



This document is the proprietary and confidential property of Kofax, Inc. Do not copy, distribute, transfer or otherwise disclose this document or the information contained herein without the prior written consent of Kofax.



## Application Note

### Windows eVRS 2.1 Settings

October 24, 2014

## Contents

Contents .....	2
Overview.....	3
Processing Modules.....	4
Module Properties.....	5
Page Detection .....	5
DO_CREATE_INDEX_IMAGE .....	5
DO_BINARIZATION and DO_ENHANCED_BINARIZATION .....	6
DO_BACKGROUND_SMOOTHING.....	6
DO_BLANK_PAGE_DETECTION .....	7
DO_COLOR_DETECTION .....	7
Color/Grayscale Brightness Adjustment.....	7
Resolution of Inconsistencies .....	8
Examples.....	9
Example 1 .....	9
Example 2 .....	9

## Overview

This document contains information on eVRS setting and properties string that can be used in KTA Scan/VRS Profile Advanced setting tab.

Most of the EVRS processing modules are disabled by default, but appropriate keywords included in the operations string enable the corresponding modules. In the operations string a leading “\_” character indicates the start of a new keyword or parameter. The trailing “\_” character may be followed by a setting value: a string in <> or one or more numeric values.

To load properties of a processing module, use the operation string “\_LoadSetting\_<PropertyName=“XXXX” Value=“XX” Comment=“XXXX”/>”. The strings XXXX and XX indicates property name, value, and comments.

## Processing Modules

Keyword	Actual String	Description and Parameter range
DO_HOLE_FILL	_DoHoleFill_	Fills punch holes on document edges in order to match the surrounding page background. Performed for dark backgrounds only.
DO_BLANK_PAGE_DETECTION	_DoBlankPageDetection_	Detects if the page is blank so that it could be deleted by the application. Capable of ignoring punch holes. In Duplex mode is capable of ignoring bleed-through from the opposite side.
DO_COLOR_DETECTION	_DoColorDetection_	Limits binarization to only those pages that do not contain any color.
DO_SKEW_CORRECTION_PAGE	_DoSkewCorrectionPage_	Deskews to page
DO_SKEW_CORRECTION_ALT	_DoSkewCorrectionAlt_	Deskews to content
DO_CROP_CORRECTION	_DoCropCorrection_	Crops to minimize the bounding box around the final position of the page after deskew.
DO_BINARIZATION	_DoBinarization_	Converts the image to binary.
DO_ENHANCED_BINARIZATION	_DoEnhancedBinarization_	Useful for creation of binary images with resolutions above 300 DPI. Combines noise reduction ability of lower resolution images with smoother character contours characteristic of higher resolutions.
DO_GRAY_OUTPUT	_DoGrayOutput_	Converts the image to grayscale.
DO_90_DEGREE_ROTATION	_Do90DegreeRotation_n	Performs image rotation in 90 degree increments or automatically, based on image content. n: 1=90 degrees counterclockwise; 2=180; 3=270 and 4=automatic.
DO_ROTATE_NONE	_Do90DegreeRotation_0	
DO_ROTATE_90	_Do90DegreeRotation_3	
DO_ROTATE_180	_Do90DegreeRotation_2	
DO_ROTATE_270	_Do90DegreeRotation_1	
DO_ROTATE_AUTO	_Do90DegreeRotation_4	
DO_EDGE_CLEANUP	_DoEdgeCleanup_	Cuts a small frame around the final image in order to clean the fringes of the page. Only works if _DoSkewCorrectionPage_ and _DoCropCorrection_ are set as well.
DO_DESPECK	_DoDespeck_n	Removes speckles on the final binary image. n indicates the maximum speck size. Applicable to binary images only.
DO_CREATE_INDEX_IMAGE	_DoCreateIndexImage_	Necessary for any work involving detection and processing of objects within images, such as graphic and text lines, pictures, punch holes, etc. See comments below.
DO_BACKGROUND_SMOOTHING	_DoBackgroundSmoothing_	Performs smoothing of page background for color and grayscale images.

## Module Properties

This section lists properties that belongs to specific processing modules listed in the previous section. Properties are loaded using the “\_LoadSetting\_” operation string.

### Page Detection

Image processing in EVRS always begins with page detection. Even if deskew to page or content, as well as crop are not requested, still page detection is performed. The reason for that is the need to know where the page is in order to detect content like punch holes within it, decide whether the page is blank, or to analyze the image only inside the page while determining how better to binarize it. In order to help page detection algorithm it is advisable to specify the range of scanner background colors by including in the scanner XML file containing all the relevant non-defaults one or more of the corresponding parameter settings, namely

Property name (case is important)	Description and Parameter range
CSkewDetect.white_bkg_threshold	Integer within [0,255] with the default =128 meaning “unknown”. See explanation below.
CSkewDetect.color_stats_error_sum_thr_white_bkg CSkewDetect.color_stats_error_sum_thr_black_bkg	Integers within [0,255]. Defines scanner background color detection sensitivity. Defaults for scanned images:  CSkewDetect.color_stats_error_sum_thr_white_bkg=60, CSkewDetect.color_stats_error_sum_thr_black_bkg=96.  Defaults for for Mobile (color input only, with local background adjustment):  CSkewDetect.color_stats_error_sum_thr_white_bkg=18, CSkewDetect.color_stats_error_sum_thr_black_bkg=36.

The minimal help can be to specify CSkewDetect.white\_bkg\_threshold. If it is set to value below 128 it means that the background of the scanner is “black”, or more specifically, its gray level can vary between zero and set value; if it is set to value above 128 it means that the background of the scanner is “white”, or more specifically, its gray level can vary between set value and 255. Better still is to specify both ends of the allowed background range in all three color channels.

Note: Some scanners add filler bands to scanned images, usually at the bottom but sometimes on other sides as well, if the scanned page is smaller than the specified paper size. If, for a variety of reasons, cropping fitting the found page is not happening (either because it was not requested or because of page detection confidence-based reject) there is a choice to either preserve or remove the filler bands. Historically in VRS the default is to preserve filler bands; in EVRS this default is to remove them. If it is necessary to mimic the behavior of VRS, it is possible to set the value of *CSkewDetect.keep\_filler\_if\_no\_crop.Boo1* to 1 (TRUE).

### DO\_CREATE\_INDEX\_IMAGE

The index image uses an 8-bit encoding of the content found within the processed image in order to exchange this information among processing modules. In the current version of EVRS

library it is still necessary to explicitly request the creation of the index image. Beginning with EVRS 2.2 this will be done automatically as necessary.

## DO\_BINARIZATION and DO\_ENHANCED\_BINARIZATION

In order to improve the quality of binary images for documents with low original resolution it is possible to replace DO\_BINARIZATION with DO\_ENHANCED\_BINARIZATION. This functionality is useful for creation of binary images with resolutions higher than 300 DPI and it results in a binary processed image with the requested high resolution (and correspondingly smoother character contours), but with lower background noise levels.

Note: there are several essential parameters controlling binarization, namely:

Property name (case is important)	Description and Parameter range
intelligent_contrast_enabled	Boolean with the default 1 (TRUE). If set to 0 will cancel Auto-Contrast.
CBinarize.Contrast_Slider_Pos.Int	Integer within [1,5] with the default =3 for scanned images and default=2 for Mobile. (See IMAGING_DEVICE_TYPE.) Controls the aggressiveness of Auto-Contrast determination.
intelligent_brightness_enabled	Boolean with the default=1 (TRUE). If set to 0 will cancel Auto-Brightness.
CBinarize.Do_Adv_Clarify.Bool	Boolean with the default=1 (TRUE). If set to 0 will cancel Intelligent Cleanup.
CBinarize.Cleanup_Slider_Pos.Int	Integer within [1,5] with the default =3 for scanned images and default=2 for Mobile. (See IMAGING_DEVICE_TYPE.) Controls the aggressiveness of Intelligent Cleanup.
Advanced_Threshold_Dot_Matrix_Image_Enable	Boolean with the default=1 (TRUE). For most contemporary documents without any dot-matrix printing should be set to 0 to save time spent on this analysis.
global_threshold	Integer within [0,255] with the default 128. Used if Auto-Brightness is off. Sets the brightness threshold for binarization of flat (non-edge) pixels: those with gray levels below 255 - <b>global_threshold</b> will become black, otherwise – white.
edge_threshold	Integer within [0,255] with the default 128. Used if Auto-Contrast is off. Sets the contrast threshold for edge strength to be used by the Dynamic Thresholding. Smaller values will create more edges.
CBinarize.Edge_Aggr.Int	Integer within [0,255] with the default 168. Used by the Dynamic Thresholding.
Threshold.alpha_strength	Integer within [0,255] with the default 128. Used if both Auto-Contrast and Auto-Brightness as well as other advanced features of binarization such as color analysis are off in order to mimic VRS4.5 binarization. Replaces the value of <b>CBinarize.Edge_Aggr.Int</b> in the Dynamic Thresholding.
VRS.Analyze.Color.Enable	Boolean with the default=0 (FALSE). If set to 1 will enable independent binarization of BLUE, GREEN, and RED channels of a color image with their combination resulting in the final binary image. Takes about 3 times more time than binarization from a grayscale image.
CBinarize.Wei_Blue_To_Gray.Int CBinarize.Wei_Green_To_Gray.Int CBinarize.Wei_Red_To_Gray.Int	If color analysis is off these 3 integer values determine the relative weights of color channels in the conversion of color to gray. The defaults for BLUE, RED, and GREEN are 4, 7 and 5 respectively, meaning that the gray value will be calculated as $GRAY=(4*BLUE+7*GREEN+5*RED)/(4+7+5)$ . These values can be used to strengthen or suppress particular colors in the image – for example, in order to make green text stronger <b>CBinarize.Wei_Green_To_Gray.Int</b> can be set to 0, or in order to suppress red in the background <b>CBinarize.Wei_Red_To_Gray.Int</b> can be set to 1 with the other 2 weights set to 0.

## DO\_BACKGROUND\_SMOOTHING

Performs smoothing of page background for color and grayscale images.

Property name (case is important)	Description and Parameter range
-----------------------------------	---------------------------------

Color.BackgroundBleach.Aggressive	Integer within [0,255] with the default 128. Controls the sensitivity of background smoothing. Higher values increase the likelihood that a pixel will be determined to be part of the image background or backgrounds if more than one.
Color.BackgroundBleach.Choice <i>Note In this property name there is a leading blank which is still preserved in the name for backward compatibility.</i>	Integer with the default = 1. 1 = Automatic: All background pixels are smoothed to the same value representing an average of the detected background color. This works well when there are only minor variations in the background colors.

## DO\_BLANK\_PAGE\_DETECTION

Detects if the page is blank so that it could be deleted by the application. Capable of ignoring punch holes. In Duplex mode is capable of ignoring bleed-through from the opposite side.

Property name (case is important)	Description and Parameter range
VRS.Ignore.Holes.Enable	Boolean with the default=0 (FALSE). If set to 1 will ignore hole marks in a page for purposes of blank page detection.
CBlkPage.Min_Count_Edge.Int CBlkPage.Min_Count_Black.Int	Controls the sensitivity of blank page detection. The range for <b>Min_Count_Edge</b> is [0, 300] with the default 75. The range for <b>Min_Count_Black</b> is [0, 90000] with the default 450. These are prorated for DPI linearly for <b>Min_Count_Edge</b> and as a square for <b>Min_Count_Black</b> . Higher values are more likely to determine the page as blank.
VRS.Blank.Page.Content.Sensitivity	Integer within [0,255] with the default 128. Controls the sensitivity of blank page detection. <b>Note</b> If the application sets this to a non-default value, the two <b>CBlkPage.Min_Count_*.Int</b> settings are adjusted from their values: they become lower for sensitivity greater than 128, higher for sensitivity smaller than 128. Higher values are more likely to determine the page as non-blank.

## DO\_COLOR\_DETECTION

With color detection, EVRS evaluates each image to determine whether or not it should be processed as a color image or a black and white image.

Property name (case is important)	Description and Parameter range
ColorDetect.DropBGColorForBitonalFG	Boolean with the default=0 (FALSE). If set to 1 will ignore the color background of a document that is otherwise black and white. For example, with a black and white document printed on color paper, you may prefer the background color to be ignored when performing color detection on the document.
ColorDetect.SmallObjEnable	Boolean with the default=0 (FALSE). If set to 1 will allow detection of small areas of color such as a date stamp or small amounts of highlighter text, so it isn't treated as black and white.
ColorDetect.SmallObjAggressiveness	Integer within [0,255]. The default=128. Specifies the small color area sensitivity: lower values decrease the likelihood of detecting small amounts of color.

## Color/Grayscale Brightness Adjustment

Property name (case is important)	Description and Parameter range
CColorAdj.Slider_Pos.Int	Integer within [0,100] with default 45. Darken or brighten color or grayscale image. The neutral value is 36. Larger values produce brighter images; lower values produce darker.

## Resolution of Inconsistencies

If the `operations_str` happens to contain contradictory requests, such as **DO\_BINARIZATION** and **DO\_GRAY\_OUTPUT**, it is necessary to resolve such conflicts. The way it is done is as follows:

- **DO\_BINARIZATION** and **DO\_ENHANCED\_BINARIZATION** take precedence over **DO\_GRAY\_OUTPUT** regardless of their order in the operations string;
- **DO\_SKEW\_CORRECTION\_PAGE** and **DO\_SKEW\_CORRECTION\_ALT** can coexist – if they are both present, then deskew to content will be done if the found page gets rejected;
- In case the same keyword or parameter name is followed by a different parameter value, for example, ***\_Do90DegreeRotation\_0\_Do90DegreeRotation\_4***, the **last** value will be used.
- All individual parameter values loaded from substrings beginning with **LOAD\_SETTING** are loaded in order they are mentioned in the **operations\_str**, so if the same parameter name is mentioned again, the **last** value will be used.

## Examples

### Example 1

```
_DoBinarization__LoadSetting_<PropertyName="CBinarize.Do_Adv_Clarify.Bool" Value="1"
Comment="DEFAULT 0" />_LoadSetting_<PropertyName="CBinarize.Cleanup_Slider_Pos.Int"
Value="3" Comment="DEFAULT 3" />
```

Performs binarization, with advanced clarity turned on, and aggressiveness set to 3.

### Example 2

```
_DoHoleFill__DoDespeck_3_DoEdgeCleanup__DoBinarization__DoBlankPageDetection__Loa
dSetting_<PropertyName="intelligent_contrast_enabled" Value="1" Comment="DEFAULT 1"
/>_LoadSetting_<PropertyName="CBlkPage.Min_Count_Edge.Int" Value="300"
Comment="DEFAULT 75" />
```

Performs binarization, hole fill, despeckle with maximum speckle size set to 3, edge cleanup, blank page detection, with auto contrast enabled, and blank page detection threshold edge detection set to 300.