Digital Forensics Workshop (CSE3156)

File Recovery & Data Craving using *formemost*

Computer Science Department, SOA University

Overview

- data recovery and data carving
- foremost

Meta data

Metadata, or "data about data," helps the operating system identify data. Metadata

includes technical information, such as the creation and modification dates and the filetype of the data

Datarecovery VS Data Carving

Data Recovery

It is the process of retrieving data from damaged, corrupted, failed, or inaccessible storage devices when it cannot be accessed normally.

Recovering files from a crashed hard drive.

Retrieving data after accidental deletion or formatting.

Restoring files from a corrupted partition or file system.

Data/file Carving

It is a technique used to extract data from a storage medium based on file signatures or patterns, without relying on the file system structure.

Recovering deleted files when the file system has been corrupted or destroyed.

Extracting files from unallocated space on a disk or from memory dumps.

Forensic investigations to recover files without relying on the file system.

Data Recovery VS Data Carving

File recovery techniques make use of the file system information and, by using this information, many files can be recovered. If the information is not correct, then it will not work.

File carving works only on raw data on the media and it is not connected with file system structure.

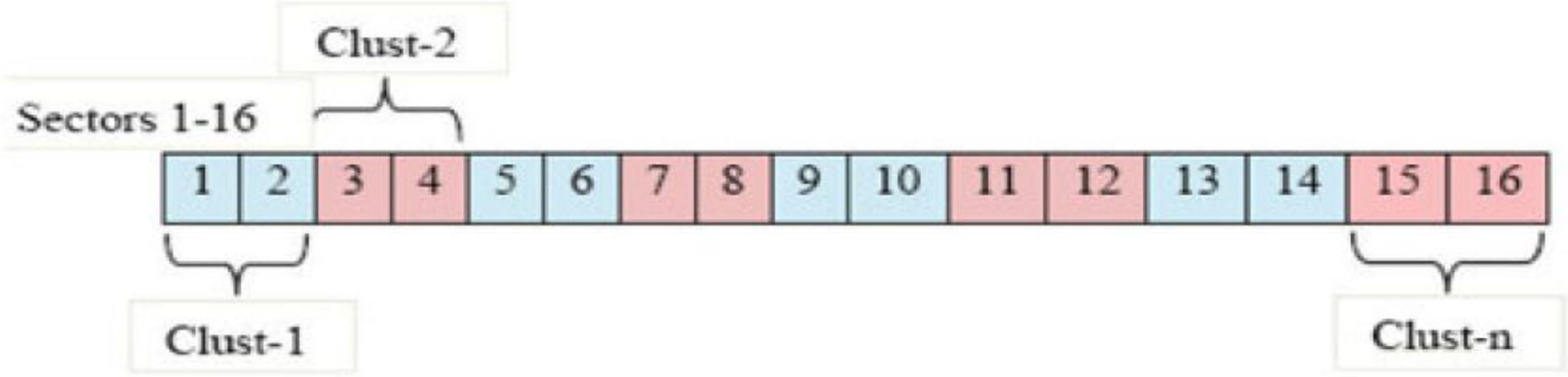
If a file header were damaged, recovery of a file would be impossible. Data carving is possible even if a file header is damaged, or if a file is fragmented or damaged.define carve

Data Carving

Lost Cluster

Lost clusters are the clusters which are allocated to a file but are not having reference in the **file allocation table.**

If a file is stored in the second cluster, its header signature is stored in first few bytes of the second cluster or third sector.

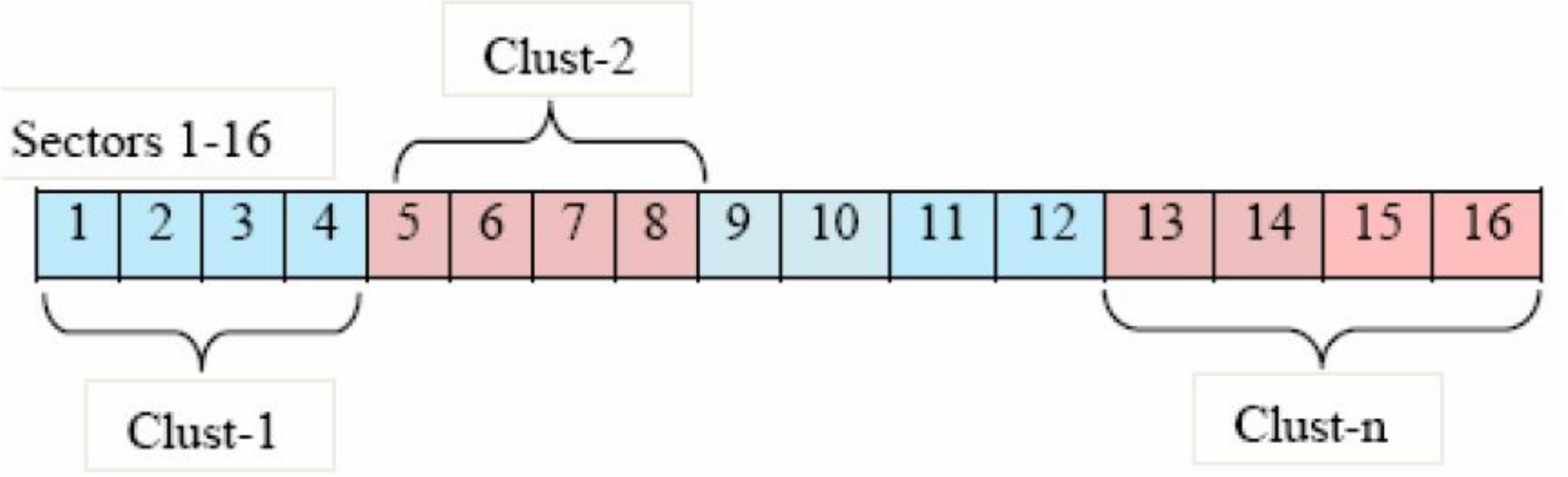


A cluster is defined as a logical unit of file storage on a hard disk.
Source

Data Carving

Lost Cluster

If the same hard disk is formatted and its cluster size is 4 sectors as shown in below figure, this search option misses the file starting at sector 3.



A cluster is defined as a logical unit of file storage on a hard disk.

Data Carving

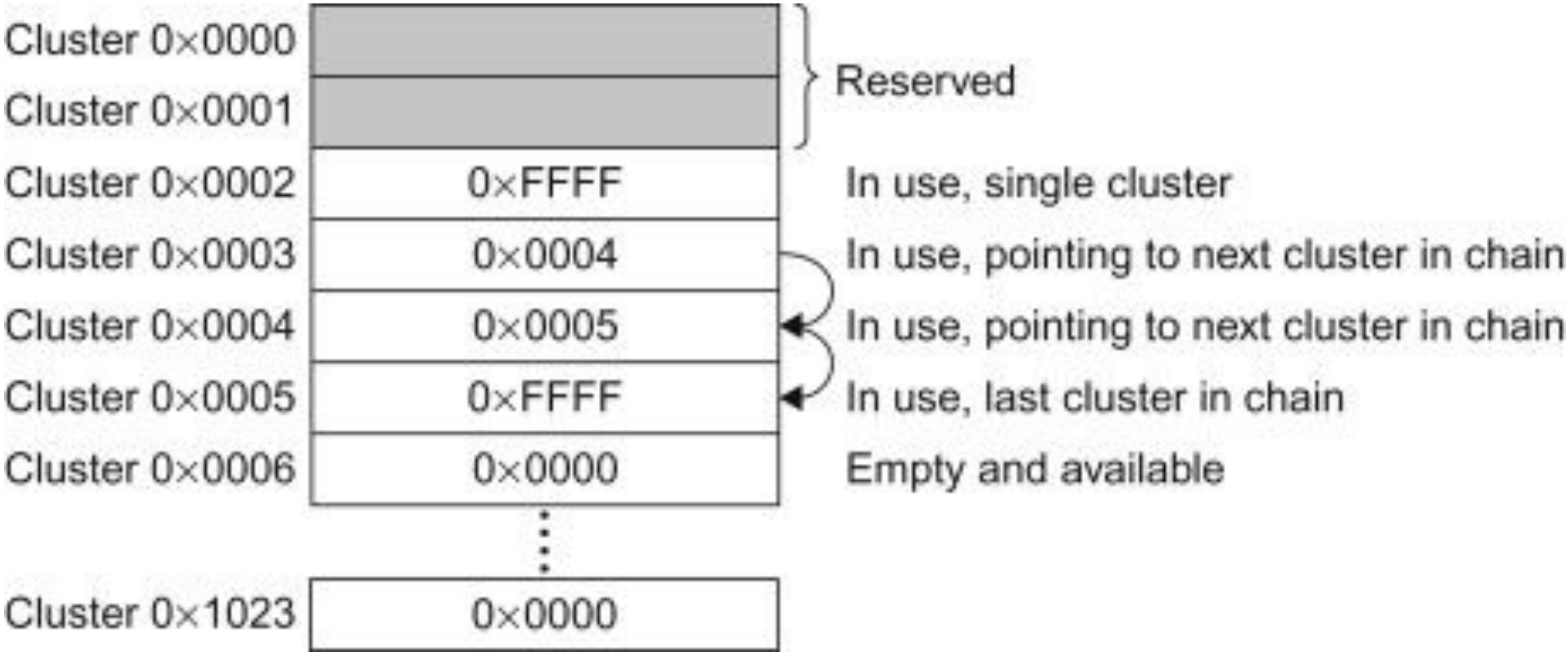
Possibility of Lost Cluster

1. Change of File Format may cause change of cluster size. Therefore lost reference in **file allocation table(FAT).**

2. files not being closed properly, from shutting down a computer without first closing an application (power failure) or from ejecting a storage medium, such as a floppy disk, from the disk drive while the drive is reading or writing.

NB: Formatting delete FAT

File allocation table.

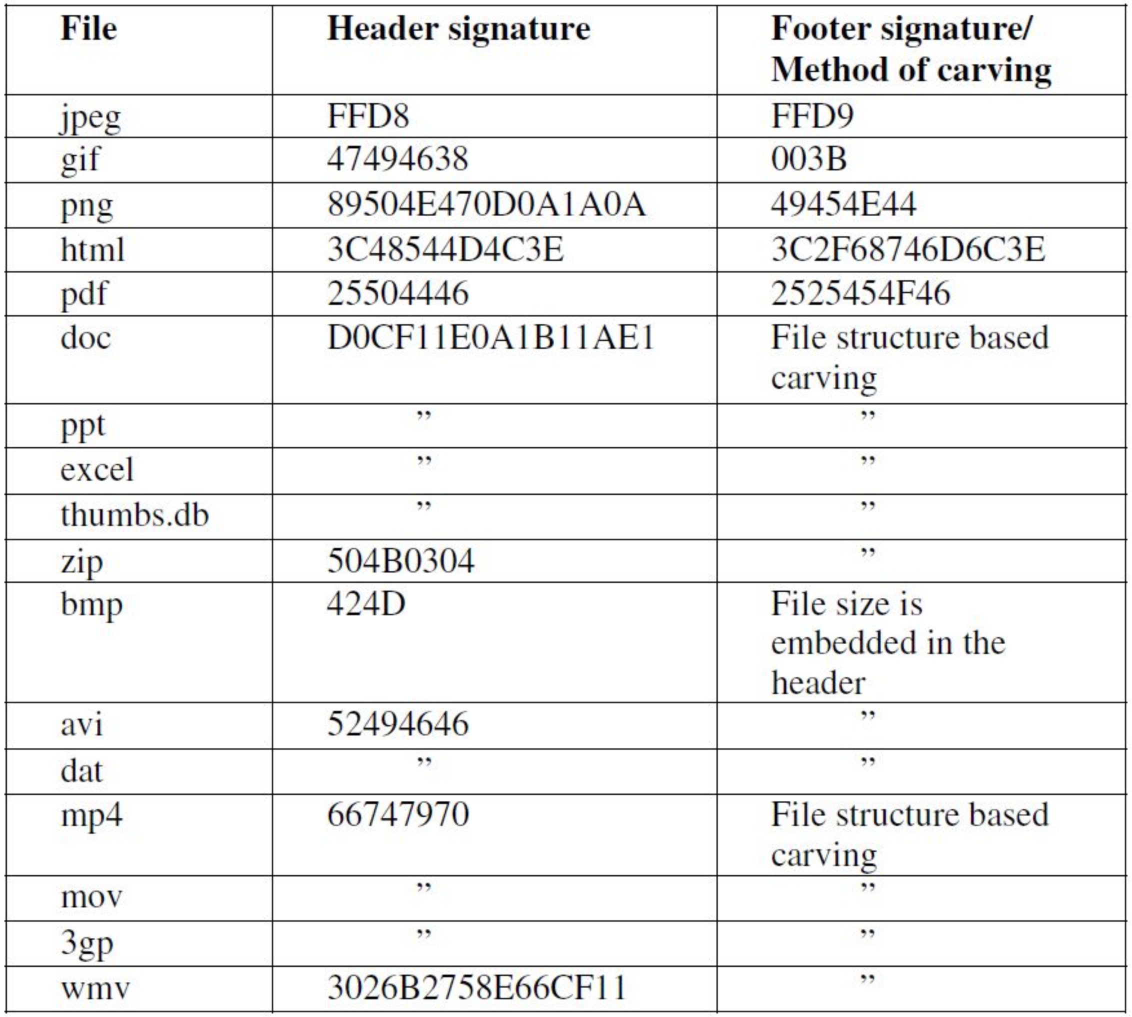


Cluster 0×0000		Reserved
Cluster 0×0001		
Cluster 0×0002	0×FFFF	In use, single cluster
Cluster 0×0003	0×0004	In use, pointing to next cluster in chain
Cluster 0×0004	0×0005	In use, pointing to next cluster in chain
Cluster 0×0005	0×FFFF	In use, last cluster in chain
Cluster 0×0006	0×0000	Empty and available
⋮	⋮	
Cluster 0×1023	0×0000	

The **FAT32** and **exFAT** file systems are popular for external storage devices like USB drives and SD cards because they are compatible across most major operating systems, including Windows, macOS, Linux, and even gaming consoles.

Source

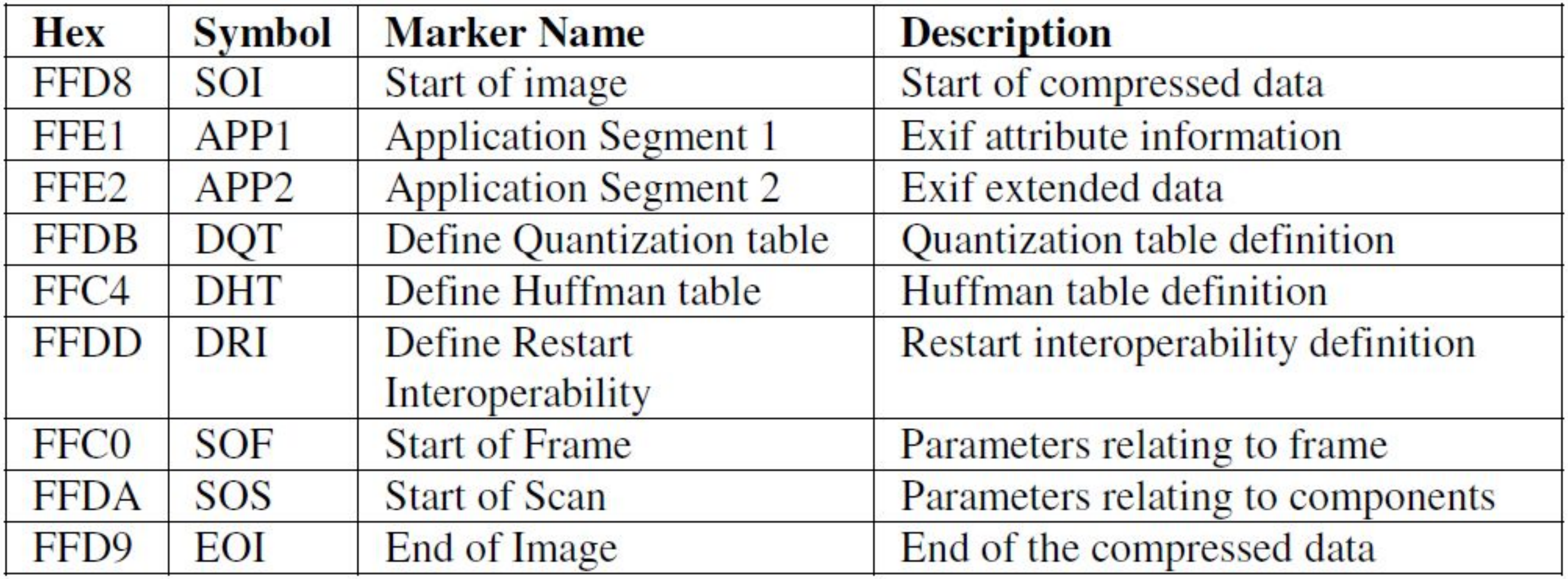
Header/Footer Carving



File	Header signature	Footer signature/ Method of carving
jpeg	FFD8	FFD9
gif	47494638	003B
png	89504E470D0A1A0A	49454E44
html	3C48544D4C3E	3C2F68746D6C3E
pdf	25504446	2525454F46
doc	D0CF11E0A1B11AE1	File structure based carving
ppt	”	”
excel	”	”
thumbs.db	”	”
zip	504B0304	”
bmp	424D	File size is embedded in the header
avi	52494646	”
dat	”	”
mp4	66747970	File structure based carving
mov	”	”
3gp	”	”
wmv	3026B2758E66CF11	”

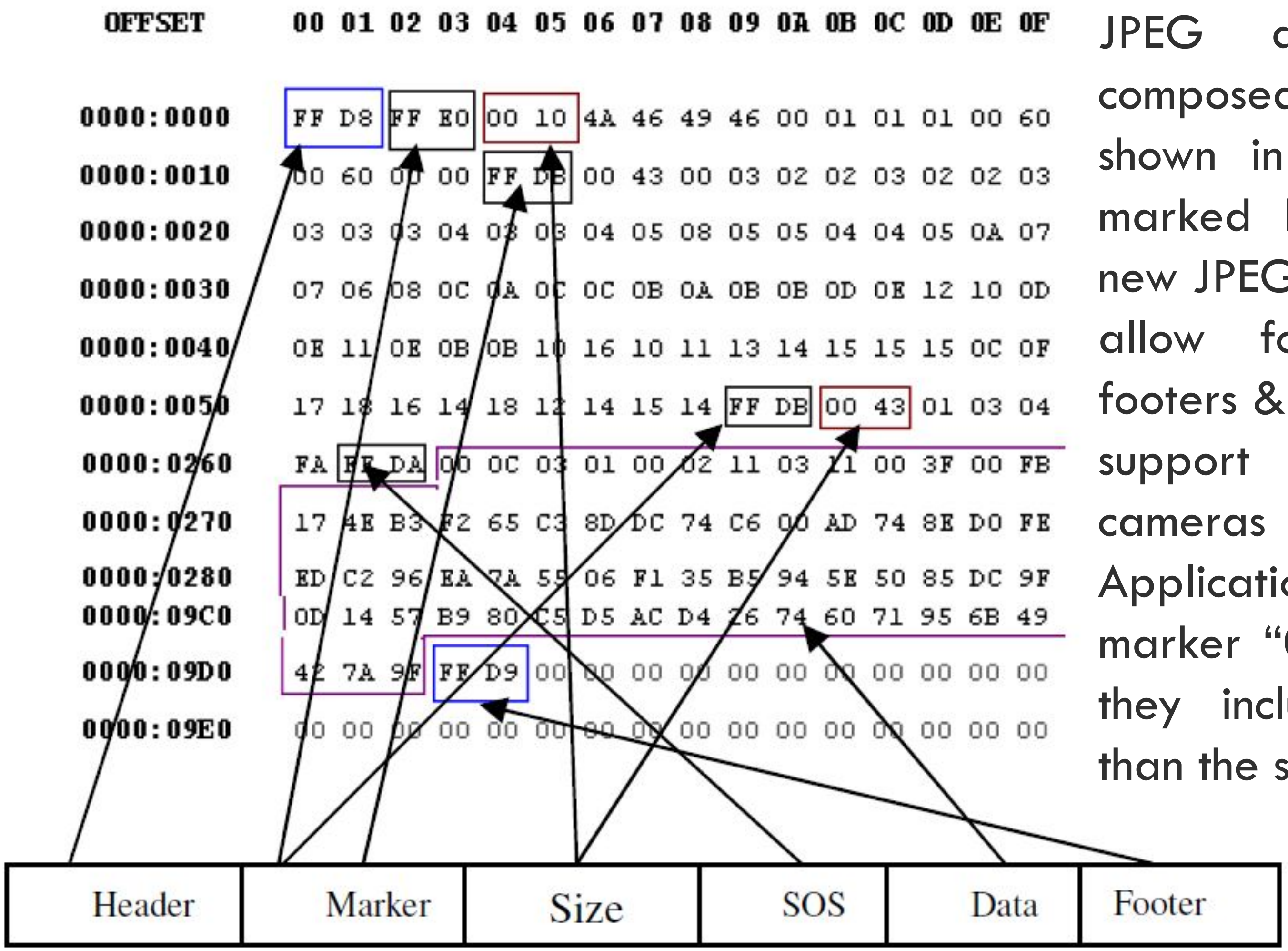
This method of carving files is used when a file has defined header and footer. Jpeg, gif, png, html, pdf etc., may fall under this category.

Header/Footer Carving



Hex	Symbol	Marker Name	Description
FFD8	SOI	Start of image	Start of compressed data
FFE1	APP1	Application Segment 1	Exif attribute information
FFE2	APP2	Application Segment 2	Exif extended data
FFDB	DQT	Define Quantization table	Quantization table definition
FFC4	DHT	Define Huffman table	Huffman table definition
FFDD	DRI	Define Restart Interoperability	Restart interoperability definition
FFC0	SOF	Start of Frame	Parameters relating to frame
FFDA	SOS	Start of Scan	Parameters relating to components
FFD9	EOI	End of Image	End of the compressed data

Header/Footer Carving

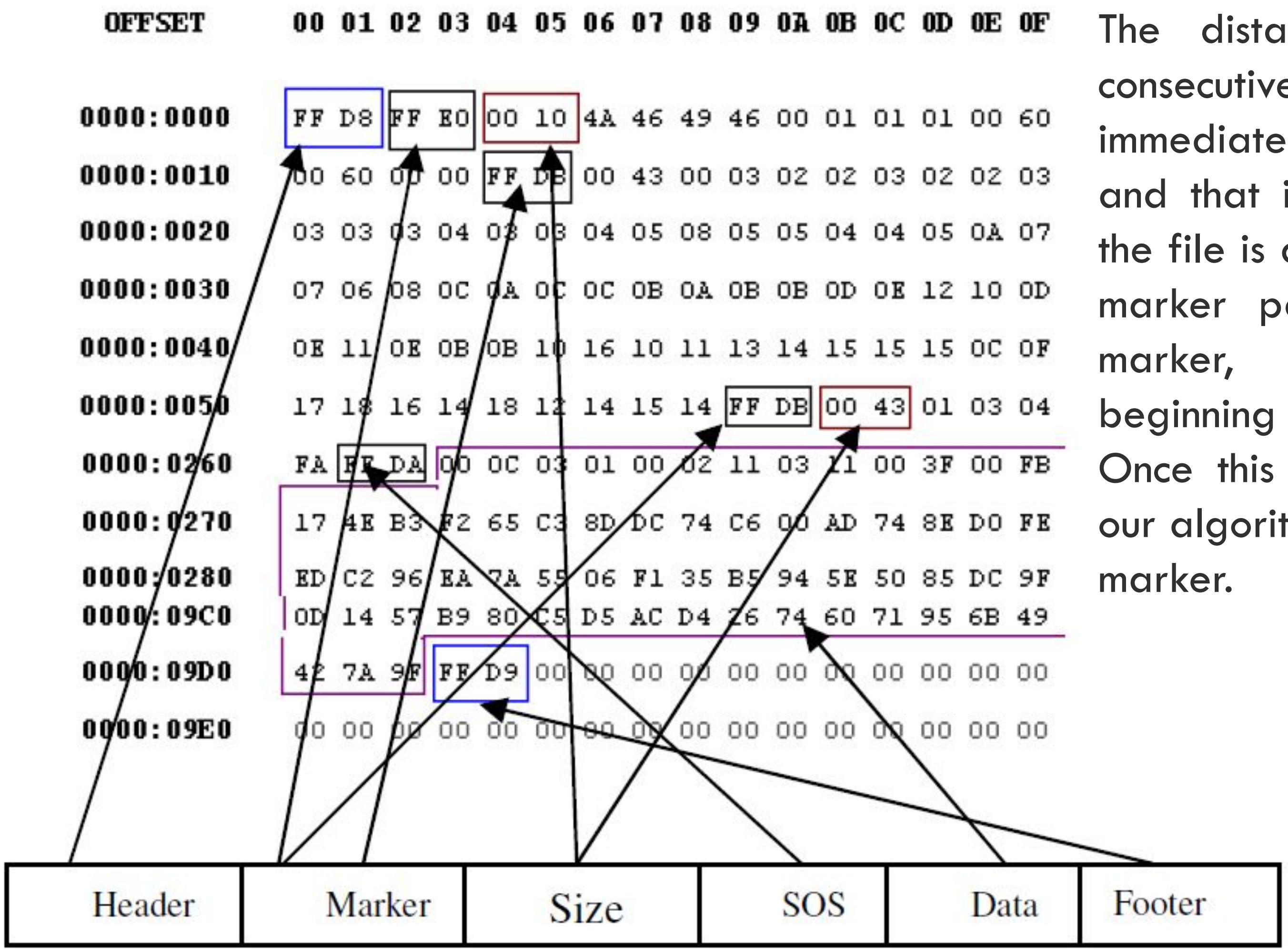


```
OFFSET      00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F

0000:0000   FF D8 FF E0 00 10 4A 46 49 46 00 01 01 01 00 60
0000:0010   00 60 00 00 FF DB 00 43 00 03 02 02 03 02 02 03
0000:0020   03 03 03 04 03 03 04 05 08 05 05 04 04 05 0A 07
0000:0030   07 06 08 0C 0A 0C 0C 0B 0A 0B 0B 0D 0E 12 10 0D
0000:0040   0E 11 0E 0B 0B 10 16 10 11 13 14 15 15 15 0C 0F
0000:0050   17 18 16 14 18 12 14 15 14 FF DB 00 43 01 03 04
0000:0260   FA FF DA 00 0C 03 01 00 02 11 03 11 00 3F 00 FB
0000:0270   17 4E B3 F2 65 C3 8D DC 74 C6 00 AD 74 8E D0 FE
0000:0280   ED C2 96 EA 7A 55 06 F1 35 B5 94 5E 50 85 DC 9F
0000:09C0   0D 14 57 B9 80 C5 D5 AC D4 26 74 60 71 95 6B 49
0000:09D0   42 7A 9F FF D9 00 00 00 00 00 00 00 00 00 00 00
0000:09E0   00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
```

JPEG data structures are composed of segments (as shown in this Table, that are marked by identifiers. As per new JPEG spe., the new formats allow for multiple headers, footers & even nested images to support thumbnails. Digital cameras often utilize the Application (APP) segment marker “0xffe1” to signify that they include more meta-data than the standard JFIF.

Header/Footer Carving



```
OFFSET      00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F

0000:0000   FF D8 FF E0 00 10 4A 46 49 46 00 01 01 01 00 60
0000:0010   00 60 00 00 FF DB 00 43 00 03 02 02 03 02 02 03
0000:0020   03 03 03 04 03 03 04 05 08 05 05 04 04 05 0A 07
0000:0030   07 06 08 0C 0A 0C 0C 0B 0A 0B 0B 0D 0E 12 10 0D
0000:0040   0E 11 0E 0B 0B 10 16 10 11 13 14 15 15 15 0C 0F
0000:0050   17 18 16 14 18 12 14 15 14 FF DB 00 43 01 03 04
0000:0260   FA FF DA 00 0C 03 01 00 02 11 03 11 00 3F 00 FB
0000:0270   17 4E B3 F2 65 C3 8D DC 74 C6 00 AD 74 8E D0 FE
0000:0280   ED C2 96 EA 7A 55 06 F1 35 B5 94 5E 50 85 DC 9F
0000:09C0   0D 14 57 B9 80 C5 D5 AC D4 26 74 60 71 95 6B 49
0000:09D0   42 7A 9F FF D9 00 00 00 00 00 00 00 00 00 00 00
0000:09E0   00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
```

The distance between any two consecutive markers is stored immediately after the first marker and that is two bytes in length. If the file is a valid JPEG then the last marker parsed will be the SOS marker, which signifies the beginning of the actual image data. Once this marker is reached then, our algorithm looks for the “0xffd9” marker.

foremost

The syntax for using foremost is as follows:

foremost -i (forensic image) -o (output folder) -options

example

foremost -i *image.dd* -o *recovery* -t file *format to recover*

In this example, we have specified the 11-carve-fat.dd file located on the desktop as the input file (-i) and specified an empty folder, named Foremost_recovery, as the output file (-o). Additionally, other switches can also be specified as needed.

foremost

Data set 1 to investigate using foremost.

Data set 2 is you pen drive can be used for foremost.

If partition of pen drive causes write block

```
sudo mount /dev/sdbx //media/your_dir
```

```
sudo chown yourusername:yourusername /media/yourusername/yourusbdrive
```

Thank You