SGS-THOMSON MICROELECTRONICS

TECHNICAL NOTE

AN OVERVIEW OF THE MPEG COMPRESSION ALGORITHM

CONTENTS **Page**

I	**MPEG-1**	2
I.1	ALGORITHM STRUCTURE AND TERMINOLOGY	2
I.2	MPEG-1 COMPRESSION ALGORITHM	2
I.3	TEMPORAL COMPRESSION	2
I.4	PREDICTION	4
I.5	INTERPOLATION	4
I.6	SPATIAL COMPRESSION	5
II	**MPEG-2**	5
II.1	FIELD PREDICTION	6
II.2	DUAL-PRIME PREDICTION	6
II.3	EXTENDED MOTION VECTOR RANGE	6
II.4	FRAME/FIELD DCT CODING	6
II.5	CONCEALMENT MOTION VECTORS	6
II.6	ALTERNATIVE ZIG-ZAG SCAN	6
II.7	INCREASED DCT COEFFICIENT PRECISION	6
II.8	ALTERNATIVE INTRA VLCs	6
II.9	NON-LINEAR QUANTIZER SCALE MAPPING	6
II.10	ALTERNATIVE IDCT MISMATCH CONTROL	6
II.11	PAN-SCAN	6

There are now two MPEG standards, MPEG-1 and MPEG-2.

The MPEG-1 standard was developed in response to industry needs for an efficient way of storing and retrieving video information on digital storage media. One inexpensive medium is the CD-ROM which can deliver data at approximately 1.2Mbps, the MPEG-1 standard was subsequently aimed at this data rate, in fact the data rate is variable and all decoders must be able to decode at rates up to 1.856Mbps. Although the standard was developed with CD-ROM in mind, other storage and transmission media can include DAT, Winchester Disk, Optical Disk, ISDN and LAN. It is also important to say that it is common for double speed CD-ROMs to be used nowadays and it is genarally accepted that any MPEG-1 decoder system should be able to decode at up to 3Mbps.

The MPEG-2 standard was conceived to support the need for Digital Video Broadcating. The standard has been generally accepted for Digital TV systems, compressing fully interlaced TV resolution pictures to a bitrate up to around 20Mbps.

A version of MPEG-2 will also be used for the next generation of HDTV systems.

Two other relevant international standards were also being developed during the work of the MPEG committee : H.261 by CCITT aimed at telecommunications applications and ISO 10918 by the ISO JPEG committee aimed at the coding of still pictures. Elements of both standards were incorporated into the MPEG standard, but subsequent development work by the committee resulted in coding elements found in neither.

Some of the participants in the MPEG committee include : INTEL, BELLCORE, DEC, IBM, JVC Corp, PHILIPS CE, SGS-THOMSON, TCE, SONY Corp, NEC Corp and MATSUSHITA IEC. These may not be the most important members of the committee but it gives an indication of the relevant importance of the MPEG standard.

I - MPEG-1

Although the MPEG-1 standard is quite flexible, the basic algorithms have been tuned to work well at data rates from 1 to 1.5 Mbps, at resolutions of about 350 by 250 Pixels at picture rates of up to 25 or 30 pictures per second. MPEG-1 codes progressively-scanned images and does not recognise the concept of interlace, interlaced source video must be converted to a non interlace format prior to encoding. The format of the coded video allows forward play and pause, typical coding and decoding methods allow random access, fast forward and reverse play also, the requirements for these functions are very much application dependent and different encoding techniques will include varying levels of flexibility to account for these functions.

Compression of the digitised video comes from the use of several techniques : Sub sampling of the chroma information to match the human visual system, differential coding to exploit spatial redundancy, motion compensation to exploit temporal redundancy, Discrete Cosine Transform (DCT) to match typical image statistics, quantization, variable length coding, entropy coding and use of interpolated pictures.

I.1 - Algorithm Structure and Terminology

The MPEG hierarchy is arranged into layers (Figure 1). This layered structure is designed for flexibility and management efficiency, each layer is intended to support a specific function i.e. the sequence layer specifies sequence parameters such as picture size, aspect ratio, picture rate, bit rate etcetera , whereas the picture layer defines parameters such as the temporal reference and picture type. This layered structure improves robustness and reduces susceptibility to data corruption.

For convenience of coding, macroblocks are divided into six blocks of component Pixels - four luma and two chroma (Cr and Cb) (Figure 2). Blocks are the basic coding unit and the DCT is applied at this block level. Each block contains 64 component Pixels arranged in an 8x8 array (Figure 3).

There are four picture types : I pictures or INTRA pictures, which are coded without reference to any other pictures; P pictures or PREDICTED pictures which are coded using motion compensation from a previous picture; B pictures or BIDIRECTIONALLY predicted pictures which are coded using interpolation from a previous and a future picture and D pictures or DC pictures in which only the low frequency component is coded and which are only intended for fast forward search mode. B and P pictures are often called Inter pictures. Some other terminology that is often used are the terms M and N, M+1 represents the number of frames between successive I and P pictures whereas N+1 represents the number of frames between successive I pictures. M and N can be varied according to different applications and requirements such as fast random access.

A typical coding scheme will contain a mix of I,P and B pictures. A typical scheme will have an I picture every 10 to 15 pictures and two B pictures between succesive I and P pictures (Figure 4).

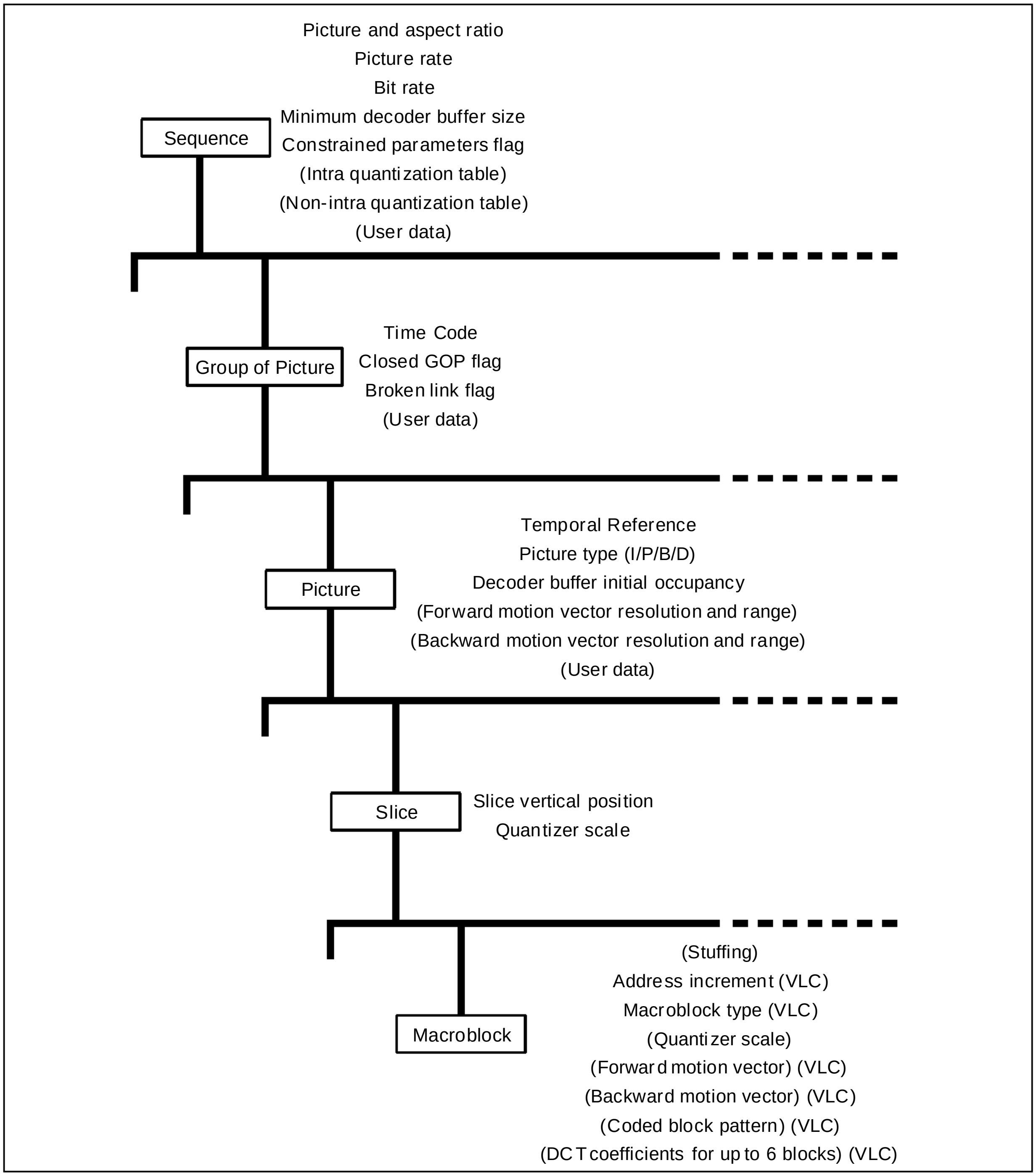
I.2 - MPEG-1 Compression Algorithm

The MPEG-1 algorithm is based around two key techniques : -temporal compression and spatial compression. Temporal compression relies upon similarity between successive pictures using prediction and motion compensation whereas spatial compression relies upon redundancy within small areas of a picture and is based around the DCT transform, quantization and entropy coding techniques.

I.3 - Temporal Compression

Inter (B and P) pictures are coded using motion compensation, primarily prediction and interpolation.

Figure 1 : MPEG Bistream Hierarchy



AN652-01.EPS

Figure 2 : Macroblock Structure

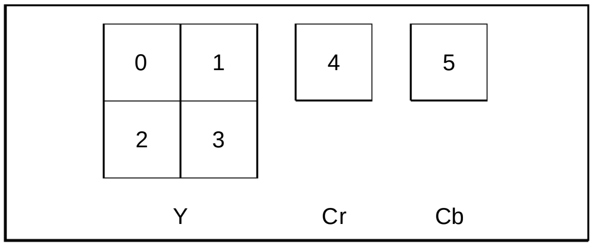


Figure 3 : Block Structure

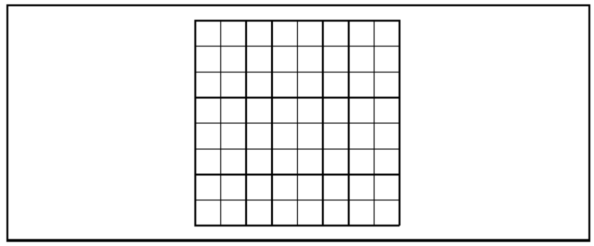
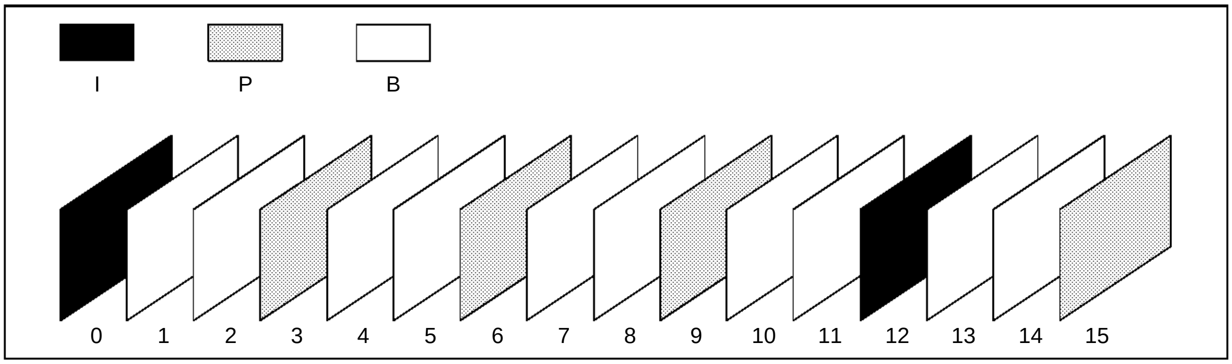


Figure 4 : Typical sequence of pictures in display order



I.4 - Prediction

The predicted picture is the previous picture modified by motion compensation. Motion vectors are calculated for each macroblock. The motion vector is applied to all four luminance blocks in the macro block. The motion vector for both chrominance blocks is calculated from the luma vector. This technique relies upon the assumption that within a macroblock the difference between successive pictures can be represented simply as a vector transform (i.e. there is very little difference between successive pictures, the key difference being in position of the Pixels) (Figure 5).

I.5 - Interpolation

Interpolation (or bidirectional prediction) generates high compression in that the picture is represented simply as an interpolation between the past and future I or P pictures (again this is performed on a macroblock level).

Pictures are not transmitted in display order but in the order in which the decoder requires them to decode the bitstream (the decoder must of course have the reference picture(s) before any interpolated or predicted pictures can be decoded).

Figure 5 : Make up of I, B and P pictures

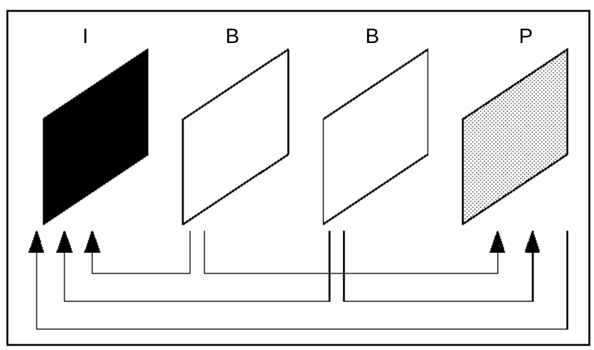
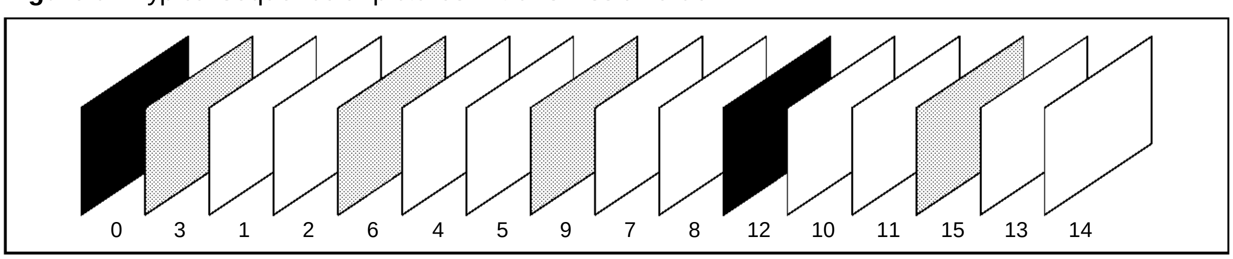


Figure 6 : Typical sequence of pictures in transmission order



I.6 - Spatial Compression

The spatial compression techniques are similar to those of JPEG , DCT, Quantization and entropy coding. The compression algorithm takes advantage of the redundancy within each block (8 x 8 Pixels).

The resulting compressed datastream is made up of a combination of spatial and temporal compression techniques which best suit the type of picture being compressed. Decoding is controlled through the use of MPEG system codes which are put into the data stream explaining how to reconstruct specific areas of picture - as shown in Figure 1.

Through a combination of techniques, MPEG-1 compression is designed to give good quality (typically similar or better quality to VHS) images from such storage media as CD-ROM. The quality is however, dependent upon the type of picture compressed and the level of redundancy within the sequence coded. Picture quality will also depend upon how well the sequence has been coded and which features are required - For Example : For fast random access, N will tend towards zero hence the quality of compression will deteriorate, if random access is not required then the number of P and B frames can increase, hence increasing the potential quality. The standard does not specify a method of compression but a syntax for the compressed data, this allows for differing compression techniques depending upon differing requirements. The decoding techniques are defined due to the nature of the compressed data stream.

This method allows for true flexibility in coding whilst retaining the format and hierarchy ensuring compatibility in the datastream and hence uniform readability.

II - MPEG-2

MPEG-2 was developed in response to the growing need for a generic coding method of moving pictures and associated sound for various applications such as digital storage media and TV broadcasting. The key applications targetted by MPEG-2 are : CATV, DBS and DAB.

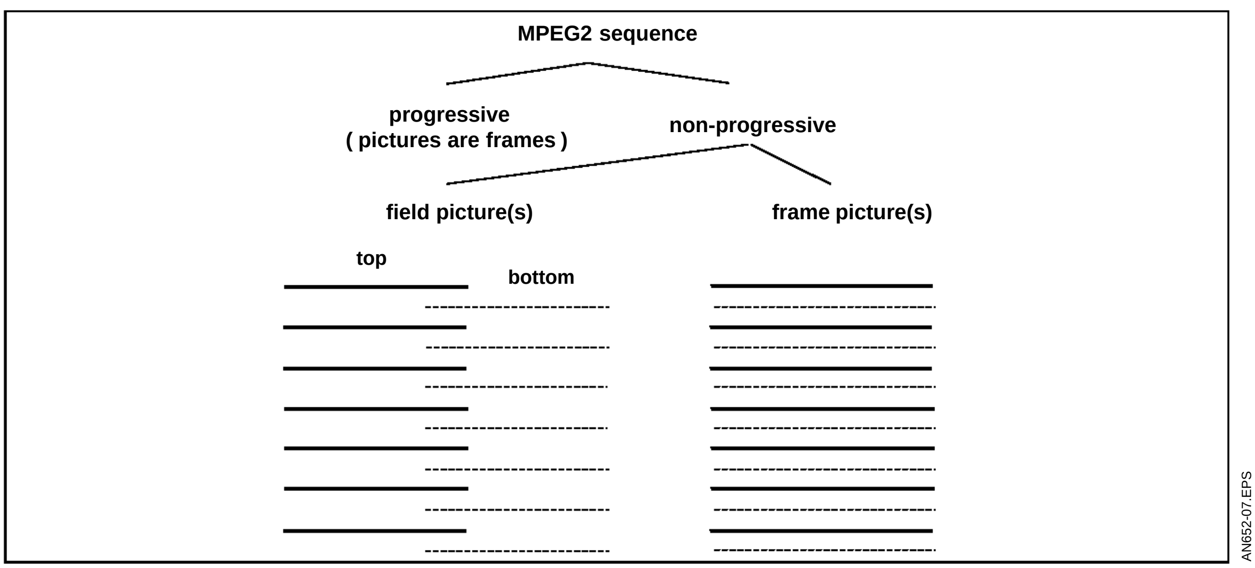
Within the standard, there are various ”levels” and ”profiles”, these are essentially a set of constraints imposed on parameters of the bistream, it is not practical today to develop systems with the ability to decode pictures as large as 2E14 pixels wide by 2E14 lines high (these sizes are within the standard). The development of the profiles and levels allows constrained MPEG-2 compatible systems, the key levels and profiles in todays systems are : MAIN PROFILE@MAINLEVEL (MP@ML), MAIN PROFILE@LOWLEVEL (MP@LL), SIMPLE PROFILE@MAIN LEVEL (SP@ML) and SIMPLE PROFILE@LOWLEVEL (SP@LL). The typical picture resolution of these images are 720x576x30fps and lower. These profiles and levels, ensure MPEG-2 compatibility between all decoders and encoders.

MPEG-2 is different from MPEG-1 in that it supports interlaced modes and more complex sampling structures (Figure 7).

The compression theory of MPEG-2 is the same as in MPEG-1, compression is achieved through the same techniques.

The inclusion of fields (i.e. interlaced frames) in MPEG-2 means that the Prediction modes in MPEG-2 are far more complex than in MPEG-1, it is here where most of the differences lie.

Figure 7 : Frame Pictures and Field Pictures



Relevant MPEG-2 Modes are as follows :

II.1 - Field Prediction

With fied prediction, the predictors in each field are independent. In a P frame picture with field prediction, the prediction is from two fields of the most recently decoded I- or P- frame. Picture prediction of a B- frame picture requires up to 4 vectors. The predictions for each field are formed by averaging the forward and backward predictors.

II.2 - Dual-prime Prediction

This prediction mode is applicable to sequences in which there are no B-pictures. Only one vector, plus a small differential vector, is sent with every macroblock. All necessary vectors are derived from these.

II.3 - Extended Motion Vector Range

In MPEG-1 the maximum motion vector range is -1024 to 1023 for full-pel resolution, and -512 to 511.5 for half-pel resolution. MPEG-2 specifies a maximum range of -2048 to 2047.5 at half-pel resolution ; full-pel resolution is not defined. MPEG-2 MP@ML specifies a vertical range of -128 to 127.5 and horizontal range of -1024 to 1023.5.

II.4 - Frame/Field DCT Coding

The luminance part of macroblock data may be reordered before the IDCT is performed on the blocks. The frame ordering is the same as that used in MPEG-1.

II.5 - Concealment Motion Vectors

The option exits in MPEG-2 to transmit motion vectors with intra macroblocks. This additional information is intended to be used when data is lost due to errors in the bitsteam. In this case the lost macroblocks are reconstructed as if they were skipped macroblocks in a forward predicted field or frame. The motion vectors are those which have been stored by the decoder in the previous row of macroblocks.

II.6 - Alternative Zig-Zag Scan

MPEG-2 allows two zig-zag block scanning patterns. Which one is used is specified in the picture layer of the bistream. The second pattern is intended to maximise the run lenghts of blocks containing data from a single field.

II.7 - Increased DCT Coefficient Precision

MPEG-2 decoders must be able to handle DCT coefficients in the range -2047 to 2048 (12 bits). The range in MPEG-1 is -256 to 255.

II.8 - Alternative Intra VLCs

The precision to which the DC coefficients of intra-coded blocks are coded can be changed every picture. Values of 8, 9 or 10 bits are allowed in MPEG-2 MP@ML (the precision is fixed at 8 bits in MPEG-1). There is also the option in MPEG-2 to select at the picture layer an alternative, more efficient, VLC table for coding the intra DCT coefficients. This selection is independent of the intra DC precision selection.

II.9 - Non-Linear Quantizer Scale Mapping

Two mappings are defined between the 5-bit quantizer scale code transmitted in the bistream and the actual quantizer scale. The first mapping is linear, as in MPEG-1. The alternative mapping is non-linear and is designed to give a finer control of quantizer scale when the values are low. Which mapping is used is defined at the picture layer of the bistream.

II.10 - Alternative IDCT Mismatch Control

The coefficients at the input to the IDCT are modified in order to avoid certain patterns of data.
This operation reduces the mismatch of outputs between different IDCT implementations, specifically in the encoder and decoder. In MPEG-1 the dequantized AC coefficients are forced to be odd before saturation. In MPEG-2 all of the coefficients of a block are summed after the saturation stage. If this sum even, the the last coefficient is changed by 1.

II.11 - Pan-Scan

Information is transmitted in the bistream which defines the dimensions and location of the displayable region within the encoded picture. The position of this window can be changed every field. This feature is useful, for example, when displaying part of a 16/9-format coded picture on a 4/3-format display. The MPEG-2 standard does not specify what should be done with this information, since it affects the display and not the decoding process.