Exiv2 - Bug #1276

fails to read any XMP metadata when duplicates present

08 Feb 2017 14:40 - Aerilius .

Status:	Closed		**Start date:**	08 Feb 2017
Priority:	Normal		**Due date:**	
Assignee:	Robin Mills		**% Done:**	100%
Category:	xmp		**Estimated time:**	2.00 hours
Target version:	0.26			

Description

When reading XMP metadata from image files with duplicate tags, exiv2 (and gexiv2) does not give any metadata output, although metadata is present.

As a consequence, features in applications using exiv2 are rendered disfunctioning if they rely on reading a specific XMP tag (even if that single tag is consistent and not duplicate). In real-world scenarios, some users do have image files with duplicate tags and they experience expected features not working although they could work.

Observation:

```
$ exiv2 -px example_duplicate_tags.jpg
Error: XMP Toolkit error 203: Duplicate property or field node
Warning: Failed to decode XMP metadata.
example_duplicate_tags.jpg: (No XMP data found in the file)
```

By contrast, `exiftool example_duplicate_tags.jpg` prints all metadata and ignores duplicates.

Expectation:

exiv2 should give access to all readable tags. If some XMP tags are inconsistent, it should only be indicated by a warning but not a denial of operation. In case of duplicates the question is open whether to return the last occuring value or the first.

Ubuntu 16.10
exiv2 0.25-3
libexiv2-14 0.25-3

History

#1 - 08 Feb 2017 16:00 - Robin Mills

- Category set to xmp

- Status changed from New to Closed

- Assignee set to Robin Mills

- Target version set to 0.26

- % Done changed from 0 to 100

- Estimated time set to 2.00 h

Thank you for reporting this and providing a test file. I have recreated your issue with Exiv2 v0.26rc1 (current trunk).

The issue is caused by duplicated keys in the XMP and this causes the embedded XMPsdk in exiv2 to panic and throw an exception. This may be valid XML, however I believe it is invalid XMP and the exception is correctly thrown. However, Adobe seem to allow this in the latest XMPsdk (XMP-Toolkit-SDK-CC201607) which we hope to support in exiv2 v0.27.

```
$ exiv2 -pX ~/Downloads/example_duplicate_tags.jpg | xmllint --format - | grep GPano: | wc
    20      20    1368
$ exiv2 -pX ~/Downloads/example_duplicate_tags.jpg | xmllint --format - | grep GPano: | sort | uniq | wc
    10      10     684
$ exiv2 -pX ~/Downloads/example_duplicate_tags.jpg | xmllint --format - | grep GPano: | sort | uniq
     <GPano:CroppedAreaImageHeightPixels>3127</GPano:CroppedAreaImageHeightPixels>
     <GPano:CroppedAreaImageWidthPixels>6288</GPano:CroppedAreaImageWidthPixels>
     <GPano:CroppedAreaLeftPixels>0</GPano:CroppedAreaLeftPixels>
     <GPano:CroppedAreaTopPixels>8</GPano:CroppedAreaTopPixels>
     <GPano:FullPanoHeightPixels>3144</GPano:FullPanoHeightPixels>
```

```
        <GPano:FullPanoWidthPixels>6288</GPano:FullPanoWidthPixels>
        <GPano:ProjectionType>equirectangular</GPano:ProjectionType>
        <GPano:SourcePhotosCount>42</GPano:SourcePhotosCount>
        <GPano:StitchingSoftware>Hugin</GPano:StitchingSoftware>
        <GPano:UsePanoramaViewer>True</GPano:UsePanoramaViewer>
$
```

Examining your test file with XMP-Toolkit-SDK-CC201607:

```
$ ~/gnu/xmpsdk/XMP-Toolkit-SDK-CC201607/samples/target/macintosh/intel_64/Release/ReadingXMP ~/Downloads/examp
le_duplicate_tags.jpg

/Users/rmills/Downloads/example_duplicate_tags.jpg is opened successfully
CreatorTool = GIMP 2.8.18
dc:title in English =
dc:title in French =

XMP dumped to XMPDump.txt
$ cat XMPDump.txt
Dumping XMPMeta object ""   (0x0)

   GPano:  http://ns.google.com/photos/1.0/panorama/   (0x80000000 : schema)
      GPano:CroppedAreaImageHeightPixels = "3127"
      GPano:CroppedAreaImageWidthPixels = "6288"
      GPano:CroppedAreaLeftPixels = "0"
      GPano:CroppedAreaTopPixels = "8"
      GPano:FullPanoHeightPixels = "3144"
      GPano:FullPanoWidthPixels = "6288"
      GPano:ProjectionType = "equirectangular"
      GPano:SourcePhotosCount = "42"
      GPano:StitchingSoftware = "Hugin"
      GPano:UsePanoramaViewer = "True"

   xmp:  http://ns.adobe.com/xap/1.0/   (0x80000000 : schema)
      xmp:CreateDate = "2017:01:06 15:58:33"
      xmp:CreatorTool = "GIMP 2.8.18"
      xmp:ModifyDate = "2017-02-06T16:20:57"

   exif:  http://ns.adobe.com/exif/1.0/   (0x80000000 : schema)
      exif:DateTimeOriginal = "2017:01:06 15:58:33"
      exif:GPSLongitude = "8,2.7893E"
      exif:GPSTimeStamp = "2017-01-06T14:58:07Z"
      exif:GPSAltitude = "550/1"
      exif:GPSAltitudeRef = "0"
      exif:ColorSpace = "1"
      exif:ExifVersion = "0230"
      exif:FlashpixVersion = "0100"
      exif:ComponentsConfiguration   (0x600 : isOrdered isArray)
         [1] = "1"
         [2] = "2"
         [3] = "3"
         [4] = "0"
      exif:UserComment   (0x1E00 : isLangAlt isAlt isOrdered isArray)
         [1] = "-
Projection: Equirectangular (2)
FOV: 360 x 179
Ev: 11,73"   (0x50 : hasLang hasQual)
               ? xml:lang = "x-default"   (0x20 : isQual)
      exif:GPSVersionID = "2.2.0.0"
      exif:GPSLatitude = "49,25.5675N"
      exif:GPSProcessingMethod = "GPS"

   tiff:  http://ns.adobe.com/tiff/1.0/   (0x80000000 : schema)
      tiff:XResolution = "1/1"
      tiff:YResolution = "1/1"
      tiff:ResolutionUnit = "0"
      tiff:YCbCrPositioning = "1"
      tiff:Make = "Xiaomi"
      tiff:Model = "Redmi Note 2"

   photoshop:  http://ns.adobe.com/photoshop/1.0/   (0x80000000 : schema)
      photoshop:DateCreated = "2017:01:06 15:58:33"
$
```

The workaround is to extract the XMP, fix it, replace it in the file. Something like this:

```
$ exiv2 -pX ~/Downloads/example_duplicate_tags.jpg | xsltproc magic.xslt | exiv2 -iX- ~/Downloads/example_dupl
icate_tags.jpg
```

The code in magic.xslt is discussed here: http://stackoverflow.com/questions/10912544/removing-duplicate-elements-with-xslt

I am very reluctant to submit a fix for this issue in Exiv2 at this time in our release cycle. I believe this will be resolved when we upgrade the version of XMPsdk (#941) which is likely to be available in 2017.

I have investigated how "duplicated" keys/values are handled in XMP-Toolkit-SDK-CC201607 by modifying one of the CroppedAreaImageHeightPixels values in your file. The first value defined is used and no warning is issued when the second (and different) value is encountered.

#2 - 08 Feb 2017 17:18 - Robin Mills

I've done more work on this. Maybe it's easy to fix (but it isn't). The exception comes from xmpsdk/src/RDFParse.cpp line 396.

```
    // Make sure that this is not a duplicate of a named node.
    if ( ! (isArrayItem | isValueNode) ) {
        if ( FindChildNode ( xmpParent, childName, kXMP_ExistingOnly ) != 0 ) {
            XMP_Throw ( "Duplicate property or field node", kXMPErr_BadXMP );
        }
    }
```

I've changed this to:

```
    if ( ! (isArrayItem | isValueNode) ) {
        if ( FindChildNode ( xmpParent, childName, kXMP_ExistingOnly ) != 0 ) {
            // #1276, don't throw
            // XMP_Throw ( "Duplicate property or field node", kXMPErr_BadXMP );
            return FindChildNode ( xmpParent, childName, kXMP_ExistingOnly )
        }
    }
```

And no more exception. *BUT* the second value of CroppedAreaImageHeightPixels is being appended to the first:

```
  Xmp.GPano.CroppedAreaImageHeightPixels      XmpText    8  31283127
```

I tried to fix that and ran into new exceptions. I know very little about the XMPsdk code and I'm very reluctant to change it. This is dangerous country. As Captain Kirk would say "We're surrounded by aliens. Beam me up, Scotty". I feel I must defer this and when we upgrade XMPsdk in v0.27, this issue will be resolved.

#3 - 08 Feb 2017 17:59 - Aerilius .

Thanks for your effort and commitment!

Files

example_duplicate_tags.jpg.zip	21 KB	08 Feb 2017	Aerilius .