A Standard Magnetic Tape Format

for

Digital Image Exchange

BRENT S. BAXTER
LEWIS E. HITCHNER
GERALD Q. MAGUIRE JR.

UTEC-81-018

UNIVERSITY OF UTAH
DEPARTMENTS OF
COMPUTER SCIENCE AND RADIOLOGY
SALT LAKE CITY, UTAH 84112

15 April 1981

This work was supported in part by: NIH Grant R01 CA 24822 and The Department of Radiology of the University of Utah

Table of Contents

ABSTRACT

This proposal describes a simple yet flexible magnetic tape format for exchanging digital image information. Multi-dimensional arrays of raster scan data are stored as sequential files, with descriptive information related to these files stored in an initial directory file. Directory information is stored as ASCII "key value pair" character strings that can be read by people as well as computer programs. Key value pairs are used for directory searches, also to identify and provide supplementary information about the images. Since the directory is relatively self explanatory it is possible in many cases to write a program for reading an AAPM tape without reference to a published standard.

This proposal conforms with applicable sections of ANSI standards X3.22-1973 and X3.39-1973, "Recorded Magnetic Tape for Information Interchange" (800 and 1600 BPI) X3.40-1976, "Unrecorded Magnetic Tape for Information Interchange", and X3.27-1969, "Standard Magnetic Tape Labels for Information Interchange".

1. Introduction

1.1 Background

The data format described in this proposal provides a simple yet powerful means for exchanging digital images stored on magnetic tape. The primary goal is to facilitate the exchange of data between groups having otherwise dissimilar equipment. Every effort has been made to keep the proposed format simple and easy to use while anticipating the need for future additions and modifications.[1] This standard is not proposed as a substitute for the working file system provided by the manufacturer of the computer equipment. Several criteria which have strongly influenced the proposal are as follows:

Simplicity — The tape format should comprise a small number of easily understood structures. Directory information should be self explanatory and should be readable by people as well as by computers. Complex data formats should be avoided.

Flexibility — It should be possible to include a variety of descriptive information in the directory and to add new items as needed. Commonly used image data types should be easily accommodated by the standard. Revisions of the standard should not obsolete tapes conforming to earlier versions of the standard. There should be no more constraints than necessary.

Addressability — Directory searches, based on matching one or more keys such as "date", "image #", "patient name", etc., should enable one to locate the image data by skipping to the required file.

Universal Readability — Computer hardware requirements, record lengths, etc., should be such as to allow the widest possible distribution of tapes meeting the standard.

These criteria are reflected in the organization of the tape and the

---

[1] This document is an outgrowth of an earlier proposal by Schimpf [2].

directory, as well as in the structure of the key value mechanism.

1.2 Hardware Considerations

Industry standard 1/2 inch, nine track tape written in odd parity is the media. The recording density may be 800 or 1600 bits per inch, and the record size must be 2048 bytes (the largest record size allowed by ANSI X3.22). Records of this length utilize approximately 75% of the tape surface at a recording density of 1600 BPI and an even greater percentage at 800 BPI. The ASCII character code is to be used.[2] Characters are to be recorded on the tape in the order they appear on a printed page. Numeric data are recorded most significant byte first. Thus, a byte swapping program may be required on some machines, particularly on DEC PDP-11 equipment. The record length (2048 bytes), the recording density, , and if an ANSI standard tape label is used, the volume ID must all be affixed to the tape reeel.

An explanation of the directory and data file structures may be found in the next section. Appendix material includes a sample tape directory, a list of suggested AAPM format type key value pairs, listings of programs to perform common operations needed to access information on the tape, and a glossary of magnetic tape terminology.

---

[2]Machines which use other codes internally generally provide conversion programs to and from ASCII, while most small machines only support ASCII.

2. Tape Organization

2.1 Directory File Structure

Descriptive information is contained in a directory at the beginning of the tape. Directory information is placed there, rather than in file headers scattered throughout the tape, to make searches and directory listings quick and easy. The directory consists of one or more 2048 byte records followed by an unused 2048 byte record followed by an end of file mark (EOF).

Directory information is in the form of key value pairs[3], with a key beginning a line, possibly with leading spaces or tabs[4] followed by a colon/equal, followed by the value, and terminated by <CR LF>. A line should be no longer than 80 characters to prevent wrap-around problems on standard width terminals. No more than one key value pair may appear on a single line of text. Both upper and lowercase text is acceptable. Lowercase and uppercase text are considered equivalent when searching.

There are three advantages to this arrangement in comparison with a fixed format approach. First, the directory can be read and understood without reference to a published standard. Second, the directory can be examined and manipulated using widely available file system utilities and editors. Third, variable length names and comments are easily accommodated.

---

[3]See [3] for details of a "name-type-value" system similar to the key value structure outlined here.

[4]The AAPM software will convert tabs to sequences of spaces before writing text on the tape. See [1] page 23 for details of how this may be done.

2.2 Directory Header

Certain key value pairs which pertain to the entire tape are located at the beginning of the directory ahead of any other directory entries. These serve to identify the tape, its owner, his institutional affiliation, the AAPM tape format version number, etc. This section of the directory begins with the "Number of records in directory" key value pair (required) followed by any of the optional key value pairs. The header is terminated by the appearance of the first "image #" key value pair or, in the case of an empty directory, by the end of the directory[5].

2.3 Directory Entries

Directory entries follow the directory header. Each entry begins with the required "image #" key value pair:

image # := N

where the number N refers to the Nth image file. Other key value pairs associated with this entry follow in any order. A directory entry is terminated by a succeeding "image #" key value pair or by the end of the directory. The directory is stored in file 0, the first image data file is found in file 1, the second is file 2, ...

---

[5]The end of the directory is at the end of the Nth record rather than at the EOF because an extra unused record is placed after the end of the directory to alleviate hardware related problems associated with appending to the directory.

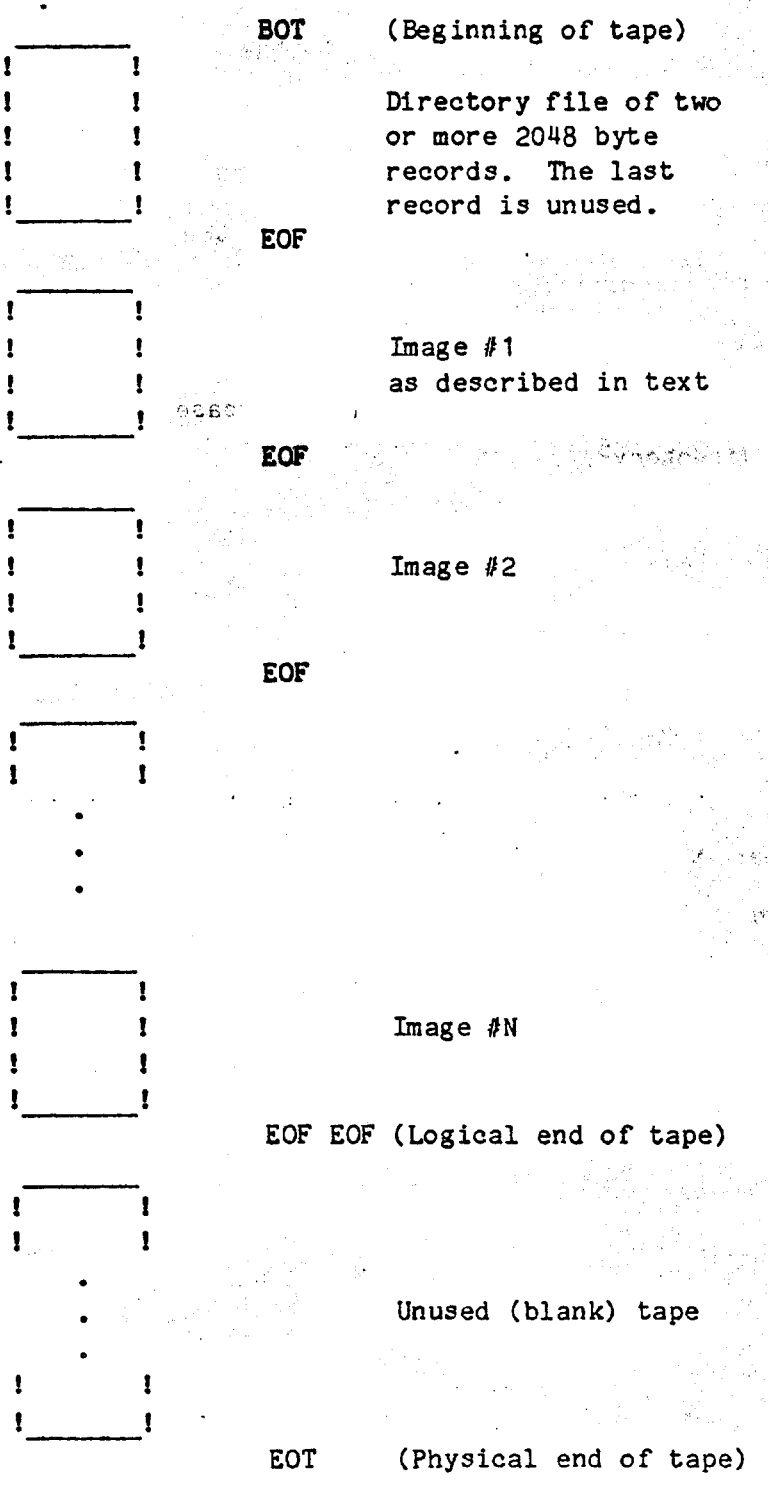

Figure 2-1: Tape Organization

2.3.1 Image Format Keys (required)

The minimum information necessary to locate individual picture elements (pixels) within image files is contained in the following required key value pairs:

```
Image # :=
Bytes per pixel :=
Number of dimensions :=
Size of dimension 1 :=
Size of dimension 2 :=
   .
   .
   .
```

Picture elements are recorded by exhausting the first index first, then the second index ...etc. In the following example pixel (27,33,3) will be found in bytes 53 and 54 of record 36 in image #1. Note that indexing begins with one.

```
Image # := 1
Bytes per pixel := 2
Number of dimensions := 3
Size of dimension 1 := 128
Size of dimension 2 := 128
Size of dimension 3 := 8
```

These keys must always be present in a file entry because they are needed to locate picture element data within an image file. Other optional annotation key value pairs may also be present as described in the next section.

Color and other multiband images may be treated as three dimensional arrays with either the first or third dimension indicating the band. If the data are arranged on the tape in a pixel interleaved fashion then the dimensions might be as follows:

```
Size of dimension 1 := 3
Size of dimension 2 := 128
Size of dimension 3 := 128
```

If bands are recorded in sequence with one band appearing in its entirety before the next. The directory entry would appear as follows:

```
Size of dimension 1 := 128
Size of dimension 2 := 128
Size of dimension 3 := 3
```

Image files too large to fit on a single reel of tape should be placed on a multi tape volume using the ANSI standard tape volume label mechanism[6].

2.3.2 Annotation Keys (optional)

Annotation keys appear following the image format keys described above, or they may be omitted at the option of the person preparing the tape. A collection of key value pairs to be used by AAPM users is included as Appendix I. Any key value pair may be used for a directory search though annotation keys will probably be most useful for the purpose. Most revisions and changes to the standard will be accomplished by updating the list of annotation keys (Appendix I).

2.3.3 Comments

Free field comments may appear anywhere in the directory except on lines containing key value pairs. Lines not containing a ":=" sequence are ignored during directory search operations.

---

[6]ANSI Standard X3.27-1969

2.3.4 Directory Inquiries

Directory searches will be done by locating the required key and examining its value. If the value matches, the image number is returned to allow skipping ahead to the proper image file. A small library of AAPM routines will be provided to assist with certain elementary directory access operations. These are:

List the directory header
: This routine simply lists the directory header. Note that this could be done on most systems with their file utilities.

List a directory entry
: This routine lists the specified directory entry. This operation could also be done on most systems with their file utilities or by means of a text editor.

Key search
: This routine searches through the directory stopping on the next match with the key value pair. The file number is returned along with the result of the search. Possible search outcomes might include: a match, no match, or no such key. Matches require exact spelling but disregard case shifts and extra tabs and spaces.

2.4 Image File Format

Data files are recorded as rectangular arrays of numeric data padded with zeros where necessary to fill out the array. A single datum is stored in one or more bytes (most significant byte first, see 2.3.1). Data compression such as run length encoding, DPCM, chain coding, etc., is not permitted to avoid complicating the task of exchanging image data.

2.5 Data Transfer Operations

A minimum number of data transfer operations are envisioned because the standard is intended primarily for data interchange. General purpose data base management is not within the scope of this standard.

Initialize the tape

This operation records the tape header information at the beginning of the directory file and fills the remainder of the directory with null characters. One extra unused record is added to the end of the directory to minimize the possibility of problems with the append operation. This extra record provides leeway for variations in recording density and inter record gap length on different tape drives.

Append an image file

An image file is appended to the tape and a directory entry is appended to the directory file.

Read an image file

The required number of image files is skipped, then data is read from the specified image file.

No explicit provision is made in the AAPM programs for deleting or updating images. These more complex operations are outside the scope of a standard intended primarily for image data exchange.

Other common operations such as rewinding and copying tapes, writing ANSI standard volume labels, etc., are to be left to utility programs supplied with each computer system.

3. Conclusion

The standardized tape file system described here provides features needed for a convenient image interchange medium between computer equipment of different manufacturers. Descriptive annotation information is provided to identify image files and to permit rapid searches of the tape for images identified by name. A limited amount of data structure information is included to indicate the number of bytes per picture element, number of rows,columns, etc. This descriptive information is recorded in a form understandable by people as well as by computers.

4. References

[1] Kernighan, B. W. and Plauger, P. J.
Software Tools.
Addison and Wesley, Reading, Massachusetts, 1976.

[2] Schimpf, J. S. .
AAPM Tape Format Standard.
Internal Draft Dated July 9, 1979. (NYU).

[3] Selfridge, P. G. .
A Flexible Data Structure for Accessory Image Information.
Technical Report TR-45, Department of Computer Science, University of Rochester, May, 1979.

I. Annotation Keys for use by the AAPM Community

The AAPM directory keys and their associated values are described in this appendix. Keys are placed at the beginning of a line followed by a colon/equal sequence, followed by the value, followed by a carriage return/line feed sequence. Note that a key must begin on a new line. Also, no more than one key may appear on a single line.

(Header.....)

Number of records in directory := <integer> (required)
This is the first key value pair on the tape. It is suggested that the directory be no more than 16 records in length to minimize problems in appending to the directory.

Tape Standard := <decimal-number>
The first release will be 1.00 and updates will be published as appropriate.

Directory header := <character string>
Description of the tape.

Institution := <character string>
Name of the institution.

Department := <character string>
Further identification of institutional affiliation.

Date created := <day>,<month>,<year> (all decimal integers)
The calendar date the tape was created.

(Directory entries .....)

Image # := <integer>	(required)
Bytes per pixel := <integer>	"
Integer of dimensions := <integer>	"
Size of dimension 1 := <integer>	"
Size of dimension 2 := <integer>	"

.
.
.

Date written := <day>,<month>,<year> (all integers)
This value is the calendar date the image
was recorded on this tape.

Writer := <character string>
The name of the person who created the file.

Grid 1 units := <decimal-number>
The grid spacing in cm. in the "X" direction

Grid 2 units := <decimal-number>
The grid spacing in cm. in the "Y" direction

Grid 3 units := <decimal-number>
The grid spacing in cm. in the "Z" direction

.
.
.

Number representation := <character string>
Number system used to record image data
Common choices might include:

Two's complement integer
Positive integer

Patient name := <character string>
Patient's name.

Patient number := <character string>
Patient's local hospital code.

Exam type := <character string>
Identifies the type of examination.

II. Sample Tape Directory

Number of records in directory := 16 (required)

Tape Standard # := 1.00

Directory header := Sample AAPM digital image tape

Institution := University of Utah Medical Center

Department := Radiology Department, Nuclear Medicine Division

Date created := 17,3,80

.....
Note that the preceding line of periods will be treated
as a comment, as will this explanation.

Image # := 1 (required)

Bytes per pixel := 2 "
Number of dimensions := 3 "
Size of dimension 1 := 128 "
Size of dimension 2 := 128 "
Size of dimension 3 := 8 "

Date created := 17,3,80
Date written := 20,3,80
Patient name := Sam Jones
Exam type := Liver spleen study

This is a very interesting case because Jones has no liver!
Note the absence of any key or colon/equal sequence in this
free field comment.

Image # := 2 (required)

This is a test image.

Bytes per pixel := 1 "
Number of dimensions := 2 "
Size of dimension 1 := 64 "
Size of dimension 2 := 64 "

Image # := 3 "

.
.
.

<EOF>

III. Descriptions of AAPM Computer Programs.

Section 2.5 summarizes the programs to be supplied with this standard. As programs are developed and tested, brief descriptions will be added to this section.

IV. Glossary of Magnetic Tape Terminology

BOT — Beginning of tape. This point is identified by a reflective marker at the beginning of a reel of magnetic tape.

BPI — The recording density in bytes per inch, also referred to as bits per inch.

BYTE — A group of 8 binary bits often used to encode one alpha/numeric character.

EOF — End of file marker. This is a pattern recorded on magnetic tape to mark the end of a file. Typically a tape drive can both sense these and skip to the next EOF to locate specific files on the tape.

EOT — End of tape. A reflective marker at the end of a reel of magnetic tape identifies the physical end of the tape.

FILE — One or more records followed by an EOF.

IRG — Inter-record gap. A short length of blank tape used to separate a file into small units that can be manipulated in the computer's memory.

RECORD — A set of data bytes between inter-record gaps. The record must be greater than some minimum set by the tape drive hardware and smaller than the attached computer's memory. This standard specifies 2048 byte records.