Chapter 2 Formatting Your Data Files

Chapter Overview

File Formats

- RGB and HSI Color Models
- Raw Raster Files
- Raw Palette Files
- HDF Files
- NCSA ImageTool Default Files

Specifying the Dimensions of Raw Raster Files

Chapter Overview

This chapter explains how NCSA ImageTool reads and displays data files. The chapter contains a detailed discussion of the file formats NCSA ImageTool can read and mentions other software packages that will enhance the use of this program.

File Formats

To understand the file formats required by NCSA ImageTool, you need to know how visualization programs represent colors. This chapter covers two of those methods. Raster, palette, and HDF files are also discussed.

RGB and HSI Color Models

The most popular method of color representation is the *RGB* (red, green, blue) model where all colors can be represented as a combination of these three colors. For example, a high concentration of red and green with very little blue forms a shade of yellow. A different color model uses a mixture of hue, saturation, and intensity. This model is called the *HSI model*. NCSA ImageTool assumes the RGB model when reading palette files.

Raw Raster Files

A raw raster file is a stream of raw, binary, 8-bit raster data in row-major order. Each 8-bit byte corresponds to a pixel in the image. The image is represented in *row-major order*; that is, the first raster line appears first in the file, succeeded by the next raster line, and so forth. Though raw raster files are easy to create, the raw raster file format is not very flexible and therefore, not highly recommended.

If you choose to use raw raster files to store your image data, you must specify the dimensions of the dataset before the file can be read by NCSA Image. The section entitled "Specifying the Dimensions of Raw Raster Files" discusses the procedures for specifying the dimensions of your dataset.

Raw Palette Files

A raw palette file is a stream of 768 bytes. Raw palette files can store 256 colors, which can be selected from a palette of over 16 million possible colors. Palette files are based on the red, green, and blue representation of color, the RGB color model. The files consist of, in order, 256 bytes of red, 256 bytes of green, and 256 bytes of blue. The 256 color palette entries are calculated by combining the *n*th element (red), the (*n*+256)th element (green), and the (*n*+512)th element (blue) to create the *n*th RGB component.

In other words, a palette file is a lookup table with 256 entries that tell which color to associate with each of the 256 possible pixel values. Each of the 256 palette entries in the palette is chosen from a master palette of 2^{24} RGB colors. Each palette entry consists of three bytes, one each for red, green, and blue; the first red component, the first green component, and the first blue component, for example, comprise the first palette entry.

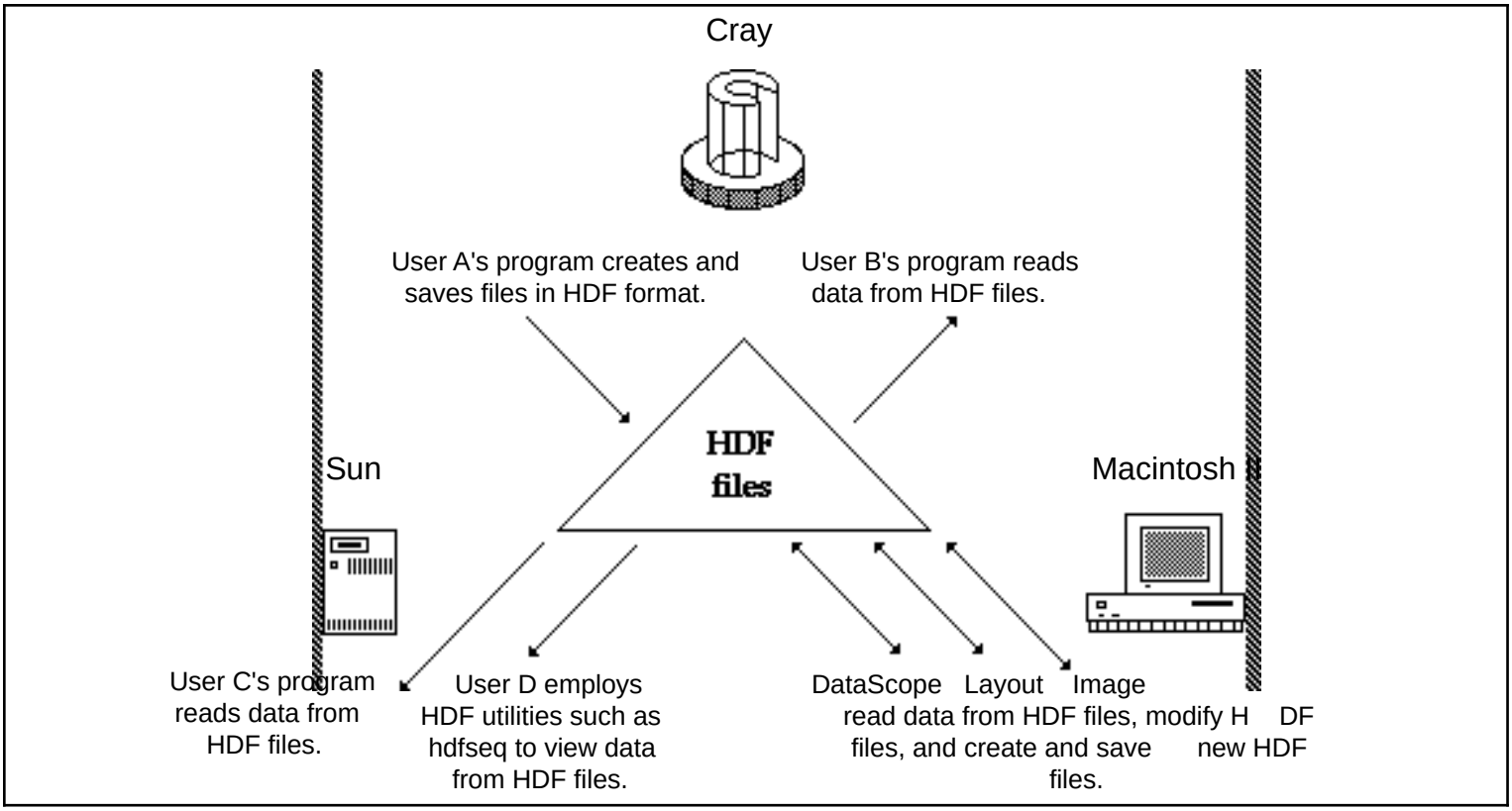
When you load a palette file, the hardware color table entries are remapped according to the new palette; that is, each color in the palette file is assigned to its corresponding entry in the hardware color table.

NOTE: The data values for user-defined palettes must be within the range of 0 to 255; however, the Macintosh reserves the palette values 0 and 255 for white and black, respectively. Consequently, you may have to scale your data to the range from 1 to 254, for NCSA Image overrides other assignments to 0 and 255 with white and black, respectively.

HDF Files

Hierarchical Data Format, or HDF, is a flexible, standard file format designed at NCSA for sharing of graphical and floating-point data among different programs and machines. This general purpose and extensible file format allows you to store raster images, their dimensions, color tables, and annotations all in the same file. In addition, you may store floating-point data, scaling information, dimensions, annotations, and units of measurement in a single file. HDF files are accessible from NCSA software for the Macintosh such as NCSA Image, NCSA DataScope, NCSA PalEdit, and NCSA Layout as well as user programs, other NCSA software, and HDF utilities. These files can be used on such machines as the CRAY X-MP/48, CRAY-2, Sun Workstation, IBM PC, and Alliant. The portability and usefulness of HDF files is shown in Figure 2.1.

Figure 2.1 HDF Environment



Why Should I Use HDF?

HDF lets you store datasets with extra file information about your data, for example, the dimensions of your image. This makes the files easier to read and manage by programs such as NCSA ImageTool and saves you the trouble of tracking this information externally.

What Information Goes into an HDF File?

Both raster images and scientific floating-point data can be stored in HDF files for use in NCSA ImageTool.

A *raster image set* contains a raster image, together with its dimensions. It may also contain a palette. You may specify that the raster image be stored in compressed or non-compressed form.

A *scientific dataset* can store scientific data in an array of 32-bit floating-point numbers of any dimension, together with information regarding the rank and size of each dimension. When storing your datasets in HDF files, you have the option of storing other information such as labels for the data and the axes, scales for the axes, and the maximum and minimum values of the data.

How Do I Create an HDF File?

Public domain software is available from NCSA for creating HDF files. NCSA's HDF libraries support both Fortran and C calls on any of the following machines: Cray (UNICOS), Sun (UNIX), Alliant (Concentrix), Macintosh, and IBM PC (MS-DOS).

The best way to store your data in an HDF file is to incorporate calls to the appropriate HDF library in the program that produces your image or scientific data. These calls can store your raw image, palette, scientific data, and other information in an HDF file in proper format.

If you have access to the HDF libraries, you can store floating-point data in your file, and read it directly. NCSA ImageTool reads floating-point data rather than 8-bit binary data whenever the former is available, but performs image processing operations on the 256-color representation (i.e., 8-bit binary data), *not* on the floating-point data from which the 256-color image was constructed.

NOTE: If you have a UNIX-based system, you can use the command line utility called r8tohdf to convert one or more raw raster images and palettes to HDF format.

Where Can I Obtain More Information about HDF?

Refer to Appendix C, "Obtaining NCSA Software," on information to obtain NCSA HDF as well as other NCSA software.

Specifying the Dimensions of Raw Raster Files

To read and display an image, NCSA Image must be able to ascertain its dimensions. If you use an image stored in an HDF file, which can contain an image along with information about the image dimensions and associated color table, NCSA Image can read the dimensions from the HDF file.

However, raw raster files do not provide this convenience. You must specify the image dimensions in the dimension areas
(Xdim, Ydim) of the control panel.