White Paper

Kofax Implementation of TIFF with JPEG Compression

June 2003

Table of Contents

Introduction 3
Default TIFF-JPEG Storage Filter 3
Default Writing of a TIFF-JPEG File 3
Default Reading of a TIFF-JPEG File 4
Default TIFF Header 4
Optional TIFF-JPEG Storage Filter 5
Writing of a TIFF-JPEG File with the Optional Storage Filter 5
Reading of a TIFF-JPEG File with the Optional Storage Filter 5
Optional TIFF Header 5
Switching Between the two Storage Filters 7
Changing the ini File 7
Guidelines 7

Introduction

Since 1992 Kofax has supported the use of a TIFF file with JPEG compression (TIFF-JPEG) for the storage of multi-bit images (color or grayscale). Kofax implemented the TIFF-JPEG storage filter in accordance with a simplified and widely accepted interpretation of the TIFF 6.0 specification for handling JPEG compression within a TIFF file. As a standard, TIFF-JPEG was non-specific and confusing, leaving developers open to interpreting the specification.

Following the TIFF 6.0 specification, there was a Draft TIFF Technical Note #2. This draft attempted to clarify the support for TIFF-JPEG; however, the draft was never brought to be accepted as a standard. Additionally, the Technical Note #2 implementation was incompatible with the existing format outlined in the TIFF 6.0 specification. Developers of storage and display technologies that wish to leverage the benefits of a TIFF file with JPEG compression now have to make an implementation decision based upon the format and interpretation they believe will be most utilized by their customer base. Unfortunately the varied formats may not work together.

Recently with advancements in scanning technologies, imaging software, and algorithms we are seeing more customers turn to the use of multi-bit images within their document management/imaging workflows. Whereas the TIFF standard for binary images is well documented and accepted by imaging developers (i.e. binary images written as TIFF files have an extremely high probability of being displayed and read by a wide variety of engines), TIFF-JPEG images written by one vendor according to their interpretation of the TIFF-JPEG specification may not be read by another vendor because of a differing interpretation/implementation of the specification. The problem this creates is one in which images scanned with one vendors implementation of TIFF-JPEG may not be able to be displayed and leveraged by another vendors repository or management system.

With the release of the 3.75 Adrenaline engines, Kofax has added a second storage filter that supports JPEG compression within a TIFF file. This second storage filter is not enabled by default but can be easily switched to be the default by following the simple instructions outlined in this document. The second storage TIFF-JPEG storage filter does not support the Draft TIFF Technical Note #2, but does support a TIFF 6.0 format that has been implemented by other vendors in the imaging market.

Kofax chose to add the second storage filter to support customers who discovered incompatibilities with the default storage filter in their imaging pilots.

Default TIFF-JPEG Storage Filter

This is the storage filter that Kofax has utilized and supported since 1992. Due to testing with the majority of content management and document management systems, we have selected this storage filter as the default simply because it was the most compatible with the widest variety of systems.

Default Writing of a TIFF-JPEG File

Kofax will write the following JPEG specific tag:

- JPEGInterchangeFormat = An offset value pointing to the start of the JPEG data

Other tags with a relationship to the JPEG data are:

- Compression = 6 (JPEG)
- PhotometricInterpretation = 6 (YCbCr)
- StripOffsets = The start of the complete JPEG data (header and compressed data) - where the SOI (Start Of Image) (FFD8) marker can be found - the beginning of that which is a complete JPEG image outside the TIFF file.

- StripByteCounts = The length of the complete JPEG image. Typically everything from the SOI to EOI (End Of Image) (FFD9).

Default Reading of a TIFF-JPEG File

Both Kofax support TIFF-JPEG formats can be read by the default TIFF-JPEG storage filter.

Default TIFF Header

The following is the TIFF header created with the default storage filter:

```
C:\imgctls\BIN>dumptiff scn0001

TIFF HEADER for scn0001.TIF:
HEADER:
000000  Byte Order            4949      'II'
000002  Version               002A'
000004  1st IFD Offset        00000008
IFD #0:
000008  Entry Count           0014
00000A  NewSubfileType        00FE 4 000001 00000000, Long   = 000000
000016  ImageWidth            0100 4 000001 00000280, Long   = 000640
000022  ImageLength           0101 4 000001 000001E0, Long   = 000480
00002E  BitsPerSample         0102 3 000003 000002FA, Short  @ 0002FA
00003A  Compression           0103 3 000001 00000006, Short  = 000006 JPEG
000046  PhotometricInterpret  0106 3 000001 00000006, Short  = 000006 YCbCr
000052  ImageDescription      010E 2 00002E 000002CC, Ascii  @ 0002CC
00005E  StripOffsets          0111 4 000001 00000300, Long   = 000768
00006A  Orientation           0112 3 000001 00000001, Short  = 000001 UprLeft
000076  SamplesPerPixel       0115 3 000001 00000003, Short  = 000003
000082  RowsPerStrip          0116 4 000001 000001E0, Long   = 000480
00008E  StripByteCounts       0117 4 000001 0000CE1F, Long   = 052767
00009A  XResolution           011A 5 000001 000002C4, Ration @ 0002C4
0000A6  YResolution           011B 5 000001 000002BC, Ration @ 0002BC
0000B2  PlanarConfiguration   011C 3 000001 00000001, Short  = 000001 Chunky
0000BE  ResolutionUnit        0128 3 000001 00000002, Short  = 000002 inch
0000CA  Software              0131 2 000038 00000284, Ascii  @ 000284
0000D6  DateTime              0132 2 000014 00000270, Ascii  @ 000270
0000E2  JPEGInterChangeForma  0201 4 000001 00000300, Long   = 000768
0000EE  ???                   80E8 3 000001 00000000, Short  = 000000
```

```
0000FA  Next IFD Offset       00000000
000270  *DateTime             "2003:05:01 11:03:08"
000284  *Software             "Kofax standard Multi-Page TIFF Storage Filter
v3
                              .03.000"
0002CC  *ImageDescription     "C:\imgctls\BIN\SCN0001.TIF adding description"
0002C4  *XResolution          100/1, 100.0000 dots/inch
0002BC  *YResolution          100/1, 100.0000 dots/inch
```

Optional TIFF-JPEG Storage Filter

This is an additional storage filter Kofax has created to assist customers who utilize retrieval and display technologies that do not support the default format.

Writing of a TIFF-JPEG File with the Optional Storage Filter

We create the following JPEG specific tags:

- YCbCrSubSampling = Value derived from JPEG header.
- ReferenceBlackWhite = 0/1, 255/1, 128/1, 255/1, 128/1, 255/1.
- JPEGProc = 1 - Indicates the JPEG was created with the Baseline Sequential Process.
- JPEGInterchangeFormat = An offset value pointing to the start of the JPEG data.
- JPEGInterchangeFormatLength = Length of the data from SOI to SOS markers.
- JPEGRestartInterval = Value derived from JPEG header, 0 if not specified in JPEG.
- JPEGQTables = Location of the Quantization tables.
- JPEGDCTables = Inside the JPEG data. The location of the Huffman DC tables.
- JPEGACTables = Same as JPEGDCTables except dealing with the Huffman AC tables rather than the DC.

Other tags with a relationship to the JPEG data are:

- Compression = 6 (JPEG).
- PhotometricInterpretation = 6 (YCbCr)
- StripOffsets = Start of the compressed JPEG data – Location of the SOS marker.
- StripByteCounts = Length of the compressed JPEG image – Length of the SOS to the EOI.

Reading of a TIFF-JPEG File with the Optional Storage Filter

Both the Optional and Default storage filters are capable of reading the other format.

Optional TIFF Header

The following is the header data created by the Optional TIFF-JPEG storage filter:

```
C:\imgctls\BIN>dumptiff scn0001
```

```
TIFF HEADER for scn0001.TIF:
HEADER:
000000  Byte Order            4949      'II'
000002  Version               002A'
000004  1st IFD Offset        00000008
IFD #0:
000008  Entry Count           001C
00000A  NewSubfileType        00FE 4 000001 00000000, Long   = 000000
000016  ImageWidth            0100 4 000001 00000280, Long   = 000640
000022  ImageLength           0101 4 000001 000001E0, Long   = 000480
00002E  BitsPerSample         0102 3 000003 000002FA, Short  @ 0002FA
00003A  Compression           0103 3 000001 00000006, Short  = 000006 JPEG
000046  PhotometricInterpret  0106 3 000001 00000006, Short  = 000006 YCbCr
000052  ImageDescription      010E 2 00002E 000002CC, Ascii  @ 0002CC
00005E  StripOffsets          0111 4 000001 00000580, Long   = 001408
00006A  Orientation           0112 3 000001 00000001, Short  = 000001 UprLeft
000076  SamplesPerPixel       0115 3 000001 00000003, Short  = 000003
000082  RowsPerStrip          0116 4 000001 000001E0, Long   = 000480
00008E  StripByteCounts       0117 4 000001 0000CB9F, Long   = 052127
00009A  XResolution           011A 5 000001 000002C4, Ration @ 0002C4
0000A6  YResolution           011B 5 000001 000002BC, Ration @ 0002BC
0000B2  PlanarConfiguration   011C 3 000001 00000001, Short  = 000001 Chunky
0000BE  ResolutionUnit        0128 3 000001 00000002, Short  = 000002 inch
0000CA  Software              0131 2 000038 00000284, Ascii  @ 000284
0000D6  DateTime              0132 2 000014 00000270, Ascii  @ 000270
0000E2  JPEGProc              0200 3 000001 00000001, Short  = 000001
0000EE  JPEGInterChangeForma  0201 4 000001 00000300, Long   = 000768
0000FA  JPEGInterChangeForma  0202 4 000001 00000280, Long   = 000640
000106  JPEGRestartInterval   0203 3 000001 00000000, Short  = 000000
000112  JPEGQtables           0207 4 000003 00000264, Long   @ 000264
00011E  JPEGDctables          0208 4 000003 00000258, Long   @ 000258
00012A  JPEGActables          0209 4 000003 0000024C, Long   @ 00024C
000136  YCbCrSubSampling      0212 3 000002 00010002, Short  = 2,1
000142  ReferenceBlockWhite   0214 5 000006 0000021C, Ration @ 00021C
00014E  ???                   80E8 3 000001 00000000, Short  = 000000
00015A  Next IFD Offset       00000000
000270  *DateTime             "2003:05:01 11:08:44"
```

```
000284  *Software              "Kofax standard Multi-Page TIFF Storage Filter
v3
                               .03.000"
0002CC  *ImageDescription      "C:\imgctls\BIN\SCN0001.TIF adding description"
0002C4  *XResolution           100/1, 100.0000 dots/inch
0002BC  *YResolution           100/1, 100.0000 dots/inch
```

Switching Between the two Storage Filters

With the 3.75 release of Adrenaline engines, Kofax provides users with the ability to switch from the Default TIFF-JPEG storage filter to the Optional storage filter. The Optional storage filter should only be used in circumstances where images generated by the default storage filter are not compatible with customer retrieval or viewing technologies.

In some cases, neither storage filter will generate images that are compatible.

To switch the storage filter from the default to the optional functionality, add the following entry to the [Defaults] section of the Kofax200.ini for the new format to be enabled.

TIFF_JPEG_Style=0

To switch back to the default TIFF-JPEG storage filter, add the following entry to the [Defaults] section of the Kofax200.ini for the new format to be enabled.

TIFF_JPEG_Style=1

Changing the ini File

When making a change to the ini file it is important that the ImageControls based application be shut down and not running. Once the change is made the new storage filter will be enabled upon reserving the scanner source. Simply un-reserving the source will not cause the ini file to reload.

Guidelines

The following guidelines could be helpful in determining your requirements:

1. If you are currently using TIFF-JPEG images created by Kofax software with no problems, we do not recommend changing to the Optional storage filter.
2. If you are currently experiencing difficulties with the default storage filter, try testing with the Optional storage filter. If that filter is more compatible with your imaging system, use the Optional way of writing the TIFF-JPEG images. If neither format works for you, see recommendation #4.
3. Prior to changing storage filters, Kofax strongly recommends customers perform tests with the images to ensure they are saved in a format compatible with their system.
4. In certain cases, neither format will be compatible with a particular repository, retrieval engine, or viewer. In those instances, Kofax recommends storing each image in a single file (no multi-page files). Use the TIFF format for binary images and JPEG format for grayscale and color images.