Image Compression and Recognition

Compression Types

Now consider an encoder and a decoder as shown in Fig.. When the encoder receives the original image file, the image file will be converted into a series of binary data, which is called the bit-stream.

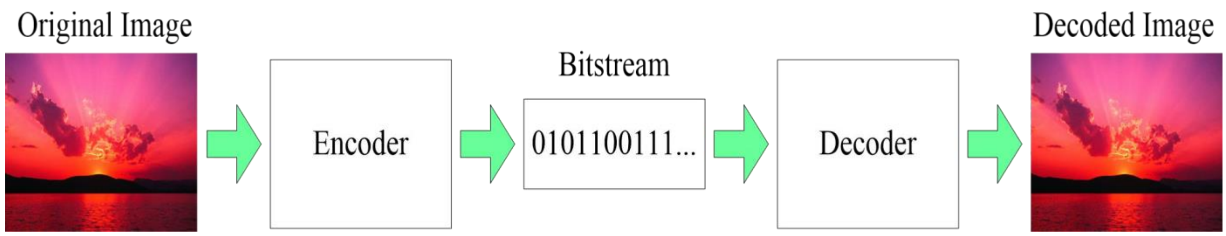

Fig.5.2.1The basic flow of image compression coding

Source: Tutorial point

The decoder then receives the encoded bit-stream and decodes it to form the decoded image. If the total data quantity of the bit-stream is less than the total data quantity of the original image, then this is called image compression.

Lossless compression can recover the exact original data after compression. It is used mainly for compressing database records, spreadsheets or word processing files.

There is no information loss, and the image can be reconstructed exactly the same as the original

Applications: Medical imagery

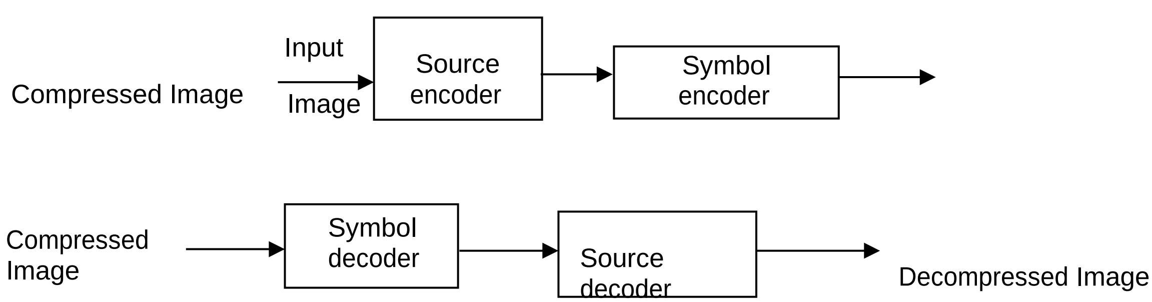

Fig5.2.2: The block diagram of encoder& decoder of Lossless compression

Source: Tutorial point

1.Source encoder is responsible for removing the coding and inter pixel redundancy and psycho visual redundancy.

2.Symbol encoder- This reduces the coding redundancy .This is the final stage of encoding process.

Source decoder- has two components

a) Symbol decoder- This performs inverse operation of symbol encoder.

b) Inverse mapping- This performs inverse operation of map.

<u>Lossless Compression technique</u>

Variable length coding (Huffman , arithmetic),

LZW coding, Bit Plane coding, Lossless Predictive coding

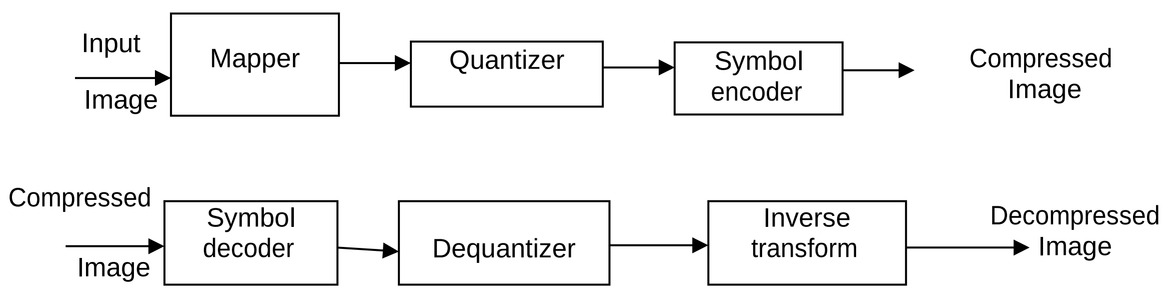

Fig5.2.3 The Block Diagram Of Encoder& Decoder Of Lossy compression

Source: Tutorial point

Lossy compression will result in a certain loss of accuracy in exchange for a Substantial increase in compression.

Lossy compression is more effective when used to compress graphic images

1. Information loss is tolerable
2. Many-to-1 mapping in compression eg. quantization
3. Applications: commercial distribution (DVD) and rate constrained environment

where lossless methods can not provide enough compression ratio

1)Mapper -this transforms the input data into non-visual format. It reduces the interpixel redundancy.

2)Quantizer - It reduces the psycho visual redundancy of the input images .This step is omitted if the system is error free.

3)Symbol encoder- This reduces the coding redundancy .This is the final stage of encoding process.

Source decoder- has two components

a)Symbol decoder- This performs inverse operation of symbol encoder.

b) Inverse mapping- This performs inverse operation of mapper.

Channel decoder-this is omitted if the system is error free.

MPEG COMPRESSION STANDARD.

Video Compression standards:

Video compression is used for compression of moving picture frames.

Based on the applications, video compression standards are grouped into two categories,

(1)Video teleconferencing standards

(2)Multimedia standards

The International Telecommunication Union (ITU) has a number of video conferencing compression standards such as, H.261, H.262, H.263 and H.320

1.H.261 is used to support full motion video transmission over T1 lines

2.H.263 is designed for very low bit rate video in the range of 10 to 30 Kbits/sec

3.H.320 is used for ISND

b.Multimedia Standards

These are used for video on demand, digital HDTV broadcasting and image/video database services.

The principal multimedia compression standards are MPEG-1, MPEG-2 and MEPG-4

1. MEPG-1 is an entertainment quality coding standard for the storage and retrieval of video on digital media such as CDROMs. It supports a bit rate of 1.5 Mbit/s. Thus it is used for compression of low resolution.

2. MPEG-2 is used for cable TV distribution and narrow channel satellite broadcasting with a bit rate of 2 to 10 Mbit/s. Thus is used for higher resolution standards, mainly for studio quality audio and video compression.

3. MPEG-4 provides 5 and 64 Kbit/s for mobile and PSTN and upto 4 Mbit/s for TV and film applications.

4. MPEG -4 provides (a) improved video compression efficiency; (b) Content based interactivity and (c) universal access including increased robustness.

MPEG Algorithm:

MPEG standard consists of both video and audio compression. The MPEG algorithm relies on two basic techniques

Block based motion compensation DCT based compression

1.The video signal is sequence of still picture frames. From the picture, slices are formed in the raster scan order as shown below,

2.For each slice the macro blocks are obtained. The macro block consists of four blocks of 8 x 8 pixels. Thus the micro block has a size of 16 x 16 pixels. If the macro block is in RGB format, then it is converted into Y Cr Cb format, where Cr and Cb represent chrominance signals. Once the sequence of macro blocks is formed, coding can take place.

3.The motion is estimated by predicting the current frame on the basis of certain previous and/or forward frame. The information sent to the decoder consists of the compressed DCT coefficients of the residual block together with the motion vector. There are three types of pictures in MPEG:Intra-frames (I) Predicted frames (P) Bidirectional predicted frames (B)

Figure demonstrates the position of the different types of pictures. Every Nth frame in the video sequence is an I-picture, and every Mth frame a P-picture. Here N=12 and M=4. The rest of the frames are B-pictures.

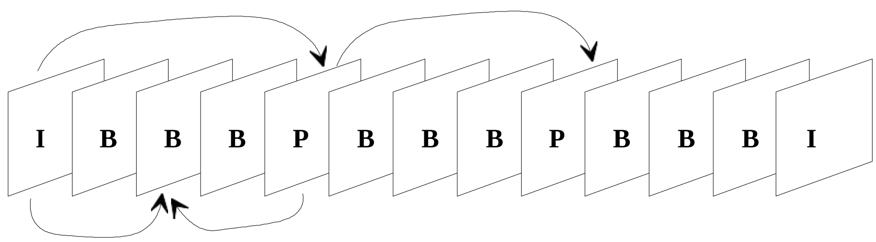

Figure5.2.4: Interframe coding in MPEG.

Source: Tutorial point

Compression of the picture types:

1. **Intra frame (I –frame)** are coded as still images by DCT algorithm. They provide access points for random access, but only with moderate compression. It is used as the reference point for the motion estimation.
2. **Predicted frame (P-frame)** are coded with reference to a past picture. The current frame is predicted on the basis of the previous I- or P-picture. The residual (difference between the prediction and the original picture) is then compressed by DCT.
3. **Bidirectional frame (B-frame)** is the compressed difference between the current frame and a prediction of it based on the previous I- or P- frame and next P-frame. Thus the decoder must have access to both he past and future reference frames. The encode frames are therefore recorded before transmission and the decoder

reconstructs and displays them in the proper sequence. Bidirectional pictures are never used as reference.

1. After the pictures are divided into 1616 macro blocks, each consisting of four 88 elementary blocks, the DCT of the macro block is calculated whose coefficients are quantized.

The choice of the prediction method is chosen for each macro block separately.

1.The intra-coded blocks are quantized differently from the predicted blocks:

Intra-coded blocks contain information in all frequencies and are quantized differently from the predicted blocks .The predicted blocks, contain mostly high frequencies and can be quantized with coarser quantization tables.

Motion estimation and compensation:

1.The prediction block in the reference frame is not necessarily in the same coordinates than the block in the current frame.

2.Because of motion in the image sequence, the most suitable predictor for the current block may exist anywhere in the reference frame.

3.The motion estimation specifies where the best prediction (best match) is found.

4.The prediction block in the reference frame is not necessarily in the same coordinates than the block in the current frame.

5.Because of motion in the image sequence, the most suitable predictor for the current block may exist anywhere in the reference frame.

6.The motion estimation specifies where the best prediction (best match) is found.

4.Motion compensation consists of calculating the difference between the reference and the current block.

5.Finally the quantized DCT coefficients are then run – length encoded in a zig-zag order and then huffmann coded to compress the data.

6.Also the encoder is designed to generate a bit stream that matches the capacity of the intended video channel. This is done by using a rate controller which adjusts the quantization parameters as a function of the occupancy of the output buffer.

7. As the buffer becomes fuller the quantization is made coarser, so that fewer bits stream into buffer.

Three basic types of encoded output frames:

1. **Intra frame or independent frame (I-frame).** An I frame is compressed independently of all previous and future video frames. Of the three possible encoded output frames, it most highly resembles a JPEG encoded image, Moreover, it is the reference point for the motion estimation needed to generate subsequent P- and B-frames.

2. **Predictive frame (P-frame).** A P-frame is the compressed difference between the current frame and a prediction of its based on the previous I-or P-frame. The difference is formed in the leftmost summer of fig. The prediction is motion compensated and typically involves sliding the decoded block in the lower part of fig. around its immediate neighborhood in the current frame and computing a measure of correlation (such as the sum of the square of the pixel-by-pixel differences). In fact, the process is often carried out in sub pixel increments (such as sliding the sub image ¼ pixels at a time), which necessitates interpolating pixel values prior to computing the correlation measure. The computed motion vector is variable length coded and transmitted as an integral part of the encoded data stream. Motion estimation is carried out on the macro block level.
3. **Bidirectional frame (B- frame).** A B- frame is the compressed difference between the current frame and a prediction of it based on the previous I- or P- frame and next- P- frame. Accordingly, the decoder must have access to both past and future reference frames. The encoded frames are therefore reordered before transmission; the decoder reconstructs and displays them in the proper sequence.