting Pages

This operation deletes your specified page(s) in one document or across multiple documents. You can perform this operation in single display view (View > Organization > Single Display) or multiple display view (**View > Organization > Multiple Display)** as well as the Thumbnails window (**View > Windows > Thumbnails**).

> **Tip**:
>
> You can also right-click on a page or within the multiple display view or Thumbnails window to delete pages.

To delete pages:

1. Select one or more pages in the same document or across multiple documents.
2. Click the **Delete** icon. The **Delete Pages** dialog box appears.



Delete Pages

3. If you selected pages within the same document, the first option, **Selected pages (current document)**, is selected. For this option, the selected pages are listed in the **Pages** field. If you selected pages across multiple documents, the second option, **Selected pages (across documents)**, would be selected. Ensure the appropriate option is selected, and then click **OK**.
4. To confirm the deletion, click **OK**.

> **Tip**:
>
> To prevent the confirmation message from appearing each time you perform this operation, select the **Suppress future delete pages messages** option. To show the message once again, select the **PaperVision Capture Desktop Button** > **Preferences** > **Confirmation Messages**, and select the applicable check box.

Inverting Page Polarity

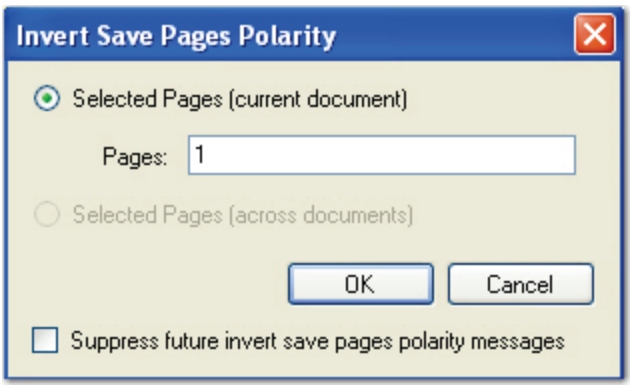
Inverting the page polarity reverses the black text on white background to white text on black background (or vice versa). When you are in single-display view, you can click the **Invert** icon to run this operation.

Inverting and Saving Page Polarity

Inverting the page polarity reverses the black text on white background to white text on black background (or vice versa). This operation saves the inverted pages prior to batch submittal.

To invert and save page polarity on one or multiple pages:

1. Navigate to the appropriate page.
2. Click the **Invert** drop-down list, and then select the **Invert Save Pages Polarity** option. The **Invert Save Pages Polarity** dialog box appears.



Invert Save Pages Polarity

3. If you selected pages within the same document, the first option, **Selected pages (current document)**, will be selected. For this option, the selected pages are listed in the **Pages** field. If you selected pages across multiple documents, the second option, **Selected pages (across documents)**, is selected. Ensure the appropriate option is selected.
4. To specify a certain page, page range, or combination thereof, enter the page numbers in the **Pages** field. For example, you can enter "2, 4-7".
5. Click **OK**.

> **Tip**:
>
> To prevent the confirmation message from appearing each time you perform this operation, select the **Suppress future invert messages** option. To show the message once again, select the **PaperVision Capture Desktop Button** > **Preferences** > **Confirmation Messages**, and select the applicable check box.

Rotating an Image 90° Clockwise

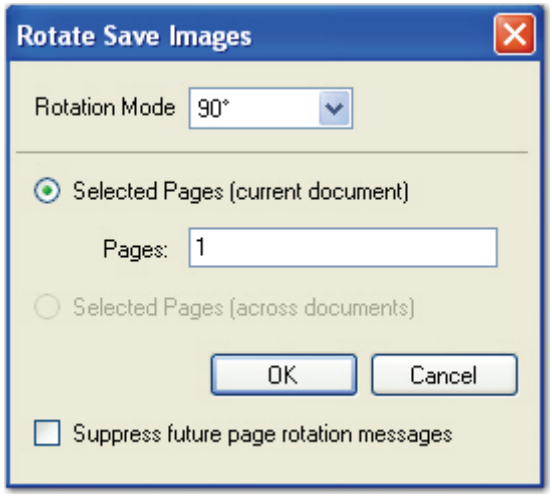
To rotate the image 90 degrees clockwise, click the **Rotate** icon.

Rotating and Saving Pages

You can rotate one or multiple images in one document or across multiple documents and save them. You can perform this operation in the Thumbnails window, single display view (**View page > Single Display**) or multiple display view (**View Page > Multiple Display**).

To rotate and save one or multiple images:

1. Select one or more pages.
2. Click the **Rotate Save Images** icon. The **Rotate Save Images** dialog box appears.



Rotate Save Images

3. From the **Rotation Mode** drop-down list, select the angle (90, 180, 270).
4. If you selected pages within the same document, the first option, **Selected pages (current document)**, will be selected. For this option, the selected pages are listed in the **Pages** field. If you selected pages across multiple documents, the second option, **Selected pages (across documents)**, is selected. Ensure the appropriate option is selected.
5. To specify a certain page, page range, or combination thereof, enter the page numbers in the **Pages** field. For example, you can enter "2, 4-7".

6. Click **OK**.

> **Tip:**
>
> To prevent the confirmation message from appearing each time you perform this operation, select the **Suppress future page rotation messages** option. To show the message once again, select the **PaperVision Capture Desktop Button** > **Preferences** > **Confirmation Messages**, and select the applicable check box.

Resetting Pages

If you have rotated or inverted a page, you can reset the page to its original view. To reset the image to its original view, click the **Reset** icon.

Saving Pages

If you rotate a page or invert its polarity, this command allows you to save its new appearance.

To re-save the page:

1. After you change the page rotation or polarity, click the **Save** icon.
2. Click **OK** to the confirmation prompt.

Miscellaneous Operations

Miscellaneous operations include custom code and match and merge operations that you can execute in the Edit page.

Custom Code

This operation is available if you have configured a custom operation in the Home page's **Job Settings**. For example, an operation can validate your indexing entries against an external database, or your indexing entries can update an external database. Click the **Custom Code** icon to run a custom operation that has been configured.

Match and Merge

If the Match and Merge command is executed, your indexing entries will be combined with entries found in the specified database (configured in the **Home page > Advanced Job Settings > Custom Code** menu). For example, you enter one required index value, a social security number, for an insurance invoice. You execute the Match and Merge command, and the remaining two index fields, First Name and Last Name, are automatically filled in.

If you are unsure of the exact index value while hand-key indexing, you can insert wildcard characters to perform a partial search against the database. For example, you can insert the percent sign (%) to specify any number of unknown characters to search for in a SQL, Sybase, or Oracle database; or, you can insert the asterisk (*) to specify any number of unknown characters to search for within a Microsoft Access database. For information on specific wildcard characters, refer to the documentation for the respective database.

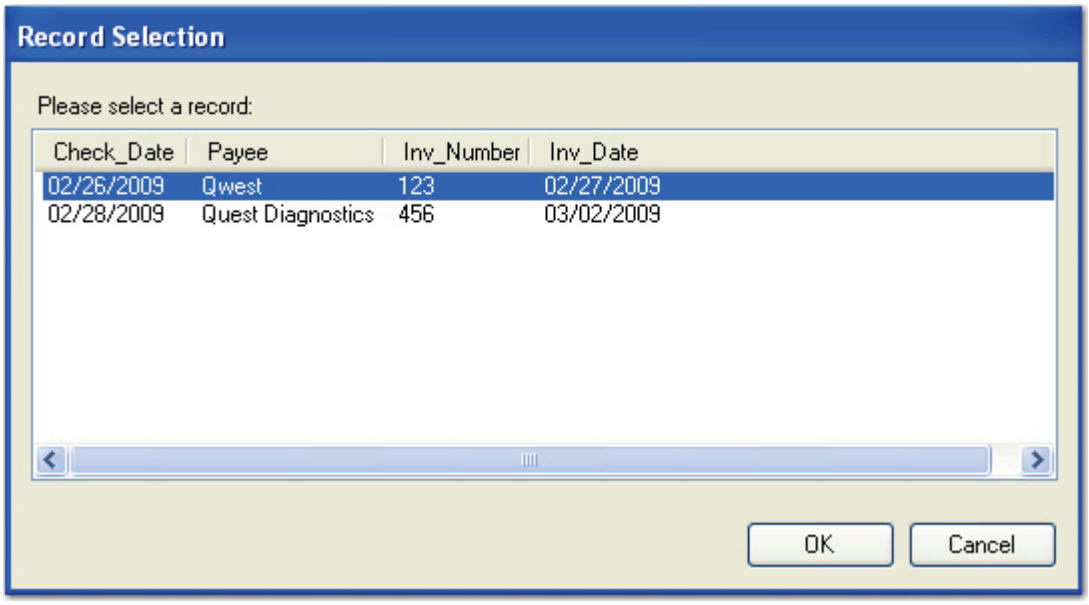
> **Note**:
>
> See the **Configuring the Match and Merge Wizard** section for information on configuring this operation.

To match and merge your index values:

1. Enter the required index values in the **Indexing** window.
2. In the **Edit** page, click the **Match and Merge**  icon. The entries found in the database populate the remaining index fields.

> **Note**:
>
> A message will appear if no matches were found.

3. If more than one entry is found, you will be presented with a list of records in the **Record Selection** dialog. Highlight the record and click **OK**.



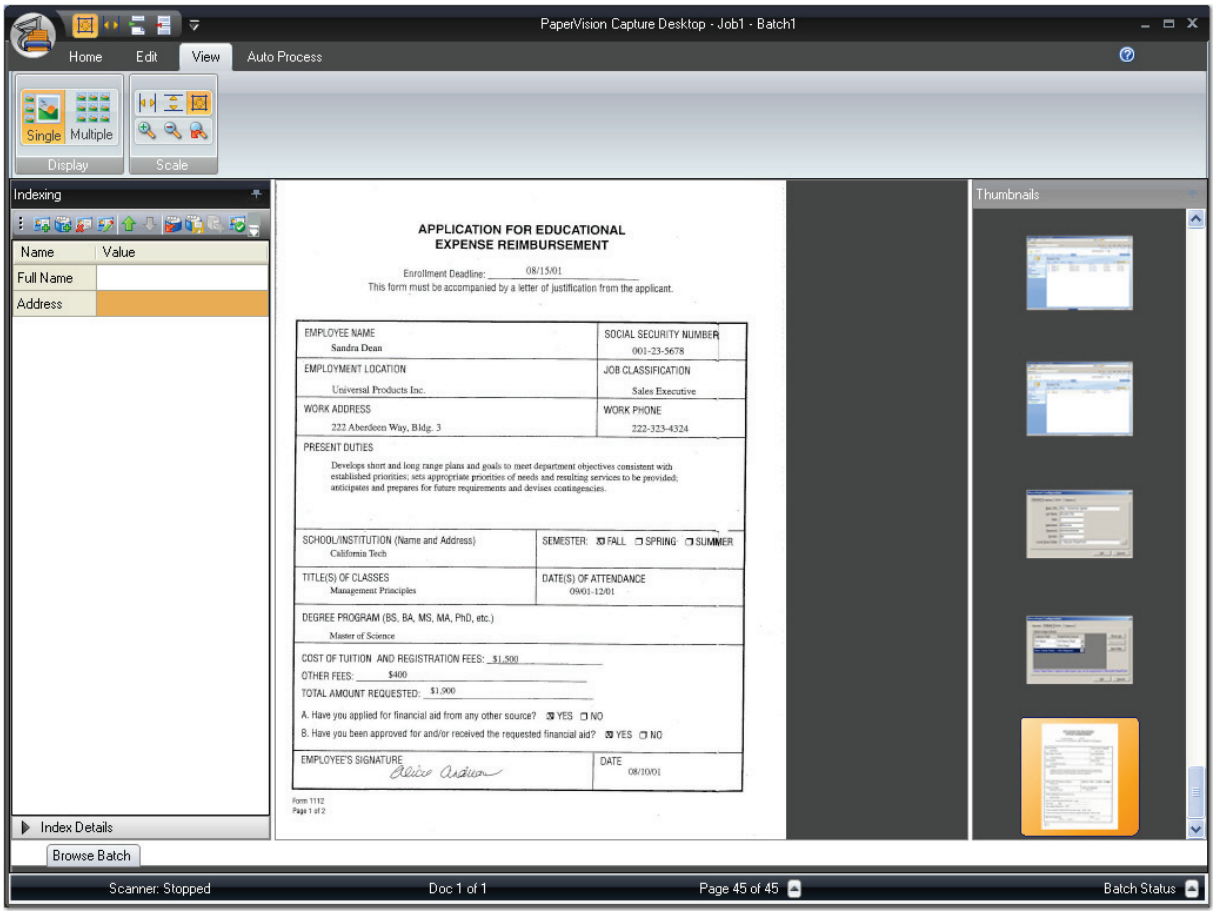
Record Selection

> **Note**:
>
> Click a column header to sort the records in ascending or descending order, and the sort order will be maintained for the duration of your logon session (i.e., the records return to their original sort order after you close the application).

4. If you do not see the correct record, click **Cancel.**

Chapter 4 - View Page

The View page contains document and page viewing operations that help you manage page display settings in PaperVision Capture Desktop. Additionally, you can alter the scaling of each page (Scale to Height, Scale to Width, and Scale to Window) and execute zoom (Zoom In and Zoom Out) operations. You can also index documents (and all associated indexing operations) within the View page.



View Page

Display

The Display toolbar group operations allow you to choose a single-page display or multiple-page display for documents in the main viewing window.

- To view one page at a time in the main window, select the **Single**  icon.
- To view multiple pages in the main window, select the **Multiple**  icon**.** To assign the number of rows and columns that appear at a time, see Display Preferences for more information.

> **Note**:
>
> Windows that you open or close in single-display and multiple-display views, such as Thumbnails, Indexing window, and Browse Batch, etc., are retained in each view.

Scale

These commands fit images to the height, width, or both height and width of the screen.

- Click the **Scale to Height**  icon to scale the image to the height of the main window.
- Click the **Scale to Width**  icon to scale the image to the width of the main window.
- Click the **Scale to Window**  icon to scale the image to fit the entire image in the main window.

Zoom Operations

- To zoom in on an area of the image, click the **Zoom In**  icon.

> **Note**:
>
> You can also use the left mouse button to draw a boundary that will zoom into an area of the image.

- To zoom out of the current view of the image, click the **Zoom Out**  icon.
- This command sets the image rotation back to its original view and displays the image using the Scale to Window view. To reset the image to its original view, click the **Reset**  icon.

Zooming and Scaling - Order of Operations

In PaperVision Capture Desktop, index zoom zones (described in the Basic Indexing Properties topic), user-defined zoom zones (Zoom In/Out), and scaling (Scale to Height/Width/Window) can be applied to images viewed in the application. Zoom zones and scaling operations are applied in the following order:

1. Any defined index zoom zones take precedence over all zoom zone or scaling settings that have been applied on images in the Operator Console. So, when you navigate to an index field defined by an index zoom zone, the index zoom zone will be applied.
2. If you enable the **Hold Zoom Regions** system preference, any zoom zone that you apply will be retained until you navigate to an index field defined with a different zoom zone.
3. If you enable the **Hold Zoom Regions** system preference, any zoom zone that you apply on the image will be retained across all index fields, pages, and documents.
4. If you disable **Hold Zoom Regions** system preference, any zoom zone that you apply on the image will not be retained across index fields, pages, and documents.
5. Finally, any scaling that you apply on the image takes precedence over the Hold Zoom Regions system preference and any zoom zone that you have applied on the image. However, scaling the image does not impact any defined index zoom zones, which take precedence over all settings and operations.

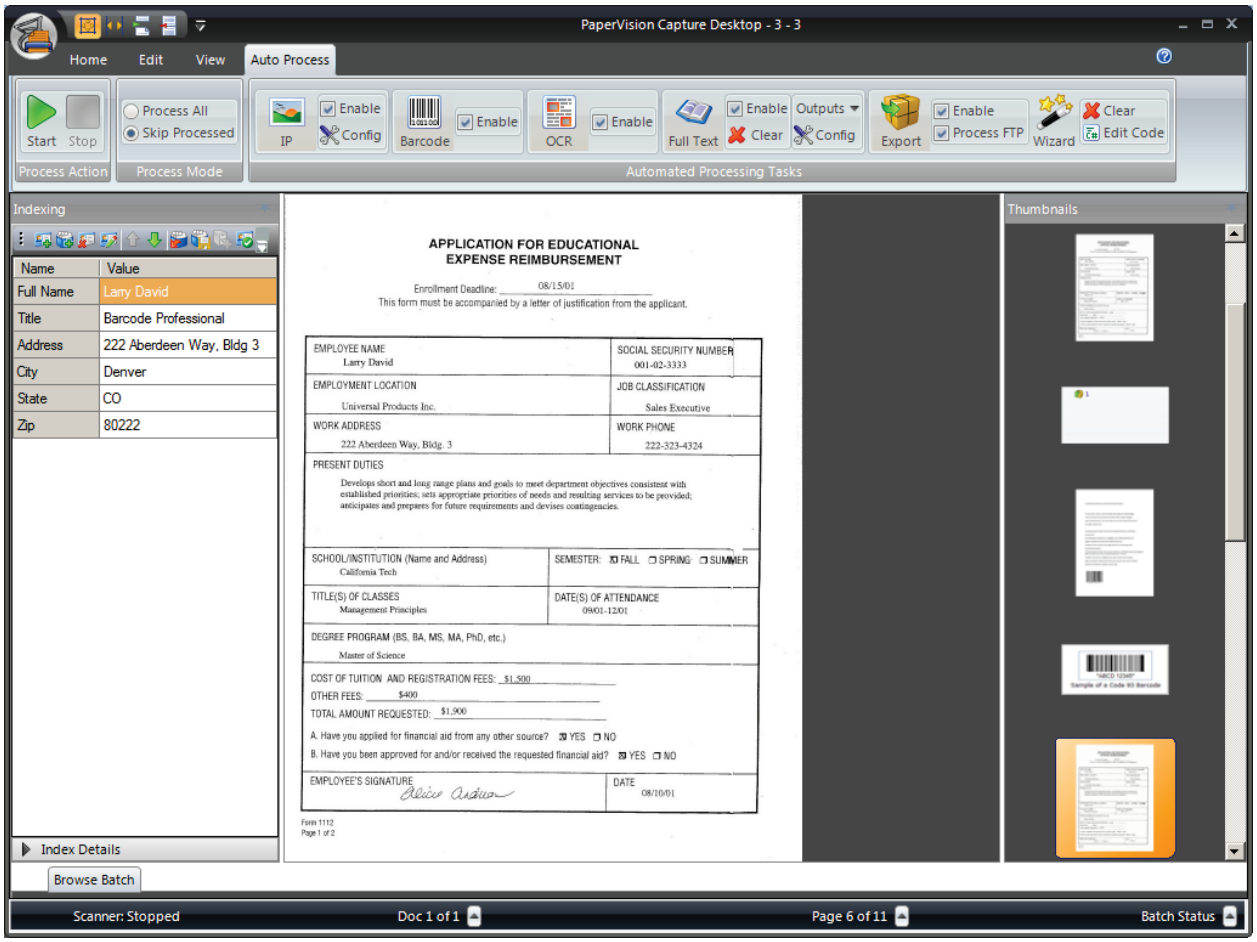
> **Note**:
>
> For more information, see the Hold Zoom Regions setting description in the **System Preferences** section of Chapter 1.

Chapter 5 - Auto Process Page

The Auto Process page allows you to select the automated processing tasks required to process the batch. When you configure the Image Processing or Full-Text automated task, a context-sensitive page will launch that contains only the operations and properties specific to the task. If you want to include the barcode and zonal OCR automated tasks as a post-process (after scanning and indexing), you must first define barcode and OCR zones in the Home page. To include a task during automated processing, select the Enable check box next to each required task. If an automated task is not yet configured, the Enable check box will not be available for selection.

In the Auto Process page, you can define PaperVision Capture Desktop exports in a wizard that contains a simple, tabbed graphical user interface. To expedite export configuration, default properties appear in drop-down menus, editable fields, and check boxes that you can easily modify.

The Process Action toolbar group contains the Start and Stop operations. The Process Mode allows you to select whether to process all documents in the batch (regardless of whether previously processed) or skip documents already processed. Your selected automated processes and settings are retained for the current job, so you will not have to repeatedly assign the same settings for subsequent batches.



Auto Process Page

Process Action (Starting/Stopping)

The Process Action toolbar group contains the Start and Stop commands that start and stop automated processing tasks.

To start and stop automated processing:

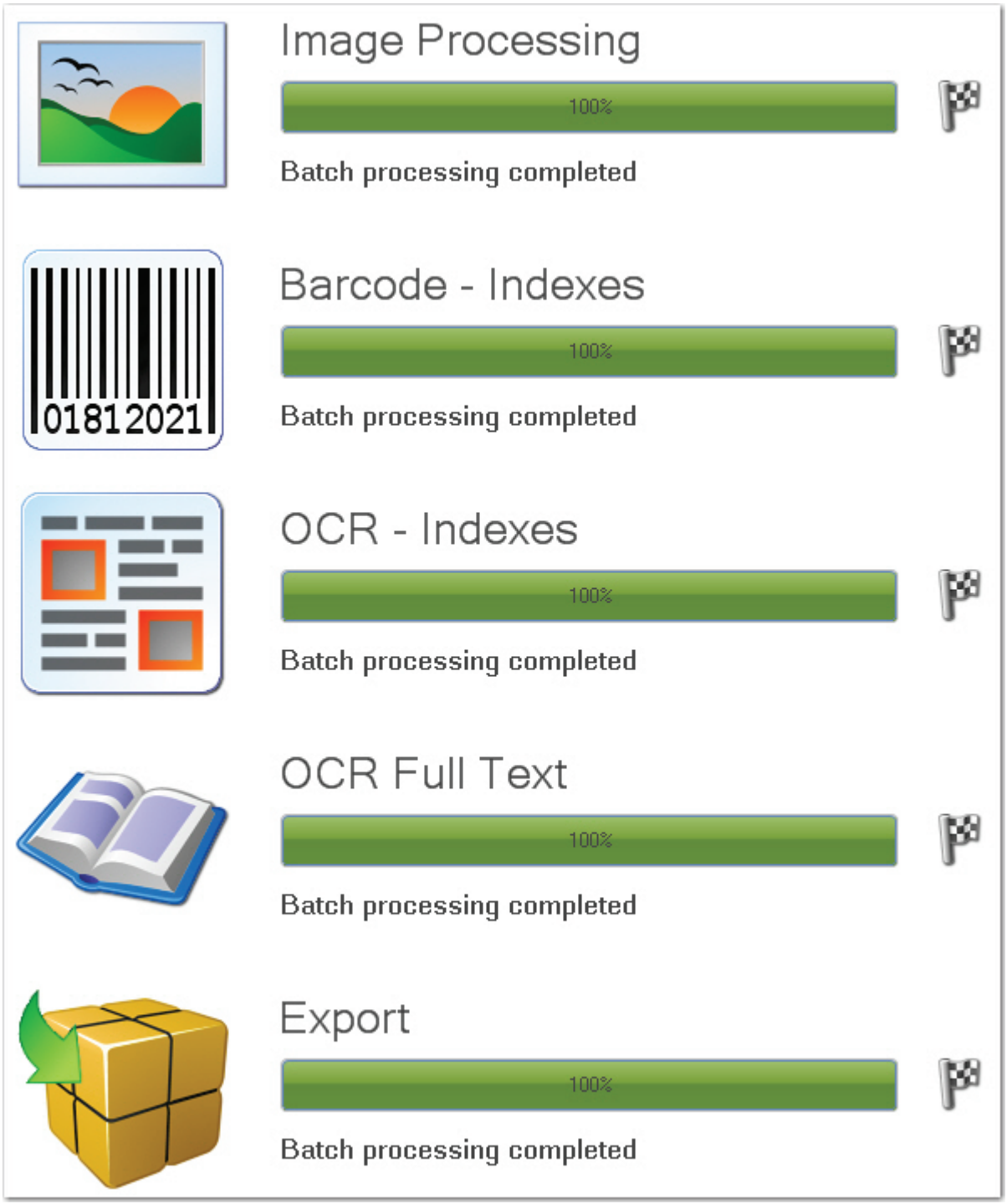
1. Ensure all required automated processing tasks have been configured.
2. Select the **Enable** check box next to all required automated tasks.
3. To execute the automated processing tasks, click the **Start**  icon.
4. To cancel (stop) an automated process, click the **Stop**  icon.

> **IMPORTANT!**
>
> If you attempt to close PaperVision Capture Desktop while automated processing is executing, you will be prompted to stop all automated processing tasks currently in progress. When you stop an automated process already in progress, the remainder of the current document will be processed, so larger documents may take several minutes to process.

While automated processing executes in the Auto Process page, the following information displays on screen:

- Whether the batch was processed successfully for each task
- Number of documents (document x of x) completed for each task
- Number of pages (page x of x) completed for each task
- If an error occurs, descriptions of each error appear on screen

> **Note**:
>
> If a task fails to process, subsequent tasks will not be processed.



Auto Process Progress Bar

5. To exit the progress page and return to the Auto Progress page, click the **Exit** 

icon.

Process Mode

The Process Mode allows you to select whether to process all documents in the batch (regardless of whether previously processed) or skip documents that were already processed. The default selection is **Skip Processed.**

If you opt to skip documents that have already been processed, a progress bar will read "Skipping Document x of x" when you execute automated processing. You cannot stop automated processing while documents are being skipped since the Stop operation only executes when documents are being processed.

Image Processing

Image processing filters help remove unnecessary borders, lines, and noise from images; they also enhance text readability and reduce file size. Additional image processing filters evaluate images, and then keep or discard them based on your defined criteria. Color detection filters identify your specified colors and convert the image to black and white or remove the page containing the color image. Binary filters can only be applied to bitonal (1 bit per pixel) images; color and grayscale are ignored.

To enable automated image processing:

1. You must first define IP filters and/or zones. See the **Image Processing Configuration** section for more information.
2. When finished with image processing configuration, return to the **Auto Process** page.
3. In the Auto Process page, click the **Enable** check box in the **IP** toolbar group.

Image Processing Automated Task - Enabled

4. To begin automated processing, click the **Start**  icon.

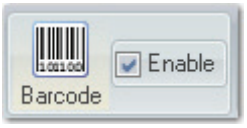
Barcode

You can use barcodes to populate new or existing index values and insert document breaks. PaperVision Capture Desktop recognizes one- and two-dimensional, black and white, and color barcodes. In the Auto Process page, you can configure a barcode reading process that executes automatically during post-processing. For information on configuring a barcode reading process while you scan (on the fly), see the **Home Page** chapter.

> **Note**:
>
> Use of the binary scaling image processing filter can improve the recognition rate of barcode detection. For more information, see the Image Processing Filters section.

To enable an automated barcoding process:

1. You must first define barcode zones and properties in the Home page. See the **Barcode Zone Configuration** section for more information.
2. When finished with barcode zone configuration, return to the **Auto Process** page.
3. In the Auto Process page, click the **Enable** check box in the **Barcode** toolbar group.



Barcode Automated Task - Enabled

4. If you are ready to execute automated processing, click the **Start**  icon.

Zonal OCR

You can customize zonal Optical Character Recognition (OCR) settings for individual index fields or to insert document breaks. Character recognition options allow you to customize how values are recognized by OCR processing.

To enable automated zonal OCR:

1. You must first define OCR zones and properties in the Home page. See the **OCR Zone Configuration** section for more information.
2. When finished with barcode zone configuration, return to the **Auto Process** page.
3. In the Auto Process page, click the **Enable** check box in the **OCR** toolbar group.

OCR Automated Task- Enabled

4. To begin automated processing, click the **Start**  icon.

Full-Text OCR

You can configure a full-text OCR automated process that reads pages of text and converts recognized results to one or multiple file types. Once configured, this automated process executes automatically during post-processing.

Extracted text is converted into various file types such as .txt, XML, and PDF. Each converter output type contains unique settings that you can configure to support your full-text OCR requirements. Prior to saving your settings, you can test and preview the full-text OCR results.

To enable automated full-text OCR:

1. You must first select and configure full-text OCR outputs. See the **Full-Text OCR Configuration** section for details on configuring each output.
2. When finished with full-text OCR configuration, return to the **Auto Process** page.
3. In the Auto Process page, click the **Enable** check box in the **Full Text** toolbar group.

Full-Text OCR Process - Applied

4. To begin automated processing, click the **Start**  icon.

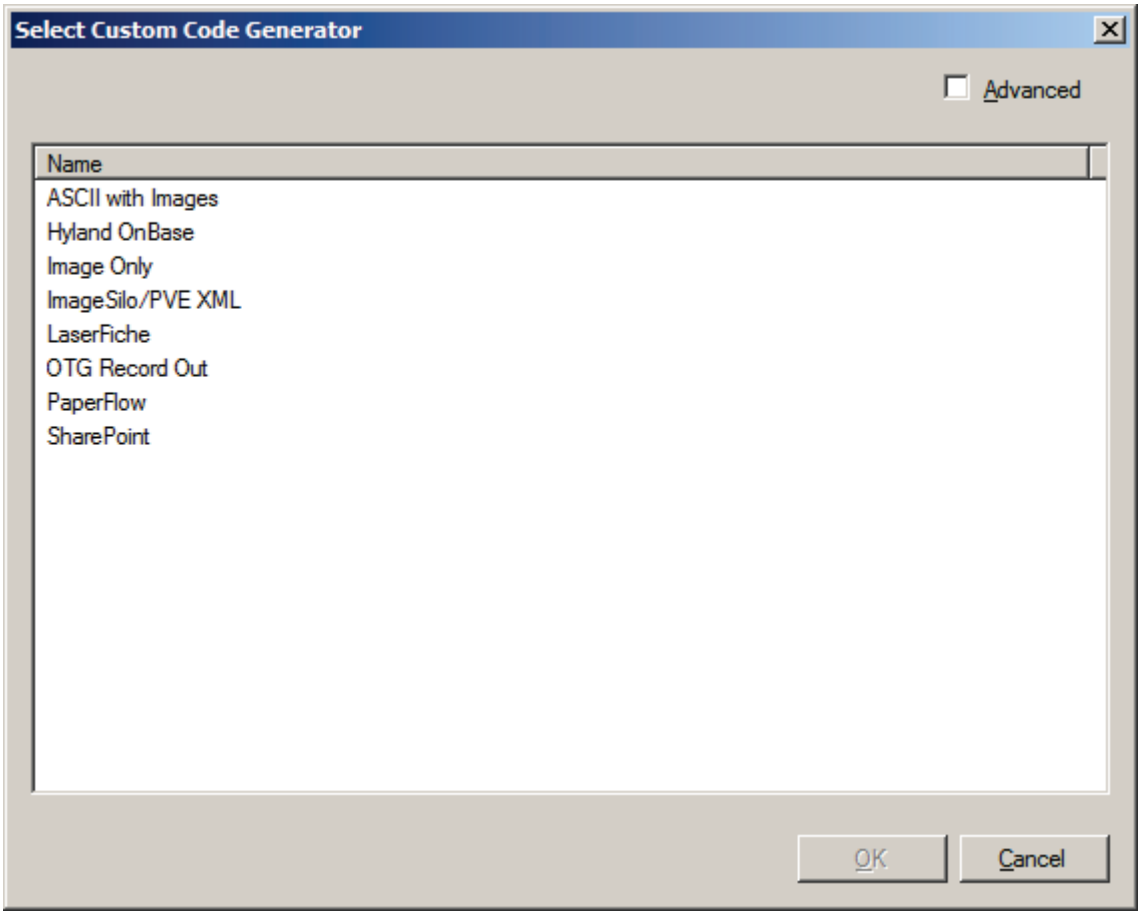
Exports

When you configure an export, you can select the custom code generator that will execute automatically during batch processing. Custom Code generators include all exports and customizable scripts (via the Export Template generator) that contain pre-written, generic code to edit and compile directly in the Script Editor window. You can configure Custom Code generators within a graphical user interface that displays only the applicable properties for your selection. Default settings are provided for each generator within drop-down menus, editable fields, and check boxes (indicating a default **True** or **False** setting). The Export Template generator provides a generic template for custom exports that you can execute automatically during batch processing.

> **IMPORTANT**:
>
> The Visual Basic programming language can only be used with the Export Template.

To configure and process exports:

1. In the Auto Process page, click the **Wizard** icon in the **Export** toolbar group. The **Select Custom Code Generator** dialog box appears, where exports are listed.



Select Custom Code Generator

> **Tip**:
>
> To view advanced settings, select the **Advanced** check box. Advanced settings include programming language options (C# or Visual Basic) and the Export Template, which contains additional pre-defined code that will automatically process batches.

2. Double-click the export, and the selected export and corresponding properties appear in tabbed dialog boxes.

> **Note**:
>
> See the **Exports** chapter for information on specific exports and their associated properties.

   - Default values and applicable index fields are provided for your reference, and drop-down menus contain only the options specific to your selected export.
   - You can manually enter file paths or browse to the appropriate directory.

3. After you have configured the export, click **OK** in the **Export Configuration** dialog box.
4. Click the **Enable** check box in the **Export** toolbar group.
5. To process the export, click the **Start**  icon in the **Process Action** toolbar group. A progress bar will appear, indicating whether the batch was successfully processed.

Export Wizard

If you have not yet manually edited the resulting code in the Script Editor, you can edit properties in your existing export by clicking the **Export Wizard**  icon in the **Export** toolbar group. The tabbed **Export Configuration** dialog box will appear, allowing you to edit previously assigned values. See the **Exports** chapter for more information on specific exports.

Editing Code

If you have already configured an export using the export wizard and you want to edit the resulting code in the Script Editor, click the **Edit Code**  icon. However, if you further edit the custom code script with the Custom Code Generator Wizard, your previous code changes in the Script Editor will be permanently lost. A warning prompt will notify you if you attempt to make export modifications in this manner.

Clearing Export Scripts and Code

You can remove your existing export script and code from the Auto Process page.

WARNING!

This operation is irreversible.

To clear the existing export:

1. In the Auto Process page, click the **Clear** icon in the **Export** toolbar group.
2. To confirm the removal, click **OK**.

FTP

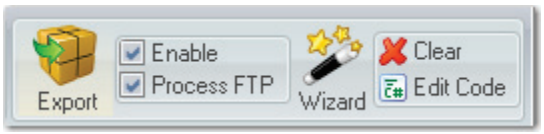
The FTP operation securely transfers your documents, images, and associated index values to an FTP site. Before processing FTP, ensure the required FTP settings have been configured in the PaperFlow or ImageSilo/PVE XML custom code generator wizard (**Auto Process page** > **Wizard** > **PaperFlow** or **ImageSilo/PVE XML**). You can execute FTP from the Auto Process page, PaperVision Capture Desktop Button, or from the Start page.

Note:

For more information on PaperFlow or ImageSilo/PVE XML export configuration, see **Chapter 9 - Exports.**

To start (and stop) the FTP process:

1. Ensure you have configured the required FTP properties for the PaperFlow or ImageSilo/PVE XML custom code generator (export).
2. In the Auto Process page, select the **Enable** and **Process FTP** settings in the **Exports** toolbar group.



Export and FTP Enabled

3. To process FTP from the Auto Process page, click the **Start** icon in the **Process Action** toolbar group.

- Alternatively, from the **PaperVision Capture Desktop Button**  , select the **FTP**  icon.
- To process FTP from the Start page, select the **FTP** tab, and then click the **Start FTP** 

button.

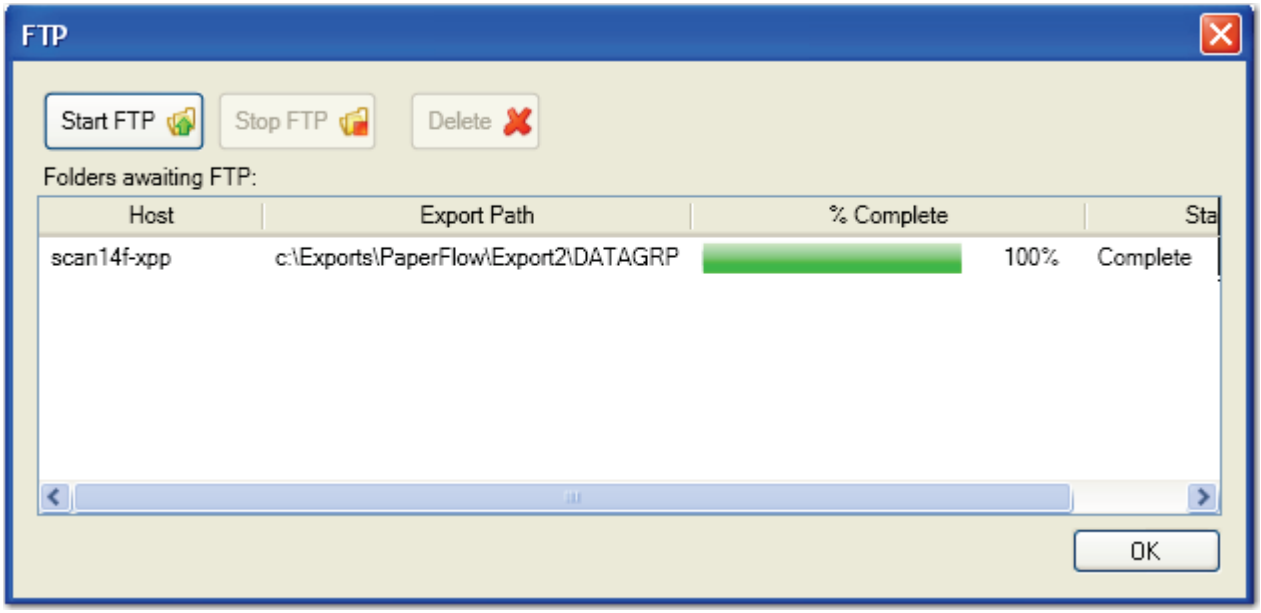
> **Note**:
> A notification will inform you if no folders are awaiting transfer.

If an error occurs during FTP transfer, select from one of the following options:

- Skip Folder and Continue
- Remove Folder and Continue
- Retry Folder (With this option, the system will attempt to FTP the same folder until it is successfully transferred to the FTP site. As a result, the remaining folders may not be transferred.)
- Cancel FTP

> **Tip**:
> If you frequently select the same option, you can select **Always perform selected action without prompt**. Alternatively, you can enable the FTP user preference, **Do Not Prompt to Select Action** (and, select the preferred action).

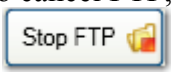
While FTP processes, the **FTP** progress bar appears:



FTP

While FTP executes, the FTP dialog box displays the following information:

- FTP host site name
- Export path
- Percentage of FTP completed
- Status of FTP (e.g., FTP uploading file x of x, complete, etc.)

4. If you do not want to cancel FTP, proceed to step 6. To cancel the FTP operation, click the **Stop FTP**  button.

5. To restart FTP, click the **Start FTP**  button.

6. If FTP finishes successfully, click **OK** in the **FTP** dialog.
7. To close the progress page, click the **Exit**  button.

Deleting Folders in FTP Queue

Before you have started the FTP process, you can delete one or more folders awaiting transfer.

To delete one or more folders:

1. In the **FTP** dialog, select one or more folders.
2. Select the **Delete** 

button.
3. Click **Yes** to confirm the deletion.

Chapter 6 - Image Processing

You can define image processing zones that will be processed while you scan documents (on the fly) in the Home page. Alternatively, you can configure image processing filters (to be applied on entire images) in the Auto Process page. Binary image processing includes filters such as border removal, crop, dilation, erosion, halftone removal, hole removal, noise removal, scaling, and others. Page deletion filters allow you to specify certain parameters that determine whether pages are retained in a batch. Additionally, you can apply color filters as well as deskew, rotation, and threshold filters. You can configure image processing properties including the file type for colored images, image processing filters, and whether to save processed images. This automated process provides you the flexibility to apply image processing filters on the entire image or within specific zones that you define.

When you configure image processing filters, you can view a side-by-side comparison of the original image alongside the filtered image (click the **Test** button in the toolbar to view the comparison). Thumbnail previews display the document's images and allow you to navigate through the document and perform basic operations including the cut/paste, copy/paste, and delete operations.

You can assign the page ranges that will be applied to each filter in the IP Filter grid (click the **Show Advanced Configuration** link in the toolbar to view the grid). Within this grid, you can also see what filters are applied to specific page ranges (Apply and Page Range columns). In addition, you can view the dimensions (in millimeters) of each image processing zone that you have applied. For information on each filter's specific properties, see the **Image Processing Filters** section.

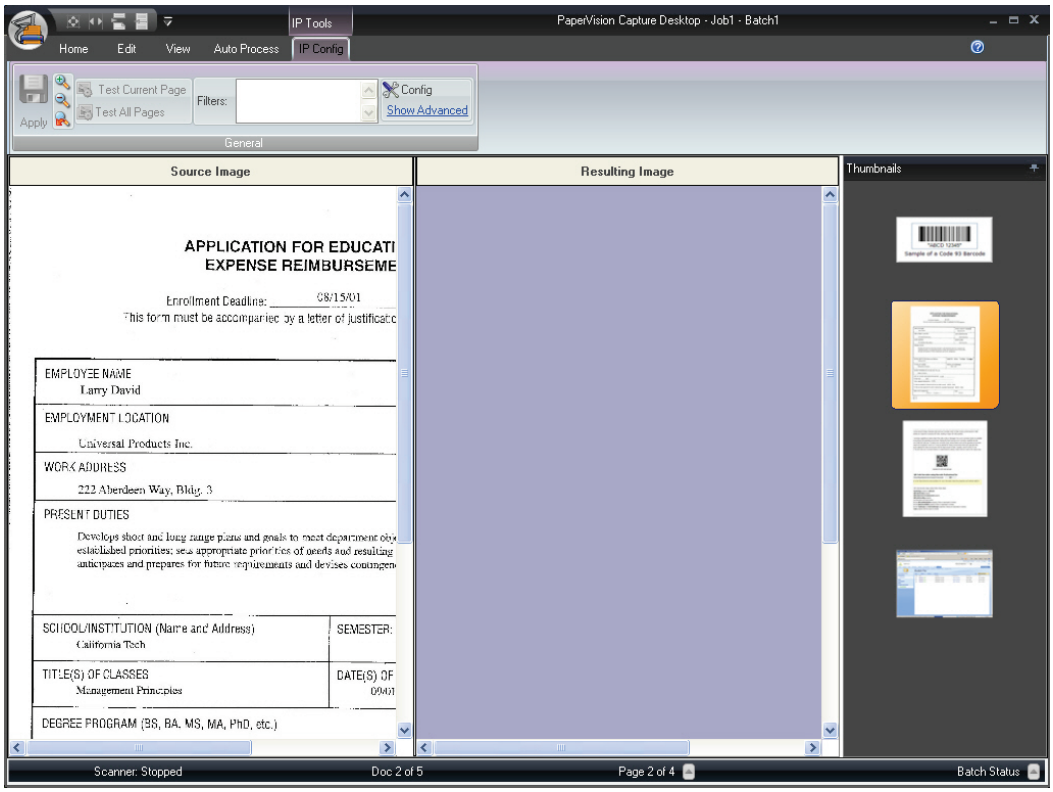


Image Processing Configuration Page

Note:

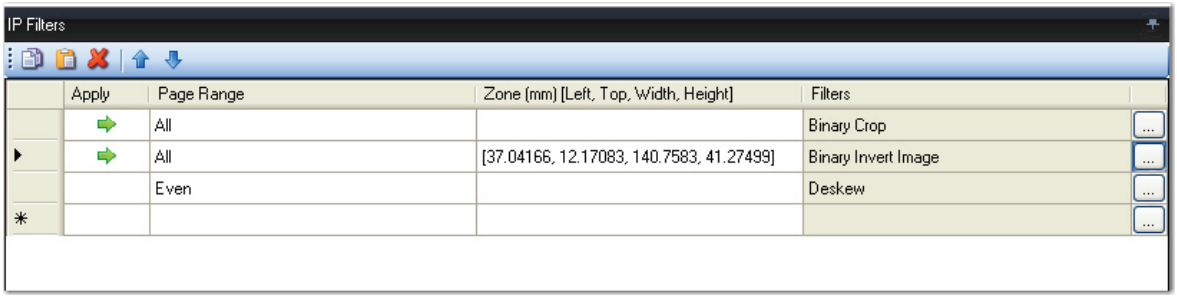
Incoming color images can have maximum dimensions of 10,000 x 10,000 pixels when they are processed through the application. Bitonal (black and white) images can have slightly larger dimensions. Larger images can be ingested into PaperVision Capture Desktop provided that no OCR will be performed on the images; no image processing will be performed on the images; or, images will not viewed as thumbnails.

Configuring Image Processing Filters

You can configure, preview, and test image processing filters before processing the batch. Zooming operations are available if you need to adjust the image. When you select the **Show Advanced Configuration** link, the IP Filters grid opens, and you can also draw and configure IP zones for certain binary filters.

The Image Processing Configuration page contains the following components:

- The **Source Image** window displays the original, unfiltered image.
- The **Resulting Image** window displays the filtered image, after you test the image.
- The **IP Filters** grid (select **Show Advanced Configuration** to see the grid) displays all page ranges and configured filters for each page range. The dockable **IP Filters** grid allows you to select the page range and apply image processing filters to specific pages or zones. Select the **Page Range** from the drop-down list (all, odd, even, or last). Or, enter the page range (e.g., 1; 1-5, 4; 1-7, etc.).



IP Filters Grid

To configure IP filters:

1. In the IP Config page, click the **Config** icon. The **Image Processing Filters** dialog appears. Filters supported in zones are marked with asterisks (*).

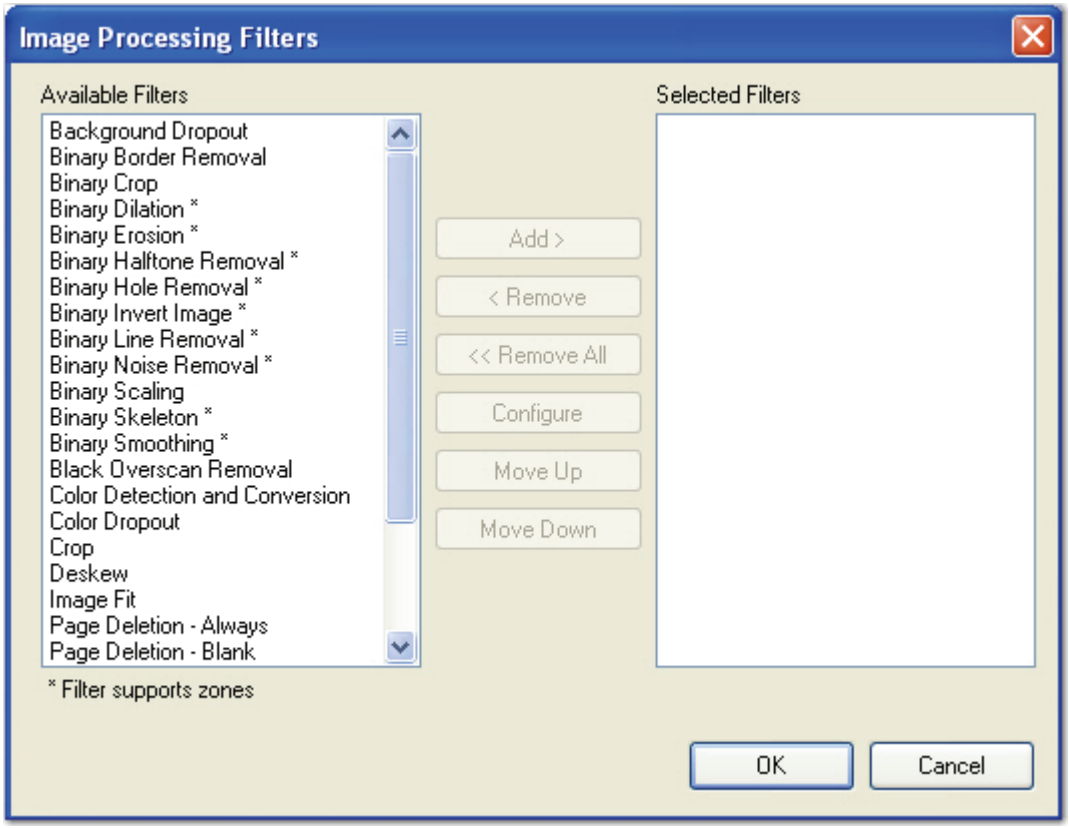


Image Processing Filters

2. From the **Available Filters** list, highlight the filter, and then click **Add**.
3. To configure a selected filter, highlight the filter in the **Selected Filters** list, and then click **Configure**.
    - To remove a filter from the **Selected Filters** list, highlight the filter, and then select the **Remove** button.
    - To remove all filters from the **Selected Filters** list, click the **Remove All** button.
    - To move a filter up or down in the **Selected Filter** list, highlight the filter, and then select **Move Up** or **Move Down**.
4. Click **OK** after you have configured all filters.
5. After you have configured all IP filters, click the **Apply** icon to save the filter settings.
6. To close and exit out of the IP Configuration page, click the **Auto Process** tab.

Removing and Reordering Filters

You can reorder a filter so that it will be processed before the other selected filters.

To remove a filter from the Selected Filter list:

1. Highlight the filter(s).
2. Click **Remove.**
3. To remove all filters, click **Remove All.**

To reorder the filters:

1. Highlight the filter(s).
2. Click **Move Up** or **Move Down.**
3. Click OK.

Drawing Image Processing Zones

You can apply certain binary image processing filters to zones within bitonal images. For example, you may want to apply the Binary Hole Removal filter only to the left two inches of a bitonal image or the Binary Invert Image to expose a specific area of a bitonal image. During IP configuration, you can use the Draw IP Zone operation to draw a zone on the image. The following binary IP filters can be applied to zones that you define on the image:

- Binary Dilation
- Binary Erosion
- Binary Halftone Removal
- Binary Hole Removal
- Binary Invert Image
- Binary Line Removal
- Binary Noise Removal
- Binary Skeleton
- Binary Smoothing

> **Note**:
>
> Descriptions for each filter can be found in the **Image Processing Filters** section.

To draw IP zones:

1. In the IP Config page, click **Show Advanced** in the toolbar.
2. To equip the cursor to draw a zone on the source image, click the **Draw IP Zone**  icon.
3. Drag the crosshair cursor around the zone on the image.
4. Adjust the zone borders if necessary.

5. In the IP Filters window, click the ellipsis button next to the **Filters** property in the first row. The **Image Processing Filters** dialog appears.

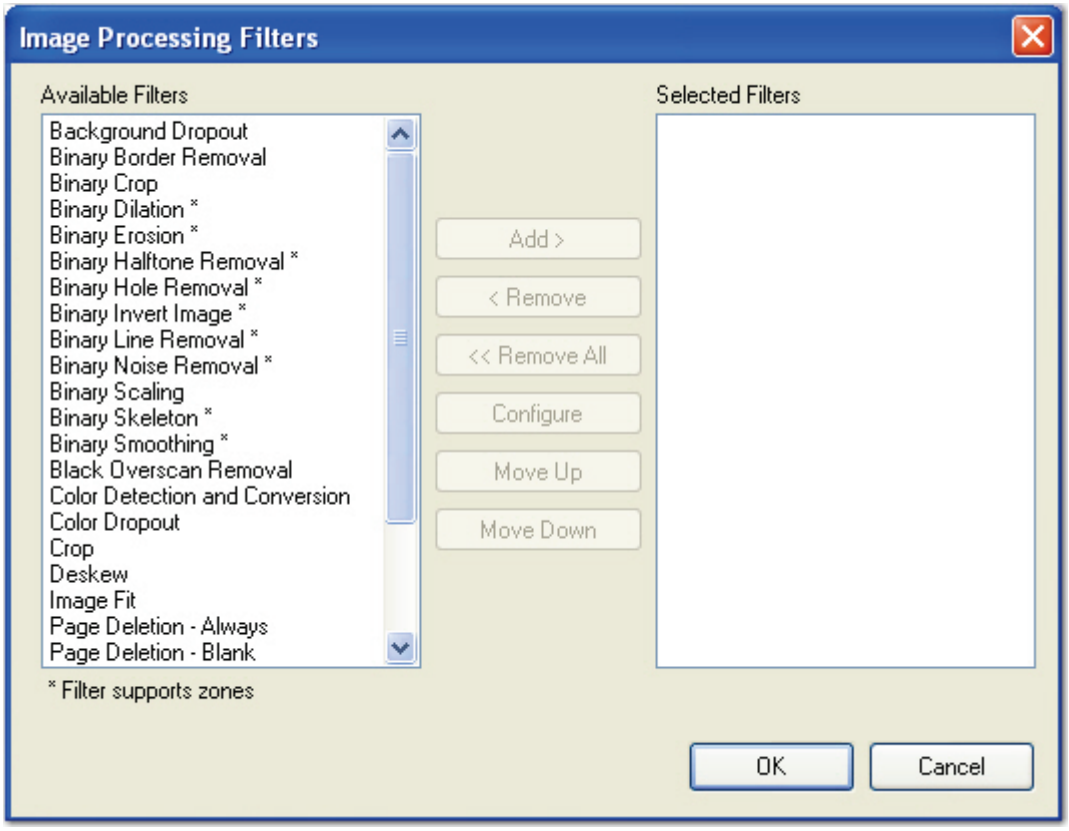


Image Processing Filters

6. The dockable **IP Filters** grid allows you to select the page range and image processing filters. If an image processing zone is configured, its dimensions (in mm) appear in the **Zone** column. Select from the **Page Range** column drop-down list (all, odd, even, or last), or enter the page range (e.g., 1; 1-5, 4; 1-7; etc.).
7. Filters supported in zones are marked with asterisks (*). From the **Available Filters** list, highlight the filter, and then click **Add**.
8. To configure a filter, highlight the filter in the **Selected Filters** list, and then click **Configure**.
9. Click **OK** after you have configured the filters. The **IP Config** screen appears once again, where you can test the zone to ensure the filters work correctly. See the next section on **Testing Image Processing Filters** for more information.
10. Click the **Apply**  icon to save the IP filter settings.

To move an IP zone:

1. Select the center of the zone until the cursor turns into a four sided arrow.
2. Move the zone to the appropriate location on the image.
3. Click the **Apply**  icon.

To remove an IP zone:

1. Select the zone.
2. Click the **Remove IP Zone**  icon.
3. Click the **Apply**  icon.

Testing Image Processing Filters

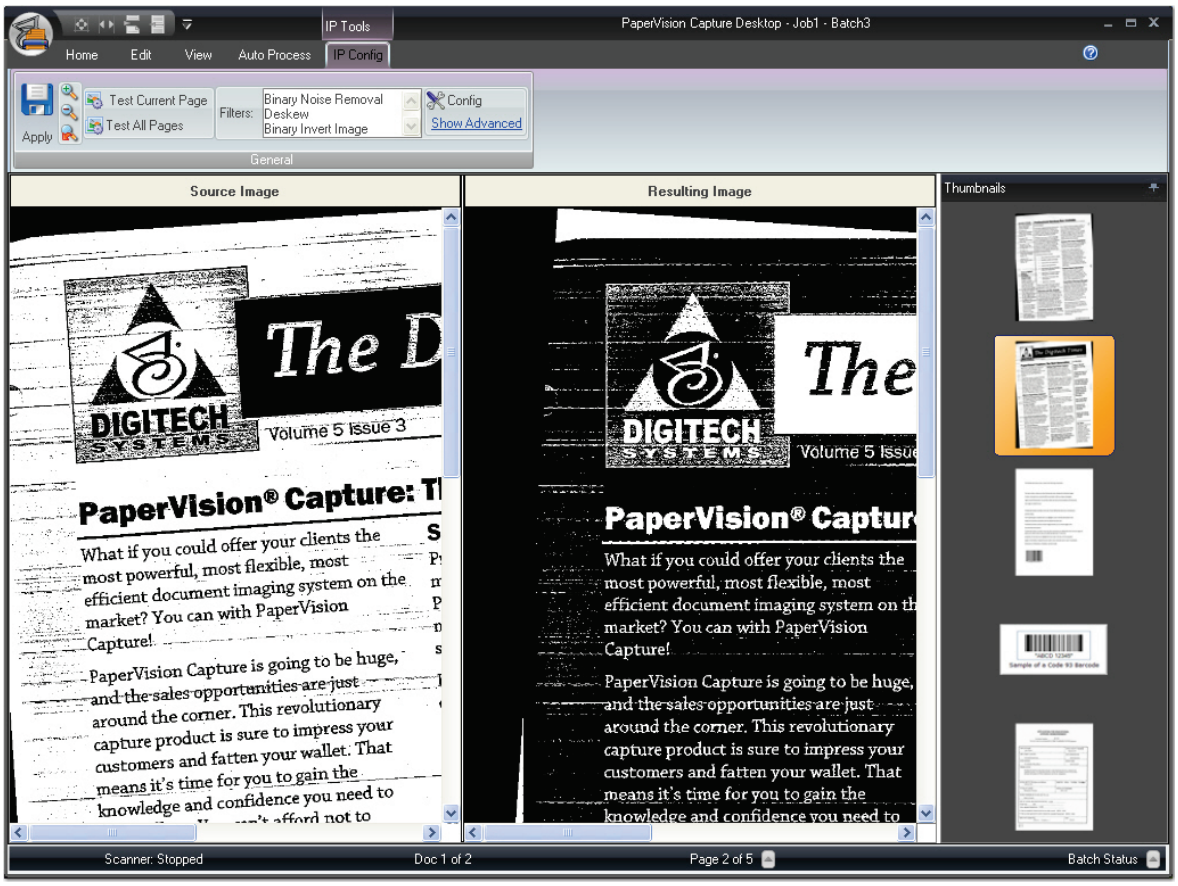
You can test and preview individual or all IP filters that are applied to pages in the document.

To test image processing filters for the current page:

1. After configuring the filters for a page, click the **Test Current Page**  icon. The filtered image appears in the **Resulting Image** window.
2. If the filter is acceptable, click the **Apply**  icon.

To test image processing filters for all pages:

1. After configuring the filters for all pages, click the **Test All Pages**  icon.
2. Navigate through the document to ensure the filters are acceptable, and adjust them if necessary.
3. If filters for all pages appear acceptable, click the **Apply**  icon.



IP Configuration - Test Current Page

Zooming Operations

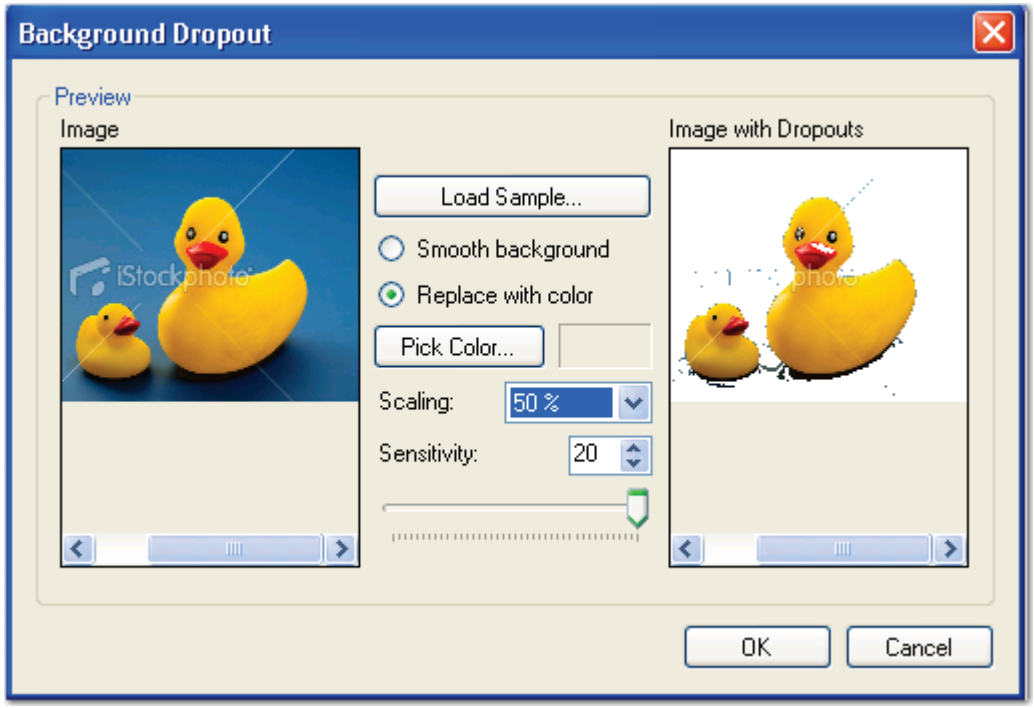
- To zoom in on the image, click the **Zoom In** icon.
- To zoom out of the image, click the **Zoom Out** icon.
- To reset the view of the image, click the **Zoom Reset** icon.

Image Processing Filters

Image Processing filters improve image quality by removing unnecessary borders, lines, and noise; enhancing text readability; and reducing file size. Additional image processing filters evaluate images, and then keep or discard them based on your defined criteria. Color detection filters identify your specified colors and convert the image to black and white or remove the page containing the color image. Binary filters can only be applied to bitonal (1 bit per pixel) images; color and grayscale are ignored.

Background Dropout

This filter is intended to be used on color images with contrasting text or a uniform background of the same color or similar colors. The background is a set of pixels of the same or similar color that covers the majority of the image, contrasting with other informative pixels. Background detection is based on the image histograms of red, green, and blue (RGB) channels. Only the margins of the image are used for histogram analysis, assuming that margins are free from any information and clearly represent the background of the image.



Background Dropout

To load a sample image and apply the Color Dropout filter:

1. Click the **Load Sample** button.
2. Browse to the directory, and then select the image.
3. Click **Open**. The image appears in the **Image** window on the left.
4. To zoom in/out on the image, select a larger/smaller percentage in the **Scaling** drop-down list.
5. To smooth the background color and make it appear more uniform, select **Smooth background**. The results appear in the **Image with Dropouts** window, so proceed to step 8.
6. Or, select **Replace with color** to replace the background color your selected color. Proceed to the next step.
7. Click the **Pick Color** button. The selected color appears next to the **Pick Color** button.
8. To apply a more noticeable background dropout, move the **Sensitivity** slider to the right, and the value increases.
   - Move it to the left to reduce the amount of dropout applied to the image, and the value decreases.
   - Or, enter a value between **-20** and **20**.
9. When you are satisfied with the results of the background dropout, select **OK**.

Binary Border Removal

The Binary Border Removal filter deletes the black edges that appear around images during scanning or photocopying. In the **Processing Limits** section, you can assign the number of millimeters (in whole or decimal numbers) that are removed from the top, bottom, left, and/or right borders. The size of the image does not change after this filter is applied; rather, white pixels replace the border's black pixels.

- **Use Same Value for All Sides** applies the value of the left border to all sides.
- **Process Inverted Images** removes the border if images appear inverted.

Before Binary Border Removal

After Binary Border Removal (also with Deskew)

Binary Crop

The Binary Crop filter allows you to assign margins to add and remove white space from the edge of the image. You can set different values for the top, bottom, left, and right margins.

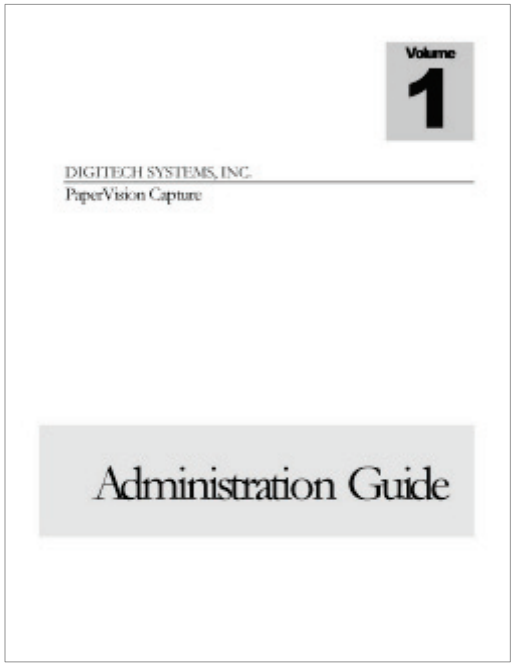
Image Margins

Positive margin values represent the white space between the edge of the image and the black pixel closest to that edge. Negative margin values crop the specified amount from the black pixel closest to the edge towards the center of the image. Enter the margin values in millimeters (in whole or decimal numbers) for the top, bottom, left, and right margins.

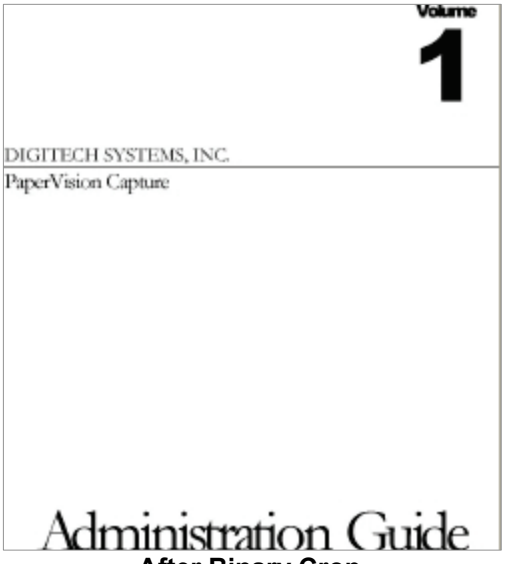
Force Symmetry

This filter assigns the same values to opposite margins. Enter a value in the **Top** field to apply the same value to the top/bottom margins. Enter a value in the **Left** field to apply the same value to the left/right margins.

> **Note**:
>
> If you enter values for the **Bottom** or **Right** fields, they are ignored.



Before Binary Crop



After Binary Crop

Binary Dilation

The Binary Dilation filter expands a black area of an image using your specified direction (horizontal, vertical, and/or diagonal) and number of times (passes) to apply the dilation. This filter can improve text legibility, but can increase file size.

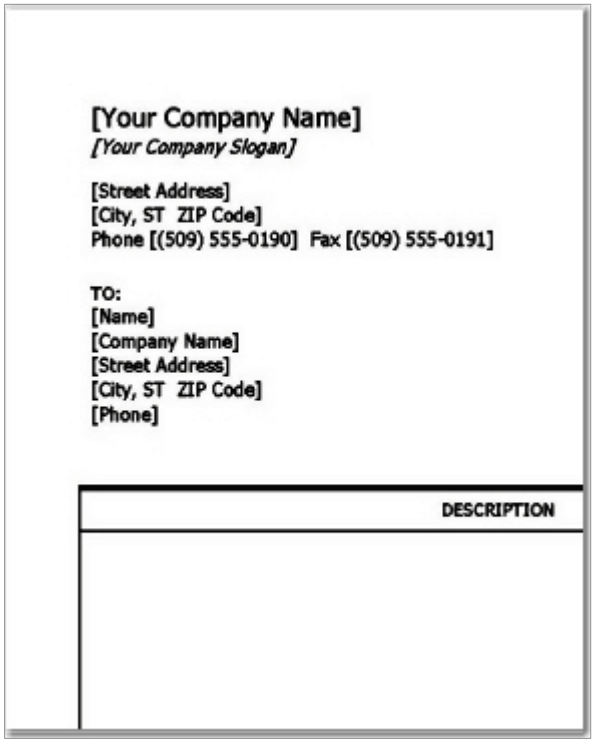

[Your Company Name]
[Your Company Slogan]

[Street Address]
[City, ST ZIP Code]
Phone [(509) 555-0190] Fax [(509) 555-0191]

TO:
[Name]
[Company Name]
[Street Address]
[City, ST ZIP Code]
[Phone]

DESCRIPTION

Before Dilation


[Your Company Name]
[Your Company Slogan]

[Street Address]
[City, ST ZIP Code]
Phone [(509) 555-0190] Fax [(509) 555-0191]

TO:
[Name]
[Company Name]
[Street Address]
[City, ST ZIP Code]
[Phone]

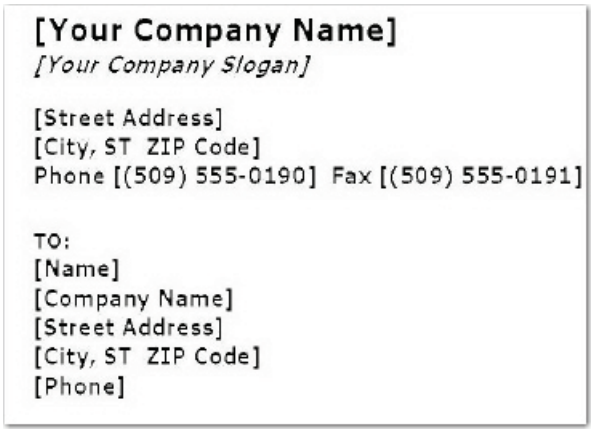
DESCRIPTION

After Dilation

Binary Erosion

The Binary Erosion filter trims an area of a black image using your specified direction (horizontal, vertical, and/or diagonal) and number of times (passes) to apply the erosion. This filter can reduce file size but causes a loss of detail in the image.

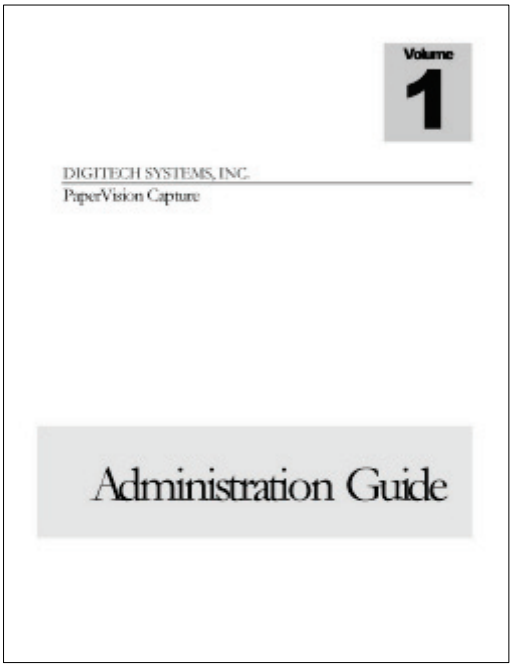
Before Erosion



After Horizontal Erosion

Binary Halftone Removal

The Binary Halftone Removal filter removes the background, such as a halftone or dither pattern, from an image.

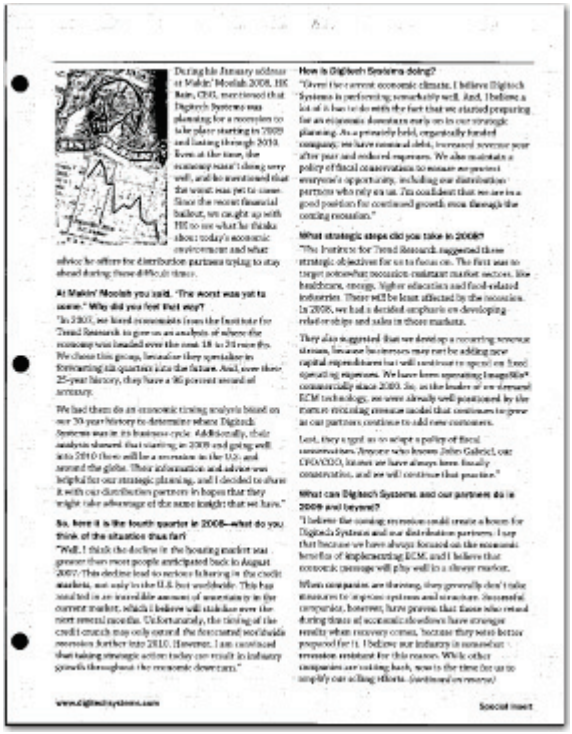


Before Binary Halftone Removal

After Binary Halftone Removal

Binary Hole Removal

The Binary Hole Removal filter identifies objects that look like binder hole punches near the edge of the image, and then deletes those objects. Objects that appear like binder hole punches that are visible in other areas of the image, such as the center, will not be removed.

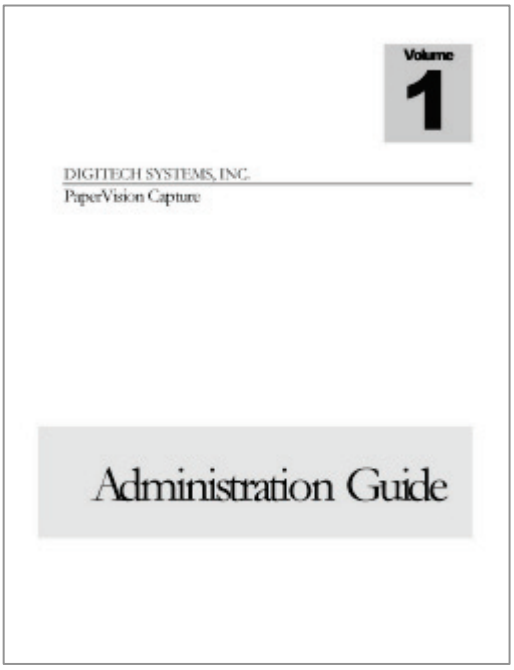


Before Binary Hole Removal

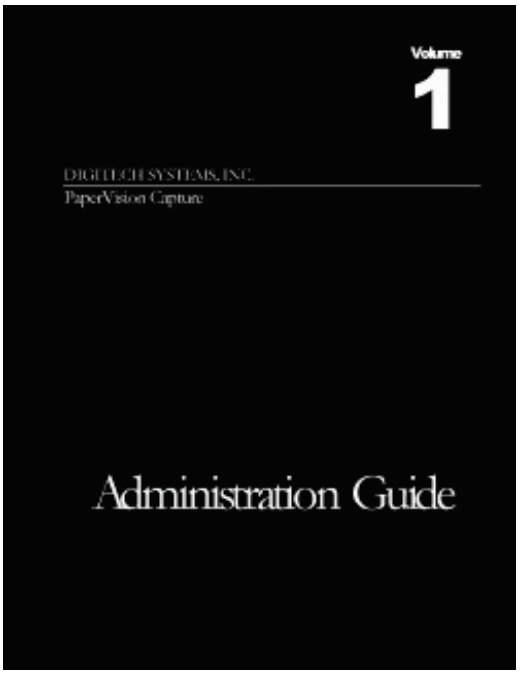
After Binary Hole Removal

Binary Invert Image

The Binary Invert Image filter reverses the polarity of the image. Black pixels become white pixels, and white pixels become black pixels.



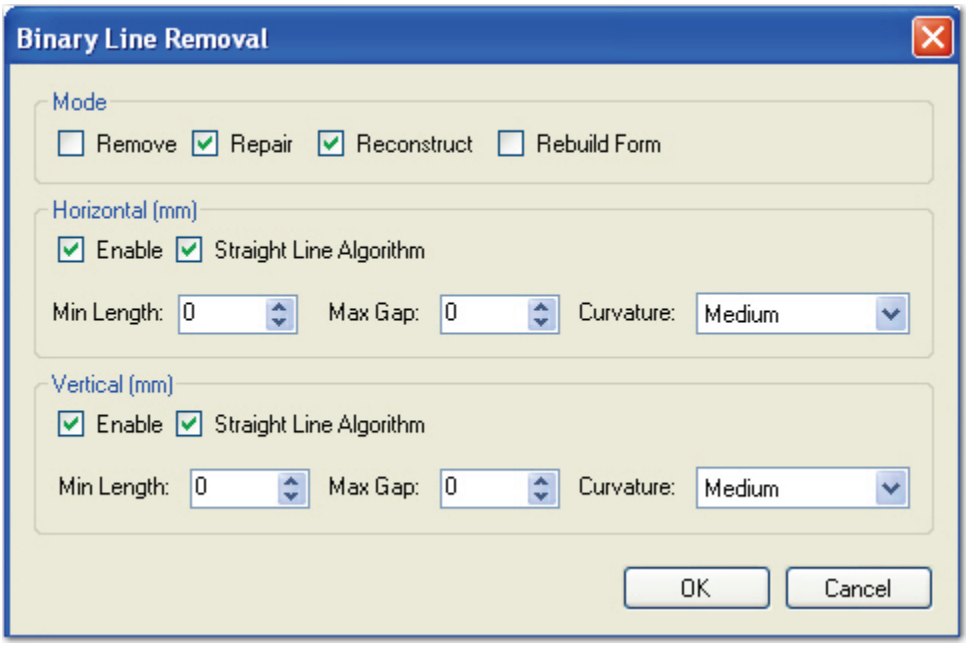
Before Binary Invert Image



After Binary Invert Image

Binary Line Removal

The Binary Line Removal filter deletes lines or reconstructs lines on a form-based image. Removing lines can reduce file size and improve OCR results.



Binary Line Removal

Mode

This setting specifies the type of line correction to perform on the page.

- **Remove Lines** takes out all objects considered as lines.
- **Repair** removes lines and repairs all graphics and text overlapped by the removed lines.
- **Reconstruct** removes lines, repairs overlapped graphics and text, and redraws straight lines in place of removed lines.
- **Rebuild Form** removes lines, redraws straight lines, and reconnects lines that were previously connected. This type of line correction is commonly used for tables and forms.

Horizontal Line Removal

Enable this setting to detect horizontal lines that will be taken out during the line removal process.

Straight Line Algorithm

The Straight Line Algorithm setting provides faster processing of straight lines that are longer than 100 pixels (suitable for forms and light paper). This setting evaluates the height or width of the bounding rectangles around line-like objects to determine if the object is a line. If this setting is not enabled, the line-like object is broken into small segments and uses the minimum length, curvature, and maximum gap to determine whether the segments comprise a line.

Minimum Length

This setting defines the minimum length in millimeters (in whole or decimal numbers) that the filter will detect as a horizontal line.

Maximum Gap

This setting defines the maximum amount of allowable white space in millimeters (in whole or decimal numbers) between two horizontal line-like objects to consider as one line.

Curvature

This setting defines the maximum allowable amount of deviation from a straight line for a horizontal line-like object to be considered a line.

- **Straight** contains a curvature value of 5.
- **Low** contains a value of 15.
- **Medium** contains a value of 30.
- **High** contains a value of 40.

Vertical Line Removal

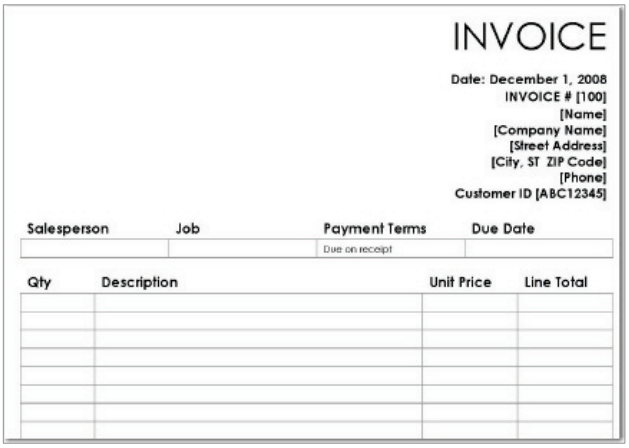
This setting detects vertical lines that will be taken out during the line removal process.

Minimum Length

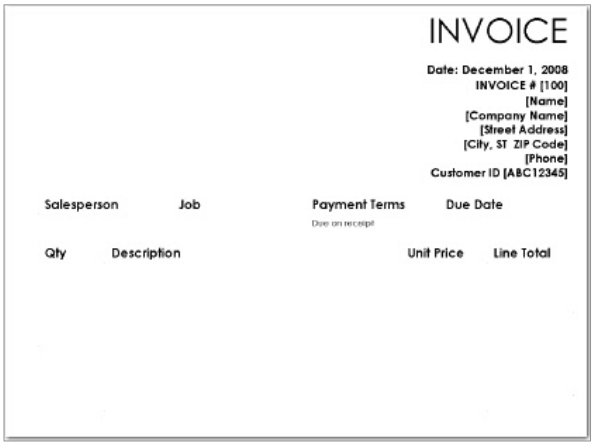
This setting defines the minimum length in millimeters (in whole or decimal numbers) that the filter will detect as a vertical line.

Maximum Gap

This setting defines the maximum amount of allowable white space in millimeters (in whole or decimal numbers) between two vertical line-like objects to be considered as one line.



Before Binary Line Removal



After Binary Line Removal

Binary Noise Removal

Noise can originate from carbon or dirt particles on scanners, fax machines, or copiers. Noise removal takes out extraneous specks from an image. If the image contains text, this filter may remove periods and dots from sentences and letters. To avoid removing essential parts of text characters, assign the Minimum Separation value to be greater than the distance between dots and the lower parts of letters. To apply cropping and noise removal to an image, perform the noise removal first for best results.

Maximum Height and Width

This setting defines the maximum height/width in millimeters (in whole or decimal numbers) of an object to be considered noise.

Maximum Area Percentage

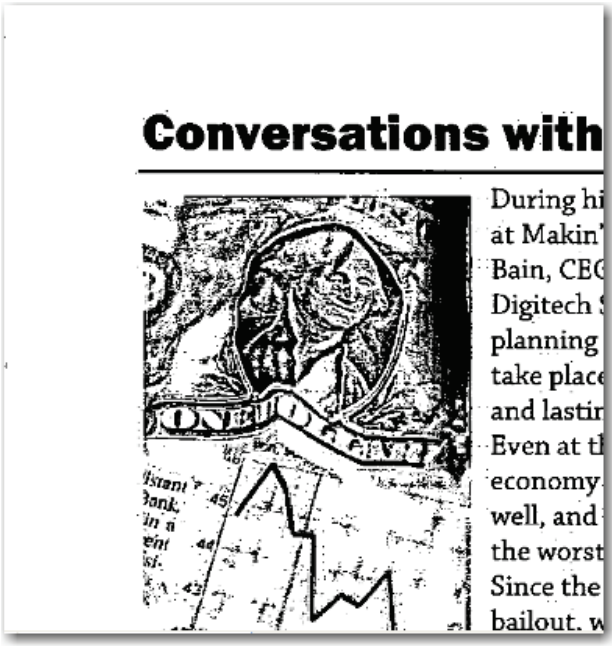
This value is defined by the specified height/width of an object to be removed as noise. The Maximum Area Percentage setting detects long narrow objects such as lines, decorative banners, and highlight areas that may appear both vertically and horizontally on a page.

For example, to remove colored banners with the dimensions 5" x 1" or 1" x 5", you can assign the Maximum Height and Maximum Width values to five inches. However, a 5" x 5" picture would also be detected as noise and removed. To avoid this problem, assign 20% so that only the banner area is detected as noise, regardless of its orientation.

Minimum Separation

This setting defines the minimum distance in millimeters (in whole or decimal numbers) that separates noisy areas from non-noisy areas of the page. A value of zero removes all noisy objects within your specified values in the Maximum Height, Maximum Width, and Area Percentage fields. Assigning a zero value may remove text elements, such as broken characters, periods, and dots above letters. Assigning a value greater than zero preserves noise-like objects near text characters and may improve OCR accuracy.

Before Binary Noise Removal

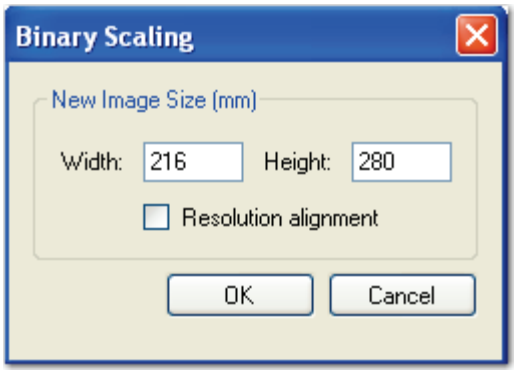


After Binary Noise Removal (and Binary Hole Removal)

Binary Scaling

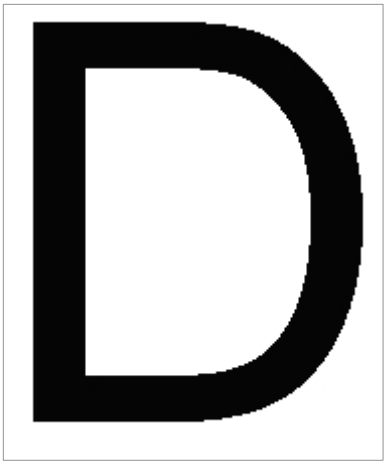
The Binary Scaling filter resizes an image while preserving the original aspect ratio. After you specify the width and height to apply to the image after scaling, its area is resized to fit within those boundaries while maintaining the aspect ratio. You can assign the resulting width and height in millimeters (in whole or decimal numbers) of the image after it is scaled. If the specified height or width value is larger than the area of the scaled image, the area is centered along the specified dimensions, and white margins are added to both sides.

The Resolution Alignment property adjusts the X (horizontal) and Y (vertical) resolutions of an image so they are equal. If the X and Y resolutions are not equal, the lower resolution is scaled up to match the higher resolution. When this setting is enabled, you cannot specify the width and height of the image.



Binary Scaling

> **Note**:
> Use of binary scaling can improve the recognition rate of barcode detection.

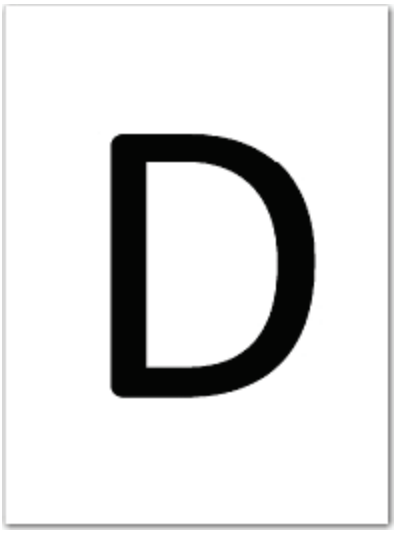


Before Binary Scaling

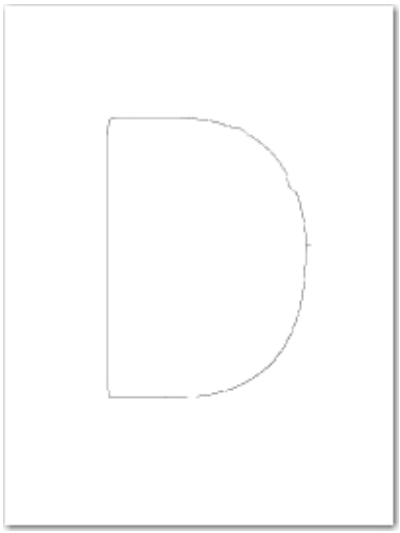
After 50% Binary Scaling

Binary Skeleton

The Binary Skeleton filter should be used with caution, since it can significantly distort the image. This filter can reduce the file size, and should only be used when performing certain types of OCR.



Before Binary Skeleton



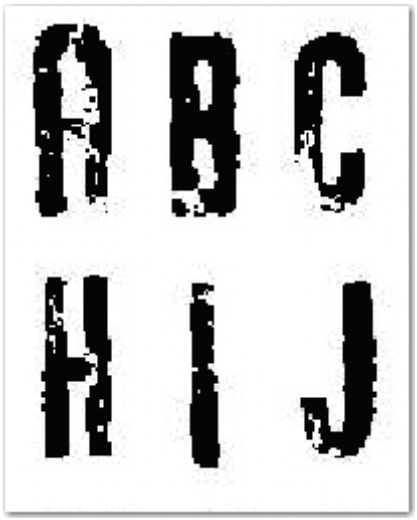
After Binary Skeleton (Zoomed 1x)

Binary Smoothing

The Binary Smoothing filter removes bumps that appear on text characters or graphics in an image. This filter looks for any pixel surrounded by five or six connected pixels of the opposite color, and then inverts that center pixel based on the filter's configuration. Smoothing improves legibility and can reduce file size without compromising detail.

- **Trim First** removes black noise pixels before white noise pixels. If this option is disabled, white noise pixels are removed before black noise pixels.
- **Corner Black** removes black noise pixels from the corners of objects in the image.
- **Corner White** removes white noise pixels from the corners of objects in the image.

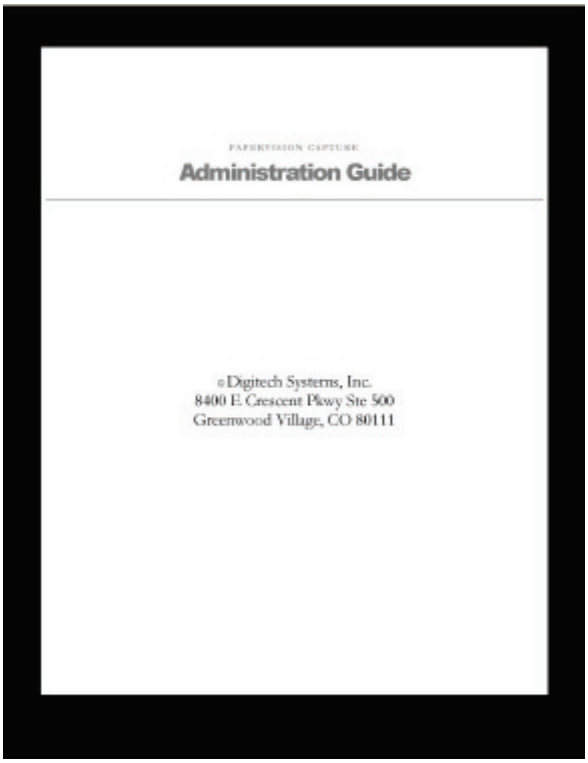
Before Binary Smoothing



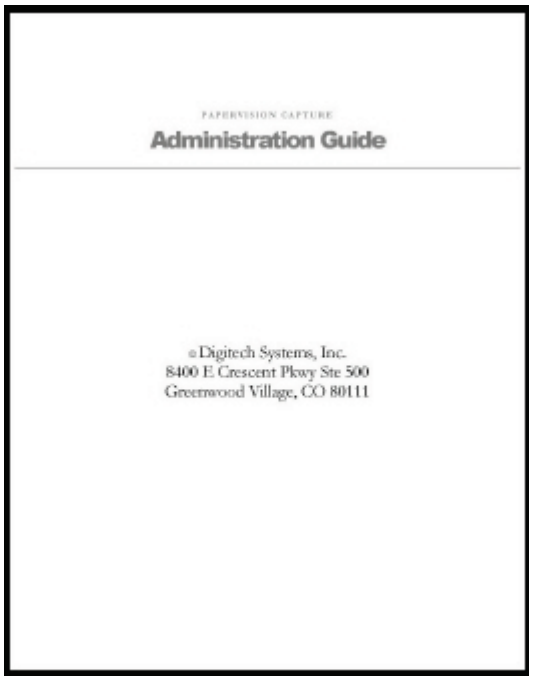
After Binary Smoothing

Black Overscan Removal

The Black Overscan Removal filter deletes the black overscan area that appears around an image produced by scanners with black borders. This filter reduces the image file size. To maximize results, apply the Deskew filter with a black fill color prior to applying the Black Overscan Removal filter.



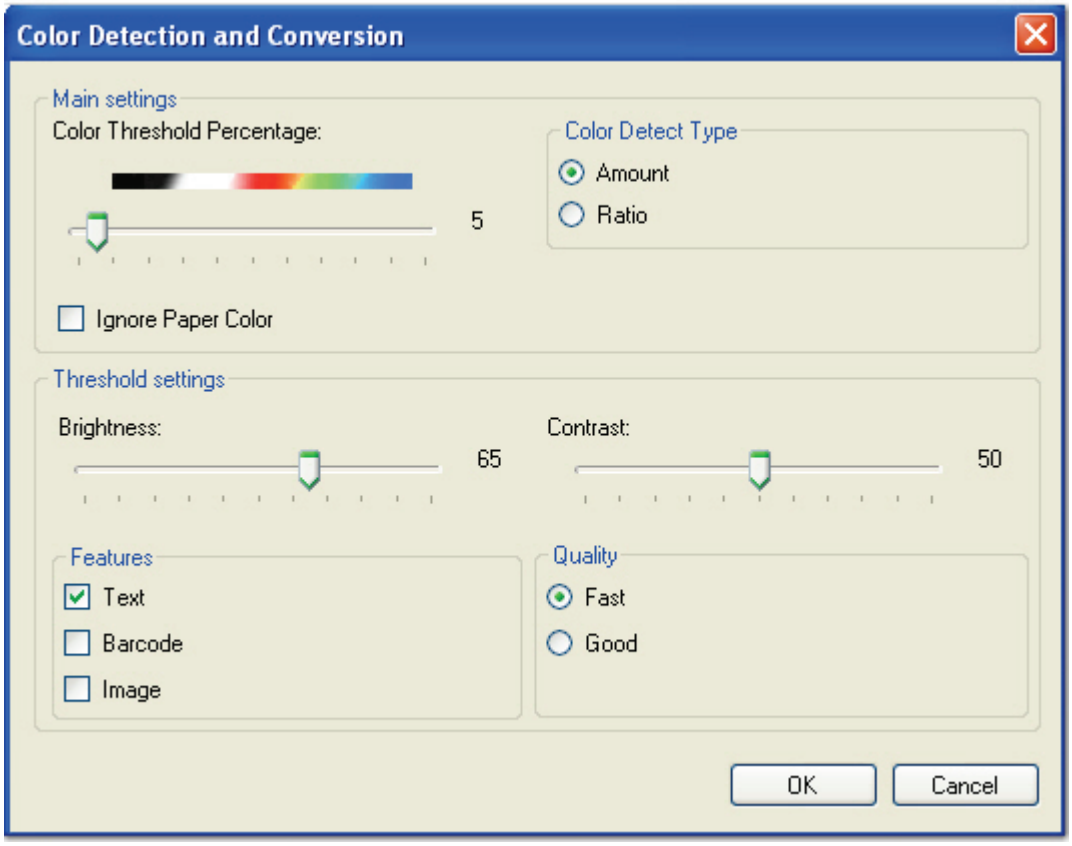
Before Black Overscan Removal



After Black Overscan Removal

Color Detection and Conversion

This filter detects the colorfulness of an image, and then returns either a binary or a color image based on your assigned threshold settings. If you enable the **Ignore Paper Color** setting, the paper's background changes to white. The filter then counts the number of white (and nearly-white) and black (and nearly-black) pixels and excludes them from the color count. The colorfulness of the image is then computed according to the selected **Color Detect Type**. If the resulting colorfulness value is less than your assigned threshold, the resulting image displays as binary (black and white).



Color Detection and Conversion

> **Note**:
>
> If the original image is more colorful than your specified threshold, the filter is not applied.

Color Threshold Percentage

This setting assigns the amount of color that an image must contain in order to be considered colorful. If you enable the **Ignore Paper Color** setting, the background color of the image changes to white before automatic color detection is performed.

Color Detect Type

The default setting, **Amount**, detects the number of color pixels in the image. The **Ratio** setting detects the ratio of color and black pixels in the image.

Brightness

Brightness defines a pixel's lightness value from black (darkest) to white (brightest). Move the slider to assign the amount of brightness to apply to binary images.

Contrast

Contrast is a measure of the rate of change of brightness in an image. A high-contrast image contains defined transitions from black to white. Move the slider to assign the amount of contrast for binary images.

Features

To preserve a specific feature in the binary image, you can select **Text, Barcode,** and/or **Image.**

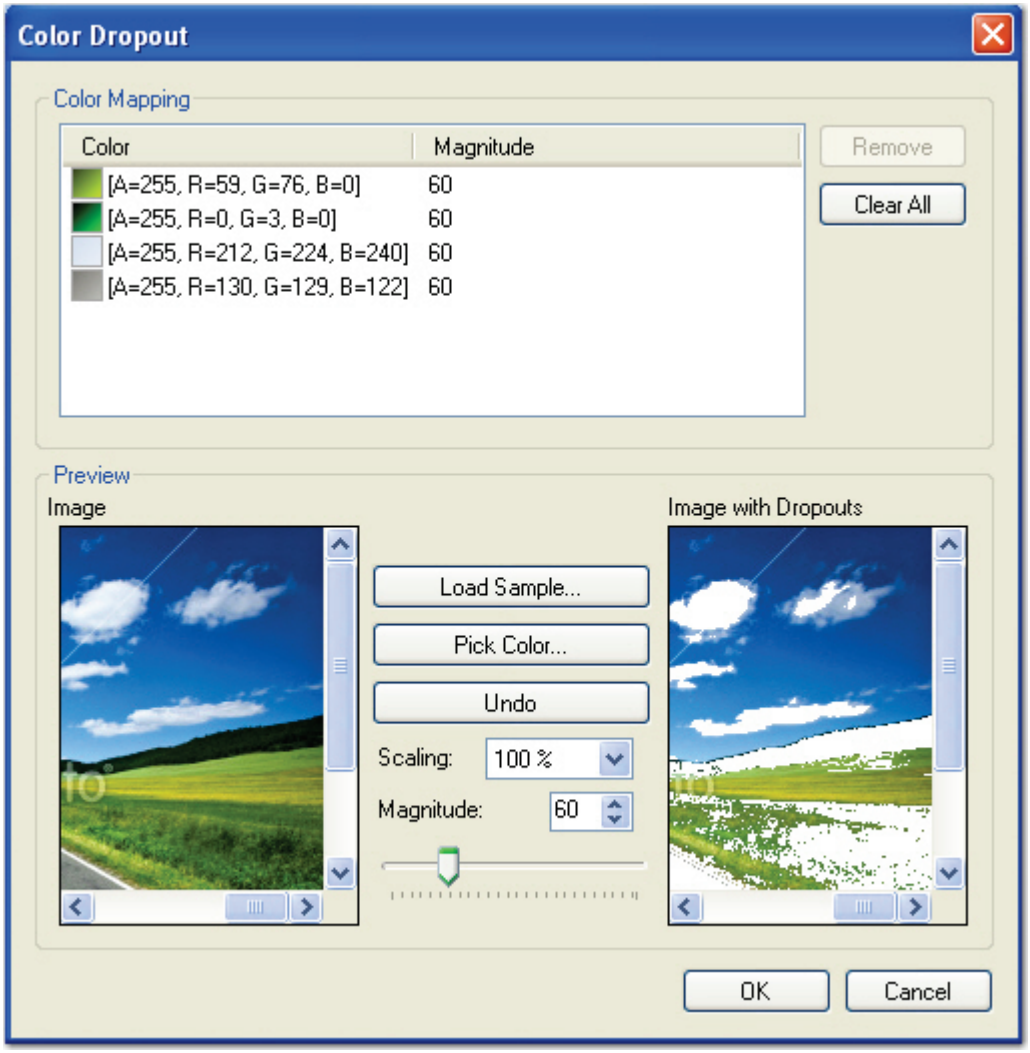
Quality

This setting specifies the quality and speed of the thresholding process.

- **Fast** causes thresholding to process quickly, and results in quality images.
- **Good** causes thresholding to process more slowly, but results in better quality images.

Color Dropout

The Color Dropout filter removes your specified colors from the image, and then displays the scanned image without your specified colors.



Color Dropout

To load a sample image and apply the Color Dropout filter:

1. Click **Load Sample Image.**
2. Browse to the directory.
3. Select the image.
4. Click **Open.**
5. To select the color to delete from the image, click the **Pick Color** button.

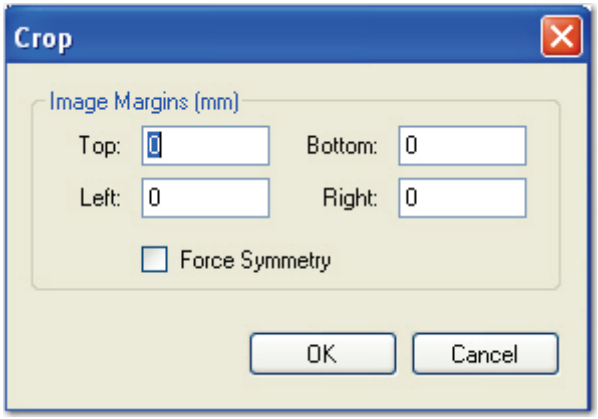
6. To undo the most recent color selections (since the last time you clicked **OK**), click the **Undo** button.

> **Note**:
>
> If the colors are not being restored, highlight the color in the **Color Mapping** section, and then click the **Remove** button on top.

7. To zoom in on the image, select a larger percentage in the **Scaling** drop-down list.
8. To apply a larger magnitude to the color dropout filter, enter a value between **1** and **255**.
   - Or, move the slider to see the effect on the image.
   - A larger magnitude value results in the removal of more adjoining colors to your selected color.
9. Click on the color to extract. The selected color appears in the **Color Mapping** list on top, along with its RGB color codes.
10. Click the **Remove** button to remove the color from the dropout list.
11. Select **Clear All** to remove all colors from the dropout list.

Crop

Cropping allows you to assign margins in millimeters (in whole or decimal numbers) to remove white space from the edge of the image. You can set different values for each margin.



Crop

Image Margins

Positive margin values represent the white space between the edge of the image and the black pixel closest to that edge. Negative margin values crop the specified amount from the black pixel closest to the edge towards the center of the image. Enter values in the **Top, Bottom, Left,** and **Right** fields to assign the margins.

Force Symmetry

This setting assigns the same values to opposite margins.

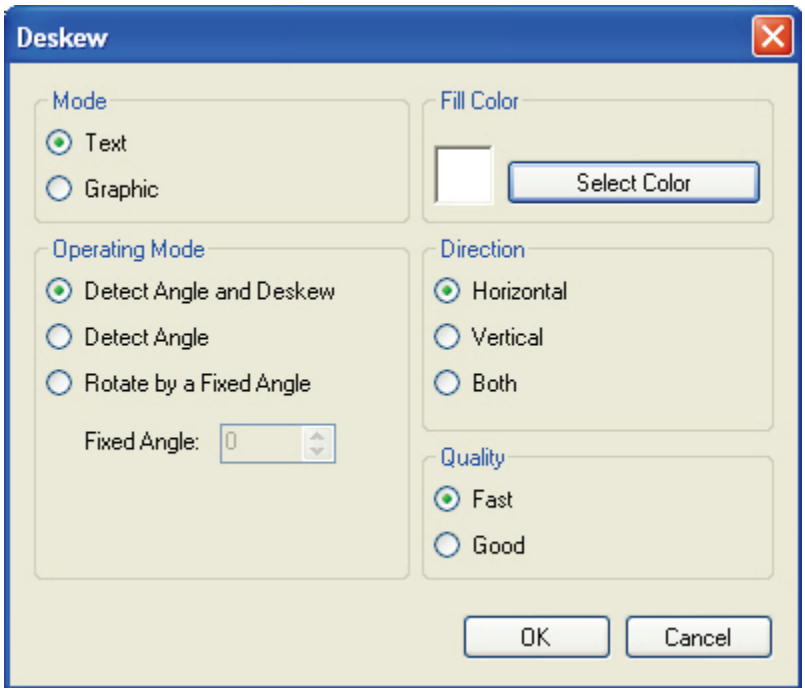
- Enter a value in the **Top** field to apply the same value to the top and bottom margins.
- Enter a value in the **Left** field to apply the same value to the left and right margins.

> **Note**:
> If you enter values for the **Bottom** or **Right** fields, they are ignored.

Deskew

Skewing can occur when the original document was fed into the scanner, fax machine, or photocopier. This filter examines the image and determines the skew angle, which is measured between the edge of the image and the horizontal or vertical axis. The filter straightens images that slant from their correct orientation.

You can rotate an image from -44.9 degrees to +44.9 degrees, in 0.1 degree increments, without detecting a skew angle. You can adjust the values most suitable for your documents.



Deskew

Mode

The Mode setting indicates whether text or graphics will be used to determine the skew angle.

- Select **Text** if pages primarily contain text with some tables and lines.
- Select **Graphics** if pages contain large blocks of black areas.

Operating Mode

- The default setting, **Detect Angle and Deskew**, automatically examines the images and determines the skew angles.
- **Rotate by a Fixed Angle** rotates the image by your specified fixed angle.
- **Detect Angle** deskews the images by a fixed number of degrees.

Fill Color

You can assign a fill color of black or white (default), which can match the color in the overscan area of the image. If the image contains a border, you can assign the fill color to match the border after the image is deskewed.

Direction

This setting indicates the image's skew angle measurement direction.

- Select **Horizontal** if only horizontal text exists in the documents.
- Select **Vertical** if only vertical text exists in the documents.
- Select **Both** if either text orientation may exist.

Quality

This setting specifies the quality and speed of the deskew process.

- **Fast** causes deskewing to process quickly, and results in quality images.
- **Good** causes the deskewing to process more slowly, but results in better quality images.

Before Deskew

After Deskew (with Binary Border Removal)

Image Fit

This filter is intended to crop images before they are processed through the Open Text Full-Text engine. The minimum and maximum width and height dimensions that can be specified are 16 x 16 to 8400 x 8400 pixels. If the image size is less than 16 x 16 pixels, white space will be added to the image from the bottom and right corners until the minimum size (16 x 16 pixels) is reached. If the image size is greater than 8400 x 8400 pixels, the image is cropped from the bottom and right corners until the maximum size is reached.

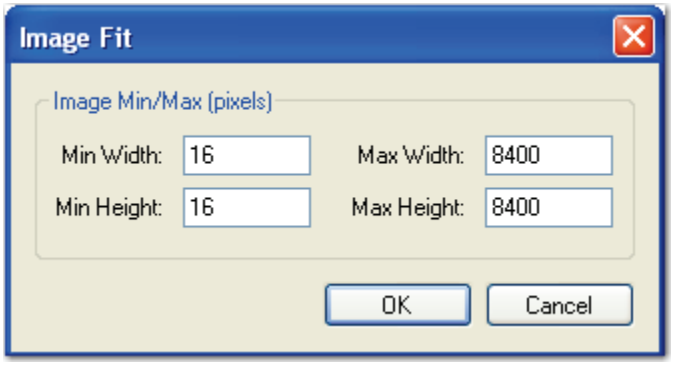


Image Fit

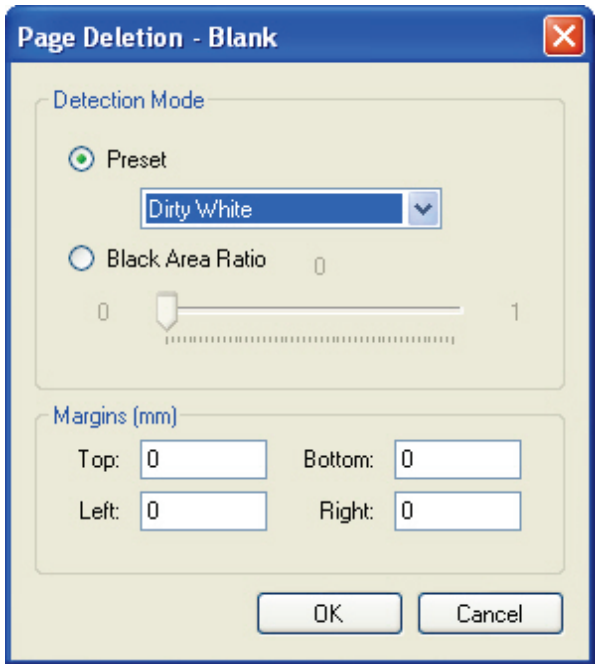
Page Deletion - Always

This filter removes the entire page from the batch.

Page Deletion - Blank

To detect blank pages in a document, one of two methods can be applied. If you apply the **Preset** method, select from the following options:

- **Dirty White**, the default setting, considers pages blank when they contain some noise.
- **One Line OK** considers pages blank when they contain one specified line of text.
- **Pristine White** considers pages blank when they contain no noise.
- **Two Lines** considers pages blank when they contain two specified lines of text.
- **Very Dirty White** considers pages blank when they contain a lot of noise.

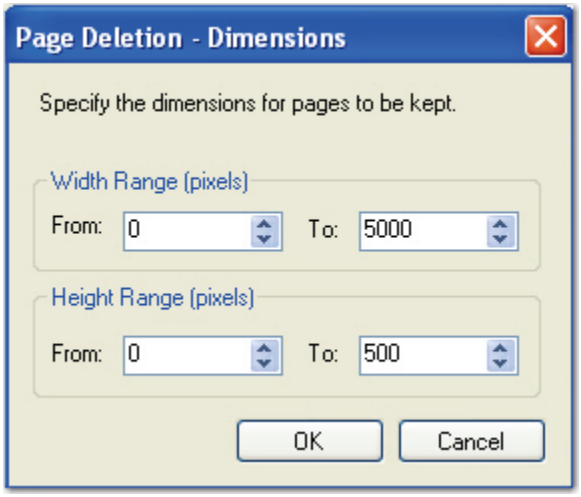


Page Deletion – Blank

If you select **Black Area Ratio,** move the slider to assign the ratio that determines when a page is blank. The ratio is calculated by dividing black pixels by the number of All Region Pixels. Enter margins in millimeters (in whole or decimal numbers) to exclude when this setting determines whether a page is blank. This filter then deletes pages detected as blank according to your specified parameters.

Page Deletion - Dimensions

This filter allows you to specify the dimensions (in pixels) of pages that will remain in the batch. Enter the width and height ranges in the **From** and **To** fields, and images with dimensions that fall outside your specified ranges will be deleted from the batch.

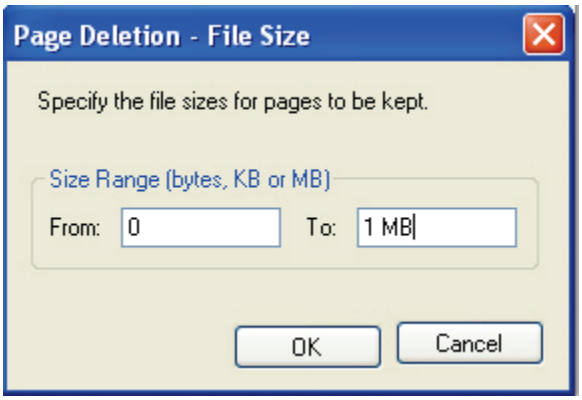


Page Deletion - Dimensions

Page Deletion – File Size

This filter allows you to specify the file size for pages that will remain in the batch. Enter the size range, including the numeric value and file size unit, in the **From** and **To** fields, and images falling outside your specified size range will be deleted from the batch.

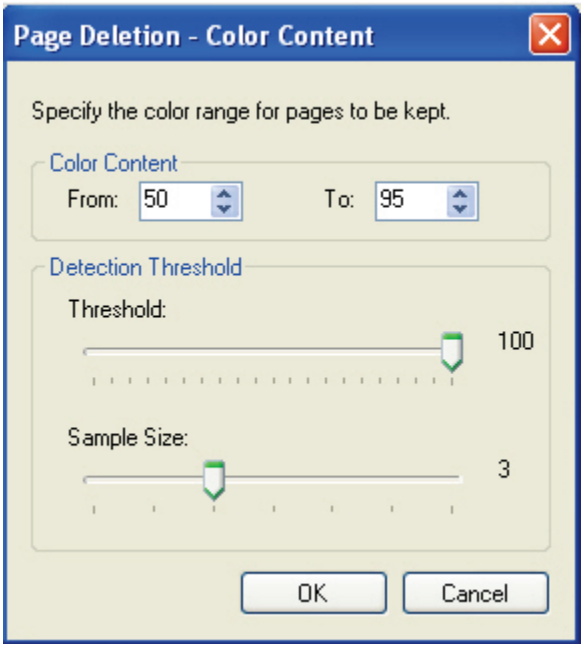
> **Note**:
>
> If you do not enter a specific file size unit (KB, MB, etc) after the numeric value, the unit defaults to bytes. Therefore, for kilobytes and megabytes, you must enter "KB" and "MB" after the numeric values.



Page Deletion – File Size

Page Deletion - Color Content

This filter allows you to assign color threshold settings that specify whether to delete color pages or non-colorful pages.

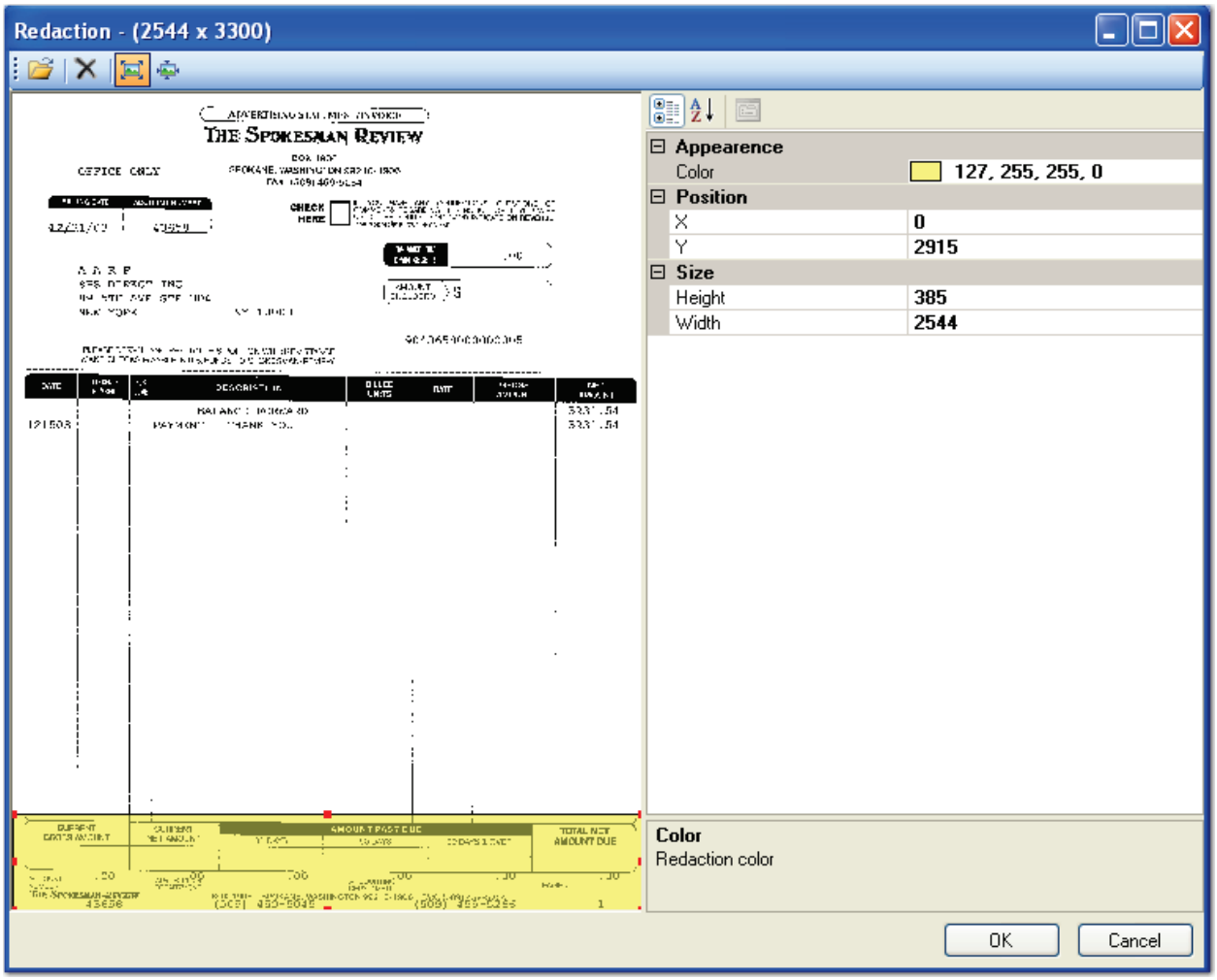


Page Deletion - Color Content

- The **Color Content** ranges between 1 and 100. Pages detected outside the specified range will be deleted.
- The **Threshold** value ranges between 1 and 100.
- The **Sample Size** value ranges between 1 and 7.

Redaction

The Redaction filter allows you to cover confidential or sensitive data on images. To ensure redactions consistently cover the same area on every image, it is recommended to test images with similar sizes that will be used in production. For your reference, the size (in pixels) of each imported image appears in the title bar.



Redaction

To import an image:

1. Click the **Import Image** icon in the toolbar.
2. In the **Open** dialog box, locate the image.
3. Click **Open**.

To adjust the image view:

- To fit the image exactly within the window, click the **Best Fit**  icon.
- To view the image in its actual size, click the **Actual Size**  icon.

Drawing Redactions

After you have imported a sample image into the Redaction window, the cursor is automatically equipped with the Redaction tool.

To draw a redaction:

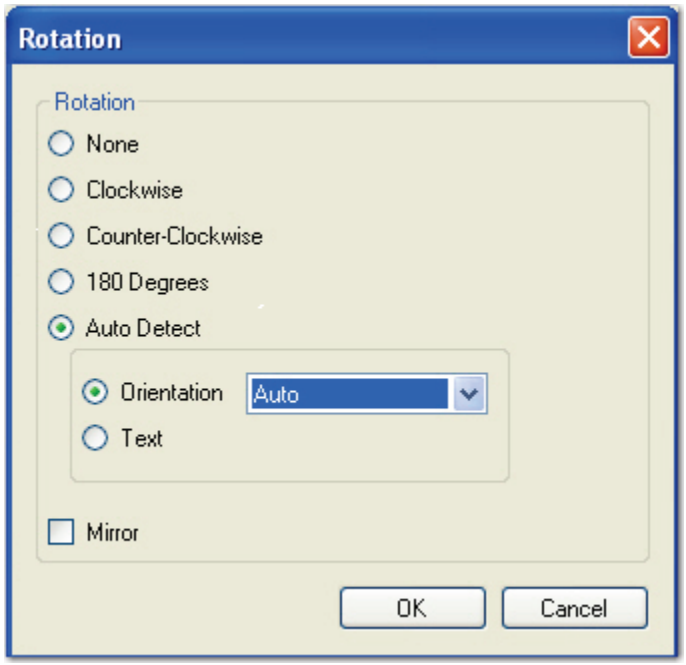
1. Drag the cursor around the area on the image. By default, a transparent rectangle appears on the image.
2. Once the redaction is drawn, the redaction properties appear in the properties grid on the right. You can edit the color, position, and size of the redaction.
   - **Color**: From the drop-down list, you can select the background color of the redaction.
   - **Position**: The X coordinate indicates the position of the redaction's upper-left corner relative to the container's left edge. The Y coordinate indicates the position of the redaction's upper-left corner relative to the container's top edge.
   - **Size**: The width and height of the redaction are specified in pixels.
3. After making necessary adjustments, click **OK** to save the redaction properties.

To delete a redaction:

1. Select the redaction.
2. Click the **Delete**  icon, or press the **Delete** key.

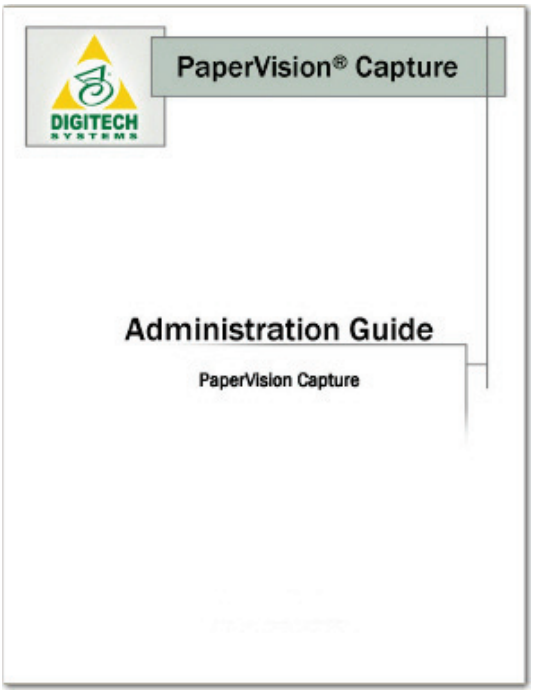
Rotation

The Rotation filter automatically rotates scanned images by your specified direction, fixed amount of degrees, or detected text orientation. The Text setting detects the image's text orientation using the Open Text Full-Text OCR engine, and then automatically rotates the image.



Rotation

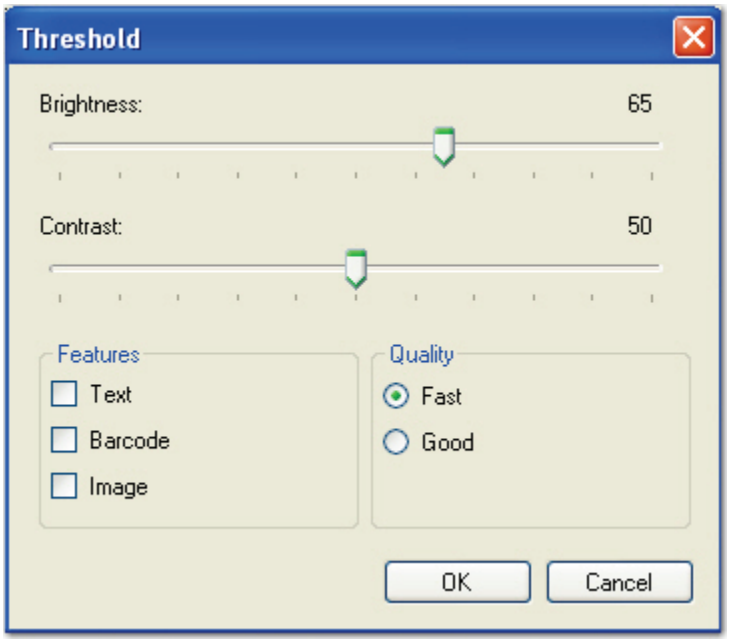
Before Rotation



After 180-Degree Rotation

Threshold

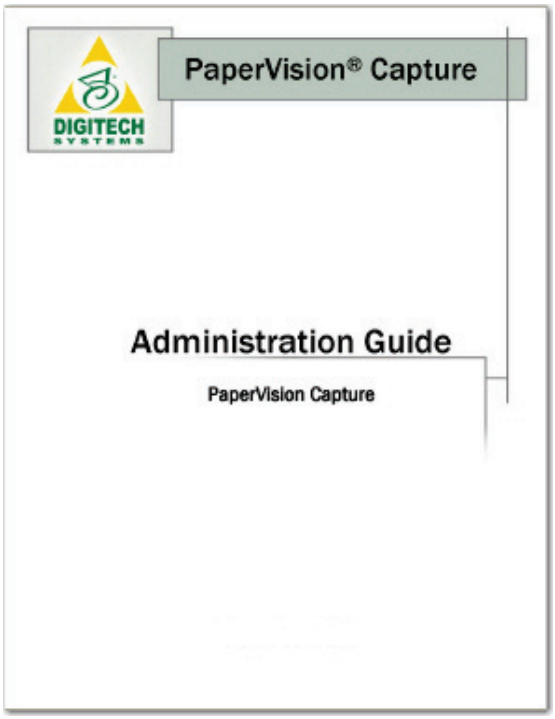
The Threshold filter converts a 24-bit color image to a binary image. The pixels in a color image that are darker than the specified Brightness and Threshold properties are converted to black. The pixels that are lighter than the threshold are converted to white.



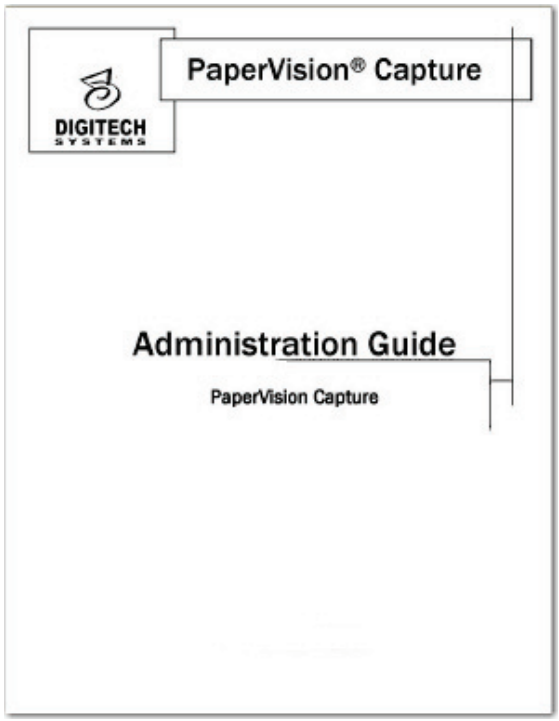
Threshold

To assign Threshold settings:

1. Move the **Brightness** slider to assign the point at which color pixels are converted to white rather than black.
2. Move the **Contrast** slider to assign the contrast of the resulting binary image.
3. To preserve a specific feature from the color image in the resulting binary image, select **Text, Barcode,** and/or **Image.**
4. Select **Fast** or **Good** thresholding quality.
   - **Fast** causes thresholding to process quickly, and results in quality images.
   - **Good** causes the thresholding to process more slowly, but results in better quality images.
5. Click **OK.**



Before Threshold



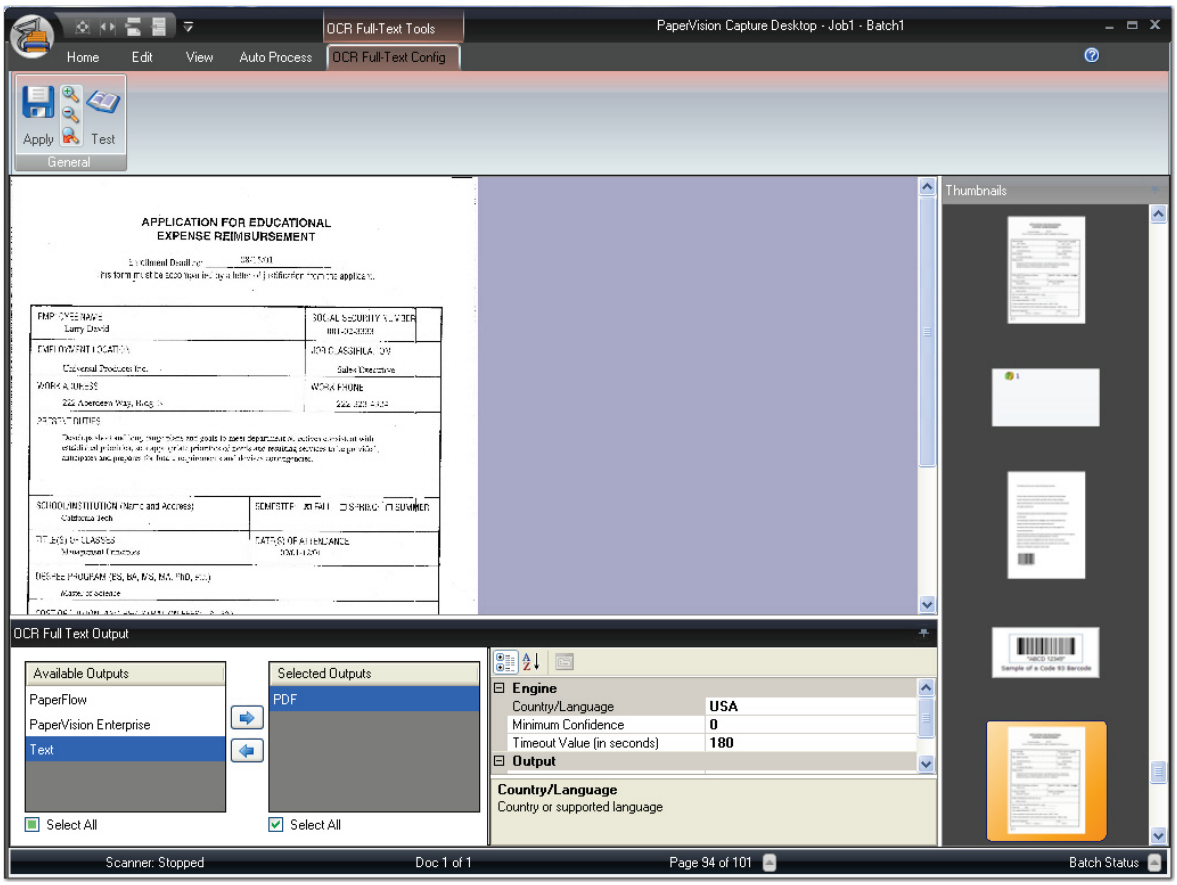
After Threshold

Chapter 7 - Full-Text OCR

In PaperVision Capture Desktop, full-text OCR processing is performed by the Open Text® engine that recognizes machine-printed text. In the **Auto Process** page, you can configure an automated process that reads pages of text and converts recognized results to one or multiple file types. Each output type contains unique settings that you can configure to support your full-text OCR requirements. During full-text processing, documents can be converted to several PDF versions, including those compatible with PDF-A, 1.4, 1.5, 1.6, and 1.7. The engine also converts documents to PaperVision Enterprise, PaperFlow, and text (.txt) output file types.

Note:

Handwritten text will not be recognized. Additionally, new line characters will be removed during OCR processing.



Full-Text OCR Configuration

Maximum Supported Image Sizes

The maximum supported image dimensions that can be processed through the Open Text engine vary with resolution. The approximate maximum width is approximately 32,000 pixels, and the maximum height is approximately 24,000 pixels. For example, the maximum supported image dimensions at 300 dpi are approximately 106 inches x 80 inches. Images that are processed through the Open Text OCR engine must contain matching horizontal and vertical resolutions.

DISCLAIMER – PLEASE READ:

These dimensions are provided only as estimates to identify size limits processing images in PaperVision Capture Desktop. Variations in technical environments may cause maximum image sizes to fluctuate across systems.

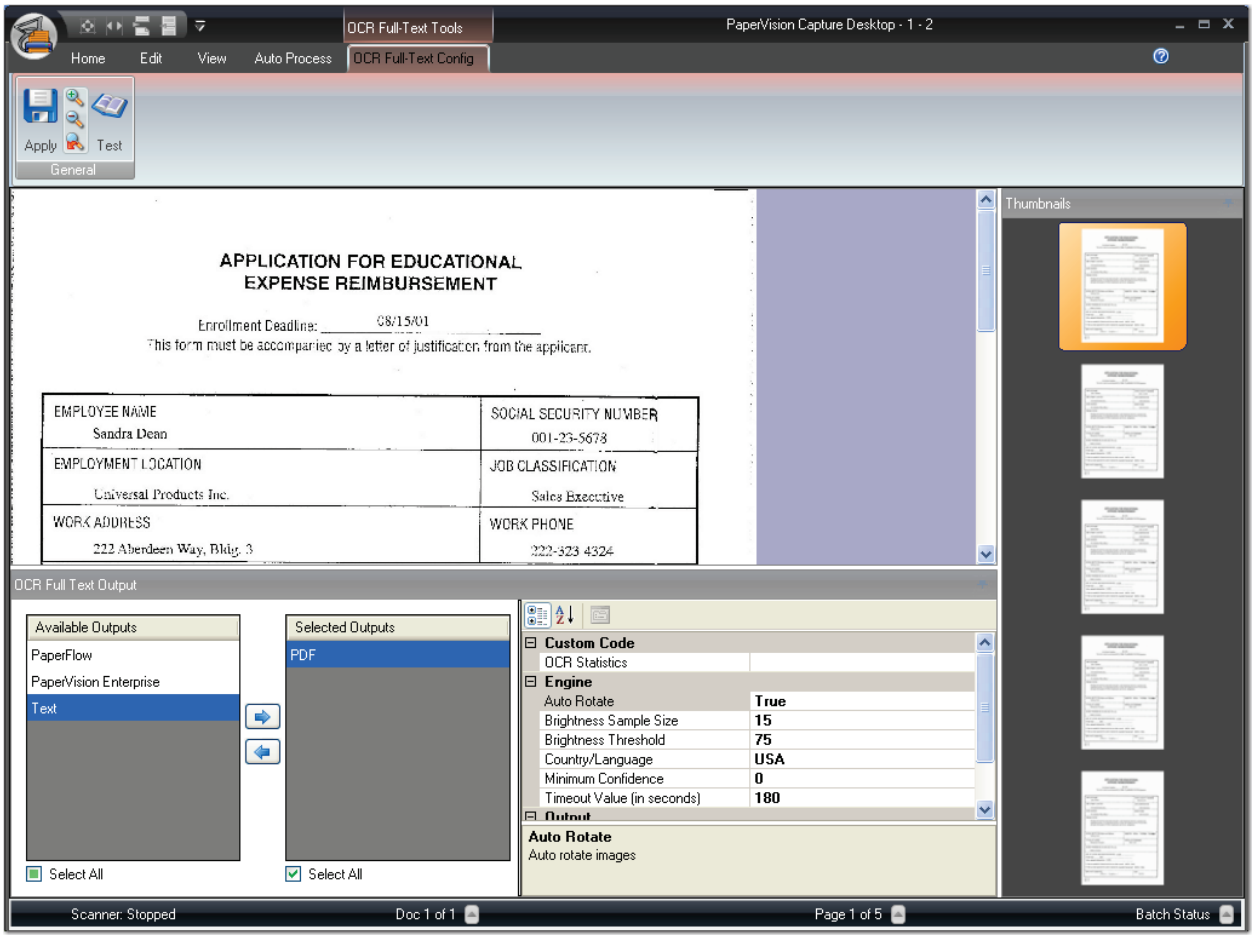
When you configure full-text OCR outputs and their associated properties, you can preview the full-text OCR results before you process the batch of documents. Thumbnail previews display the document's images and allow you to navigate through the document and perform basic operations including the cut/paste, copy/paste, and delete operations.

Configuring and Processing Full-Text OCR

Within the Auto Process page, you can configure an automated full-text OCR process to read pages of text and convert recognized results to various file formats.

To configure full-text OCR:

1. In the Auto Process page, click the **Config** icon in the **Automated Processing Tasks** toolbar group. The **Full-Text OCR Configuration** context-sensitive page appears.



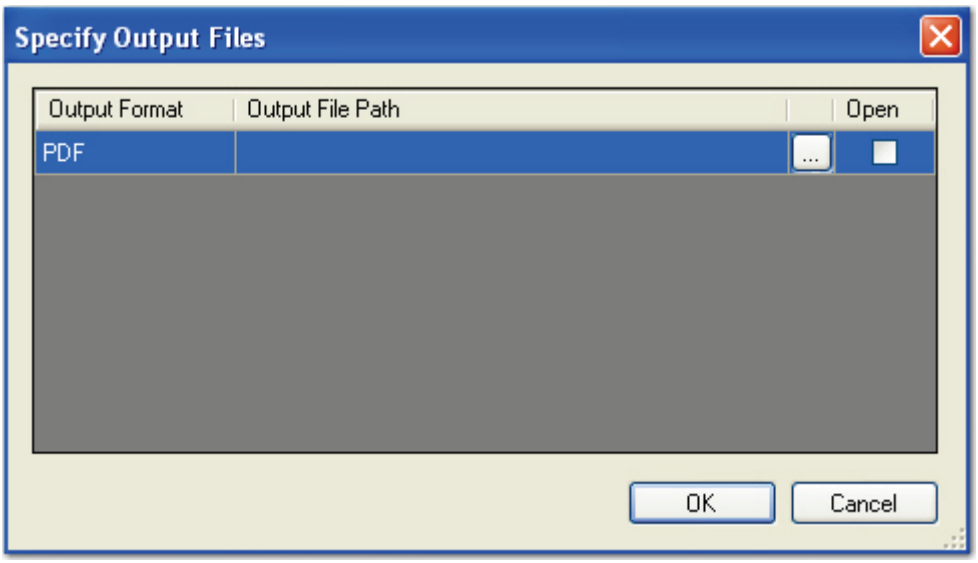
Full-Text OCR Configuration

2. In the **OCR Full Text Output** window, double-click the output in the **Available Outputs** list. Your selection populates the **Selected Outputs** list and the properties grid.

> **Note**:
>
> You can select more than one output.

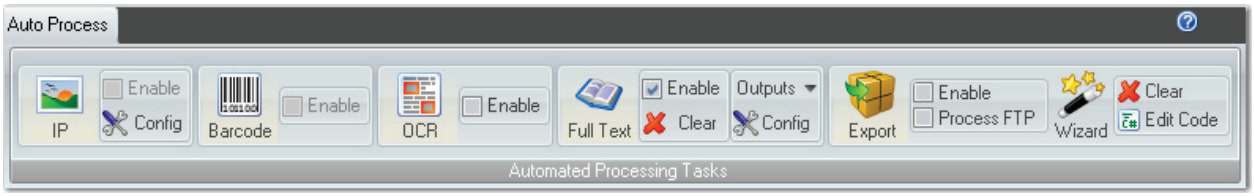
3. To remove one or more selected outputs, highlight the output(s) in the **Selected Outputs** list, and then click the left arrow.

4. To test the full-text output for your selections, click the **Test** icon. The **Specify Output Files** dialog box appears.



Specify Output Files

5. Enter the output file path where the full-text OCR results will reside.
   - Or, click the ellipsis button to browse to the output file location.
   - If you browsed to the output file location, enter the file name in the **Save As** dialog box, and then click **Save**.
6. If you want to view the results once they are tested, select the **Open** check box.
7. Click **OK**. The full-text OCR engine will process the results. If you opted to open the resulting output file, it will open in its respective application or editor.
8. If the resulting file is not acceptable, adjust the OCR page properties and/or the converter's properties, and run the test again.
9. Click the **Apply** icon to save your full-text OCR converter settings. The **Auto Process** page appears once again.
10. In the Auto Process page, select the **Enable** check box next the Full Text OCR automated task.



Automated Processing Tasks (Full Text)

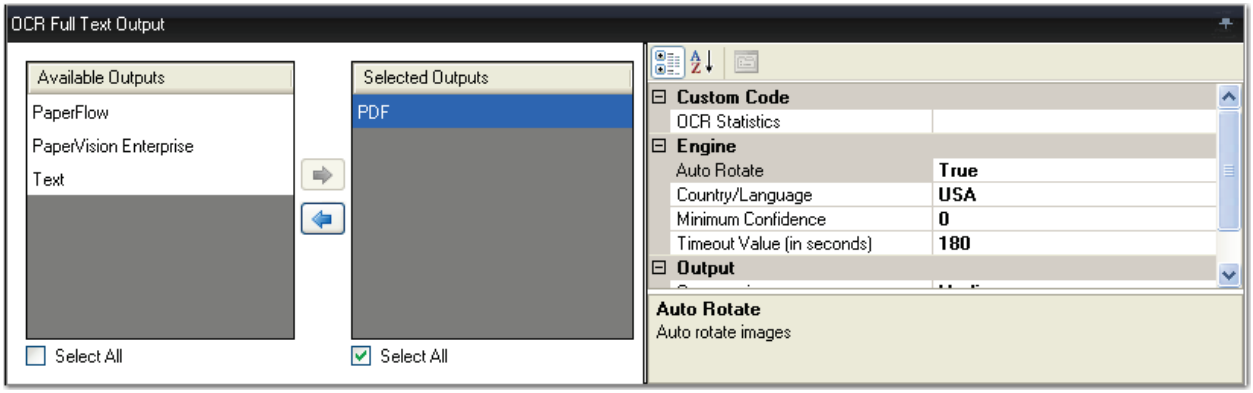
11. To process the batch (with full-text OCR output now enabled), click the **Start** icon. A progress screen appears, and displays the progress of full-text OCR output file. If any errors occur during processing, they will appear in this screen.
12. Click the **Exit** button to close the progress screen.

Testing Full-Text OCR Output

Prior to starting automated processing, you can verify that the current page's text can be read successfully. You can opt to open the output file for the selected output type.

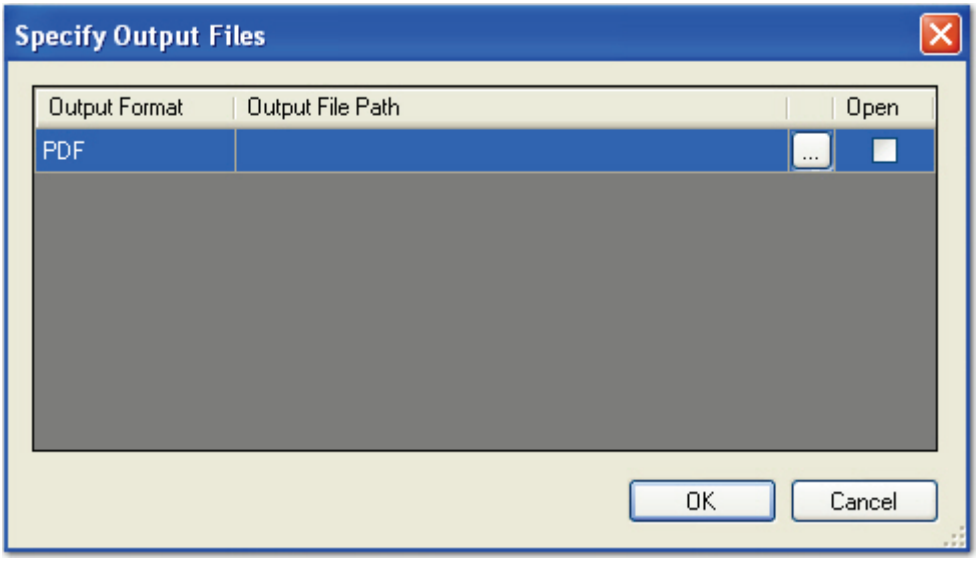
To test full-text OCR for the current page:

1. Navigate to the appropriate page.
2. In the **Full-Text OCR Config** page, select the output from the **Available Outputs** list. The selected output's properties populate the **Full-Text OCR Output** grid.



Full-Text OCR Output

3. Click the **Test** icon. The **Specify Output Files** dialog box appears.



Specify Output Files

4. Enter the output file path where the full-text OCR results will reside. Proceed to step 5.
5. Or, click the ellipsis button to browse to the location. Proceed to the next step.
6. If you browsed to the file location, enter the file name in the **Save As** dialog box, and then click **Save**.
7. To view the results, select the **Open** check box.
8. Click **OK**. The full-text OCR engine will process the results. If you opted to open the resulting output file, it will open in its respective application or editor.
9. If the resulting file is not acceptable, adjust the output properties, and run the test again.
10. If the resulting file is acceptable, click the **Apply** icon to save the settings.

Zooming Operations

If you need to adjust the view of the page, you can use the zoom operations in the toolbar.

- To zoom in on an area of the image, click the **Zoom In** icon.
- To zoom out of the current view of the image, click the **Zoom Out** icon.
- To reset the image to its original view, click the **Zoom Reset** icon.

Supported Output File Types

PaperVision Capture Desktop supports the following full-text OCR output file types:

- **PDF**: The PDF output produces a searchable PDF (.pdf) file compatible with your specified PDF version.
- **PaperFlow**: The PaperFlow output is a text-based full-text output file that you can subsequently import into PaperFlow. You can select the supported country/country group/language/language group/character set. You can also assign the Minimum Confidence, Timeout Value, and Rejection Symbol.
- **PaperVision Enterprise**: The PaperVision Enterprise output is a text-based full-text output file that you can subsequently import into PaperVision Enterprise. You can select the supported country/country group/language/language group/character set. You can also assign the Minimum Confidence, Timeout Value, and Rejection Symbol.
- **Text**: The Text output produces a text (.txt) file.

Auto Rotate

By default, this property is set to **True**, and the Open Text Full-Text OCR engine may automatically rotate some images in order to recognize text. If you do not want the Open Text Full-Text OCR engine to automatically rotate images prior to text recognition, set this property to **False**.

> **Note**:
>
> Since the engine may automatically rotate some images in order to recognize text, the resulting output images may also be rotated.

Brightness Sample Size

This value (indicating both width and height) specifies the rectangle size used to calculate the brightness threshold. You can specify a value between 1 and 32, and the default value is 15.

> **Note**:
>
> Smaller brightness sample sizes may cause the OCR engine to recognize extraneous noise on the image.

Brightness Threshold

You can assign a brightness threshold value (between 0 and 255) for the image. The default value is 75.

Country/Language

When you select from the **Country/Language** property, your selection may reflect not only a country or language, but country groups (e.g., Western Europe), language groups (e.g., Latin), and character sets (e.g., OCR). Each country corresponds to one or more languages, and countries are automatically expanded into language sets (e.g., German corresponds to the German language; Switzerland corresponds to the German, French, Italian, and Rhaeto-Romantic languages).

Specific languages are also available for selection under the Country/Language property (e.g., English, German, Dutch, Italian, etc.). It is recommended to narrow your selection as much as possible since OCR recognition may become slower with a greater number of selected countries or languages. It is also recommended to select a country rather than a language or country group (e.g., Western Europe, South America, Scandinavia) since the recognition of certain types of addresses and money transfer forms may improve.

Note:

You cannot select the OCR character set individually; it must be selected with another language, language group, country, or country group.

Language Groups

If you select a language group, it is recommended to select only one, since they encompass multiple languages, countries, and code pages:

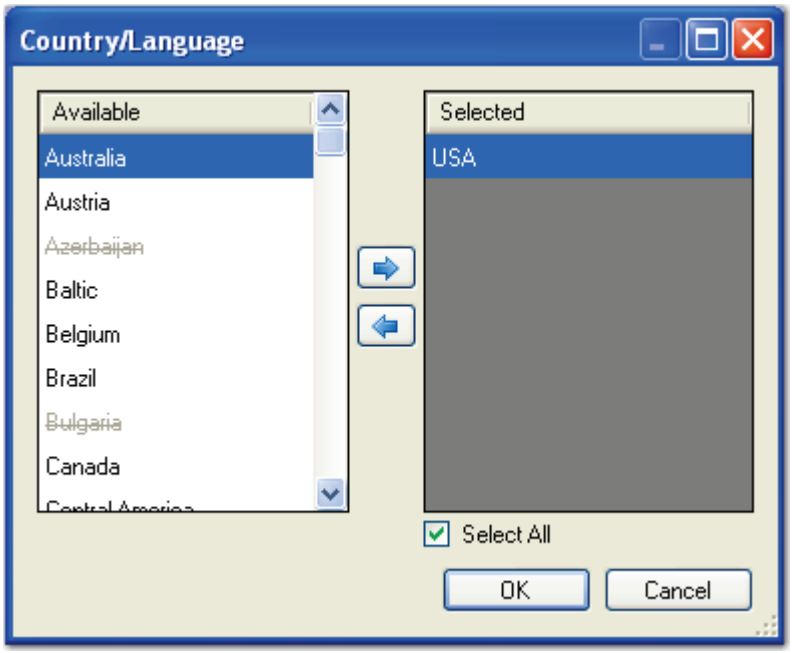
1. Cyrillic (Code page 1251)
2. Greek (Code page 1253)
3. Latin (Code page 1250, 1252, 1254, and 1257; i.e., Central Europe, Western Europe, Turkey, and Baltic)
4. Azerbaijanian

Note:

For language groups, recognition results are always represented by Unicode characters. The English character set (A-Z, a-z) is implicitly available with all country-language selections, even Greek or Cyrillic.

To select a country or language for full-text OCR output:

1. After selecting an output type, click the ellipsis button to the right of the Country/Language property. The **Country/Language** dialog box appears.



Country/Language

> **Note**:
>
> If a country or language appears crossed out, it does not belong to the same code page as the selected country or language. Therefore, countries or languages containing strikethroughs cannot be added to the **Selected** list.

2. Highlight one or more countries/languages from the **Available** list, and then click the right arrow.
3. To remove one or more selections from the **Selected** list, highlight the countries/languages, and then click the left arrow.
4. When finished with your selections, click **OK**.

> **Note**:
>
> For a list of the supported countries, languages, country groups, and character sets, see **Appendix B.**

Minimum Confidence

The confidence level reflects the reliability of the OCR recognition results. Values range from zero (the default setting), the lowest confidence level, to 255, the highest confidence level indicating the most reliable recognition results. Characters with lower confidence levels than your specified value will display as the rejection symbol, which is the tilde (~) character by default.

Timeout Value (sec)

This property allows you to define the maximum amount of time that the Open Text OCR engine processes a single image before it fails. By default, this property is set to 180 seconds (3 minutes). You can assign a timeout between one second and 3,600 seconds (1 hour).

> **Note**:
>
> Raising the timeout setting may increase the amount of time to process all images.

Compression

You can set the level of compression applied to PDF outputs. The higher the compression, the smaller the output file size. The default level of compression is medium. You can select from the following compression levels:

- None (no compression will be applied)
- Low (low level of compression is applied)
- Medium (medium level of compression is applied)
- High (highest level of compression is applied)

PDF Version

You can select the compatible PDF version for PDF output files. The following versions are supported by the full-text OCR engine:

- PDF/A: Format for long-term archiving of electronic documents - with Level B compliance in Part 1 (1b)
- PDF 1.4: Acrobat 5.0
- PDF 1.5: Acrobat 6.0
- PDF 1.6: Acrobat 7.0
- PDF 1.7: Acrobat 8 and 9

Rejection Symbol

This property represents rejected characters in output documents. A rejected character is not recognized by the active OCR recognition engine configuration. The default value is the Tilde character (~). Only a single character can be entered in this field.

> **Tip**:
>
> To prevent unrecognized characters from appearing in output documents, leave this field blank.

Chapter 8 - Custom Code Configuration

PaperVision Capture Desktop's custom code engine enables you to write C# or VB.NET code that users can execute at any time. You can write custom code to validate/manipulate/export data, access a database, interface with an external line-of-business application, or implement other desired functionality. Additionally, Digitech Systems provides a .NET Application Programming Interface (API) that you can use for read/write access to batch metadata, documents, images, output OCR data, and index values. You can insert your code in two areas: (1) events that are automatically fired by the application and (2) exports that move captured images and data out of the application. For more information on configuring code in these two areas, see **Chapter 2 - Home Page (Job Configuration section)** and **Chapter 9 - Exports**.

Custom Code Event Arguments

Each custom code event exposes an argument parameter that is specific to the given event type. Within your code, you can access these arguments to read event-specific data and to configure settings. For example, your code can change a property that determines the action that is triggered in PaperVision Capture Desktop after the event. The event-specific arguments are listed below.

> **Note**:
>
> The following classes are derived from the .NET System.Data.DataSet class and support all DataSet properties and functions. Additionally, DataSets are mapped to index values in the Indexing window.

Add Page Event - CCustomCodeNewImageEventArgs

The Add Page event uses the CCustomCodeNewImageEventArgs class to pass every scanned image to custom code. Use of this argument is illustrated in the InspectBeforeAddPage sample script:

```
CCustomCodeNewImageEventArgs args = base.Parameter as
CCustomCodeNewImageEventArgs;
```

1. `Image.Attributes` (hashtable containing the following image attributes):
    a. `PageSide: string` (indicates the side of the page as "Front" or "Back")
    b. `DriverName: string` (indicates the name of the scanner driver)
2. `PageTags: TagInfo[]`

    This property can be used to specify one or more page tags to be added after the page has been appended to the batch. Tags added to a break page (based on job configuration settings to delete break pages) will be ignored.

Barcode Detected Event - BarcodeReadEventArgs

The Barcode Detected event uses the BarcodeReadEventArgs class to pass every barcode's data (from each barcode zone) to the custom code. This event is triggered each time a barcode is successfully detected during scanning (multiple barcodes can be detected per page).

The following properties are located within the custom code:

1. `BarcodeItem Properties`

   These properties contain all barcode data, including barcode value, location, size, orientation, and barcode type.

2. `PageTags: TagInfo[]`

   This property can be used to specify one or more page tags to be added after the page has been appended to the batch. Tags added to a break page (based on job configuration settings to delete break pages) will be ignored.

Custom Code Execution Event - ManualCustomCodeEventArgs

The Custom Code Execution event uses the ManualCustomCodeEventArgs class to pass index values to the manual custom code event. This event is triggered when the user triggers the Custom Code operation in the Edit page.

```
ManualCustomCodeEventArgs args = base.Parameter as
ManualCustomCodeEventArgs;
```

Index Populated Event - IndexPopulateEventArgs

The Index Populated event uses the IndexPopulateEventArgs class to pass index values to the custom code. This event is triggered when an index value is populated.

```
IndexPopulateEventArgs args = base.Parameter as IndexPopulateEventArgs
```

Index Validate Event - IndexValidateEventArgs

The Index Validate event uses the IndexValidateEventArgs class to pass index values to the custom code. This event is triggered once the user proceeds or tabs to the next index field in the Indexing window.

```
IndexValidateEventArgs args = base.Parameter as IndexValidateEventArgs;
```

OCR Statistics Event - OCRFullTextPageProcessedEventArgs

The OCR Statistics custom code event uses the OCRFullTextPageProcessedEventArgs class to pass Open Text full-text data from each page (per selected output format) to the custom code. For each output type, this event is triggered once a page has been converted to PDF, PaperVision Enterprise, PaperFlow, or Text full-text output.

The following properties are located within the custom code:

1. `DocumentId: string`
2. `PageId: Guid`
3. `PageIndex: int32`
4. `OCRWords: int32`

   The OCRWords property contains the following variables:

```
internal OCRCharacter[] characters = new OCRCharacter[] { };
        internal Int32 line = 0;
        internal System.Drawing.Point location = new
System.Drawing.Point();
        internal System.Drawing.Size size = new
System.Drawing.Size();
```

The `OCRCharacter` variable contains the following properties:

```
public System.Drawing.Point Location
{
    get
    {
        return location;
    }
}
public System.Drawing.Size Size
{
    get
    {
        return size;
    }
}
public Byte Confidence
{
    get
    {
        return confidence;
    }
}
public Char Code
{
    get
    {
        return code;
    }

}
public bool Rejected
{
    get
    {
        return rejected;
    }
}
public Char[] Alternatives
{
    get
    {
        return alternatives;
    }
}
```

5. RecognitionTime: int32 (milliseconds)
6. AdditionalValues: Hashtable
7. ConverterName: string

OCR Statistics Event - OCRIndexZoneProcessedEventArgs

The OCR Statistics custom code event uses the OCRIndexZoneProcessedEventArgs class to pass index values populated by Open Text OCR zones to the custom code. This event is triggered once the contents of an Open Text OCR zone populate an index value.

The following properties are located within the custom code:

1. `DocumentId: string`
2. `PageId: Guid`
3. `PageIndex: int32`
4. `OCRWords: int32`

The **`OCRWords`** property contains the following variables:

```
internal OCRCharacter[] characters = new OCRCharacter[] { };
internal Int32 line = 0;
internal System.Drawing.Point location = new System.Drawing.Point();
internal System.Drawing.Size size = new System.Drawing.Size();
```

The OCRCharacter variable contains the following properties:

```
public System.Drawing.Point Location
{
    get
    {
        return location;
    }
}
public System.Drawing.Size Size
{
    get
    {
        return size;
    }
}
public Byte Confidence
{
    get
    {
        return confidence;
    }
}
public Char Code
{
    get
    {
        return code;
    }

}
public bool Rejected
{
    get
    {
        return rejected;
    }
}
public Char[] Alternatives
{
    get
    {
        return alternatives;
    }
}
```

5. RecognitionTime: int32 (milliseconds)
6. AdditionalValues: Hashtable
7. FieldName: string

OCR Statistics Event - OCRMarkSenseZoneProcessedEventArgs

The OCR Statistics custom code event uses the OCRMarkSenseZoneProcessedEventArgs class to pass auto document break zone statistics to the custom code. This event is triggered when an Open Text OCR zone inserts an auto document break page between documents.

The following properties are located within the custom code:

1. `DocumentId: string`
2. `PageId: Guid`
3. `PageIndex: int32`
4. `OCRWords: int32`

The `OCRWords` property contains the following variables:

```
internal OCRCharacter[] characters = new OCRCharacter[] { };
        internal Int32 line = 0;
        internal System.Drawing.Point location = new
System.Drawing.Point();
        internal System.Drawing.Size size = new
System.Drawing.Size();
```

The `OCRCharacter` variable contains the following properties:

```
public System.Drawing.Point Location
{
    get
    {
        return location;
    }
}
public System.Drawing.Size Size
{
    get
    {
        return size;
    }
}
public Byte Confidence
{
    get
    {
        return confidence;
    }
}
public Char Code
{
    get
    {
        return code;
    }

}
public bool Rejected
{
    get
    {
        return rejected;
    }
}
public Char[] Alternatives
{
    get
    {
        return alternatives;
    }
}
```

5. RecognitionTime: int32 (milliseconds)

AdditionalValues: Hashtable

Saving Indexes Event - IndexSaveEventArgs

The Saving Indexes event uses the IndexSaveEventArgs class to pass index values to custom code. The Saving Indexes event is triggered as index values are saved to the batch. This class contains the BatchNavigation enumeration property that determines which document opens immediately after indexes are saved.

```
IndexSaveEventArgs args = base.Parameter as IndexSaveEventArgs;
```

> **Note**:
>
> By default, the Saving Indexes event proceeds to the next document.

Within your custom code, you can use the following constants to set the BatchNavigation enumeration property:

1. None: Remains on current document
2. NextDoc: Proceeds to next document
3. PreviousDoc: Returns to previous document
4. LastDoc: Proceeds to last document in batch
5. FirstDoc: Returns to first document in batch

For example, you can configure the BatchNavigation enumeration property to remain on the current document after index values are saved:

```
args.BatchNavigation = BatchNavigation.None;
```

User-Defined Custom Code Generators

If you would like to display your own customized custom code generators you must complete the following steps:

1. Build an assembly that contains a class implementing the ICustomCodeGenerator interface. This interface is located within **DSI.Capture.ScriptingLibrary.dll**.
2. Copy the generated assembly into the **\CCGenerators** folder where PaperVision Capture Desktop was installed.
3. Your custom code generator must implement the following interface:

```
public interface ICustomCodeGenerator
{
    /// <summary>
    /// Returns name of Generator
    /// </summary>
    /// <returns></returns>
    string GetName();
    /// <summary>
    /// Returns description of Generator
    /// </summary>
    /// <returns></returns>
    string GetDescription();
    /// <summary>
    /// Returns supported languages
    /// </summary>
    /// <returns></returns>
    Language[] GetSupportedLanguages();
    /// <summary>
    /// Generates the CustomCode object
    /// </summary>
    /// <param name="options"></param>
    /// <returns></returns>
    CustomCode GenerateCustomCode(CustomCodeGeneratorOptions
options);
```

```
            /// <summary>
            /// Returns tags
            /// </summary>
            /// <returns></returns>
            CustomCodeGeneratorTags GetTags();
        }
        public sealed class CustomCodeGeneratorTags
        {
            /// <summary>
            /// Is this an export?
            /// </summary>
            public bool IsExport = false;
            /// <summary>
            /// Is this automated?
            /// </summary>
            public bool IsAutomated = false;
            /// <summary>
            /// Is this manual?
            /// </summary>
            public bool IsManual = false;
            /// <summary>
            /// Is there a UI for this?
            /// </summary>
            public bool CanDesign = false;
        }
      public sealed class CustomCodeGeneratorOptions
        {
            private IWin32Window parent = null;
            private Language language = Language.CSharp;
            private PVBatch batch = null;
            private ICustomCodeGenerator generator = null;  //Optional
generator to be used (frmSelectGenerator will not be displayed)
            private object parameter = null;   //Custom Code to be passed
to generator (optional). This can be used for Design
```

```
        /// <summary>
        /// Parent window
        /// </summary>
        public IWin32Window Parent
        {
            get
            {
                return parent;
            }
            set
            {
                parent = value;
            }
        }

        /// <summary>
        /// Language
        /// </summary>
        public Language Language
        {
            get
            {
                return language;
            }
            set
            {
                if (value == ScriptingLibrary.Language.Unknown)
                    throw (new Exception("Property does not support
Language.Unknown!"));
                language = value;
            }
        }
```

```
/// <summary>
/// Batch
/// </summary>
public PVBatch Batch
{
    get
    {
        return batch;
    }
    set
    {
        batch = value;
    }
}

/// <summary>
/// Generator
/// </summary>
public ICustomCodeGenerator Generator
{
    get
    {
        return generator;
    }
    set
    {
        generator = value;
    }
}
```

```
        /// <summary>
        /// (Internal usage only)
        /// </summary>
        public object Parameter
        {
            get
            {
                return parameter;
            }
            set
            {
                parameter = value;
            }
        }
    }
```

Digitech Systems' API

Digitech Systems' API is accessible from within the Script Editor. The API provides classes for reading/writing documents and indexes within the current batch. For more information on the Digitech Systems API, launch the **PVCaptureBatchAPI.chm** Help file that is found within the **Docs** directory where PaperVision Capture Desktop was installed. This help file provides Microsoft Developer's Network (MSDN)-style documentation on our DSI.Capture.API namespace, including code samples.

Custom code samples are found in the **Library\Samples** directory (as text or XML files), where PaperVision Capture was installed. You can cut and paste the code directly into the Script Editor for a Custom Code step. The following code samples are included:

- **AddPrefixValuetoBatchDocumentIndexes** iterates through all documents comprising a batch and appends prefixes to index values.
- **CalltoCustomAssembly** demonstrates one way to call out to code in your own assembly.
- **CopyIndexValues** duplicates an index value from a source document to one or more subsequent documents.
- **DisplayBatchPageCount** displays the total number of pages in the batch (designed to be run from a manual custom code event).
- **ImportASCIIwithImages** imports images and index information from a number of other document imaging systems.

> **Note**:
>
> Constants at the beginning of the script must be configured in order for the operator to execute the script successfully.

- **InspectBeforeAddPage** examines the physical dimensions of a scanned image and inserts a document break if the page is detected as an envelope.
- **MatchandMergeOnIndexValidate** executes custom code that will look up and populate index values when the operator enters a index value and tabs to the next field.
- **OCRFullTextPageStatistics** records Open Text Full-Text OCR statistics per selected output. Statistics are recorded when the Open Text Full-Text OCR step processes a page and converts the page to the selected output (s).
- **OCRIndexZoneStatistics** records Open Text Zonal OCR statistics when an Open Text OCR zone populates an index value.
- **OCRMarkSenseZoneStatistics** records Open Text Zonal OCR statistics when an Open Text OCR zone inserts an auto document break page between documents.
- **OpenBatchCustomCode** shows how code can be fired when a batch is opened. To attach to this event, a new "Open Batch" event was added to the Job Configuration properties.
- **MultiPageTIFFConversion** divides a multi-page TIFF into separate images (one image per page).

- **SetScanDate** automatically sets a scan date index value (document creation date) into the batch for every document. The document's creation date is the date/time the document entered the batch. The date/time value is stored in Universal Time Coordinated (UTC), also known as Greenwich Mean Time (GMT). For example, Denver, Colorado's UTC time at 2:00 PM on April 9, 2009 will display as "04/09/2009 20:00:00". To change the date/time value to your local time zone instead of UTC, change the code in line 46 to:

```
if (!this.Batch.TrySetIndexValue(id, "ScanDate",
documentCreatedDate.ToLocalTime(), true, out error))
```

- **ValidateIndex** provides an example of how to validate an index field value.

Batch Property

Within your custom code, you can access the Digitech Systems API via the Batch property. The Batch property is of the type DSI.Capture.API.Batch and represents the primary entry point for the Digitech Systems API.

For example, to insert a new document to a batch within your CallHandler method (C# in this case), you can type:

```
this.Batch.TryInsertDocument(/*see API documentation for
parameters*/)
```

Another approach would be to call out to your own assembly and pass the instance of the Batch object to your code (again, the instance is available as the "Batch" property inside the pre-written "Code" class.) This approach would allow you to use Visual Studio for coding. Then, at run time, you would need to ensure that your assembly is located in the same directory as the PaperVision Capture Desktop executables.

Additional API Functions

In addition to the API Functions documented in the **PVCaptureBatchAPI.chm** help file, the API functions described in this section can be used within your custom code.

Custom Code/Export Functions
`protected string[] GetPageFiles(string documentID)` Returns path values for all images contained in a document (from all pages)
`protected Stream GetFileStream(PVFile file)` Returns the stream for a specified PVFile
`protected Stream[]GetDocumentStreams(string documentID)` Returns an array of streams for all files contained in a document (from all pages)
`protected Stream[] GetDocumentStreams(string documentID, string jobStepName, bool bitonal)` Returns streams for all files contained in a document (from all pages) based on job step name and bitonal option
`protected void CopyStreamToDisk(Stream stream, string path)` Copies content of a stream to disk
`public string[] CopyFilesToDisk(string documentID, string rootPath)` Copies all files from a document (from all pages) to a folder and returns an array for all image path values
`protected void SetPersistValue(string key, string value, string rootPath)` Copies all files from a document (from all pages) to a folder based on job step name and bitonal option
`protected string Get PersistValue(string key, string rootPath)` Reads persisted value for a key
`protected string GetNextLockedPath(string root, Int32 maxExportSize, bool exclusive)` Returns the next available path (path is locked before it is returned)
`protected string GetNextLockedPath(string root, Int32 maxExportSize)` Returns the next available path (path is locked before it is returned)

Custom Code/Export Functions (continued)

```
String GetNextLockedPath(string root, Int32 maxExportSize,
ExcludePathDelegate excludeFunction, bool exclusive)
```

Returns the next available path (path is locked before it is returned)

```
protected void UnlockPath(string path)
```

Deletes lock for a specified path

```
void ClearRootPath(string path)
```

Deletes all folders containing empty subfolders for all folders listed under 'path'

```
protected void SetExportComplete(string path)
```

Flags folder as complete by dropping export.complete file

```
protected bool IsExportComplete(string path)
```

Checks whether export folder is flagged as complete

```
protected bool IsExported(string documentID)
```

Checks whether document was previously exported

```
protected bool SetExported(string documentID)
```

Sets the document's exported status

```
protected void DeleteDocument(string documentID)
```

Deletes document after it has been exported

```
protected void SetStatus(string status, Int32 percentage)
```

Returns percentage of custom code that has been executed

```
protected Int32 GetNonExportedDocumentCount()
```

Returns number of non-exported documents

```
{
        Int32 count = 0;
        foreach (string id in GetDocumentIDs())
        {
            if (!IsExported(id))
                ++count;
        }
        return count;
    }
```

Full-Text OCR Functions

```
protected string[] GetPageText (string filePath)
```

Returns text for each page

```
protected string[] GetOCRFiles (string documentID, string
stepName, string converterCode)
```

Returns Full-Text OCR files that belong to a specific converter

```
string[] GetOCRFiles (string documentID, string stepName,
string converterCode, string path)
```

Writes Full-Text OCR files belonging to a specific converter to directory 'path'

> **Important**!
>
> The caller is responsible for post-processing clean-up if the files are not required.

Image Processing Functions

```
string ConvertImages(string[] sourceFiles, string
destinationFile, ConvertFileType convertFileType)
```

Converts one or more images to a single destination image file and returns the actual path under which the file was saved

```
Int32 GetPageCount(string sourceFile)
```

Returns the number of pages found in a multi-page image

```
string GetPageImage(string sourceFile, Int32 pageIndex,
string destinationFile, OutputFileType outputFileType)
```

Retrieves specific image referenced by a specific page index in a multi-page image

```
protected string[] GetPageFiles(string documentID)
```

Returns a path value for all images belonging to a document (from all pages)

```
bool IsMultipageFormat(ConvertFileType convertFileType)
```

Determines if the passed file type supports multi-page format

PVBatch Helper Functions
`Int32 GetBlankIndexCount()` Returns the number of blank indices
`string[] GetAvailableFields()` Returns the set of fields that can be written to
`string GetIndexValue(string fieldname)` Returns the field value for the specified field name
`void SetIndexValue(string fieldname, string fieldValue)` Assigns a field value for a specified field name **Note**: This function cannot be used with a detail set field; otherwise, an exception will result. Also, when called from within an Index Validate event, this function can only be used for the target index.
`string[] GetDetailSetFields()` Returns the field names of the detail set in Match and Merge
`void AssignDetailSet(DataRow row)` Assigns a detail set field in automated match and merge using a single passed DataRow
`void AssignDetailSet(DataSet dataset)` Assigns detail set values from a DataSet (returned from the database) - used in match and merge
`void AssignDetailSet(DataRow row, DataSet indices)` Assigns a detail set from a passed DataRow value (manual match and merge) - detail set is not written to the batch; instead, it is written to the indices DataSet which passed from the UI
`void AssignDetailSet(DataSet dataset, DataSet indices)` Assigns detail set values to passed indices (manual match and merge)
`void UpdateCurrentIndex(DataRow row)` Updates the current index value from the passed DataRow - row is retrieved from a dataset populated by the SQL database (match and merge)
`Bool IsFieldDetailSet(string fieldName)` Checks whether the specified field is a detail set field
`PVIndexMetadata GetIndexMetadata(string fieldName)` Returns metadata for an index

PVBatch Helper Functions (continued)
`bool IsFieldEmpty(string fieldName)` Checks whether a field is empty
`string GetMappedColumn(string fieldName)` Returns the mapped column to a specific field name (match and merge)
`DataTable GetMapping()` Returns a mapping table between indices and table columns (match and merge)
`string GetWhereClause()` Generates a WHERE clause to be used in the SQL query (match and merge)
`string GetWhereClause(DataRow row)` Generates a WHERE clause to be used in the SQL query that uses the values in DataRow to add conditions (match and merge)
`string[] GetDocumentIDs()` Returns list of document id values
`PVPage[] GetPages(string documentID)` Returns a list of pages for a specific document
`string GetPath(PVFile file)` Returns a path for a specified file
`PVIndex[] GetIndices(string documentID)` Returns a list of indices for a specific document
`PVDetailSet[] GetDetailSets(string documentID)` Returns the DetailSet values for a specific document
`PVFile GetPreferredFile(PVPage, string jobStepName, bool bitonal)` Returns the file that matches the bitonal value; otherwise, the first file in the array is returned
`string GetExtenstion(string imagePath)` Returns the extension of an image path

Enumerations

The enumerations described in this section can be used within your custom code.

public enum ConvertFileType

This enumeration is used by the ConvertImages() function and specifies the conversion types that will be applied to one or more images.

```
  {
        /// <summary>
        /// No file conversion (returns image input path and appends an
extension if not passed in destinationFile variable)
        /// </summary>
        CVT_NO_CONVERSION,
        /// <summary>
        /// TIFF with Group IV and/or medium JPEG compression (single- or
multi-page)
        /// </summary>
        CVT_TIFF_G4_MEDJPG,
        /// <summary>
        /// TIFF with Group IV and/or LZW compression (single- or multi-
page)
        /// </summary>
        CVT_TIFF_G4_LZW,
        /// <summary>
        /// TIFF with no compression (single- or multi-page)
        /// </summary>
        CVT_TIFF_NONE,
        /// <summary>
        /// PDF with Group IV and/or medium JPEG compression (single- or
multi-page)
        /// </summary>
        CVT_PDF_G4_MEDJPG,
        /// <summary>
        /// PDF with Group IV and/or LZW compression (single- or multi-
page, and image-only PDFs)
        /// </summary>
        CVT_PDF_G4_LZW,
```

```
        /// <summary>
        /// JPEG with medium JPEG compression (single-page only)
        /// </summary>
        CVT_JPG_MEDJPG,
        /// <summary>
        /// GIF (single-page only)
        /// </summary>
        CVT_GIF,
        /// <summary>
        /// BMP (single-page only)
        /// </summary>
        CVT_BMP,
        /// <summary>
        /// PNG (single-page only)
        /// </summary>
        CVT_PNG
       /// <summary>
        /// JPEG 2000
        /// </summary>
        CVT_JPG2000

    }
```

public enum OutputFileType

This enumeration is used by the GetPageImage() function, and specifies the output file types when single pages are retrieved from a multi-page image.

```
{
        /// <summary>
        /// JPEG
        /// </summary>
        OFT_JPG
        /// <summary>
        /// TIFF
        /// </summary>
        OFT_TIFF
        /// <summary>
        /// Bitmap
        /// </summary>
        OFT_BMP
  }
```

public enum UIRefreshLevel

This enumeration synchronizes the user interface with any changes made to the batch via custom code. Setting the UIRefreshLevel in custom code forces the user interface to refresh the selected component specified by the enumeration value (None, Index, CurrentDocumentIndexes, etc.). If you use either the Index Populated or Index Validate Custom Code Event to change an index value, the Indexing window will remain synchronized using the **UIRefreshLevel.Index** value.

```
{
    /// <summary>
    /// no UI refresh required
    /// </summary>
    None = 0x00,
    /// <summary>
    /// index field needs to be refreshed (i.e., via IndexValidate or
  IndexPopulate event)
    /// </summary>
    Index = 0x01,
    /// <summary>
    /// all indexes for current document need to be refreshed (does not
  apply to Match and Merge)
    /// </summary>
    CurrentDocumentIndexes = 0x02,
    /// <summary>
    /// current page needs to be refreshed
    /// </summary>
    SinglePage = 0x04,
    /// <summary>
    /// multiple pages need to be refreshed
    /// </summary>
    MultiPage = 0x08
   }
```

Public Properties

The public properties listed in this section can be used within your custom code:

```
 /// <summary>
/// Batch object
/// </summary>
public PVBatch Batch

/// <summary>
/// Parent window
/// </summary>
public Control Parent

/// <summary>
/// Control referencing the current index
/// </summary>
public Control Control

/// <summary>
/// Used to pass optional parameters
/// </summary>
public object Parameter

/// <summary>
/// Code result that returns status of custom code execution
/// </summary>
public CodeResult CodeResult

/// <summary>
/// PDF Resolution used when importing PDF files
/// </summary>
public Int32 PDFResolution

/// <summary>
/// PDF Smoothing option used when importing PDF files
/// </summary>
public PDFSmoothing PDFSmoothing
```

Debugging Custom Code

Custom code that you enter in the Script Editor is compiled on the fly by the PaperVision Capture Desktop application so there is no way to debug or step through this code at run time. However, if you write code in your own assemblies and call out to these pre-compiled assemblies, then you can debug this code by attaching your debugger to the CaptureDesktop.exe process.

WARNING!

Do not attempt to debug code in a production environment. Doing so may adversely impact system performance and have unpredictable impacts on customer data and end-user functionality.

Match and Merge Wizard

When you edit your job settings from the Home page, you can configure the Match and Merge Wizard to match PaperVision Capture Desktop indexes to external database tables, spreadsheets, and other files. The Match and Merge Wizard is where you configure the connection properties, field mapping, and additional settings.

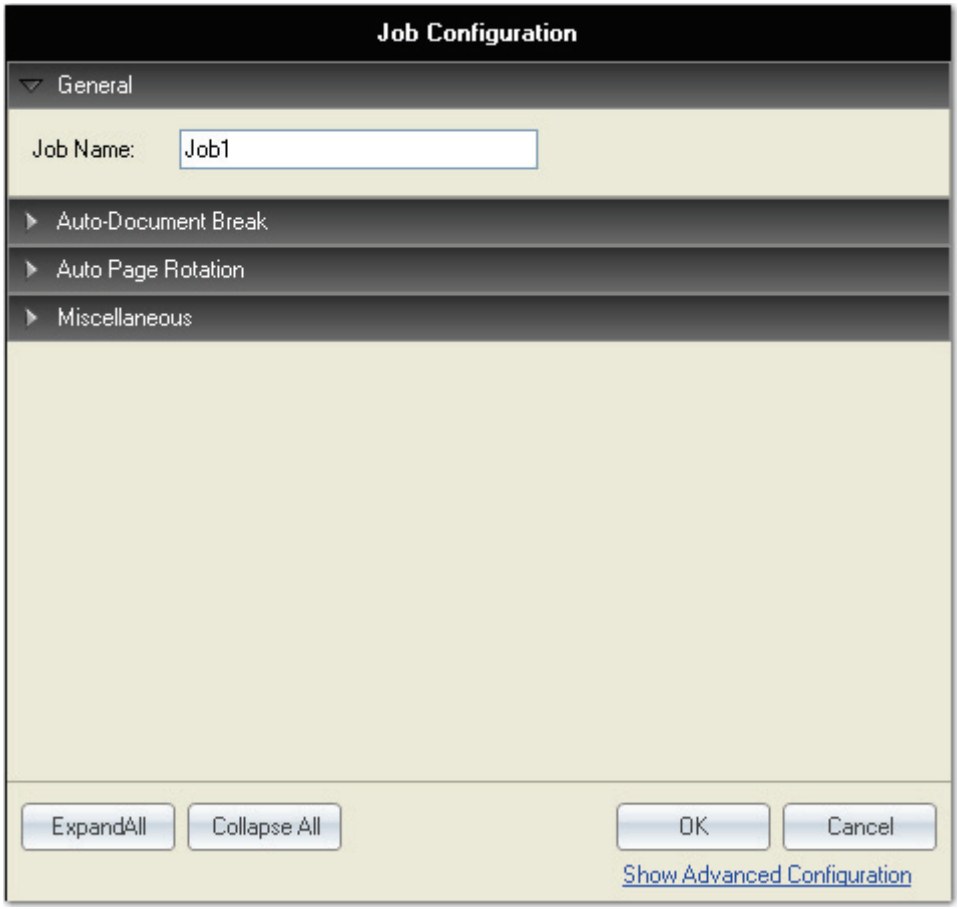
> **Note:**
>
> Ensure that the lookup table and columns for the database have been configured and indexes have been defined before launching the Match and Merge Wizard.

Configuring the Match and Merge Wizard

The Match and Merge Wizard provides a graphical user interface to guide you through the configuration of your database server parameters, field mapping, and other settings.

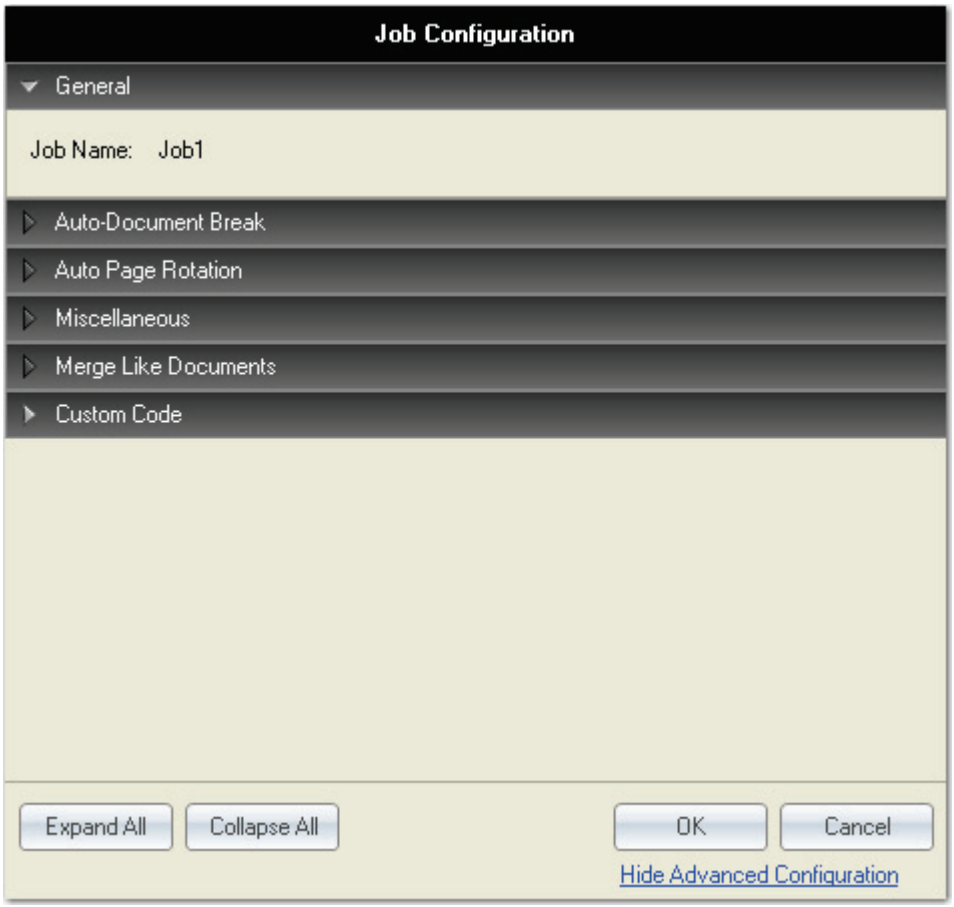
To configure the Match and Merge Wizard:

1. In the **Home** page, click the **Job**  icon in the **Settings** toolbar group. The **Job Configuration** window appears.



Job Configuration

2. Click the **Show Advanced Configuration** link in the lower right. The **Advanced Job Configuration** settings appear.

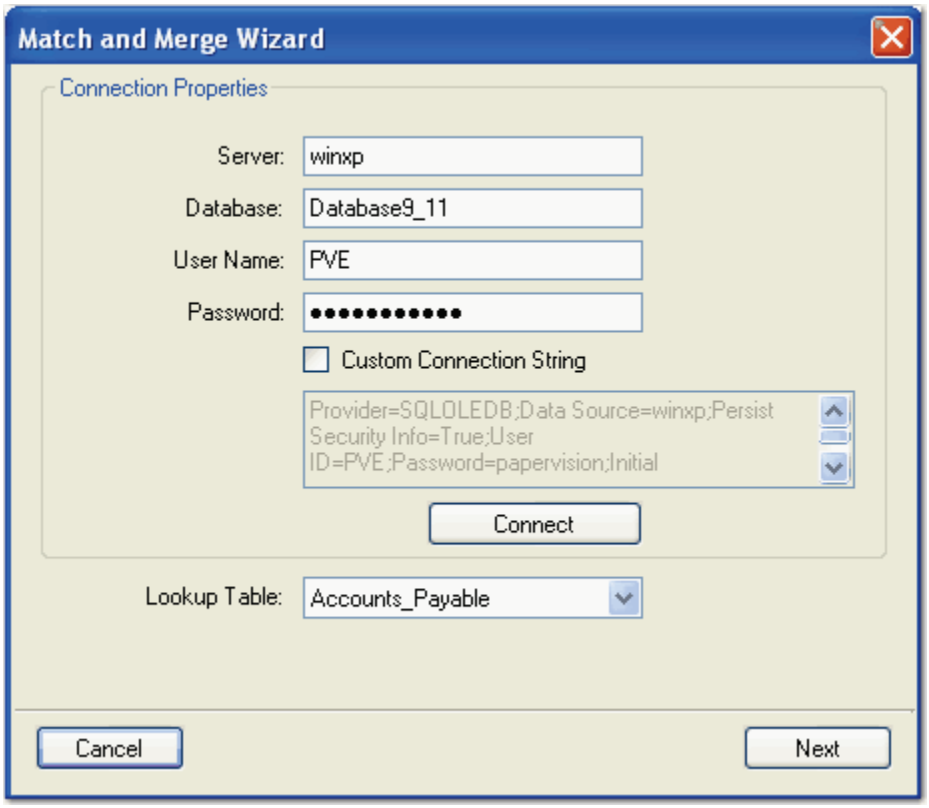


Advanced Job Configuration

3. Expand the **Custom Code** menu. The **Custom Code** events appear.

Custom Code

4. Click the **Add** icon next to **Match and Merge**. The **Match and Merge Wizard - Connection Properties** screen appears after you launch the wizard. You can configure the database connection properties including the database server and name, user name and password, and database lookup table.



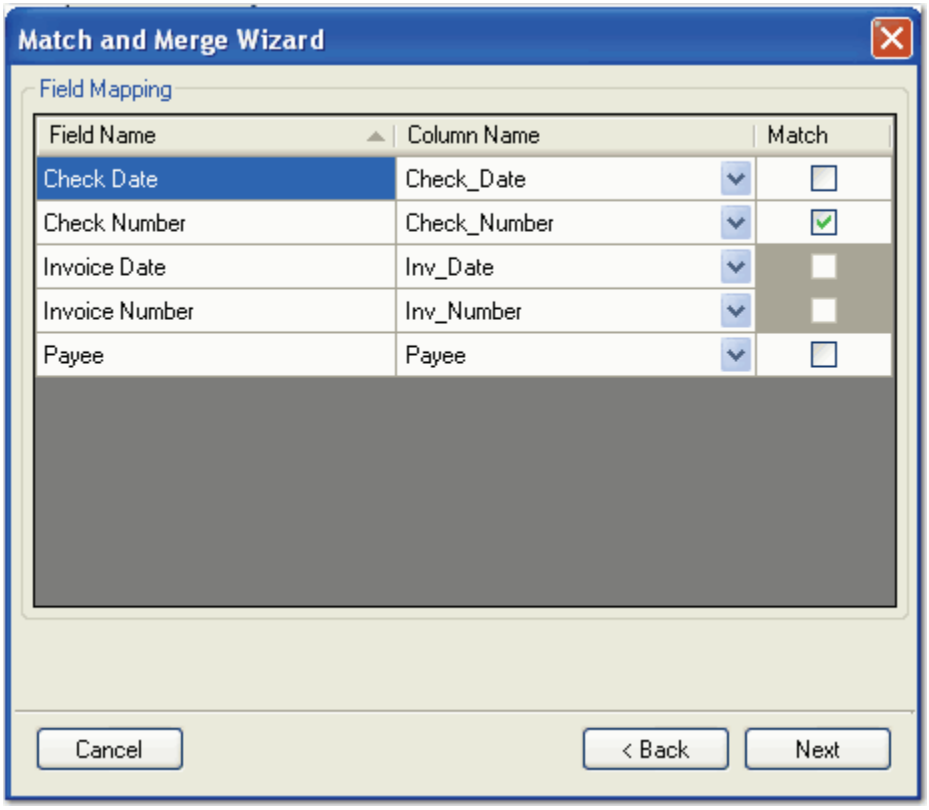
Connection Properties

5. In the **Connection Properties** screen, enter the database server and database name where Match and Merge will be performed.
6. Enter the user name and password for the database server connection.

> **Note**:
>
> If you leave the **User Name** and **Password** fields blank, the database connection will use the Windows Authentication credentials. Entering a user name and password for the database will supercede the Windows Authentication credentials.

7. To insert a custom connection string, select the check box and edit the string in the window.
8. Click the **Connect** button to test the connection to the database. Once connected, the **Lookup Table** drop-down menu will populate.
9. Click the **Lookup Table** drop-down menu to select the database table used for lookups.

10. Click **Next**, and the **Field Mapping** screen appears.



Field Mapping

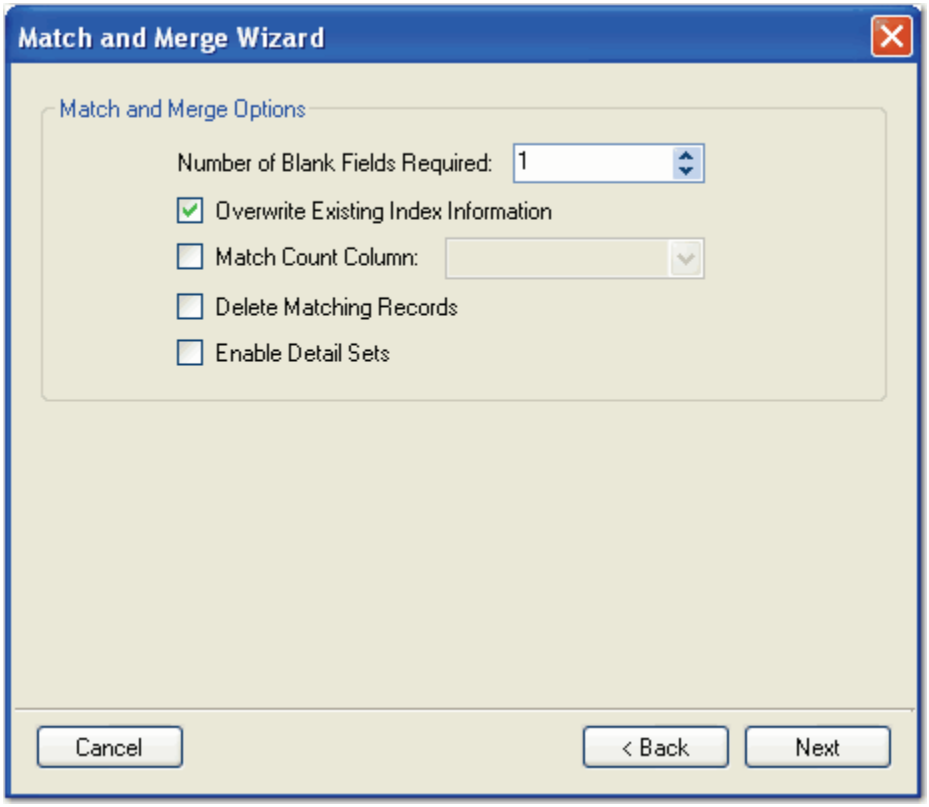
11. The Field Mapping screen allows you to match the columns in the database to the field names (indexes) that you defined. Click the **Column Name** drop-down menu(s) to select the database column name that will match the field name(s). If one of the index fields should not be matched, do not map it to the Column Name. When the user executes the Merge Index Values command, only the mapped fields will be populated in the Indexing window.

> **Note:**
>
> Field names are synonymous with indexes that have been defined.

12. After selecting the column names, click the **Match** check box(es). Detail fields are denoted with shaded Match columns and cannot be selected to match.
    - In the example above, the Check Number index value, entered by the operator, will be matched with the corresponding Check_Number column in the database.
    - Once the operator executes the Merge Index Values command, the corresponding Check Date, Invoice Date, Invoice Number, and Payee are populated in the Indexing window.
    - If the operator does not know the exact index value during hand-key indexing, the operator can insert wildcard characters to perform a partial search against a database. For example, the operator can insert the percent sign (%) to specify any number of unknown characters to search for in a SQL, Sybase, or Oracle database; the operator can insert the asterisk (*) to specify any number of unknown characters to search for within a Microsoft Access database.

> **Note:**
>
> All fields with the **Match** column selected must be populated prior to running the Merge Index Values command.

13. Click **Next**, and the **Match and Merge Options** screen appears.



Match and Merge Options

14. The Match and Merge Options screen contains additional parameters that define the match and merge process. In the **Number of Blank Fields Required** field, enter the number of fields that must be blank in order for PaperVision Capture to attempt to match during the custom code execution.
    - For example, you assign the **Number of Blank Fields Required** a value of two. If only one field is left blank before the Match and Merge is executed, PaperVision Capture will not match because at least two fields were not blank.
    - Valid values range from zero to the number of database columns that are defined. For example, if you have five database columns defined, you can enter a value from zero to five.
15. If you select the **Overwrite Existing Index Information** check box, the Match and Merge values will overwrite the existing index entries already populated in the batch.
16. The Match Count Column setting applies only to integer data type columns in the database. Select the **Match Count Column** check box if the match count should increment in the database by one each time a match is encountered. If you enable this setting, choose the database column from the drop-down menu. If you enable this setting, choose the database column from the drop-down menu.
17. Select the **Delete Matching Records** check box to remove the matching record from the database once it is found during the match and merge process.

    Note:

    You can only enable the **Match Count Column** or the **Delete Matching Records** setting, but not both.

18. For manual indexing, select the **Enable Detail Sets** check box if the detail fields should be populated when the operator enters the index fields.

    Note:

    If you do not select this check box, the operator is presented with a pick list of data that meets the index field criteria. The operator then selects the appropriate record, and the detail fields are populated according to the selected record.

    When you define a Custom Code step to run an automated Match and Merge process:
    - If you select the check box, all detail fields are automatically populated (e.g., if five rows of data meet your criteria, five detail sets are populated).
    - Conversely, if you do not select the check box, the detail fields populate with data from the first row of results.
19. Click **Next**, which opens the last screen of the wizard.
20. Click **Finish**, which opens the Script Editor so you can make edits to the code if necessary.
21. Click **OK**.

Matching and Merging with Text Files

If you are using custom code to match and merge index fields with a text file, you can control how data is handled in the lookup table. If the text file contains dates, currency, or decimal data, for example, you can manipulate how data is formatted by creating a schema information (Schema.ini) file and placing it in the same directory where the text file resides. If you do not define how date columns are handled, date values will be imported in the DateTime format. Information on how to create Schema.ini files can be found in the Microsoft Software Developer's Network:

http://msdn.microsoft.com/en-us/library/ms709353(VS.85).aspx

Chapter 9 - Exports

In the Auto Process page, you can determine how images are sent to external applications, web sites, folders, and other locations. For example, you can export images that you have scanned, cropped, and indexed, and subsequently upload them to an external web site or application, such as ImageSilo, PaperVision Enterprise, and Microsoft SharePoint. Or, you can simply export your scanned documents as PDF files to a specified location on your local drive or network drive.

You can assign properties that can be configured within a graphical user interface organized in tabbed dialog boxes including the General, Indexes, OCR, Options, and FTP tabs. To expedite the configuration of each export script, default properties are provided to you in drop-down menus, editable fields, and check boxes that you can easily modify.

As exports are executed, they are appended to the first available destination folder based on sequence number and maximum export size (defined by the **Max Export Size** property). When the maximum export size is reached, exports will be appended to the next available folder. As a result, all subsequent processes will append exports to the next available folder.

For example, four folders are available for the **Root Path** property, and the **Max Export Size** is defined as 600 MB:

1. Folder_1: 600 MB
2. Folder_2: 400 MB
3. Folder_3: 600 MB
4. Folder_4: 100 MB

Since the maximum export size has been reached in Folder_1, Folder_2 will be used as the export folder.

Export Custom Code Configuration

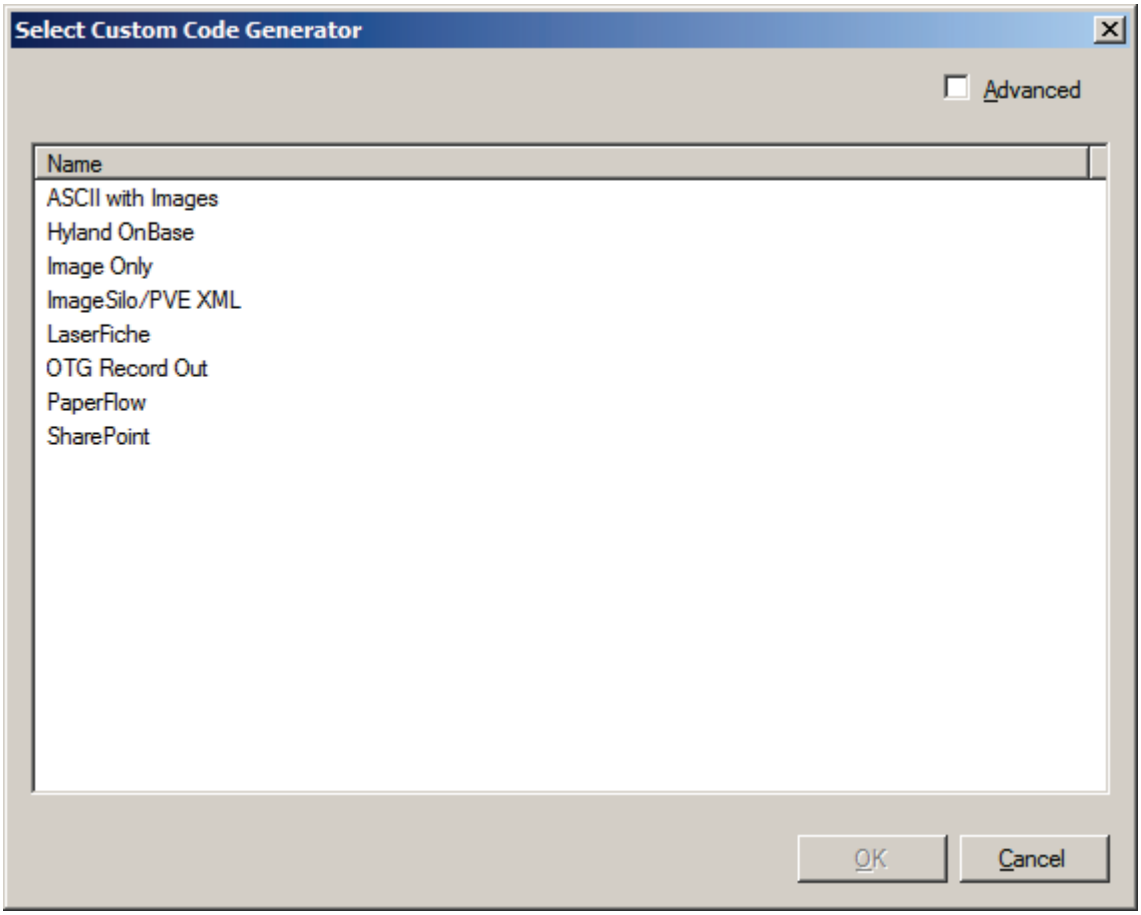
In PaperVision Capture Desktop, exports are also driven by the custom code engine and allow further opportunity for customization.

To configure an export:

1. In the Auto Process page, click the **Wizard** 

icon in the **Export** toolbar group. The **Select Custom Code Generator** dialog box appears, where exports are listed.

> **Tip**:
>
> To view advanced settings, select the **Advanced** check box. Advanced settings include programming language options (C# or Visual Basic) and the Export Template, which contains additional pre-defined code that will automatically process batches.



Select Custom Code Generator

2. Select the export, and then click **OK**. The selected export and corresponding properties appear in tabbed dialog boxes.
    - Default values and applicable index fields are provided for your reference, and drop-down menus contain only the options specific to your selected export.
    - You can manually enter file paths or browse to the appropriate directory.
3. After you have configured the properties for the selected export, click **OK**.
4. Prior to processing the export, click the **Enable** check box in the **Export** toolbar group.
5. To process the export, click the **Start**  icon in the **Process Action** toolbar group. A progress bar will appear, indicating whether the batch was successfully processed.

Export Definitions

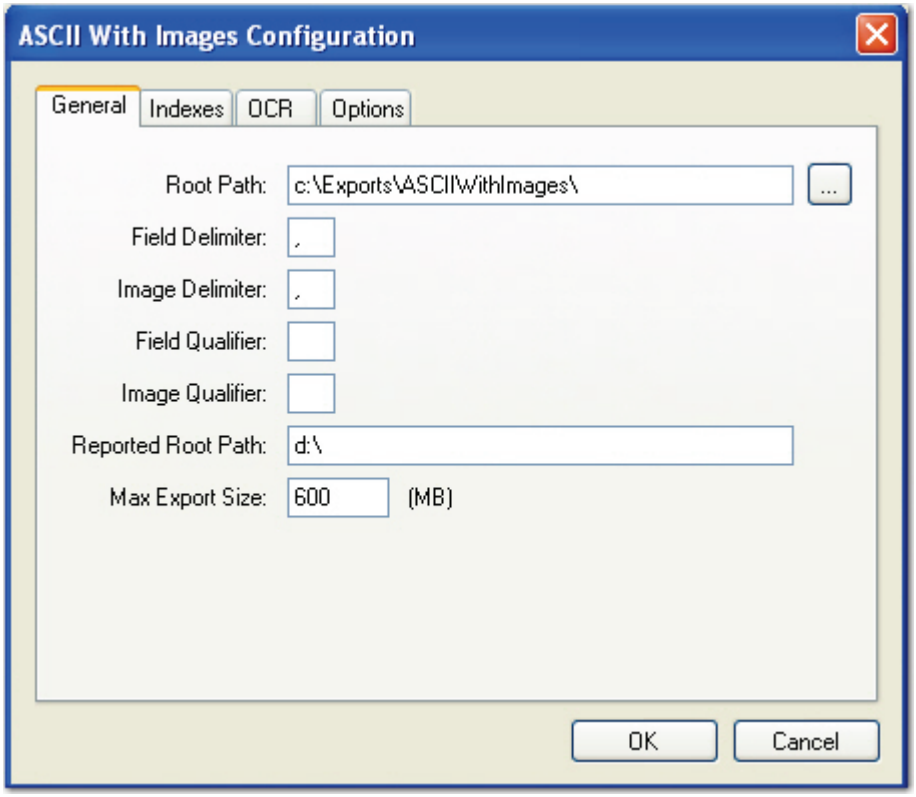
PaperVision Capture Desktop contains specific definitions that can be configured within a graphical user interface. When you configure an export from the **Select Custom Code Generator** dialog box, properties specific to that export will appear in tabbed dialog boxes. To expedite the configuration of each export script, default properties are provided for you in drop-down menus, editable fields, and check boxes that you can modify.

ASCII with Images

The ASCII with Images export creates an ASCII text file containing images that can be imported into other systems. The format of the file is completely customizable.

To configure the ASCII with Images export:

1. From the **Select Custom Code Generator** dialog box, double-click the ASCII with Images export, and the tabbed **ASCII with Images Configuration - General** dialog box appears.



ASCII with Images Configuration - General

Default values, paths, and other default settings are provided for your reference, and drop-down menus contain only the options specific to your selected generator. In addition, you can browse to the appropriate directories instead of manually entering file paths.

2. Assign the appropriate properties in the **Indexes, OCR,** and **Options** tabs. Descriptions for constant values appearing in the resulting export script are listed below.
3. When you have finished configuring the export, click **OK**.

ASCII with Images - General

When you configure properties in the General tab, you can modify constant values that appear in the resulting export script. Properties and resulting script constants that are available for configuration are described below:

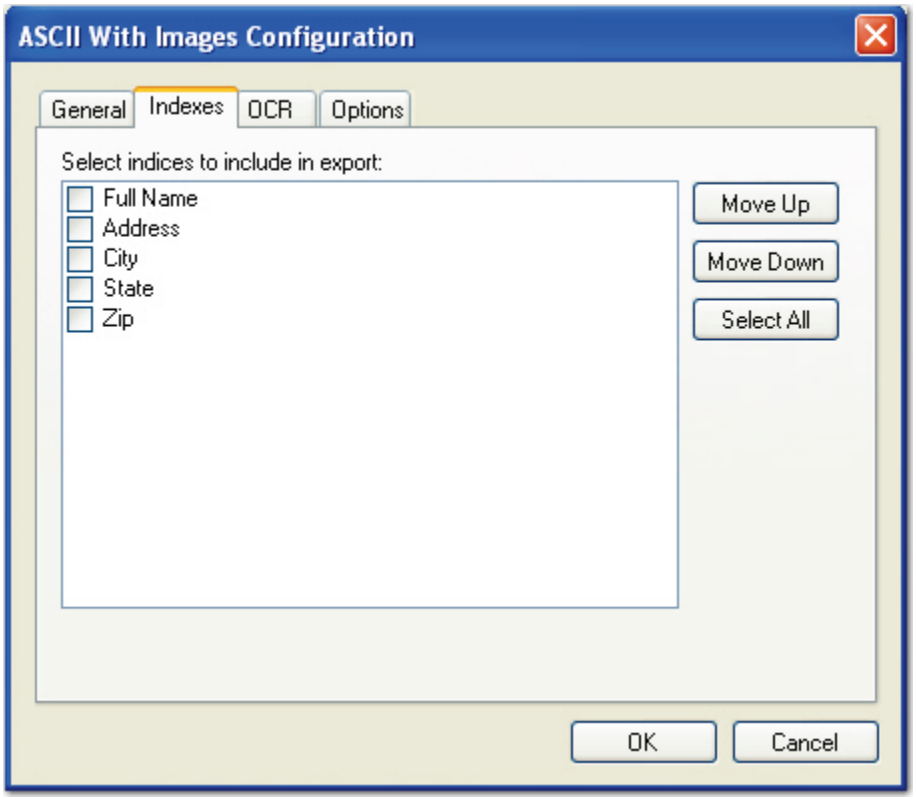
- **Root Path:** This is the location where the exports will be created once you have started the automated process. Script constant: ROOT_PATH
- **Field Delimiter**: This customizable delimiter separates index values, page number/counts, and image sizes. Script constant: FIELD_DELIMITER
- **Image Delimiter**: This customizable delimiter separates images when exporting using multi-line indexing and converting to single-page images. Script constant: IMAGE_DELIMITER
- **Field Qualifier**: This property contains the characters that surround the field values. By default, quotation marks will appear. Script constant: FIELD_QUALIFIER
- **Image Qualifier**: This property contains the characters that surround the image values. By default, quotation marks will appear. Script constant: IMAGE_QUALIFER
- **Reported Root Path**: The path referenced in the export file originates from this location, not the ROOT_PATH. Script constant: REPORTED_ROOT_PATH
- **Max Export Size**: This property indicates the maximum export file size (in MB), which defaults to a value of "600". Script constant: MAX_EXPORT_SIZE

> **Note**:
>
> If the Root Path is blank, the export will be written to the directory where the application was installed (e.g., C:\Program Files\Digitech Systems\PaperVision Capture Desktop). If the Reported Root Path is blank, the resulting export script will display a blank value for the REPORTED_ROOT_PATH.

ASCII with Images - Indexes

In the Indexes tab, you can select the index values that will appear in the export by double-clicking within the appropriate check boxes. Alternatively, click the **Select All** button to include all indexes in the export. You can also click **Deselect All** to remove all selections. To change the order in which the indexes display, single-click an index name (to highlight it), and then click the **Move Up** or **Move Down** buttons.

> **Tip**:
>
> Single-click an index name to move it up or down the list. Double-click an index name to include it in the export.



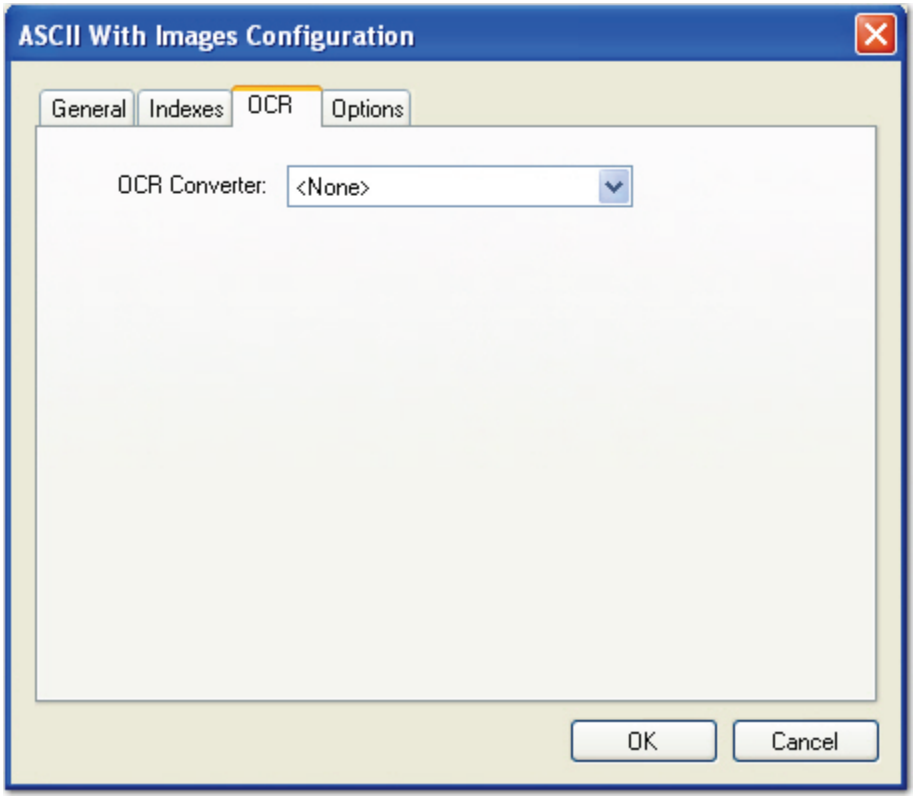
ASCII with Images Configuration - Indexes

To edit the indexes in the resulting export script, you can modify the INDICES_TO_INCLUDE constant in the export script:

- **Indices to Include**: Your selections determine what index values are included in the export file. If you leave this array blank, no indices are included. In the resulting script, you can enter the name of the index value(s) between quotation marks, and separate each index value with a comma. Script constant: INDICES_TO_INCLUDE

ASCII with Images - OCR

When you configure the properties in the OCR tab, you can modify constant values that appear in the resulting export script.



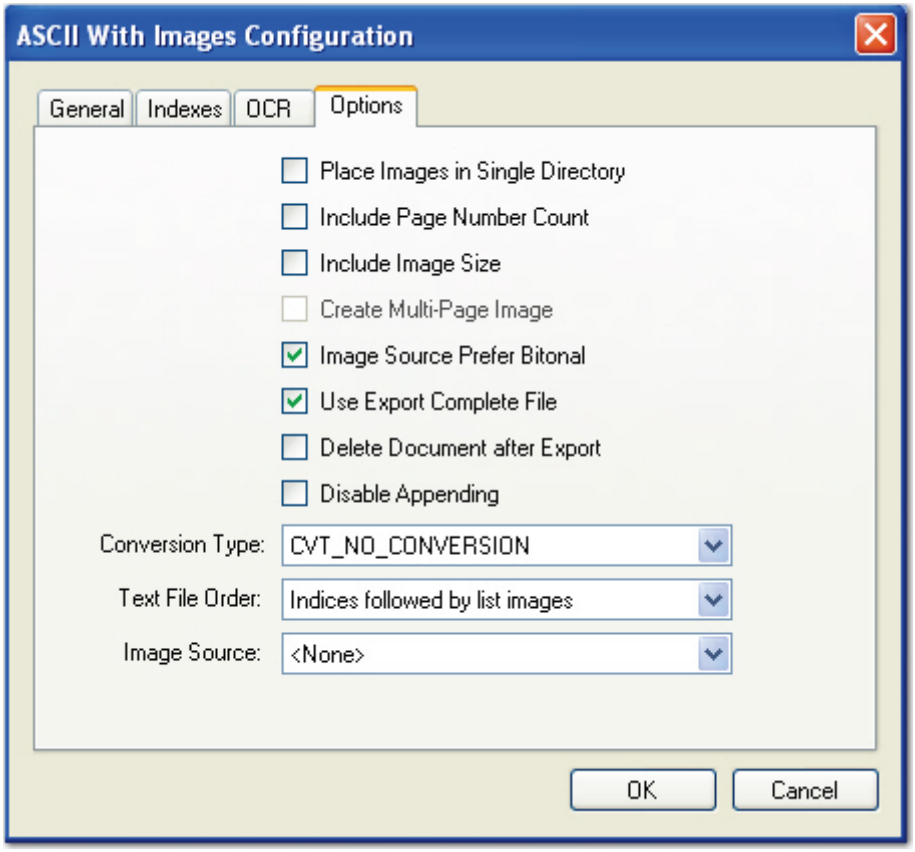
ASCII with Images Configuration - OCR

The OCR Converter property and resulting script constant are described below:

- **OCR Converter:** Your selection specifies the OCR converter code, such as text, PDF, etc., whose output format is used to export full-text data. No value is defined by default, so images will be retrieved instead of full-text output files. For a list of converter codes, see the **PVCaptureBatchAPI.chm** help file's **PVBatch.TryGetOCRFiles Method** topic found within the Docs directory where PaperVision Capture Desktop was installed. Script constant: **OCR_CONVERTER_CODE**

ASCII with Images - Options

When you configure properties in the Options tab, you can modify constant values that appear in the resulting export script.



ASCII with Images Configuration - Options

Properties and resulting script constants that are available for configuration are described below:

- **Place Images in Single Directory:** If set to **False**, the images will be placed in subdirectories at the ROOT_PATH (maximum of 1000 images per directory). If set to **True**, the images will be placed directly in the ROOT_PATH folder. Script constant: PLACE_IMAGES_IN_SINGLE_DIR
- **Include Page Number Count**: This determines whether the page number or page count of the document should be added as an additional field in the export. If set to **False**, when exporting in a multi-line format and creating single-page images, this value will match the page number of the document. If set to **True**, the value will match the total number of pages in the document. Script constant: INCLUDE_PAGE_NUMBER_COUNT

ASCII with Images - Options

- **Include Image Size:** This property determines whether the image file size is added as an additional field in the export. If set to **True**, this value will match the image size referenced on that line of the export file when exporting using a multi-line format and creating single-page images. If set to **False**, this value will match the size of the first page in the document. Script constant: INCLUDE_IMAGE_SIZE
- **Create Multi-Page Image:** Used in conjunction with CONVERSION_TYPE, this constant determines whether exported images are single-page or multi-page. Script constant: CREATE_MULTI_PAGE_IMAGE
- **Image Source Prefer Bitonal**: This property is applicable to dual-stream scanners and determines whether to export bitonal or color images. When set to **True**, which is the default setting, bitonal images will be exported. Script constant: IMG_SRC_PREFER_BITONAL_IMAGES
- **Use Export Complete File**: This property, set to **True** by default, generates an "export.complete" file once an export has reached its maximum file size, so data will no longer be appended to the export. When set to **False**, the "export.complete" file is not generated, so data may be appended to export folders that have not reached their maximum size. Script constant: USE_EXPORT_COMPLETE_FILE
- **Delete Document After Export:** This property specifies whether documents are deleted after they have been exported (set to **False** by default). Script constant: DELETE_DOCUMENT_AFTER_EXPORT
- **Disable Appending**: This property is set to **False** by default. When set to **True**, exported images will not be appended to export folders whose maximum file sizes have not been reached. Script constant: DISABLE_APPENDING
- **Conversion Type:** This property determines the type of image file created during the export. The default value, CVT_NO_CONVERSION, does not convert images during the export. If exporting to a format that supports both single and multi-page images, you must set the CREATE_MULTI_PAGE_IMAGE constant to **True** if you want to create multi-page images; otherwise single page images will result. For example, if you set this to CVT_TIFF_G4_MEDJPG, a TIFF image is created during the export. If the source image is binary, it will create a TIFF using Group 4 compression; if the source image is color (JPG or BMP), it will create a TIFF using Medium JPEG compression. See the **Custom Code Configuration** chapter on Enumerations for more information. Script constant: CONVERSION_TYPE

ASCII with Images - Options

- **Text File Order:** This property determines how the export file is formatted. You can select from the following options:

    a. **IndicesFollowedByListImages**: This option creates a single row for each document with indexes listed first, followed by image files.

    b. **ListImagesFollowedByIndices**: This option creates a single row for each document with images listed first, followed by the index values.

    c. **MultiLineIndicesFollowedBySingleImage**: This option creates one row of index values for every image created during the export. If multiple image files are created for a single document, multiple rows of identical index values will be created, each referencing a different page of the document. This will be formatted with index values followed by images.

    d. **MultiLineImagesFollowedByIndices**: One row of index values for every image created during the export. If multiple image files are created for a single document, multiple rows of identical index values will be created, each referencing a different page of the document. This will be formatted with images followed by index values.

    Script constant: TEXT_FILE_ORDER

- **Image Source:** This property determines whether original images (pre-image processing) or modified images (post-image processing) are used for the export. The Pre-Image Processing option uses the original image that has not been modified by image processing, and the Post-Image Processing option uses the image that has been modified by image processing. The default selection, <None>, uses the most recent image prior to exporting. Script constant: IMG_SRC

Hyland OnBase

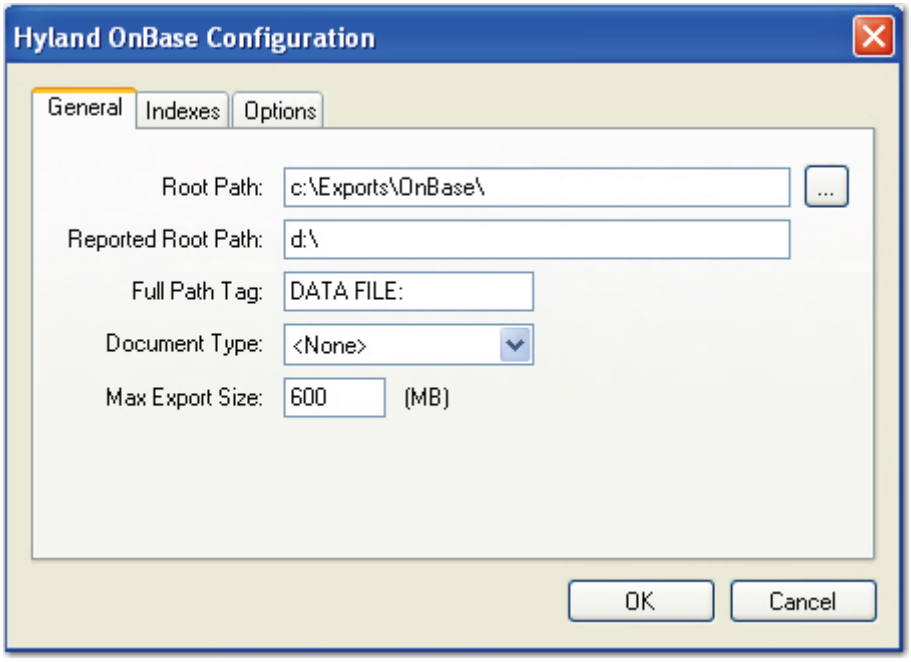
The Hyland OnBase export creates an ASCII text file and single-page TIFF images that can be imported into the Hyland OnBase system. The following settings must be configured in the Hyland OnBase system prior to importing any PaperVision Capture exports:

1. The Document Import Processor separator must be set to **New Line.**
2. The field delimiter must be set to **None.**
3. The field type must be set to **Tagged Fields.**

> **Note:**
>
> If the PaperVision Capture job contains dates, the Hyland OnBase date format settings must match the date field format for that job.

To configure the Hyland OnBase export:

1. From the **Select Custom Code Generator** dialog box, double-click the **Hyland OnBase** export, and the tabbed **Hyland OnBase Configuration** dialog box appears.



Hyland OnBase Configuration

Default values, paths, and other default settings are provided for your reference, and drop-down menus contain only the options specific to your selected generator. In addition, you can browse to the appropriate directories instead of manually entering file paths.

2. Assign the appropriate properties in the **Indexes** and **Options** tabs. Descriptions for constant values appearing in the resulting export script are listed below.
3. When you have finished configuring the export, click **OK**.

Hyland OnBase - General

When you configure properties in the General tab, you can modify constant values that appear in the resulting export script. Properties and resulting script constants that are available for configuration are described below:

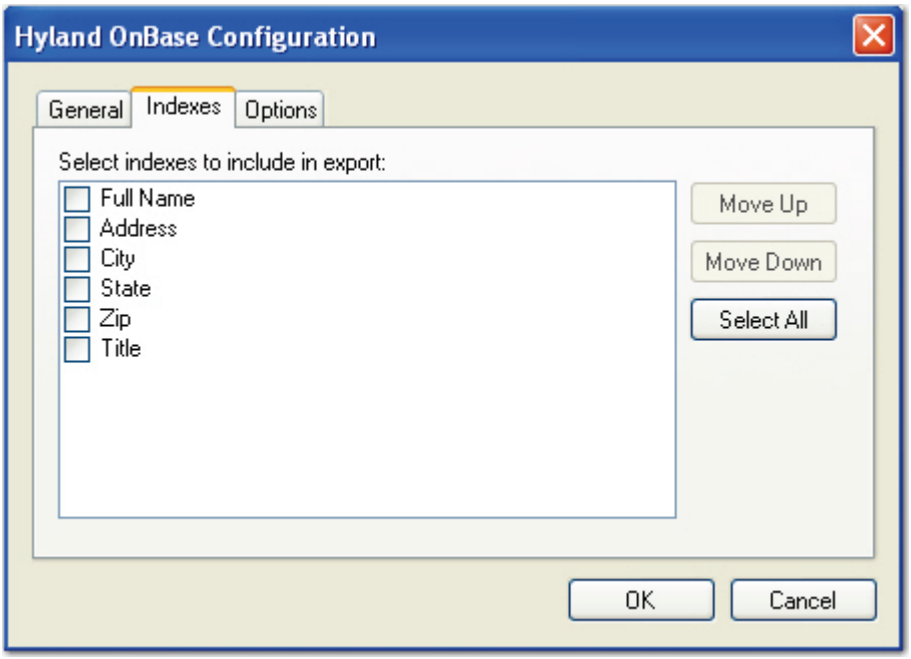
- **Root Path:** This is the location where the exports will be created once you have started the automated process. Script constant: ROOT_PATH
- **Reported Root Path**: The path referenced in the export file originates from this location, not the ROOT_PATH. Script constant: REPORTED_ROOT_PATH
- **Full Path Tag**: This tag precedes the REPORTED_ROOT_PATH in the export file. Script constant: FULL_PATH_TAG
- **Document Type**: This is the specified field name for the index value that should populate the DOCUMENT TYPE field in the export. Script constant: DOCUMENT_TYPE
- **Max Export Size**: This property indicates the maximum export file size (in MB), which defaults to a value of "600". Script constant: MAX_EXPORT_SIZE

> **Note**:
>
> If the Root Path is blank, the export will be written to the directory where the application was installed (e.g., C:\Program Files\Digitech Systems\PaperVision Capture Desktop). If the Reported Root Path is blank, the resulting export script will display a blank value for the REPORTED_ROOT_PATH.

Hyland OnBase - Indexes

In the Indexes tab, you can select the index values that will appear in the export by double-clicking within the appropriate check boxes. Alternatively, click the **Select All** button to include all indexes in the export. You can also click **Deselect All** to remove all selections. To change the order in which the indexes display, single-click an index name (to highlight it), and then click the **Move Up** or **Move Down** buttons.

> **Tip**:
>
> Single-click an index name to move it up or down the list. Double-click an index name to include it in the export.



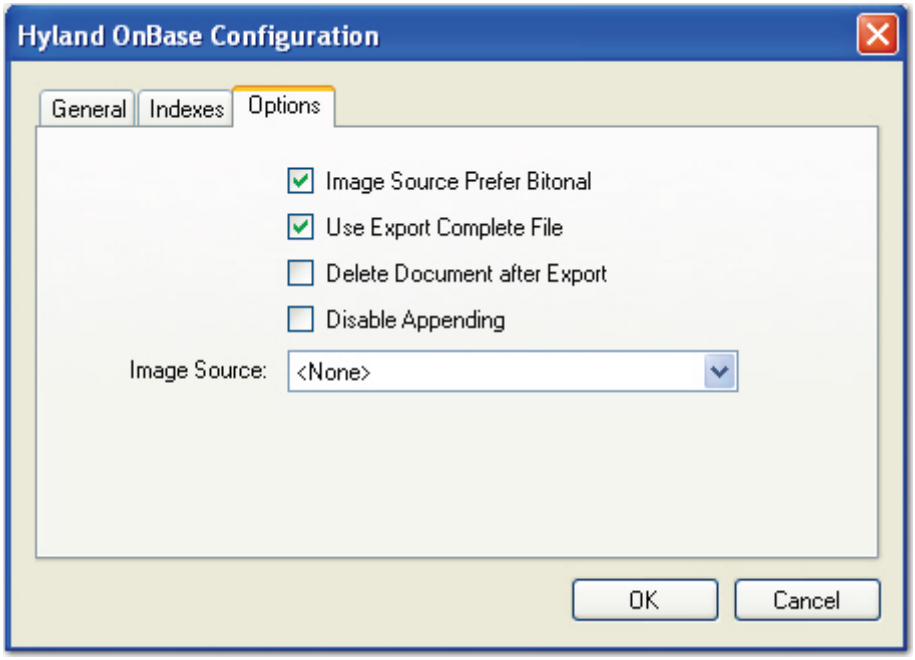
Hyland OnBase Configuration - Indexes

To edit the indexes in the resulting export script, you can modify the INDICES_TO_INCLUDE constant in the export script.

- **Indices to Include**: Your selections determine what index values are included in the export file. In the script, enter the name of the index value(s) between quotation marks, and separate each index value with a comma. Script constant: INDICES_TO_INCLUDE

Hyland OnBase - Options

When you configure properties in the Options tab, you can modify constant values that appear in the resulting export script.



Hyland OnBase Configuration - Options

Properties and resulting script constants that are available for configuration are described below:

- **Image Source Prefer Bitonal:** This property is applicable to dual-stream scanners and determines whether to export bitonal or color images. When set to **True**, which is the default setting, bitonal images will be exported. Script constant: IMG_SRC_PREFER_BITONAL_IMAGES
- **Use Export Complete File**: This property, set to **True** by default, generates an "export.complete" file once an export has reached its maximum file size, so data will no longer be appended to the export. When set to **False**, the "export.complete" file is not generated, so data may be appended to export folders that have not reached their maximum size. Script constant: USE_EXPORT_COMPLETE_FILE
- **Delete Document After Export:** This property specifies whether documents are deleted after they have been exported (set to **False** by default). Script constant: DELETE_DOCUMENT_AFTER_EXPORT
- **Disable Appending:** This property is set to **False** by default. When set to **True**, exported images will not be appended to export folders whose maximum file sizes have not been reached. Script constant: DISABLE_APPENDING

Hyland OnBase - Options

- **Image Source:** This property determines whether original images (pre-image processing) or modified images (post-image processing) are used for the export. The Pre-Image Processing option uses the original image that has not been modified by image processing, and the Post-Image Processing option uses the image that has been modified by image processing. The default selection, **<None>**, uses the most recent image prior to exporting. Script constant: IMG_SRC

Image Only

The Image Only export creates image files that are named after a specific index field. Any subdirectories containing those image files are named after other index fields (optional). Single-page image file formats will be names with an ”-X” at the end of the file name where ”X” denotes the page number.

To configure the Image Only export:

1. From the **Select Custom Code Generator** dialog box, double-click the **Image Only** generator, and the tabbed **Image Only Configuration** dialog box appears.

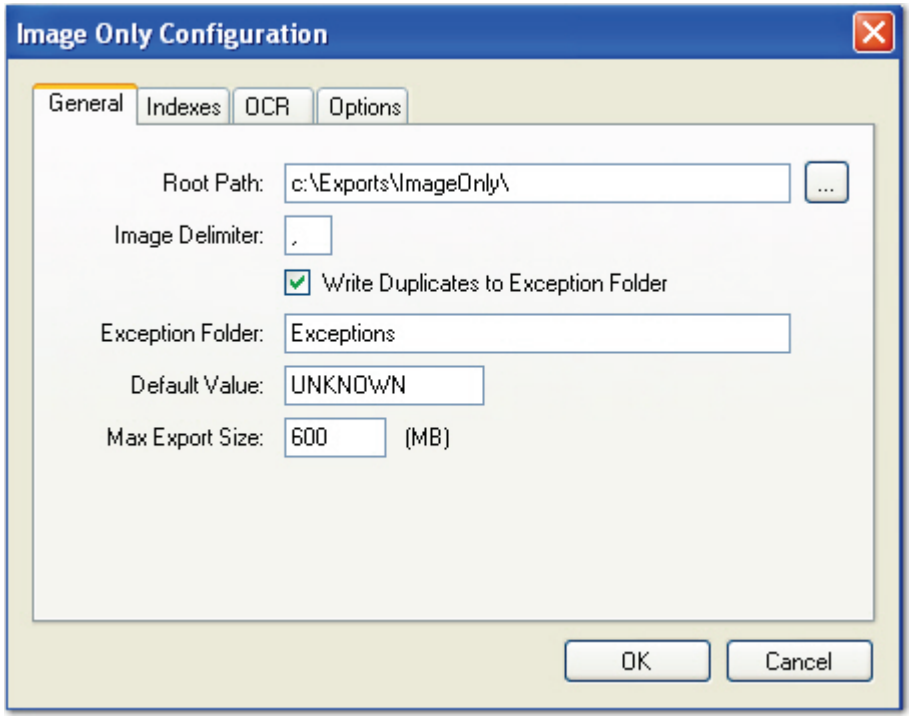


Image Only Configuration - General

Default values, paths, and other default settings are provided for your reference, and drop-down menus contain only the options specific to your selected generator. In addition, you can browse to the appropriate directories instead of manually entering file paths.

2. Assign the appropriate properties in the **Indexes** and **Options** tabs. Descriptions for constant values appearing in the resulting export script are listed below.
3. When you have finished configuring the export, click **OK**.

Image Only - General

When you configure properties in the General tab, you can modify constant values that appear in the resulting export script. Properties and resulting script constants that are available for configuration are described below:

- **Root Path:** This is the location where the exports will be created once you have started the automated process. Script constant: ROOT_PATH
- **Image Delimiter**: This customizable delimiter separates images when exporting using multi-line indexing and converting to single-page images. Script constant: IMAGE_DELIMITER
- **Write Duplicates to Exception Folder**: If duplicate files are created in the same directory during the export and this is set to **False**, PaperVision Capture will not copy the duplicate files into the EXCEPTION_FOLDER directory. If set to **True**, duplicate files are placed in the EXCEPTION_FOLDER instead.

  Script constant: WRITE_DUPLICATES_TO_EXCEPTION_FOLDER

  > **Note**:
  >
  > Files in the EXCEPTION_FOLDER directory will display with "_#" appended to the file name, where "#" is a unique incrementing number starting with "1". This appending process prevents the exception files from being overwritten in the directory.

- **Exception Folder**: If WRITE_DUPLICATES_TO_EXCEPTION_FOLDER is **True** and multiple images with the same file name are created in the same directory, duplicates will be placed in this folder at the ROOT_PATH instead of overwriting the existing file of that name. Script constant: EXCEPTION_FOLDER
- **Default Value**: As the export script executes, invalid characters are stripped from index fields, possibly resulting in blank fields. By default, the resulting DEFAULT_VALUE for these blank fields is defined as "Unknown". Script constant: DEFAULT_VALUE
- **Max Export Size**: This property indicates the maximum export file size (in MB), which defaults to a value of "600". Script constant: MAX_EXPORT_SIZE

  > **Note**:
  >
  > If the Root Path is blank, the export will be written to the directory where the application was installed (e.g., C:\Program Files\Digitech Systems\PaperVision Capture Desktop).

Image Only - Indexes

In the Indexes tab, you can select the index values that will appear in the export by double-clicking within the appropriate check boxes. Alternatively, click the **Select All** button to include all indexes in the export. You can also click **Deselect All** to remove all selections. To change the order in which the indexes display, single-click an index name (to highlight it), and then click the **Move Up** or **Move Down** buttons.

> **Tip**:
>
> Single-click an index name to move it up or down the list. Double-click an index name to include it in the export.

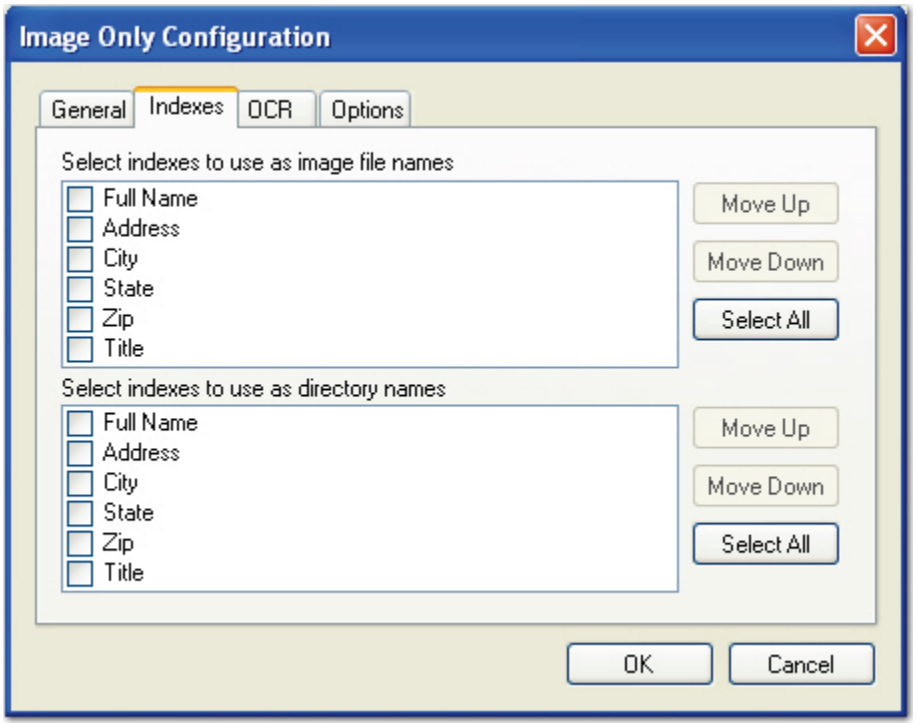


Image Only Configuration - Indexes

To edit the indexes in the resulting export script, you can modify the **IMAGE_INDICES** and **FOLDER_INDICES** constants described below.

- **Image Indices**: Images created during the export will be named based on the index fields mapped in the IMAGE_INDICES field. If multiple index fields are mapped, the IMAGE_DELIMITER will be used to separate the fields in the name of the file. If no fields are mapped, it will use a standard 8-digit incrementing file name. Script constant: IMAGE_INDICES

> **Note**:
>
> Image file names are pulled from a single index field configured in the IMAGE_INDICES field. Any subdirectories are also configured similarly. Index fields should not contain characters that create invalid file names or directory names.

- **Folder Indices**: Images created during the export will be placed in named folders based on the FOLDER_INDICES. The first mapped field will match the first folder, the second mapped field will match the name of the subfolder, etc. If no fields are mapped, the images will be placed directly in the ROOT_PATH. Script constant: FOLDER_INDICES

Image Only - OCR

When you select the OCR Converter in the OCR tab, you can modify constant values that appear in the resulting export script.

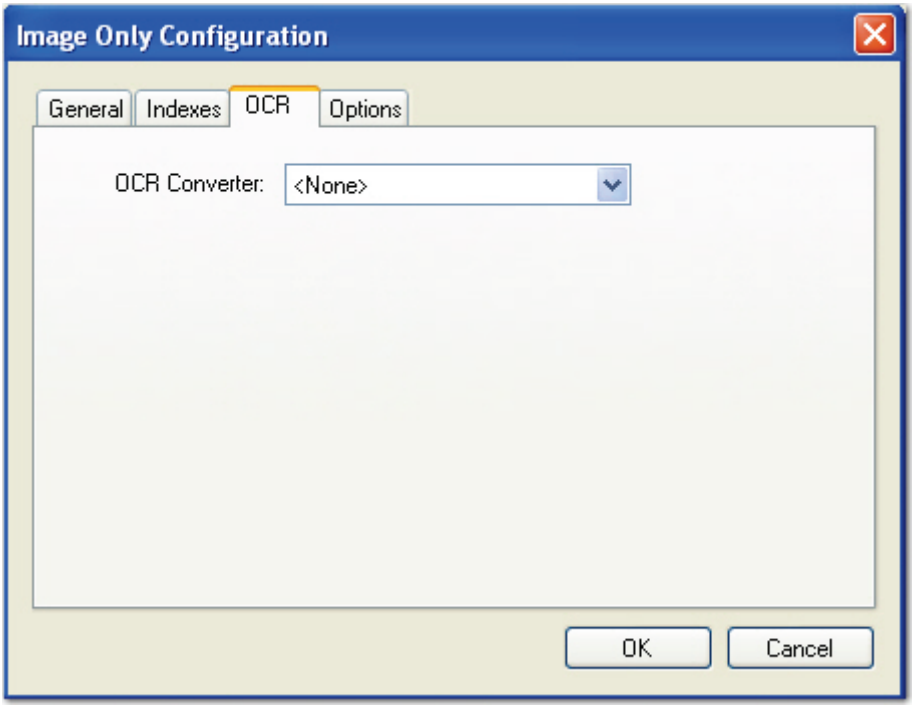


Image Only Configuration - OCR

- **OCR Converter:** This property specifies the OCR converter code, such as text, PDF, etc., whose output format is used to export full-text data. No value is defined by default, so images will be retrieved instead of full-text output files. For a list of converter codes, see the **PVCaptureBatchAPI.chm** help file's **PVBatch.TryGetOCRFiles Method** topic found within the Docs directory where PaperVision Capture Desktop was installed. Script constant: OCR_CONVERTER_NAME

Image Only - Options

When you configure the properties in the Options tab, you can modify constant values that appear in the resulting export script.

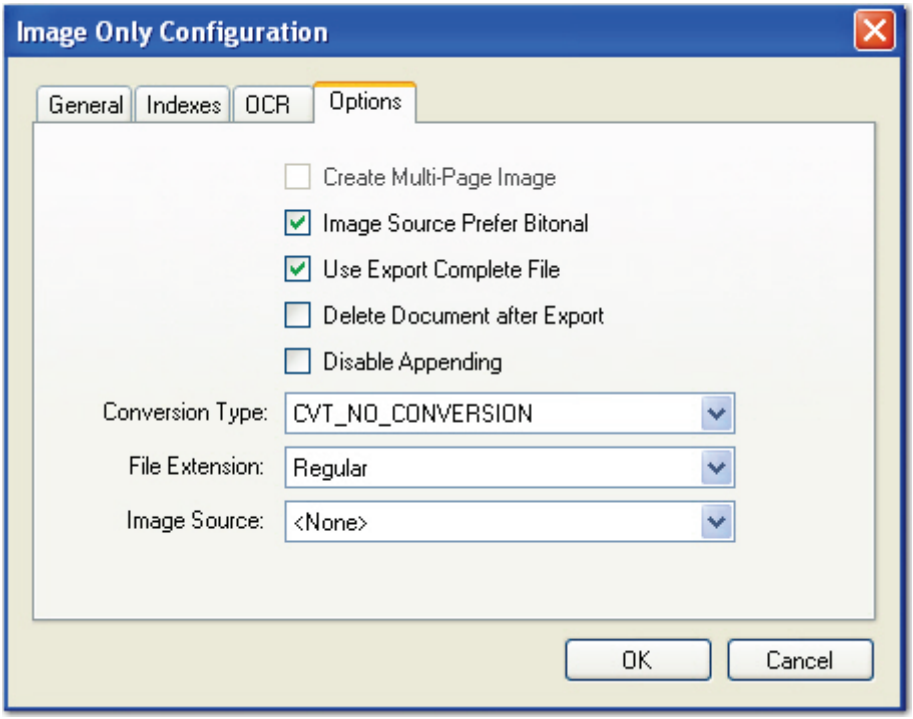


Image Only Configuration - Options

Properties and resulting script constants that are available for configuration are described below:

- **Create Multi-Page Image:** Used in conjunction with CONVERSION_TYPE, this constant determines whether exported images are single-page or multi-page. Script constant: CREATE_MULTI_PAGE_IMAGE
- **Image Source Prefer Bitonal**: This property is applicable to dual-stream scanners and determines whether to export bitonal or color images. When set to **True**, which is the default setting, bitonal images will be exported. Script constant: IMG_SRC_PREFER_BITONAL_IMAGES
- **Use Export Complete File**: This property, set to **True** by default, generates an "export.complete" file once an export has reached its maximum file size, so data will no longer be appended to the export. When set to **False**, the "export.complete" file is not generated, so data may be appended to export folders that have not reached their maximum size. Script constant: USE_EXPORT_COMPLETE_FILE
- **Delete Document After Export:** This property specifies whether documents are deleted after they have been exported (set to **False** by default). Script constant: DELETE_DOCUMENT_AFTER_EXPORT

- **Disable Appending**: This property is set to **False** by default. When set to **True**, exported images will not be appended to export folders whose maximum file sizes have not been reached. Script constant: DISABLE_APPENDING
- **Conversion Type:** This constant determines the type of image file created during the export. The default value, CVT_NO_CONVERSION, does not convert images during the export. If exporting to a format that supports both single and multi-page images, you must set the CREATE_MULTI_PAGE_IMAGE constant to **True** if you want to create multi-page images; otherwise single page images will result. For example, if you set this to CVT_TIFF_G4_MEDJPG, a TIFF image is created during the export. If the source image is binary, it will create a TIFF using Group 4 compression; if the source image is color (JPG or BMP), it will create a TIFF using Medium JPEG compression. See the **Custom Code Configuration** chapter on Enumerations for more information. Script constant: CONVERSION_TYPE
- **File Extension**: This constant determines whether the file extension or page number will be assigned to the file type created during the export.
    a. **Regular**: This option uses the original file extension (.tif, .jpg, etc.).
    b. **PageNumberStartingZero**: This option uses the page number for the file extension, starting with 0 (e.g., -0, -1, etc.).
    c. **PageNumberStartingOne**: This option uses the page number for file extension, starting with 1 (e.g., -1, -2, etc.).
    d. **PageNumberStartingZeroWithPadding**: This option uses the page number for file extension, starting with 000 (e.g., -000, -001, etc.).
    e. **PageNumberStartingOneWithPadding**: This option uses the page number for file extension, starting with 001 (e.g., -001, -002, etc.).

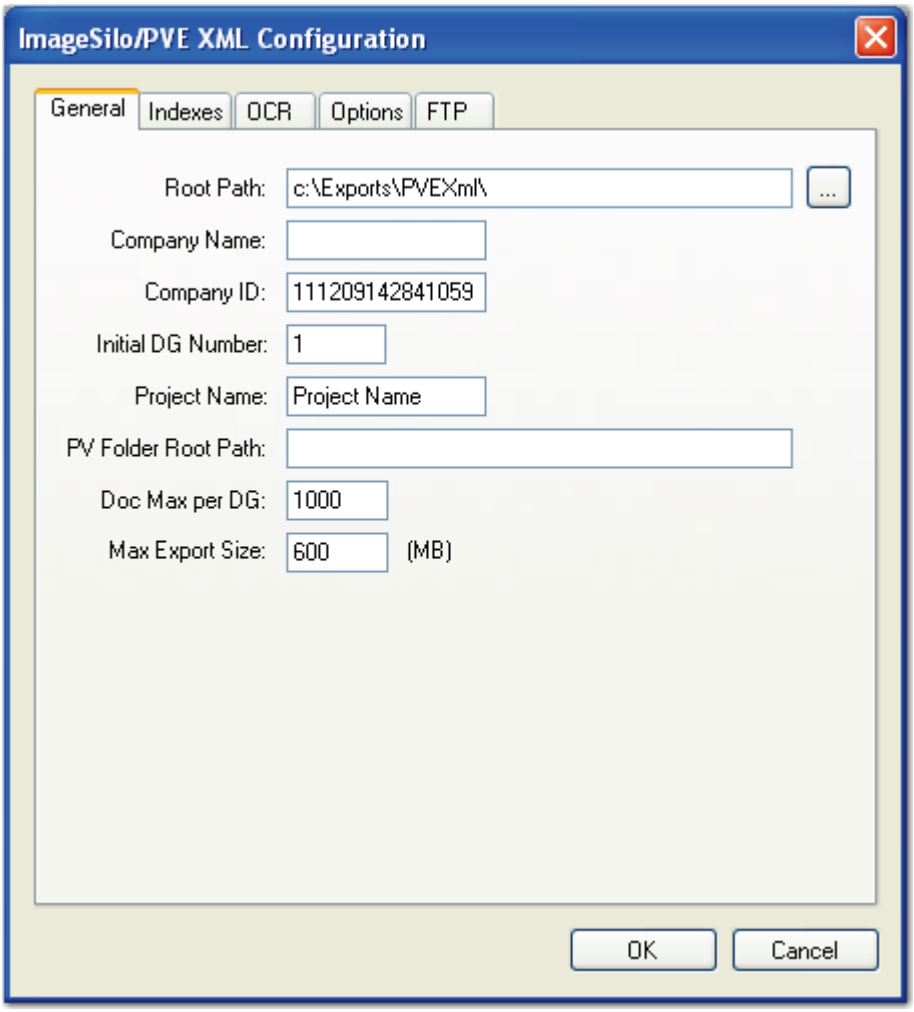
  Script constant: FILE_EXTENSION
- **Image Source:** This constant determines whether original images (pre-image processing) or modified images (post-image processing) are used for the export. The Pre-Image Processing option uses the original image that has not been modified by image processing, and the Post-Image Processing option uses the image that has been modified by image processing. The default selection, <None>, uses the most recent image prior to exporting. Script constant: IMG_SRC

ImageSilo/PVE XML

The ImageSilo/PVE XML export creates an export that can be used to import batches into an ImageSilo or PaperVision Enterprise site.

To configure the ImageSilo/PVE XML export:

1. From the **Select Custom Code Generator** dialog box, double-click the **ImageSilo/PVE XML** export, and the tabbed **ImageSilo/PVE XML Configuration - General** dialog box appears.



ImageSilo/PVE XML Configuration - General

Default values, paths, and other properties are provided for your reference, and drop-down menus contain options specific to your selected generator. In addition, you can browse to some directories or manually enter file paths. Descriptions for all properties are listed below.

2. Proceed to the **Indexes**, **OCR**, **Options**, and **FTP** tabs to modify the appropriate properties. Descriptions for constant values appearing in the resulting export script are listed below.
3. When you have finished configuring the export, click **OK**.

ImageSilo/PVE XML - General

When you configure properties in the General tab, you can modify constant values that appear in the resulting export script. Properties and resulting script constants that are available for configuration are described below.

- **Root Path:** This is the location where the exports will be created once you have started the automated process. Script constant: ROOT_PATH

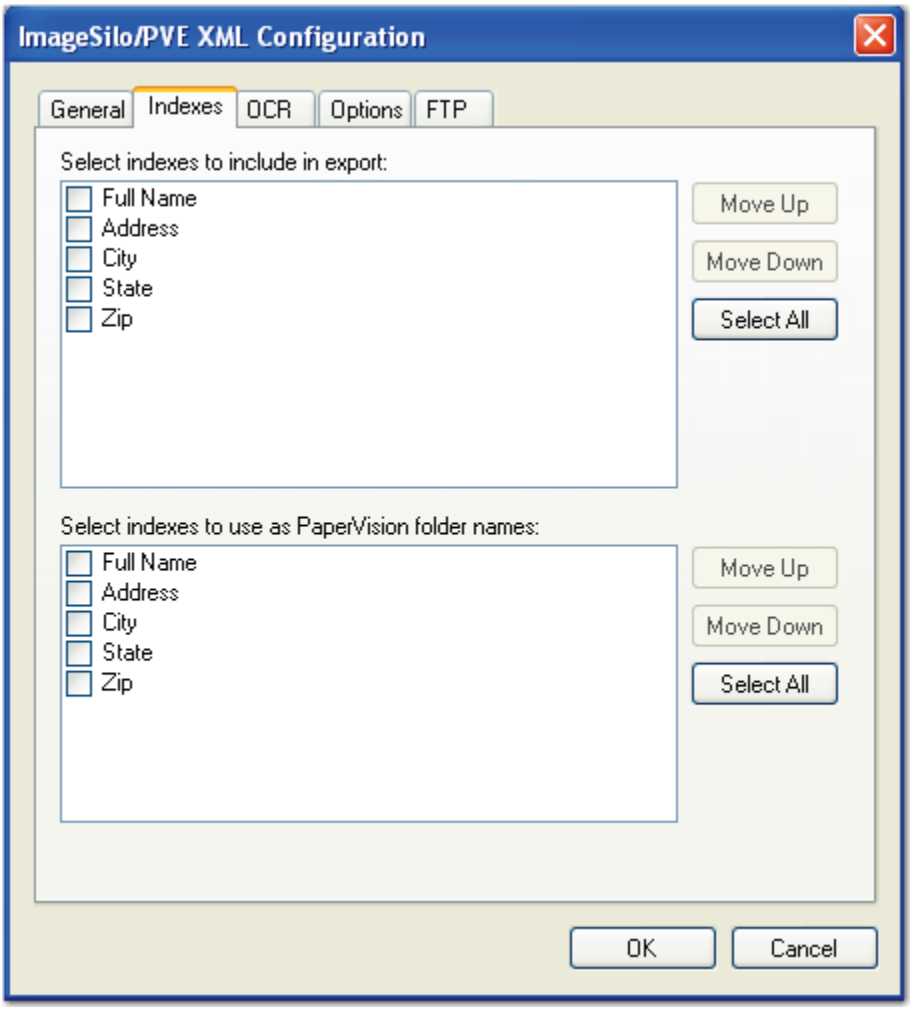
> **Note:**
>
> If the Root Path is blank, the export will be written to the directory where the application was installed (e.g., C:\Program Files\Digitech Systems\PaperVision Capture Desktop).

- **Company Name**: This required field is the name of your company or department, and is blank by default. Script constant: COMPANY_NAME
- **Company ID**: This property is the ID of your company or department. The default value is set to the identifier, "yymmddhhnnssms". Script constant: COMPANY_ID
- **Initial DG Number**: This property represents the initial Data Group number used by PaperVision Enterprise. The default value is "1". Script constant: INITIAL_DATA_GROUP_NUMBER
- **Project Name**: This property indicates the name of your project. The default value is set to "Project Name". Script constant: PROJECT_NAME
- **PV Folder Root Path**: This property specifies the root path containing all folders (used in the Folder view in PaperVision Enterprise). Enter the root path between the quotes (e.g., C:\\Exports\\PVEXml\\FolderRootPath\\). Script constant: PV_FOLDER_ROOT_PATH
- **Doc Max per DG**: This property indicates the maximum number of documents per data group. The default value is "1000", which is the recommended value for XML files**.** Script Constant: DOCUMENT_MAX_PER_DATAGROUP
- **Max Export Size**: This property indicates the maximum export file size (in MB), which defaults to a value of "600". Script constant: MAX_EXPORT_SIZE

ImageSilo/PVE XML - Indexes

In the Indexes tab, you can select the index values that will appear in the export by double-clicking within the appropriate check boxes. Alternatively, click the **Select All** button to include all indexes in the export. You can also click **Deselect All** to remove all selections. To change the order in which the indexes display, single-click an index name (to highlight it), and then click the **Move Up** or **Move Down** buttons.

> **Tip**:
>
> Single-click an index name to move it up or down the list. Double-click an index name to include it in the export.



ImageSilo/PVE XML Configuration - Indexes

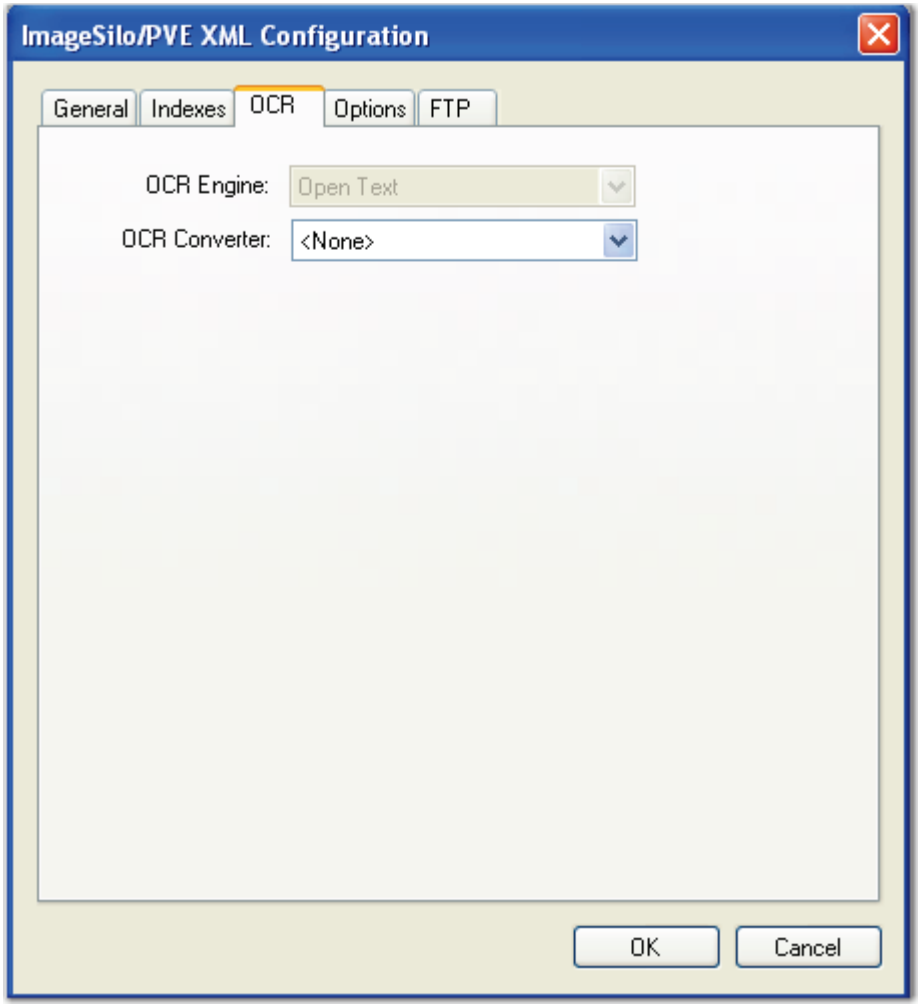
ImageSilo/PVE XML - Indexes

To edit the indexes in the resulting export script, you can modify the **INDICES_TO_INCLUDE** and **PV_FOLDER_INDICES** script constants described below:

- **Select indexes to include in export:** Your selection determines the index values included in the export file. To include all indices, leave the array blank. Script constant: INDICES_TO_INCLUDE
- **Select indexes to use as PaperVision Folder names**: Your selection determines the index value(s) representing each folder (used in the Folder view in PaperVision Enterprise). If you leave the array blank, no index values will be included. Script constant: PV_FOLDER_INDICES

ImageSilo/PVE XML - OCR

When you select the OCR converter in the OCR tab, you can modify the OCR_CONVERTER_CODE constant value that appears in the resulting export script.



ImageSilo/PVE XML Configuration - OCR

- **OCR Converter:** This selection specifies the OCR converter code, including PaperFlow, PaperVision, PDF and Text, whose output format is used to export full-text data. When no value is defined (default setting), both images and associated full-text data will be exported. For example, to export PDF searchable images, the following line in the XML script would read:

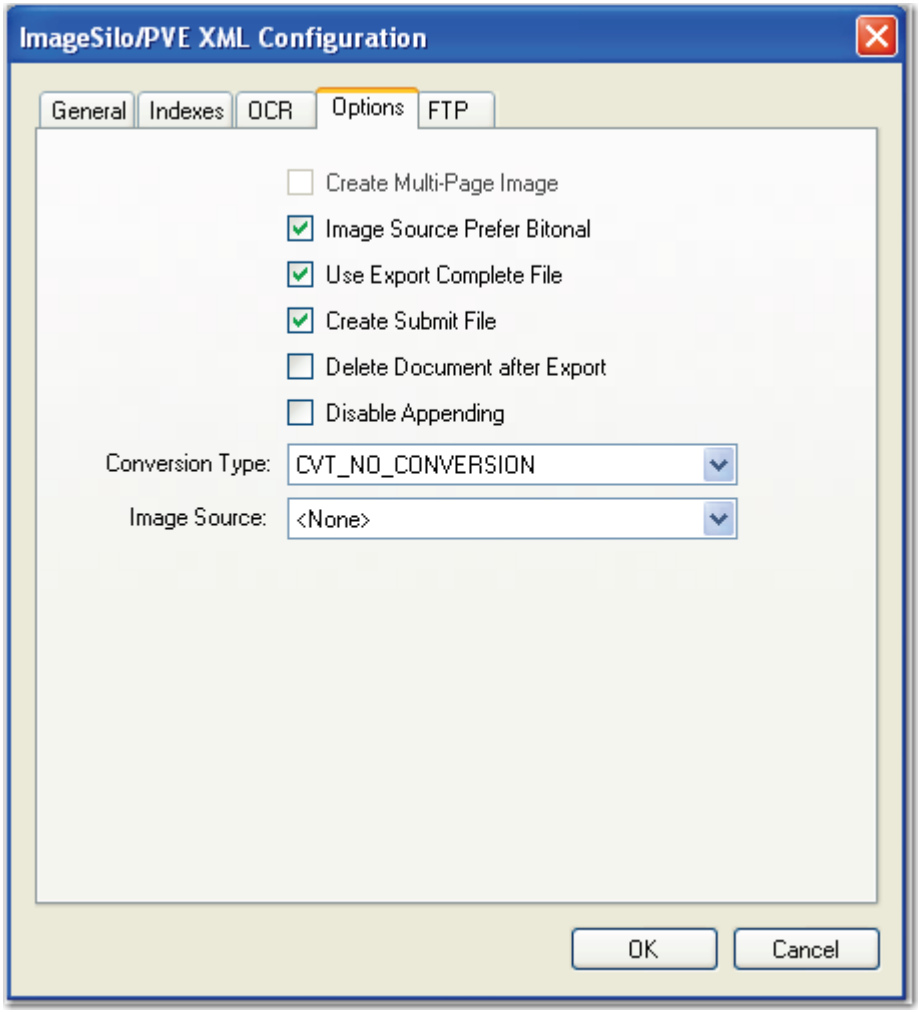
  ```
  private const string OCR_CONVERTER_CODE = "PDF";
  ```

  > **Note**:
  >
  > For a list of converter codes, see the **PVCaptureBatchAPI.chm** help file's **PVBatch.TryGetOCRFiles Method** topic found within the **Docs** directory where PaperVision Capture Desktop was installed.

  Script constant: OCR_CONVERTER_CODE

ImageSilo/PVE XML - Options

When you configure the properties in the Options tab, you can modify constant values that appear in the resulting export script.



ImageSilo/PVE XML Configuration - Options

Properties and resulting script constants available for configuration are described below:

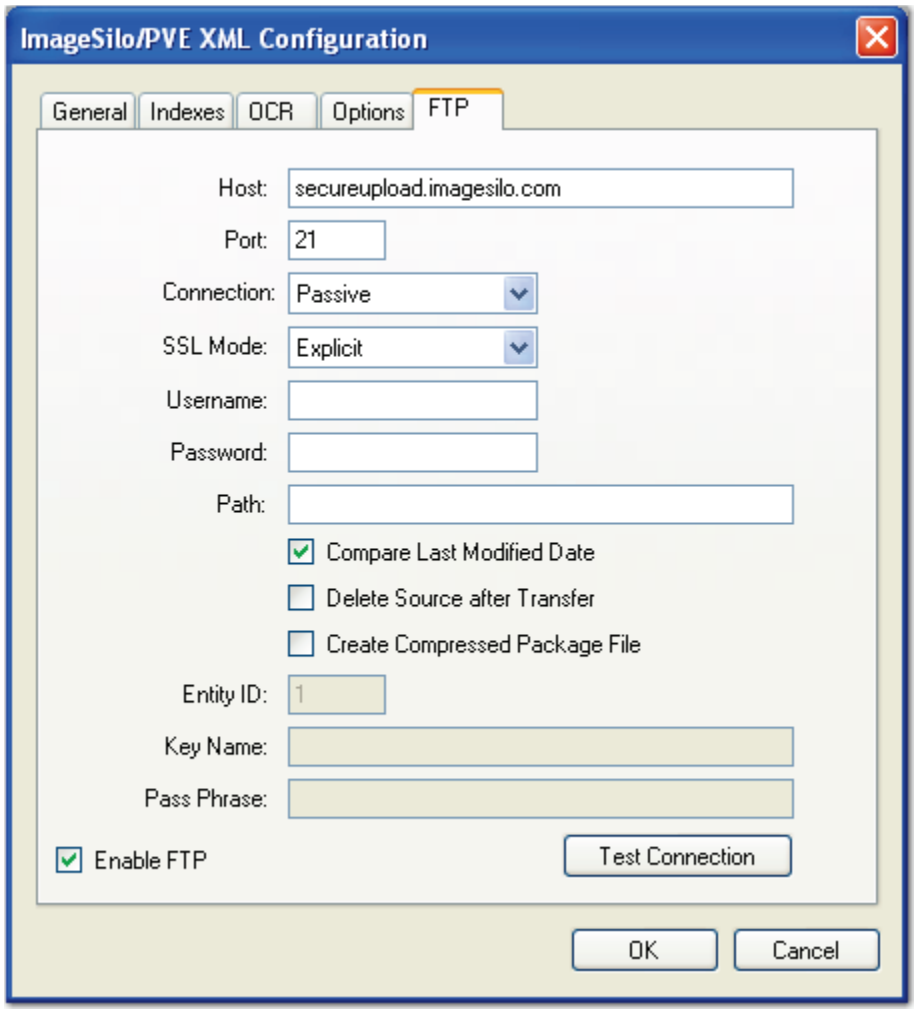
- **Create Multi-Page Image:** Used in conjunction with CONVERSION_TYPE, this constant determines whether exported images are single-page or multi-page. Script constant: CREATE_MULTI_PAGE_IMAGE
- **Image Source Prefer Bitonal:** This property is applicable to dual-stream scanners and determines whether to export bitonal or color images. When set to **True**, which is the default setting, bitonal images will be exported. Script constant: IMG_SRC_PREFER_BITONAL_IMAGES

ImageSilo/PVE XML - Options

- **Use Export Complete File:** This property, set to **True** by default, generates an "export.complete" file once an export has reached its maximum file size, so data will no longer be appended to the export. When set to **False**, the "export.complete" file is not generated, so data may be appended to export folders that have not reached their maximum size. Script constant: USE_EXPORT_COMPLETE_FILE
- **Create Submit File:** Enable this option to automatically generate a DATAGRP.SUBMIT file. If you are importing the data group into PaperVision Enterprise via a Monitored Import Path or via Data Transfer Manager, this file is required before the import can run in PaperVision Enterprise. Script constant: CREATE_SUBMIT_FILE
- **Delete Document After Export**: This property specifies whether documents are deleted after they have been exported (set to **False** by default). Script constant: DELETE_DOCUMENT_AFTER_EXPORT
- **Disable Appending:** This property is set to **False** by default. When set to **True**, exported images will not be appended to export folders whose maximum file sizes have not been reached. Script constant: DISABLE_APPENDING
- **Conversion Type**: This property determines the type of image file created during the export. The default value, CVT_NO_CONVERSION, does not convert images during the export. If exporting to a format that supports both single and multi-page images, you must set the CREATE_MULTI_PAGE_IMAGE constant to **True** if you want to create multi-page images; otherwise single page images will result. For example, if you set this to CVT_TIFF_G4_MEDJPG, a TIFF image is created during the export. If the source image is binary, it will create a TIFF using Group 4 compression; if the source image is color (JPG or BMP), it will create a TIFF using Medium JPEG compression. For a list of file types that can be converted to during the export, see the Enumerations section in the **Custom Code Configuration** chapter. Script constant: CONVERSION_TYPE
- **Image Source:** This property determines whether original images (pre-image processing) or modified images (post-image processing) are used for the export. The Pre-Image Processing option uses the original image that has not been modified by image processing, and the Post-Image Processing option uses the image that has been modified by image processing. The default selection, <None>, uses the most recent image prior to exporting. Script constant: IMG_SRC

ImageSilo/PVE XML - FTP

The ImageSilo/PVE export contains settings that enable you to securely transfer data to an FTP site. Original data files can be transferred in their original state, or they can be placed in a compressed package file. You can configure FTP properties using the Custom Code Generator Wizard, and you can modify constant values in the resulting export script.



ImageSilo/PVE XML Configuration - FTP

Properties and resulting script constants that are available for configuration are described below:

- **Host:** This property specifies the FTP host site name used for the export. Script constant: FTP_HOST
- **Port**: This property specifies the command port number that will be used to connect to the remote FTP server. FTP communications are typically initiated on port 21. Script constant: FTP_PORT

ImageSilo/PVE XML - FTP

- **Connection:** This property specifies the type of connection that will be created. During an active connection, the remote FTP server specifies the data port number that will be used. During a passive connection, PaperVision Capture Desktop specifies the data port number that will be used. Script constant: FTP_CONNECTION
- **SSL Mode**: This export supports fully encrypted FTP communications using SSL (also known as FTPS). The remote FTP server must also support this feature in order to take advantage of the export's capabilities.

    1. **Automatic** SSL indicates that the server will use SSL encryption, but will attempt to automatically determine whether to use Implicit or Explicit SSL.
    2. **Implicit** SSL indicates that the SSL negotiation will start immediately after the FTP connection is established.
    3. **Explicit** SSL indicates the connection will be established in plain text and then explicitly starts the SSL negotiation.
    4. **None** (no SSL encryption) indicates that a standard FTP, non-encrypted session connection will be used.

    Script Constant: FTP_ENCRYPTION
- **Username**: This property specifies the user name that will be used to authenticate to the remote FTP server. Script Constant: FTP_USERNAME
- **Password:** This property specifies the password that will be used to authenticate to the remote FTP server. If desired, you can expose the password in the Script Editor by inserting the tilde character (~) prefix before the password (e.g., ~password). Script constant: FTP_PASSWORD
- **Path**: This property specifies the folder name on the FTP site that stores the exported data. By default, this field is blank, and will write data to the user's home directory as specified by the FTP server.

    For example, other possible paths include the following:

    1. / (root)
    2. FolderA (subdirectory under home directory)
    3. /FolderA (subfolder under root path)

    Script constant: FTP_PATH

ImageSilo/PVE XML - FTP

- **Compare Last Modified Date**: For an operation type related to data groups or package files, the agent will automatically record the last modified date of the file that is being processed. When the same job is processed (and potentially the same file), the last modified date of the previous run is compared to the current, last modified date. If the file has not changed, it will not be processed again.

  For data group processing, this will also allow users to perform incremental data group processing. Once the data group has been changed, any data group files (i.e., images) that have a modified date/time greater than or equal to the previous run's database (i.e., DATAGRP.MDB or DATAGRP.XML) last modified date/time will be processed. Script constant: FTP_COMPARE_LAST_MODIFIED_DATE
- **Delete Source after Transfer**: Once the data has been successfully transferred, this option allows the agent to delete the source data. Script constant: FTP_DELETE_SOURCE_AFTER_EXPORT
- **Create Compressed Package File**: When pushing data groups or files to a remote site, you can increase transfer speed by sending a single, large file rather than hundreds or thousands of small files. This option causes the agent to create a compressed package file that increases transfer speeds and security (if encryption is enabled). Script constant: FTP_ENABLE_PACKAGE
- **Entity ID**: When the export is configured to create compressed package files, the Entity ID and Encryption values are placed into the package file to allow the remote PaperFlow system to decrypt the data. This constant specifies the ID of the remote entity whose encryption key will be used to decrypt the package file. Script constant: FTP_ENTITY_ID
- **Key Name:** This property specifies the name of the encryption key used to decrypt the package file. Script constant: FTP_KEY_NAME
- **Pass Phrase:** For compressed package files, this property specifies a user-defined pass phrase that is passed through a SHA-2 algorithm (Secure Hashing Algorithm) to generate a 256-bit hash. Script constant: FTP_PASS_PHRASE
- **Enable FTP:** This property specifies whether FTP has been enabled for the export. Script constant: FTP_ENABLE

Testing FTP Connections

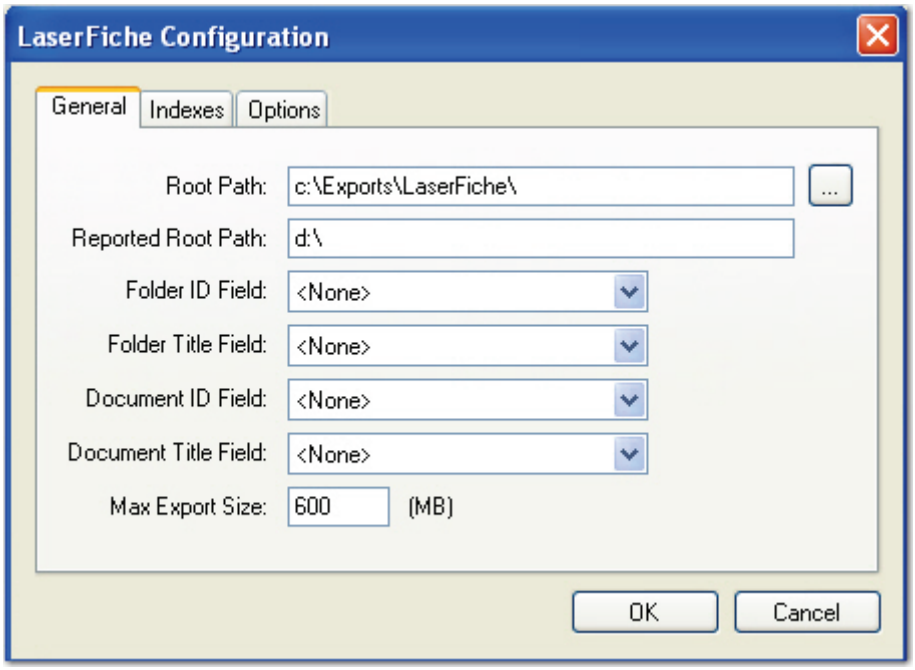
After you have configured the FTP settings, click the **Test Connection** button to ensure the connection is valid. If you successfully connected to the site, click **OK** in the **Success** prompt.

LaserFiche

The LaserFiche export creates an ASCII text file and single-page TIFF images that can be imported into the LaserFiche system using the LaserFiche List Import Feature.

To configure the LaserFiche export:

1. From the **Select Custom Code Generator** dialog box, double-click the **LaserFiche** export, and the tabbed **LaserFiche Configuration** dialog box appears.



LaserFiche Configuration - General

Default values, paths, and other default settings are provided for your reference, and drop-down menus contain only the options specific to your selected generator. In addition, you can browse to the appropriate directories instead of manually entering file paths.

2. Assign the appropriate properties in the **Indexes** and **Options** tabs. Descriptions for constant values appearing in the resulting export script are listed below.
3. When you have finished configuring the export, click **OK**.

LaserFiche - General

When you configure properties in the General tab, you can modify constant values that appear in the resulting export script. Properties and resulting script constants that are available for configuration are described below:

- **Root Path:** This is the location where the exports will be created once you have started the automated process. Script constant: ROOT_PATH
- **Reported Root Path**: The path referenced in the export file originates from this location, not the ROOT_PATH. Script constant: REPORTED_ROOT_PATH

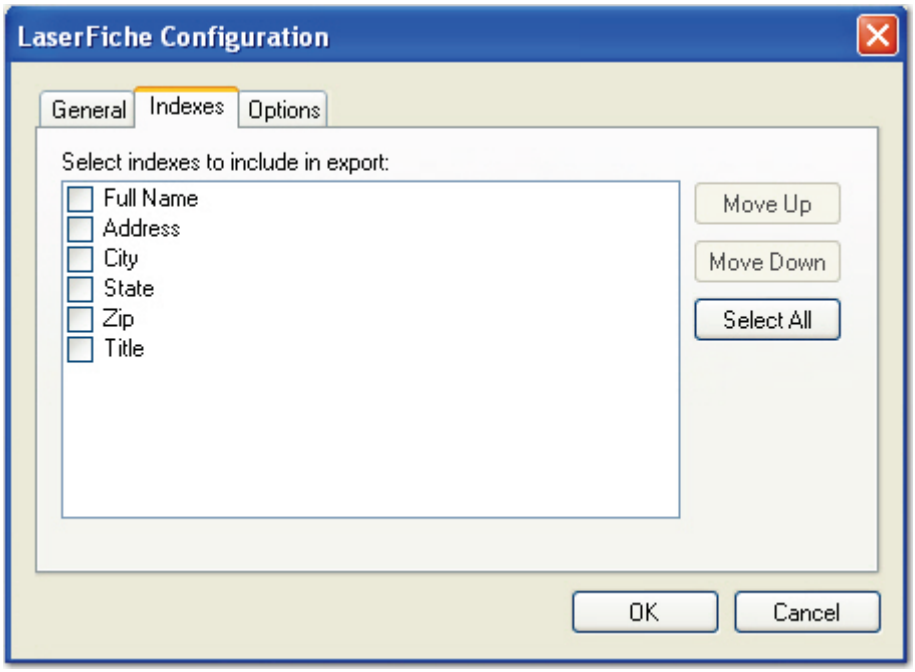
> **Note**:
>
> If the Root Path is blank, the export will be written to the directory where the application was installed (e.g., C:\Program Files\Digitech Systems\PaperVision Capture Desktop). If the Reported Root Path is blank, the resulting export script will display a blank value for the REPORTED_ROOT_PATH.

- **Folder ID Field**: This field name specifies the index value that populates the FOLDER ID field in the export. Script constant: FOLDER_ID_FIELD_NAME
- **Folder Title Field**: This field name specifies the index value that populates the FOLDER TITLE field in the export. Script constant: FOLDER_TITLE_FIELD_NAME
- **Document ID Field**: This field name specifies the index value that populates the DOCUMENT ID field in the export. Script constant: DOCUMENT_ID_FIELD_NAME
- **Document Title Field**: This field name specifies the index value that populates the DOCUMENT TITLE field in the export. Script constant: DOCUMENT_TITLE_FIELD_NAME
- **Max Export Size**: This property indicates the maximum export file size (in MB), which defaults to a value of "600". Script constant: MAX_EXPORT_SIZE

LaserFiche - Indexes

In the Indexes tab, you can select the index values that will appear in the export by double-clicking within the appropriate check boxes. Alternatively, click the **Select All** button to include all indexes in the export. You can also click **Deselect All** to remove all selections. To change the order in which the indexes display, single-click an index name (to highlight it), and then click the **Move Up** or **Move Down** buttons.

> **Tip**:
>
> Single-click an index name to move it up or down the list. Double-click an index name to include it in the export.



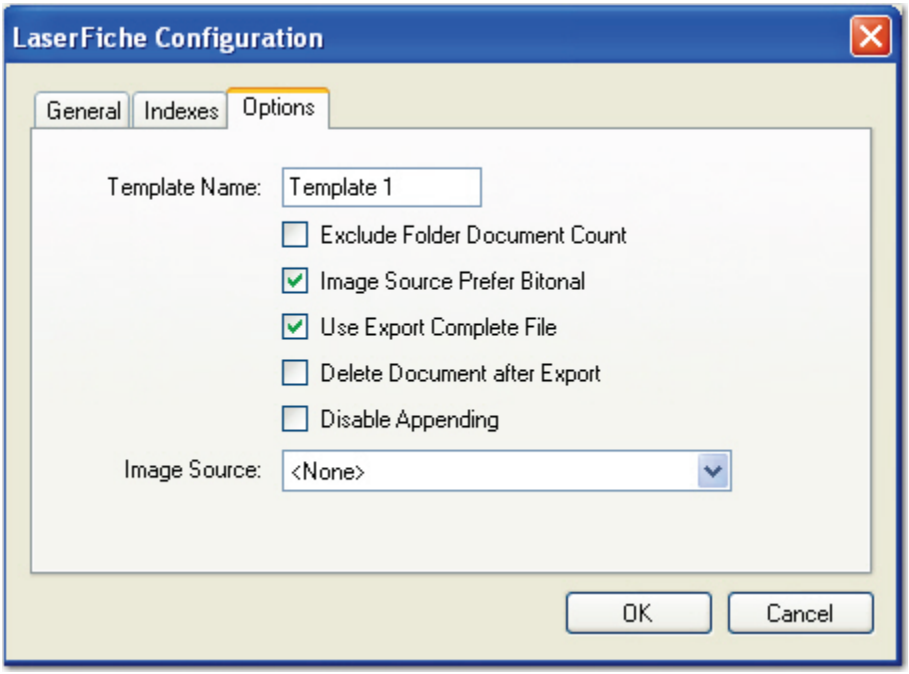
LaserFiche Configuration - Indexes

To edit the indexes in the resulting export script, you can modify the INDICES_TO_INCLUDE constant described below.

- **Indices to Include**: Your selections determine what index values are included in the export file. In the script, enter the name of the index value(s) between quotation marks, and separate each index value with a comma. Script constant: INDICES_TO_INCLUDE

LaserFiche - Options

When you configure properties in the Options tab, you can modify constant values that appear in the resulting export script.



LaserFiche Configuration - Options

Properties and resulting script constants that are available for configuration are described below:

- **Template Name**: This specified value will populate the TEMPLATE NAME field in the export. Script constant: TEMPLATE_NAME
- **Exclude Folder Document Count**: When set to **True**, an incrementing number can be appended to the FOLDER line of the export. It will increment from 1 to 2, etc, for each new document. If set to **False**, no numbers are appended to the FOLDER line of the export. Script constant: EXCLUDE_FOLDER_DOCUMENT_COUNT
- **Image Source Prefer Bitonal:** This property is applicable to dual-stream scanners and determines whether to export bitonal or color images. When set to **True**, which is the default setting, bitonal images will be exported. Script constant: IMG_SRC_PREFER_BITONAL_IMAGES
- **Use Export Complete File:** This property, set to **True** by default, generates an "export.complete" file once an export has reached its maximum file size, so data will no longer be appended to the export. When set to **False**, the "export.complete" file is not generated, so data may be appended to export folders that have not reached their maximum size. Script constant: USE_EXPORT_COMPLETE_FILE
- **Delete Document After Export:** This property specifies whether documents are deleted after they have been exported (set to **False** by default). Script constant: DELETE_DOCUMENT_AFTER_EXPORT

LaserFiche - Options

- **Disable Appending**: This property is set to **False** by default. When set to **True**, exported images will not be appended to export folders whose maximum file sizes have not been reached. Script constant: DISABLE_APPENDING
- **Image Source:** This property determines whether original images (pre-image processing) or modified images (post-image processing) are used for the export. The Pre-Image Processing option uses the original image that has not been modified by image processing, and the Post-Image Processing option uses the image that has been modified by image processing. The default selection, **<None**>, uses the most recent image prior to exporting. Script constant: IMG_SRC

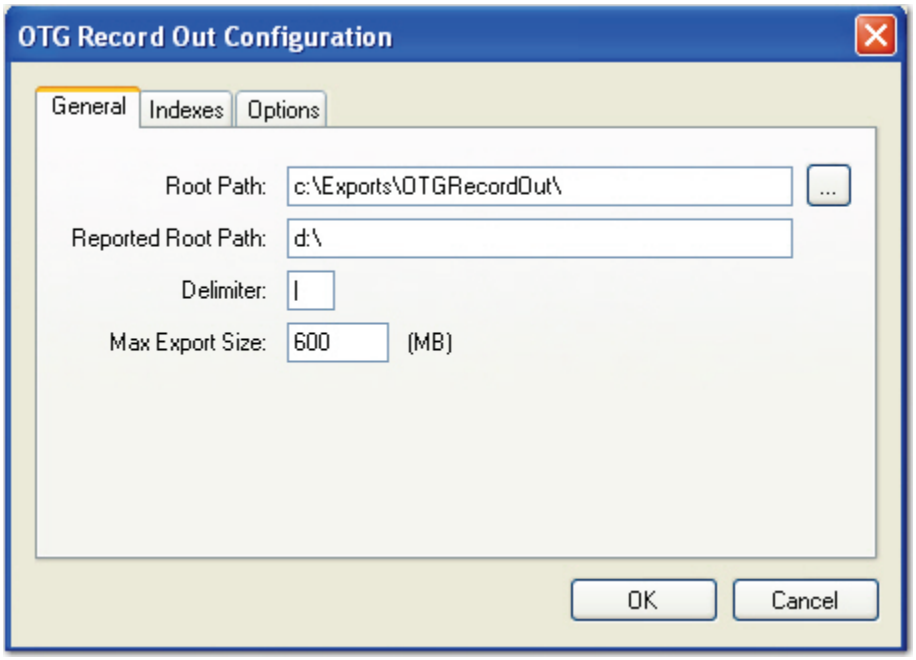
OTG Record Out

The OTG Record Out export creates a valid OTG Record-Out file and its associated images. This can be imported into the OTG Application Extender system using the OTG RDS.

> **Note:**
>
> Ensure that date formats for the PaperVision Capture job correspond with date formats configured in OTG and that all appropriate index values have been defined.

To configure the OTG Record Out export:

1. From the **Select Custom Code Generator** dialog box, double-click the **OTG Record Out** export, and the tabbed **OTG Record Out Configuration** dialog box appears.



OTG Record Out Configuration - General

   Default values, paths, and other default settings are provided for your reference, and drop-down menus contain only the options specific to your selected generator. In addition, you can browse to the appropriate directories instead of manually entering file paths.

2. Assign the appropriate properties in the Indexes and Options tabs. Descriptions for constant values appearing in the resulting export script are listed below.
3. When you have finished configuring the export, click **OK**.

OTG Record Out - General

When you configure properties in the General tab, you can modify constant values that appear in the resulting export script. Properties and resulting script constants that are available for configuration are described below:

- **Root Path**: This is the location where the exports will be created once you have started the automated process. Script constant: ROOT_PATH
- **Reported Root Path**: The path referenced in the export file originates from this location, not the ROOT_PATH. Script constant: REPORTED_ROOT_PATH

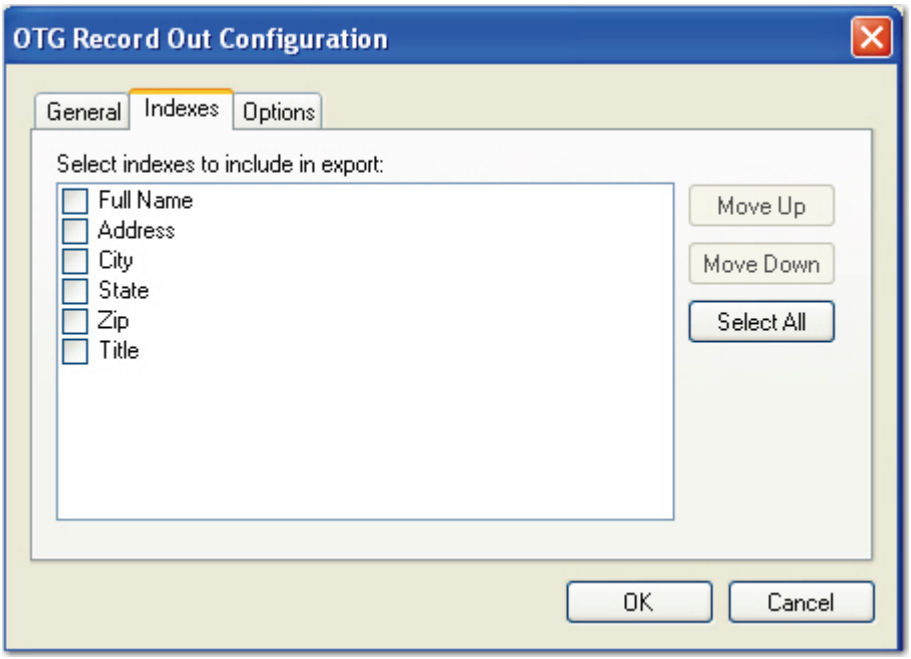
> **Note**:
>
> If the Root Path is blank, the export will be written to the directory where the application was installed (e.g., C:\Program Files\Digitech Systems\PaperVision Capture Desktop). If the Reported Root Path is blank, the resulting export script will display a blank value for the REPORTED_ROOT_PATH.

- **Delimiter**: This property specifies the character that will delimit index values in the export file. Script constant: DELIMITER
- **Max Export Size**: This property indicates the maximum export file size (in MB), which defaults to a value of "600". Script constant: MAX_EXPORT_SIZE

OTG Record Out - Indexes

In the Indexes tab, you can select the index values that will appear in the export by double-clicking within the appropriate check boxes. Alternatively, click the **Select All** button to include all indexes in the export. You can also click **Deselect All** to remove all selections. To change the order in which the indexes display, single-click an index name (to highlight it), and then click the **Move Up** or **Move Down** buttons.

> **Tip**:
>
> Single-click an index name to move it up or down the list. Double-click an index name to include it in the export.



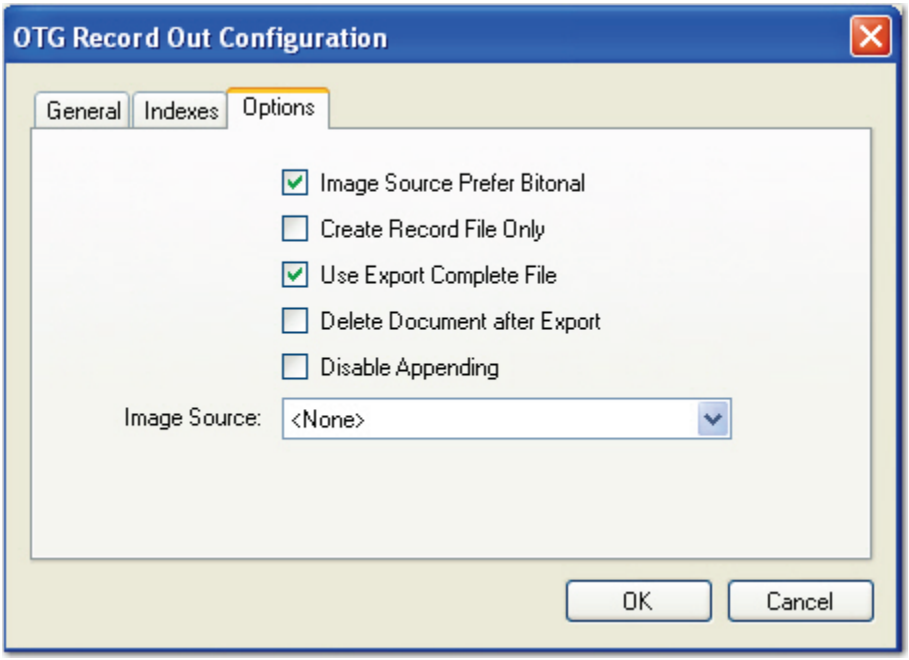
OTG Record Out Configuration - Indexes

To edit the indexes in the resulting export script, you can modify the INDICES_TO_INCLUDE constant described below.

- **Indices to Include**: Your selections determine what index values are included in the export file. In the script, enter the name of the index value(s) between quotation marks, and separate each index value with a comma. Script constant: INDICES_TO_INCLUDE

OTG Record Out - Options

When you configure the properties in the Options tab, you can modify constant values that appear in the resulting export script.



OTG Record Out Configuration - Options

Properties and resulting script constants that are available for configuration are described below:

- **Image Source Prefer Bitonal**: This property is applicable to dual-stream scanners and determines whether to export bitonal or color images. When set to **True**, which is the default setting, bitonal images will be exported. Script constant: IMG_SRC_PREFER_BITONAL_IMAGES
- **Create Record File Only**: If set to **True**, a RECORD.TXT file will be created, but no images will be created during the export. Script Constant: CREATE_RECORD_FILE_ONLY
- **Use Export Complete File**: This property, set to **True** by default, generates an "export.complete" file once an export has reached its maximum file size, so data will no longer be appended to the export. When set to **False**, the "export.complete" file is not generated, so data may be appended to export folders that have not reached their maximum size. Script constant: USE_EXPORT_COMPLETE_FILE
- **Delete Document After Export**: This property specifies whether documents are deleted after they have been exported (set to **False** by default). Script constant: DELETE_DOCUMENT_AFTER_EXPORT
- **Disable Appending**: This property is set to **False** by default. When set to **True**, exported images will not be appended to export folders whose maximum file sizes have not been reached. Script constant: DISABLE_APPENDING

OTG Record Out - Options

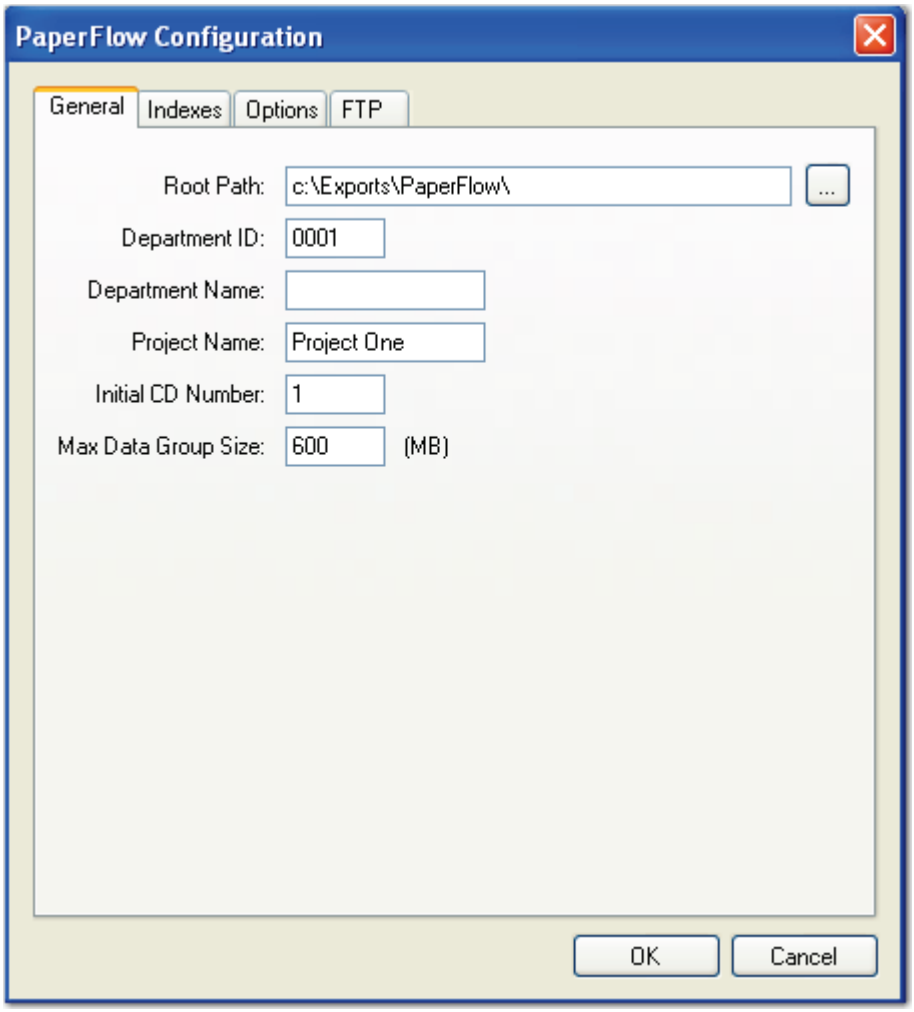
- **Image Source**: This property determines whether original images (pre-image processing) or modified images (post-image processing) are used for the export. The Pre-Image Processing option uses the original image that has not been modified by image processing, and the Post-Image Processing option uses the image that has been modified by image processing. The default selection, **<None>**, uses the most recent image prior to exporting. Script constant: IMG_SRC

PaperFlow

The PaperFlow export can be used to import batches into PaperFlow, OCRFlow™, or QCFlow™.

To configure the PaperFlow export:

1. From the **Select Custom Code Generator** dialog box, double-click the **PaperFlow** export, and the tabbed **PaperFlow Configuration - General** dialog box appears.



PaperFlow Configuration - General

Default values, paths, and other default settings are provided for your reference, and drop-down menus contain only the options specific to your selected generator. In addition, you can browse to the appropriate directories instead of manually entering file paths.

2. Assign the appropriate properties in the **Indexes**, **Options**, and **FTP** tabs. Descriptions for constant values appearing in the resulting export script are listed below.
3. When you have finished configuring the export, click **OK**.

PaperFlow - General

When you configure properties in the General tab, you can modify constant values that appear in the resulting export script. Properties and resulting script constants that are available for configuration are described below:

- **Root Path:** This is the location where the exports will be created once you have started the automated process. Script constant: ROOT_PATH

  > **Note**:
  >
  > If the Root Path is blank, the export will be written to the directory where the application was installed (e.g., C:\Program Files\Digitech Systems\PaperVision Capture Desktop).

- **Department ID**: This value is uniquely assigned to each client for which the export is generated. The default value is "0001". This value must be four characters in length, consisting only of numbers and letters. Special characters (e.g., space, "^", ":", "*", "?", "<", ">", etc.) are invalid. Script constant: DEPT_ID
- **Department Name:** This value is uniquely assigned to each client or department and is a required field. The default value is blank. Script constant: DEPT_NAME
- **Project Name:** This value is uniquely assigned to each client or department. The default value is "Project". Script constant: PROJECT_NAME
- **Initial CD Number:** This value can be used to export to a CD. The default value is "1". Script constant: INITIAL_CD_NUMBER

  If you change this value after you have already performed a PaperFlow export, the new value will not be reflected in exported data groups unless you remove the ”//” comment codes. The ”Reset CD Number?” code should appear as follows in the export script:

  ```
  if (!PVUtilities.TrySetCustomCounter(DEPT_ID + "_" +
  PROJECT_NAME, INITIAL_CD_NUMBER, out error))

  throw (new Exception("Unable to reset custom counter: " +
  error.Message));
  ```

  After you remove the comment codes, you must run the export to reset the counter. The next data group that is created will reflect your new INITIAL_CD_NUMBER value. Lastly, to ensure that new data groups increment properly from the new INITIAL_CD_NUMBER, you must insert the ”\\” comment codes once again:

  ```
  //if (!PVUtilities.TrySetCustomCounter(DEPT_ID + "_" +
  PROJECT_NAME, INITIAL_CD_NUMBER, out error))

  //throw (new Exception("Unable to reset custom counter: " +
  error.Message));
  ```

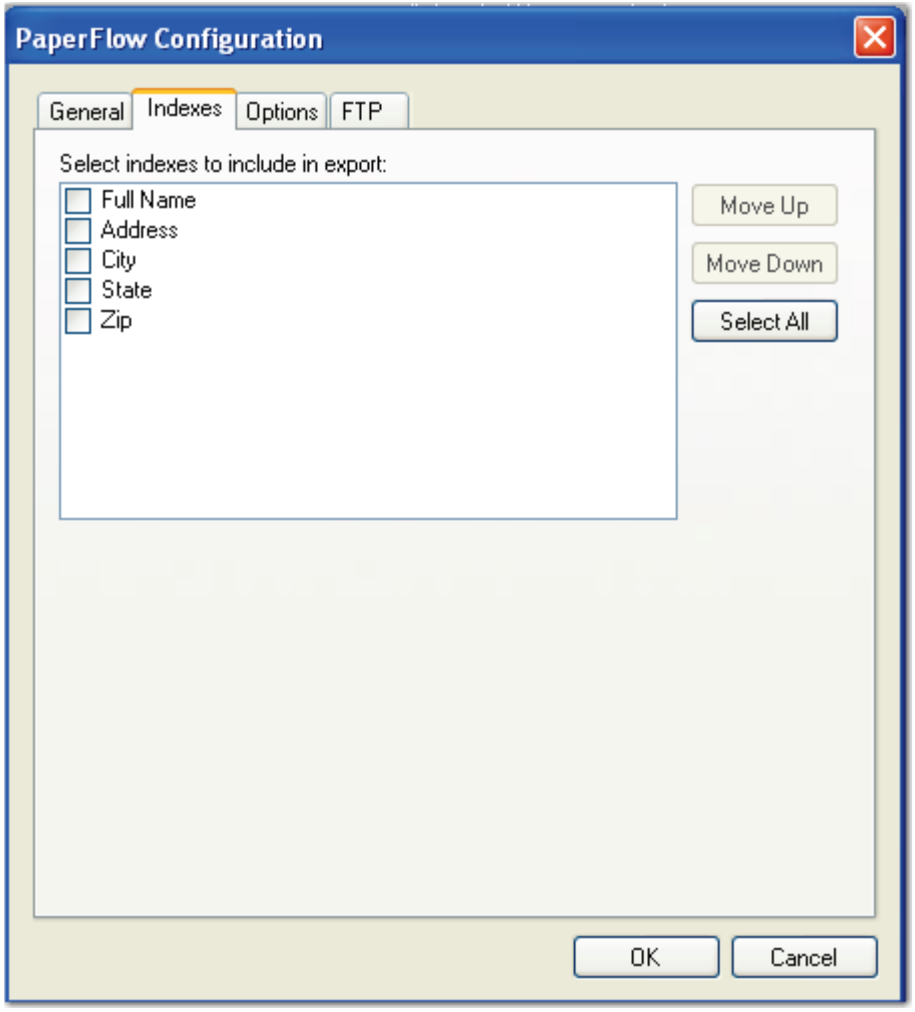
> **Note**:
>
> You must export to a directory that does not contain existing data groups. Otherwise, the system will attempt to append to data groups whose maximum size has not been reached, and the new INITIAL_CD_NUMBER value may be ignored or other unexpected results may occur.

- **Max Data Group Size**: This indicates the maximum size (in MB) that a data group can reach before a new data group begins. The default value is "600", the standard CD size. Script constant: MAX_DATAGROUP_SIZE

PaperFlow - Indexes

In the Indexes tab, you can select the index values that will appear in the export by double-clicking within the appropriate check boxes. Alternatively, click the **Select All** button to include all indexes in the export. You can also click **Deselect All** to remove all selections. To change the order in which the indexes display, single-click an index name (to highlight it), and then click the **Move Up** or **Move Down** buttons.

> **Tip**:
>
> Single-click an index name to move it up or down the list. Double-click an index name to include it in the export.



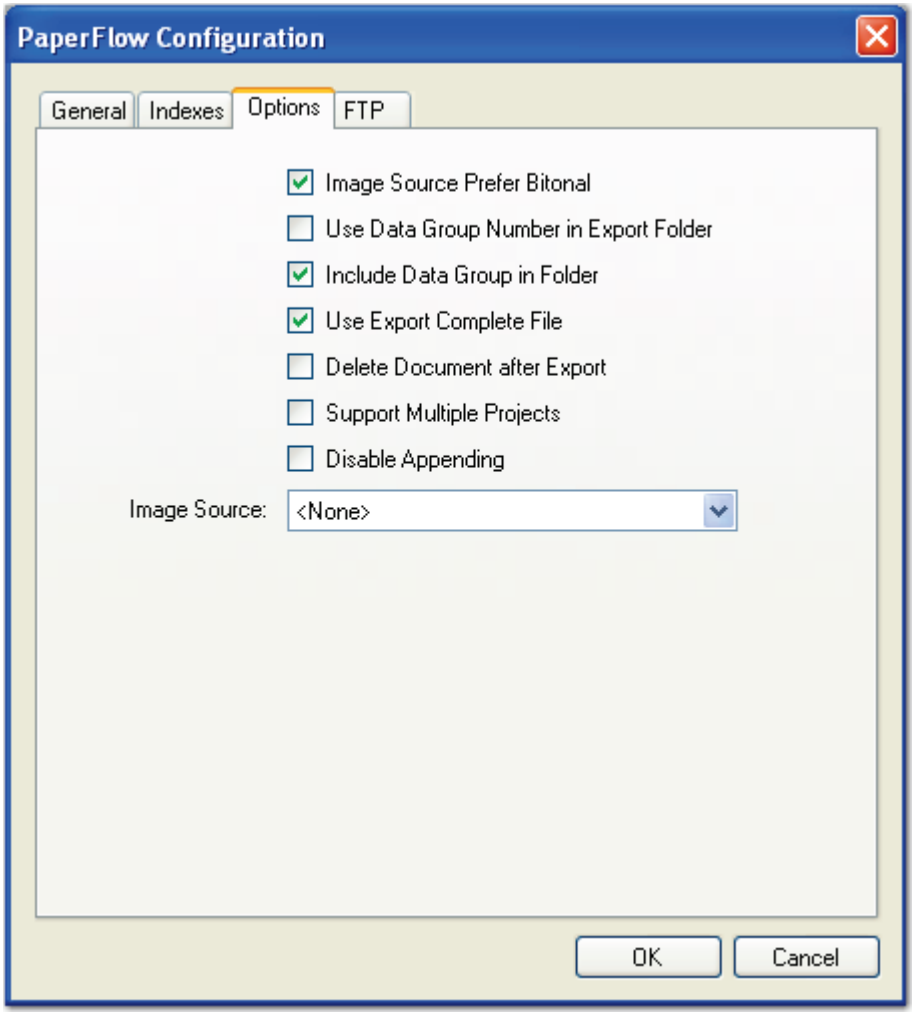
PaperFlow Configuration - Indexes

To edit the indexes in the resulting export script, you can modify the INDICES_TO_INCLUDE constant described below:

- **Indices to Include**: Your selections determine what index values are included in the export file. In the script, enter the name of the index value(s) between quotation marks, and separate each index value with a comma. To include all indices, leave the array blank. Script constant: INDICES_TO_INCLUDE

PaperFlow - Options

When you configure properties in the Options tab, you can modify constant values that appear in the resulting export script.



PaperFlow Configuration - Options

Properties and resulting script constants that are available for configuration are described below:

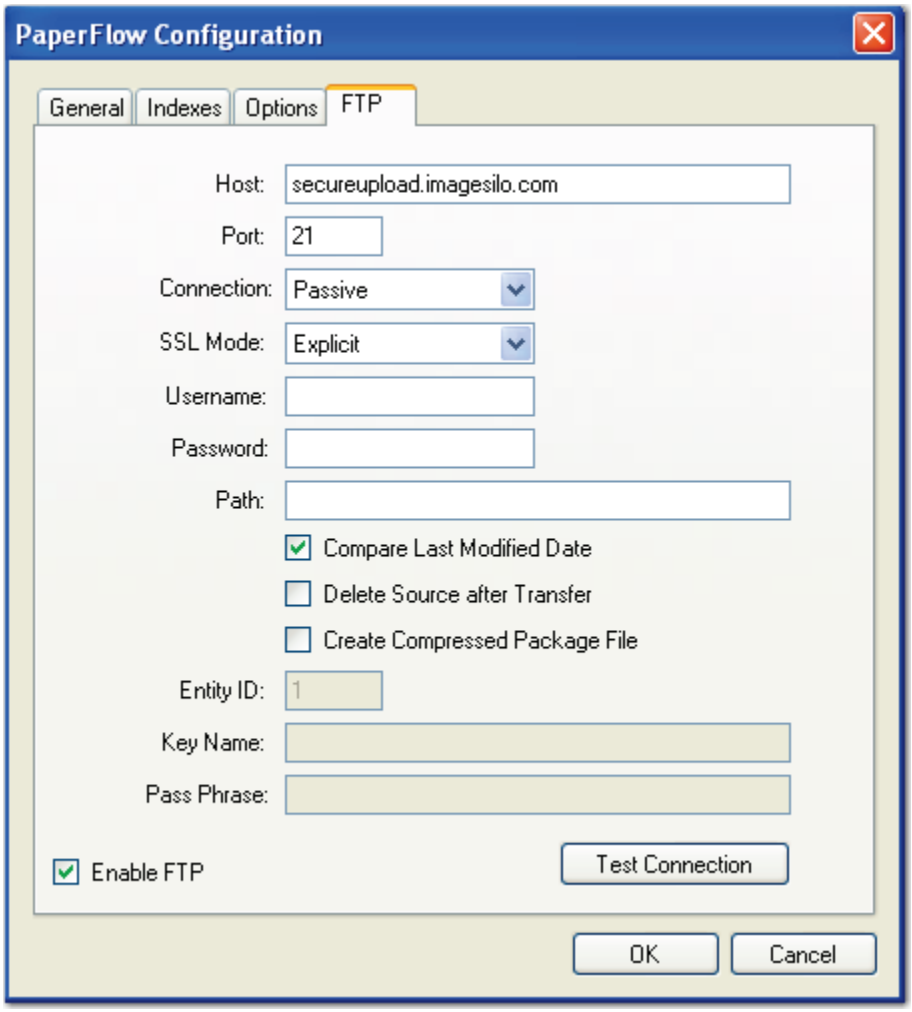
- **Image Source Prefer Bitonal:** This property is applicable to dual-stream scanners and determines whether to export bitonal or color images. When set to **True**, which is the default setting, bitonal images will be exported. Script constant: IMG_SRC_PREFER_BITONAL_IMAGES
- **Use Data Group Number in Export Folder**: When set to **True**, the parent export directory will be organized by data group name instead of export number. Script constant: USE_DATAGROUP_NUMBER_IN_EXPORT_FOLDER

PaperFlow - Options

- **Include Data Group in Folder**: When set to **True** (default setting), a folder named "DATAGRP" is created under the directory in which the export data is copied (e.g.,<root>\<export#>\DATAGRP\<export data>). When set to **False**, the "DATAGRP" folder is not created. Script constant: INCLUDE_DATAGROUP_IN_FOLDER
- **Use Export Complete File:** This property, set to **True** by default, generates an "export.complete" file once an export has reached its maximum file size, so data will no longer be appended to the export. When set to **False**, the "export.complete" file is not generated, so data may be appended to export folders that have not reached their maximum size. Script constant: USE_EXPORT_COMPLETE_FILE
- **Delete Document After Export**: This property specifies whether documents are deleted after they have been exported (set to False by default). Script constant: DELETE_DOCUMENT_AFTER_EXPORT
- **Support Multiple Projects:** When set to **True**, multiple Department IDs will be exported to the same folder, creating a single MDB file. When set to **False** (default setting), one Department ID will be exported to a single folder. Script Constant: SUPPORT_MULTIPLE_PROJECTS
- **Disable Appending:** This property is set to **False** by default. When set to **True**, exported images will not be appended to export folders whose maximum file sizes have not been reached. Script constant: DISABLE_APPENDING
- **Image Source:** This property determines whether original images (pre-image processing) or modified images (post-image processing) are used for the export. The Pre-Image Processing option uses the original image that has not been modified by image processing, and the Post-Image Processing option uses the image that has been modified by image processing. The default selection, <None>, uses the most recent image prior to exporting. Script constant: IMG_SRC

PaperFlow - FTP

The PaperFlow and ImageSilo/PVE XML exports contain FTP settings that enable you to securely transfer data to FTP sites. Original data files can be transferred in their original state, or they can be placed in a compressed package file. In the Auto Process page, the FTP Enabled option will be selected after you have configured the PaperFlow or ImageSilo/PVE XML export for FTP. When you configure properties in the FTP tab, you can modify constant values that appear in the resulting export script.



PaperFlow Configuration - FTP

Properties and resulting script constants that are available for configuration are described below:

- **Host**: This property specifies the FTP host site name used for the export. Script constant: FTP_HOST
- **Port**: This property specifies the command port number that will be used to connect to the remote FTP server. FTP communications are typically initiated on port 21. Script constant: FTP_PORT

PaperFlow - FTP

- **Connection**: This property specifies the type of connection that will be created. During an active connection, the remote FTP server specifies the data port number that will be used. During a passive connection, PaperVision Capture specifies the data port number that will be used. Script constant: FTP_CONNECTION
- **SSL Mode**: This export supports fully encrypted FTP communications using SSL (also known as FTPS). The remote FTP server must also support this feature in order to take advantage of the export's capabilities.

  1. **Automatic SSL** indicates that the server will use SSL encryption, but will attempt to automatically determine whether to use Implicit or Explicit SSL.
  2. **Implicit SSL** indicates that the SSL negotiation will start immediately after the FTP connection is established.
  3. **Explicit SSL** indicates the connection will be established in plain text and then explicitly starts the SSL negotiation.
  4. **None** (no SSL encryption) indicates that a standard FTP, non-encrypted session connection will be used.

  Script constant: FTP_ENCRYPTION
- **Username**: This property specifies the user name that will be used to authenticate to the remote FTP server. Script constant: FTP_USERNAME
- **Password**: This property specifies the password that will be used to authenticate to the remote FTP server. Script constant: FTP_PASSWORD
- **Path**: This property specifies the folder name on the FTP site that stores the exported data. By default, this field is blank, and will write data to the user's home directory as specified by the FTP server.

  For example, other possible paths include the following:

  1. / (root)
  2. FolderA (subdirectory under home directory)
  3. /FolderA (subfolder under root path)

  Script constant: FTP_PATH
- **Compare Last Modified Date**: For an operation type related to data groups or package files, the agent will automatically record the last modified date of the file that is being processed. When the same job is processed (and potentially the same file), the last modified date of the previous run is compared to the current, last modified date. If the file has not changed, it will not be processed again.

  For data group processing, this will also allow users to perform incremental data group processing. Once the data group has been changed, any data group files (i.e., images) that have a modified date/time greater than or equal to the previous run's database (i.e., DATAGRP.MDB or DATAGRP.XML) last modified date/time will be processed. Script constant: FTP_COMPARE_LAST_MODIFIED_DATE

PaperFlow - FTP

- **Delete Source After Transfer**: Once the data has been successfully transferred, this property allows the agent to delete the source data. Script constant: FTP_DELETE_SOURCE_AFTER_EXPORT
- **Create Compressed Package File**: When pushing data groups or files to a remote site, you can increase transfer speed by sending a single, large file rather than hundreds or thousands of small files. This option causes the agent to create a compressed package file that increases transfer speeds and security (if encryption is enabled). Script constant: FTP_ENABLE_PACKAGE
- **Entity ID**: When the export is configured to create compressed package files, the Entity ID and Encryption values are placed into the package file to allow the remote PaperFlow system to decrypt the data. This constant specifies the ID of the remote entity whose encryption key will be used to decrypt the package file. Script constant: FTP_ENTITY_ID
- **Key Name**: This property specifies the name of the encryption key used to decrypt the package file. Script constant: FTP_KEY_NAME
- **Pass Phrase**: For compressed package files, this property specifies a user-defined pass phrase that is passed through a SHA-2 algorithm (Secure Hashing Algorithm) to generate a 256-bit hash. Script constant: FTP_PASS_PHRASE
- **Enable FTP:** This property specifies whether FTP has been enabled for the export. Script constant: FTP_ENABLE

Testing FTP Connections

After you have configured the FTP settings, click the **Test Connection** button to ensure the connection is valid.

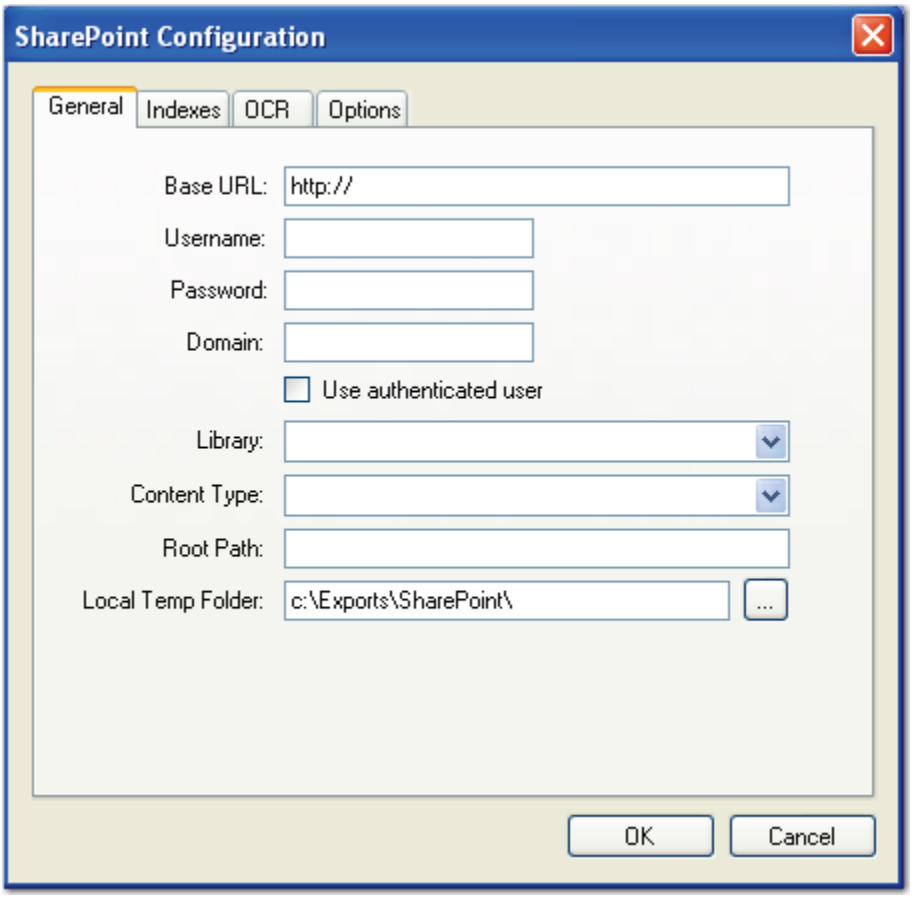
SharePoint

The SharePoint export creates a file that can be used to export PaperVision Capture Desktop data into a Microsoft® SharePoint® site.

> **Note**:
>
> Only Microsoft SharePoint 2007 (on Windows Server 2003 or 2008) or Microsoft SharePoint 2010 (on Windows Server 2008) are supported for this export.

To configure the SharePoint export:

1. From the **Select Custom Code Generator** dialog box, double-click the SharePoint export, and the tabbed **SharePoint Configuration - General** dialog box appears.



SharePoint Configuration - General

2. You must configure all properties (described in the next section) in the **General** tab.
3. Proceed to the **Indexes** tab. If you entered valid SharePoint data, you can map PaperVision Capture index field names to SharePoint columns.

> **Note**:
>
> An error message will inform you when you have entered invalid SharePoint data.

4. If applicable, map the appropriate index field names to SharePoint columns.
5. Proceed to the **OCR** and **Options** tabs to modify the appropriate properties that are described below.
6. When you have finished configuring the export, click **OK**.

When you configure properties in the General tab, you can modify constant values that appear in the resulting export script. Properties and resulting script constants that are available for configuration are described below:

- **Base URL**: This property specifies the Microsoft SharePoint host site name and port used for the export. Script constant: SHAREPOINT_BASE_URL
- **Username**: This property specifies the Microsoft SharePoint user name. Script constant: SHAREPOINT_USERNAME
- **Password:** This property specifies the Microsoft SharePoint user's password. By default, the SharePoint password is encrypted in the Script Editor. If desired, you can expose the password in the Script Editor by inserting the tilde (~) prefix before the password (e.g., ~password). Script Constant: SHAREPOINT_PASSWORD
- **Domain:** This property specifies the Microsoft SharePoint domain name. Script constant: SHAREPOINT_DOMAIN

  > **Note**:
  >
  > If you select the **Authenticated User** option, the database connection will use Windows Authentication credentials. Entering a user name and password for the database will supercede the Windows Authentication credentials.

- **Library:** This property specifies the Microsoft SharePoint list name (Document Library). Script constant: SHAREPOINT_LISTNAME
- **Content Type**: If applicable, select the SharePoint content type. If content types have been created in the SharePoint library, they will appear in this list. Script Constant: CONTENT_TYPE

  > **Note**:
  >
  > For more information, see the next section on **Content Types.**

- **Root Path**: This is the location on your SharePoint Server where the folders will be created once the automation service processes the step. If you do not specify a value for the Root Path property, no folders will be created on the SharePoint Server.
- **Local Temp Folder**: This property specifies the local folder path where the Microsoft SharePoint export is temporarily stored on your local machine prior to moving to the Microsoft SharePoint site. Script constant: TEMP_FOLDER

  > **Note:**
  >
  > If the Local Temp Folder is blank, the export will be written to the directory where the application was installed (e.g., C:\Program Files\Digitech Systems\PaperVision Capture Desktop).

Content Types

When exporting documents to a SharePoint site, you can optionally link documents to content types. Content types contain limited subsets of index fields in a SharePoint library. For example, a Financial Documents SharePoint library can contain three content types including Purchase Orders, Invoices, and Expense Reports. Each content type can be associated with a specific subset of index fields. Document content types, the default selection, include all index fields in the library. Content types are independent of file types, so one content type can be applied to multiple file types, such as Microsoft Word documents, Excel spreadsheets, and PowerPoint presentations.

For example, Purchase Orders, Invoices, and Expense Reports content types in a Financial Documents library can be associated with the following index fields:

Content Type	Check Number	Check Date	Company Name	PO Number	PO Date	Invoice Number	Invoice Date	Amount
Purchase Orders			x	x	x			x
Invoices	x	x	x	x	x	x	x	x
Expense Reports			x			x		x

Information on SharePoint 2007 and 2010 content types, respectively, can be found in the following sites:

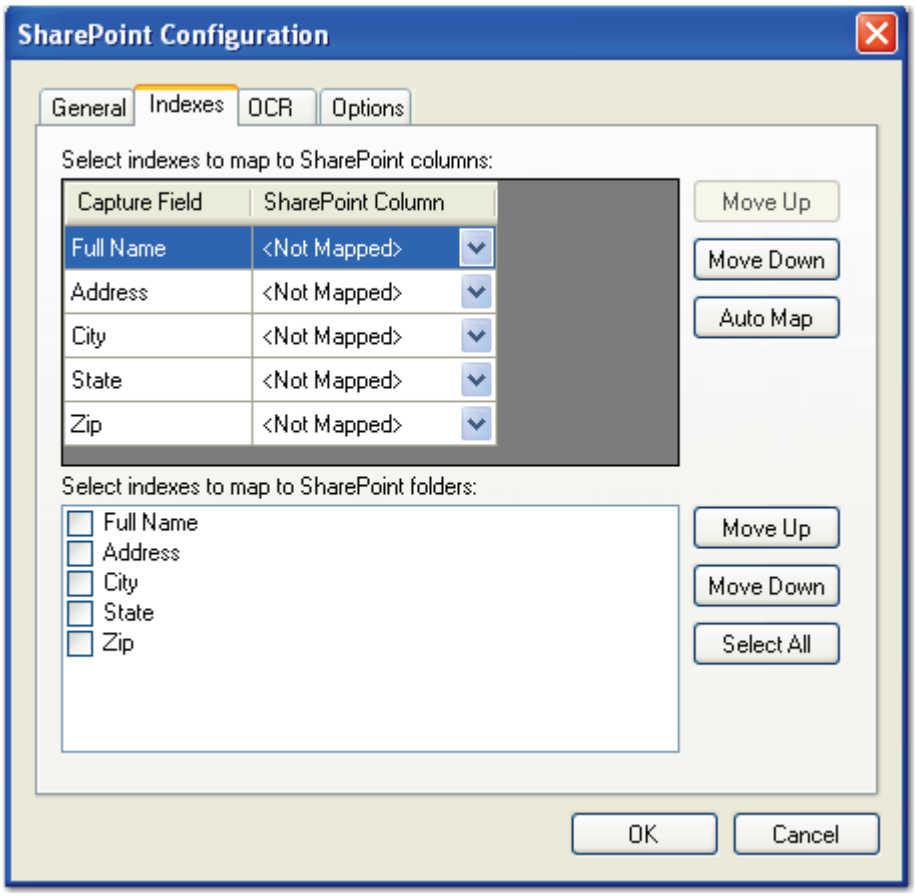
- **http://technet.microsoft.com/en-us/library/cc262735(office.12).aspx**
- **http://technet.microsoft.com/en-us/library/cc262735.aspx**

SharePoint - Indexes

In the Indexes tab, you can map PaperVision Capture Desktop index field names to SharePoint column names. PaperVision Capture Desktop index field names appear in the left column. From the **SharePoint Column** drop-down list, select the column name that maps to the PaperVision Capture index field name. To automatically map an index field to a similarly-named Microsoft SharePoint column, click the **Auto Map** button. To edit the indexes in the resulting export script, you can modify the INDICES_TO_INCLUDE constant, which is described below.

Alternatively, if a SharePoint column does not exist, you can create a new column that will be mapped to the corresponding index field. To do this, select **<Create New>** from the **SharePoint Column** drop down list.

> **Note**:
>
> Some PaperVision Capture Desktop index field types may not be supported in Microsoft SharePoint. Therefore, some index fields may not be mapped to SharePoint columns in the export.



SharePoint Configuration - Indexes

- **Select Indexes to map to SharePoint columns**: Your selections determine the index values mapped from PaperVision Capture Desktop to Microsoft SharePoint columns.

These columns must already be defined in your Microsoft SharePoint list. To provide a mapping between fields, the following format is required:

```
<Capture Field>:<SharePoint>
Example 1: "Field1", "Field 2", "Field 3", etc.
```

> **Note**:
>
> This format can be used when the same field names exist in both PaperVision Capture and your Microsoft SharePoint site.

```
Example 2: "Field1:Field1", "Field2:Field2:", etc.
```

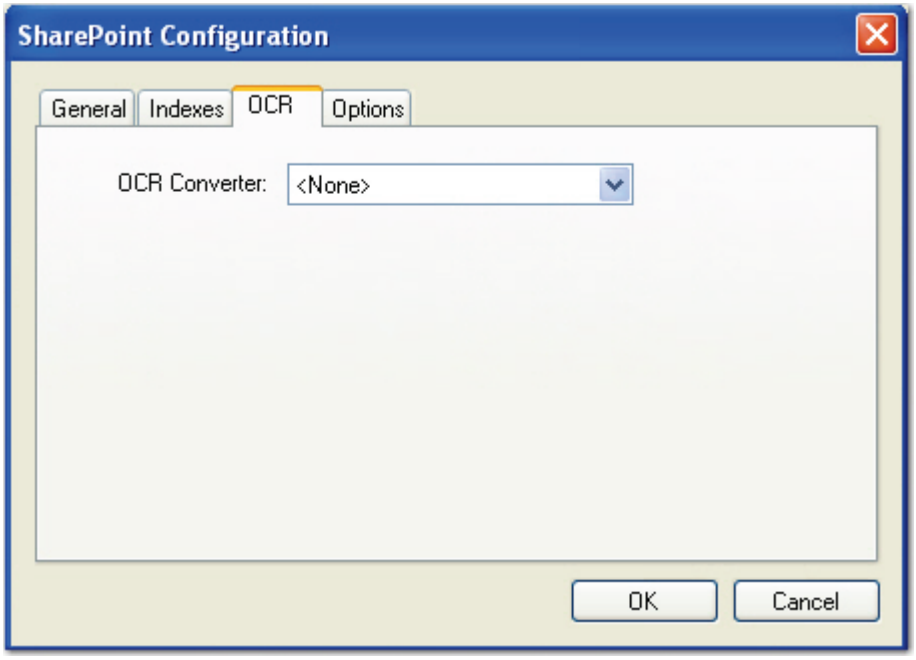
> **Note**:
>
> This constant is optional, so when an empty array is assigned to COLUMN_INDICES, Microsoft SharePoint's metadata is not populated.

Script constant: COLUMN_INDICES

- **Select Indexes to map to SharePoint folders**: Your selections determine the index values mapped from PaperVision Capture Desktop to Microsoft SharePoint folders. These folders must already be defined in your Microsoft SharePoint list. Script constant: FOLDER_INDICES

SharePoint - OCR

When you select the OCR converter in the OCR tab, you can modify the OCR_CONVERTER_CODE constant value that appears in the resulting export script.



SharePoint Configuration - OCR

- **OCR Converter:** This selection specifies the OCR converter code, including PaperFlow, PaperVision, PDF and Text, whose output format is used to export full-text data. When no value is defined (default setting), both images and associated full-text data will be exported. For example, to export PDF searchable images, the following line in the XML script would read:

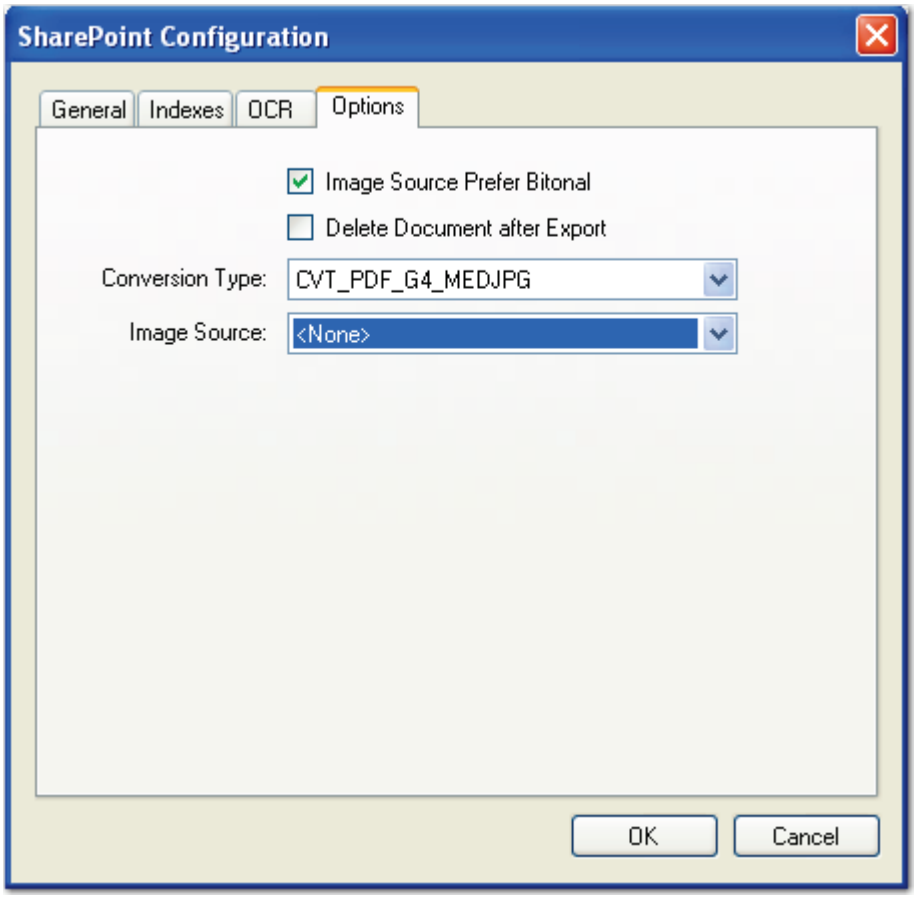
  ```
  private const string OCR_CONVERTER_CODE = "PDF";
  ```

  > **Note**:
  >
  > For a list of converter codes, see the PVCaptureBatchAPI.chm help file's **PVBatch.TryGetOCRFiles Method** topic found within the **Docs** directory where PaperVision Capture was installed.

  Script constant: OCR_CONVERTER_CODE

SharePoint - Options

When you configure properties in the Options tab, you can modify constant values that appear in the export script. Properties and resulting script constants that are available for configuration are described below.



SharePoint Configuration - Options

- **Image Source Prefer Bitonal**: This property is applicable to dual-stream scanners and determines whether to export bitonal or color images. When set to **True**, which is the default setting, bitonal images will be exported. Script constant: IMG_SRC_PREFER_BITONAL_IMAGES
- **Delete Document After Export**: This property specifies whether documents are deleted after they have been exported (set to **False** by default). Script constant: DELETE_DOCUMENT_AFTER_EXPORT

SharePoint - Options

- **Conversion Type**: This property determines the type of image file created during the export. The default value, CVT_NO_CONVERSION, does not convert images during the export. If exporting to a format that supports both single and multi-page images, you must set the CREATE_MULTI_PAGE_IMAGE constant to **True** if you want to create multi-page images; otherwise single page images will result. For example, if you set this to CVT_TIFF_G4_MEDJPG, a TIFF image is created during the export. If the source image is binary, it will create a TIFF using Group 4 compression; if the source image is color (JPG or BMP), it will create a TIFF using Medium JPEG compression. For a list of file types that can be converted to during the export, see the Enumerations section in the **Custom Code Configuration** chapter. Script constant: CONVERSION_TYPE

  > **Note**:
  >
  > Some PaperVision Capture Desktop index field types may not be supported in Microsoft SharePoint, so some field mapping may not be included in the export.

- **Image Source:** This property determines whether original images (pre-image processing) or modified images (post-image processing) are used for the export. The Pre-Image Processing option uses the original image that has not been modified by image processing, and the Post-Image Processing option uses the image that has been modified by image processing. The default selection, **<None>**, uses the most recent image prior to exporting. Script constant: IMG_SRC

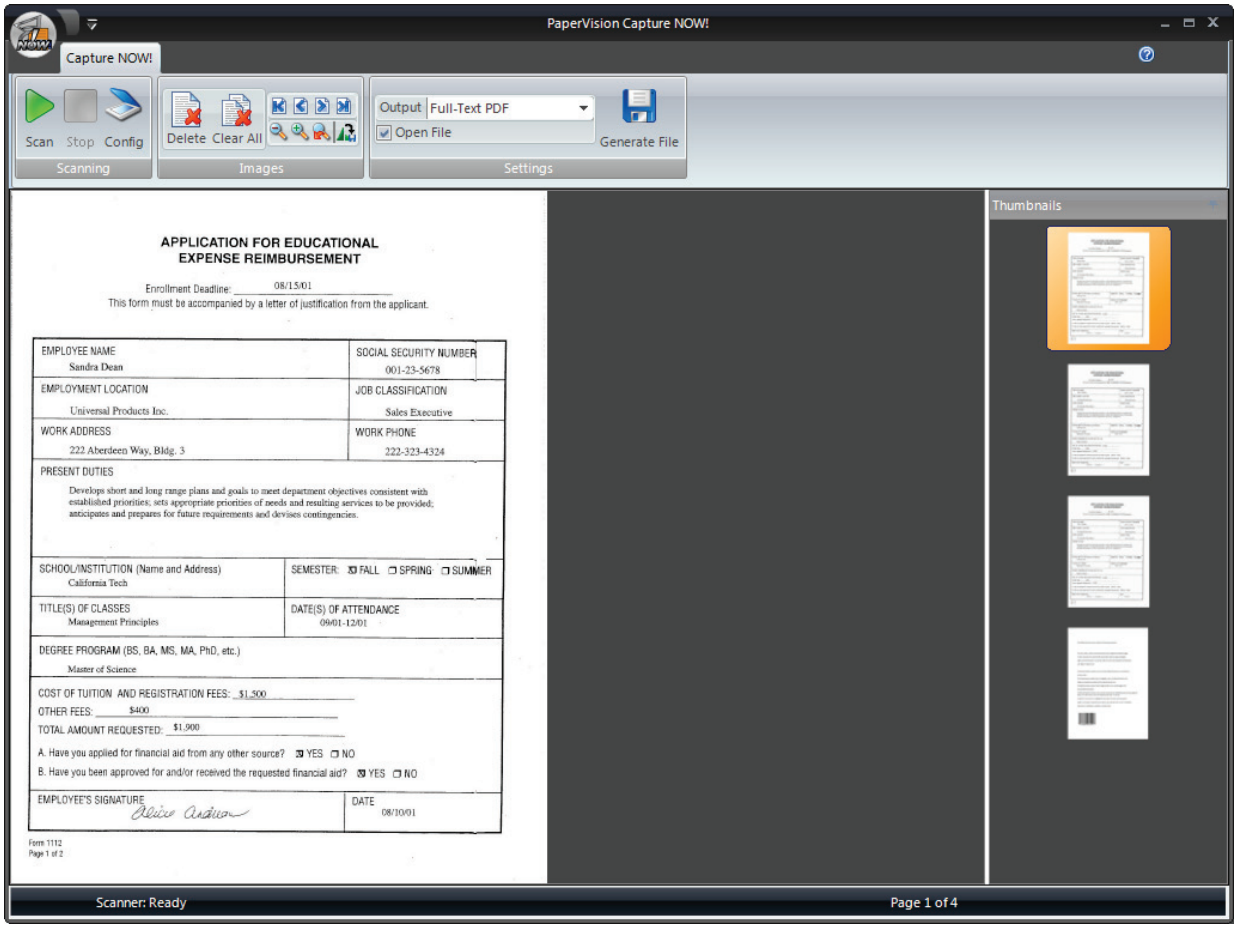
Chapter 10 - PaperVision Capture NOW!

The PaperVision Capture NOW! application enables you to quickly scan and output images to full-text searchable PDF, image only PDF, and single- or multi-page TIFF files. You can scan and output images in a few simple steps, without having to create new jobs and batches to process the images. With PaperVision Capture NOW!, you simply configure your scanner, scan pages, and select your output format and output directory.

Launching PaperVision Capture NOW!

To launch the PaperVision Capture NOW! application, double-click the **PaperVision Capture NOW!** icon located on your desktop. The Capture NOW! page will appear, where you can scan and output your documents.

> **Note**:
>
> The PaperVision Capture NOW! graphical user interface application (launched via the PaperVision Capture NOW! icon) and PaperVision Capture Desktop cannot be run simultaneously on the same machine.



PaperVision Capture NOW!

Configuring the Scanner

Before you initially scan documents, you must first configure a scanner by selecting the **Config**  icon in the toolbar. The Config operation allows you to assign scanner settings. See the **Scanner Settings** section for details on each setting.

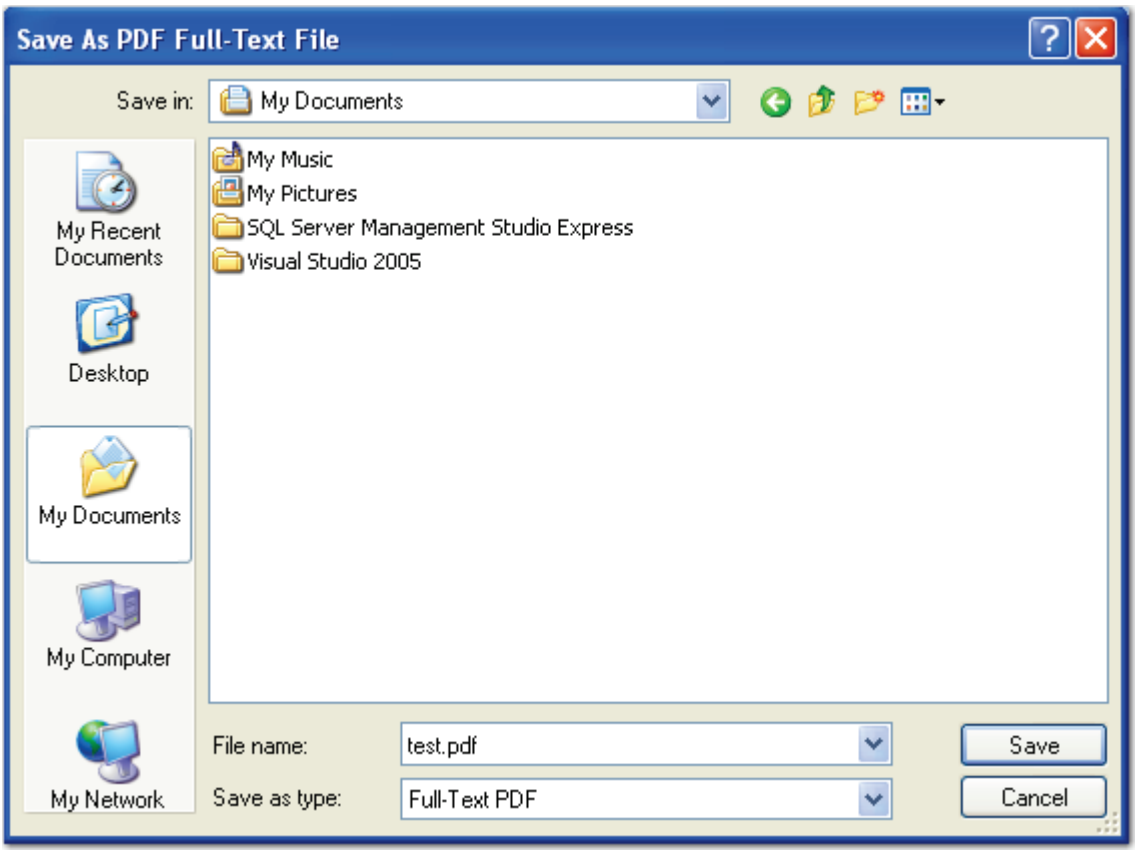
Scanning and Generating Output Files

After you have configured your scanner, you are ready to scan images and output them as PDF or TIFF files.

To scan and generate output files:

1. To scan the images, click the Scan  icon.
2. Once the images have been scanned, select from the following **Output** options:
   - Full-Text PDF
   - Image Only PDF
   - TIFF
3. Optionally, select **Open File** if you want the resulting file to open automatically.

4. To generate the output file, click the **Generate File** icon. The Windows **Save As** dialog appears.



Save As Full-Text PDF File

5. Enter the file name, and then click **Save**. If you selected the **Open File** option, the resulting file will open automatically.
6. Images are not automatically removed from the main viewing window, allowing you to view the resulting file and make changes (delete images, re-scan images, etc.) if necessary. To remove the images from the main viewing window, click the **Clear All** icon.

Stopping the Scanning Process

To stop scanning, click the **Stop** icon. If you stop the scanning process and do not want to save any scanned images, click the **Clear All** icon before proceeding.

Deleting Images

You can delete one or more images in the PaperVision Capture NOW! application.

To delete an image:

1. Select one or more images in the **Thumbnails** window.
2. Click the **Delete**  icon.
3. Click **Yes** to confirm the deletion.

To clear all images:

1. Click the **Clear All**  icon.
2. Click **Yes** to confirm the deletion.

Page Navigation Operations	
First Page 	Displays the first page of the current document Hot Key: Home
Previous Page 	Displays the previous page of the current document Hot Key: Page Up
Next Page 	Displays the next page of the current document Hot Key: Page Down
Last Page 	Displays the last page of the current document Hot Key: End
Jump To Page 	Retrieves your specified page number Hot Key: J

Zooming and Rotating Operations

- To zoom in on the image, click the **Zoom In**  icon.
- To zoom out of the image, click the **Zoom Out**  icon.
- To reset the view of the image, click the **Zoom Reset**  icon.
- To rotate and save one or more images 90 degrees clockwise, select the image(s) in the Thumbnails window, and then click the **Rotate and Save**  icon.

About

Click the **PaperVision Capture NOW! Button**  , and then select **About** to display the product name, current released version of the application, license information, and copyright information.

Appendix A - Additional Help Resources

At Digitech Systems, we provide multiple resources to help find answers to your questions.

Technical Support

Contact our legendary customer support staff Monday through Friday between the hours of 8 a.m. and 6 p.m. Central Time for answers to your questions about our products.

Direct: (402)484-7777

Toll-free: (877)374-3569

Email: support@digitechsystems.com

Help on the Web

MyDSI is an interactive tool for all Digitech Systems customers. Log in to MyDSI to download product updates, license purchased software, view support contract renewals, and check the status of your software support cases and requests.

User Forums

Join the Digitech Systems user forums to exchange answers and ideas with other users in our moderated community.

Knowledge Base

Search our extensive database for articles on all Digitech Systems products. This feature is accessible to everyone.

Appendix B - Supported Countries/Languages (Groups)/ Character Sets

If you narrow the search for specific languages or countries, the OCR engine will process more rapidly during OCR recognition. The following table displays the supported countries, languages, country groups, language groups, and character sets available in PaperVision Capture Desktop.

Supported Countries/Languages (Groups)/Character Sets
Australia
Austria
Azerbaijan
Baltic
Belgium
Brazil
Bulgaria
Canada
Central America
Central Europe
Croatia
Cyrillic
Czech
Denmark
Estonia

Supported Countries/Languages (Groups)/Character Sets
Finland
France
Germany
Great Britain
Greece
Hungary
Ireland
Italy
Liechtenstein
Lithuania
Luxembourg
Netherlands
New Zealand
Norway
Poland
Portugal
Romania
Russia

Supported Countries/Languages (Groups)/Character Sets
Scandinavia
Slovakia
South Africa
South America
South America Spanish
Spain
Sweden
Switzerland
Turkey
Western Europe
OCR
Afrikaans
Albanian
Azerbaijani Latin
Basque
Bosnian Latin
Bulgarian
Catalan

Supported Countries/Languages (Groups)/Character Sets
Croatian
Czech Language
Danish
Dutch
English
Estonian
Faroese
Finnish
French
Frisian
German
Greek
Guarani
Hani
Hungarian
Icelandic
Indonesian
Irish

Supported Countries/Languages (Groups)/Character Sets
Italian
Kirundi
Latin
Latvian
Lithuanian
Luxembourgish
Malay
Norwegian
Polish
Portuguese
Quechua
Rhaeto Romanic
Romanian
Russian
Rwanda
Serbian Latin
Shona

Supported Countries/Languages (Groups)/Character Sets
Slovak
Slovenian
Somali
Sorbian
Spanish
Swahili
Swedish
Turkish
Wolof
Xhosa
Zulu

Appendix C - PaperVision Capture NOW! Command Line Parameters

The PaperVision Capture NOW! application can be run from the command line to automatically execute scanning and file conversion operations, without requiring user interaction in a graphical user interface. Other applications, hardware, and network devices can use the parameters (also known as switches) provided in this appendix to communicate with PaperVision Capture NOW!. For example, you could assign command line switches to communicate directly with a scanner, so that users only need to press a button on the scanner to initiate PaperVision Capture NOW! operations. In this scenario, the scanner must be configured (via the Configure Scanner operation in PaperVision Capture Desktop or PaperVision Capture NOW!) before you assign command line switches.

Once operations are completed, output is displayed in the form of text lines in the command line interface. You can enter combinations of command line switches to simultaneously execute multiple operations. Command line switches are case-insensitive.

> **Note**:
>
> If Capture NOW! is in Demo mode, each image will be watermarked. If a purchased Capture NOW! license has been activated, images will not be watermarked.

The following command line switches can be used to run PaperVision Capture NOW!:

Command Line Switch	Description
CaptureDesktop.exe /pvcnow	* Launches PaperVision Capture NOW! application (graphical user interface); *No optional parameters are available for this switch.
CaptureDesktop.exe /pvcauto	Launches PaperVision Capture NOW! application in auto mode (no graphical user interface)

Optionally, the following command line switches can be used in conjunction with the <pvcauto> switch:

Command Line Switch	Description
/file:"<file path and name.extension>"	This switch specifies a valid Windows path (including file name) of the resulting output file. Local and UNC paths are valid. If this switch is not assigned, the user will be prompted with a "Save As" dialog when scanning completes. If any spaces exist in the file path, the path must begin and end with double quotes (") to indicate that spaces exist. For example, valid paths and file names are: /file:"c:\Capture Now Outputs\Output.pdf" /file:c:\Outputs\Output.pdf
/output:	This switch specifies the resulting output file type. Available outputs include: 1. fulltextpdf (full-text searchable PDF) 2. imageonlypdf (image only PDF) 3. tiff (single- or multi-page tiff) If this switch is not assigned, the default full-text searchable PDF file will be created. **Note**: If you specify the /file:"<file path and name.extension>" switch, the extension must match the output type. Otherwise, documents will not be scanned. In the following example, a TIFF file will be created: CaptureDesktop.exe /pvcauto /file:"c:\Capture Now Outputs\Test.tiff" /output:tiff
/open	When this switch is assigned, the resulting output file will open automatically after it is created. If this switch is not assigned, the file will not open automatically.
/successprompt	This switch determines whether the user receives a success message once the file is created. If this switch is present, a success message will appear on the user's screen once the file is created and saved. If the switch is not entered, no message will appear.

> **Note**:
>
> If multiple switches are assigned, you must insert a single space between each switch.

For example, the switch below automatically creates a full-text searchable PDF named Test and saves the file to a folder called Capture Now Outputs on the c: drive. In addition, the resulting file will open automatically after it has been created and the success prompt has displayed for the user.

```
CaptureDesktop.exe /pvcauto /file:"c:\Capture Now Outputs\Test.pdf" /open /successprompt
```

To create a TIFF file instead, change the following switches:

```
CaptureDesktop.exe /pvcauto /file:"c:\Capture Now Outputs\Test.tiff" /output:tiff /open
/successprompt
```

Exit Codes

The following exit codes are included with PaperVision Capture NOW!. If an exit code appears in the command line interface, you can refer to the following table for each code's description:

Exit Code	Description
0	The operation has been completed successfully.
1	An instance of PaperVision Capture NOW! is already running on the workstation.
4	The scanner has not been configured. **Note**: You must first configure the scanner via the **Config**  icon in the **Scanning** toolbar group (in PaperVision Capture NOW!'s main page or in PaperVision Capture Desktop's Home page) prior to running PaperVision Capture NOW! with the command line interface.
8	DSI ImageSpewer is an invalid scanner.
16	A general application exception error has occurred. For details, refer to the Windows Event Log.
32	An invalid command parameter has been entered.
64	An invalid output parameter has been entered.
128	An invalid file path has been entered.
256	An invalid "/open" switch has been entered.
512	An invalid "/successprompt" switch has been entered.
1024	The scanner has a paper jam.
2048	The scanner has an empty feeder.
4096	A general scanner error has occurred.
8192	A scanner exception error has occurred. For details, refer to the Windows Event Log.

Index

E

F

H

I

J

K

L

M

N

O

P