

PART NUMBER, BUILD NUMBER

Part#	1041/29
Build#	11.1.6.94

NEW FEATURES AND IMPROVEMENTS

POSSIBILITY TO EXTRACT AND TO ADD ATTACHMENTS FROM PDF

From this release it is possible to extract attachments from input PDF file and to add attachments to the output PDF file.

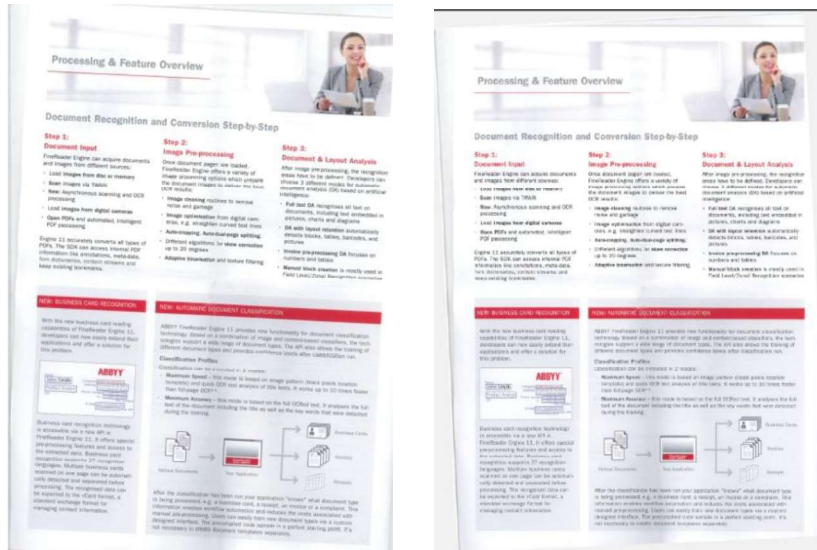
Each attachment is represented by PDFAttachment object that exposes methods which allow you to access the attached file by saving it on disk or into the global memory.

PDFAttachment object also provides access to the original file name, description added by the author and the type of binding of the attachment. The binding can take values from PDFAttachmentBindingEnum: PAB\_Annotation and PAB\_Document. PAB\_Annotation means that attached file is associated with a specific annotation on a specific page. PAB\_Document value is set if the file is attached to the whole document. Note that for an attachment which is added via FineReader Engine API binding value is always PAB\_Document.

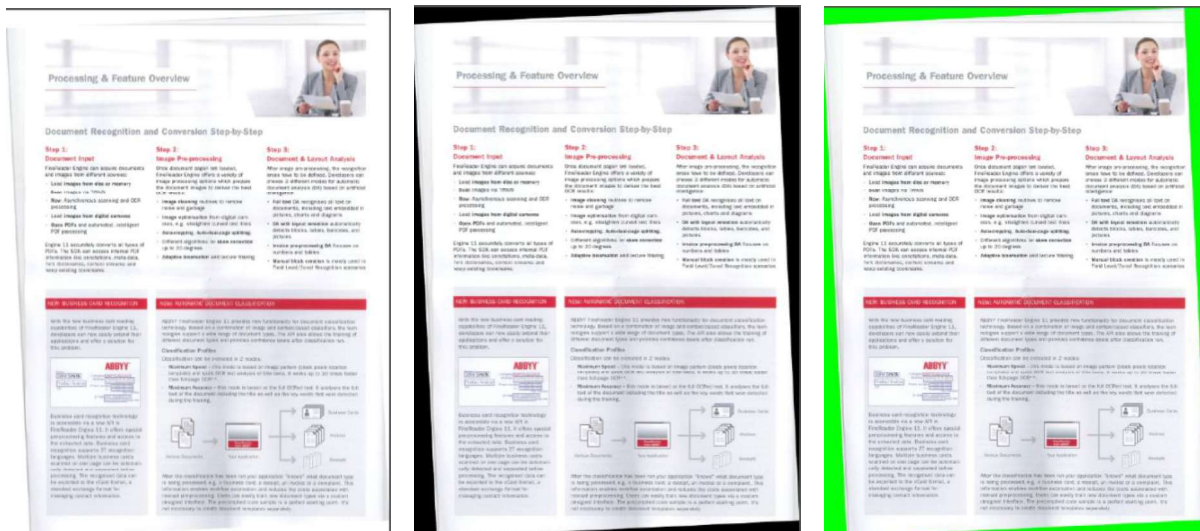
All the attachments are contained in FRDocument::PDFAttachments. They are extracted from the input PDF document during opening, or you can add your own files to be attached to the output PDF file during export. To attach all the files of this collection to the output PDF file, it is necessary set the IPDFExportFeatures::WriteSourceAttachments property to TRUE.

NEW PROPERTY IPREPAREIMAGE::BACKGROUNDFILLINGCOLOR TO FILL AREAS ADDED AFTER SKEW CORRECTION

After skew correction supplementary areas are added to the edges of document. The color of these areas is set automatically by FineReader Engine:



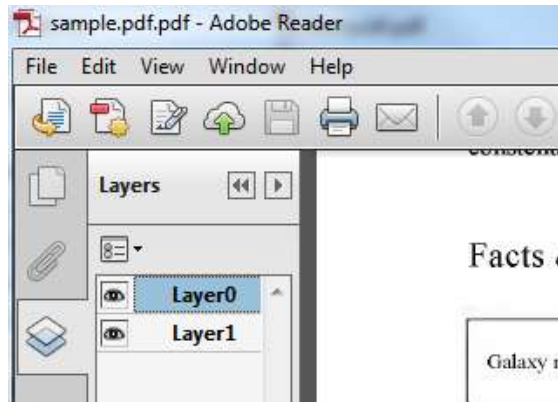
Now it is possible to specify the color used for filling the areas manually with `IPPrepareImageMode::BackgroundFillingColor` property:



The default value of this property is -1, which means that the color is determined by ABBYY FineReader Engine automatically (HelpDesk request 401344).

## POSSIBILITY TO SEE PDF LAYERS IN PDF VIEWERS

A new property `IPDFMRCParams::AssignPdfLayersToMrCPlanes` was added in this release. This property manages the availability of PDF layers in the output PDF file. If you set it to TRUE, PDF layers are assigned to MRC image planes, so that in some PDF viewers such as Adobe Acrobat you can choose which layers to view:



This property is only taken into account for image-only and image-on-text PDF files.

The default value of this property is FALSE.

---

### PARALLEL EXPORT TO PDF AND TO PPTX

Now export of multi-page documents to PDF and to PPTX will be performed in parallel mode. This feature increases processing speed of multi-page documents in parallel processing scenario using Batch Processor or FRDocument. In the table below you can see the number of pages processed per minute and exported to PDF in this release compared to the previous release:

	Release 3 with parallel export	Release 2 without parallel export
Processing with <b>FRDocument</b>	111	79
Processing with <b>FRDocument</b> (with PageFlushingPolicy = PFP_KeepInMemory)	143	96
Processing using Batch Processor	117	82

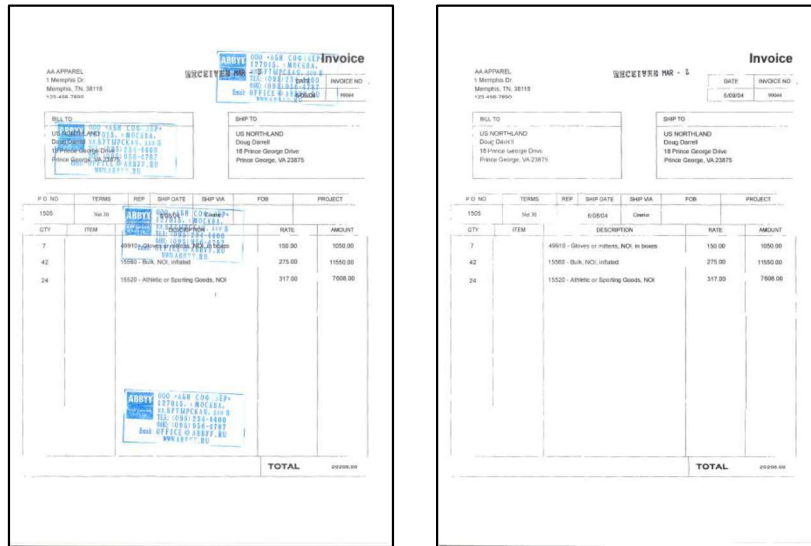
The processor of the testing machine is Intel® Core™ i5-3450 (3.10GHz, 4 physical cores) with 8 GB of RAM, the number of simultaneously run processes is 4. Performance was tested on 300 English-language images, with the settings of the DocumentArchiving\_Speed predefined profile.

---

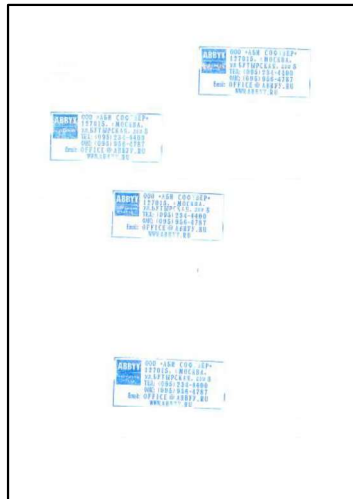
### NEW METHOD IIMAGEDOCUMENT::REMOVECOLOROBJECTSEX

This feature can be useful for documents with black text and white background that contain colored elements that don't need to be recognized. For example, for monochromic documents with colored stamps only black-and-white layer can be recognized and the colored elements can be added to output document.

ImageDocument::RemoveColorObjectsEx method allows you to remove color objects of specified hues from original document



The hues of removed objects are set by HSL representation. The method allows replacing removed objects with the specified color. It also allows saving a separate image containing only the extracted objects:



## NEW ADVANCED LANGUAGE DETECTION MODE

New property RecognizerParams::LanguageDetectionMode was added on this release. This property will be useful for recognition of documents the language of which is not known to you. If language detection mode is on, the

recognition languages are selected from the list of languages specified in the TextLanguage property. Language detection was significantly improved from the previous release and now it turns on automatically if it is necessary.

This property has three modes: TSPV\_Auto, TSPV\_No and TSPV\_Yes. The default value of this property is TSPV\_Auto. In this mode ABBYY FineReader Engine will automatically determine if this processing mode should be used, depending on the situation.

Auto detection will be useful only if TextLanguage set contains a combination of CJK and European languages. New feature will increase the speed and improve the quality of recognition for documents with European languages in the scenario of processing of input batch that contains documents only with CJK languages and documents only with European languages.

In the scenario when you know that all languages, specified in TextLanguage set, are present in the document, that you process, we recommend setting RecognizerParams::LanguageDetectionMode to TSPV\_No.

Old property RecognizerParams::DetectLanguage is marked as deprecated.

---

## NEW PREDEFINED FILTER FNF\_PDF IN FONTNAMESFILTERSENUM

FontNamesEnum enumeration constants describe predefined filters of font family names. These filters specify the set of fonts to be used during document synthesis. A new value FNF\_PDF was added to FontNamesFiltersEnum. If this filter is applied, document synthesis uses font families the names of which are specified in the resources of the input PDF file. However, the fonts themselves are not extracted from the PDF file; they need to be installed on the workstation to be used.

---

## NEW PROPERTY IFRDOCUMENT::PDFFONTNAMES

New read-only property IFRDocument::PDFFontName returns the collection of fonts extracted from PDF file.

---

## SKEW CORRECTION IN PREPROCESS STAGE

Now it is possible correct skew at preprocessing stage instead of during image opening.

CorrectSkew property was added to PagePreprocessingParams object. The type of skew correction is defined by the CorrectSkewMode property.

---

## ADVANCED IEMPTYEX METHOD FOR CHECKING IF THE PAGE IS EMPTY

IsEmptyEx is a new method which allows you to specify additional parameters during empty page detection.

With EmptyPageDetectionParams object you can define the number of letters and text objects that a page can contain and still be considered empty. It is possible to set maximum black percentage and to specify if the page must be searched for barcodes. You can also set the page rectangle, so that any garbage on the margins does not affect the result.

A method IEngine::CreateEmptyPageDetectionParams was included in API to create the EmptyPageDetectionParams object.

---

## WIBUKEY SUPPORT

WibuKey protection hardware keys are supported in this release. It is necessary to use new pricelists to generate licenses with license Storage CodeMeter Key. HQ will add new pricelists for each particular regional office to Registration Server upon request. You can address directly to HQ product analysts.

---

## POSSIBILITY TO EXPORT TIFF FILES WITH ONE STRIP

A strip is a subsection of the image composed of one or more rows. A TIFF image may be composed of one or more strips. Now it is possible to deliberately produce TIFF file composed of one strip by setting `ITiffExtendedParams::WriteSingleStrip` to `TRUE`. The parameters for TIFF files producing can be used `Image::WriteToFile` method.

That feature was implemented as some archive writers don't support valid TIFF files with several strips (HelpDesk request 421941).

---

## NEW PROPERTY RECOGNITION SET

A property `ITextLanguage::RecognitionSet` returns the full letter set used for recognition with this `TextLanguage`, combining all letter sets of its base languages and additional letter sets.

(HelpDesk requests 375306, 420377, 300708).

---

## DEACTIVATION IN SILENT MODE

Now it is possible to deactivate licenses in silent mode via License Manager by entering the parameter `/SilentDeactivation` through the command line (HelpDesk request 401395).

---

## IMPROVED CRM\_CONTENTONLY MODE

New technology is used in `CRM_Content` Mode for document analysis of PDF files. In `CRM_ContentOnly` mode only content of the source PDF file is used (the image is not rasterized for recognition as it was before, all the information about text layer, images, separators, etc. is taken directly from the PDF). This mode is designed for PDF that contain not only raster elements. This changing helps to preserve the original layout of the source document and improve the quality of the output picture. This feature is particularly useful for conversion from PDF to formats of MS Office.

Comparing to the previous release `CRM_Content` Mode works faster and some shortcomings and bugs of this mode were eliminated.

To use the feature you can set `IObjectsExtractionParams::SourceContentReuseMode` to `CRM_ContentOnly` and use these parameters during analysis stage.

---

## CORRECT DISPLAY OF PDF FILES WITH PMINGLIU/MINGLIU FONTS IN PDF VIEWERS

Now all exported PDF files with PMingLiU and MingLiU fonts are displayed correctly in all PDF viewers. Previously there were problems with display due to errors of some viewers. As a workaround these fonts are now embedded to the output PDF file.

Please note that if for a Chinese-language document the PMingLiU/MingLiU font is used, it will be embedded into output PDF file regardless of the value of the property IPDFExportFeatures::FontEmbeddingMode (HelpDesk request 414470).

---

## INFORMATION ABOUT ADDING COMMENTS IN PROFILES IN HELP FILE

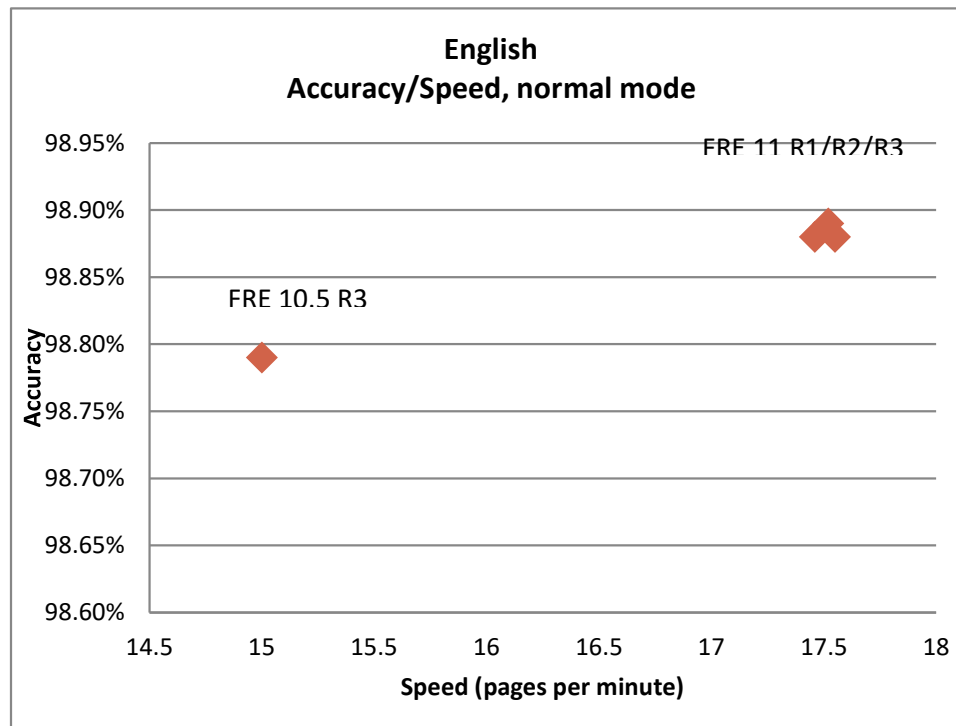
The information and an example on how to add comments, were in the section Working with Profiles Comments can be added by starting a line with a semicolon (HelpDesk request 406671).

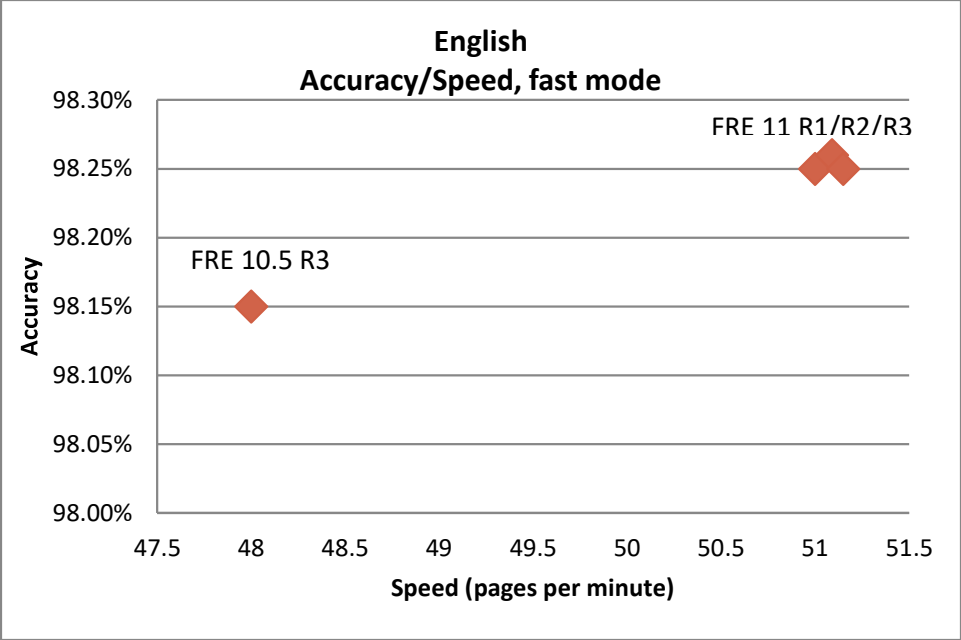
## PERFORMANCE RESULTS

This section contains performance results of FRE 11 R3 comparing to the previous releases and the latest version of FRE 10.5. The processor of the testing machine is Intel® Core™ 2 Duo CPU E6750 (2.66GHz, 2 physical cores) with 3,9 GB of RAM.

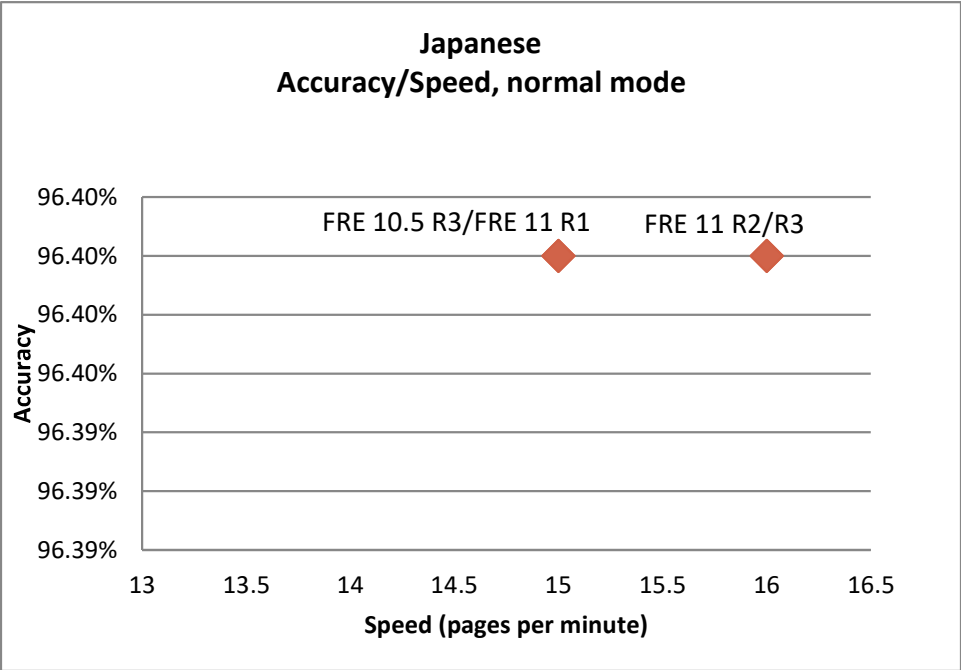
---

## ENGLISH

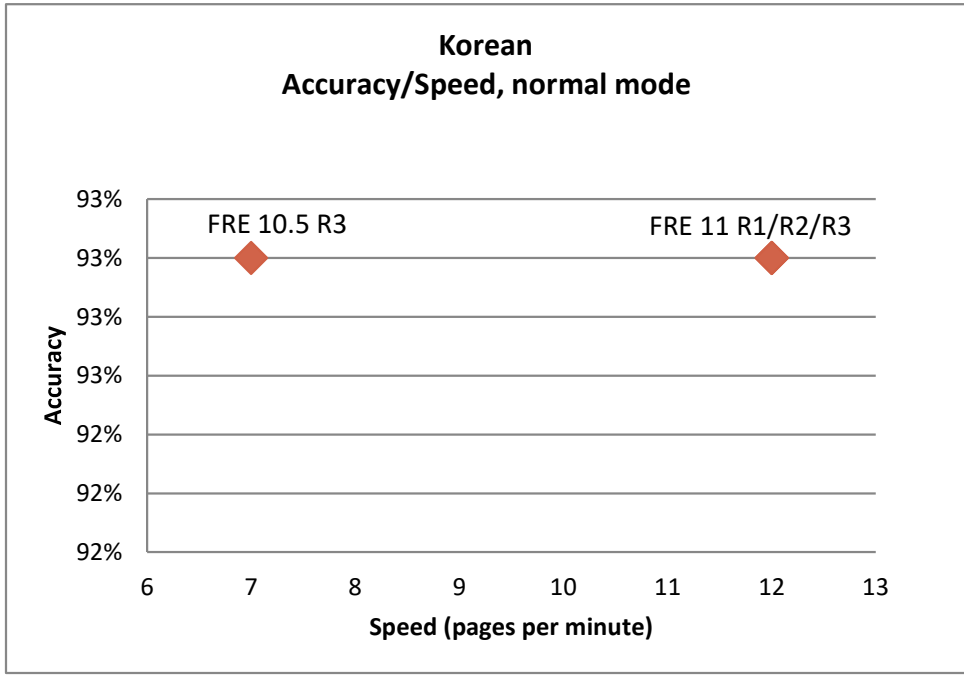




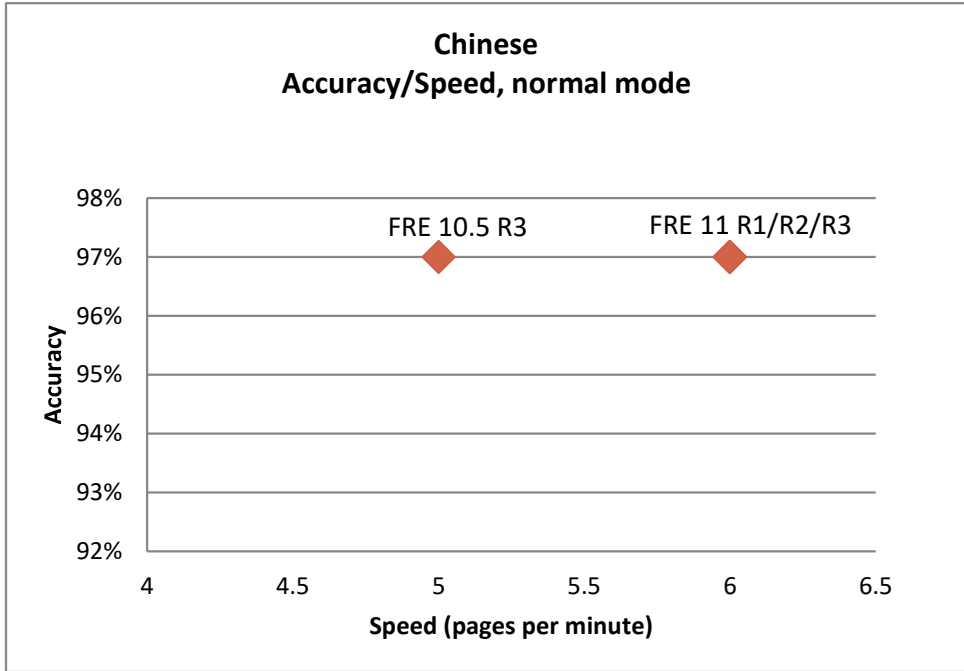
JAPANESE



KOREAN



CHINESE



## FIXED BUGS

This section contains a list of bugs reported by customers that have been fixed.

Four-point scale will help you to evaluate the severity of each issue, enabling you to make informed decision on how important updates are for your system.

<b>Critical</b>	A bug that causes crashes or hangings of software. Critical bugs can include access violations, internal program errors, stack overflow, out of memory or other exceptions that can lead to program failure.
<b>Major</b>	A bug that does not cause program failure but affects major functionality of a feature or impairs the system's performance. Major bugs can include disparity of the feature functionality to the internal specifications, memory leaks or data corruption.
<b>Minor</b>	A bug that leads to feature malfunctioning or affects minor functionality of the software. Minor bugs can include recognition errors, missing or lost objects, wrong color detection, incorrect document analysis, license counter errors, etc.
<b>Trivial</b>	A cosmetic issue that does not affect the functionality of the product but can cause inconveniences. Trivial bugs can include Help file errors, log errors, incomplete information in error messages, etc.

The following table contains bugs fixed in this release sorted in descending order of severity. If the bugs have workarounds, root causes or side effects, they will be mentioned in the Description section.

Severity	Description	Subsystem	HD #	Office
Critical	The project hangs on the IFRDocument::Recognize method with the following settings:  RecognizerParams::TextLanguage = English, CMC7 and RecognizerParams::TextTypes = (TT_MICR_CMC7   TT_Normal).	Recognizer	405817	US
Critical	The project hangs on the IFRDocument::Recognize method after adding of two blocks with hand-printed text (FMT_PartitionedFrame and FMT_SimpleComb).	Handprint Recognizer	405607	EU
Critical	Unhandled access violation during call of IMultiProcessingParams::RecognitionProcessesCount  <b>Root cause:</b> interface IMultiProcessingParams lacked "dual" attribute.	API	405818	EU
			335175	US
Critical	IPE: .\src\ProcessorsCallsHandler.cpp, 74. during analysis stage with MultiProcessingMode set to MPM_Parallel on specific image.	DA	412308	EU

Critical	IPE: .\src\PrepareImageEmulator.cpp, 59 on the IEngine::PrepareDib method on specific image.	API	419854	US
Critical	IPE: Division by zero during analysis stage on specific image in multiprocessing mode.	DA	426015	EU
Critical	IPE: .\Src\TableVerticalSplitter.cpp, 141 during analysis stage on specific image in multiprocessing mode.	DA	426015	EU
Critical	An error "...Data\Arabic.oce was not found." on analysis of CJK documents  <b>Workaround:</b> leave Arabic.oce in the Data folder in the distribution	DA	418831	EU
Major	During export to PDF with PDFAComplianceMode = PCM_Pdfa_1a and FontEmbeddingMode = FEM_Embed an error occurs "Could not embed the font 'Helvetica'. Embedding is not allowed for this font."	PDFToolkit	407718	EU
Major	ImageOnly PDF files exported with MRC are fuzzy.	Image	416816 423131	US
Major	Documents exported to PDF with PDFAComplianceMode = PCM_Pdfa_1a and PDFAComplianceMode = PCM_Pdfa_1b don't pass online validation on <a href="http://www.pdf-tools.com/pdf/validate-pdf-online.aspx">http://www.pdf-tools.com/pdf/validate-pdf-online.aspx</a> .	Export	418803	US
Major	With IRecognizerParams::SaveCharacterRecognitionVariants = true only one character recognition variant is saved	Recognizer	413018	US
Major	An error "An incorrect structure was found in the PDF file" in Adobe Acrobat XI PRO during page insertion (Insert Pages... > From File...). The problem is reproduced on exported tagged PDF with links.	PDFToolkit	419801	US
Major	Chinese text written with MingLIU and PMingLIU fonts isn't displayed in Adobe Reader.  <b>Workaround:</b> set FontEmbeddingMode = FEM_Embed  <b>Root cause:</b> Adobe Reader doesn't display Chinese text written with MingLIU and PMingLIU fonts if they are not embedded.	PDFToolkit	414470	EU
Minor	In results of export with ObjectsExtractionParams::SourceContentReuseMode=CRM_Content Only the size and the boldness of print are different from those of text content in PDF file.	Synthesis	398687	RU
Minor	During call of FEngine::LoadModule with FREM_EuropeanPatterns parameter page counter decreases on one page.	API	416835	US

Minor	The content of IParagraph::Words sometimes differs from the content of IParagraph::Text.	Synthesis	410630	EU
Minor	The tabulation isn't saved in table cells.	Synthesis	302091	US
Minor	Extra spaces appear on the sides of letter "i" in the recognition results of addresses in envelopes in Italian.	Recognizer	364232	EU
Minor	Some text is lost during analysis stage on specific documents.	DA	407460	EU
Minor	A table on specific document is detected as three text blocks. <b>Workaround:</b> define table block manually	DA	412314	
Trivial	User profiles are case sensitive. The property specified in user profile isn't change if it is defined with case errors in a profile. In this release the property doesn't depend on the case.	API	406671	US

## KNOWN ISSUES AND WORKAROUNDS

### SOME API IS NOT IMPLEMENTED

The following API is not implemented in FRE 10 and FRE 11:

- IFootnoteSeries::IsNumberingWithSuperscript. Always returns "false".
- IFootnoteSeries::PositionOnPage. Always returns "FPPT\_SingleColumnSection".
- IFootnoteSeries::PositionInDocument. Always returns "FPDT\_PageEnd".
- IFootnoteSeries::HasSeparator. Always returns "true".
- ITextPicture::ColumnNumber. Always returns "0".
- ICharParams::IsWordStart. Always returns "false". It is true only for character parameters got through IWordRecognitionVariants interface.
- IIncut::TextWrapping. Always returns "TW\_Undefined".
- IRunningTitlesSeriesText::HasSeparator. Always returns "false".

The implementation is not planned.

### IWORDS COLLECTION IS EMPTY AND IWORDRECOGNITIONVARIANTS AND ICHARACTERRECOGNITIONVARIANT RETURN NULL AFTER PROCESSING

The issue is reproduced after the recognition with methods of IEngine and IDocumentAnalyzer objects, developed only for one-page documents processing (e.g. IEngine::ProcessPage, IDocumentAnalyzer::PreprocessAnalyzeRecognizePage or IDocumentAnalyzer::RecognizePage). After the recognition IWords collection is empty and IWords collection properties aren't filled up, IWordRecognitionVariants and CharacterRecognitionVariant collections return NULL.

**Workaround:** use IFRDocument object or IFRPage object for processing.

This issue will be fixed in the next maintenance release.

---

## THE TEXT DOESN'T FIT THE SIZE OF CELLS DURING EXPORT TO XLSX.

The text doesn't fit the cells when it is exported to XLSX in ExactCopy mode.

---

## WRONG DETECTION OF FOOTERS AND HEADERS

In some cases extra footers and headers are detected on the processed documents.

---

## INJECTTEXTLAYER METHOD WORKS INCORRECTLY ON SOME JAPANESE IMAGES

New IFREngine::InjectTextLayer method that processes the input PDF file and creates and injects the text layer created from the recognized text in a searchable PDF file has a shortcoming. On some pictures with Japanese text the injected text layer contains wrong characters.

---

## IPDFMRCPARAMS::MONOCHROMETEXT DOESN'T WORK CORRECTLY

Different algorithms of compressions during export to PDF are used with IPDFMRCPParams::MonochromeText set to TRUE (HelpDesk request 416986).

---

## ERRORS DURING OPENING OF PDF FILE

Unknown errors during opening of PDF file (HelpDesk requests 414136, 395847).

---

## CSS FILES DURING EXPORT TO HTML

During export to HTML apart from .html files, .css files appears in the output results.

---

## PDF/A VALIDATION REPORT

The following issues are detected by Adobe Acrobat 11.0.6 (Preflight 11.0.6) for PDF/A files produced by this release of FRE 11:

1. Adobe Acrobat rarely detects an error for exported PDF/A files "Text cannot be mapped to Unicode". Exported files don't pass validation.
2. Adobe Acrobat rarely detects an error for exported PDF/A -1a files "Syntax problem: Array with more than 8191 elements The issue is reproduced on files with a lot of pages. Exported files don't pass validation.
3. Adobe Acrobat detects an error "Text is mapped to Unicode Private Use Area but no ActualText entry is present" for exported PDF/A-2a и PDF/A-3a in the scenario of conversion from PDF to PDF/A with SourceContentReuseMode set to CRM\_ContentOnly.