

Exporting SAS/GRAPH to Other Packages What Format Should I Choose? (Version 6 of the SAS System)

Peter Ruzsa presenting, SAS Institute Inc.

I. Introduction

(Paraphrased from Graphics File Formats by David Kay and John R. Levine).

Graphics from SAS/GRAPH can be stored in basically two different representations: the Vector format family and the Bitmap format family, (also called raster or pixel maps). These families have distinct advantages and disadvantages depending on what you wish to do. Before we get started, a few definitions are in order.

A. Definitions

1. Vector files

Vector representation means describing a picture as a series of discrete shapes or lines, such as a circle, a dashed line, square or polygon, and text. Some of these objects will also have descriptions of how these shapes are filled. Strictly speaking, the word "vector" actually refers to a line, but the popular meaning includes any shape in a graph. In this paper, page description languages such as Encapsulated PostScript are also included in this family. File types included in this family are CGM (Computer Graphics Metafile), HPGL (Hewlett Packard Graphics Language), EPS (Encapsulated PostScript), and WMF, (Windows Metafile). EPS, CGM, and WMF have bitmap representations as well, but are generally not used for exporting. Some of these formats can also be combined, which will be described later. These types of formats are best used for graphical charts such as bar charts or plots.

2. Bitmap files

Bitmap representation means that an picture has basically been broken into a grid of very small elements called pixels. The lightness and color value of each element or piece is recorded individually. Typically, the data point position in the grid or field determine what the pixel represents. In other words, the data points map to the image, which is where the name bitmap comes from. Filetypes of this family include GIF (Graphics Interchange Format), TIFF (Tagged Image File Format), PCX, (PC Paintbrush), and several others. This type of format is best used for photographic pictures and other complex graphical images.

B. Comparative Advantages and Disadvantages of Bitmap vs. Vector Formats

1. Advantages of Bitmaps

A. Ease of Manufacture

Bitmaps can record just about any conceivable graph such as a picture because any picture can be broken into a grid. They can be colorful and they can be animated in some cases such as in an Animated GIF file. They can also be easy to obtain via any screen capture program or digital camera. In SAS/GRAPH, you can simply select FILE=>Export on most platforms to create a bitmap file.

2. Disadvantages of Bitmaps

A. Image Size

A high resolution color image file can be many megabytes in size depending on the color depth that it was stored in.

B. Depth Complexity

The ability to display these images varies from computer to computer depending on the video display and the color depth set on the video display driver itself or the imaging software.

C. Lack of flexibility**1. Editing capability**

The pixels in a bitmap do not have any relationship to one another. So, if you want to manipulate any portion of the image, the editing program has to do something complex, such as finding all of the pixels that match the same color. Even this operation can often be inadequate for some editing.

2. Fixed Resolution

When an image is created, it is created with a certain number of pixels across and down. If you attempt to enlarge an image, the pixels will become large enough to see their rectangular shape, an effect known as *aliasing* or *staircasing*. Some graphics programs can compensate for this problem by interpolating between the pixels (known as *antialiasing*). Likewise, when you shrink an image, it often loses resolution because the program has to decide what pixels to throw away to make it fit. If you decide to resize it again after shrinking it, the image can be fuzzy and aliased, and shaded areas may appear patterned. The size of an image can also change depending on the device it is displayed or printed on.

1. Advantages of Vector Files**a. Efficiency of Storage**

Rather than having to store an object as a collection of disconnected pixels, a vector file can store an object as a command. For example, a line could be described as two end points, or text be described as an actual text string.

b. Ease of Manipulation

Because each object is discrete, they can be changed easily by selecting them. Moving them on the page is also much easier.

c. Variable Resolution and Sizing

Because each object is a description rather than a collection of pixels, resolution is generally not a problem. The graph can be resized easily and displayed or printed on just about any device without the loss of resolution.

2. Disadvantages of Vector Files

Vector files have more limitations in what they can store, such as charts, plots or maps. If you attempted to store a picture in this manner, you would be storing thousands of very small polygons.

II. What Kind Of Formats Can SAS/GRAPH Generate?

SAS/GRAPH can generate output in several types of vector files on all platforms, and several bitmap formats on most platforms (the exceptions are the Mainframe and VAX Release 6.09e releases, which can only generate GIF files).

The following vector formats can be generated from SAS/GRAPH for export.

The first four languages are most commonly used and will be discussed here.

CGM	Computer Graphics Metafile
PS and EPS	PostScript and Encapsulated PostScript
WMF	Windows Metafile – (available on Windows Platforms only)
HP-GL	Hewlett Packard Graphics Language
PCL	Page Control Language

The following bitmap formats can be generated from SAS/GRAPH for export:

BMP	Microsoft Windows Device Independent Bitmap v 1.0 and 2.0
DIB	Device Independent Bitmap
EPSI	Encapsulated PostScript Image Interchange
GIF	Graphics Interchange Format – both 87, 89 and animated versions
MET	OS/2 Metafile (OS/2 only)
PBM	Portable Pixmap
TIFF	Tagged Image File Format (monochrome and color)
WMF	Windows Metafile Image (Windows only)
XBM XPM	X11 Bitmap and Pixmap (Unix only)

III. An Overview of the Vector File Formats

A. Computer Graphics Metafiles (CGM)

The Computer Graphics Metafile is intended to be a vendor neutral and hardware neutral interchange format between packages. CGM is a very rich format providing support for most hardware shapes including lines, circles, arcs, polygons, and text. It has three encoding variants: Binary, which is used in 99 % of all applications and is generally unreadable; Character, which contains a compressed

form of the language in character; and CLEAR TEXT, which generates a verbose version.

Multiple pictures can be placed in a single metafile. The beginning of the Computer Graphics Metafile generally contains a BEGIN METAFILE header, followed by a BEGIN PICTURE for each picture contained in it. The end of each picture will have an END PICTURE at the end of the METAFILE.

B. Windows Metafiles

The Windows Metafile was invented by Microsoft® Corporation as a way of transporting graphics between Windows applications. A Windows Metafile is a list of Microsoft Windows graphical function calls. Only a subset of the total calls is allowed, but it is a large subset containing most of the calls. A Windows metafile contains a short header followed by some records. Each record corresponds to a Windows graphics device interface call (GDI). Because of its Windows-dependant nature, the SAS system can create this type of file on Windows versions of SAS. Enhanced Windows Metafiles can also be created interactively in the SAS Release 6.12 on Windows Platforms. The following table documents where you can get more information on using Windows Metafiles with your package:

Package	Document
Microsoft Office 97	TS252X
Microsoft Office 95	TS352C
Lotus SmartSuite 97	TS252Y
Corel WordPerfect Suite 8	TS252Z
Any other release of software	TS352C

C. Encapsulated PostScript

Encapsulated PostScript, or EPS is the major form of PostScript used to exchange graphics from one package to another. It cannot have more than one page of information in it. SAS/GRAPH can generate this format in two different ways depending on the platform. You can either use the PSEPSF, PSLEPSFC, or similar SAS/GRAPH based drivers, or on the PC or UNIX platforms you can use the WINPRTC or XPRINTC drivers with Host based drivers provided by the vendor. Encapsulated PostScript can also have several variants as far as the format involved:

1. A vector file with PostScript commands in them (this can be generated on any SAS platform).
2. A bitmap file that usually is in a TIFF style format, (this can be generated on PC, Alpha/VMS, and UNIX platforms).
3. A combination of both. Normally, the bitmap or WMF part is called a header. This requires Ghostview to display the graph and which will be discussed later.

D. Hewlett Packard Graphics Language (HP-GL)

HP-GL is the official command language for Hewlett Packard Pen Plotters. Because of the great acceptance of HP Plotters and Printers in the workplace, it has become a standard that is widely used in CAD and Architectural drawing applications. SAS/GRAPH also has several different drivers that can generate HP-GL of several different varieties and complexities. This graphics language originated in the mid seventies with the HP7400 and HP7500 series of plotters. The language itself is comprised primarily of readable ASCII characters that can be easily decoded. In 1990 HP-GL/2 arrived to support LaserJet and DeskJet printers. One of the main additions to HP-GL/2 is a palette extension to support color DeskJet printers and the Dual Context extension

which allows HP-GL/2 to work with LaserJet printers more smoothly.

III. Bitmap Formats

SAS/GRAPH can also generate bitmap files in several formats. The most popular formats are

GIF (Graphics Interchange Format), BMP (Windows Bitmap), TIFF (Tagged Image File Format), and JPEG (Joint Photographers Export Group). These can be generated interactively and in batch. The GIF file format is very useful for Web applications.

The following bitmap formats can be generated from SAS/GRAPH for export:

BMP	Microsoft Windows Device Independent Bitmap v 1.0 and 2.0
DIB	Device Independent Bitmap
EPSI	Encapsulated PostScript Image Interchange
GIF	Graphics Interchange Format - both 87, 89 and animated versions
JPG	Joint Photographers Export Group
MET	OS/2 Metafile (OS/2 only)
PBM	Portable Pixmap
TIFF	Tagged Image File Format monochrome and color
XBM XPM	X11 Bitmap and Pixmap (Unix only)

IV. Creating Graphics for the Web

Note: an excellent web page for this information is located at:

<http://www.sas.com/rnd/web/publish.html>

The most commonly used format for exporting graphics to the Web is the GIF format. There are 4 different types of GIF drivers: the GIF interactive export driver which is selected by selecting **File=>Export** from the screen; the IMGGIF driver; the GIF driver; and the GIFANIM, which generates animated GIF files.

V. Future Enhancements

Version 7 of the SAS system will have several enhancements, including:

1. A HTML driver that generates a GIF file with HTML. This will be very useful for batch jobs that create HTML pages on the fly.
2. A WEBFRAME driver that creates one or more GIF files and several HTML files that are used to display the GIF images with thumbnail links.
3. The ability to drill down on specific portions of a graph that is created.
4. Object Linking and Embedding between PC SAS and other applications.