PDF

PDF Days Europe 2021

OpenType Color Fonts in PDF

Implementing emoji and other colored symbols in PDF

Thomas Merz, PDFlib GmbH, www.pdflib.com

About

- I started PDFlib GmbH when Deutsche Mark still existed
- We produce PDF-related software and market it on all continents
- Developer components for creating and processing PDF
- Co-founder of the PDF/A Competence Center, now PDF Association
- Several years board member and chair of the Technical Working Group
- www.pdflib.com

Topics

- Four OpenType color font flavors
- Emoji and Unicode
- Color fonts in PDF

Why Color Fonts?

- All stand-alone font formats so far are monochromatic
- Glyphs can be colorized when used, but not inside the font
- »Color is the new Black«
  - wide-spread use of emoji on mobile devices
  - color fonts are mainly driven by Web and mobile applications
  - artistic use of color in fonts
- Simplified exchange of emoji due to Unicode standardization (emoji are a big step back in the long history of writing…)

Commercial Font Samples
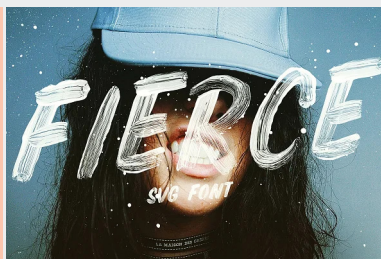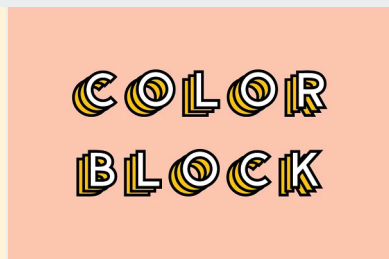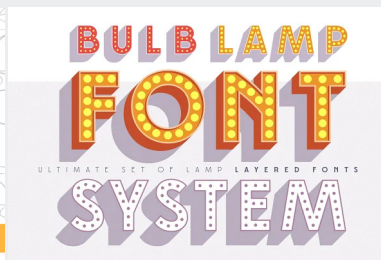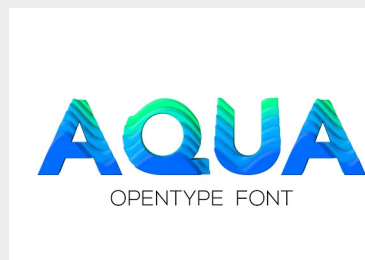
Color Font Standards

- Four vendors invented competing color font formats
- All four formats ended up in ISO 14496-22:2019
  - Information technology — Coding of audio-visual objects — Part 22: Open Font Format
- Existence of a standard doesn't mean that everybody supports it

General Properties

- Color fonts need much more space than monochromatic fonts
- 3 of 4 color font formats don't support hinting for improved rendering at small sizes
- Color fonts don't work on all systems and applications
- Color fonts may contain a combination of mono and color glyphs
- Mono versions of color glyphs are usually present for compatibility with old environments

- Transparency may be used in glyphs
  - for typographic effects
  - to simplify combination of glyph layers
    (e.g. transparent glasses for all skin tones)

Adobe/Mozilla OpenType SVG

- Glyph descriptions use the power of SVG
  - vector graphics
  - transparency
  - smooth shadings (gradients)
  - patterns and masks
- SVG may also include PNG and JPEG raster images

Adobe/Mozilla OpenType SVG

- Some SVG features are restricted for security and performance reasons
  - SVG Integration (www.w3.org/TR/svg-integration)
  - disables scripting, external references, interactivity
  - support not required in applications: CSS, animation, filter, mask, pattern
- Minor SVG features added
  - e.g. for selecting color from a palette

Adobe/Mozilla OpenType SVG

- Reference fonts:
  - Adobe EmojiOneColor: 1836 glyphs, 4MB
  - Adobe TrajanColor-Concept: 1366 glyphs, 8MB
- OpenType SVG support in operating systems
  - Windows 10 Anniversary Update (2016), macOS 10.14 (2018)

Microsoft OpenType COLR

- Layered Glyphs:
  - each color glyph is built from multiple mono layers
  - each layer is assigned color and transparency
- Hinting is possible (unlike all other formats)
- Relevant font tables:
  - »COLR«: list of mono glyphs used to construct color glyphs
  - »CPAL«: one or more RGBA color palettes, e.g. for light or dark background
- One of the CPAL palettes can be selected when using the font
  - different palettes for dark and light background
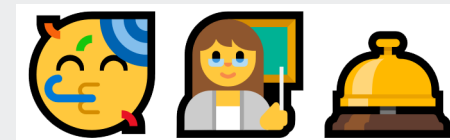 - palette selection not yet supported in applications
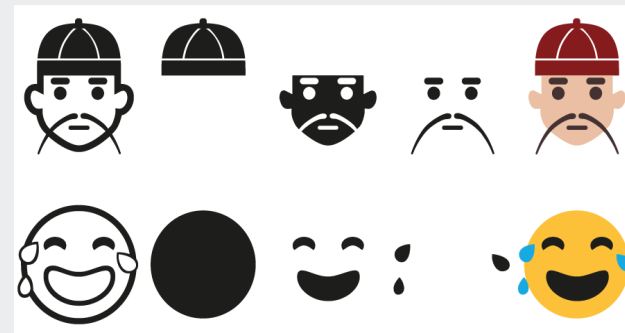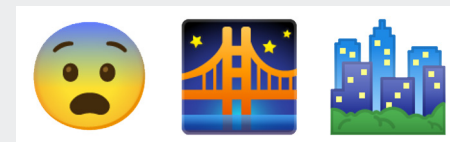
Microsoft OpenType COLR

- COLR table version 0
  - glyph parts are stacked and can use solid fills and transparency
- COLR table version 1
  - will be introduced in OpenType 1.9
  - gradient fills, affine transformations, masks, blending modes
  - color information may be variable, e.g. number of color stops in a gradient
  - requires much more complex implementation than version 0

Microsoft OpenType COLR

- Reference font Microsoft Segoe UI Emoji: 3665 glyphs, 2MB
- OpenType COLR support in operating systems
  - Windows 8.1 (2013), macOS 10.13 (2017)

Apple OpenType SBIX

- SBIX = Standard Bitmap Graphics table
- Glyphs in common image formats PNG, JPEG, TIFF
- Scaling to large glyph sizes is problematic
- Mono glyphs for applications without color support are optional

Apple OpenType SBIX

- Reference font Apple Color Emoji: 2600 glyphs, 156MB
- OpenType SBIX support in operating systems
  - Windows 10 Anniversary Update (2016), macOS 10.7 (2011)
  - but Apple Color Emoji doesn't work in Windows 10

Google OpenType CBDT

- CBDT = Color Bitmap Data Table
- Raster image data: PNG with alpha or uncompressed 32-bit RGBA
- Scaling to large glyph sizes is problematic
- No »glyf« table for mono glyph descriptions required
  - no fallback for applications which cannot handle color glyphs
- Reference font NotoColorEmoji: 3600 glyphs, 9MB
- OpenType CBDT support in operating systems
  - Windows 10 Anniversary Update (2016), Android
 - but NotoColorEmoji doesn't work in Windows 10 for lack of »glyf« table

Emoji in Unicode

- Emoji are standardized in Unicode
  - First emoji in Unicode 6.0 (2010)
  - Unicode 14.0 (2021-09) adds 112 emoji
  - Unicode 14.0 contains 3633 emoji
- Unicode Technical Standard #51 specifies details of emoji processing
  - modifiers, sort order
  - emoji presentation vs. text presentation (when to use color or mono)
  - zero-width joiner (ZWJ) sequences

Emoji Sequences

ZWJ grouping
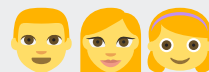
skin tone
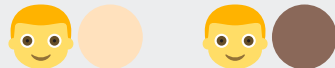
emoji flag

ZWJ gender

skin tone and gender

ZWJ hair components

ZWJ role: man farmer

Emoji Sequences

- Additional features for emoji with zero-width joiner U+200D (ZWJ):
  - the ZWJ is invisible and combines characters
  - sequence and combinations vary some aspect of a character
  - gender, hair components, role (occupation), flags, keycaps
- Amazing list at unicode.org/emoji/charts/emoji-zwj-sequences.html
- Fonts support emoji sequences via OpenType feature tables
  - usually the »ccmp« feature (glyph composition/decomposition)
  - applications should should enable »ccmp« feature by default

A longer Role Sequence

- How to create a female dark-skinned scientist:

  start with U+1F469 WOMAN

  add U+1F3FF EMOJI MODIFIER FITZPATRICK TYPE-6

  add U+200D ZERO WIDTH JOINER

  add U+1F52C MICROSCOPE
  - right – she doesn't use a microscope, but that's the font designer's liberty
  - create farmers, teachers, painters, cooks, astronauts and many other roles
  - see unicode.org/emoji/charts/emoji-zwj-sequences.html

Color Fonts in PDF

- PDF 2.0 doesn't support any of the four color font formats
- Alternative approach: rasterize color glyphs – many disadvantages
- Type 3 fonts:
  - very flexible
  - arbitrary contents: raster, vector, shadings, glyphs from other fonts
  - transparency
  - can use intrinsic glyph color or inherit color from outside
  - all glyph descriptions can share common resources like, e.g. color spaces, shadings
  - ToUnicode CMap defines Unicode mapping of Type 3 glyphs for text extraction
  - glyphs created by emoji sequences must be encoded by a sequence

Disadvantages of Type 3 Fonts

- Type 3 fonts exist only in PDF documents; cannot be exchanged
- Not supported in operating systems
  - PDF viewer cannot use »local fonts«
- No hinting
- 8-bit encoding, i.e. max. 255 glyphs per font instance
  - font objects may contain an arbitrary number of glyphs
  - glyphs cannot be addressed by Unicode value
  - Large fonts can be segmented in subfonts with max. 255 glyphs
- Glyph names required; inefficient for large fonts

Type 3 Font Subsetting

- Subsetting is not defined for Type 3 fonts
  - are all glyphs embedded or only a subset?
  - compare with the original OpenType color font
  - use random subset prefix in the font name?
- Skipping unused color glyphs is important to keep file size down
  - typically only a few emoji of a large set are used
  - Microsoft Segoe UI Emoji with 1 glyph: 5.5 KB
  - ...with 10 glyphs: 10 KB
  - ...with full set of 3665 glyphs: 2.6 MB
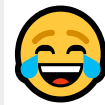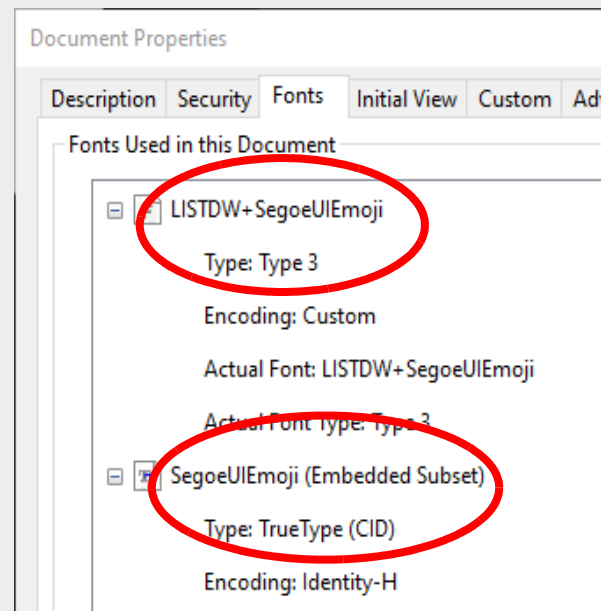
Converting COLR Fonts to PDF

- Type 3 fonts derived from COLR fonts:
  - Type 3 font references glyph parts (layers) from the OpenType font
  - OpenType font must be embedded in addition to the Type 3 font
  - hinting is preserved
  - OpenType font contains more glyphs than Type 3 font because of glyph parts
  - Segoe UI Emoji: ca. 12000 full and partial glyphs for ca. 3700 final glyphs

Color Fonts in PDF Standards

- PDF/A-1 and PDF/X-3:
  - transparent glyphs not allowed
  - a single glyph can spoil standard conformance
- PDF/A-1a/2a/2u/3a/3u/4 and PDF/UA-1:
  - proper Unicode mappings must be provided
  - application must handle supplementary Unicode characters > U+FFFF
- Color management:
  - SVG, COLR and CBDT: sRGB required
  - SBIX: sRGB optional

Accessibility and Emoji

- How to present emoji with Assistive Technology (AT)?
- All Unicode characters have descriptive names

  - e.g. U+1F494 BROKEN HEART 💔, U+1F602 FACE WITH TEARS OF JOY 😂
  - names are also useful when searching emoji on the virtual keyboard
- Assistive Technology can use these names
- The »Read Out Loud« feature in Acrobat DC uses Unicode names
  - proper ToUnicode CMap required in PDF

Summary

- PDF doesn't support OpenType color fonts natively
- Nevertheless they can be used due to the versatility of Type 3 fonts

Thanks for your Attention!

- OpenType specification: docs.microsoft.com/en-us/typography/opentype/spec/
- ISO 14496-22:2015: Information technology—Coding of audio-visual objects—Part 22: Open Font Format
- Unicode Resources for emoji: unicode.org/emoji/techindex.html
- Realtime emoji use on Twitter: emojitracker.com/
- »Read Out Loud« emoji test page (distributed with the presentation slides)
- PDFlib 10 (available soon) supports color fonts: www.pdflib.com