Network Working Group
Request for Comments: 1314

A. Katz
D. Cohen
ISI
April 1992

A File Format for the Exchange of Images in the Internet

Status of This Memo

This document specifies an IAB standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "IAB Official Protocol Standards" for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Abstract

This document defines a standard file format for the exchange of fax-like black and white images within the Internet. It is a product of the Network Fax Working Group of the Internet Engineering Task Force (IETF).

The standard is:

** The file format should be TIFF-B with multi-page files supported. Images should be encoded as one TIFF strip per page.

** Images should be compressed using MMR when possible. Images may also be MH or MR compressed or uncompressed. If MH or MR compression is used, scan lines should be "byte-aligned".

** For maximum interoperability, image resolutions should either be 600, 400, or 300 dpi; or else be one of the standard Group 3 fax resolutions (98 or 196 dpi vertically and 204 dpi horizontally).

Note that this specification is self contained and an implementation should be possible without recourse to the TIFF references, and that only the specific TIFF documents cited are relevant to this specification. Updates to the TIFF documents do not change this specification.

Experimentation with this file format specified here is encouraged.

1. Introduction

The purpose of this document is to define a standard file format for exchange of black and white images using the Internet. Since many organizations have already started to accumulate and exchange scanned documents it is important to reach agreement about an interchange file format in order to promote and facilitate the exchange and distribution of such documents. These images may originate from scanners, software, or facsimile (fax) machines. They may be manipulated by software, communicated, shared, duplicated, displayed, printed by laser printers, or faxed.

This file format provides for the uniform transfer of high quality images at a reasonable cost and with reasonable speed whether these files are generated by scanners, totally by software (e.g., text-to-fax, bitmap-to-fax, OCR, etc), or by fax. Also the intent of this document is to remain compatible with future moves to multi-level (i.e., gray-scale), higher resolution, or color images. The format proposed here is supported by both commercially available hardware and commercial and public domain software for most popular platforms in current use.

The file format for images is a totally separate issue from how such files are to be communicated. For example, FTP or SMTP could be used to move an image file from one host to another, although there are complications in the use of SMTP as currently implemented due to file size and the need to move binary data. (There is currently a proposal for removing these limitations from SMTP and in particular extending it to allow binary data. See reference [1].)

One major potential application of the communications format defined here is to allow images to be sent to fax machines using the Internet. It is intended that one or more separate companion documents will be formulated to address the issues of standardization in the areas of protocols for transmitting images through the Internet and the issues of addressing fax machines and routing faxes. Just as the exchange format is separate from the transmission mechanism, it is also separate from how hosts store images.

This document specifies a common exchange format; it does not require a host to store images in the format specified here, only to convert between the host's local image storage formats and the exchange format defined here for the purpose of exchanging images with other hosts across the network.

This standard specifies the use of TIFF (Tagged Image File Format, see below) as a format for exchange of image files. This is not a specific image encoding, but a framework for many encoding

techniques, that can be used within the TIFF framework. For example, within TIFF it is possible to use MMR (the data encoding of CCITT Group 4 fax, see below), MH or MR (the data encodings of CCITT Group 3 fax), or other encoding methods.

Which encoding technique to use is not specified here. Instead, with time the encoding schemes used by most document providers will emerge as the de-facto standard. Therefore, we do not declare any as "the standard data encoding scheme," just as we do not declare that English is the standard publication language. (However, we expect that most document providers will use MMR in the immediate future because it offers much better compression ratios than MH or MR.)

Similarly, TIFF does not require that an image be communicated at a specific resolution. Resolution is a parameter in the TIFF descriptive header. We do suggest that images now be sent using one of a set of common resolutions in the interests of interoperability, but the format accommodates other resolutions that may be required by specialized applications or changing technologies.

Occasionally, image files will have to be converted, such as in the case where a document that was scanned at 400 dpi is to be printed on a 300 dpi printer. This conversion could be performed by the document provider, by the consumer, or by a third party. This document specifies neither who performs the conversion, nor which algorithms should be used to accomplish it.

Note that this standard does not attempt to define an exchange format for all image types that may be transmitted in the Internet. Nothing in this standard precludes it from being used for other image type such as gray-scale (e.g., JPEG) or color images but, for the purposes of standardization, the scope of this document is restricted to monochromatic bitmapped images.

The developers of this standard recognize that it may have a limited lifespan as Office Document Architecture (ODA) matures and comes into use in the Internet; ultimately the class of images covered by this standard will likely be subsumed by the more general class of images supported by the ODA standards. However, at present, there does not appear to be a sufficient installed base of ODA compliant software and the ODA standards are not fully mature. This standard is intended to fill the need for a common image transfer format until ODA is ready. Finally, we believe that it should be possible to automatically map images encoded in the format specified here into a future ODA-based image interchange format, thus providing a reasonable transition path to these future standards.

2. Relationship to Fax

Transmission of facsimile (fax) images over phone lines is becoming increasingly widespread. The standard of most fax machines in the U.S. is CCITT Group 3 (G3), specified in Recommendations T.4 and T.30 [2] and in EIA Standards EIA-465 and EIA-466. G3 faxes are 204 dots per inch (dpi) horizontally and 98 dpi (196 dpi optionally, in fine-detail mode) vertically. Since G3 neither assumes error free transmission nor retransmits when errors occur, the encoding scheme used is differential only over small segments never exceeding 2 lines at standard resolution or 4 lines for fine-detail. (The incremental G3 encoding scheme is called two-dimensional and the number of lines so encoded is specified by a parameter called k.)

CCITT Group 4 fax (G4) is defined by the T.400 and T.500 series of Recommendations as well as Recommendation T.6 [2]. It provides for 400 dpi (both vertical and horizontal) and is a fully two-dimensional encoding scheme (k is infinite) called MMR (Modified Modified READ, where READ stands for: Relative Element Address Designate). G4 assumes an error free transmission medium (generally an X.25 Public Data Network, or PDN). Because of this, G4 is not in widespread use in the U.S. today.

The traditional fax bundles together four independent issues:

(1) Data presentation and compression;
(2) Data transmission;
(3) Image input from paper ("scanning"); and
(4) Image output to paper ("printing").

This bundling supports, for example, the high quality CCITT Group 4 (G4) images (400x400 dpi) but only over X.25 public data networks with error correction, and similarly it supports the mid-quality CCITT Group 3 (204x98 and 204x196 dpi) but only over phone voice circuits (the Switched Telephone Network, or STN) without error correction. This bundling does not support the use of any other data transmission capabilities (e.g., FTP over LANs and WANs), nor asynchrony between the scanning and the printing, nor image storage, nor the use of the popular laser printers for output (even though they are perfectly capable of doing so).

In conventional fax, images are never stored. In today’s computer network environment, a better model is:

(1) Images are scanned into files or created by software;
(2) These image files are stored, manipulated, or communicated;
(3) Images in a file are printed or displayed.

The only feature of the CCITT fax that should be used is the encoding technique (preferably MMR, but with MR or MH allowed) which may be implemented with a variety of fax-oriented chips at low cost due to the popularity of fax.

"Sending a fax" means both encoding (and decoding) the fax images as well as transmitting the data. Since the Internet ALREADY provides several mechanisms for data transmission (in particular, FTP for general file transmission), it is unnecessary to use the data transmission methods specified in the CCITT standard. Within the Internet, each fax image should be stored in a file and these files could be transferred (e.g., using FTP, SMTP, RPC-based methods, etc.).

Fax machines should be considered just as scanners and printers are, as I/O devices between paper and files; but not as a transmission means. Higher quality Group 4 images are thus supported at low cost, while enjoying the freedom to use any computerized file transfer and duplication mechanism, standard laser printers, multiple printing (possibly at multiple remote sites) of the same image without having to rescan it physically, and a variety of software for various processing of these images, such as OCR and various drawing programs. We should be able to interoperate with files created by fax machines, scanners, or software and to be able to print all of them on fax machines or on laser printers.

The CCITT Recommendations assume realtime communications between fax machines and do not therefore specify any kind of fax file format. We propose using TIFF [3] which seems to be emerging as a standard, for encapsulation of encoded images. Because they assume realtime communications, the CCITT fax protocols require negotiations to take place between the sender and receiver. For example, they negotiate whether to use two-dimensional coding (and with what k parameter) and what (if any) padding there is between scan lines.

In our approach, the image in the file is already compressed in a particular manner. If it is to be sent to an ordinary fax machine using a fax board/modem, that board will perform the negotiations with the receiving fax machine. In the cases where the receiver cannot handle the type of compression used in the file, it will be necessary to convert the image to another compression scheme before transmission. (Most fax cards seem to either store images using the default values of the parameters which are negotiated or in a format which can quickly be converted to this. With currently available hardware and software, any necessary format conversion should be easy to accomplish.)

In conventional fax, if the compression used for a particular image

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.