Before the image data is ever loaded when a JPEG image is selected for viewing the markers must be read.  In a JPEG image, the very first marker is the SOI, or Start Of Image, marker.  This is the first "hey, I'm a JPEG" declaration by the file.  The JPEG standard, as written by the Joint Picture Expert's Group, specified the JPEG interchange format.  This format had several shortcomings for which the JFIF (JPEG File Interchange Format) was an attempted remedy.  The JFIF is the format used by almost all JPEG file readers/writers.  It tells the image readers, "Hey, I'm a JPEG that almost anyone can understand."

Most markers will have additional information following them.  When this is the case, the marker and its associated information is referred to as a "header."   In a header the marker is immediately followed by two bytes that indicate the length of the information, in bytes, that the header contains.  The two bytes that indicate the length are always included in that count.

A marker is prefixed by FF (hexadecimal).  The marker/header information that follows does not specify all known markers, just the essential ones for baseline JPEG.

A component is a specific color channel in an image.  For instance, an RGB image contains three components; Red, Green, and Blue.

© 1998 by James R. Weeks

Start of Image (SOI) marker -- two bytes (FFD8)

JFIF marker (FFE0)

|  |
| --- |
|  |

* length -- two bytes
* identifier -- five bytes: 4A, 46, 49, 46, 00 (the ASCII code equivalent of a zero terminated "JFIF" string)
* version -- two bytes: often 01, 02
  + the most significant byte is used for major revisions
  + the least significant byte for minor revisions
* units -- one byte: Units for the X and Y densities
  + 0 => no units, X and Y specify the pixel aspect ratio
  + 1 => X and Y are dots per inch
  + 2 => X and Y are dots per cm
* Xdensity -- two bytes
* Ydensity -- two bytes
* Xthumbnail -- one byte: 0 = no thumbnail
* Ythumbnail -- one byte: 0 = no thumbnail
* (RGB)n -- 3n bytes: packed (24-bit) RGB values for the thumbnail pixels, n = Xthumbnail \* Ythumbnail

Define Quantization table marker (FFDB)

|  |
| --- |
|  |

* the first two bytes, the length, after the marker indicate the number of bytes, including the two length bytes, that this header contains
* until the length is exhausted (loads two quantization tables for baseline JPEG)
  + the precision and the quantization table index -- one byte: precision is specified by the higher four bits and index is specified by the lower four bits
    - precision in this case is either 0 or 1 and indicates the precision of the quantized values; 8-bit (baseline) for 0 and  up to 16-bit for 1
  + the quantization values -- 64 bytes
    - the quantization tables are stored in zigzag format

Define Huffman table marker (FFC4)

|  |
| --- |
|  |

* the first two bytes, the length, after the marker indicate the number of bytes, including the two length bytes, that this header contains
* until length is exhausted (usually four Huffman tables)
  + index -- one byte: if >15 (i.e. 0x10 or more) then an AC table, otherwise a DC table
  + bits -- 16 bytes
  + Huffman values -- # of bytes = the sum of the previous 16 bytes

Start of frame marker (FFC0)

|  |
| --- |
|  |

* the first two bytes, the length, after the marker indicate the number of bytes, including the two length bytes, that this header contains
* P -- one byte: sample precision in bits (usually 8, for baseline JPEG)
* Y -- two bytes
* X -- two bytes
* Nf -- one byte: the number of components in the image
  + 3 for color baseline JPEG images
  + 1 for grayscale baseline JPEG images
* Nf times:
  + Component ID -- one byte
  + H and V sampling factors -- one byte: H is first four bits and V is second four bits
  + Quantization table number-- one byte

The H and V sampling factors dictate the final size of the component they are associated with. For instance, the color space defaults to YCbCr and the H and V sampling factors for each component, Y, Cb, and Cr, default to 2, 1, and 1, respectively (2 for both H and V of the Y component, etc.) in the Jpeg-6a library by the Independent Jpeg Group. While this does mean that the Y component will be twice the size of the other two components--giving it a higher resolution, the lower resolution components are quartered in size during compression in order to achieve this difference. Thus, the Cb and Cr components must be quadrupled in size during decompression.

Start of Scan marker (FFDA)

|  |
| --- |
|  |

* the first two bytes, the length, after the marker indicate the number of bytes, including the two length bytes, that this header contains
* Number of components, n -- one byte: the number of components in this scan
* n times:
  + Component ID -- one byte
  + DC and AC table numbers -- one byte: DC # is first four bits and AC # is last four bits
* Ss -- one byte
* Se -- one byte
* Ah and Al -- one byte

Comment marker (FFFE)

|  |
| --- |
|  |

* the first two bytes, the length, after the marker indicate the number of bytes, including the two length bytes, that this header contains
* whatever the user wants

End of Image (EOI) marker (FFD9)

|  |
| --- |
|  |

* the very last marker

------------------------------------------------

JPEG is rather complex in this aspect, so we shall just give an overview of the basic principles (see the JPEG Book, chapter 7 for the full picture).

JPEG data is divided into *segments*, each of which starts with a 2-byte *marker*.

All markers are byte-aligned - they start on the byte boundaries of the transmission/storage medium. Any variable-length data which precedes a marker is padded with extra ones to achieve this.

The first byte of each marker is FFH FFH . The second byte defines the type of marker.

To allow for recovery in the presence of errors, it must be possible to detect markers without decoding all of the intervening data. Hence markers must be unique. To achieve this, if an FFH FFH byte occurs in the middle of a segment, an extra 00H 00H *stuffed* byte is inserted after it and 00H 00H is never used as the second byte of a marker.

Some important markers in the order they are often used are:

| **Name** | **Code (hex)** | **Purpose** |
| --- | --- | --- |
| SOI | FFD8 | Start of image. |
| COM | FFFE | Comment (segment ignored by decoder). Lseg Lseg , <Text comments> |
| DQT | FFDB | Define quantisation table(s). Lseg Lseg , < Qlum Qlum , Qchr Qchr . > |
| SOF0 SOF0 | FFC0 | Start of Baseline DCT frame. Lseg Lseg , <Frame size, no. of components (colours), sub-sampling factors, Q-table selectors> |
| DHT | FFC4 | Define Huffman table(s). Lseg Lseg , <DC Size and AC (Run,Size) tables for each component> |
| SOS | FFDA | Start of scan. Lseg Lseg , <Huffman table selectors for each component> <Entropy coded DCT blocks> |
| EOI | FFD9 | End of image. |

In [table 1](http://www.videotechnology.com/jpeg/j1.html#table1) the data which follows each marker is shown between <> brackets. The first 2-byte word of most segments is the length (in bytes) of the segment, Lseg Lseg . The length of <Entropy coded DCT blocks>, which forms the main bulk of the compressed data, is not specified explicitly, since it may be determined by decoding the entropy codes. This also allows the data to be transmitted with minimal delay, since it is not necessary to determine the total length of the compressed data before any of the DCT block data can be sent.

Long blocks of entropy-coded data are rather prone to being corrupted by transmission errors. To mitigate the worst aspects of this, Restart Markers (FFD0 . FFD7) may be included at regular intervals (say at the start of each row of DCT blocks in the image) so that separate parts of the entropy coded stream may be decoded independently of errors in other parts. The restart interval, if required, is defined by a DRI (FFDD) marker segment. There are 8 restart markers, which are used in sequence, so that if one (or more) is corrupted by errors, its absence may be easily detected.

The use of multiple scans within each image frame and multiple frames within a given image allows many variations on the ordering and interleaving of the compressed data. For example:

* Chrominance and luminance components may be sent in separate scans or interleaved into a single scan.
* Lower frequency DCT coefs may be sent in one or more scans before higher frequency coefs.
* Coarsely quantised coefs may be sent in one or more scans before finer (refinement) coefs.
* A coarsely sampled frame of the image may be sent initially and then the detail may be progressively improved by adding differentially-coded correction frames of increasing resolution.