

How can I tell if a file is an image of some type?

 devblogs.microsoft.com/oldnewthing/20150430-00

April 30, 2015



Raymond Chen

A customer wanted to know if there is some standard way of determining from the full path to a file whether it is a photo or image. They are writing a program that transfers a bunch of files to a back-end server, and they want to treat photos differently from other files.

You can use [the `AssocGetPerceivedType` function](#) to classify a file into one of a variety of categories. You can [consult the document to see the full list](#), but it's things like "image", "audio", "video", "document". This information is obtained by studying the file extension and looking up [the registered perceived type](#).

Let's take the function out for a spin:

```
#define UNICODE
#define _UNICODE
#include <windows.h>
#include <shlwapi.h>
#include <stdio.h> // horrors! mixing C and C++!

int __cdecl wmain(int argc, wchar_t **argv)
{
    PERCEIVED perceived;
    PERCEIVEDFLAG flag;
    PWSTR pszType;
    if (argc > 1 && SUCCEEDED(AssocGetPerceivedType(
        PathFindExtensionW(argv[1]),
        &perceived, &flag, &pszType)))
    {
        wprintf(L"Type is %ls\n", pszType);
        if (perceived == PERCEIVED_TYPE_IMAGE) {
            wprintf(L"Hey, that's an image!\n");
        }
        CoTaskMemFree(pszType);
    }
    return 0;
}
```

Run this program and give a file name, or just an extension (with the dot) as the command line parameter. It will tell you the perceived type and include a special message if the type is an image.

But let's look at the customer's question again. It's not clear whether they are trying to identify files by file format or by classification. For example, suppose the file in question is a TIFF image. The `AssocGetPerceivedType` function will report this as an image, because, well, it's an image. But that may not be a file format that their back-end server supports.

- If they wanted to know whether the file is a PNG, GIF, or JPG because those are the image formats supported by their back-end server, then they need to check for those specific extensions (and possibly even sniff file contents if they are paranoid).
- If they care only that the file represents some sort of image (possibly in a format their program does not know how to parse), because they want to, say, upload all images into a Pictures folder regardless of the image format, then they should use the Perceived Type.

The customer thanked us for the pointer to the `AssocGetPerceivedType` function and especially for the clarifying remarks. "It turns out that the feature specification was not clear on the definition of 'image file', so that's something we will need to resolve ourselves. But the information you provided will definitely solve our problem, once we figure out what our problem is!"

Raymond Chen

Follow

