
Subject: Re: Basic character set analyzer
Posted by [mirek](#) on Mon, 09 Feb 2009 07:42:10 GMT
[View Forum Message](#) <> [Reply to Message](#)

cbpporter wrote on Sun, 08 February 2009 06:45luzr wrote on Sat, 07 February 2009 15:53
I mean, for missing char in Arial, I would check some sans-serif font first rather than trying e.g. Times New Roman..

I've given this issues a lot of thought, but I don't think there is a good way to do this.

Well, first possibility is fixed table of substitutions (the number of widely used fonts is quite small IMO).

Second is PANOSE:

<http://en.wikipedia.org/wiki/PANOSE>

[http://msdn.microsoft.com/en-us/library/ms534014\(VS.85\).aspx](http://msdn.microsoft.com/en-us/library/ms534014(VS.85).aspx)

(-> Win32 provides PANOSE number)

- not sure whether FreeType supports PANOSE... Also, another question is whether fonts actually support it

Third: Perform matching based on visual appearance. That is a bit though, but still possible.

Can be based on raster data processing or even curve processing....

E.g. examining "l" character raster, it should be quite easy to identify serifs, width of central line (and perhaps some other aspects I do not see yet...). By comparing width of 'l' and 'm', you should be able to tell the font is monospace (that is, if this is not stored as attribute somewhere in the font).

Mirek
