Java, C#, Visual Basic .NET Developer’s Guide to

Asprise OCR SDK 5.0

Prepared by: ASPRISE

May 2014

ALL RIGHTS RESERVED, ASPRISE © 2014

www.asprise.com
# Table of Contents

1 **Introduction** ................................................................................................................. 1  
   1.1 About OCR ................................................................................................................. 1  
   1.2 About Asprise OCR SDK ......................................................................................... 1  
   1.3 Features of Asprise OCR ......................................................................................... 1  
   1.4 Compatibility of Asprise OCR SDK .......................................................................... 2  

2 **Programming with Asprise OCR in Java** ................................................................. 4  
   2.1 Asprise OCR for Java Installation ............................................................................. 4  
   2.2 Development Environment Setup ........................................................................... 4  
   2.3 Jump Start .................................................................................................................. 4  
   2.4 Input and Output ......................................................................................................... 5  
      2.4.1 Plain-text Format ................................................................................................. 6  
      2.4.2 XML Format ........................................................................................................ 6  
      2.4.3 PDF Format ........................................................................................................ 6  
   2.5 Other Usages .............................................................................................................. 7  
      2.5.1 Recognizes text only or barcode only ............................................................... 7  
      2.5.2 Perform OCR on part of the image .................................................................. 8  
      2.5.3 Perform OCR on multiple input files in one shot ........................................... 8  
      2.5.4 Perform OCR on a certain page from the specified TIFF file ...................... 8  
      2.5.5 Perform OCR on a PDF input file .................................................................... 8  
   2.6 Multi-threading with Asprise OCR ........................................................................... 9
2.7  Software Packaging and Distribution

3  Programming with Asprise OCR in C#

3.1  Asprise OCR for .NET Installation

3.2  Sample Projects

3.3  Jump Start

3.4  Input and Output

3.4.1  Plain-text Format

3.4.2  XML Format

3.4.3  PDF Format

3.5  Other Usages

3.5.1  Recognizes text only or barcode only

3.5.2  Perform OCR on part of the image

3.5.3  Perform OCR on multiple input files in one shot

3.5.4  Perform OCR on a certain page from the specified TIFF file

3.6  Multi-threading with Asprise OCR

3.7  Software Packaging and Distribution

4  Programming with Asprise OCR in Visual Basic .NET

4.1  Asprise OCR for .NET Installation

4.2  Sample Projects

4.3  Jump Start

4.4  Input and Output

4.4.1  Plain-text Format
6.2.1 Sample Use ....................................................................................................................................... 31
6.2.2 Supported Image Formats ........................................................................................................... 31
6.2.3 Compatibility ................................................................................................................................... 32
6.2.4 Software Packaging and Distribution ........................................................................................... 32
1 Introduction

1.1 About OCR

OCR (Optical Character Recognition) is the technology that allows you to transform images (e.g., images scanned from paper documents) into editable text-based computer files.

1.2 About Asprise OCR SDK

Embedded a high performance OCR engine, Asprise OCR SDK is OCR software development kit that can be used with Java, C#, Visual Basic (VB). Net, Delphi, Borland C, etc. on multiple platforms – Windows, Linux, Mac OS, Solaris and AIX.

1.3 Features of Asprise OCR

An incomplete list of features offered by Asprise OCR:

- **Highest Level of Accuracy**
  Asprise OCR can easily recognize difficult documents of poor image quality

- **Excellent Format Retention**
  Text layouts on the input documents are preserved;

- **High Speed**
  Asprise OCR uses optimized OCR engine to perform excellent recognition in very short time;

- **Ease of Use**
  We strive to make the developer’s life easier. Complex parameter configurations are removed from Asprise OCR SDK. You only have to supply the image document. Asprise OCR can intelligently determine the best setting internally.

- **Barcode Recognition**
  Beside characters (letters and numbers), Asprise OCR can recognize almost every kind of bar code. You can choose to recognize barcode or characters or both.
Currently, the following bar code formats are supported:

- CODE 128 (128b, 128C, 128raw)
- EAN 8 EAN 13
- UPC
- code 3 of 9
- code interleaved 2 of 5
- QR code

1.4 Compatibility of Asprise OCR SDK

Currently the following OS are supported:

<table>
<thead>
<tr>
<th>OS</th>
<th>Evaluation version availability</th>
<th>Licensed version availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows XP 32bit</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
</tr>
<tr>
<td>Windows XP 64bit</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
</tr>
<tr>
<td>Windows Vista 32bit</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
</tr>
<tr>
<td>Windows Vista 64bit</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
</tr>
<tr>
<td>Windows Server 32bit</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
</tr>
<tr>
<td>Windows Server 64bit</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
</tr>
<tr>
<td>Windows 7 32bit</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
</tr>
<tr>
<td>Windows 7 64bit</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
</tr>
<tr>
<td>Windows 8 32bit</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
</tr>
<tr>
<td>Windows 8 64bit</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
<td>Yes (Java, VB, C, C++, VB.NET &amp; C#)</td>
</tr>
<tr>
<td>Platform</td>
<td>Support</td>
<td>Tool Support</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>----------------</td>
</tr>
<tr>
<td>MacOS X PowerPC</td>
<td>Yes</td>
<td>Java &amp; command line tool</td>
</tr>
<tr>
<td>MacOS X Intel</td>
<td>Yes</td>
<td>Java &amp; command line tool</td>
</tr>
<tr>
<td>Linux 32bit</td>
<td>Yes</td>
<td>Java &amp; command line tool</td>
</tr>
<tr>
<td>Linux 64bit</td>
<td>Yes</td>
<td>Java &amp; command line tool</td>
</tr>
<tr>
<td>Solaris SPARC</td>
<td>Yes</td>
<td>Java &amp; command line tool</td>
</tr>
<tr>
<td>Solaris x86</td>
<td>Yes</td>
<td>Java &amp; command line tool</td>
</tr>
<tr>
<td>AIX</td>
<td>No</td>
<td>Please evaluate on other platforms</td>
</tr>
<tr>
<td>HP-UX</td>
<td>No</td>
<td>Please evaluate on other platforms</td>
</tr>
</tbody>
</table>
2 Programming with Asprise OCR in Java

2.1 Asprise OCR for Java Installation

First, make sure that you have already installed Java runtime version 1.5 or above on your system.

Download a copy of Asprise OCR SDK from http://www.asprise.com/product/ocr. Simply unzip it to an empty folder. Let refer this folder as OCR_HOME.

The file organization of Asprise OCR SDK distribution is as follows:

OCR_HOME
+--- images [folder, containing sample image documents]
+--- javadoc [Java docs]
+--- aocr.jar [Contains Asprise OCR classes]
+--- Asprise-OCR-SDK-Developers-Guide.pdf [This document]
+--- runDemox.bat/sh [OCR demos on different images]

2.2 Development Environment Setup

After you have obtained and unzipped the Asprise OCR SDK kit, you need to setup your development environment in order to develop Java applications with Asprise OCR. To do so, you only need to:

- Put aocr.jar into your class path

2.3 Jump Start

The following code demonstrates the basic usage of Asprise OCR:

```java
import com.asprise.ocr.Ocr

Ocr.setUp(); // one time setup
Ocr ocr = new Ocr(); // create a new OCR engine
ocr.startEngine("eng", Ocr.SPEED_FASTEST); // English
```
String s = ocr.recognize(new File[] {new File("test.png")}, Ocr.RECOGNIZE_TYPE_ALL, Ocr.OUTPUT_FORMAT_PLAINTEXT);
System.out.println("Result: "+ s);
ocr.stopEngine();

Line 1: Imports the Ocr class;

Line 5: Creates a new Ocr engine that recognizes English in fastest speed setting;

Line 8: All the OCR work is done here. The recognizeAll method of the com.asprise.util.ocr.OCR class recognizes all the characters and barcodes from the image and output them as a string.

Line 7 Performs the OCR on the specified file. The output format is set as plain text. Other supported formats are: XML (Ocr.OUTPUT_FORMAT_XML) and searchable PDF(Ocr.OUTPUT_FORMAT_PDF).

2.4 Input and Output

For the sample OCR code in above section, if the input is the figure below:

Asprise OCR and Barcode Recognition
High performance, royalty-free OCR and barcode recognition on Windows, Linux, Mac OS and Unix.
Asprise OCR (optical character recognition) and barcode recognition SDK offers a high performance library for you to equip your Java applications (Java applets, web applications, Swing/javadoc components, JEE enterprise applications), C#/VB.NET applications, and CO/Python applications with functionality of extracting text and barcode information from scanned documents.

Convert Images To Searchable PDF
With a few lines of code, you can convert various formats of images such as JPEG, PNG, TIFF, and PDF into searchable PDF files.

All Popular Barcode Formats

Then the output will be:
2.4.1 Plain-text Format

Asprise OCR and Barcode Recognition

High performance, royalty-free OCR and barcode recognition on Windows, ...

[QR-Code: www.asprise.com]
[CODE-128: Asprise].

The last two lines represents bar code information extract. Note both the format and content of the barcode are enclosed in '[[ ]]' pairs.

2.4.2 XML Format

If you set the output format as OUTPUT_FORMAT_XML, you’ll get:

```xml
<?xml version="1.0"?>
<asprise-ocr input="test.png">
  <page no="1" width="2400" height="3200" words="122" confidence="88">
    <block type="text" x="321" y="254" width="1438" height="76" words="5" confidence="86">
      Asprise OCR and Barcode Recognition
    </block>
    ...
    <block type="barcode" subtype="QR-Code" x="1397" y="1484" width="441" height="441" confidence="1">
      www.asprise.com
    </block>
    <block type="barcode" subtype="CODE-128" x="348" y="1489" width="583" height="145" confidence="34">
      Asprise
    </block>
  </page>
</asprise-ocr>
```

Note that the coordinate information and confidence are available in the output.

2.4.3 PDF Format

If you set the output format as OUTPUT_FORMAT_PDF, you need specify the target output file as following:

```java
ocr.recognize(new File[] {new File("test.png")},
    Ocr.RECOGNIZE_TYPE_ALL, Ocr.OUTPUT_FORMAT_PDF,
    Ocr.PROP_PDF_OUTPUT_FILE, "searchable.pdf",
    Ocr.PROP_PDF_OUTPUT_TEXT_VISIBLE, true);
```

Once the OCR done, you can open the PDF output file with any PDF viewer and perform searches:
Note recognized text and barcodes are rendered in orange on top of the image. To make them transparent, you simply set:

```
Ocr.PROP_PDF_OUTPUT_TEXT_VISIBLE, false
```

### 2.5 Other Usages

This section covers other essential functions offered by the Ocr class.

#### 2.5.1 Recognizes text only or barcode only

To save OCR time, you can choose to OCR text or barcode only:

**Text only:**

```java
String s = ocr.recognize(new File[] {new File("test.png")},
    Ocr.RECOGNIZE_TYPE_TEXT, Ocr.OUTPUT_FORMAT_PLAINTEXT);
```

**Barcode only:**

```java
String s = ocr.recognize(new File[] {new File("test.png")},
    Ocr.RECOGNIZE_TYPE_BARCODE, Ocr.OUTPUT_FORMAT_PLAINTEXT);
```
2.5.2 Perform OCR on part of the image

In some cases, you might not want to OCR the whole image. In that case, you can OCR on part of the image to save time:

```java
String s = ocr.recognize("C:/test.png", -1, 0, 0, 400, 200, Ocr.RECOGNIZE_TYPE_TEXT, Ocr.OUTPUT_FORMAT_PLAINTEXT);
```

The above code OCR the top left part of the image with width 400 pixels and height 200 pixels.

2.5.3 Perform OCR on multiple input files in one shot

You can perform OCR on multiple files in one shot:

```java
String s = ocr.recognize("C:/test.png;C:/test2.jpg", -1, 0, 0, 400, 200, Ocr.RECOGNIZE_TYPE_TEXT, Ocr.OUTPUT_FORMAT_PLAINTEXT);
```

2.5.4 Perform OCR on a certain page from the specified TIFF file

A TIFF file may contain multiple pages. If you need to recognize only a certain page, you can specify the page number as following:

```java
String s = ocr.recognize("C:/test.png", 2, -1, -1, -1, -1, Ocr.RECOGNIZE_TYPE_TEXT, Ocr.OUTPUT_FORMAT_PLAINTEXT);
```

Note 2 means the second page (the page number of the first page is 1).

2.5.5 Perform OCR on a PDF input file

You use the following method to perform OCR on a PDF input file:

```java
String s = ocr.recognize("C:/test.pdf", -1, 100, 100, 400, 200, Ocr.RECOGNIZE_TYPE_TEXT, Ocr.OUTPUT_FORMAT_PLAINTEXT);
```
2.6 Multi-threading with Asprise OCR

Most of modern computers run on multi-core CPUs. Multi-threading can significantly reduce the runtime by making full use of the processor power. It is easy to do multi-threading with Asprise OCR.

You can implement thread pool pattern\(^1\) by yourself with the `com.asprise.ocr.Ocr` class. Note that each thread must have its own dedicated OCR engine instance.

However, you do not need to start from scratch as Asprise OCR offer `com.asprise.ocr.OcrExecutorService` class that wraps OCR functionality in an executor service\(^2\). OcrExecutorService frees you from writing tricky thread management code.

Code sample:

```java
OcrExecutorService oes =
    new OcrExecutorService("eng", Ocr.SPEED_FASTEST, 4); // 4 threads
List<Future<String>> futures = oes.invokeAll(Arrays.asList(
    new OcrExecutorService.OcrCallable(
        new File[] {new File("test1.png")},
        Ocr.RECOGNIZE_TYPE_ALL, Ocr.OUTPUT_FORMAT_XML),
    new OcrExecutorService.OcrCallable(
        new File[] {new File("test2.png")},
        Ocr.RECOGNIZE_TYPE_ALL, Ocr.OUTPUT_FORMAT_XML)
));
System.out.println("Result of test1.png: " + futures.get(0).get());
System.out.println("Result of test2.png: " + futures.get(1).get());
oes.shutdown(); // stops all OCR engines and disposes all threads
```

2.7 Software Packaging and Distribution

So you have successfully developed your Java applications with Asprise OCR. It's time to distribute your programs to end users. First, make sure you are an authorized licensee registered with Asprise. To purchase a license, please visit:
http://www.asprise.com/product/ocr

---

\(^1\) http://en.wikipedia.org/wiki/Thread_pool_pattern

\(^2\) http://docs.oracle.com/javase/1.5.0/docs/api/java/util/concurrent/ExecutorService.html
You need to distribute aocr.jar along with your other jar files.
3 Programming with Asprise OCR in C#

3.1 Asprise OCR for .NET Installation

First, make sure that you have already installed .NET framework version 4.0 or above on your system.

Download a copy of Asprise OCR SDK from http://www.asprise.com/product/ocr. Simply unzip it to an empty folder. Let refer this folder as OCR_HOME.

The file organization of Asprise OCR SDK distribution is as follows:

OCR_HOME
   +--- images [folder, containing sample image documents]
   +--- sample-projects [sample .NET projects]
   +--- aocr.dll, aocr_x64.dll [OCR dll]
   +--- Asprise-OCR-SDK-Developers-Guide.pdf [This document]

3.2 Sample Projects

Sample projects show you how to invoke Asprise OCR functionalities using runnable code.

After you have obtained and unzipped the Asprise OCR SDK kit, you can open the sample projects using Visual Studio 2010 or later versions.

The asprise-ocr-api project is a library project that provides OCR API access for the .NET world. You can access it in any C# or Visual Basic project by simply adding a reference:
Running the C# sample project, you’ll see the window as following:

Select an image and click the ‘OCR’ button to perform OCR.

### 3.3 Jump Start

The following code demonstrates the basic usage of Asprise OCR:

```csharp
using asprise_ocr_api;
AspriseOCR.SetUp();
AspriseOCR ocr = new AspriseOCR();
ocr.StartEngine("eng", AspriseOCR.SPEED_FASTEST);
string s = ocr.Recognize("C:\path\img.jpg", -1, -1, -1, -1, -1, AspriseOCR.RECOGNIZE_TYPE_ALL, AspriseOCR.OUTPUT_FORMAT_PLAINTEXT);
```
8     // process more images here ...
9
10    ocr.StopEngine();

Line 3: performs one-time setup if it has not been done;

Lines 4 & 5: Creates a new Ocr engine that recognizes English in fastest speed setting;

Line 7: All the OCR work is done here. The *Recognize* method of the *AspriseOCR* class recognizes all the characters and barcodes from the image and output them as plain text. Other supported output formats are: XML (*AspriseOCR.OUTPUT_FORMAT_XML*) and searchable PDF (*AspriseOCR.OUTPUT_FORMAT_PDF*).

### 3.4 Input and Output

For the sample OCR code in above section, if the input is the figure below:

*Asprise OCR and Barcode Recognition*

High performance, royalty-free OCR and barcode recognition on Windows, Linux, Mac OS and Unix.

Asprise OCR (optical character recognition) and barcode recognition SDK offers a high performance library for you to equip your Java applications (Java applets, web applications, Swing/JavaFX components, J2EE enterprise applications), C#/.NET applications, and C/C++/Python applications with functionality of extracting text and barcode information from scanned documents.

**Convert Images To Searchable PDF**

With a few lines of code, you can convert various formats of images such as JPEG, PNG, TIFF, and PDF into searchable PDF files.

**All Popular Barcode Formats**


![Barcode Image]

Then the output will be:

### 3.4.1 Plain-text Format

*Asprise OCR and Barcode Recognition*
High performance, royalty-free OCR and barcode recognition on Windows, ...

The last two lines represents bar code information extract. Note both the format and content of the barcode are enclosed in '[[]]' pairs.

### 3.4.2 XML Format

If you set the output format as OUTPUT_FORMAT_XML, you’ll get:

```xml
<?xml version="1.0"?>
<asprise-ocr input="test.png">
  <page no="1" width="2400" height="3200" words="122" confidence="88">
    <block type="text" x="321" y="254" width="1438" height="76" words="5" confidence="86">
      Asprise OCR and Barcode Recognition
    </block>
    ...
    <block type="barcode" subtype="QR-Code" x="1397" y="1484" width="441" height="441" confidence="1">
      www.asprise.com
    </block>
    <block type="barcode" subtype="CODE-128" x="348" y="1489" width="583" height="145" confidence="34">
      Asprise
    </block>
  </page>
</asprise-ocr>
```

Note that the coordinate information and confidence are available in the output.

### 3.4.3 PDF Format

If you set the output format as OUTPUT_FORMAT_PDF, you need specify the target output file as following:

```csharp
string s = ocr.Recognize("C:\path\img.jpg", -1, -1, -1, -1, -1, AspriseOCR.RECOGNIZE_TYPE_ALL, AspriseOCR.OUTPUT_FORMAT_PDF, AspriseOCR.PROP_PDF_OUTPUT_FILE, "c:\output.pdf", AspriseOCR.PROP_PDF_OUTPUT_TEXT_VISIBLE, "true", // render in color AspriseOCR.PROP_PDF_OUTPUT_IMAGE_FORCE_BW, "true");
```

Once the OCR done, you can open the PDF output file with any PDF viewer and perform searches:
3.5 Other Usages

This section covers other essential functions offered by the AspriseOCR class.

3.5.1 Recognizes text only or barcode only

To save OCR time, you can choose to OCR text or barcode only:

Text only:

```csharp
string s = ocr.Recognize("C:\path\img.jpg", -1, -1, -1, -1, -1,
AspriseOCR.RECOGNIZE_TYPE_TEXT, AspriseOCR.OUTPUT_FORMAT_PLAINTEXT);
```

Barcode only:

```csharp
string s = ocr.Recognize("C:\path\img.jpg", -1, -1, -1, -1, -1,
AspriseOCR.RECOGNIZE_TYPE_BARCODE, AspriseOCR.OUTPUT_FORMAT_PLAINTEXT);
```
3.5.2 Perform OCR on part of the image

In some cases, you might not want to OCR the whole image. In that case, you can OCR on part of the image to save time:

```csharp
string s = ocr.Recognize("C:\\path\\img.jpg", -1, 0, 0, 400, 200,
AspriseOCR.RECOGNIZE_TYPE_ALL, AspriseOCR.OUTPUT_FORMAT_PLAINTEXT);
```

The above code OCR the top left part of the image with width 400 pixels and height 200 pixels.

3.5.3 Perform OCR on multiple input files in one shot

You can perform OCR on multiple files in one shot:

```csharp
string s = ocr.Recognize("C:\img1.jpg;C:\img2.png", -1, -1, -1, -1,
-1, AspriseOCR.RECOGNIZE_TYPE_ALL,
AspriseOCR.OUTPUT_FORMAT_PLAINTEXT);
```

3.5.4 Perform OCR on a certain page from the specified TIFF file

A TIFF file may contain multiple pages. If you need to recognize only a certain page, you can specify the page number as following:

```csharp
string s = ocr.Recognize("C:\img1.tif", 2, -1, -1, -1, -1,
AspriseOCR.RECOGNIZE_TYPE_ALL, AspriseOCR.OUTPUT_FORMAT_PLAINTEXT);
```

Note 2 means the second page (the page number of the first page is 1).

3.6 Multi-threading with Asprise OCR

Most of modern computers run on multi-core CPUs. Multi-threading can significantly reduce the runtime by making full use of the processor power. It is easy to do multi-threading with Asprise OCR.
You can implement thread pool pattern\(^3\) by yourself or you can use Microsoft’s default implementation of thread pooling\(^4\). Note that each thread must have its own dedicated OCR engine instance.

### 3.7 Software Packaging and Distribution

So you have successfully developed your C# applications with Asprise OCR. It’s time to distribute your programs to end users. First, make sure you are an authorized licensee registered with Asprise. To purchase a license, please visit:

http://www.asprise.com/product/ocr

You need to distribute aocr.dll along with your own binary code. Dll files should be put into one of the folders in system PATH.

---


4  Programming with Asprise OCR in Visual Basic .NET

4.1 Asprise OCR for .NET Installation

First, make sure that you have already installed .NET framework version 4.0 or above on your system.

Download a copy of Asprise OCR SDK from http://www.asprise.com/product/ocr. Simply unzip it to an empty folder. Let refer this folder as OCR_HOME.

The file organization of Asprise OCR SDK distribution is as follows:

OCR_HOME
|-- images [folder, containing sample image documents]
   `-- sample-projects [sample .NET projects]
   `-- aocr.dll, aocr_x64.dll [OCR dll]
   `-- Asprise-OCR-SDK-Developers-Guide.pdf [This document]

4.2 Sample Projects

Sample projects show you how to invoke Asprise OCR functionalities using runnable code.

After you have obtained and unzipped the Asprise OCR SDK kit, you can open the sample projects using Visual Studio 2010 or later versions.

The asprise-ocr-api project is a library project that provides OCR API access for the .NET world. You can access it in any C# or Visual Basic project by simply adding a reference:
Running the C# sample project, you’ll see the window as following:

Select an image and click the ‘OCR’ button to perform OCR.

4.3 Jump Start

The following code demonstrates the basic usage of Asprise OCR:

```
Imports asprise_ocr_api
Private ocr As AspriseOCR
AspriseOCR.SetUp()
ocr = New AspriseOCR()
ocr.StartEngine("eng", AspriseOCR.SPEED_FASTEST)
```
Dim s As String = ocr.Recognize("C:\img.jpg", -1, -1, -1, -1, -1, AspriseOCR.RECOGNIZE_TYPE_ALL, AspriseOCR.OUTPUT_FORMAT_PLAINTEXT)
' process more images here ...
ocr.StopEngine()

Line 5: performs one-time setup if it has not been done;

Lines 6 & 7: Creates a new Ocr engine that recognizes English in fastest speed setting;

Line 9: All the OCR work is done here. The Recognize method of the AspriseOCR class recognizes all the characters and barcodes from the image and output them as plain text. Other supported output formats are: XML (AspriseOCR.OUTPUT_FORMAT_XML) and searchable PDF(AspriseOCR.OUTPUT_FORMAT_PDF).

4.4 Input and Output

For the sample OCR code in above section, if the input is the figure below:

Asprise OCR and Barcode Recognition
High performance, royalty-free OCR and barcode recognition on Windows, Linux, Mac OS and Unix.

Asprise OCR (optical character recognition) and barcode recognition SDK offers a high performance library for you to equip your Java applications (Java applets, web applications, Swing/javadoc components, JEE enterprise applications), C#/.NET applications, and C/C++-Python applications with functionality of extracting text and barcode information from scanned documents.

Convert Images To Searchable PDF
With a few lines of code, you can convert various formats of images such as JPEG, PNG, TIFF, and PDF into searchable PDF files.

All Popular Barcode Formats

Then the output will be:
4.4.1 Plain-text Format

Asprise OCR and Barcode Recognition

High performance, royalty-free OCR and barcode recognition on Windows, ...

[[QR-Code: www.asprise.com]]
[[CODE-128: Asprise]].

The last two lines represents bar code information extract. Note both the format and content of the barcode are enclosed in '[[ ]]' pairs.

4.4.2 XML Format

If you set the output format as OUTPUT_FORMAT_XML, you’ll get:

```xml
<?xml version="1.0"?><asprise-ocr input="test.png"><page no="1" width="2400" height="3200" words="122"
confidence="88">  <block type="text" x="321" y="254" width="1438" height="76"
words="5" confidence="86">Asprise OCR and Barcode Recognition</block>
...  <block type="barcode" subtype="QR-Code" x="1397" y="1484"
width="441" height="441" confidence="1">www.asprise.com</block>
  <block type="barcode" subtype="CODE-128" x="348" y="1489"
width="583" height="145" confidence="34">Asprise</block>
</page></asprise-ocr>
```

Note that the coordinate information and confidence are available in the output.

4.4.3 PDF Format

If you set the output format as OUTPUT_FORMAT_PDF, you need specify the target output file as following:

```csharp
s = ocr.Recognize("C:\img.jpg", -1, -1, -1, -1, -1,
AspriseOCR.RECOGNIZE_TYPE_ALL, AspriseOCR.OUTPUT_FORMAT_PDF,
AspriseOCR.PROP_PDF_OUTPUT_FILE, "C:\output.pdf",
AspriseOCR.PROP_PDF_OUTPUT_TEXT_VISIBLE, "true", // render in color
AspriseOCR.PROP_PDF_OUTPUT_IMAGE_FORCE_BW, "true")
```

Once the OCR done, you can open the PDF output file with any PDF viewer and perform searches:
Note recognized text and barcodes are rendered in orange on top of the image. To make them transparent, you simply set:

AspriseOCR.PROP_PDF_OUTPUT_TEXT_VISIBLE, "false"

4.5 Other Usages

This section covers other essential functions offered by the AspriseOCR class.

4.5.1 Recognizes text only or barcode only

To save OCR time, you can choose to OCR text or barcode only:

Text only:

s = ocr.Recognize("C:\\img.jpg", -1, -1, -1, -1, -1, AspriseOCR.RECOGNIZE_TYPE_TEXT, AspriseOCR.OUTPUT_FORMAT_PLAINTEXT)

Barcode only:

s = ocr.Recognize("C:\\img.jpg", -1, -1, -1, -1, -1, AspriseOCR.RECOGNIZE_TYPE_ALL, AspriseOCR.OUTPUT_FORMAT_PLAINTEXT)
4.5.2 Perform OCR on part of the image

In some cases, you might not want to OCR the whole image. In that case, you can OCR on part of the image to save time:

```python
s = ocr.Recognize("C:\img.jpg", -1, 0, 0, 400, 200,
AspriseOCR.RECOGNIZE_TYPE_ALL, AspriseOCR.OUTPUT_FORMAT_PLAINTEXT)
```

The above code OCR the top left part of the image with width 400 pixels and height 200 pixels.

4.5.3 Perform OCR on multiple input files in one shot

You can perform OCR on multiple files in one shot:

```python
s = ocr.Recognize("C:\img.jpg;C:\2.jpg", -1, -1, -1, -1, -1,
AspriseOCR.RECOGNIZE_TYPE_ALL, AspriseOCR.OUTPUT_FORMAT_PLAINTEXT)
```

4.5.4 Perform OCR on a certain page from the specified TIFF file

A TIFF file may contain multiple pages. If you need to recognize only a certain page, you can specify the page number as following:

```python
s = ocr.Recognize("C:\img.tif", 2, -1, -1, -1, -1,
AspriseOCR.RECOGNIZE_TYPE_ALL, AspriseOCR.OUTPUT_FORMAT_PLAINTEXT)
```

Note 2 means the second page (the page number of the first page is 1).

4.6 Multi-threading with Asprise OCR

Most of modern computers run on multi-core CPUs. Multi-threading can significantly reduce the runtime by making full use of the processor power. It is easy to do multi-threading with Asprise OCR.
You can implement thread pool pattern\(^5\) by yourself or you can use Microsoft’s default implementation of thread pooling\(^6\). Note that each thread must have its own dedicated OCR engine instance.

### 4.7 Software Packaging and Distribution

So you have successfully developed your Visual Basic .NET applications with Asprise OCR. It’s time to distribute your programs to end users. First, make sure you are an authorized licensee registered with Asprise. To purchase a license, please visit: [http://www.asprise.com/product/ocr](http://www.asprise.com/product/ocr)

You need to distribute aocr.dll along with your own binary code. Dll files should be put into one of the folders in system PATH.


5  Support and Professional Services

5.1 Support Web Site

http://www.asprise.com/product/ocr

5.2 Basic Support

Our team provides basic support for general Asprise OCR developers. Email your technical questions to support@asprise.com. Most of enquiries get reply within 5 business days although some may take longer. Please consider to subscribe our premium support service in order to get your problems solved quickly.

5.3 Premium Support Services + Updates

For premium support service subscribers: your requests will be of our top priority and your email will be replied within 1 business day. We’ll send you software update promptly.

5.4 Professional Services

We provide custom professional services such as software integration, OCR font training, etc.
6 Appendix: Java Image Acquisition Components

The image acquisition UI components are not part of Asprise OCR library. However, based on our customers’ experience, if you need to build a front-end for OCR, they are invaluable and could save you a lot of time. Otherwise, you may skip this chapter.

6.1 JImageDialog

JImageDialog is an image acquisition UI component that allows the user to load images and to perform basic image editing tasks. If you are developing some applications that require the user to select/edit/input images, then JImageDialog will make your life extremely easy – and more importantly, the user experience will be improved dramatically.
Let say you want to build an album application, the user is required to supply photos (i.e. images). You put a button on your panel. When the user click the button, JImageDialog is brought up – now the user can select existing pictures files from his or her computer or acquire images from digital cameras or scanners. And the user can edit images before putting it into the album.

6.1.1 Advantages

- Multiple image sources supported: local computer, digital cameras, scanners and the web;
- Multiple image formats: read and write BMP, PNG, JPG, GIF, PCT, PSD and many other formats;
- Platform/Virtual machine independent: Any platform, any Java virtual machine (version 1.3 or above);
- Powerful features: rotation, flipping, scaling, clipping, etc.
- User friendly as well as developer friendly.

The user can load images from local computer or the web, he or she can also acquire images from digital cameras and scanners. After the image has been loaded, the user can rotate, clip, flip, and scale the image. The image has been loaded and edited, the user can save the image or select the image - which will be used in your applications.

6.1.2 Sample Uses

**Modal (synchronous) mode**

```java
1. JImageDialog dialog = new JImageDialog(frame, "Sample", true);
   // Modal dialog
2. BufferedImage image = dialog.showDialog();
3. ...
```

Line 1 constructs the image dialog.

Line 2 brings up the image dialog and waiting for user's selection/acquisition.
Besides using JImageDialog in synchronous mode, you can also use it in:

**Asynchronous mode**

```java
public class JImageDialogSample extends JPanel implements JImageDialogListener {
    ...
    BufferedImage image;

    // Displays selected image if any.
    public void paintComponent(Graphics g) {
        super.paintComponent(g); // Paint background.
        if(image != null)
            g.drawImage(image, 0, 0, null);
    }

    // Sets image and refreshes the panel.
    public void setImage(BufferedImage image) {
        this.image = image;
        setPreferredSize(getPreferredSize());
        revalidate();
        repaint();
    }

    // Methods in JImageDialogListener
    // When the user presses cancel button, this method will be called.
    public void onCancel() {
        setImage(null);
    }

    // When the user presses the selection button, will be invoked.
    public void onImageSet(BufferedImage image) {
        setImage(image);
    }

    ...
}
```

JImageDialogSample imagePanel = new JImageDialogSample();
JImageDialog dialog = new JImageDialog();
dialog.addImageDialogListener(imagePanel);
dialog.showDialog();

Line 1-30 implements a JImageDialogListener.
Line 33 constructs the listener.

Line 35 constructs the dialog.

Line 36 registers the listener the dialog

Line 37 brings up the dialog

When the user acquires an image and selects it, JImageDialog's listeners will be notified. In this case, `imagePanel.onImageSet(BufferedImage image)` will be called and thus the panel will display the selected image. If the user cancels the selection, `onCancel()` will be called instead.

Sample application: `com.asprise.util.ui.JImageDialogSample`

### 6.1.3 Supported Image Formats

The following table shows image formats supported by JImageDialog:

<table>
<thead>
<tr>
<th>Formats</th>
<th>File extensions</th>
<th>READ</th>
<th>WRITE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adobe Photoshop</td>
<td>*.psd</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Bitmap, Windows/OS2</td>
<td>*.bmp, *.dib</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Cursor</td>
<td>*.cur</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graphics Interchange Format</td>
<td>*.gif</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Icon</td>
<td>*.ico</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>JPEG</td>
<td>*.jpg, *.jpeg</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Macintosh PICT Format</td>
<td>*.pict, *.pct</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>PCX Format</td>
<td>*.pcx</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Portable Network Graphics</td>
<td>*.png</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Sun Raster Format</td>
<td>*.ras</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Tag Image File Format</td>
<td>*.tif, *.tiff</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Targa</td>
<td>*.tga</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>X Bitmap</td>
<td>*.xbm</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>X PixMap</td>
<td>*.xpm</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

On any Java platforms (version 1.3 or above), JImageDialog supports the above formats (using its own library to read/write image files). JImageDialog intelligently selects the best way to read or write files – e.g. on Java 1.4, it may invoke ImageIO to
see whether a file can be read or written; if the ImageIO can do the job then JImageDialog will let it do; otherwise, JImageDialog will use its own library to access the file.

Note: You can only read/write image files from the JImageDialog UI component with unlicensed image acquisition UI component package. If you want to access image files from your Java code and/or to perform other advanced operations, you need to obtain an affordable license from LAB Asprise!

6.1.4 Compatibility

All Java runtimes with version 1.3 or above.

6.1.5 Software Packaging and Distribution

Mandatory:  
  jid.jar, JTwain.jar

6.2 JImageFileChooser

JImageFileChooser is an extended JFileChooser that supports image preview and image information extraction.
When the user clicks an image file, its preview and associated information will be displayed to assist the user to select the proper image.

### 6.2.1 Sample Use

```java
JFileChooser fc = new JImageFileChooser(lastDirectory);
fc.addChoosableFileFilter(JImageFileChooser.getImageFileFilter());
int returnVal = fc.showOpenDialog(frame);
...
```

Line 1 creates the image file chooser;

Line 2 set the file filter.

You can use it as normal JFileChooser, and it improves the user experience greatly.

### 6.2.2 Supported Image Formats

Please refer to Supported Image Formats in JImageDialog section.

Note: You can only preview image files from the JImageFileChooser UI component with unlicensed image acquisition UI component package. If you want to read/write
image files from your Java code with the package and/or to perform other advanced operations, you need to obtain an affordable license from LAB Asprise!

6.2.3 Compatibility

All operating systems; All Java runtimes with version 1.2 or above.

6.2.4 Software Packaging and Distribution

Mandatory:  jid.jar
Third-Party License Information

This SDK may contain third-party code. Asprise is required to provide the following notices. Note, however, that the SDK license that accompanied this product determines your right to use the Asprise program, including the third-party software, and the terms contained in the following notices do not change those rights.

This section is common across all SDKs offered by Asprise. This SDK may not contain all components listed below.

Apache License Version 2.0

TERMS AND CONDITIONS FOR USE, REPRODUCTION, AND DISTRIBUTION

1. Definitions.

“License” shall mean the terms and conditions for use, reproduction, and distribution as defined by Sections 1 through 9 of this document.

“Licensor” shall mean the copyright owner or entity authorized by the copyright owner that is granting the License.

“Legal Entity” shall mean the union of the acting entity and all other entities that control, are controlled by, or are under common control with that entity. For the purposes of this definition, “control” means (i) the power, direct or indirect, to cause the direction or management of such entity, whether by contract or otherwise, or (ii) ownership of fifty percent (50%) or more of the outstanding shares, or (iii) beneficial ownership of such entity.

“You” (or “Your”) shall mean an individual or Legal Entity exercising permissions granted by this License.

“Source” form shall mean the preferred form for making modifications, including but not limited to software source code, documentation source, and configuration files.

“Object” form shall mean any form resulting from mechanical transformation or translation of a Source form, including but not limited to compiled object code, generated documentation, and conversions to other media types.

“Work” shall mean the work of authorship, whether in Source or Object form, made available under the License, as indicated by a copyright notice that is included in or attached to the work (an example is provided in the Appendix below).

“Derivative Works” shall mean any work, whether in Source or Object form, that is based on (or derived from) the Work and for which the editorial revisions, annotations, elaborations, or other modifications represent, as a whole, an original work of authorship. For the purposes of this License, Derivative Works shall not include works that remain separable from, or merely link (or bind by name) to the interfaces of, the Work and Derivative Works thereof.

“Contribution” shall mean any work of authorship, including the original version of the Work and any modifications or additions to that Work or Derivative Works thereof, that is intentionally submitted to Licensor for inclusion in the Work by the copyright owner or by an individual or Legal Entity authorized to submit on behalf of the copyright owner. For the purposes of this definition, “submitted” means any form of electronic, verbal, or written communication sent to the Licensor or its representatives, including but not limited to communication on electronic mailing lists, source code control systems, and issue tracking systems that are managed by, or on behalf of, the Licensor for the purpose of discussing and improving the Work, but excluding communication that is conspicuously marked or otherwise designated in writing by the copyright owner as “Not a Contribution.”

“Contributor” shall mean Licensor and any individual or Legal Entity on behalf of whom a Contribution has been received by Licensor and subsequently incorporated within the Work.

2. Grant of Copyright License. Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable copyright license to reproduce, prepare Derivative Works of, publicly display, publicly perform, sublicense, and distribute the Work and such Derivative Works in Source or Object form.

3. Grant of Patent License. Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable (except as stated in this section) patent license to make, have made, use, offer to sell, sell, import, and otherwise transfer the Work, where such license applies only to those patent claims licensable by such Contributor that are necessarily infringed by their Contribution(s) alone or by combination of their Contribution(s) with the Work to which such Contribution(s) was submitted. If You institute patent litigation against any entity (including a cross-claim or counterclaim in a lawsuit) alleging that the Work or a Contribution incorporated within the Work constitutes direct or contributory patent infringement, then any patent licenses granted to You under this License for that Work shall terminate as of the date such litigation is filed.

4. Redistribution. You may reproduce and distribute copies of the Work or Derivative Works thereof in any medium, with or without modifications, and in Source or Object form, provided that You meet the following conditions:

You must give any other recipients of the Work or Derivative Works a copy of this License; and
You must cause any modified files to carry prominent notices stating that You changed the files; and
You must retain, in the Source form of any Derivative Works that You distribute, all copyright, patent, trademark, and attribution notices from the Source form of the Work, excluding those notices that do not pertain to any part of the Derivative Works; and
If the Work includes a “NOTICE” text file as part of its distribution, then any Derivative Works that You distribute must include a readable copy of the attribution notices contained within such NOTICE file, excluding those notices that do not pertain to any part of the Derivative Works; and
If the Work is modified, You shall include a statement in the modified work denoting that such modifications were made.

NOTICE: This software relies on the following components: [list of components].
NOTICE text from the Work, provided that such additional attribution
notices cannot be construed as modifying the License.

You may add Your own copyright statement to Your modifications and may
provide additional or different license terms and conditions for use,
reproduction, or distribution of Your modifications, or for any such Derivative
Works as a whole, provided You use, reproduction, and distribution of the
Work otherwise complies with the conditions stated in this License.
5. Submission of Contributions. Unless You explicitly state otherwise, any
Contribution intentionally submitted for inclusion in the Work by You to the
Licensor shall be under the terms and conditions of this License, without any
additional terms or conditions. Notwithstanding the above, nothing herein
shall supersede or modify the terms of any separate license agreement you may
have executed with Licensor regarding such Contributions.
6. Trademarks. This License does not grant permission to use the trade names,
trademarks, service marks, or product names of the Licensor, except as
required for reasonable and customary use in describing the origin of the
Work and reproducing the content of the NOTICE file.
7.Disclaimer of Warranty. Unless required by applicable law or agreed to in
writing, Licensor provides the Work (and each Contributor provides its
Contributions) on an “AS IS” BASIS, WITHOUT WARRANTIES OR
CONDITIONS OF ANY KIND, either express or implied, including,
without limitation, any warranties or conditions of TITLE, NONINFRINGEMENT,
MERCHANTABILITY, or FITNESS FOR A PARTICULAR PURPOSE. You are solely responsible for determining the
appropriateness of using or redistributing the Work and assume any risks
associated with Your exercise of permissions under this License.
8. Limitation of Liability. In no event and under no legal theory, whether in
tort (including negligence), contract, or otherwise, unless required by
applicable law (such as deliberate and grossly negligent acts) or agreed to in
writing, shall any Contributor be liable to You for damages, including any
direct, indirect, special, incidental, or consequential damages of any kind
arising as a result of this License or out of the use or inability to use the Work
(including but not limited to damages for loss of goodwill, work stoppage,
computer failure or malfunction, or any and all other commercial damages or
losses), even if such Contributor has been advised of the possibility of such
damages.
9. Accepting Warranty or Additional Liability. While redistributing the Work
or Derivative Works thereof, You may choose to offer, and charge a fee for,
acceptance of support, warranty, indemnity, or other liability obligations
and/or rights consistent with this License. However, in accepting such
obligations, You may act only on Your own behalf and on Your sole
responsibility, not on behalf of any other Contributor, and only if You agree
to indemnify, defend, and hold each Contributor harmless for any liability
incurred by, or claims asserted against, such Contributor by reason of your
accepting any such warranty or additional liability.
Apache Jakarta Commons Lang
Copyright 2001-2007 The Apache Software Foundation

This product includes software developed by
The Apache Software Foundation (http://www.apache.org/).
Apache log4j

Copyright 2007 The Apache Software Foundation

This product includes software developed at
The Apache Software Foundation (http://www.apache.org/).

ICU License - ICU 1.8.1 and later

COPYRIGHT AND PERMISSION NOTICE

Copyright (c) 1995-2009 International Business Machines Corporation and others

All rights reserved.

Permission is hereby granted, free of charge, to any person obtaining a copy of
this software and associated documentation files (the “Software”), to deal in
the Software without restriction, including without limitation the rights to
use, copy, modify, merge, publish, distribute, and/or sell copies of the
Software, and to permit persons to whom the Software is furnished to do so,
provided that the above copyright notice(s) and this permission notice appear
in all copies of the Software and that both the above copyright notice(s) and
this permission notice appear in supporting documentation.

THE SOFTWARE IS PROVIDED “AS IS”, WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT
LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OF
THIRD PARTY RIGHTS. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR HOLDERS INCLUDED IN THIS NOTICE BE LIABLE
FOR ANY CLAIM, OR ANY SPECIAL INDIRECT OR CONSEQUENTIAL DAMAGES, OR ANY DAMAGES WHATSOEVER
RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER
TORTIOUS ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE.

Copyright (C) 2001 Leptonica. All rights reserved.

Redistribution and use in source and binary forms, with or without
modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright
notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above
copyright notice, this list of conditions and the following
disclaimer in the documentation and/or other materials provided with the distribution.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS
AND CONTRIBUTORS
“AS IS” AND ANY EXPRESS OR IMPLIED WARRANTIES,
INCLUDING, BUT NOT
LIMITED TO, THE IMPLIED WARRANTIES OF
MERCHANTABILITY AND FITNESS FOR
A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT
SHALL ANY
CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT,
INCIDENTAL, SPECIAL,
EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO,
PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR
PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY
OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING
NEGligence or OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS
SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Copyright (c) 1988-1997 Sam Leffler
Copyright (c) 1991-1997 Silicon Graphics, Inc.

Permission to use, copy, modify, distribute, and sell this software and
its documentation for any purpose is hereby granted without fee, provided
that (i) the above copyright notices and this permission notice appear in
all copies of the software and related documentation, and (ii) the names of
Sam Leffler and Silicon Graphics may not be used in any advertising or
publicity relating to the software without the specific, prior written
permission of Sam Leffler and Silicon Graphics.

THE SOFTWARE IS PROVIDED "AS-IS" AND WITHOUT
WARRANTY OF ANY KIND,
EXPRESS, IMPLIED OR OTHERWISE, INCLUDING WITHOUT
LIMITATION, ANY
WARRANTY OF MERCHANTABILITY OR FITNESS FOR A
PARTICULAR PURPOSE.

IN NO EVENT SHALL SAM LEFFLER OR SILICON GRAPHICS BE
LIABLE FOR
ANY SPECIAL, INCIDENTAL, INDIRECT OR CONSEQUENTIAL
DAMAGES OF ANY KIND,
OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF
USE, DATA OR PROFITS,
WHETHER OR NOT ADVISED OF THE POSSIBILITY OF DAMAGE,
AND ON ANY THEORY OF
LIABILITY, ARISING OUT OF OR IN CONNECTION WITH THE
USE OR PERFORMANCE
OF THIS SOFTWARE.

Copyright (C) 1994-1998, Thomas G. Lane.
* Modified 2003-2010 by Guido Vollbeding.
* This file is part of the Independent JPEG Group’s software.

SAX is free!

In fact, it’s not possible to own a license to SAX, since it’s been placed in
the public domain.

No Warranty

Because SAX is released to the public domain, there is no warranty for the
design or for the software implementation, to the extent permitted by
applicable law. Except when otherwise stated in writing the copyright holders
and/or other parties provide SAX “as is” without warranty of any kind, either
expressed or implied, including, but not limited to, the implied warranties
of merchantability and fitness for a particular purpose. The entire risk as
to the quality and performance of SAX is with you. Should SAX prove
defective, you assume the cost of all necessary servicing, repair or
correction.

In no event unless required by applicable law or agreed to in writing will
any copyright holder, or any other party who may modify and/or redistribute
SAX, be liable to you for damages, including any general, special, incidental
or consequential damages arising out of the use or inability to use SAX
(including but not limited to loss of data or data being rendered inaccurate
or losses sustained by you or third parties or a failure of the SAX to
operate with any other programs), even if such holder or other party has been
advised of the possibility of such damages.

Copyright Disclaimers

This page includes statements to that effect by David Megginson, who would
have been able to claim copyright for the original work. SAX 1.0

Version 1.0 of the Simple API for XML (SAX), created collectively by the
membership of the XML-DEV mailing list, is hereby released into the public
domain.

No one owns SAX: you may use it freely in both commercial and non-
commercial
applications, bundle it with your software distribution, include it on a
CD-ROM, list the source code in a book, mirror the documentation at your
own
web site, or use it in any other way you see fit.

David Megginson, sax@megginson.com
1998-05-11

SAX 2.0

I hereby abandon any property rights to SAX 2.0 (the Simple API for XML),
and
release all of the SAX 2.0 source code, compiled code, and documentation
contained in this distribution into the Public Domain. SAX comes with NO
WARRANTY or guarantee of fitness for any purpose.

David Megginson, david@megginson.com
2000-05-05

Tesseract Open Source
Copyright Ray Smith

This product includes software Licensed under the Apache License, Version
2.0.

Unicode Terms of Use

For the general privacy policy governing access to this site, see the Unicode
Privacy Policy. For trademark usage, see the Unicode® Consortium Name and
Trademark Usage Policy.

A. Unicode Copyright.
1. Copyright © 1991-2011 Unicode, Inc. All rights reserved.

2. Certain documents and files on this website contain a legend indicating that "Modification is permitted." Any person is hereby authorized, without fee, to modify such documents and files to create derivative works conforming to the Unicode Standard, subject to Terms and Conditions herein.

3. Any person is hereby authorized, without fee, to view, use, reproduce, and distribute all documents and files solely for informational purposes in the creation of products supporting the Unicode Standard, subject to the Terms and Conditions herein.

4. Further specifications of rights and restrictions pertaining to the use of the particular set of data files known as the "Unicode Character Database" can be found in Exhibit 1.

5. Each version of the Unicode Standard has further specifications of rights and restrictions of use. For the book editions (Unicode 5.0 and earlier), these are found on the back of the title page. The online code charts carry specific restrictions. All other files, including online documentation of the core specification for Unicode 6.0 and later, are covered under these general Terms of Use.

6. No license is granted to "mirror" the Unicode website where a fee is charged for access to the "mirror" site.

7. Modification is not permitted with respect to this document. All copies of this document must be verbatim.

B. Restricted Rights Legend. Any technical data or software which is licensed to the United States of America, its agencies and/or instrumentalities under this Agreement is commercial technical data or commercial computer software developed exclusively at private expense as defined in FAR 2.101, or DFARS 252.227-7014 (June 1995), as applicable. For technical data, use, modification, or disclosure by the Government is subject to restrictions as set forth in DFARS 202.227-7015 Technical Data, Commercial and Items (Nov 1995) and this Agreement. For Software, in accordance with FAR 12-212 or DFARS 227-7202, as applicable, use, modification or disclosure by the Government is subject to the restrictions set forth in this Agreement.

C. Warranties and Disclaimers.
1. This publication and/or website may include technical or typographical errors or other inaccuracies. Changes are periodically added to the information herein; these changes will be incorporated in new editions of the publication and/or website. Unicode may make improvements and/or changes in the product(s) and/or program(s) described in this publication and/or website at any time.

2. If this file has been purchased on magnetic or optical media from Unicode, Inc. the sole and exclusive remedy for any claim will be exchange of the defective media within ninety (90) days of original purchase.

3. EXCEPT AS PROVIDED IN SECTION C.2, THIS PUBLICATION AND/OR SOFTWARE IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND EITHER EXPRESS, IMPLIED, OR STATUTORY, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NON-INFRINGEMENT. UNICODE AND ITS LICENSORS ASSUME NO RESPONSIBILITY FOR ERRORS OR OMISSIONS IN THIS PUBLICATION AND/OR SOFTWARE OR OTHER DOCUMENTS WHICH ARE REFERENCED BY OR LINKED TO THIS PUBLICATION OR THE UNICODE WEBSITE.

D. Waiver of Damages. In no event shall Unicode or its licensors be liable for any special, incidental, indirect or consequential damages of any kind, or any damages whatsoever, whether or not Unicode was advised of the possibility of the damage, including, without limitation, those resulting from the following: loss of use, data or profits, in connection with the use, modification or distribution of this information or its derivatives.

E. Trademarks & Logos.
1. The Unicode Word Mark and the Unicode Logo are trademarks of Unicode, Inc. "The Unicode Consortium" and "Unicode, Inc." are trade names of Unicode, Inc. Use of the information and materials found on this website indicates your acknowledgement of Unicode, Inc.’s exclusive worldwide rights in the Unicode Word Mark, the Unicode Logo, and the Unicode trade names.

2. The Unicode Consortium Name and Trademark Usage Policy ("Trademark Policy") are incorporated herein by reference and you agree to abide by the provisions of the Trademark Policy, which may be changed from time to time in the sole discretion of Unicode, Inc.

3. All third party trademarks referenced herein are the property of their respective owners.

Miscellaneous.
1. Jurisdiction and Venue. This server is operated from a location in the State of California, United States of America. Unicode makes no representation that the materials are appropriate for use in other locations. If you access this server from other locations, you are responsible for compliance with local laws. This Agreement, all use of this site and any claims and damages resulting from use of this site are governed solely by the laws of the State of California without regard to any principles which would apply the laws of a different jurisdiction. The user agrees that any disputes regarding this site shall be resolved solely in the courts located in Santa Clara County, California. The user agrees said courts have personal jurisdiction and agree to waive any right to transfer the dispute to any other forum.

2. Modification by Unicode. Unicode shall have the right to modify this Agreement at any time by posting it to this site. The user may not assign any part of this Agreement without Unicode’s prior written consent.

3. Taxes. The user agrees to pay any taxes arising from access to this
We protect your rights with a two-step method: (1) we copyright the library, and (2) we offer you this license, which gives you legal permission to copy, distribute and/or modify the library.

To protect each distributor, we want to make it very clear that there is no warranty for the free library. Also, if the library is modified by someone else and passed on, the recipients should know that what they have is not the original version, so that the original author’s reputation will not be affected by problems that might be introduced by others.

Finally, software patents pose a constant threat to the existence of any free program. We wish to make sure that a company cannot effectively restrict the users of a free program by obtaining a restrictive license from a patent holder. Therefore, we insist that any patent license obtained for a version of the library must be consistent with the full freedom of use specified in this license.

Most GNU software, including some libraries, is covered by the ordinary GNU General Public License. This license, the GNU Lesser General Public License, applies to certain designated libraries, and is quite different from the ordinary General Public License. We use this license for certain libraries in order to permit linking those libraries into non-free programs.

When a program is linked with a library, whether statically or using a shared library, the combination of the two is legally speaking a combined work, a derivative of the original library. The ordinary General Public License therefore permits such linking only if the entire combination fits its criteria of freedom. The Lesser General Public License permits more lax criteria for linking other code with the library.

We call this license the "Lesser" General Public License because it does Less to protect the user's freedom than the ordinary General Public License. It also provides other free software developers Less of an advantage over competing non-free programs. These disadvantages are the reason we use the ordinary General Public License for many libraries. However, the Lesser license provides advantages in certain special circumstances.

For example, on rare occasions, there may be a special need to encourage the widest possible use of a certain library, so that it becomes a de-facto standard. To achieve this, non-free programs must be allowed to use the library. A more frequent case is that a free library does the same job as widely used non-free libraries. In this case, there is little to gain by limiting the free library to free software only, so we use the Lesser General Public License.

In other cases, permission to use a particular library in non-free programs enables a greater number of people to use a large body of free software. For example, permission to use the GNU C Library in non-free programs enables many more people to use the whole GNU operating system, as well as its variant, the GNU/Linux operating system.

Although the Lesser General Public License is Less protective of the users’ freedom, it does ensure that the user of a program that is linked with the Library has the freedom and the wherewithal to run that program using a modified version of the Library.

The precise terms and conditions for copying, distribution and modification follow. Pay close attention to the difference between a "work based on the library" and a "work that uses the library". The former contains code derived...
from the library, whereas the latter must be combined with the library in order to run.

TERMS AND CONDITIONS FOR COPYING, DISTRIBUTION AND MODIFICATION

0. This License Agreement applies to any software library or other program which contains a notice placed by the copyright holder or other authorized party saying it may be distributed under the terms of this Lesser General Public License (also called "this License"). Each licensee is addressed as "you".

A "library" means a collection of software functions and/or data prepared so as to be conveniently linked with application programs (which use some of those functions and data) to form executables.

The "Library", below, refers to any such software library or work which has been distributed under these terms. A "work based on the Library" means either the Library or any derivative work under copyright law: that is to say, a work containing the Library or a portion of it, either verbatim or with modifications and/or translated straightforwardly into another language. (Hereinafter, translation is included without limitation in the term "modification").

"Source code" for a work means the preferred form of the work for making modifications to it. For a library, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the library.

Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of running a program using the Library is not restricted, and output from such a program is covered only if its contents constitute a work based on the Library (independent of the use of the Library in a tool for writing it). Whether that is true depends on what the Library does and what the program that uses the Library does.

1. You may copy and distribute verbatim copies of the Library’s complete source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice and disclaimer of warranty; keep intact all the notices that refer to this License and to the absence of any warranty; and distribute a copy of this License along with the Library.

You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2. You may modify your copy or copies of the Library or any portion of it, thus forming a work based on the Library, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:

   a) The modified work must itself be a software library.
   b) You must cause the files modified to carry prominent notices stating that you changed the files and the date of any change.
   c) You must cause the whole of the work to be licensed at no charge to all third parties under the terms of this License.
   d) If a facility in the modified Library refers to a function or a table of data to be supplied by an application program that uses the facility, other than as an argument passed when the facility is invoked, then you must make a good faith effort to ensure that, in the event an application does not supply such function or table, the facility still operates, and performs whatever part of its purpose remains meaningful.
   (For example, a function in a library to compute square roots has a purpose that is entirely well-defined independent of the application. Therefore, Subsection 2d requires that any application-supplied function or table used by this function must be optional: if the application does not supply it, the square root function must still compute square roots.)

   These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Library, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Library, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

   Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative works based on the Library.

   In addition, mere aggregation of another work not based on the Library with the Library (or with a work based on the Library) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3. You may opt to apply the terms of the ordinary GNU General Public License instead of this License to a given copy of the Library. To do this, you must alter all the notices that refer to this License, so that they refer to the ordinary GNU General Public License, version 2, instead of to this License. (If a newer version than version 2 of the ordinary GNU General Public License has appeared, then you can specify that version instead if you wish.) Do not make any other change in these notices.

   Once this change is made in a given copy, it is irreversible for that copy, so the ordinary GNU General Public License applies to all subsequent copies and derivative works made from that copy.

   This option is useful when you wish to copy part of the code of the Library into a program that is not a library.

4. You may copy and distribute the Library (or a portion or derivative of it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange.

   If distribution of object code is made by offering access to copy from a designated place, then offering equivalent access to copy the source code from the same place satisfies the requirement to distribute the source code, even though third parties are not compelled to copy the source along with the object code.

5. A program that contains no derivative of any portion of the Library, but is designed to work with the Library by being compiled or linked with it, is called a "work that uses the Library". Such a work, in isolation, is not a
derivative work of the Library, and therefore falls outside the scope of this License.

However, linking a “work that uses the Library” with the Library creates an executable that is a derivative of the Library (because it contains portions of the Library), rather than a “work that uses the library”. The executable is therefore covered by this License. Section 6 states terms for distribution of such executables.

When a “work that uses the Library” uses material from a header file that is part of the Library, the object code for the work may be a derivative work of the Library even though the source code is not. Whether this is true is especially significant if the work can be linked without the Library, or if the work is itself a library. The threshold for this to be true is not precisely defined by law.

If such an object file uses only numerical parameters, data structure layouts and accessors, and small macros and small inline functions (ten lines or less in length), then the use of the object file is unrestricted, regardless of whether it is legally a derivative work. (Executables containing this object code plus portions of the Library will still fall under Section 6.)

Otherwise, if the work is a derivative of the Library, you may distribute the object code for the work under the terms of Section 6. Any executables containing that work also fall under Section 6, whether or not they are linked directly with the Library itself.

6. As an exception to the Sections above, you may also combine or link a “work that uses the Library” with the Library to produce a work containing portions of the Library, and distribute that work under terms of your choice, provided that the terms permit modification of the work for the customer’s own use and reverse engineering for debugging such modifications.

You must give prominent notice with each copy of the work that the Library is used in it and that the Library and its use are covered by this License. You must supply a copy of this License. If the work during execution displays copyright notices, you must include the copyright notice for the Library among them, as well as a reference directing the user to the copy of this License. Also, you must do one of these things:

a) Accompany the work with the complete corresponding machine-readable source code for the Library including whatever changes were used in the work (which must be distributed under Sections 1 and 2 above); and, if the work is an executable linked with the Library, with the complete machine-readable “work that uses the Library”, as object code and/or source code, so that the user can modify the Library and then relink to produce a modified executable containing the modified Library. (It is understood that the user who changes the contents of definitions files in the Library will not necessarily be able to recompile the application to use the modified definitions.)

b) Use a suitable shared library mechanism for linking with the Library. A suitable mechanism is one that (1) uses at run time a copy of the library already present on the user’s computer system, rather than copying library functions into the executable, and (2) will operate properly with a modified version of the library, if the user installs one, as long as the modified version is interface-compatible with the version that the work was made with.

c) Accompany the work with a written offer, valid for at least three years, to give the same user the materials specified in Subsection 6a, above, for a charge no more than the cost of performing this distribution.

d) If distribution of the work is made by offering access to copy from a designated place, offer equivalent access to copy the above specified materials from the same place.

e) Verify that the user has already received a copy of these materials or that you have already sent this user a copy.

For an executable, the required form of the “work that uses the Library” must include any data and utility programs needed for reproducing the executable from it. However, as a special exception, the materials to be distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.

It may happen that this requirement contradicts the license restrictions of other proprietary libraries that do not normally accompany the operating system. Such a contradiction means you cannot use both them and the Library together in an executable that you distribute.

7. You may place library facilities that are a work based on the Library side-by-side in a single library together with other library facilities not covered by this License, and distribute such a combined library, provided that the separate distribution of the work based on the Library and of the other library facilities is otherwise permitted, and provided that you do these two things:

a) Accompany the combined library with a copy of the same work based on the Library, uncombined with any other library facilities. This must be distributed under the terms of the Sections above.

b) Give prominent notice with the combined library of the fact that part of it is a work based on the Library, and explaining where to find the accompanying uncombined form of the same work.

8. You may not copy, modify, sublicense, link with, or distribute the Library except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, link with, or distribute the Library is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

9. You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Library or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or distributing the Library (or any work based on the Library), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Library or works based on it.

10. Each time you redistribute the Library (or any work based on the Library), the recipient automatically receives a license from the original licensor to copy, distribute, link with or modify the Library subject to these terms and conditions. You may not impose any further restrictions on the recipients’ exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties with this License.

11. If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot distribute so as to satisfy simultaneously your obligations under this License and any other pertinent
Copyright 1995-2013 Jean-loup Gailly and Mark Adler

The zlib compression library is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

The zlib compression library is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with the zlib compression library; if not, write to the Free Software Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston, MA 02110-1234 USA

zlib.h -- interface of the ‘zlib’ general purpose compression library version 1.2.8, April 28th, 2013

Copyright (C) 1995-2013 Jean-loup Gailly and Mark Adler

This software is provided `as-is`, without any express or implied warranty. In no event will the authors be held liable for any damages arising from the use of this software.

Permission is granted to anyone to use this software for any purpose, including commercial applications, and to alter it and redistribute it freely, subject to the following restrictions:

1. The origin of this software must not be misrepresented; you must not claim that you wrote the original software. If you use this software
in a product, an acknowledgment in the product documentation would be
appreciated but is not required.

2. Altered source versions must be plainly marked as such, and must not be
misrepresented as being the original software.

3. This notice may not be removed or altered from any source distribution.

Jean-loup Gailly        Mark Adler
jloup@gzip.org          madler@alumni.caltech.edu