



# HP Smart Document Scan Software compression schemes and file sizes

## Table of contents

Introduction	2
Compression schemes supported in HP Smart Document Scan Software	2
Compression Scheme Types	2
JPEG	2
PNG	3
LZW	3
TIFF Group 3 and 4 (CCITT Group 3 and 4)	3
JBIG2 lossy and lossless	3
MRC	3
Scanned Image File Size	3
File size for 24bit color (KB)	4
File size for 8 bit grayscale (KB)	4
File size for 1bit halftone (KB)	5
File size for 1bit black-and-white (KB)	5
Summary	6

# Introduction

This document provides a description of each of the various compression schemes utilized by the most common file formats supported within the HP Smart Document Scan Software included with HP Scanjet scanners.

## Compression schemes supported in HP Smart Document Scan Software

HP Smart Document Scan Software makes it easy to convert stacks of hard-copy documents and save them as electronic files in a variety of popular file formats including PDF, JPEG, TIFF, Microsoft® Word, and more. The software also provides a wide selection of compression schemes to meet various customer requirements for balancing between file size, image quality, and file access compatibility.

The table below shows the compression schemes utilized by some file formats that HP Smart Document Scan Software provides. Use the table below to quickly compare how image quality and file size are related to the different compression schemes. Color modes detected by the auto color detection feature also fall into the same respective color category.

Color mode	File type												
	PDF						JPEG			TIFF			PNG
	Normal			High Compression (MRC)			Compression	Image quality	File size	Compression	Image quality	File size	PNG
Color or gray	JPEG	Good ↑ Best	Small ↑ Large	MRC	Good ↑ Best	Small ↑ Large	JPEG	Good ↑ Best	Small ↑ Large	JPEG LZW Uncompressed	Good ↑ Best	Small ↑ Large	
Black-and-white or Halftone	JBIG2 Lossy JBIG2 Lossless TIFF G4	Good ↑ Best	Small ↑ Large	JBIG2 Lossy JBIG2 Lossless TIFF G4	Good ↑ Best	Small ↑ Large	-	-	-	TIFF G4 TIFF G3 LZW Uncompressed	Best	Small ↑ Large	

The lossy compression compresses data by discarding unimportant parts and is mostly used to compress multimedia data, such as still images. HP Smart Document Scan Software uses lossy compression in most compression schemes. The software also provides adjustment capability to each lossy compression to help customers balance file size and image quality.

## Compression scheme types

Compression schemes help reduce the resources being used to store or transfer scanned documents or images more efficiently. Below is more detail on each of the supported compression scheme types.

### JPEG

JPEG is the most popular image compression method. It is suitable for all kinds of images, and is compatible with almost all viewing application tools. As a lossy compression scheme, the compression ratio can be achieved in a very wide range (from 1/4 to 1/200), however the image quality decreases while increasing the compression ratio for a smaller file size. JPEG supports color and gray image compressions.

## PNG

PNG is a file format that uses a lossless compression scheme suitable for all kinds of color and gray images, and is compatible with some popular file readers. With PNG, image quality does not degrade but a high compression ratio is hard to achieve. PNG supports color and gray image compressions.

## LZW

LZW is a lossless compression scheme that is suitable for all kinds of images, and compatible with some popular viewing application tools. Image quality does not degrade but a high compression ratio is hard to achieve (e.g. around 1/2 of the original). LZW supports color, gray, and black-and-white image compressions.

## TIFF Group 3 and 4 (CCITT Group 3 and 4)

TIFF Group 3 and 4 are lossless compression schemes compatible with most TIFF readers, however some older TIFF readers may not be able to read Group 4 files. Image quality does not degrade. Group 4 is a newer standard than Group 3 and normally achieves a higher compression ratio than Group 3, but may not be compatible with some older version TIFF readers. TIFF Group 3 and 4 supports black-and-white image compression only.

## JBIG2 lossy and lossless

JBIG2 lossy and lossless is a newer compression standard than TIFF Group 3 and 4, and also achieves higher compression ratio than TIFF Group 3 and 4. This compression scheme is compatible with some popular file readers. With the lossless mode, the image quality does not degrade, however using the lossy mode provides the most efficient balanced results between image quality and file size when compressing binary document images constructed with fewer types of fonts. JBIG2 lossy and lossless supports black-and-white image compression only.

## MRC

MRC is a compression model that intelligently segments image objects of a scanned image into layers and compresses each layer separately using all available compression schemes for effective results. It may take slightly longer than other compression schemes for the processing but the difference is negligible with recent PC processing power. This compression model is suitable for all kinds of images but is only supported by PDF file format and only compatible with a few popular PDF readers. MRC PDF provides efficient compression results with balanced image quality. MRC supports color, gray, and black-and-white image compressions.

# Scanned image file size

HP Smart Document Scan Software also provides a selection of compression levels (Best/Better/Good) to help customers balance file size, image quality, and file format compatibility for supported compression schemes. For instance, when a lossy compression scheme is selected a smaller file size normally results in a lower image quality. In editing situations, Best and Better levels are more suitable, while the Good level is more suitable for transferring files.

PDF file format supports various compression schemes, below is a table showing how PDF file size varies with different compression settings. These numbers are only for reference, since an actual file size of a scanned image can be impacted by a large number of factors including, document contents, scan condition settings (scan color mode, scan resolution, tone curve setting, image processing features on/off), file format, and compression schemes and levels. Figure A, letter-size ISO24712 standard chart, shown below, was used as the scan target for the results shown in the tables on the following pages.

Figure A



**File size for 24bit color (KB)**

Resolution		Quality					
		Normal Compression			High Compression (MRC)		
		Good	Medium	Best	Good	Medium	Best
75 ppi		60	79	133	21	25	69
100 ppi		94	126	210	20	25	97
150 ppi		171	226	387	22	32	184
200 ppi		249	326	566	23	33	199
240 ppi		314	409	745	18	25	85
300 ppi		453	607	1150	21	32	122
400 ppi	Fast mode	620	821	1567	29	45	246
	Quality mode	666	899	1747	30	48	247
500 ppi	Fast mode	874	1183	2255	39	71	340
	Quality mode	930	1275	2455	40	72	346
600 ppi	Fast mode	1220	1666	3109	67	133	470
	Quality mode	1333	1864	3515	56	124	473

**File size for 8 bit grayscale (KB)**

Resolution		Quality					
		Normal Compression			High Compression (MRC)		
		Good	Medium	Best	Good	Medium	Best
75 ppi		54	74	100	16	17	24
100 ppi		86	117	156	12	14	25
150 ppi		157	209	290	12	15	42
200 ppi		228	303	427	13	16	53
240 ppi		290	393	553	11	13	18
300 ppi		427	595	857	12	14	23
400 ppi	Fast mode	592	811	1130	14	18	73
	Quality mode	646	885	1282	15	19	68
500 ppi	Fast mode	833	1164	1643	18	24	106
	Quality mode	897	1245	1825	18	25	98
600 ppi	Fast mode	1171	1642	2304	32	41	146
	Quality mode	1292	1775	2681	24	34	128

**File size for 1bit halftone (KB)**

Resolution		Quality					
		Normal Compression			High Compression (MRC)		
		Good	Medium	Best	Good	Medium	Best
75 ppi					32		
100 ppi					46		
150 ppi					76		
200 ppi					131		
240 ppi					182		
300 ppi					220		
400 ppi	Fast mode				357		
	Quality mode				355		
500 ppi	Fast mode				530		
	Quality mode				524		
600 ppi	Fast mode				741		
	Quality mode				741		

**File size for 1bit black-and-white (KB)**

Resolution		Quality					
		Normal Compression			High Compression (MRC)		
		Good	Medium	Best	Good	Medium	Best
75 ppi		9		11		9	
100 ppi		10		15		10	
150 ppi		10		23		10	
200 ppi		12		34		12	
240 ppi		13		43		13	
300 ppi		18		58		18	
400 ppi	Fast mode	19		74		19	
	Quality mode	23		79		23	
500 ppi	Fast mode	24		93		24	
	Quality mode	32		106		32	
600 ppi	Fast mode	28		100		28	
	Quality mode	63		151		63	

## Summary

Many different drivers and software applications are included with HP Scanjet scanners. This provides added value for HP customers and helps ensure that the right tools are available to meet a variety of scanning needs.

---

© Copyright 2012 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Microsoft is a U.S. registered trademarks of Microsoft Corporation.

4AA4-2053ENW, May 2012

