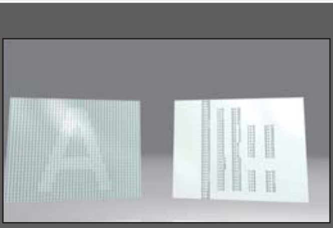
MPEG-4 AVC/H.264 codec technology explanation

First, what is compression? We will discuss by using faxes as an example ①. When sending an image, written on paper, via fax, the image is broken up into white and black sections and then encoded. For example, the black section is encoded to 1 and the white section is encoded to 0. Fax uses the “Run-Length Encoding” technique to reduce the size of the data. This process is called compression.



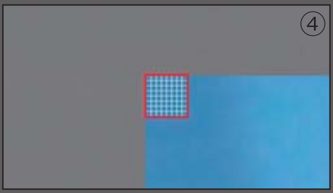
Video compression technology has been making significant progress for years. The MPEG-2 standard was a significant development and was standardized in 1995. MPEG-2 has been used in HDV camcorders, and was a breakthrough compression technology at that time. But, technology has advanced with the times, and MPEG-4 AVC/H.264 became standardized. MPEG-4 AVC/H.264 is at least twice as efficient in compression as MPEG-2. The compression algorithm used in MPEG-4 AVC/H.264 is basically the same as that used in MPEG-2. They use inter-frame prediction, quantization, entropy coding, etc. But in MPEG-4 AVC/H.264, these tools have been improved significantly and new tools, such as arithmetic coding and filters, have been added. Further, by using various modes appropriate for the characteristics of the images, MPEG-4 AVC/H.264 achieves compression efficiency far better than that of MPEG-2.

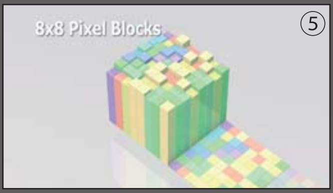
②
High Efficiency
H.264
MPEG2

DCT (Discrete Cosine Transform)

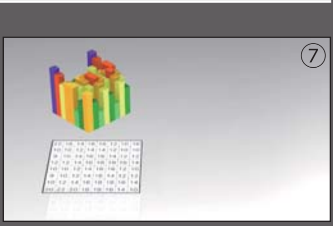
DCT is a transformation technique to compress original input images. Pixels are broken up into small blocks and then a compression coding process is applied to each block (③ to ⑥).

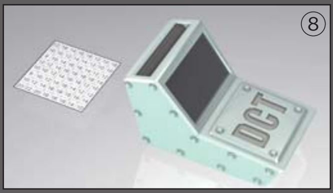
In the example, the size of the block is 8 x 8 pixels. DCT is a compression coding technique, based on a mathematical (discrete cosine) operation to transform an image into its frequency components (low frequency, high frequency, etc.). In simple terms, DCT transforms the method of expressing an image from “pixel blocks to frequency components,” preserving al l informat ion contained in the original input image; thus, it is a reversible process.

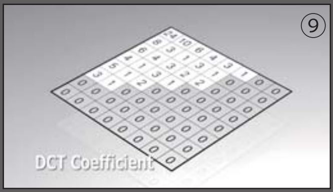


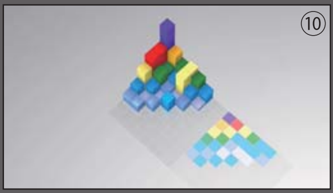


This illustrates the DCT process (⑦ to ⑩).



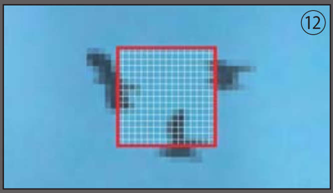


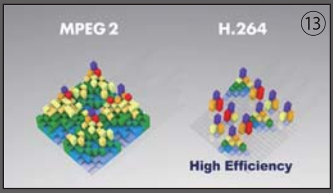




This is an image expressed in frequency components by applying DCT to the original input image signal ⑪.

What is the difference between MPEG-2 and MPEG-4 AVC/H.264 in DCT? As this example shows, you can see the difference by looking at mixed blocks with large and small color changes ⑫. A block layer, which is a processing unit in DCT, is always 8 x 8 pixels in MPEG-2, but is adjustable up to 4 x 4 pixels in MPEG-4 AVC/H.264. Therefore, MPEG-4 AVC/H.264 can provide far more efficient compression than MPEG-2 ⑬.

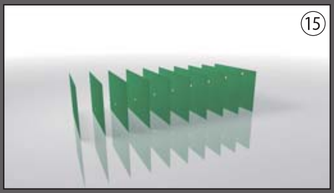


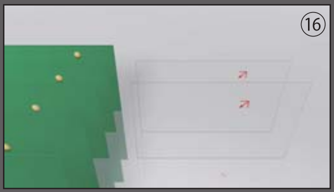


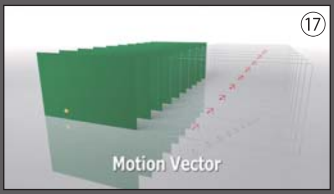
MPEG-4 AVC/H.264 codec technology explanation

Motion Compensation

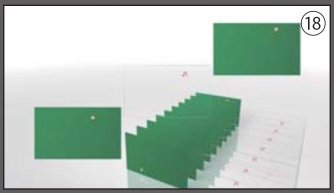
Many motion-picture coding techniques, such as MPEG-2 and MPEG-4 AVC/H.264, use motion compensation where image frames are broken up into blocks and movement is predict based on pre-coded frames. Look at the movement of the ball (⑭ to ⑮). The ball in the first frame moves toward down left. A motion vector is information about how an object moved ⑯ . For a block to be coded, prediction images are searched for the most similar block, and the motion between these blocks is represented by a motion vector and prediction error information ⑰ .

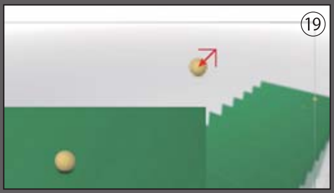




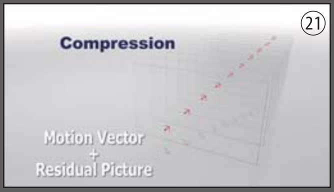


Prediction images are produced by using motion vectors (⑱ to ⑲), and then each prediction error is computed ⑳ . By doing so, efficient compression coding is achieved ㉑ .

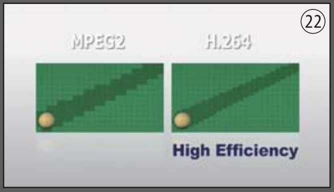




⑳
Residual Picture



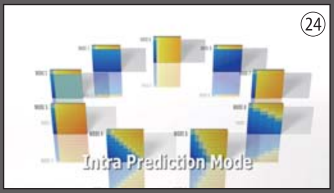
In actual motion compensation, an image is partitioned into small blocks (for example, the block size used in MPEG-2 is fixed at 16 x 16 pixels), and a motion vector is defined for each block. MPEG-4 AVC/H.264 allows smaller block sizes than that of MPEG-2 for prediction; therefore ㉒ by using an optimal size as required, MPEG-4 AVC/H.264 can offer much smoother motion compensation than MPEG-2.



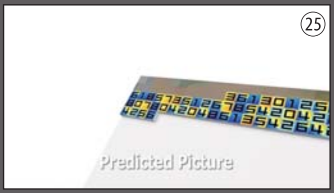
Intra-Prediction

AVC-Intra uses intra-prediction based on similarities among adjacent pixels to compress an image. More specifically, from an original input image, the pixel value in an 8x8 block to be coded is predicted and produced by using pre-coded adjacent pixel values ㉓ .

A predict ion mode can be selected from 9 modes for luminance signals and 4 modes for chroma signals ㉔ . The figure shows 9 intra prediction modes for luminance signals.



AVC-Intra determines an appropriate mode for each block. It also has a mode optimization algorithm to prevent image dubbing deterioration. Therefore, AVC-Intra can produce highly accurate intra-predicted images (㉕ to ㉗).



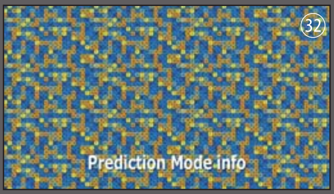
MPEG-4 AVC/H.264 codec technology explanation

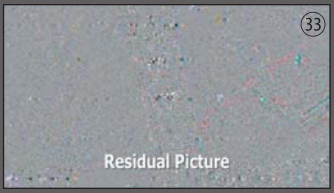
Subtracting the intra-predicted image from the original input image ㉘ produces a residual image ㉙. This is the residual image ㉚.

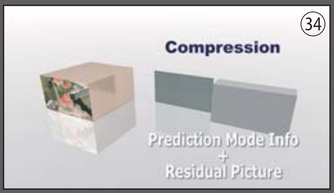
㉙
Subtraction

㉚
Residual Picture

Recording this original input image ㉛ will generate a large volume of data. AVC-Intra compresses ㉜ and records this intra-prediction mode information and the DCT of the residual image ㉝ together. Because this highly accurate intra-prediction can reduce the amount of data of the residual image, AVC-Intra can achieve high efficiency even with Intra (I-only) compression ㉞. (In addition, because Intra-predict ion predicts within the limits of a single frame, it has an advantage over inter-frame prediction in preventing the deterioration of prediction accuracy even for volatile movement.







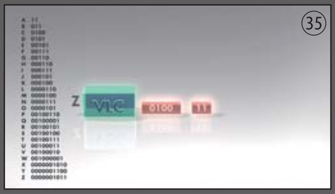
CABAC (Context Adaptive Binary Arithmetic Coding)

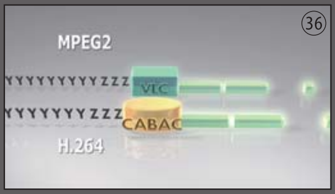
Entropy encoding is a method to process digital signals. Both MPEG-2 and MPEG-4 AVC/H.264 use entropy encoding but they differ significantly in method and characteristics.

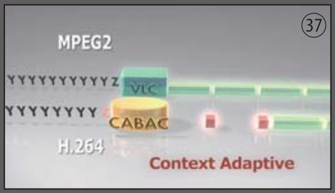
In MPEG-2, coding of each alphabetical character is done by using VLC (Variable Length Coding), which is based on a fixed table, so that Signal A is coded into a short code but Signal Z is coded into this very long code ㉟.

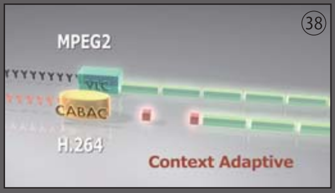
On the other hand, in MPEG-4 AVC/H.264, entropy encoding is performed adaptively to the characteristics of an image to achieve high compression efficiency. In particular, CABAC used in AVC-Intra50 is designed to adjust operations precisely to each image so that data is compressed stably and efficiently (㊱ to ㊳).

Context-Adaptive means “adapting to surrounding circumstances.” By using this Context-Adaptive technique in arithmetic coding, CABAC can achieve ideal entropy encoding efficiency close to the theoretical limit.


㉟
Z
VLC
0100
11


㊱
MPEG2
YYYYYYYYYYZZZ
VLC
YYYYYYYYZZZ
CABAC
H.264


㊲
MPEG2
YYYYYYYYYYZ
VLC
YYYYYYYY
CABAC
H.264
Context Adaptive


㊳
MPEG2
YYYYYYY
VLC
CABAC
H.264
Context Adaptive