

Exiv2 - Feature #1024

Provide regular expression support for the exiv2 -g feature

10 Jan 2015 21:13 - Robin Mills

Status:	Closed	Start date:	10 Jan 2015
Priority:	Normal	Due date:	
Assignee:	Robin Mills	% Done:	100%
Category:	samples	Estimated time:	4.00 hours
Target version:	0.26		
Description			
Thomas has done some work on this as a side task in #917 . He attached the file. Here are my comments on his suggestion:			
Good work. Thanks very much. I didn't know about <regex.h> - I think this must be quite a new addition to standard C++. A wonderful new addition.			
I will integrate this to the trunk and add a build flag HAVE_REGEX. Exiv2(.exe) will degrade to earlier behaviour (string match) when <regex.h> is not available. I will update the man page src/exiv2.1.			
According to MSDN docs, <regex.h> is supported by Visual Studio 2010 and later. I'll investigate if earlier MSVC editions can use GCC's /usr/include/regex.h which is on the build machines for Cygwin and MinGW builds. However MSVC will have to be able to build OK when <regex.h> is not available as this will be "standard" on older windows platforms.			
I should add something to the test suite. I'll update exiv2 -v -V (--verbose --version) to output the value of HAVE_REGEX. Our current use of exiv2 -v -V in our test suite is limited to the new webready code, and test/build-test.py. However exiv2 -v -V provides us with the potential for the test suite to morph appropriately.			
<pre>509 rmills@rmillssmbp:~/gnu/exiv2/trunk \$ exiv2 -pa -g "L..s" -g Make ~/Pictures/Christmas/ElTeide/ DSC_4056.jpg Exif.Image.Make Ascii 18 NIKON CORPORATION Exif.Photo.MakerNote Undefined 14328 (Binary value suppressed) Exif.MakerNote.Offset Long 1 876 Exif.MakerNote.ByteOrder Ascii 3 II Exif.Nikon3.LensType Byte 1 D G VR Exif.Nikon3.Lens Rational 4 18-250mm F3.5-6.3 Exif.Nikon3.LensFStops Undefined 4 4.16667 Exif.NikonLd3.LensIDNumber Byte 1 Sigma 18-250mm F3.5-6.3 DC OS Macro HS M Exif.NikonLd3.LensFStops Byte 1 F4.2 510 rmills@rmillssmbp:~/gnu/exiv2/trunk \$</pre>			
Thomas has pointed out that it may be desirable to clean up regular expressions which are out of scope. I don't feel we need to bother with this for the exiv2(.exe) sample application.			
I did not bother cleaning up the compiled regexps on shutdown, assuming that they will be removed on program termination anyway. But if there is a special mode of the program (e.g. a lib mode?) where the keys_ member goes out of scope, the cleanup function regfree should be called for each of its members before that happens.			
Thanks very much to Thomas for this contribution.			
Related issues:			
Related to Exiv2 - Feature #917: Modify exiv2/actions.cpp return -3/253 when ...		Closed	11 Aug 2013
Related to Exiv2 - Feature #1053: Add option -K Key (--key Key) to specify on...		Closed	08 Apr 2015

Associated revisions

Revision 3562 - 11 Jan 2015 17:24 - Robin Mills

#1024. Thank you, Thomas for this very useful contribution.

Revision 3563 - 11 Jan 2015 17:25 - Robin Mills

#1024. Adding file missing from r3562.

Revision 3564 - 11 Jan 2015 21:38 - Robin Mills

#1024. Add CMake support for regex. Fixed unused variables in version.cpp

Revision 3686 - 13 Apr 2015 12:33 - Robin Mills

#1024 and #1053. Changed option REG_EXTENDED and REG_BASIC to support ^ as a begin marker.

Revision 4075 - 08 Dec 2015 09:27 - Robin Mills

#1024. Support for C++11 #include <regex>. --grep keys may have an optional trailer /i to indicate to ignore case.

History

#1 - 11 Jan 2015 17:48 - Robin Mills

- Status changed from Assigned to Resolved

Submitted [r3562](#) and [r3563](#). Thank you, Thomas for this contribution.

Currently MinGW detects <regex.h> and fails to link libregex.a. I don't know why not, so I've forced MinGW to undefine EXV_HAVE_REGEX. There are warnings about unused variables in src/version.cpp on all platforms. And I haven't added anything special to the test suite for the -g feature. For the moment, this is "good enough". I will cleanup the warnings before we ship v0.25. Everything builds and passes the test suite on the build server.

#2 - 11 Jan 2015 22:21 - Robin Mills

[r3564](#). CMake support. Fixed unused variables in src/version.cpp.

I installed Visual Studio 2012 and VS 2013 Windows Desktop Express on the build server. Very pleased to say that the msvc2005 environment builds exiv2 with both 2012 and 2013 and the test suite passes.

Neither 2012 nor 2013 include <regex.h>. I tried copying the Cygwin <regex.h> to <exiv2dir>/include/exiv2/regex.h. It doesn't compile because regex.h requires other header files which are not available with Visual Studio. I'm not going to do any further work on this. -g degrades to a string match for all MSVC builds.

#3 - 17 Jan 2015 09:53 - T Modes

Robin Mills wrote:

Neither 2012 nor 2013 include <regex.h>.

Yes, but both have <regex> from C++11. Maybe this could be an alternative.

#4 - 17 Jan 2015 10:27 - Robin Mills

Thanks T. I expected it to be there. Perhaps I have to set something in the msvc project files to say "use C++11" if available.

We're in the end-game of getting v0.25 ready for release and I'm reluctant to start messing about with msvc. I'm happy to stick to "-g match" degrades to match on MSVC and MinGW builds.

If you know how to fix this, I'll be happy to accept a patch.

Robin

#5 - 17 Jan 2015 11:27 - Robin Mills

I've looked at this on Visual Studio 2013 Windows Desktop Express Edition. Indeed #include <regex> works and it will be possible to use <regex> instead of "regex.h".

Adopting #include <regex> is not pain free because the standard <regex> is not source compatible with "regex.h". The standard uses std::regex instead of ::regex_t. And, from inspection, I believe there will be other code changes. For example, the macro REGEX_NOTBOL in "regex.h" becomes enum match_flag_type { ... match_not_bol ... }

For sure, all of this can be changed. However, we should really sense three regex support conditions: <regex>, "regex.h" and 'no regex' (which Safari keeps changing to reggae).

I'm reluctant to divert energy into this for v0.25.

#6 - 13 Apr 2015 09:45 - Robin Mills

- Status changed from Resolved to Assigned

I've asked Thomas to do more work on this to detect the three build conditions for the autotools and CMake build environments. I've asked Thomas B to fix msvc2005 when Thomas S is complete.

- 0) No regex support (-g = simple match)
- 1) <regex.h> # GCC
- 2) <regex> # C++ 11

Additionally, I discovered that ^ and \$ are not effective in -g options (see [#1053](#) and forum discussion). Thomas has correctly identified this as using REGEX_NOTBOT and REGEX_NOTEOL. I've submitted a fix for this [r3684](#) and [r3685](#). Thomas - you are welcome to change my fix if you wish.

I've tried expressions such as -g /tag/i and these are ineffective. We should define the syntax of the -g option in the man page src/exiv2.1

Thomas S, can you also add to the test suite to ensure -g performs as expected. You can determine the level of regex support (0,1 or 2) with the following command which uses regex!

```
536 rmills@rmillsmbp:~ $ exiv2 -vVg regex
exiv2 0.24 001800 (64 bit build)
have_regex=1
537 rmills@rmillsmbp:~ $
```

Thanks very much to Thomas S and Thomas B for working on this.

#7 - 19 Apr 2015 22:11 - Robin Mills

- Target version changed from 0.25 to 0.26

I'm going to defer this for v0.26. Thanks to Thomas S, -g/--grep is supported on GCC. C++ 11 compilers will have to wait for v0.26 when Thomas has time available to work on this.

#8 - 08 Dec 2015 09:27 - Robin Mills

- Status changed from Assigned to Resolved
- Assignee changed from Thomas Schmidt to Robin Mills
- % Done changed from 0 to 100
- Estimated time set to 4.00 h

Fix submitted: [r4075](#)

The --grep feature will use include <regex> when built with c++11. On other platforms, <regex.h> will be used if available, else string match. The pattern can be "embellished" with the optional trailer /i = ignore case.

```
673 rmills@rmillsmbp:~/gnu/exiv2/trunk $ exiv2 -pa --grep gps1/i test/data/exiv2-bug1122.exv
Exif.GPSInfo.GPSLatitudeRef      Ascii      2  North
Exif.GPSInfo.GPSLatitude         Rational   3  52deg 3.81700'
Exif.GPSInfo.GPSLongitudeRef     Ascii      2  East
Exif.GPSInfo.GPSLongitude        Rational   3  1deg 13.81940'
674 rmills@rmillsmbp:~/gnu/exiv2/trunk $
```

You can detect the configuration with exiv2 -vV (verbose version).

```
667 rmills@rmillsmbp:~/gnu/exiv2/trunk $ exiv2 -vV --grep cplusplus --grep regex
exiv2 0.25 001900 (64 bit build)
cplusplus=199711 <----- __cplusplus preprocessor value
cplusplus11=0 <---- not C++11
have_regex=1 <---- has <regex.h>
668 rmills@rmillsmbp:~/gnu/exiv2/trunk $
```

#9 - 22 Dec 2015 04:47 - Robin Mills

- Status changed from Resolved to Closed